Stochastic Calculus for Finance II: Continuous-Time Models
Solution of Exercise Problems

Yan Zeng

Version 1.0.8, last revised on 2015-03-13.

Abstract

This is a solution manual for Shreve [14]. If you find any typos/errors or have any comments, please email me at zypublic@hotmail.edu. This version skips Exercise 7.1, 7.2, 7.5–7.9.

Contents

1 General Probability Theory

★ **Comments**:

Example 1.1.4 of the textbook illustrates the paradigm of extending probability measure from a σ-algebra consisting of finitely many elements to a σ-algebra consisting of infinitely many elements. This procedure is typical of constructing probability measures and its validity is justified by Carathéodary's Extension Theorem (see, for example, Durrett[4, page 402]).

► **Exercise 1.1.** using the properties of Definition 1.1.2 for a probability measure $\mathbb{P}$, show the following.
(i) If $A \in \mathcal{F}$, $B \in \mathcal{F}$, and $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$.

Proof. $\mathbb{P}(B) = \mathbb{P}((B-A) \cup A) = \mathbb{P}(B-A) + P(A) \geq \mathbb{P}(A)$. □

(ii) If $A \in \mathcal{F}$ and $\{A_n\}_{n=1}^{\infty}$ is a sequence of sets in $\mathcal{F}$ with $\lim_{n\to\infty} \mathbb{P}(A_n) = 0$ and $A \subset A_n$ for every n, then $\mathbb{P}(A) = 0$. (This property was used implicitly in Example 1.1.4 when we argued that the sequence of all heads, and indeed any particular sequence, must have probability zero.)

Proof. According to (i), $\mathbb{P}(A) \leq \mathbb{P}(A_n)$, which implies $\mathbb{P}(A) \leq \lim_{n\to\infty} \mathbb{P}(A_n) = 0$. So $0 \leq \mathbb{P}(A) \leq 0$. This means $\mathbb{P}(A) = 0$. □

► **Exercise 1.2.** The infinite coin-toss space Ω_∞ of Example 1.1.4 is *uncountably infinite*. In other words, we cannot list all its elements in a sequence. To see that this is impossible, suppose there were such a sequential list of all elements of Ω_∞:

$$\begin{aligned} \omega^{(1)} &= \omega_1^{(1)}\omega_2^{(1)}\omega_3^{(1)}\omega_4^{(1)}\cdots, \\ \omega^{(2)} &= \omega_1^{(2)}\omega_2^{(2)}\omega_3^{(2)}\omega_4^{(2)}\cdots, \\ \omega^{(3)} &= \omega_1^{(3)}\omega_2^{(3)}\omega_3^{(3)}\omega_4^{(3)}\cdots, \\ &\vdots \end{aligned}$$

An element that does not appear in this list is the sequence whose first component is H is $\omega_1^{(1)}$ is T and is T if $\omega_1^{(1)}$ is H, whose second component is H if $\omega_2^{(2)}$ is T and is T if $\omega_2^{(2)}$ is H, whose third component is H if $\omega_3^{(3)}$ is T and is T if $\omega_3^{(3)}$ is H, etc. Thus, the list does not include every element of Ω_∞.

Now consider the set of sequences of coin tosses in which the outcome on each even-numbered toss matches the outcome of the toss preceding it, i.e.,

$$A = \{\omega = \omega_1\omega_2\omega_3\omega_4\omega_5\cdots; \omega_1 = \omega_2, \omega_3 = \omega_4, \cdots\}.$$

(i) Show that A is uncountably infinite.

Proof. We define a mapping ϕ from A to Ω_∞ as follows: $\phi(\omega_1\omega_2\cdots) = \omega_1\omega_3\omega_5\cdots$. Then ϕ is one-to-one and onto. So the cardinality of A is the same as that of Ω_∞, which means in particular that A is uncountably infinite. □

(ii) Show that, when $0 < p < 1$, we have $\mathbb{P}(A) = 0$.

Proof. Let $A_n = \{\omega = \omega_1\omega_2\cdots : \omega_1 = \omega_2, \cdots, \omega_{2n-1} = \omega_{2n}\}$. Then $A_n \downarrow A$ as $n \to \infty$. So

$$\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} [\mathbb{P}(\omega_1 = \omega_2)\cdots\mathbb{P}(\omega_{2n-1} = \omega_{2n})] = \lim_{n\to\infty} [p^2 + (1-p)^2]^n.$$

Since $p^2 + (1-p)^2 \leq \max\{p, 1-p\}[p + (1-p)] < 1$ for $0 < p < 1$, we have $\lim_{n\to\infty}(p^2 + (1-p)^2)^n = 0$. This implies $\mathbb{P}(A) = 0$. □

► **Exercise 1.3.** Consider the set function $\mathbb{P}$ defined for every subset of $[0,1]$ by the formula that $\mathbb{P}(A) = 0$ if A is a finite set and $\mathbb{P}(A) = \infty$ if A is an infinite set. Show that $\mathbb{P}$ satisfies (1.1.3)-(1.1.5), but $\mathbb{P}$ does not have the countable additivity property (1.1.2). We see then that the finite additivity property (1.1.5) does not imply the countable additivity property (1.1.2).

Proof. Clearly $\mathbb{P}(\emptyset) = 0$. For any A and B, if both of them are finite, then $A \cup B$ is also finite. So $\mathbb{P}(A \cup B) = 0 = \mathbb{P}(A) + \mathbb{P}(B)$. If at least one of them is infinite, then $A \cup B$ is also infinite. So $\mathbb{P}(A \cup B) = \infty = \mathbb{P}(A) + \mathbb{P}(B)$. Similarly, we can prove $\mathbb{P}(\cup_{n=1}^N A_n) = \sum_{n=1}^N \mathbb{P}(A_n)$, even if A_n's are not disjoint.

To see countable additivity property doesn't hold for $\mathbb{P}$, let $A_n = \{\frac{1}{n}\}$. Then $A = \cup_{n=1}^\infty A_n$ is an infinite set and therefore $\mathbb{P}(A) = \infty$. However, $\mathbb{P}(A_n) = 0$ for each n. So $\mathbb{P}(A) \neq \sum_{n=1}^\infty \mathbb{P}(A_n)$. □

► **Exercise 1.4.**

(i) Construct a standard normal random variable Z on the probability space $(\Omega_\infty, \mathcal{F}_\infty, \mathbb{P})$ of Example 1.1.4 under the assumption that the probability for head is $p = \frac{1}{2}$. (Hint: Consider Examples 1.2.5 and 1.2.6)

Solution. By Example 1.2.5, we can construct a random variable X on the coin-toss space, which is uniformly distributed on $[0,1]$. For the strictly increasing and continuous function $N(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi$, we let $Z = N^{-1}(X)$. Then $\mathbb{P}(Z \leq a) = \mathbb{P}(X \leq N(a)) = N(a)$ for any real number a, i.e. Z is a standard normal random variable on the coin-toss space $(\Omega_\infty, \mathcal{F}_\infty, \mathbb{P})$. □

(ii) Define a sequence of random variables $\{Z_n\}_{n=1}^\infty$ on Ω_∞ such that

$$\lim_{n\to\infty} Z_n(\omega) = Z(\omega) \text{ for every } \omega \in \Omega_\infty$$

and, for each n, Z_n depends only on the first n coin tosses. (This gives us a procedure for approximating a standard normal random variable by random variables generated by a finite number of coin tosses, a useful algorithm for Monte Carlo simulation).

Solution. Define

$$X_n = \sum_{i=1}^n \frac{1}{2^i} 1_{\{\omega_i = H\}}.$$

Then $X_n(\omega) \to X(\omega)$ for every $\omega \in \Omega_\infty$ where X is defined as in Example 1.2.5. So $Z_n = N^{-1}(X_n) \to Z = N^{-1}(X)$ for every ω. Clearly Z_n depends only on the first n coin tosses and $\{Z_n\}_{n\geq 1}$ is the desired sequence. □

► **Exercise 1.5.** When dealing with double Lebesgue integrals, just as with double Riemann integrals, the order of integration can be reversed. The only assumption required is that the function being integrated be either nonnegative or integrable. Here is an application of this fact.

Let X be a nonnegative random variable with cumulative distribution function $F(x) = \mathbb{P}\{X \leq x\}$. Show that

$$\mathbb{E}X = \int_0^\infty (1 - F(x))dx$$

by showing that

$$\int_\Omega \int_0^\infty 1_{[0,X(\omega))}(x) dx d\mathbb{P}(\omega)$$

is equal to both $\mathbb{E}X$ and $\int_0^\infty (1 - F(x))dx$.

Proof. First, by the information given by the problem, we have

$$\int_\Omega \int_0^\infty 1_{[0,X(\omega))}(x) dx d\mathbb{P}(\omega) = \int_0^\infty \int_\Omega 1_{[0,X(\omega))}(x) d\mathbb{P}(\omega) dx.$$

The left side of this equation equals to

$$\int_\Omega \int_0^{X(\omega)} dx dP(\omega) = \int_\Omega X(\omega) dP(\omega) = \mathbb{P}[X].$$

The right side of the equation equals to

$$\int_0^\infty \int_\Omega 1_{\{x<X(\omega)\}} d\mathbb{P}(\omega) dx = \int_0^\infty \mathbb{P}(x < X) dx = \int_0^\infty (1 - F(x)) dx.$$

So $\mathbb{E}[X] = \int_0^\infty (1 - F(x))dx$. □

► **Exercise 1.6.** Let u be a fixed number in $\mathbb{R}$, and define the convex function $\varphi(x) = e^{ux}$ for all $x \in \mathbb{R}$. Let X be a normal random variable with mean $\mu = \mathbb{E}X$ and standard deviation $\sigma = [\mathbb{E}(X-\mu)^2]^{\frac{1}{2}}$, i.e., with density

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

(i) Verify that

$$\mathbb{E}e^{uX} = e^{u\mu + \frac{1}{2}u^2\sigma^2}.$$

Proof.

$$\begin{aligned}
\mathbb{E}\left[e^{uX}\right] &= \int_{-\infty}^{\infty} e^{ux} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2 - 2\sigma^2 ux}{2\sigma^2}} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{[x-(\mu+\sigma^2 u)]^2 - (2\sigma^2 u\mu + \sigma^4 u^2)}{2\sigma^2}} dx \\
&= e^{u\mu + \frac{\sigma^2 u^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{[x-(\mu+\sigma^2 u)]^2}{2\sigma^2}} dx \\
&= e^{u\mu + \frac{\sigma^2 u^2}{2}}
\end{aligned}$$

□

(ii) Verify that Jensen's inequality holds (as it must):

$$\mathbb{E}\varphi(X) \geq \varphi(\mathbb{E}X).$$

Proof. $\mathbb{E}[\varphi(X)] = \mathbb{E}\left[e^{uX}\right] = e^{u\mu + \frac{u^2\sigma^2}{2}} \geq e^{u\mu} = \varphi(\mathbb{E}[X])$. □

► **Exercise 1.7.** For each positive integer n, define f_n to be the normal density with mean zero and variance n, i.e.,

$$f_n(x) = \frac{1}{\sqrt{2n\pi}} e^{-\frac{x^2}{2n}}.$$

(i) What is the function $f(x) = \lim_{n\to\infty} f_n(x)$?

Solution. Since $|f_n(x)| \leq \frac{1}{\sqrt{2n\pi}}$, $f(x) = \lim_{n\to\infty} f_n(x) = 0$. □

(ii) What is $\lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(x)dx$?

Solution. By the change of variable formula, $\int_{-\infty}^{\infty} f_n(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$. So we must have

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(x)dx = 1.$$

$\square$

(iii) Note that

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(x)dx \neq \int_{-\infty}^{\infty} f(x)dx.$$

Explain why this does not violate the Monotone Convergence Theorem, Theorem 1.4.5.

Solution. This is not contradictory with the Monotone Convergence Theorem because $\{f_n\}_{n\geq 1}$ doesn't increase to 0. $\square$

▶ **Exercise 1.8. (Moment-generating function).** Let X be a nonnegative random variable, and assume that

$$\varphi(t) = \mathbb{E}e^{tX}$$

is finite for every $t \in \mathbb{R}$. Assume further that $\mathbb{E}\left[Xe^{tX}\right] < \infty$ for every $t \in \mathbb{R}$. The purpose of this exercise is to show that $\varphi'(t) = \mathbb{E}\left[Xe^{tX}\right]$ and, in particular, $\varphi'(0) = \mathbb{E}X$.

We recall the definition of derivative:

$$\varphi'(t) = \lim_{s\to t} \frac{\varphi(t)-\varphi(s)}{t-s} = \lim_{s\to t} \frac{\mathbb{E}e^{tX} - \mathbb{E}e^{sX}}{t-s} = \lim_{s\to t} \mathbb{E}\left[\frac{e^{tX}-e^{sX}}{t-s}\right].$$

The limit above is taken over a *continuous* variable s, but we can choose a sequence of numbers $\{s_n\}_{n=1}^{\infty}$ converging to t and compute

$$\lim_{s_n\to t} \mathbb{E}\left[\frac{e^{tX}-e^{s_nX}}{t-s_n}\right],$$

where now we are taking a limit of the expectation of the *sequence* of random varibles

$$Y_n = \frac{e^{tX}-e^{s_nX}}{t-s_n}.$$

If this limit turns out to be the same, regardless of how we choose the sequence $\{s_n\}_{n=1}^{\infty}$ that converges to t, then this limit is also the same as $\lim_{s\to t} \mathbb{E}\left[\frac{e^{tX}-e^{sX}}{t-s}\right]$ and is $\varphi'(t)$.

The Mean Value Theorem from calculus states that if $f(t)$ is a differentiable function, then for any two numbers s and t, there is a number θ between s and t such that

$$f(t) - f(s) = f'(\theta)(t-s).$$

If we fix $\omega \in \Omega$ and define $f(t) = e^{tX(\omega)}$, then this becomes

$$e^{tX(\omega)} - e^{sX(\omega)} = (t-s)X(\omega)e^{\theta(\omega)X(\omega)}, \tag{1.9.1}$$

where $\theta(\omega)$ is a number depending on ω (i.e., a random variable lying between t and s).

(i) Use the Dominated Convergence Theorem (Theorem 1.4.9) and equation (1.9.1) to show that

$$\lim_{n\to\infty} \mathbb{E}Y_n = \mathbb{E}\left[\lim_{n\to\infty} Y_n\right] = \mathbb{E}\left[Xe^{tX}\right]. \tag{1.9.2}$$

This establishes the desired formula $\varphi'(t) = \mathbb{E}\left[Xe^{tX}\right]$.

Proof. By (1.9.1), $|Y_n| = \left|\frac{e^{tX}-e^{s_nX}}{t-s_n}\right| = |Xe^{\theta_nX}| = Xe^{\theta_nX} \leq Xe^{\max\{2|t|,1\}X}$. The last inequality is by $X \geq 0$ and the fact that θ_n is between t and s_n, and hence smaller than $\max\{2|t|,1\}$ for n sufficiently large. So by the Dominated Convergence Theorem, $\varphi'(t) = \lim_{n\to\infty} \mathbb{E}[Y_n] = \mathbb{E}[\lim_{n\to\infty} Y_n] = \mathbb{E}[Xe^{tX}]$. $\square$

(ii) Suppose the random variable X can take both positive and negative values and $\mathbb{E}e^{tX} < \infty$ and $\mathbb{E}\left[|X|e^{tX}\right] < \infty$ for every $t \in \mathbb{R}$. Show that once again $\varphi'(t) = \mathbb{E}\left[Xe^{tX}\right]$. (Hint: Use the notation (1.3.1) to write $X = X^+ - X^-$.)

Proof. Since $\mathbb{E}[e^{tX^+}1_{\{X\geq 0\}}] + \mathbb{E}[e^{-tX^-}1_{\{X<0\}}] = \mathbb{E}[e^{tX}] < \infty$ for every $t \in \mathbb{R}$, $\mathbb{E}[e^{t|X|}] = \mathbb{E}[e^{tX^+}1_{\{X\geq 0\}}] + \mathbb{E}[e^{-(-t)X^-}1_{\{X<0\}}] < \infty$ for every $t \in \mathbb{R}$. Similarly, we have $\mathbb{E}[|X|e^{t|X|}] < \infty$ for every $t \in \mathbb{R}$. So, similar to (i), we have $|Y_n| = |Xe^{\theta_n X}| \leq |X|e^{\max\{2t,1\}|X|}$ for n sufficiently large, and by the Dominated Convergence Theorem, $\varphi'(t) = \lim_{n\to\infty}\mathbb{E}[Y_n] = \mathbb{E}[\lim_{n\to\infty} Y_n] = \mathbb{E}[Xe^{tX}]$. □

► **Exercise 1.9.** Suppose X is a random variable on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, A is a set in $\mathcal{F}$, and for every Borel subset B of $\mathbb{R}$, we have

$$\int_A 1_B(X(\omega))d\mathbb{P}(\omega) = \mathbb{P}(A)\cdot\mathbb{P}\{X \in B\}. \tag{1.9.3}$$

Then we say that X is *independent* of the event A.

Show that if X is independent of an event A, then

$$\int_A g(X(\omega))d\mathbb{P}(\omega) = \mathbb{P}(A)\cdot\mathbb{E}g(X)$$

for every nonnegative, Borel-measurable function g.

Proof. If $g(x)$ is of the form $1_B(x)$, where B is a Borel subset of $\mathbb{R}$, then the desired equality is just (1.9.3). By the linearity of Lebesgue integral, the desired equality also holds for simple functions, i.e. g of the form $g(x) = \sum_{i=1}^n 1_{B_i}(x)$, where each B_i is a Borel subset of $\mathbb{R}$. Since any nonnegative, Borel-measurable function g is the limit of an increasing sequence of simple functions, the desired equality can be proved by the Monotone Convergence Theorem. □

► **Exercise 1.10.** Let $\mathbb{P}$ be the uniform (Lebesgue) measure on $\Omega = [0,1]$. Define

$$Z(\omega) = \begin{cases} 0 & \text{if } 0 \leq \omega < \frac{1}{2}, \\ 2 & \text{if } \frac{1}{2} \leq \omega \leq 1. \end{cases}$$

For $A \in \mathcal{B}[0,1]$, define

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega)d\mathbb{P}(\omega).$$

(i) Show that $\widetilde{\mathbb{P}}$ is a probability measure.

Proof. If $\{A_i\}_{i=1}^\infty$ is a sequence of disjoint Borel subsets of $[0,1]$, then by the Monotone Convergence Theorem, $\widetilde{\mathbb{P}}(\cup_{i=1}^\infty A_i)$ equals to

$$\int 1_{\cup_{i=1}^\infty A_i} Z d\mathbb{P} = \int \lim_{n\to\infty} 1_{\cup_{i=1}^n A_i} Z d\mathbb{P} = \lim_{n\to\infty}\int 1_{\cup_{i=1}^n A_i} Z d\mathbb{P} = \lim_{n\to\infty}\sum_{i=1}^n \int_{A_i} Z d\mathbb{P} = \sum_{i=1}^\infty \widetilde{\mathbb{P}}(A_i).$$

Meanwhile, $\widetilde{\mathbb{P}}(\Omega) = 2\mathbb{P}([\frac{1}{2}, 1]) = 1$. So $\widetilde{\mathbb{P}}$ is a probability measure. □

(ii) Show that if $\mathbb{P}(A) = 0$, then $\widetilde{\mathbb{P}}(A) = 0$. We say that $\widetilde{\mathbb{P}}$ is *absolutely continuous* with respect to $\mathbb{P}$.

Proof. If $\mathbb{P}(A) = 0$, then $\widetilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P} = 2\int_{A\cap[\frac{1}{2},1]} d\mathbb{P} = 2\mathbb{P}(A \cap [\frac{1}{2}, 1]) = 0$. □

(iii) Show that there is a set A for which $\widetilde{\mathbb{P}}(A) = 0$ but $\mathbb{P}(A) > 0$. In other words, $\widetilde{\mathbb{P}}$ and $\mathbb{P}$ are not equivalent.

Proof. Let $A = [0, \frac{1}{2})$. □

▶ **Exercise 1.11.** In Example 1.6.6, we began with a standard normal random variable X under a measure $\mathbb{P}$. According to Exercise 1.6, this random variable has the moment-generating function

$$\mathbb{E}e^{uX} = e^{\frac{1}{2}u^2} \text{ for all } u \in \mathbb{P}.$$

The moment-generating function of a random variable determines its distribution. In particular, any random variable that has moment-generating function $e^{\frac{1}{2}u^2}$ must be standard normal.

In Example 1.6.6, we also defined $Y = X + \theta$, where θ is a constant, we set $Z = e^{-\theta X - \frac{1}{2}\theta^2}$, and we defined $\widetilde{\mathbb{P}}$ by the formula (1.6.9):

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega)d\mathbb{P}(\omega) \text{ for all } A \in \mathcal{F}.$$

We showed by considering its cumulative distribution function that Y is a standard normal random variable under $\widetilde{\mathbb{P}}$. Give another proof that Y is standard normal under $\widetilde{\mathbb{P}}$ by verifying the moment-generating function formula

$$\widetilde{\mathbb{E}}e^{uY} = e^{\frac{1}{2}u^2} \text{ for all } u \in \mathbb{R}.$$

Proof.

$$\widetilde{\mathbb{E}}\left[e^{uY}\right] = \mathbb{E}\left[e^{uY}Z\right] = \mathbb{E}\left[e^{uX+u\theta}e^{-\theta X - \frac{\theta^2}{2}}\right] = e^{u\theta - \frac{\theta^2}{2}}E\left[e^{(u-\theta)X}\right] = e^{u\theta - \frac{\theta^2}{2}}e^{\frac{(u-\theta)^2}{2}} = e^{\frac{u^2}{2}}.$$

□

▶ **Exercise 1.12.** In Example 1.6.6, we began with a standard normal random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and defined the random variable $Y = X + \theta$, where θ is a constant. We also defined $Z = e^{-\theta X - \frac{1}{2}\theta^2}$ and used Z as the Radon-Nikodým derivative to construct the probability measure $\widetilde{\mathbb{P}}$ by the formula (1.6.9):

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega)d\mathbb{P}(\omega), \text{ for all } A \in \mathcal{F}.$$

Under $\widetilde{\mathbb{P}}$, the random variable Y was shown to be standard normal.

We now have a standard normal random variable Y on the probability space $(\Omega, \mathcal{F}, \widetilde{\mathbb{P}})$, and X is related to Y by $X = Y - \theta$. By what we have just stated, with X replaced by Y and θ replaced by $-\theta$, we could define $\widehat{Z} = e^{\theta Y - \frac{1}{2}\theta^2}$ and then use $\widehat{Z}$ as a Radon-Nikodým derivative to construct a probability measure $\widehat{\mathbb{P}}$ by the formula

$$\widehat{\mathbb{P}}(A) = \int_A \widehat{Z}(\omega)d\widetilde{\mathbb{P}}(\omega) \text{ for all } A \in \mathcal{F},$$

so that, under $\widehat{\mathbb{P}}$, the random variable X is standard normal. Show that $\widehat{Z} = \frac{1}{Z}$ and $\widehat{\mathbb{P}} = \mathbb{P}$.

Proof. First, $\widehat{Z} = e^{\theta Y - \frac{\theta^2}{2}} = e^{\theta(X+\theta) - \frac{\theta^2}{2}} = e^{\frac{\theta^2}{2} + \theta X} = Z^{-1}$. Second, for any $A \in \mathcal{F}$, $\widehat{\mathbb{P}}(A) = \int_A \widehat{Z}d\widetilde{\mathbb{P}} = \int (1_A\widehat{Z})Zd\mathbb{P} = \int 1_A d\mathbb{P} = \mathbb{P}(A)$. So $\mathbb{P} = \widehat{\mathbb{P}}$. In particular, X is standard normal under $\widehat{\mathbb{P}}$, since it's standard normal under $\mathbb{P}$. □

▶ **Exercise 1.13 (Change of measure for a normal random variable).** A nonrigorous but informative derivation of the formula for the Radon-Nikodým derivative $Z(\omega)$ in Example 1.6.6 is provided by this exercise. As in that example, let X be a standard normal random variable on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $Y = X + \theta$. Our goal is to define a strictly positive random variable $Z(\omega)$ so that when we set

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega)d\widetilde{\mathbb{P}}(\omega) \text{ for all } A \in \mathcal{F}, \tag{1.9.4}$$

the random variable Y under $\widetilde{\mathbb{P}}$ is standard normal. If we fix $\overline{\omega} \in \Omega$ and choose a set A that contains $\overline{\omega}$ and is "small," then (1.9.4) gives

$$\widetilde{\mathbb{P}}(A) \approx Z(\overline{\omega})\mathbb{P}(A),$$

where the symbol $\approx$ means "is approximately equal to." Dividing by $\mathbb{P}(A)$, we see that

$$\frac{\widetilde{\mathbb{P}}(A)}{\mathbb{P}(A)} \approx Z(\overline{\omega})$$

for "small" sets A containing $\overline{\omega}$. We use this observation to identify $Z(\overline{\omega})$.

With $\overline{\omega}$ fixed, let $x = X(\overline{\omega})$. For $\epsilon > 0$, we define $B(x,\epsilon) = \left[x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}\right]$ to be the closed interval centered at x and having length ϵ. Let $y = x + \theta$ and $B(y,\epsilon) = \left[y - \frac{\epsilon}{2}, y + \frac{\epsilon}{2}\right]$.

(i) Show that

$$\frac{1}{\epsilon}\mathbb{P}\{X \in B(x,\epsilon)\} \approx \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{X^2(\overline{\omega})}{2}\right\}.$$

Proof.

$$\frac{1}{\epsilon}\mathbb{P}(X \in B(x,\epsilon)) = \frac{1}{\epsilon}\int_{x-\frac{\epsilon}{2}}^{x+\frac{\epsilon}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \approx \frac{1}{\epsilon}\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \epsilon = \frac{1}{\sqrt{2\pi}} e^{-\frac{X^2(\bar{\omega})}{2}}.$$

□

(ii) In order for Y to be a standard normal random variable under $\widetilde{\mathbb{P}}$, show that we must have

$$\frac{1}{\epsilon}\widetilde{\mathbb{P}}\{Y \in B(y,\epsilon)\} \approx \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{Y^2(\overline{\omega})}{2}\right\}.$$

Proof. Similar to (i). □

(iii) Show that $\{X \in B(x,\epsilon)\}$ and $\{Y \in B(y,\epsilon)\}$ are the same set, which we call $A(\overline{\omega},\epsilon)$. This set contains $\overline{\omega}$ and is "small" when $\epsilon > 0$ is small.

Proof. $\{X \in B(x,\epsilon)\} = \{X \in B(y-\theta,\epsilon)\} = \{X + \theta \in B(y,\epsilon)\} = \{Y \in B(y,\epsilon)\}$. □

(iv) Show that

$$\frac{\widetilde{\mathbb{P}}(A)}{\mathbb{P}(A)} \approx \exp\left\{-\theta X(\overline{\omega}) - \frac{1}{2}\theta^2\right\}.$$

The right-hand side is the value we obtained for $Z(\overline{\omega})$ in Example 1.6.6.

Proof. By (i)-(iii), $\frac{\widetilde{\mathbb{P}}(A)}{\mathbb{P}(A)}$ is approximately

$$\frac{\frac{\epsilon}{\sqrt{2\pi}} e^{-\frac{Y^2(\bar{\omega})}{2}}}{\frac{\epsilon}{\sqrt{2\pi}} e^{-\frac{X^2(\bar{\omega})}{2}}} = e^{-\frac{Y^2(\bar{\omega}) - X^2(\bar{\omega})}{2}} = e^{-\frac{(X(\bar{\omega})+\theta)^2 - X^2(\bar{\omega})}{2}} = e^{-\theta X(\bar{\omega}) - \frac{\theta^2}{2}}.$$

□

▶ **Exercise 1.14 (Change of measure for an exponential random variable).** Let X be a nonnegative random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the *exponential distribution*, which is

$$\mathbb{P}\{X \le a\} = 1 - e^{-\lambda a},\ a \ge 0,$$

where λ is a positive constant. Let $\tilde{\lambda}$ be another positive constant, and define

$$Z = \frac{\tilde{\lambda}}{\lambda} e^{-(\tilde{\lambda}-\lambda)X}$$

Define $\widetilde{\mathbb{P}}$ by

$$\widetilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P} \text{ for all } A \in \mathcal{F}.$$

(i) Show that $\widetilde{\mathbb{P}}(\Omega) = 1$.

Proof.

$$\widetilde{\mathbb{P}}(\Omega) = \int \frac{\widetilde{\lambda}}{\lambda} e^{-(\widetilde{\lambda}-\lambda)X} d\mathbb{P} = \frac{\widetilde{\lambda}}{\lambda} \int_0^\infty e^{-(\widetilde{\lambda}-\lambda)x} \lambda e^{-\lambda x} dx = \int_0^\infty \widetilde{\lambda} e^{-\widetilde{\lambda} x} dx = 1.$$

□

(ii) Compute the cumulative distribution function

$$\widetilde{\mathbb{P}}\{X \leq a\} \text{ for } a \geq 0$$

for the random variable X under the probability measure $\widetilde{\mathbb{P}}$.

Solution.

$$\widetilde{\mathbb{P}}(X \leq a) = \int_{\{X \leq a\}} \frac{\widetilde{\lambda}}{\lambda} e^{-(\widetilde{\lambda}-\lambda)X} d\mathbb{P} = \int_0^a \frac{\widetilde{\lambda}}{\lambda} e^{-(\widetilde{\lambda}-\lambda)x} \lambda e^{-\lambda x} dx = \int_0^a \widetilde{\lambda} e^{-\widetilde{\lambda} x} dx = 1 - e^{-\widetilde{\lambda} a}.$$

□

► **Exercise 1.15 (Provided by Alexander Ng).** Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and assume X has a density function $f(x)$ that is positive for every $x \in \mathbb{R}$. Let g be a strictly increasing, differentiable function satisfying

$$\lim_{y \to -\infty} g(y) = -\infty, \lim_{y \to \infty} g(y) = \infty,$$

and define the random variable $Y = g(X)$.

Let $h(y)$ be an arbitrary nonnegative function satisfying $\int_{-\infty}^{\infty} h(y) dy = 1$. We want to change the probability measure so that $h(y)$ is the density function for the random variable Y. To do this, we define

$$Z = \frac{h(g(X))g'(X)}{f(X)}.$$

(i) Show that Z is nonnegative and $\mathbb{E}Z = 1$. Now define $\widetilde{\mathbb{P}}$ by

$$\widetilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P} \text{ for all } A \in \mathcal{F}.$$

Proof. Clearly $Z \geq 0$. Furthermore, we have

$$\mathbb{E}\{Z\} = \mathbb{E}\left\{\frac{h(g(X))g'(X)}{f(X)}\right\} = \int_{-\infty}^{\infty} \frac{h(g(x))g'(x)}{f(x)} f(x) dx = \int_{-\infty}^{\infty} h(g(x)) dg(x) = \int_{-\infty}^{\infty} h(u) du = 1.$$

□

(ii) Show that Y has density h under $\widetilde{\mathbb{P}}$.

Proof.

$$\widetilde{\mathbb{P}}(Y \leq a) = \int_{\{g(X) \leq a\}} \frac{h(g(X))g'(X)}{f(X)} d\mathbb{P} = \int_{-\infty}^{g^{-1}(a)} \frac{h(g(x))g'(x)}{f(x)} f(x) dx = \int_{-\infty}^{g^{-1}(a)} h(g(x)) dg(x).$$

By the change of variable formula, the last equation above equals to $\int_{-\infty}^{a} h(u) du$. So Y has density h under $\widetilde{\mathbb{P}}$. □

2 Information and Conditioning

▶ **Exercise 2.1.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a general probability space, and suppose a random variable X on this space is measurable with respect to the trivial σ-algebra $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Show that X is not random (i.e., there is a constant c such that $X(\omega) = c$ for all $\omega \in \Omega$). Such a random variable is called *degenerate*.

Proof. For any real number a, we have $\{X \le a\} \in \mathcal{F}_0 = \{\emptyset, \Omega\}$. So $\mathbb{P}(X \le a)$ is either 0 or 1. Since $\lim_{a\to\infty} \mathbb{P}(X \le a) = 1$ and $\lim_{a\to-\infty} \mathbb{P}(X \le a) = 0$, we can find a number x_0 such that $\mathbb{P}(X \le x_0) = 1$ and $\mathbb{P}(X \le x) = 0$ for any $x < x_0$. So

$$\mathbb{P}(X = x_0) = \lim_{n\to\infty} \mathbb{P}\left(x_0 - \frac{1}{n} < X \le x_0\right) = \lim_{n\to\infty}\left(\mathbb{P}(X \le x_0) - \mathbb{P}\left(X \le x_0 - \frac{1}{n}\right)\right) = 1.$$

Since $\{X = x_0\} \in \{\emptyset, \Omega\}$, we conclude $\{X = x_0\} = \Omega$. □

▶ **Exercise 2.2.** Independence of random variables can be affected by changes of measure. To illustrate this point, consider the space of two coin tosses $\Omega_2 = \{HH, HT, TH, TT\}$, and let stock prices be given by

$$S_0 = 4, S_1(H) = 8, S_1(T) = 2,$$
$$S_2(HH) = 16, S_2(HT) = S_2(TH) = 4, S_2(TT) = 1.$$

Consider the two probability measures given by

$$\widetilde{\mathbb{P}}(HH) = \frac{1}{4}, \widetilde{\mathbb{P}}(HT) = \frac{1}{4}, \widetilde{\mathbb{P}}(TH) = \frac{1}{4}, \widetilde{\mathbb{P}}(TT) = \frac{1}{4},$$
$$\mathbb{P}(HH) = \frac{4}{9}, \mathbb{P}(HT) = \frac{2}{9}, \mathbb{P}(TH) = \frac{2}{9}, \mathbb{P}(TT) = \frac{1}{9}.$$

Define the random variable

$$X = \begin{cases} 1 & \text{if } S_2 = 4, \\ 0 & \text{if } S_2 \neq 4. \end{cases}$$

(i) List all the sets in $\sigma(X)$.

Solution. $\sigma(X) = \{\emptyset, \Omega, \{HT, TH\}, \{TT, HH\}\}$. □

(ii) List all the sets in $\sigma(S_1)$.

Solution. $\sigma(S_1) = \{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}\}$. □

(iii) Show that $\sigma(X)$ and $\sigma(S_1)$ are independent under the probability measure $\widetilde{\mathbb{P}}$.

Proof. $\widetilde{\mathbb{P}}(\{HT, TH\} \cap \{HH, HT\}) = \widetilde{\mathbb{P}}(\{HT\}) = \frac{1}{4}$, $\widetilde{\mathbb{P}}(\{HT, TH\}) = \widetilde{\mathbb{P}}(\{HT\}) + \widetilde{\mathbb{P}}(\{TH\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, and $\widetilde{\mathbb{P}}(\{HH, HT\}) = \widetilde{\mathbb{P}}(\{HH\}) + \widetilde{\mathbb{P}}(\{HT\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. So we have

$$\widetilde{\mathbb{P}}(\{HT, TH\} \cap \{HH, HT\}) = \widetilde{\mathbb{P}}(\{HT, TH\})\widetilde{\mathbb{P}}(\{HH, HT\}).$$

Similarly, we can work on other elements of $\sigma(X)$ and $\sigma(S_1)$ and show that $\widetilde{\mathbb{P}}(A \cap B) = \widetilde{\mathbb{P}}(A)\widetilde{\mathbb{P}}(B)$ for any $A \in \sigma(X)$ and $B \in \sigma(S_1)$. So $\sigma(X)$ and $\sigma(S_1)$ are independent under $\widetilde{\mathbb{P}}$. □

(iv) Show that $\sigma(X)$ and $\sigma(S_1)$ are not independent under the probability measure $\mathbb{P}$.

Proof. $\mathbb{P}(\{HT, TH\} \cap \{HH, HT\}) = \mathbb{P}(\{HT\}) = \frac{2}{9}$, $\mathbb{P}(\{HT, TH\}) = \frac{2}{9} + \frac{2}{9} = \frac{4}{9}$ and $\mathbb{P}(\{HH, HT\}) = \frac{4}{9} + \frac{2}{9} = \frac{6}{9}$. So

$$\mathbb{P}(\{HT, TH\} \cap \{HH, HT\}) \neq \mathbb{P}(\{HT, TH\})\mathbb{P}(\{HH, HT\}).$$

Hence $\sigma(X)$ and $\sigma(S_1)$ are not independent under $\mathbb{P}$. □

(v) Under $\mathbb{P}$, we have $\mathbb{P}\{S_1 = 8\} = \frac{2}{3}$ and $\mathbb{P}\{S_1 = 2\} = \frac{1}{3}$. Explain intuitively why, if you are told that $X = 1$, you would want to revise your estimate of the distribution of S_1.

Solution. Because S_1 and X are not independent under the probability measure $\mathbb{P}$, knowing the value of X will affect our opinion on the distribution of S_1. □

▶ **Exercise 2.3 (Rotating the axes).** Let X and Y be independent standard normal random variables. Let θ be a constant, and define random variables

$$V = X\cos\theta + Y\sin\theta \text{ and } W = -X\sin\theta + Y\cos\theta.$$

Show that V and W are independent standard normal random variables.

Proof. We note (V, W) are jointly Gaussian. In order to prove their independence, it suffices to show they are uncorrelated. Indeed,

$$\begin{aligned} \mathbb{E}[VW] &= \mathbb{E}[-X^2\sin\theta\cos\theta + XY\cos^2\theta - XY\sin^2\theta + Y^2\sin\theta\cos\theta] \\ &= -\sin\theta\cos\theta + 0 + 0 + \sin\theta\cos\theta \\ &= 0. \end{aligned}$$

□

▶ **Exercise 2.4.** In Example 2.2.8, X is a standard normal random variable and Z is an independent random variable satisfying

$$\mathbb{P}\{Z = 1\} = \mathbb{P}\{Z = -1\} = \frac{1}{2}.$$

We defined $Y = XZ$ and showed that Y is standard normal. We established that although X and Y are uncorrelated, they are not independent. In this exercise, we use moment-generating functions to show that Y is standard normal and X and Y are not independent.

(i) Establish the joint moment-generating function formula

$$\mathbb{E}e^{uX+vY} = e^{\frac{1}{2}(u^2+v^2)} \cdot \frac{e^{uv} + e^{-uv}}{2}.$$

Solution.

$$\begin{aligned} \mathbb{E}\left[e^{uX+vY}\right] &= \mathbb{E}\left[e^{uX+vXZ}\right] \\ &= \mathbb{E}\left[e^{uX+vXZ}|Z=1\right]P(Z=1) + \mathbb{E}\left[e^{uX+vXZ}|Z=-1\right]P(Z=-1) \\ &= \frac{1}{2}\mathbb{E}\left[e^{uX+vX}\right] + \frac{1}{2}E\left[e^{uX-vX}\right] \\ &= \frac{1}{2}\left[e^{\frac{(u+v)^2}{2}} + e^{\frac{(u-v)^2}{2}}\right] \\ &= e^{\frac{u^2+v^2}{2}} \cdot \frac{e^{uv} + e^{-uv}}{2}. \end{aligned}$$

□

(ii) Use the formula above to show that $\mathbb{E}e^{vY} = e^{\frac{1}{2}v^2}$. This is the moment-generating function for a standard normal random variable, and thus Y must be a standard normal random variable.

Proof. Let $u = 0$, we are done. □

(iii) Use the formula in (i) and Theorem 2.2.7(iv) to show that X and Y are not independent.

Proof. $\mathbb{E}\left[e^{uX}\right] = e^{\frac{u^2}{2}}$ and $\mathbb{E}\left[e^{vY}\right] = e^{\frac{v^2}{2}}$. So $\mathbb{E}\left[e^{uX+vY}\right] \neq \mathbb{E}\left[e^{uX}\right]\mathbb{E}\left[e^{vY}\right]$. Therefore X and Y cannot be independent. □

▶ **Exercise 2.5.** Let (X, Y) be a pair of random variables with joint density function

$$f_{X,Y}(x,y) = \begin{cases} \frac{2|x|+y}{\sqrt{2\pi}} \exp\left\{-\frac{(2|x|+y)^2}{2}\right\} & \text{if } y \geq -|x|, \\ 0 & \text{if } y < -|x|. \end{cases}$$

Show that X and Y are standard normal random variables and that they are uncorrelated but not independent.

Proof. The density $f_X(x)$ of X can be obtained by

$$f_X(x) = \int f_{X,Y}(x,y)dy = \int_{\{y\geq -|x|\}} \frac{2|x|+y}{\sqrt{2\pi}} e^{-\frac{(2|x|+y)^2}{2}} dy = \int_{\{\xi \geq |x|\}} \frac{\xi}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

The density $f_Y(y)$ of Y can be obtained by

$$\begin{aligned} f_Y(y) &= \int f_{XY}(x,y)dx \\ &= \int 1_{\{|x|\geq -y\}} \frac{2|x|+y}{\sqrt{2\pi}} e^{-\frac{(2|x|+y)^2}{2}} dx \\ &= \int_{0\vee(-y)}^{\infty} \frac{2x+y}{\sqrt{2\pi}} e^{-\frac{(2x+y)^2}{2}} dx + \int_{-\infty}^{0\wedge y} \frac{-2x+y}{\sqrt{2\pi}} e^{-\frac{(-2x+y)^2}{2}} dx \\ &= \int_{0\vee(-y)}^{\infty} \frac{2x+y}{\sqrt{2\pi}} e^{-\frac{(2x+y)^2}{2}} dx + \int_{\infty}^{0\vee(-y)} \frac{2x+y}{\sqrt{2\pi}} e^{-\frac{(2x+y)^2}{2}} d(-x) \\ &= 2\int_{|y|}^{\infty} \frac{\xi}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d(\frac{\xi}{2}) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}. \end{aligned}$$

So both X and Y are standard normal random variables. Since $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$, X and Y are not independent. However, if we set $F(t) = \int_t^\infty \frac{u^2}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$, we have

$$\begin{aligned} \mathbb{E}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy 1_{\{y\geq -|x|\}} \frac{2|x|+y}{\sqrt{2\pi}} e^{-\frac{(2|x|+y)^2}{2}} dxdy \\ &= \int_{-\infty}^{\infty} xdx \int_{-|x|}^{\infty} y \frac{2|x|+y}{\sqrt{2\pi}} e^{-\frac{(2|x|+y)^2}{2}} dy \\ &= \int_{-\infty}^{\infty} xdx \int_{|x|}^{\infty} (\xi - 2|x|) \frac{\xi}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi \\ &= \int_{-\infty}^{\infty} xdx \left(\int_{|x|}^{\infty} \frac{\xi^2}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi - 2|x| \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \right) \\ &= \int_0^\infty x \int_x^\infty \frac{\xi^2}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi dx + \int_{-\infty}^0 x \int_{-x}^\infty \frac{\xi^2}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi dx \\ &= \int_0^\infty xF(x)dx + \int_{-\infty}^0 xF(-x)dx \\ &= 0. \end{aligned}$$

□

▶ **Exercise 2.6.** Consider a probability space Ω with four elements, which we call a, b, c, and d (i.e., $\Omega = \{a, b, c, d\}$). The σ-algebra $\mathcal{F}$ is the collection of all subsets of Ω; i.e., the sets in $\mathcal{F}$ are

$$\begin{gathered}\Omega, \{a,b,c\}, \{a,b,d\}, \{a,c,d\}, \{b,c,d\}, \\ \{a,b\}, \{a,c\}, \{a,d\}, \{b,c\}, \{b,d\}, \{c,d\}, \\ \{a\}, \{b\}, \{c\}, \{d\}, \emptyset.\end{gathered}$$

We define a probability measure $\mathbb{P}$ by specifying that

$$\mathbb{P}\{a\} = \frac{1}{6}, \mathbb{P}\{b\} = \frac{1}{3}, \mathbb{P}\{c\} = \frac{1}{4}, \mathbb{P}\{d\} = \frac{1}{4},$$

and, as usual, the probability of every other set in $\mathcal{F}$ is the sum of the probabilities of the elements in the set, e.g., $\mathbb{P}\{a,b,c\} = \mathbb{P}\{a\} + \mathbb{P}\{b\} + \mathbb{P}\{c\} = \frac{3}{4}$.

We next define two random variables, X and Y, by the formulas

$$\begin{gathered}X(a) = 1, X(b) = 1, X(c) = -1, X(d) = -1, \\ Y(a) = 1, Y(b) = -1, Y(c) = 1, Y(d) = -1.\end{gathered}$$

We then define $Z = X + Y$.

(i) List the sets in $\sigma(X)$.

Solution. $\sigma(X) = \{\emptyset, \Omega, \{a,b\}, \{c,d\}\}$. □

(ii) Determine $\mathbb{E}[Y|X]$ (i.e., specify the values of this random variable for a, b, c, and d). Verify that the partial-averaging property is satisfied.

Solution.

$$\begin{aligned}
\mathbb{E}[Y|X] &= \sum_{\alpha \in \{1,-1\}} \mathbb{E}[Y|X=\alpha] 1_{\{X=\alpha\}} \\
&= \sum_{\alpha \in \{1,-1\}} \frac{\mathbb{E}[Y 1_{\{X=\alpha\}}]}{\mathbb{P}(X=\alpha)} 1_{\{X=\alpha\}} \\
&= \frac{1 \cdot \mathbb{P}(Y=1, X=1) - 1 \cdot \mathbb{P}(Y=-1, X=1)}{\mathbb{P}(X=1)} 1_{\{X=1\}} \\
&\quad + \frac{1 \cdot \mathbb{P}(Y=1, X=-1) - 1 \cdot \mathbb{P}(Y=-1, X=-1)}{\mathbb{P}(X=-1)} 1_{\{X=-1\}} \\
&= \frac{1 \cdot \mathbb{P}(\{a\}) - 1 \cdot \mathbb{P}(\{b\})}{\mathbb{P}(\{a,b\})} 1_{\{X=1\}} + \frac{1 \cdot \mathbb{P}(\{c\}) - 1 \cdot \mathbb{P}(\{d\})}{\mathbb{P}(\{c,d\})} 1_{\{X=-1\}} \\
&= -\frac{1}{3} 1_{\{X=1\}}.
\end{aligned}$$

To verify the partial-averaging property, we note

$$\mathbb{E}\left[\mathbb{E}[Y|X] 1_{\{X=1\}}\right] = -\frac{1}{3}\mathbb{E}\left[1_{\{X=1\}}\right] = -\frac{1}{3}\mathbb{P}(\{a,b\}) = -\frac{1}{6}$$

and

$$\mathbb{E}\left[Y 1_{\{X=1\}}\right] = \mathbb{P}(X=Y=1) - \mathbb{P}(Y=-1, X=1) = \mathbb{P}(\{a\}) - \mathbb{P}(\{b\}) = -\frac{1}{6}.$$

Similarly,

$$\mathbb{E}\left[\mathbb{E}[Y|X] 1_{\{X=-1\}}\right] = 0$$

and

$$\mathbb{E}\left[Y 1_{\{X=-1\}}\right] = \mathbb{P}(Y=1, X=-1) - \mathbb{P}(Y=-1, X=-1) = \mathbb{P}(\{c\}) - \mathbb{P}(\{d\}) = 0.$$

Together, we can conclude the partial-averaging property hodls. □

(iii) Determine $\mathbb{E}[Z|X]$. Again, verify the partial-averaging property.

Solution.

$$\mathbb{E}[Z|X] = X + \mathbb{E}[Y|X] = \frac{2}{3}1_{\{X=1\}} - 1_{\{X=-1\}}.$$

Verification of the partial-averaging property is skipped. □

(iv) Compute $\mathbb{E}[Z|X] - \mathbb{E}[Y|X]$. Citing the appropriate properties of conditional expectation from Theorem 2.3.2, explain why you get X.

Solution. $\mathbb{E}[Z|X] - \mathbb{E}[Y|X] = \mathbb{E}[Z - Y|X] = \mathbb{E}[X|X] = X$, where the first equality is due to linearity (Theorem 2.3.2(i)) and the last equality is due to "taking out what is known" (Theorem 2.3.2(ii)). □

► **Exercise 2.7.** Let Y be an integrable random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Based on the information in $\mathcal{G}$, we can form the estimate $\mathbb{E}[Y|\mathcal{G}]$ of Y and define the error of the estimation $\text{Err} = Y - \mathbb{E}[Y|\mathcal{G}]$. This is a random variable with expectation zero and some variance $\text{Var}(\text{Err})$. Let X be some other $\mathcal{G}$-measurable random variable, which we can regard as another estimate of Y. Show that

$$\text{Var}(\text{Err}) \leq \text{Var}(Y - X).$$

In other words, the estimate $\mathbb{E}[Y|\mathcal{G}]$ minimizes the variance of the error among all estimates based on the information in $\mathcal{G}$. (Hint: Let $\mu = \mathbb{E}(Y - X)$. Compute the variance of $Y - X$ as

$$\mathbb{E}[(Y - X - \mu)^2] = \mathbb{E}\left[\left((Y - \mathbb{E}[Y|\mathcal{G}]) + (\mathbb{E}[Y|\mathcal{G}] - X - \mu)\right)^2\right].$$

Multiply out the right-hand side and use iterated conditioning to show the cross-term is zero.)

Proof. Let $\mu = \mathbb{E}[Y - X]$ and $\xi = \mathbb{E}[Y - X - \mu|\mathcal{G}]$. Note ξ is $\mathcal{G}$-measurable, we have

$$\begin{aligned}
\text{Var}(Y - X) &= \mathbb{E}[(Y - X - \mu)^2] \\
&= \mathbb{E}\left\{[(Y - E[Y|\mathcal{G}]) + (\mathbb{E}[Y|\mathcal{G}] - X - \mu)]^2\right\} \\
&= \text{Var}(\text{Err}) + 2\mathbb{E}\left[(Y - \mathbb{E}[Y|\mathcal{G}])\xi\right] + E[\xi^2] \\
&= \text{Var}(\text{Err}) + 2\mathbb{E}\left[Y\xi - \mathbb{E}[Y\xi|\mathcal{G}]\right] + E[\xi^2] \\
&= \text{Var}(\text{Err}) + \mathbb{E}[\xi^2] \\
&\geq \text{Var}(\text{Err}).
\end{aligned}$$

□

► **Exercise 2.8.** Let X and Y be integrable random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $Y = Y_1 + Y_2$, where $Y_1 = \mathbb{E}[Y|X]$ is $\sigma(X)$-measurable and $Y_2 = Y - \mathbb{E}[Y|X]$. Show that Y_2 and X are uncorrelated. More generally, show that Y_2 is uncorrelated with every $\sigma(X)$-measurable random variable.

Proof. It suffices to prove the more general case. For any $\sigma(X)$-measurable random variable ξ, $\mathbb{E}[Y_2\xi] = \mathbb{E}[(Y - E\{Y|X\})\xi] = \mathbb{E}[Y\xi - E\{Y|X\}\xi] = \mathbb{E}[Y\xi] - \mathbb{E}[Y\xi] = 0$.

□

▲ **Exercise 2.9.** Let X be a random variable.

(i) Give an example of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a random variable X defined on this probability space, and a function f so that the σ-algebra generated by $f(X)$ is not the trivial σ-algebra $\{\emptyset, \Omega\}$ but is strictly smaller than the σ-algebra generated by X.

Solution. Consider the dice-toss space similar to the coin-toss space. Then a typical element ω in this space is an infinite sequence $\omega_1\omega_2\omega_3\cdots$, with $\omega_i \in \{1, 2, \cdots, 6\}$ $(i \in \mathbb{N})$. We define $X(\omega) = \omega_1$ and $f(x) = 1_{\{\text{odd integers}\}}(x)$. Then it's easy to see

$$\sigma(X) = \{\emptyset, \Omega, \{\omega : \omega_1 = 1\}, \cdots, \{\omega : \omega_1 = 6\}\}$$

and $\sigma(f(X))$ equals to

$$\{\emptyset, \Omega, \{\omega : \omega_1 = 1\} \cup \{\omega : \omega_1 = 3\} \cup \{\omega : \omega_1 = 5\}, \{\omega : \omega_1 = 2\} \cup \{\omega : \omega_1 = 4\} \cup \{\omega : \omega_1 = 6\}\}.$$

So $\{\emptyset, \Omega\} \subsetneq \sigma(f(X)) \subsetneq \sigma(X)$, and each of these containment is strict. □

(ii) Can the σ-algebra generated by $f(X)$ ever be strictly larger than the σ-algebra generated by X?

Solution. No. $\sigma(f(X)) \subset \sigma(X)$ is always true. □

► **Exercise 2.10.** Let X and Y be random variable (on some unspecified probability space $(\Omega, \mathcal{F}, \mathbb{P})$), assume they have a joint density $f_{X,Y}(x, y)$, and assume $\mathbb{E}|Y| < \infty$. In particular, for every Borel subset C of $\mathbb{R}^2$, we have

$$\mathbb{P}\{(X, Y) \in C\} = \int_C f_{X,Y}(x, y)dxdy.$$

In elementary probability, one learns to compute $\mathbb{E}[Y|X = x]$, which is a *nonrandom* function of the *dummy variable* x, by the formula

$$\mathbb{E}[Y|X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy, \tag{2.6.1}$$

where $f_{Y|X}(y|x)$ is the *conditional density* defined by

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)}.$$

The denominator in this expression, $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, \eta)d\eta$, is the *marginal density* of X, and we must assume it is strictly positive for every x. We introduce the symbol $g(x)$ for the function $\mathbb{E}[Y|X = x]$ defined by (2.6.1); i.e.,

$$g(x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy = \int_{-\infty}^{\infty} \frac{y f_{X,Y}(x, y)}{f_X(x)}dy$$

In measure-theoretic probability, conditional expectation is a *random variable* $\mathbb{E}[Y|X]$. This exercise is to show that when there is a joint density for (X, Y), this random variable can be obtained by substituting the random variable X in place of the dummy variable x in the function $g(x)$. In other words, this exercise is to show that

$$\mathbb{E}[Y|X] = g(X).$$

(We introduced the symbol $g(x)$ in order to avoid the mathematically confusing expression $\mathbb{E}[Y|X = X]$.)

Since $g(X)$ is obviously $\sigma(X)$-measurable, to verify that $\mathbb{E}[Y|X] = g(X)$, we need only check that the partial-averaging property is satisfied. For every Borel-measurable function h mapping $\mathbb{R}$ to $\mathbb{R}$ and satisfying $\mathbb{E}|h(X)| < \infty$, we have

$$\mathbb{E}h(X) = \int_{-\infty}^{\infty} h(x)f_X(x)dx. \tag{2.6.2}$$

This is Theorem 1.5.2 in Chapter 1. Similarly, if h is a function of both x and y, then

$$\mathbb{E}h(X, Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x, y)f_{X,Y}(x, y)dxdy \tag{2.6.3}$$

whenever (X, Y) has a joint density $f_{X,Y}(x, y)$. You may use both (2.6.2) and (2.6.3) in your solution to this problem.

Let A be a set in $\sigma(X)$. By the definition of $\sigma(X)$, there is a Borel subset B of $\mathbb{R}$ such that $A = \{\omega \in \Omega; X(\omega) \in B\}$ or, more simply, $A = \{X \in B\}$. Show the partial-averaging property

$$\int_A g(X) d\mathbb{P} = \int_A Y d\mathbb{P}.$$

Proof.

$$\begin{aligned}
\int_A g(X) dP &= \mathbb{E}[g(X)1_B(X)] = \int_{-\infty}^{\infty} g(x)1_B(x)f_X(x)dx = \int\int \frac{yf_{X,Y}(x,y)}{f_X(x)}dy 1_B(x)f_X(x)dx \\
&= \int\int y1_B(x)f_{X,Y}(x,y)dxdy = \mathbb{E}[Y1_B(X)] = \mathbb{E}[YI_A] = \int_A Y d\mathbb{P}.
\end{aligned}$$

□

▶ **Exercise 2.11.**
(i) Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let W be a nonnegative $\sigma(X)$-measurable random variable. Show there exists a function g such that $W = g(X)$. (Hint: Recall that every set in $\sigma(X)$ is of the form $\{X \in B\}$ for some Borel set $B \subset \mathbb{R}$. Suppose first that W is the indicator of such a set, and then use the standard machine.)

Proof. We can find a sequence $\{W_n\}_{n\geq 1}$ of $\sigma(X)$-measurable simple functions such that $W_n \uparrow W$. Each W_n can be written as the form $\sum_{i=1}^{K_n} a_i^n 1_{A_i^n}$, where A_i^n's belong to $\sigma(X)$ and are disjoint. So each A_i^n can be written as $\{X \in B_i^n\}$ for some Borel subset B_i^n of $\mathbb{R}$, i.e. $W_n = \sum_{i=1}^{K_n} a_i^n 1_{\{X\in B_i^n\}} = \sum_{i=1}^{K_n} a_i^n 1_{B_i^n}(X) = g_n(X)$, where $g_n(x) = \sum_{i=1}^{K_n} a_i^n 1_{B_i^n}(x)$. Define $g = \limsup g_n$, then g is a Borel function. By taking upper limits on both sides of $W_n = g_n(X)$, we get $W = g(X)$. □

(ii) Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let Y be a nonnegative random variable on this space. We do not assume that X and Y have a joint density. Nonetheless, show there is a function g such that $\mathbb{E}[Y|X] = g(X)$.

Proof. Note $\mathbb{E}[Y|X]$ is $\sigma(X)$-measurable. By (i), we can find a Borel function g such that $\mathbb{E}[Y|X] = g(X)$. □

3 Brownian Motion

▶ **Exercise 3.1.** According to Definition 3.3.3(iii), for $0 \leq t < u$, the Brownian motion increment $W(u) - W(t)$ is independent of the σ-algebra $\mathcal{F}(t)$. Use this property and property (i) of that definition to show that, for $0 \leq t < u_1 < u_2$, the increment $W(u_2) - W(u_1)$ is also indepdent of $\mathcal{F}_t$.

Proof. We have $\mathcal{F}(t) \subset \mathcal{F}(u_1)$ and $W(u_2) - W(u_1)$ is independent of $\mathcal{F}(u_1)$. So in particular, $W(u_2) - W(u_1)$ is independent of $\mathcal{F}(t)$. □

▶ **Exercise 3.2.** Let $W(t)$, $t \geq 0$, be a Brownian motion, and let $\mathcal{F}(t)$, $t \geq 0$, be a filtration for this Brownian motion. Show that $W^2(t) - t$ is a martingale. (Hint: For $0 \leq s \leq t$, write $W^2(t)$ as $(W(t) - W(s))^2 + 2W(t)W(s) - W^2(s)$.)

Proof. $\mathbb{E}[W^2(t) - W^2(s)|\mathcal{F}(s)] = \mathbb{E}[(W(t) - W(s))^2 + 2W(t)W(s) - 2W^2(s)|\mathcal{F}(s)] = t - s + 2W(s)\mathbb{E}[W(t) - W(s)|\mathcal{F}(s)] = t - s$. Simple algebra gives $\mathbb{E}[W^2(t) - t|\mathcal{F}(s)] = W^2(s) - s$. So $W^2(t) - t$ is a martingale. □

▶ **Exercise 3.3 (Normal kurtosis).** The *kurtosis* of a random variable is defined to be the ratio of its fourth central moment to the square of its variance. For a normal random variable, the kurtosis is 3. This fact was used to obtain (3.4.7). This exercise verifies this fact.

Let X be a normal random variable with mean μ, so that $X - \mu$ has mean zero. Let the variance of X, which is also the variance of $X - \mu$, be σ^2. In (3.2.13), we computed the moment-generating function of

$X-\mu$ to be $\varphi(u) = \mathbb{E}e^{u(X-\mu)} = e^{\frac{1}{2}u^2\sigma^2}$, where u is a real variable. Differentiating this function with respect to u, we obtain

$$\varphi'(u) = \mathbb{E}\left[(X-\mu)e^{u(X-\mu)}\right] = \sigma^2 u e^{\frac{1}{2}\sigma^2 u^2}$$

and, in particular, $\varphi'(0) = \mathbb{E}(X-\mu) = 0$. Differentiating again, we obtain

$$\varphi''(u) = \mathbb{E}\left[(X-\mu)^2 e^{u(X-\mu)}\right] = (\sigma^2 + \sigma^4 u^2)e^{\frac{1}{2}\sigma^2 u^2}$$

and, in particular, $\varphi''(0) = \mathbb{E}[(X-\mu)^2] = \sigma^2$. Differentiate two more times and obtain the normal kurtosis formula $\mathbb{E}[(X-\mu)^4] = 3\sigma^4$.

Solution.

$$\varphi^{(3)}(u) = 2\sigma^4 u e^{\frac{1}{2}\sigma^2 u^2} + (\sigma^2 + \sigma^4 u^2)\sigma^2 u e^{\frac{1}{2}\sigma^2 u^2} = e^{\frac{1}{2}\sigma^2 u^2}(3\sigma^4 u + \sigma^6 u^3),$$

and

$$\varphi^{(4)}(u) = \sigma^2 u e^{\frac{1}{2}\sigma^2 u^2}(3\sigma^4 u + \sigma^6 u^3) + e^{\frac{1}{2}\sigma^2 u^2}(3\sigma^4 + 3\sigma^6 u^2) = e^{\frac{1}{2}\sigma^2 u^2}(6\sigma^6 u^2 + \sigma^8 u^4 + 3\sigma^4).$$

So $\mathbb{E}[(X-\mu)^4] = \varphi^{(4)}(0) = 3\sigma^4$. □

▶ **Exercise 3.4 (Other variations of Brownian motion).** Theorem 3.4.3 asserts that if T is a positive number and we choose a partition Π with points $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$, then as the number n of partition points approaches infinity and the length of the longest subinterval $\|\Pi\|$ approaches zero, the sample quadratic variation

$$\sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2$$

approaches T for almost every path of the Brownian motion W. In Remark 3.4.5, we further showed that $\sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))(t_{j+1} - t_j)$ and $\sum_{j=0}^{n-1}(t_{j+1} - t_j)^2$ have limit zero. We summarize these facts by the multiplication rules

$$dW(t)dW(t) = dt, dW(t)dt = 0, dtdt = 0. \tag{3.10.1}$$

(i) Show that as the number m of partition points approaches infinity and the length of the longest subinterval approaches zero, the sample first variation

$$\sum_{j=0}^{n-1}|W(t_{j+1}) - W(t_j)|$$

approaches ∞ for almost every path of the Brownian motion W. (Hint:

$$\sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2 \le \max_{0\le k\le n-1}|W(t_{k+1}) - W(t_k)| \cdot \sum_{j=0}^{n-1}|W(t_{j+1}) - W(t_j)|.)$$

Proof. Assume there exists $A \in \mathcal{F}$, such that $\mathbb{P}(A) > 0$ and for every $\omega \in A$,

$$\limsup_n \sum_{j=0}^{n-1}|W_{t_{j+1}} - W_{t_j}|(\omega) < \infty.$$

Then for every $\omega \in A$,

$$\begin{aligned}
\sum_{j=0}^{n-1}(W_{t_{j+1}} - W_{t_j})^2(\omega) &\le \max_{0\le k\le n-1}|W_{t_{k+1}} - W_{t_k}|(\omega) \cdot \sum_{j=0}^{n-1}|W_{t_{j+1}} - W_{t_j}|(\omega) \\
&\le \max_{0\le k\le n-1}|W_{t_{k+1}} - W_{t_k}|(\omega) \cdot \limsup_n \sum_{j=0}^{n-1}|W_{t_{j+1}} - W_{t_j}|(\omega) \\
&\to 0,
\end{aligned}$$

since by uniform continuity of continuous functions over a closed interval, $\lim_{n\to\infty}\max_{0\le k\le n-1}|W_{t_{k+1}} - W_{t_k}|(\omega) = 0$. This is a contradiction with $\lim_{n\to\infty}\sum_{j=0}^{n-1}(W_{t_{j+1}} - W_{t_j})^2 = T$ a.s.. □

(ii) Show that as the number n of partition points approaches infinity and the length of the longest subinterval approaches zero, the sample cubic variation

$$\sum_{j=0}^{n-1} |W(t_{j+1}) - W(t_j)|^3$$

approaches zero for almost every path of the Brownian motion W.

Proof. We note by an argument similar to (i),

$$\sum_{j=0}^{n-1} |W(t_{j+1}) - W(t_j)|^3 \le \max_{0\le k\le n-1} |W(t_{k+1}) - W(t_k)| \cdot \sum_{j=0}^{n-1} (W(t_{j+1}) - W(t_j))^2 \to 0$$

as $n \to \infty$. □

▶ **Exercise 3.5 (Black-Scholes-Merton formula).** Let the interest rate r and the volatility $\sigma > 0$ be constant. Let

$$S(t) = S(0)e^{\left(r-\frac{1}{2}\sigma^2\right)t+\sigma W(t)}$$

be a geometric Brownian motion with mean rate of return r, where the initial stock price $S(0)$ is positive. Let K be a positive constant. Show that, for $T > 0$,

$$\mathbb{E}\left[e^{-rT}(S(T) - K)^+\right] = S(0)N(d_+(T, S(0))) - Ke^{-rT}N(d_-(T, S(0))),$$

where

$$d_\pm(T, S(0)) = \frac{1}{\sigma\sqrt{T}}\left[\log\frac{S(0)}{K} + \left(r \pm \frac{\sigma^2}{2}\right)T\right],$$

and N is the cumulative standard normal distribution function

$$N(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{1}{2}z^2}dz = \frac{1}{\sqrt{2\pi}}\int_{-y}^{\infty} e^{-\frac{1}{2}z^2}dz.$$

Proof.

$$\begin{aligned}
& \mathbb{E}\left[e^{-rT}(S_T - K)^+\right] \\
= \quad & e^{-rT}\int_{\frac{1}{\sigma}\left[\ln\frac{K}{S_0}-(r-\frac{1}{2}\sigma^2)T\right]}^{\infty}\left(S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma x} - K\right)\frac{e^{-\frac{x^2}{2T}}}{\sqrt{2\pi T}}dx \\
= \quad & e^{-rT}\int_{\frac{1}{\sigma\sqrt{T}}\left[\ln\frac{K}{S_0}-(r-\frac{1}{2}\sigma^2)T\right]}^{\infty}\left(S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}y} - K\right)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}dy \\
= \quad & S_0 e^{-\frac{1}{2}\sigma^2 T}\int_{\frac{1}{\sigma\sqrt{T}}\left[\ln\frac{K}{S_0}-(r-\frac{1}{2}\sigma^2)T\right]}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}+\sigma\sqrt{T}y}dy - Ke^{-rT}\int_{\frac{1}{\sigma\sqrt{T}}\left[\ln\frac{K}{S_0}-(r-\frac{1}{2}\sigma^2)T\right]}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy \\
= \quad & S_0\int_{\frac{1}{\sigma\sqrt{T}}\left[\ln\frac{K}{S_0}-(r-\frac{1}{2}\sigma^2)T\right]-\sigma\sqrt{T}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{\xi^2}{2}}d\xi - Ke^{-rT}N\left(\frac{1}{\sigma\sqrt{T}}\left(\ln\frac{S_0}{K} + (r - \frac{1}{2}\sigma^2)T\right)\right) \\
= \quad & Ke^{-rT}N(d_+(T, S_0)) - Ke^{-rT}N(d_-(T, S_0)).
\end{aligned}$$

□

▶ **Exercise 3.6.** Let $W(t)$ be a Brownian motion and let $\mathcal{F}(t)$, $t \ge 0$, be an associated filtration.

(i) For $\mu \in \mathbb{R}$, consider the *Brownian motion with drift* μ:

$$X(t) = \mu t + W(t).$$

Show that for any Borel-measurable function $f(y)$, and for any $0 \le s < t$, the function

$$g(x) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{\infty} f(y) \exp\left\{ -\frac{(y-x-\mu(t-s))^2}{2(t-s)} \right\} dy$$

satisfies $\mathbb{E}[f(X(t))|\mathcal{F}(s)] = g(X(s))$, and hence X has the Markov property. We may rewrite $g(x)$ as $g(x) = \int_{-\infty}^{\infty} f(y)p(\tau, x, y)dy$, where $\tau = t - s$ and

$$p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}} \exp\left\{ -\frac{(y-x-\mu\tau)^2}{2\tau} \right\}$$

is the *transition density* for Brownian motion with drift μ.

Proof.

$$\begin{aligned}
\mathbb{E}[f(X_t)|\mathcal{F}_t] &= \mathbb{E}[f(W_t - W_s + a)|\mathcal{F}_s]|_{a=W_s+\mu t} \\
&= \mathbb{E}[f(W_{t-s} + a)]|_{a=W_s+\mu t} \\
&= \int_{-\infty}^{\infty} f(x + W_s + \mu t) \frac{e^{-\frac{x^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dx \\
&= \int_{-\infty}^{\infty} f(y) \frac{e^{-\frac{(y-W_s-\mu s-\mu(t-s))^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dy \\
&= g(X_s).
\end{aligned}$$

So $\mathbb{E}[f(X_t)|\mathcal{F}_s] = \int_{-\infty}^{\infty} f(y)p(t-s, X_s, y)dy$ with $p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{(y-x-\mu\tau)^2}{2\tau}}$. □

(ii) For $\nu \in \mathbb{R}$ and $\sigma > 0$, consider the *geometric Brownian motion*

$$S(t) = S(0)e^{\sigma W(t) + \nu t}.$$

Set $\tau = t - s$ and

$$p(\tau, x, y) = \frac{1}{\sigma y \sqrt{2\pi\tau}} \exp\left\{ -\frac{\left(\log\frac{y}{x} - \nu\tau\right)^2}{2\sigma^2\tau} \right\}.$$

Show that for any Borel-measurable function $f(y)$ and for any $0 \le s < t$ the function

$$g(x) = \int_0^{\infty} f(y)p(\tau, x, y)dy^{1}$$

satisfies $\mathbb{E}[f(S(t))|\mathcal{F}(s)] = g(S(s))$ and hence S has the Markov property and $p(\tau, x, y)$ is its transition density.

Proof. $\mathbb{E}[f(S_t)|\mathcal{F}_s] = \mathbb{E}[f(S_0 e^{\sigma X_t})|\mathcal{F}_s]$ with $\mu = \frac{v}{\sigma}$. So by (i),

$$\begin{aligned}
\mathbb{E}[f(S_t)|\mathcal{F}_s] &= \int_{-\infty}^{\infty} f(S_0 e^{\sigma y}) \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{[y-X_s-\mu(t-s)]^2}{2(t-s)}} dy \\
&\overset{S_0 e^{\sigma y}=z}{=} \int_0^{\infty} f(z) \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{\left[\frac{1}{\sigma}\ln\frac{z}{S_0} - \frac{1}{\sigma}\ln\frac{S_s}{S_0} - \mu(t-s)\right]^2}{2}} \frac{dz}{\sigma z} \\
&= \int_0^{\infty} f(z) \frac{e^{-\frac{\left[\ln\frac{z}{S_s} - v(t-s)\right]^2}{2\sigma^2(t-s)}}}{\sigma z\sqrt{2\pi(t-s)}} dz \\
&= \int_0^{\infty} f(z)p(t-s, S_s, z)dz \\
&= g(S_s).
\end{aligned}$$

□

[1] The textbook wrote $\int_0^{\infty} h(y)p(\tau, x, y)dy$ by mistake.

► **Exercise 3.7.** Theorem 3.6.2 provides the Laplace transform of the density of the first passage time for Brownian motion. This problem derives the analogous formula for Brownian motions with drift. Let W be a Brownian motion. Fix $m > 0$ and $\mu \in \mathbb{R}$. For $0 \leq t < \infty$, define

$$\begin{aligned} X(t) &= \mu t + W(t), \\ \tau_m = \min\{t \geq 0; X(t) = m\}. \end{aligned}$$

As usual, we set $\tau_m = \infty$ if $X(t)$ never reaches the level m. Let σ be a positive number and set

$$Z(t) = \exp\left\{\sigma X(t) - \left(\sigma\mu + \frac{1}{2}\sigma^2\right) t\right\}.$$

(i) Show that $Z(t)$, $t \geq 0$, is a martingale.

Proof.

$$\begin{aligned}
& \mathbb{E}\left[\frac{Z_t}{Z_s}\middle| \mathcal{F}_s\right] \\
= & \mathbb{E}\left[\exp\left\{\sigma(W_t - W_s) + \sigma\mu(t-s) - \left(\sigma\mu + \frac{\sigma^2}{2}\right)(t-s)\right\}\right] \\
= & \exp\left\{\sigma\mu(t-s) - \left(\sigma\mu + \frac{\sigma^2}{2}\right)(t-s)\right\} \cdot \int_{-\infty}^{\infty} e^{\sigma\sqrt{t-s}\cdot x} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx \\
= & \exp\left\{-\frac{\sigma^2}{2}(t-s)\right\} \cdot \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\sigma\sqrt{t-s})^2}{2}}}{\sqrt{2\pi}} dx \cdot \exp\left\{\frac{\sigma^2}{2}(t-s)\right\} \\
= & 1.
\end{aligned}$$

□

(ii) Use (i) to conclude that

$$\mathbb{E}\left[\exp\left\{\sigma X(t\wedge\tau_m) - \left(\sigma\mu + \frac{1}{2}\sigma^2\right)(t\wedge\tau_m)\right\}\right] = 1,\ t \geq 0.$$

Proof. By optional stopping theorem, $\mathbb{E}[Z_{t\wedge\tau_m}] = \mathbb{E}[Z_0] = 1$, that is,

$$\mathbb{E}\left[\exp\left\{\sigma X_{t\wedge\tau_m} - \left(\sigma\mu + \frac{\sigma^2}{2}\right) t\wedge\tau_m\right\}\right] = 1.$$

□

(iii) Now suppose $\mu \geq 0$. Show that, for $\sigma > 0$,

$$\mathbb{E}\left[\exp\left\{\sigma m - \left(\sigma\mu + \frac{1}{2}\sigma^2\right)\tau_m\right\} 1_{\{\tau_m < \infty\}}\right] = 1.$$

Proof. If $\mu \geq 0$ and $\sigma > 0$, $Z_{t\wedge\tau_m} \leq e^{\sigma m}$. By bounded convergence theorem,

$$\mathbb{E}[1_{\{\tau_m<\infty\}} Z_{\tau_m}] = \mathbb{E}[\lim_{t\to\infty} Z_{t\wedge\tau_m}] = \lim_{t\to\infty} \mathbb{E}[Z_{t\wedge\tau_m}] = 1,$$

since on the event $\{\tau_m = \infty\}$, $Z_{t\wedge\tau_m} \leq e^{\sigma m - \frac{1}{2}\sigma^2 t} \to 0$ as $t \to \infty$. Therefore, $\mathbb{E}\left[e^{\sigma m - (\sigma\mu + \frac{\sigma^2}{2})\tau_m} 1_{\{\tau_m<\infty\}}\right] = 1$. Let $\sigma \downarrow 0$, by bounded convergence theorem, we have $\mathbb{P}(\tau_m < \infty) = 1$. Let $\sigma\mu + \frac{\sigma^2}{2} = \alpha$, we get

$$\mathbb{E}[e^{-\alpha\tau_m}] = e^{-\sigma m} = e^{m\mu - m\sqrt{2\alpha+\mu^2}}.$$

□

(iv) Show that if $\mu > 0$, then $\mathbb{E}\tau_m < \infty$. Obtain a formula for $\mathbb{E}\tau_m$. (Hint: Differentiate the formula in (iii) with respect to α.)

Proof. We note for $\alpha > 0$, $\mathbb{E}[\tau_m e^{-\alpha\tau_m}] < \infty$ since $xe^{-\alpha x}$ is bounded on $[0, \infty)$. So by an argument similar to Exercise 1.8, $\mathbb{E}[e^{-\alpha\tau_m}]$ is differentiable and

$$\frac{\partial}{\partial\alpha}\mathbb{E}[e^{-\alpha\tau_m}] = -\mathbb{E}[\tau_m e^{-\alpha\tau_m}] = e^{m\mu - m\sqrt{2\alpha+\mu^2}}\frac{-m}{\sqrt{2\alpha+\mu^2}}.$$

Let $\alpha \downarrow 0$, by monotone increasing theorem, $\mathbb{E}[\tau_m] = \frac{m}{\mu} < \infty$ for $\mu > 0$. □

(v) Now suppose $\mu < 0$. Show that, for $\sigma > -2\mu$,

$$\mathbb{E}\left[\exp\left\{\sigma m - \left(\sigma\mu + \frac{1}{2}\sigma^2\right)\tau_m\right\}1_{\{\tau_m<\infty\}}\right] = 1.$$

Use this fact to show that $\mathbb{P}\{\tau_m < \infty\} = e^{-2m|\mu|}$,[2] which is strictly less than one, and to obtain the Laplace transform

$$\mathbb{E}e^{-\alpha\tau_m} = e^{m\mu - m\sqrt{2\alpha+\mu^2}} \text{ for all } \alpha > 0.$$

Proof. By $\sigma > -2\mu > 0$, we get $\sigma\mu + \frac{\sigma^2}{2} > 0$. Then $Z_{t\wedge\tau_m} \le e^{\sigma m}$ and on the event $\{\tau_m = \infty\}$, $Z_{t\wedge\tau_m} \le e^{\sigma m - (\frac{\sigma^2}{2}+\sigma\mu)t} \to 0$ as $t \to \infty$. Therefore,

$$\mathbb{E}\left[e^{\sigma m - (\sigma\mu + \frac{\sigma^2}{2})\tau_m}1_{\{\tau_m<\infty\}}\right] = \mathbb{E}[\lim_{t\to\infty} Z_{t\wedge\tau_m}] = \lim_{t\to\infty}\mathbb{E}[Z_{t\wedge\tau_m}] = 1.$$

Let $\sigma \downarrow -2\mu$, we then get $\mathbb{P}(\tau_m < \infty) = e^{2\mu m} = e^{-2|\mu|m} < 1$. Set $\alpha = \sigma\mu + \frac{\sigma^2}{2} > 0$ and we have

$$\mathbb{E}[e^{-\alpha\tau_m}] = \mathbb{E}[e^{-\alpha\tau_m}1_{\{\tau_m<\infty\}}] = e^{-\sigma m} = e^{m\mu - m\sqrt{2\alpha+\mu^2}}.$$

□

► **Exercise 3.8.** This problem presents the convergence of the distribution of stock prices in a sequence of binomial models to the distribution of geometric Brownian motion. In contrast to the analysis of Subsection 3.2.7, here we allow the interest rate to be different from zero.

Let $\sigma > 0$ and $r \ge 0$ be given. For each positive integer n, we consider a binomial model taking n steps per unit time. In this model, the interest rate per period is $\frac{r}{n}$, the up factor is $u_n = e^{\sigma/\sqrt{n}}$, and the down factor is $d_n = e^{-\sigma/\sqrt{n}}$. The risk-neutral probabilities are then

$$\widetilde{p}_n = \frac{\frac{r}{n} + 1 - e^{-\sigma/\sqrt{n}}}{e^{\sigma/\sqrt{n}} - e^{-\sigma/\sqrt{n}}},\ \widetilde{q}_n = \frac{e^{\sigma/\sqrt{n}} - \frac{r}{n} - 1}{e^{\sigma/\sqrt{n}} - e^{-\sigma/\sqrt{n}}}.$$

Let t be an arbitrary positive rational number, and for each positive integer n for which nt is an integer, define

$$M_{nt,n} = \sum_{k=1}^{nt} X_{k,n},$$

where $X_{1,n}, \cdots, X_{nt,n}$[3] are independent, identically distributed random variables with

$$\widetilde{\mathbb{P}}\{X_{k,n} = 1\} = \widetilde{p}_n,\ \widetilde{\mathbb{P}}\{X_{k,n} = -1\} = \widetilde{q}_n, k = 1, \cdots, nt^4.$$

[2]The textbook wrote $e^{-2x|\mu|}$ by mistake.
[3]The textbook wrote $X_{n,n}$ by mistake.
[4]The textbook wrote n by mistake.

The stock price at time t in this binomial model, which is the result of nt steps from the initial time, is given by (see (3.2.15) for a similar equation)

$$
\begin{aligned}
S_n(t) &= S(0)u_n^{\frac{1}{2}(nt+M_{nt,n})}d_n^{\frac{1}{2}(nt-M_{nt,n})} \\
&= S(0)\exp\left\{\frac{\sigma}{2\sqrt{n}}(nt+M_{nt,n})\right\}^5\exp\left\{-\frac{\sigma}{2\sqrt{n}}(nt-M_{nt,n})\right\} \\
&= S(0)\exp\left\{\frac{\sigma}{\sqrt{n}}M_{nt,n}\right\}.
\end{aligned}
$$

This problem shows that as $n \to \infty$, the distribution of the sequence of random variables $\frac{\sigma}{\sqrt{n}}M_{nt,n}$ appearing in the exponent above converges to the normal distribution with mean $\left(r - \frac{1}{2}\sigma^2\right)t$ and variance $\sigma^2 t$. Therefore, the limiting distribution of $S_n(t)$ is the same as the distribution of the geometric Brownian motion $S(0)\exp\{\sigma W(t) + (r - \frac{1}{2}\sigma)t\}$ at time t.

(i) Show that the moment-generating function $\varphi_n(u)$ of $\frac{1}{\sqrt{n}}M_{nt,n}$ is given by

$$
\varphi_n(u) = \left[e^{\frac{u}{\sqrt{n}}}\left(\frac{\frac{r}{n}+1-e^{-\sigma/\sqrt{n}}}{e^{\sigma/\sqrt{n}}-e^{-\sigma/\sqrt{n}}}\right) - e^{-\frac{u}{\sqrt{n}}}\left(\frac{\frac{r}{n}+1-e^{\sigma/\sqrt{n}}}{e^{\sigma/\sqrt{n}}-e^{-\sigma/\sqrt{n}}}\right)\right]^{nt}.
$$

Proof.

$$
\begin{aligned}
\varphi_n(u) &= \widetilde{\mathbb{E}}\left[e^{u\frac{1}{\sqrt{n}}M_{nt,n}}\right] = \left(\widetilde{\mathbb{E}}\left[e^{\frac{u}{\sqrt{n}}X_{1,n}}\right]\right)^{nt} = (e^{\frac{u}{\sqrt{n}}}\widetilde{p}_n + e^{-\frac{u}{\sqrt{n}}}\widetilde{q}_n)^{nt} \\
&= \left[e^{\frac{u}{\sqrt{n}}}\left(\frac{\frac{r}{n}+1-e^{-\frac{\sigma}{\sqrt{n}}}}{e^{\frac{\sigma}{\sqrt{n}}}-e^{-\frac{\sigma}{\sqrt{n}}}}\right) + e^{-\frac{u}{\sqrt{n}}}\left(\frac{-\frac{r}{n}-1+e^{\frac{\sigma}{\sqrt{n}}}}{e^{\frac{\sigma}{\sqrt{n}}}-e^{-\frac{\sigma}{\sqrt{n}}}}\right)\right]^{nt}.
\end{aligned}
$$

□

(ii) We want to compute

$$
\lim_{n\to\infty}\varphi_n(u) = \lim_{x\downarrow 0}\varphi_{\frac{1}{x^2}}(u),
$$

where we have made the change of variable $x = \frac{1}{\sqrt{n}}$. To do this, we will compute $\log\varphi_{\frac{1}{x^2}}(u)$ and then take the limit as $x \downarrow 0$. Show that

$$
\log\varphi_{\frac{1}{x^2}}(u) = \frac{t}{x^2}\log\left[\frac{(rx^2+1)\sinh ux + \sinh(\sigma-u)x}{\sinh\sigma x}\right]
$$

(the definitions are $\sinh z = \frac{e^z-e^{-z}}{2}$, $\cosh z = \frac{e^z+e^{-z}}{2}$), and use the formula

$$
\sinh(A-B) = \sinh A\cosh B - \cosh A\sinh B
$$

to rewrite this as

$$
\log\varphi_{\frac{1}{x^2}}(u) = \frac{t}{x^2}\log\left[\cosh ux + \frac{(rx^2+1-\cosh\sigma x)\sinh ux}{\sinh\sigma x}\right].
$$

Proof.

$$
\varphi_{\frac{1}{x^2}}(u) = \left[e^{ux}\left(\frac{rx^2+1-e^{-\sigma x}}{e^{\sigma x}-e^{-\sigma x}}\right) - e^{-ux}\left(\frac{rx^2+1-e^{\sigma x}}{e^{\sigma x}-e^{-\sigma x}}\right)\right]^{\frac{t}{x^2}}.
$$

So,

$$\begin{aligned}
\log \varphi_{\frac{1}{x^2}}(u) &= \frac{t}{x^2}\log\left[\frac{(rx^2+1)(e^{ux}-e^{-ux})+e^{(\sigma-u)x}-e^{-(\sigma-u)x}}{e^{\sigma x}-e^{-\sigma x}}\right] \\
&= \frac{t}{x^2}\log\left[\frac{(rx^2+1)\sinh ux+\sinh(\sigma-u)x}{\sinh \sigma x}\right] \\
&= \frac{t}{x^2}\log\left[\frac{(rx^2+1)\sinh ux+\sinh \sigma x\cosh ux-\cosh\sigma x\sinh ux}{\sinh \sigma x}\right] \\
&= \frac{t}{x^2}\log\left[\cosh ux+\frac{(rx^2+1-\cosh\sigma x)\sinh ux}{\sinh\sigma x}\right].
\end{aligned}$$

□

(iii) Use the Taylor series expansions

$$\cosh z = 1+\frac{1}{2}z^2+O(z^4), \sinh z = z+O(z^3),$$

to show that

$$\cosh ux+\frac{(rx^2+1-\cosh\sigma x)\sinh ux}{\sinh\sigma x} = 1+\frac{1}{2}u^2x^2+\frac{rux^2}{\sigma}-\frac{1}{2}ux^2\sigma+O(x^4). \tag{3.10.2}$$

The notation $O(x^j)$ is used to represent terms of the order x^j.

Proof.

$$\begin{aligned}
& \cosh ux+\frac{(rx^2+1-\cosh\sigma x)\sinh ux}{\sinh\sigma x} \\
=\ & 1+\frac{u^2x^2}{2}+O(x^4)+\frac{(rx^2+1-1-\frac{\sigma^2x^2}{2}+O(x^4))(ux+O(x^3))}{\sigma x+O(x^3)} \\
=\ & 1+\frac{u^2x^2}{2}+\frac{(r-\frac{\sigma^2}{2})ux^3+O(x^5)}{\sigma x+O(x^3)}+O(x^4) \\
=\ & 1+\frac{u^2x^2}{2}+\frac{(r-\frac{\sigma^2}{2})ux^3(1+O(x^2))}{\sigma x(1+O(x^2))}+O(x^4) \\
=\ & 1+\frac{u^2x^2}{2}+\frac{rux^2}{\sigma}-\frac{1}{2}\sigma ux^2+O(x^4).
\end{aligned}$$

□

(iv) Use the Taylor series expansion $\log(1+x)=x+O(x^2)$ to compute $\lim_{x\downarrow 0}\log\varphi_{\frac{1}{x^2}}(u)$. Now explain how you know that the limiting distribution for $\frac{\sigma}{\sqrt{n}}M_{nt,n}$ is normal with mean $\left(r-\frac{1}{2}\sigma^2\right)t$ and variance $\sigma^2 t$.

Solution.

$$\log\varphi_{\frac{1}{x^2}} = \frac{t}{x^2}\log\left(1+\frac{u^2x^2}{2}+\frac{ru}{\sigma}x^2-\frac{\sigma ux^2}{2}+O(x^4)\right) = \frac{t}{x^2}\left(\frac{u^2x^2}{2}+\frac{ru}{\sigma}x^2-\frac{\sigma ux^2}{2}+O(x^4)\right).$$

So $\lim_{x\downarrow 0}\log\varphi_{\frac{1}{x^2}}(u) = t(\frac{u^2}{2}+\frac{ru}{\sigma}-\frac{\sigma u}{2})$, and $\widetilde{\mathbb{E}}\left[e^{u\frac{1}{\sqrt{n}}M_{nt,n}}\right] = \varphi_n(u)\to\frac{1}{2}tu^2+t(\frac{r}{\sigma}-\frac{\sigma}{2})u$. By the one-to-one correspondence between distribution and moment generating function[6], $(\frac{1}{\sqrt{n}}M_{nt,n})_n$ converges to a Gaussian random variable with mean $t(\frac{r}{\sigma}-\frac{\sigma}{2})$ and variance t. Hence $(\frac{\sigma}{\sqrt{n}}M_{nt,n})_n$ converges to a Gaussian random variable with mean $t(r-\frac{\sigma^2}{2})$ and variance $\sigma^2 t$. □

[6] This correspondence holds only under certain conditions, which normal distribution certainly satisfies. For details, see Shiryaev [12, page 293-296].

▶ **Exercise 3.9 (Laplace transform of first passage density).** The solution to this problem is long and technical. It is included for the sake of completeness, but the reader may safely skip it.

Let $m > 0$ be given, and define

$$f(t,m) = \frac{m}{t\sqrt{2\pi t}} \exp\left\{-\frac{m^2}{2t}\right\}.$$

According to (3.7.3) in Theorem 3.7.1, $f(t,m)$ is the density in the variable t of the first passage time $\tau_m = \min\{t \geq 0; W(t) = m\}$, where W is a Brownian motion without drift. Let

$$g(\alpha,m) = \int_0^\infty e^{-\alpha t} f(t,m)dt, \alpha > 0,$$

be the Laplace transform of the density $f(t,m)$. This problem verifies that $g(\alpha,m) = e^{-m\sqrt{2\alpha}}$, which is the formula derived in Theorem 3.6.2.

(i) For $k \geq 1$, define

$$\alpha_k(m) = \frac{1}{\sqrt{2\pi}} \int_0^\infty t^{-k/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt,$$

so $g(\alpha,m) = ma_3(m)$. Show that

$$\begin{aligned} g_m(\alpha,m) &= a_3(m) - m^2 a_5(m), \\ g_{mm}(\alpha,m) &= -3ma_5(m) + m^3 a_7(m). \end{aligned}$$

Proof. We note

$$\begin{aligned} \frac{\partial}{\partial m} a_k(m) &= \frac{1}{\sqrt{2\pi}} \int_0^\infty t^{-k/2} \frac{\partial}{\partial m} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt \\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty t^{-k/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} \left(-\frac{m}{t}\right) dt \\ &= -ma_{k+2}(m). \end{aligned}$$

So

$$g_m(\alpha,m) = \frac{\partial}{\partial m}[ma_3(m)] = a_3(m) - m^2 a_5(m),$$

and

$$g_{mm}(\alpha,m) = -ma_5(m) - 2ma_5(m) + m^3 a_7(m) = -3ma_5(m) + m^3 a_7(m).$$

□

(ii) Use integration by parts to show that

$$a_5(m) = -\frac{2\alpha}{3} a_3(m) + \frac{m^2}{3} a_7(m).$$

Proof. For $k > 2$ and $\alpha > 0$, we have

$$\begin{aligned} a_k(m) &= \frac{1}{\sqrt{2\pi}} \int_0^\infty t^{-k/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt \\ &= -\frac{2}{k-2} \cdot \frac{1}{\sqrt{2\pi}} \int_0^\infty \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt^{-(k-2)/2} \\ &= -\frac{2}{k-2} \cdot \frac{1}{\sqrt{2\pi}} \left[\exp\left\{-\alpha t - \frac{m^2}{2t}\right\} t^{-(k-2)/2} \Big|_0^\infty \right. \\ &\quad \left. - \int_0^\infty t^{-(k-2)/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} \left(-\alpha + \frac{m^2}{2t^2}\right) dt \right] \\ &= -\frac{2}{k-2} \cdot \frac{1}{\sqrt{2\pi}} \left[\alpha \int_0^\infty t^{-(k-2)/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt - \frac{m^2}{2} \int_0^\infty t^{-(k+2)/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt \right] \\ &= -\frac{2\alpha}{k-2} a_{k-2}(m) + \frac{m^2}{k-2} a_{k+2}(m). \end{aligned}$$

Plug in $k = 5$, we obtain $a_5(m) = -\frac{2\alpha}{3}a_3(m) + \frac{m^2}{3}a_7(m)$. $\square$

(iii) Use (i) and (ii) to show that g satisfies the second-order ordinary differential equation

$$g_{mm}(\alpha, m) = 2\alpha g(\alpha, m).$$

Proof. We note

$$g_{mm}(\alpha, m) = -3ma_5(m) + m^3 a_7(m) = 2\alpha m a_3(m) - m^3 a_7(m) + m^3 a_7(m) = 2\alpha m a_3(m) = 2\alpha g(\alpha, m).$$

$\square$

(iv) The general solution to a second-order ordinary differential equation of the form

$$ay''(m) + by'(m) + cy(m) = 0$$

is

$$y(m) = A_1 e^{\lambda_1 m} + A_2 e^{\lambda_2 m},$$

where λ_1 and λ_2 are roots of the *characteristic equation*

$$a\lambda^2 + b\lambda + c = 0.$$

Here we are assuming that these roots are distinct. Find the general solution of the equation in (iii) when $\alpha > 0$. This solution has two undetermined parameters A_1 and A_2, and these may depend on α.

Solution. The characteristic equation of $g_{mm}(\alpha, m) = 2\alpha g(\alpha, m)$ is

$$\lambda^2 - 2\alpha = 0.$$

So $\lambda_1 = \sqrt{2\alpha}$ and $\lambda_2 = -\sqrt{2\alpha}$, and the general solution of the equation in (iii) is

$$A_1 e^{\sqrt{2\alpha}m} + A_2 e^{-\sqrt{2\alpha}m}.$$

$\square$

(v) Derive the bound

$$g(\alpha, m) \le \frac{m}{\sqrt{2\pi}} \int_0^m \sqrt{\frac{m}{t}} t^{-3/2} \exp\left\{-\frac{m^2}{2t}\right\} dt + \frac{1}{\sqrt{2\pi m}} \int_m^\infty e^{-\alpha t} dt$$

and use it to show that, for every $\alpha > 0$,

$$\lim_{m \to \infty} g(\alpha, m) = 0.$$

Use this fact to determine one of the parameters in the general solution to the equation in (iii).

Solution. We have

$$\begin{aligned}
g(\alpha, m) &= ma_3(m) \\
&= \frac{m}{\sqrt{2\pi}} \int_0^\infty t^{-3/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt \\
&= \frac{m}{\sqrt{2\pi}} \left(\int_0^m t^{-3/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt + \int_m^\infty t^{-3/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt \right) \\
&\le \frac{m}{\sqrt{2\pi}} \left(\int_0^m t^{-3/2} \exp\left\{-\alpha t - \frac{m^2}{2t}\right\} dt + \int_m^\infty m^{-3/2} \exp\left\{-\alpha t\right\} dt \right) \\
&= \frac{m}{\sqrt{2\pi}} \int_0^m e^{-\alpha t} t^{-3/2} \exp\left\{-\frac{m^2}{2t}\right\} dt + \frac{1}{\sqrt{2\pi m}} \int_m^\infty e^{-\alpha t} dt.
\end{aligned}$$

Over the interval $(0,m]$, $e^{-\alpha t}$ is monotone decreasing with a range of $[e^{-\alpha m},1)$ and $\sqrt{\frac{m}{t}}$ is monotone decreasing with a range of $[1,\infty)$. so $e^{-\alpha t}<\sqrt{\frac{m}{t}}$ over the interval $(0,m]$. This gives the desired bound

$$g(\alpha,m)\le\frac{m}{\sqrt{2\pi}}\int_0^m\sqrt{\frac{m}{t}}t^{-3/2}\exp\left\{-\frac{m^2}{2t}\right\}dt+\frac{1}{\sqrt{2\pi m}}\int_m^\infty e^{-\alpha t}dt.$$

Now, let $m\to\infty$. Obviously $\lim_{m\to\infty}\frac{1}{\sqrt{2\pi m}}\int_m^\infty e^{-\alpha t}dt=0$, while the first term through the change of variable $-\frac{1}{t}=u$ satisfies

$$\begin{aligned}\frac{m}{\sqrt{2\pi}}\int_0^m\sqrt{\frac{m}{t}}t^{-3/2}\exp\left\{-\frac{m^2}{2t}\right\}dt &= \frac{m^{3/2}}{\sqrt{2\pi}}\int_0^m t^{-2}\exp\left\{-\frac{m^2}{2t}\right\}dt\\ &= \frac{m^{3/2}}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{1}{m}}e^{\frac{m^2}{2}u}du\\ &= \frac{m^{3/2}}{\sqrt{2\pi}}\cdot\frac{e^{-m/2}}{m^2/2}\to 0\end{aligned}$$

as $m\to\infty$. Combined, we can conclude that for every $\alpha>0$, $\lim_{m\to\infty}g(\alpha,m)=\lim_{m\to\infty}(A_1e^{\sqrt{2\alpha}m}+A_2e^{-\sqrt{2\alpha}m})=0$, which requires $A_1=0$ as a necessary condition. Therefore, $g(\alpha,m)=A_2e^{-\sqrt{2\alpha}m}$.

□

(vi) Using first the change of variable $s=t/m^2$ and then the change of variable $y=1/\sqrt{s}$, show that

$$\lim_{m\downarrow 0}g(\alpha,m)=1.$$

Using this fact to determine the other parameter in the general solution to the equation in (iii).

Solution. Following the hint of the problem, we have

$$\begin{aligned}g(\alpha,m) &= \frac{m}{\sqrt{2\pi}}\int_0^\infty t^{-3/2}e^{-\alpha t-\frac{m^2}{2t}}dt\\ &\overset{s=t/m^2}{=} \frac{m}{\sqrt{2\pi}}\int_0^\infty (sm^2)^{-3/2}e^{-\alpha m^2 s-\frac{1}{2s}}m^2ds\\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty s^{-3/2}e^{-\alpha m^2 s-\frac{1}{2s}}ds\\ &\overset{y=1/\sqrt{s}}{=} \frac{1}{\sqrt{2\pi}}\int_0^\infty y^3e^{-\alpha m^2\frac{1}{y^2}-\frac{y^2}{2}}\frac{2}{y^3}dy\\ &= \frac{2}{\sqrt{2\pi}}\int_0^\infty e^{-\alpha m^2\frac{1}{y^2}-\frac{y^2}{2}}dy.\end{aligned}$$

So by dominated convergence theorem,

$$\lim_{m\downarrow 0}g(\alpha,m)=\frac{2}{\sqrt{2\pi}}\int_0^\infty\lim_{m\downarrow 0}e^{-\alpha m^2\frac{1}{y^2}-\frac{y^2}{2}}dy=\frac{2}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{y^2}{2}}dy=1.$$

We already proved in (v) that $g(\alpha,m)=A_2e^{-\sqrt{2\alpha}m}$. So we must have $A_2=1$ and hence $g(\alpha,m)=e^{-\sqrt{2\alpha}m}$.

□

4 Stochastic Calculus

★ **Comments**:

1) To see how we obtain the solution to the Vasicek interest rate model (equation (4.4.33)), we recall the method of *integrating factors*, a technique often used in solving first-order linear ordinary differential equations (see, for example, Logan [8, page 74]). Starting from (4.4.32)

$$dR(t) = (\alpha - \beta R(t))dt + \sigma dW(t),$$

we move the term containing $R(t)$ to the left side of the equation and multiple both sides by the integrating factor $e^{\beta t}$:

$$e^{\beta t}[R(t)d(\beta t) + dR(t)] = e^{\beta t}[\alpha dt + \sigma dW(t)].$$

Applying Itô's formula lets us know the left side of the equation is

$$d(e^{\beta t}R(t)).$$

Integrating from 0 to t on both sides, we obtain

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

This gives us (4.4.33)

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

2) The distribution of CIR interest rate process $R(t)$ for each positive t can be found in reference [41] of the textbook (Cox, J. C., INGERSOLL, J. E., AND ROSS, S. (1985) A theory of the term structure of interest rates, *Econometrica* **53**, 373-384).

3) Regarding to the comment at the beginning of Section 4.5.2, "Black, Scholes, and Merton argued that the value of this call at any time should depend on the time (more precisely, on the time to expiration) and on the value of the stock price at that time", we note it is justified by risk-neutral pricing and Markov property:

$$C(t) = \widetilde{\mathbb{E}}[e^{-r(T-t)}(S_T - K)^+|\mathcal{F}(t)] = \widetilde{\mathbb{E}}[e^{-r(T-t)}(S_T - K)^+|S_t] = c(t, S_t).$$

For details, see Chapter 5.

4) For the PDE approach to solving the Black-Scholes-Merton equation, see Wilmott [15] for details.

► **Exercise 4.1.** Suppose $M(t)$, $0 \le t \le T$, is a martingale with respect to some filtration $\mathcal{F}(t)$, $0 \le t \le T$. Let $\Delta(t)$, $0 \le t \le T$, be a simple process adapted to $\mathcal{F}(t)$ (i.e., there is a partition $\Pi = \{t_0, t_1, \cdots, t_n\}$ of $[0,T]$ such that, for every j, $\Delta(t_j)$ is $\mathcal{F}(t_j)$-measurable and $\Delta(t)$ is constant in t on each subinterval $[t_j, t_{j+1})$). For $t \in [t_k, t_{k+1})$, define the stochastic integral

$$I(t) = \sum_{j=0}^{k-1}\Delta(t_j)[M(t_{j+1}) - M(t_j)] + \Delta(t_k)[M(t) - M(t_k)].$$

We think of $M(t)$ as the price of an asset at time t and $\Delta(t_j)$ as the number of shares of the asset held by an investor between times t_j and t_{j+1}. Then $I(t)$ is the capital gains that accrue to the investor between times 0 and t. Show that $I(t)$, $0 \le t \le T$, is a martingale.

Proof. Fix t and for any $s < t$, we assume $s \in [t_m, t_{m+1})$ for some m.

Case 1. $m = k$. Then $I(t) - I(s) = \Delta_{t_k}(M_t - M_{t_k}) - \Delta_{t_k}(M_s - M_{t_k}) = \Delta_{t_k}(M_t - M_s)$. So $\mathbb{E}[I(t) - I(s)|\mathcal{F}_t] = \Delta_{t_k}\mathbb{E}[M_t - M_s|\mathcal{F}_s] = 0$.

Case 2. $m < k$. Then $t_m \le s < t_{m+1} \le t_k \le t < t_{k+1}$. So

$$\begin{aligned} I(t) - I(s) &= \sum_{j=m}^{k-1}\Delta_{t_j}(M_{t_{j+1}} - M_{t_j}) + \Delta_{t_k}(M_s - M_{t_k}) - \Delta_{t_m}(M_s - M_{t_m}) \\ &= \sum_{j=m+1}^{k-1}\Delta_{t_j}(M_{t_{j+1}} - M_{t_j}) + \Delta_{t_k}(M_t - M_{t_k}) + \Delta_{t_m}(M_{t_{m+1}} - M_s). \end{aligned}$$

Hence

$$
\begin{aligned}
& \mathbb{E}[I(t)-I(s)|\mathcal{F}_s] \\
= & \sum_{j=m+1}^{k-1} \mathbb{E}[\Delta_{t_j}\mathbb{E}[M_{t_{j+1}}-M_{t_j}|\mathcal{F}_{t_j}]|\mathcal{F}_s] + \mathbb{E}[\Delta_{t_k}\mathbb{E}[M_t - M_{t_k}|\mathcal{F}_{t_k}]|\mathcal{F}_s] + \Delta_{t_m}\mathbb{E}[M_{t_{m+1}}-M_s|\mathcal{F}_s] \\
= & 0.
\end{aligned}
$$

Combined, we can conclude $I(t)$ is a martingale. □

▶ **Exercise 4.2.** Let $W(t)$, $0 \le t \le T$, be a Brownian motion, and let $\mathcal{F}(t)$, $0 \le t \le T$, be an associated filtration. Let $\Delta(t)$, $0 \le t \le T$, be a *nonrandom* simple process (i.e., there is a partition $\Pi = \{t_0, t_1, \cdots, t_n\}$ of $[0,T]$ such that for every j, $\Delta(t_j)$ is a nonrandom quantity and $\Delta(t) = \Delta(t_j)$ is constant in t on the subinterval $[t_j, t_{j+1})$). For $t \in [t_k, t_{k+1})$, define the stochastic integral

$$
I(t) = \sum_{j=0}^{k-1} \Delta(t_j)[W(t_{j+1}) - W(t_j)] + \Delta(t_k)[W(t) - W(t_k)].
$$

(i) Show that whenever $0 \le s < t \le T$, the increment $I(t) - I(s)$ is independent of $\mathcal{F}(s)$. (Simplification: If s is between two partition points, we can always insert s as an extra partition point. Then we can relabel the partition points so that they are still called t_0, t_1, $\cdots$, t_n, but with a larger value of n and now with $s = t_k$ for some value of k. Of course, we must set $\Delta(s) = \Delta(t_{k-1})$ so that Δ takes the same value on the interval $[s, t_{k+1})$ as on the interval $[t_{k-1}, s)$. Similarly, we can insert t as an extra partition point if it is not already one. Consequently, to show that $I(t) - I(s)$ is independent of $\mathcal{F}(s)$ for all $0 \le s < t \le T$, it suffices to show that $I(t_k) - I(t_l)$ is independent of $\mathcal{F}_l$ whenever t_k and t_l are two partition points with $t_l < t_k$. This is all you need to do.)

Proof. We follow the simplification in the hint and consider $I(t_k) - I(t_l)$ with $t_l < t_k$. Then $I(t_k) - I(t_l) = \sum_{j=l}^{k-1} \Delta(t_j)[W(t_{j+1}) - W(t_j)]$. Since $\Delta(t)$ is a non-random process and $W(t_{j+1}) - W(t_j) \perp \mathcal{F}(t_j) \supset \mathcal{F}(t_l)$ for $j \ge l$, we must have $I(t_k) - I(t_l) \perp \mathcal{F}(t_l)$. □

(ii) Show that whenever $0 \le s < t \le T$, the increment $I(t) - I(s)$ is a normally distributed random variable with mean zero and variance $\int_s^t \Delta^2(u)du$.

Proof. We use the notation in (i) and it is clear that $I(t_k) - I(t_l)$ is normal since it is a linear combination of independent normal random variables. Furthermore, $\mathbb{E}[I(t_k) - I(t_l)] = \sum_{j=l}^{k-1} \Delta_{t_j}\mathbb{E}[W(t_{j+1}) - W(t_j)] = 0$ and $\mathrm{Var}(I(t_k) - I(t_l)) = \sum_{j=l}^{k-1} \Delta^2(t_j)\mathrm{Var}[W(t_{j+1}) - W(t_j)] = \sum_{j=l}^{k-1} \Delta^2(t_j)(t_{j+1} - t_j) = \int_{t_l}^{t_k} \Delta_u^2 du$. □

(iii) Use (i) and (ii) to show that $I(t)$, $0 \le t \le T$, is a martingale.

Proof. By (i), $\mathbb{E}[I(t) - I(s)|\mathcal{F}(s)] = \mathbb{E}[I(t) - I(s)] = 0$, for $s < t$, where the last equality is due to (ii). □

(iv) Show that $I^2(t) - \int_0^t \Delta^2(u)du$, $0 \le t \le T$, is a martingale.

Proof. For $s < t$,

$$
\begin{aligned}
& \mathbb{E}\left[I^2(t) - \int_0^t \Delta_u^2 du - \left(I^2(s) - \int_0^s \Delta_u^2 du \right) \middle| \mathcal{F}_s \right] \\
= & \mathbb{E}\left[(I(t) - I(s))^2 + 2I(t)I(s) - 2I^2(s) | \mathcal{F}_s \right] - \int_s^t \Delta_u^2 du \\
= & \mathbb{E}[(I(t) - I(s))^2] + 2I(s)\mathbb{E}[I(t) - I(s)|\mathcal{F}_s] - \int_s^t \Delta_u^2 du \\
= & \int_s^t \Delta_u^2 du + 0 - \int_s^t \Delta_u^2 du \\
= & 0.
\end{aligned}
$$

$\square$

▶ **Exercise 4.3.** We now consider a case in which $\Delta(t)$ in Exercise 4.2 is simple but random. In particular, let $t_0 = 0$, $t_1 = s$, and $t_2 = t$, and let $\Delta(0)$ be nonrandom and $\Delta(s) = W(s)$. Which of the following assertions is true?

(i) $I(t) - I(s)$ is independent of $\mathcal{F}(s)$.

Solution. We first note

$$\begin{aligned} I(t) - I(s) &= \Delta(0)[W(t_1) - W(0)] + \Delta(t_1)[W(t_2) - W(t_1)] - \Delta(0)[W(t_1) - W(0)] \\ &= \Delta(t_1)[W(t_2) - W(t_1)] \\ &= W(s)[W(t) - W(s)]. \end{aligned}$$

So $I(t) - I(s)$ is not independent of $\mathcal{F}(s)$, since $W(s) \in \mathcal{F}(s)$. $\square$

(ii) $I(t) - I(s)$ is normally distributed. (Hint: Check if the fourth moment is three times the square of the variance; see Exercise 3.3 of Chapter 3.)

Solution. Following the hint, recall from (i), we know $I(t) - I(s) = W(s)[W(t) - W(s)]$. So

$$\mathbb{E}[(I(t) - I(s))^4] = \mathbb{E}[W_s^4]\mathbb{E}[(W_t - W_s)^4] = 3s^2 \cdot 3(t-s)^2 = 9s^2(t-s)^2$$

and

$$3\left(\mathbb{E}[(I(t) - I(s))^2]\right)^2 = 3\left(\mathbb{E}[W_s^2]\mathbb{E}[(W_t - W_s)^2]\right)^2 = 3s^2(t-s)^2.$$

Since $\mathbb{E}[(I(t) - I(s))^4] \neq 3\left(\mathbb{E}[(I(t) - I(s))^2]\right)^2$, $I(t) - I(s)$ is not normally distributed. $\square$

(iii) $\mathbb{E}[I(t)|\mathcal{F}(s)] = I(s)$.

Solution. $\mathbb{E}[I(t) - I(s)|\mathcal{F}(s)] = W(s)\mathbb{E}[W(t) - W(s)|\mathcal{F}(s)] = 0$. So $\mathbb{E}[I(t)|\mathcal{F}(s)] = I(s)$ is true. $\square$

(iv) $\mathbb{E}[I^2(t) - \int_0^t \Delta^2(u)du|\mathcal{F}(s)] = I^2(s) - \int_0^s \Delta^2(u)du$.

Solution.

$$\begin{aligned} & \mathbb{E}\left[I^2(t) - \int_0^t \Delta^2(u)du - \left(I^2(s) - \int_0^s \Delta^2(u)du\right)\Big|\mathcal{F}(s)\right] \\ = \ & \mathbb{E}\left[(I(t) - I(s))^2 + 2I(t)I(s) - 2I^2(s) - \int_s^t \Delta^2(u)du\Big|\mathcal{F}(s)\right] \\ = \ & \mathbb{E}\left[(I(t) - I(s))^2 + 2I(s)(I(t) - I(s)) - \int_s^t W^2(s)1_{(s,t]}(u)du\Big|\mathcal{F}(s)\right] \\ = \ & \mathbb{E}[W^2(s)(W(t) - W(s))^2 + 2\Delta(0)W^2(s)(W(t) - W(s)) - W^2(s)(t-s)|\mathcal{F}(s)] \\ = \ & W^2(s)\mathbb{E}[(W(t) - W(s))^2] + 2\Delta(0)W^2(s)\mathbb{E}[W(t) - W(s)|\mathcal{F}(s)] - W^2(s)(t-s) \\ = \ & W^2(s)(t-s) - W^2(s)(t-s) \\ = \ & 0. \end{aligned}$$

So $\mathbb{E}[I^2(t) - \int_0^t \Delta^2(u)du|\mathcal{F}(s)] = I^2(s) - \int_0^s \Delta^2(u)du$ is true. $\square$

▶ **Exercise 4.4 (Stratonovich integral).** Let $W(t)$, $t \geq 0$, be a Brownian motion. Let T be a fixed positive number and let $\Pi = \{t_0, t_1, \cdots, t_n\}$ be a partition of $[0, T]$ (i.e., $0 = t_0 < t_1 < \cdots < t_n = T$). For each j, define $t_j^* = \frac{t_j + t_{j+1}}{2}$ to be the midpoint of the interval $[t_j, t_{j+1}]$.

(i) Define the *half-sample quadratic variation* corresponding to Π to be

$$Q_{\Pi/2} = \sum_{j=0}^{n-1}(W(t_j^*) - W(t_j))^2.$$

Show that $Q_{\Pi/2}$ has limit $\frac{1}{2}T$ as $\|\Pi\| \to 0$. (Hint: It suffices to show $\mathbb{E}Q_{\Pi/2} = \frac{1}{2}T$ and $\lim_{\|\Pi\|\to 0} \operatorname{Var}(Q_{\Pi/2}) = 0$.

Proof. Following the hint, we first note that

$$\mathbb{E}[Q_{\Pi/2}] = \sum_{j=0}^{n-1} \mathbb{E}[(W(t_j^*) - W(t_j))^2] = \sum_{j=0}^{n-1} (t_j^* - t_j) = \sum_{j=0}^{n-1} \frac{t_{j+1} - t_j}{2} = \frac{T}{2}.$$

Then, by noting $W(t_j^*) - W(t_j)$ is equal to $W(t_j^* - t_j) = W\left(\frac{t_{j+1}-t_j}{2}\right)$ in distribution and $\mathbb{E}[(W^2(t) - t)^2] = \mathbb{E}[W^4(t) - 2tW^2(t) + t^2] = 3\mathbb{E}[W^2(t)]^2 - 2t^2 + t^2 = 2t^2$, we have

$$\begin{aligned}
& \operatorname{Var}(Q_{\Pi/2}) \\
= \; & \mathbb{E}\left[\left(\sum_{j=0}^{n-1} \left(W(t_j^*) - W(t_j)\right)^2 - \frac{T}{2}\right)^2\right] \\
= \; & \mathbb{E}\left[\left(\sum_{j=0}^{n-1} \left(W(t_j^*) - W(t_j)\right)^2 - \sum_{j=0}^{n-1} \frac{t_{j+1} - t_j}{2}\right)^2\right] \\
= \; & \sum_{j,\,k=0}^{n-1} \mathbb{E}\left[\left(\left(W(t_j^*) - W(t_j)\right)^2 - \frac{t_{j+1} - t_j}{2}\right)\left((W(t_k^*) - W(t_k))^2 - \frac{t_{k+1} - t_k}{2}\right)\right] \\
= \; & \sum_{j,\,k=0, j\neq k}^{n-1} \mathbb{E}\left[\left(\left(W(t_j^*) - W(t_j)\right)^2 - \frac{t_{j+1} - t_j}{2}\right)\left((W(t_k^*) - W(t_k))^2 - \frac{t_{k+1} - t_k}{2}\right)\right] \\
& + \sum_{j=0}^{n-1} \mathbb{E}\left[\left((W(t_j^*) - W(t_j))^2 - \frac{t_{j+1} - t_j}{2}\right)^2\right] \\
= \; & 0 + \sum_{j=0}^{n-1} \mathbb{E}\left[\left(W^2(t_j^* - t_j) - \frac{t_{j+1} - t_j}{2}\right)^2\right] \\
= \; & \sum_{j=0}^{n-1} 2 \cdot \left(\frac{t_{j+1} - t_j}{2}\right)^2 \\
\leq \; & \frac{T}{2} \max_{1\leq j\leq n} |t_{j+1} - t_j| \to 0.
\end{aligned}$$

Combined, we can conclude $\lim_{\|\Pi\|\to 0} Q_{\Pi/2} \to \frac{T}{2}$ in $\mathcal{L}^2(\mathbb{P})$. □

(ii) Define the Stratonovich integral of $W(t)$ with respect to $W(t)$ to be

$$\int_0^T W(t) \circ dW(t) = \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} W(t_j^*)(W(t_{j+1}) - W(t_j)). \tag{4.10.1}$$

In contrast to the Itô integral $\int_0^T W(t)dW(t) = \frac{1}{2}W^2(T) - \frac{1}{2}T$ of (4.3.4), which evaluates the integrand at the left endpoint of each subinterval $[t_j, t_{j+1}]$, here we evaluate the integrand at the midpoint t_j^*. Show that

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

(Hint: Write the approximating sum in (4.10.1) as the sum of an approximating sum for the Itô integral $\int_0^T W(t)dW(t)$ and $Q_{\Pi/2}$. The approximating sum for the Itô integral is the one corresponding to the partition $0 = t_0 < t_0^* < t_1 < t_1^* < \cdots < t_{n-1}^* < t_n = T$, not the partition Π.)

Proof. Following the hint and using the result in (i), we have

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

By the construction of Itô integral,

$$\lim_{\|\Pi^*\|\to 0} \sum_{j=0}^{n-1} \left\{W(t_j^*)[W(t_{j+1}) - W(t_j^*)] + W(t_j)[W(t_j^*) - W(t_j)]\right\} = \int_0^T W(t)dW(t),$$

where Π^* is the partition $0 = t_0 < t_0^* < t_1 < t_1^* < \cdots < t_{n-1}^* < t_n = T$. (i) already shows

$$\lim_{\|\Pi\|} \sum_{j=0}^{n-1}[W(t_j^*) - W(t_j)]^2 = \frac{T}{2} = \frac{T}{2}.$$

Combined, we conclude

$$\int_0^T W(t) \circ dW(t) = \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} W(t_j^*)(W(t_{j+1}) - W(t_j)) = \int_0^T W(t)dW(t) + \frac{T}{2} = \frac{1}{2}W^2(T).$$

$\square$

▶ **Exercise 4.5 (Solving the generalized geometric Brownian motion equation).** Let $S(t)$ be a positive stochastic process that satisfies the generalized geometric Brownian motion differential equation (see Example 4.4.8)

$$dS(t) = \alpha(t)S(t)dt + \sigma(t)S(t)dW(t), \tag{4.10.2}$$

where $\alpha(t)$ and $\sigma(t)$ are processes adapted to the filtration $\mathcal{F}(t)$, $t \geq 0$, associated with the Brownian motion $W(t)$, $t \geq 0$. In this exercise, we show that $S(t)$ must be given by formula (4.4.26) (i.e., that formula provides the only solution to the stochastic differential equation (4.10.2)). In the process, we provide a method for solving this equation.

(i) Using (4.10.2) and the Itô-Doeblin formula, compute $d\log S(t)$. Simplify so that you have a formula for $d\log S(t)$ that does not involve $S(t)$.

Solution.

$$d\log S_t = \frac{dS_t}{S_t} - \frac{1}{2}\frac{d\langle S\rangle_t}{S_t^2} = \frac{2S_t dS_t - d\langle S\rangle_t}{2S_t^2} = \frac{2S_t(\alpha_t S_t dt + \sigma_t S_t dW_t) - \sigma_t^2 S_t^2 dt}{2S_t^2} = \sigma_t dW_t + \left(\alpha_t - \frac{1}{2}\sigma_t^2\right)dt.$$

$\square$

(ii) Integrate the formula you obtained in (i), and then exponentiate the answer to obtain (4.4.26).

Solution.

$$\log S_t = \log S_0 + \int_0^t \sigma_s dW_s + \int_0^t \left(\alpha_s - \frac{1}{2}\sigma_s^2\right)ds.$$

So $S_t = S_0 \exp\left\{\int_0^t \sigma_s dW_s + \int_0^t \left(\alpha_s - \frac{1}{2}\sigma_s^2\right)ds\right\}$. $\square$

▶ **Exercise 4.6.** Let $S(t) = S(0)\exp\left\{\sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right)t\right\}$ be a geometric Brownian motion. Let p be a positive constant. Compute $d(S^p(t))$, the differential of $S(t)$ raised to the power p.

Solution. Without loss of generality, we assume $p \neq 1$. Since $(x^p)' = px^{p-1}$, $(x^p)'' = p(p-1)x^{p-2}$, we have

$$\begin{aligned} d(S_t^p) &= pS_t^{p-1}dS_t + \frac{1}{2}p(p-1)S_t^{p-2}d\langle S\rangle_t \\ &= pS_t^{p-1}(\alpha S_t dt + \sigma S_t dW_t) + \frac{1}{2}p(p-1)S_t^{p-2}\sigma^2 S_t^2 dt \\ &= S_t^p\left[p\alpha dt + p\sigma dW_t + \frac{1}{2}p(p-1)\sigma^2 dt\right] \\ &= pS_t^p\left[\sigma dW_t + \left(\alpha + \frac{p-1}{2}\sigma^2\right)dt\right]. \end{aligned}$$

□

▶ **Exercise 4.7.**
(i) Compute $dW^4(t)$ and then write $W^4(T)$ as the sum of an ordinary (Lebesgue) integral with respect to time and an Itô integral.

Solution. $dW_t^4 = 4W_t^3 dW_t + \frac{1}{2}\cdot 4\cdot 3W_t^2 d\langle W\rangle_t = 4W_t^3 dW_t + 6W_t^2 dt$. So $W_T^4 = 4\int_0^T W_t^3 dW_t + 6\int_0^T W_t^2 dt$. □

(ii) Take expectations on both sides of the formula you obtained in (i), use the fact that $\mathbb{E}W^2(t) = t$, and derive the formula $\mathbb{E}W^4(T) = 3T^2$.

Solution. $\mathbb{E}[W_T^4] = 4\mathbb{E}\left[\int_0^T W_t^3 dW_t\right] + 6\int_0^T \mathbb{E}[W_t^2]dt = 6\int_0^T t dt = 3T^2$. □

(iii) Use the method of (i) and (ii) to derive a formula for $\mathbb{E}W^6(T)$.

Solution. $dW_t^6 = 6W_t^5 dW_t + \frac{1}{2}\cdot 6\cdot 5W_t^4 dt$. So $W_T^6 = 6\int_0^T W_t^5 dW_t + 15\int_0^T W_t^4 dt$. Hence $\mathbb{E}[W_T^6] = 15\int_0^T 3t^2 dt = 15T^3$. □

▶ **Exercise 4.8 (Solving the Vasicek equation).** The Vasicek interest rate stochastic differential equation (4.4.32) is

$$dR(t) = (\alpha - \beta R(t))dt + \sigma dW(t),$$

where α, β, and σ are positive constants. The solution to this equation is given in Example 4.4.10. This exercise shows how to derive this solution.
(i) Use (4.4.32) and the Itô-Doeblin formula to compute $d(e^{\beta t}R(t))$. Simplify it so that you have a formula for $d(e^{\beta t}R(t))$ that does not involve $R(t)$.

Solution. $d(e^{\beta t}R_t) = \beta e^{\beta t}R_t dt + e^{\beta t}dR_t = e^{\beta t}[\beta R_t dt + (\alpha - \beta R(t))dt + \sigma dW(t)] = e^{\beta}(\alpha dt + \sigma dW(t))$. □

(ii) Integrate the equation you obtain in (i) and solve for $R(t)$ to obtain (4.4.33).

Solution.

$$e^{\beta t}R_t = R_0 + \int_0^t e^{\beta s}(\alpha ds + \sigma dW_s) = R_0 + \frac{\alpha}{\beta}(e^{\beta t} - 1) + \sigma\int_0^t e^{\beta s}dW_s.$$

Therefore $R_t = R_0 e^{-\beta t} + \frac{\alpha}{\beta}(1 - e^{-\beta t}) + \sigma\int_0^t e^{-\beta(t-s)}dW_s$. □

▶ **Exercise 4.9.** For a European call expiring at time T with strike price K, the Black-Scholes-Merton price at time t, if the time-t stock price is x, is

$$c(t,x) = xN(d_+(T-t,x)) - Ke^{-r(T-t)}N(d_-(T-t,x)),$$

where

$$\begin{aligned} d_+(\tau,x) &= \frac{1}{\sigma\sqrt{\tau}}\left[\log\frac{x}{K} + \left(r + \frac{1}{2}\sigma^2\right)\tau\right] \\ d_-(\tau,x) &= d_+(\tau,x) - \sigma\sqrt{\tau}, \end{aligned}$$

and $N(y)$ is the cumulative standard normal distribution

$$N(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{z^2}{2}}dz = \frac{1}{\sqrt{2\pi}}\int_{-y}^{\infty} e^{-\frac{z^2}{2}}dz.$$

The purpose of this exercise is to show that the function c satisfies the Black-Scholes-Merton partial differential equation

$$c_t(t,x) + rxc_x(t,x) + \frac{1}{2}\sigma^2x^2c_{xx}(t,x) = rc(t,x), 0 \le t < T, x > 0, \tag{4.10.3}$$

the *terminal condition*

$$\lim_{t\uparrow T} c(t,x) = (x-K)^+, x > 0, x \ne K, \tag{4.10.4}$$

and the *boundary conditions*

$$\lim_{x\downarrow 0} c(t,x) = 0, \lim_{x\to\infty}[c(t,x) - (x - e^{-r(T-t)}K)] = 0, 0 \le t < T. \tag{4.10.5}$$

Equation (4.10.4) and the first part of (4.10.5) are usually written more simply but less precisely as

$$c(T,x) = (x-K)^+, x \ge 0$$

and

$$c(t,0) = 0, 0 \le t \le T.$$

For this exercise, we abbreviate $c(t,x)$ as simply c and $d_\pm(T-t,x)$ as simply $d_\pm$.

(i) Verify first the equation

$$Ke^{-r(T-t)}N'(d_-) = xN'(d_+). \tag{4.10.6}$$

Proof.

$$\begin{aligned} Ke^{-r(T-t)}N'(d_-) &= Ke^{-r(T-t)}\frac{e^{-\frac{d_-^2}{2}}}{\sqrt{2\pi}} \\ &= Ke^{-r(T-t)}\frac{e^{-\frac{(d_+-\sigma\sqrt{T-t})^2}{2}}}{\sqrt{2\pi}} \\ &= Ke^{-r(T-t)}e^{\sigma\sqrt{T-t}d_+}e^{-\frac{\sigma^2(T-t)}{2}}N'(d_+) \\ &= Ke^{-r(T-t)}\frac{x}{K}e^{(r+\frac{\sigma^2}{2})(T-t)}e^{-\frac{\sigma^2(T-t)}{2}}N'(d_+) \\ &= xN'(d_+). \end{aligned}$$

□

(ii) Show that $c_x = N(d_+)$. This is the *delta* of the option. (Be careful! Remember that d_+ is a function of x.)

Proof. By equation (4.10.6) from (i),

$$\begin{aligned}
c_x &= N(d_+) + xN'(d_+)\frac{\partial}{\partial x}d_+(T-t,x) - Ke^{-r(T-t)}N'(d_-)\frac{\partial}{\partial x}d_-(T-t,x) \\
&= N(d_+) + xN'(d_+)\frac{\partial}{\partial x}d'_+(T-t,x) - xN'(d_+)\frac{\partial}{\partial x}d_+(T-t,x) \\
&= N(d_+).
\end{aligned}$$

□

(iii) Show that

$$c_t = -rKe^{-r(T-t)}N(d_-) - \frac{\sigma x}{2\sqrt{T-t}}N'(d_+).$$

This is the *theta* of the option.

Proof.

$$\begin{aligned}
c_t &= xN'(d_+)\frac{\partial}{\partial t}d_+(T-t,x) - rKe^{-r(T-t)}N(d_-) - Ke^{-r(T-t)}N'(d_-)\frac{\partial}{\partial t}d_-(T-t,x) \\
&= xN'(d_+)\frac{\partial}{\partial t}d_+(T-t,x) - rKe^{-r(T-t)}N(d_-) - xN'(d_+)\left[\frac{\partial}{\partial t}d_+(T-t,x) + \frac{\sigma}{2\sqrt{T-t}}\right] \\
&= -rKe^{-r(T-t)}N(d_-) - \frac{\sigma x}{2\sqrt{T-t}}N'(d_+).
\end{aligned}$$

□

(iv) Use the formulas above to show that c satisfies (4.10.3).

Proof.

$$\begin{aligned}
& c_t + rxc_x + \frac{1}{2}\sigma^2x^2c_{xx} \\
=\ & -rKe^{-r(T-t)}N(d_-) - \frac{\sigma x}{2\sqrt{T-t}}N'(d_+) + rxN(d_+) + \frac{1}{2}\sigma^2x^2N'(d_+)\frac{\partial}{\partial x}d_+(T-t,x) \\
=\ & rc - \frac{\sigma x}{2\sqrt{T-t}}N'(d_+) + \frac{1}{2}\sigma^2x^2N'(d_+)\frac{1}{\sigma\sqrt{T-t}x} \\
=\ & rc.
\end{aligned}$$

□

(v) Show that for $x > K$, $\lim_{t\uparrow T} d_\pm = \infty$, but for $0 < x < K$, $\lim_{t\uparrow T} d_\pm = -\infty$. Use these equalities to derive the terminal condition (4.10.4).

Proof. For $x > K$, $d_+(T-t,x) > 0$ and $\lim_{t\uparrow T} d_+(T-t,x) = \lim_{\tau\downarrow 0} d_+(\tau,x) = \infty$. $\lim_{t\uparrow T} d_-(T-t,x) = \lim_{\tau\downarrow 0} d_-(\tau,x) = \lim_{\tau\downarrow 0}\left(\frac{1}{\sigma\sqrt{\tau}}\ln\frac{x}{K} + \frac{1}{\sigma}(r+\frac{1}{2}\sigma^2)\sqrt{\tau} - \sigma\sqrt{\tau}\right) = \infty$. Similarly, $\lim_{t\uparrow T} d_\pm = -\infty$ for $x \in (0,K)$. Also it's clear that $\lim_{t\uparrow T} d_\pm = 0$ for $x = K$. So

$$\lim_{t\uparrow T} c(t,x) = xN(\lim_{t\uparrow T} d_+) - KN(\lim_{t\uparrow T} d_-) = \begin{cases} x-K, & \text{if } x > K \\ 0, & \text{if } x \leq K \end{cases} = (x-K)^+.$$

□

(vi) Show that for $0 \leq t < T$, $\lim_{x\downarrow 0} d_\pm = -\infty$. Use this fact to verify the first part of boundary condition (4.10.5) as $x \downarrow 0$.

Proof. We note

$$\lim_{x\downarrow 0} d_+ = \frac{1}{\sigma\sqrt{\tau}}\left[\lim_{x\downarrow 0}\log\frac{x}{K} + \left(r+\frac{1}{2}\sigma^2\right)\tau\right] = -\infty$$

and

$$\lim_{x\downarrow 0} d_- = \lim_{x\downarrow 0} d_+ - \sigma\sqrt{\tau} = -\infty.$$

So for $t \in [0, T)$,

$$\lim_{x\downarrow 0} c(t,x) = \lim_{x\downarrow 0} xN(\lim_{x\downarrow 0} d_+(T-t,x)) - Ke^{-r(T-t)}N(\lim_{x\downarrow 0} d_-(T-t,x)) = 0.$$

$\square$

(vii) Show that for $0 \leq t < T$, $\lim_{x\to\infty} d_\pm = \infty$. Use this fact to verify the second part of boundary condition (4.10.5) as $x \to \infty$. In this verification, you will need to show that

$$\lim_{x\to\infty}\frac{N(d_+)-1}{x^{-1}} = 0.$$

This is an indeterminate form $\frac{0}{0}$, and L'Hôspital's rule implies that this limit is

$$\lim_{x\to\infty}\frac{\frac{d}{dx}\left[N(d_+)-1\right]}{\frac{d}{dx}x^{-1}}.$$

Work out this expression and use the fact that

$$x = K\exp\left\{\sigma\sqrt{T-t}d_+ - (T-t)\left(r+\frac{1}{2}\sigma^2\right)\right\}$$

to write this expression solely in terms of d_+ (i.e., without the appearance of any x except the x in the argument of $d_+(T-t,x)$). Then argue that the limit is zero as $d_+ \to \infty$.

Proof. For $t \in [0,T)$, it is clear $\lim_{x\to\infty} d_\pm = \infty$. Following the hint, we note

$$\lim_{x\to\infty} x(N(d_+)-1) = \lim_{x\to\infty}\frac{N'(d_+)\frac{\partial}{\partial x}d_+}{-x^{-2}} = \lim_{x\to\infty}\frac{N'(d_+)\frac{1}{\sigma\sqrt{T-t}}}{-x^{-1}}.$$

By the expression of d_+, we have $x = K\exp\left\{\sigma\sqrt{T-t}d_+ - (T-t)(r+\frac{1}{2}\sigma^2)\right\}$. So

$$\lim_{x\to\infty} x(N(d_+)-1) = \lim_{x\to\infty} N'(d_+)\frac{-x}{\sigma\sqrt{T-t}} = \lim_{d_+\to\infty}\frac{e^{-\frac{d_+^2}{2}}}{\sqrt{2\pi}}\frac{-Ke^{\sigma\sqrt{T-t}d_+-(T-t)(r+\frac{1}{2}\sigma^2)}}{\sigma\sqrt{T-t}} = 0.$$

Therefore

$$\begin{aligned}
& \lim_{x\to\infty}[c(t,x) - (x - e^{-r(T-t)}K)] \\
= \ & \lim_{x\to\infty}[xN(d_+) - Ke^{-r(T-t)N(d_-)} - x + Ke^{-r(T-t)}] \\
= \ & \lim_{x\to\infty}[x(N(d_+)-1) + Ke^{-r(T-t)}(1-N(d_-))] \\
= \ & \lim_{x\to\infty} x(N(d_+)-1) + Ke^{-r(T-t)}(1-N(\lim_{x\to\infty} d_-)) \\
= \ & 0.
\end{aligned}$$

$\square$

**************************** UPDATE STOPPED HERE ************************************

▶ **Exercise 4.10 (Self-financing trading).** The fundamental idea behind no-arbitrage pricing is to reproduce the payoff of a derivative security by trading in the underlying asset (which we call a stock) and the money market account. In discrete time, we let X_k denote the value of the hedging portfolio at time k and let Δ_k denote the number of shares of stock held between times k and $k+1$. Then, at time k, after rebalancing (i.e., moving from a position of Δ_{k-1} to a position Δ_k in the stock), the amount in the money market account is $X_k - S_k\Delta_k$. The value of the portfolio at time $k+1$ is

$$X_{k+1} = \Delta_k S_{k+1} + (1+r)(X_k - \Delta_k S_k). \tag{4.10.7}$$

This formula can be rearranged to become

$$X_{k+1} - X_k = \Delta_k(S_{k+1} - S_k) + r(X_k - \Delta_k S_k), \tag{4.10.8}$$

which says that the gain between time k and time $k+1$ is the sum of the capital gain on the stock holdings, $\Delta_k(S_{k+1} - S_k)$, and the interest earnings on the money market account, $r(X_k - \Delta_k S_k)$. The continuous-time analogue of (4.10.8) is

$$dX(t) = \Delta(t)dS(t) + r(X(t) - \Delta(t)S(t))dt. \tag{4.10.9}$$

Alternatively, one could define the value of a share of the money market account at time k to be

$$M_k = (1+r)^k$$

and formulate the discrete-time model with two processes, Δ_k as before and Γ_k denoting the number of shares of the money market account held at time k after rebalancing. Then

$$X_k = \Delta_k S_k + \Gamma_k M_k, \tag{4.10.10}$$

so that (4.10.7) becomes

$$X_{k+1} = \Delta_k S_{k+1} + (1+r)\Gamma_k M_k = \Delta_k S_{k+1} + \Gamma_k M_{k+1}. \tag{4.10.11}$$

Subtracting (4.10.10) from (4.10.11), we obtain in place of (4.10.8) the equation

$$X_{k+1} - X_k = \Delta_k(S_{k+1} - S_k) + \Gamma_k(M_{k+1} - M_k), \tag{4.10.12}$$

which says that the gain between time k and time $k+1$ is the sum of the capital gain on stock holdings, $\Delta_k(S_{k+1} - S_k)$, and the earnings from the money market investment, $\Gamma_k(M_{k+1} - M_k)$.

But Δ_k and Γ_k cannot be chosen arbitrarily. The agent arrives at time

(i)

Proof. We show (4.10.16) + (4.10.9) $\iff$ (4.10.16) + (4.10.15), i.e. assuming X has the representation $X_t = \Delta_t S_t + \Gamma_t M_t$, "continuous-time self-financing condition" has two equivalent formulations, (4.10.9) or (4.10.15). Indeed, $dX_t = \Delta_t dS_t + \Gamma_t dM_t + (S_t d\Delta_t + dS_t d\Delta_t + M_t d\Gamma_t + dM_t d\Gamma_t)$. So $dX_t = \Delta_t dS_t + \Gamma_t dM_t \iff S_t d\Delta_t + dS_t d\Delta_t + M_t d\Gamma_t + dM_t d\Gamma_t = 0$, i.e. (4.10.9) $\iff$ (4.10.15). □

(ii)

Proof. First, we clarify the problems by stating explicitly the given conditions and the result to be proved. We assume we have a portfolio $X_t = \Delta_t S_t + \Gamma_t M_t$. We let $c(t, S_t)$ denote the price of call option at time t and set $\Delta_t = c_x(t, S_t)$. Finally, we assume the portfolio is self-financing. The problem is to show

$$rN_t dt = \left[c_t(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 c_{xx}(t, S_t)\right] dt,$$

where $N_t = c(t, S_t) - \Delta_t S_t$.

Indeed, by the self-financing property and $\Delta_t = c_x(t, S_t)$, we have $c(t, S_t) = X_t$ (by the calculations in Subsection 4.5.1-4.5.3). This uniquely determines Γ_t as

$$\Gamma_t = \frac{X_t - \Delta_t S_t}{M_t} = \frac{c(t, S_t) - c_x(t, S_t)S_t}{M_t} = \frac{N_t}{M_t}.$$

Moreover,

$$\begin{aligned} dN_t &= \left[c_t(t, S_t)dt + c_x(t, S_t)dS_t + \frac{1}{2}c_{xx}(t, S_t)d\langle S_t\rangle_t\right] - d(\Delta_t S_t) \\ &= \left[c_t(t, S_t) + \frac{1}{2}c_{xx}(t, S_t)\sigma^2 S_t^2\right]dt + [c_x(t, S_t)dS_t - d(X_t - \Gamma_t M_t)] \\ &= \left[c_t(t, S_t) + \frac{1}{2}c_{xx}(t, S_t)\sigma^2 S_t^2\right]dt + M_t d\Gamma_t + dM_t d\Gamma_t + [c_x(t, S_t)dS_t + \Gamma_t dM_t - dX_t]. \end{aligned}$$

By self-financing property, $c_x(t, S_t)dt + \Gamma_t dM_t = \Delta_t dS_t + \Gamma_t dM_t = dX_t$, so

$$\left[c_t(t, S_t) + \frac{1}{2}c_{xx}(t, S_t)\sigma^2 S_t^2\right]dt = dN_t - M_t d\Gamma_t - dM_t d\Gamma_t = \Gamma_t dM_t = \Gamma_t r M_t dt = rN_t dt.$$

□

4.11.

Proof. First, we note $c(t, x)$ solves the Black-Scholes-Merton PDE with volatility σ_1:

$$\left(\frac{\partial}{\partial t} + rx\frac{\partial}{\partial x} + \frac{1}{2}x^2\sigma_1^2\frac{\partial^2}{\partial x^2} - r\right)c(t, x) = 0.$$

So

$$c_t(t, S_t) + rS_t c_x(t, S_t) + \frac{1}{2}\sigma_1^2 S_t^2 c_{xx}(t, S_t) - rc(t, S_t) = 0,$$

and

$$\begin{aligned} dc(t, S_t) &= c_t(t, S_t)dt + c_x(t, S_t)(\alpha S_t dt + \sigma_2 S_t dW_t) + \frac{1}{2}c_{xx}(t, S_t)\sigma_2^2 S_t^2 dt \\ &= \left[c_t(t, S_t) + \alpha c_x(t, S_t)S_t + \frac{1}{2}\sigma_2^2 S_t^2 c_{xx}(t, S_t)\right]dt + \sigma_2 S_t c_x(t, S_t)dW_t \\ &= \left[rc(t, S_t) + (\alpha - r)c_x(t, S_t)S_t + \frac{1}{2}S_t^2(\sigma_2^2 - \sigma_1^2)c_{xx}(t, S_t)\right]dt + \sigma_2 S_t c_x(t, S_t)dW_t. \end{aligned}$$

Therefore

$$\begin{aligned} dX_t &= \Big[rc(t, S_t) + (\alpha - r)c_x(t, S_t)S_t + \frac{1}{2}S_t^2(\sigma_2^2 - \sigma_1^2)\sigma_{xx}(t, S_t) + rX_t - rc(t, S_t) + rS_t c_x(t, S_t) \\ &\quad -\frac{1}{2}(\sigma_2^2 - \sigma_1^2)S_t^2 c_{xx}(t, S_t) - c_x(t, S_t)\alpha S_t\Big]dt + [\sigma_2 S_t c_x(t, S_t) - c_x(t, S_t)\sigma_2 S_t]dW_t \\ &= rX_t dt. \end{aligned}$$

This implies $X_t = X_0 e^{rt}$. By X_0, we conclude $X_t = 0$ for all $t \in [0, T]$.

□

4.12. (i)

Proof. By (4.5.29), $c(t,x) - p(t,x) = x - e^{-r(T-t)}K$. So $p_x(t,x) = c_x(t,x) - 1 = N(d_+(T-t,x)) - 1$, $p_{xx}(t,x) = c_{xx}(t,x) = \frac{1}{\sigma x\sqrt{T-t}}N'(d_+(T-t,x))$ and

$$\begin{aligned}
p_t(t,x) &= c_t(t,x) + re^{-r(T-t)}K \\
&= -rKe^{-r(T-t)}N(d_-(T-t,x)) - \frac{\sigma x}{2\sqrt{T-t}}N'(d_+(T-t,x)) + rKe^{-r(T-t)} \\
&= rKe^{-r(T-t)}N(-d_-(T-t,x)) - \frac{\sigma x}{2\sqrt{T-t}}N'(d_+(T-t,x)).
\end{aligned}$$

□

(ii)

Proof. For an agent hedging a short position in the put, since $\Delta_t = p_x(t,x) < 0$, he should short the underlying stock and put $p(t,S_t) - p_x(t,S_t)S_t (> 0)$ cash in the money market account. □

(iii)

Proof. By the put-call parity, it suffices to show $f(t,x) = x - Ke^{-r(T-t)}$ satisfies the Black-Scholes-Merton partial differential equation. Indeed,

$$\left(\frac{\partial}{\partial t} + \frac{1}{2}\sigma^2 x^2\frac{\partial^2}{\partial x^2} + rx\frac{\partial}{\partial x} - r\right)f(t,x) = -rKe^{-r(T-t)} + \frac{1}{2}\sigma^2x^2\cdot 0 + rx\cdot 1 - r(x - Ke^{-r(T-t)}) = 0.$$

Remark: The Black-Scholes-Merton PDE has many solutions. Proper boundary conditions are the key to uniqueness. For more details, see Wilmott [15]. □

4.13.

Proof. We suppose (W_1, W_2) is a pair of local martingale defined by SDE

$$\begin{cases} dW_1(t) = dB_1(t) \\ dW_2(t) = \alpha(t)dB_1(t) + \beta(t)dB_2(t). \end{cases} \tag{1}$$

We want to find $\alpha(t)$ and $\beta(t)$ such that

$$\begin{cases} (dW_2(t))^2 = [\alpha^2(t) + \beta^2(t) + 2\rho(t)\alpha(t)\beta(t)]dt = dt \\ dW_1(t)dW_2(t) = [\alpha(t) + \beta(t)\rho(t)]dt = 0. \end{cases} \tag{2}$$

Solve the equation for $\alpha(t)$ and $\beta(t)$, we have $\beta(t) = \frac{1}{\sqrt{1-\rho^2(t)}}$ and $\alpha(t) = -\frac{\rho(t)}{\sqrt{1-\rho^2(t)}}$. So

$$\begin{cases} W_1(t) = B_1(t) \\ W_2(t) = \int_0^t \frac{-\rho(s)}{\sqrt{1-\rho^2(s)}}dB_1(s) + \int_0^t \frac{1}{\sqrt{1-\rho^2(s)}}dB_2(s) \end{cases} \tag{3}$$

is a pair of independent BM's. Equivalently, we have

$$\begin{cases} B_1(t) = W_1(t) \\ B_2(t) = \int_0^t \rho(s)dW_1(s) + \int_0^t \sqrt{1-\rho^2(s)}dW_2(s). \end{cases} \tag{4}$$

□

4.14. (i)

Proof. Clearly $Z_j \in \mathcal{F}_{t_{j+1}}$. Moreover

$$E[Z_j|\mathcal{F}_{t_j}] = f''(W_{t_j})E[(W_{t_{j+1}} - W_{t_j})^2 - (t_{j+1} - t_j)|\mathcal{F}_{t_j}] = f''(W_{t_j})(E[W^2_{t_{j+1}-t_j}] - (t_{j+1} - t_j)) = 0,$$

since $W_{t_{j+1}} - W_{t_j}$ is independent of $\mathcal{F}_{t_j}$ and $W_t \sim N(0,t)$. Finally, we have

$$\begin{aligned}
E[Z_j^2|\mathcal{F}_{t_j}] &= [f''(W_{t_j})]^2 E[(W_{t_{j+1}} - W_{t_j})^4 - 2(t_{j+1} - t_j)(W_{t_{j+1}} - W_{t_j})^2 + (t_{j+1} - t_j)^2|\mathcal{F}_{t_j}] \\
&= [f''(W_{t_j})]^2 (E[W^4_{t_{j+1}-t_j}] - 2(t_{j+1} - t_j)E[W^2_{t_{j+1}-t_j}] + (t_{j+1} - t_j)^2) \\
&= [f''(W_{t_j})]^2 [3(t_{j+1} - t_j)^2 - 2(t_{j+1} - t_j)^2 + (t_{j+1} - t_j)^2] \\
&= 2[f''(W_{t_j})]^2 (t_{j+1} - t_j)^2,
\end{aligned}$$

where we used the independence of Browian motion increment and the fact that $E[X^4] = 3E[X^2]^2$ if X is Gaussian with mean 0. □

(ii)

Proof. $E[\sum_{j=0}^{n-1} Z_j] = E[\sum_{j=0}^{n-1} E[Z_j|\mathcal{F}_{t_j}]] = 0$ by part (i). □

(iii)

Proof.

$$\begin{aligned}
Var[\sum_{j=0}^{n-1} Z_j] &= E[(\sum_{j=0}^{n-1} Z_j)^2] \\
&= E[\sum_{j=0}^{n-1} Z_j^2 + 2 \sum_{0 \le i < j \le n-1} Z_i Z_j] \\
&= \sum_{j=0}^{n-1} E[E[Z_j^2|\mathcal{F}_{t_j}]] + 2 \sum_{0 \le i < j \le n-1} E[Z_i E[Z_j|\mathcal{F}_{t_j}]] \\
&= \sum_{j=0}^{n-1} E[2[f''(W_{t_j})]^2 (t_{j+1} - t_j)^2] \\
&= \sum_{j=0}^{n-1} 2E[(f''(W_{t_j}))^2](t_{j+1} - t_j)^2 \\
&\le 2 \max_{0 \le j \le n-1} |t_{j+1} - t_j| \cdot \sum_{j=0}^{n-1} E[(f''(W_{t_j}))^2](t_{j+1} - t_j) \\
&\to 0,
\end{aligned}$$

since $\sum_{j=0}^{n-1} E[(f''(W_{t_j}))^2](t_{j+1} - t_j) \to \int_0^T E[(f''(W_t))^2]dt < \infty$. □

4.15. (i)

Proof. B_i is a local martingale with

$$(dB_i(t))^2 = \left(\sum_{j=1}^d \frac{\sigma_{ij}(t)}{\sigma_i(t)} dW_j(t) \right)^2 = \sum_{j=1}^d \frac{\sigma_{ij}^2(t)}{\sigma_i^2(t)} dt = dt.$$

So B_i is a Brownian motion. □

(ii)

Proof.

$$\begin{aligned} dB_i(t)dB_k(t) &= \left[\sum_{j=1}^{d} \frac{\sigma_{ij}(t)}{\sigma_i(t)} dW_j(t)\right]\left[\sum_{l=1}^{d} \frac{\sigma_{kl}(t)}{\sigma_k(t)} dW_l(t)\right] \\ &= \sum_{1\leq j,\, l\leq d} \frac{\sigma_{ij}(t)\sigma_{kl}(t)}{\sigma_i(t)\sigma_k(t)} dW_j(t)dW_l(t) \\ &= \sum_{j=1}^{d} \frac{\sigma_{ij}(t)\sigma_{kj}(t)}{\sigma_i(t)\sigma_k(t)} dt \\ &= \rho_{ik}(t)dt. \end{aligned}$$

□

4.16.

Proof. To find the m independent Brownion motion $W_1(t), \cdots, W_m(t)$, we need to find $A(t) = (a_{ij}(t))$ so that

$$(dB_1(t), \cdots, dB_m(t))^{tr} = A(t)(dW_1(t), \cdots, dW_m(t))^{tr},$$

or equivalently

$$(dW_1(t), \cdots, dW_m(t))^{tr} = A(t)^{-1}(dB_1(t), \cdots, dB_m(t))^{tr},$$

and

$$\begin{aligned} & (dW_1(t), \cdots, dW_m(t))^{tr}(dW_1(t), \cdots, dW_m(t)) \\ = \;& A(t)^{-1}(dB_1(t), \cdots, dB_m(t))^{tr}(dB_1(t), \cdots, dB_m(t))(A(t)^{-1})^{tr}dt \\ = \;& I_{m\times m}dt, \end{aligned}$$

where $I_{m\times m}$ is the $m \times m$ unit matrix. By the condition $dB_i(t)dB_k(t) = \rho_{ik}(t)dt$, we get

$$(dB_1(t), \cdots, dB_m(t))^{tr}(dB_1(t), \cdots, dB_m(t)) = C(t).$$

So $A(t)^{-1}C(t)(A(t)^{-1})^{tr} = I_{m\times m}$, which gives $C(t) = A(t)A(t)^{tr}$. This motivates us to define A as the square root of C. Reverse the above analysis, we obtain a formal proof. □

4.17.

Proof. We will try to solve all the sub-problems in a single, long solution. We start with the general X_i:

$$X_i(t) = X_i(0) + \int_0^t \theta_i(u)du + \int_0^t \sigma_i(u)dB_i(u), \; i = 1, 2.$$

The goal is to show

$$\lim_{\epsilon\downarrow 0} \frac{C(\epsilon)}{\sqrt{V_1(\epsilon)V_2(\epsilon)}} = \rho(t_0).$$

First, for $i = 1, 2$, we have

$$\begin{aligned} M_i(\epsilon) &= E[X_i(t_0+\epsilon) - X_i(t_0)|\mathcal{F}_{t_0}] \\ &= E\left[\int_{t_0}^{t_0+\epsilon} \Theta_i(u)du + \int_{t_0}^{t_0+\epsilon} \sigma_i(u)dB_i(u)|\mathcal{F}_{t_0}\right] \\ &= \Theta_i(t_0)\epsilon + E\left[\int_{t_0}^{t_0+\epsilon} (\Theta_i(u) - \Theta_i(t_0))du|\mathcal{F}_{t_0}\right]. \end{aligned}$$

By Conditional Jensen's Inequality,

$$\left|E\left[\int_{t_0}^{t_0+\epsilon}(\Theta_i(u)-\Theta_i(t_0))du|\mathcal{F}_{t_0}\right]\right| \le E\left[\int_{t_0}^{t_0+\epsilon}|\Theta_i(u)-\Theta_i(t_0)|du|\mathcal{F}_{t_0}\right]$$

Since $\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}|\Theta_i(u)-\Theta_i(t_0)|du \le 2M$ and $\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}|\Theta_i(u)-\Theta_i(t_0)|du = 0$ by the continuity of Θ_i, the Dominated Convergence Theorem under Conditional Expectation implies

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}E\left[\int_{t_0}^{t_0+\epsilon}|\Theta_i(u)-\Theta_i(t_0)|du|\mathcal{F}_{t_0}\right] = E\left[\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}|\Theta_i(u)-\Theta_i(t_0)|du|\mathcal{F}_{t_0}\right] = 0.$$

So $M_i(\epsilon) = \Theta_i(t_0)\epsilon + o(\epsilon)$. This proves (iii).

To calculate the variance and covariance, we note $Y_i(t) = \int_0^t \sigma_i(u)dB_i(u)$ is a martingale and by Itô's formula $Y_i(t)Y_j(t) - \int_0^t \sigma_i(u)\sigma_j(u)du$ is a martingale ($i = 1, 2$). So

$$\begin{aligned}
& E[(X_i(t_0+\epsilon)-X_i(t_0))(X_j(t_0+\epsilon)-X_j(t_0))|\mathcal{F}_{t_0}] \\
= \; & E\left[\left(Y_i(t_0+\epsilon)-Y_i(t_0)+\int_{t_0}^{t_0+\epsilon}\Theta_i(u)du\right)\left(Y_j(t_0+\epsilon)-Y_j(t_0)+\int_{t_0}^{t_0+\epsilon}\Theta_j(u)du\right)|\mathcal{F}_{t_0}\right] \\
= \; & E\left[(Y_i(t_0+\epsilon)-Y_i(t_0))\left(Y_j(t_0+\epsilon)-Y_j(t_0)\right)|\mathcal{F}_{t_0}\right] + E\left[\int_{t_0}^{t_0+\epsilon}\Theta_i(u)du\int_{t_0}^{t_0+\epsilon}\Theta_j(u)du|\mathcal{F}_{t_0}\right] \\
& + E\left[(Y_i(t_0+\epsilon)-Y_i(t_0))\int_{t_0}^{t_0+\epsilon}\Theta_j(u)du|\mathcal{F}_{t_0}\right] + E\left[(Y_j(t_0+\epsilon)-Y_j(t_0))\int_{t_0}^{t_0+\epsilon}\Theta_i(u)du|\mathcal{F}_{t_0}\right] \\
= \; & I + II + III + IV.
\end{aligned}$$

$$I = E[Y_i(t_0+\epsilon)Y_j(t_0+\epsilon) - Y_i(t_0)Y_j(t_0)|\mathcal{F}_{t_0}] = E\left[\int_{t_0}^{t_0+\epsilon}\sigma_i(u)\sigma_j(u)\rho_{ij}(t)dt|\mathcal{F}_{t_0}\right].$$

By an argument similar to that involved in the proof of part (iii), we conclude $I = \sigma_i(t_0)\sigma_j(t_0)\rho_{ij}(t_0)\epsilon + o(\epsilon)$ and

$$\begin{aligned}
II &= E\left[\int_{t_0}^{t_0+\epsilon}(\Theta_i(u)-\Theta_i(t_0))du\int_{t_0}^{t_0+\epsilon}\Theta_j(u)du|\mathcal{F}_{t_0}\right] + \Theta_i(t_0)\epsilon E\left[\int_{t_0}^{t_0+\epsilon}\Theta_j(u)du|\mathcal{F}_{t_0}\right] \\
&= o(\epsilon) + (M_i(\epsilon)-o(\epsilon))M_j(\epsilon) \\
&= M_i(\epsilon)M_j(\epsilon) + o(\epsilon).
\end{aligned}$$

By Cauchy's inequality under conditional expectation (note $E[XY|\mathcal{F}]$ defines an inner product on $L^2(\Omega)$),

$$\begin{aligned}
III &\le E\left[|Y_i(t_0+\epsilon)-Y_i(t_0)|\int_{t_0}^{t_0+\epsilon}|\Theta_j(u)|du|\mathcal{F}_{t_0}\right] \\
&\le M\epsilon\sqrt{E[(Y_i(t_0+\epsilon)-Y_i(t_0))^2|\mathcal{F}_{t_0}]} \\
&\le M\epsilon\sqrt{E[Y_i(t_0+\epsilon)^2-Y_i(t_0)^2|\mathcal{F}_{t_0}]} \\
&\le M\epsilon\sqrt{E[\int_{t_0}^{t_0+\epsilon}\Theta_i(u)^2du|\mathcal{F}_{t_0}]} \\
&\le M\epsilon\cdot M\sqrt{\epsilon} \\
&= o(\epsilon)
\end{aligned}$$

Similarly, $IV = o(\epsilon)$. In summary, we have

$$E[(X_i(t_0+\epsilon)-X_i(t_0))(X_j(t_0+\epsilon)-X_j(t_0))|\mathcal{F}_{t_0}] = M_i(\epsilon)M_j(\epsilon) + \sigma_i(t_0)\sigma_j(t_0)\rho_{ij}(t_0)\epsilon + o(\epsilon) + o(\epsilon).$$

This proves part (iv) and (v). Finally,

$$\lim_{\epsilon\downarrow 0}\frac{C(\epsilon)}{\sqrt{V_1(\epsilon)V_2(\epsilon)}} = \lim_{\epsilon\downarrow 0}\frac{\rho(t_0)\sigma_1(t_0)\sigma_2(t_0)\epsilon + o(\epsilon)}{\sqrt{(\sigma_1^2(t_0)\epsilon + o(\epsilon))(\sigma_2^2(t_0)\epsilon + o(\epsilon))}} = \rho(t_0).$$

This proves part (vi). Part (i) and (ii) are consequences of general cases. $\square$

4.18. (i)

Proof.

$$d(e^{rt}\zeta_t) = (de^{-\theta W_t - \frac{1}{2}\theta^2 t}) = -e^{-\theta W_t - \frac{1}{2}\theta^2 t}\theta dW_t = -\theta(e^{rt}\zeta_t)dW_t,$$

where for the second "=", we used the fact that $e^{-\theta W_t - \frac{1}{2}\theta^2 t}$ solves $dX_t = -\theta X_t dW_t$. Since $d(e^{rt}\zeta_t) = re^{rt}\zeta_t dt + e^{rt}d\zeta_t$, we get $d\zeta_t = -\theta\zeta_t dW_t - r\zeta_t dt$. $\square$

(ii)

Proof.

$$\begin{aligned}
d(\zeta_t X_t) &= \zeta_t dX_t + X_t d\zeta_t + dX_t d\zeta_t \\
&= \zeta_t(rX_t dt + \Delta_t(\alpha - r)S_t dt + \Delta_t \sigma S_t dW_t) + X_t(-\theta\zeta_t dW_t - r\zeta_t dt) \\
&\quad +(rX_t dt + \Delta_t(\alpha - r)S_t dt + \Delta_t\sigma S_t dW_t)(-\theta\zeta_t dW_t - r\zeta_t dt) \\
&= \zeta_t(\Delta_t(\alpha - r)S_t dt + \Delta_t\sigma S_t dW_t) - \theta X_t\zeta_t dW_t - \theta\Delta_t\sigma S_t\zeta_t dt \\
&= \zeta_t\Delta_t\sigma S_t dW_t - \theta X_t \zeta_t dW_t.
\end{aligned}$$

So $\zeta_t X_t$ is a martingale. $\square$

(iii)

Proof. By part (ii), $X_0 = \zeta_0 X_0 = E[\zeta_T X_t] = E[\zeta_T V_T]$. (This can be seen as a version of risk-neutral pricing, only that the pricing is carried out under the actual probability measure.) $\square$

4.19. (i)

Proof. B_t is a local martingale with $[B]_t = \int_0^t sign(W_s)^2 ds = t$. So by Lévy's theorem, B_t is a Brownian motion. $\square$

(ii)

Proof. $d(B_t W_t) = B_t dW_t + sign(W_t)W_t dW_t + sign(W_t)dt$. Integrate both sides of the resulting equation and the expectation, we get

$$E[B_t W_t] = \int_0^t E[sign(W_s)]ds = \int_0^t E[1_{\{W_s \geq 0\}} - 1_{\{W_s < 0\}}]ds = \frac{1}{2}t - \frac{1}{2}t = 0.$$

$\square$

(iii)

Proof. By Itô's formula, $dW_t^2 = 2W_t dW_t + dt$. $\square$

(iv)

Proof. By Itô's formula,

$$\begin{aligned}
d(B_t W_t^2) &= B_t dW_t^2 + W_t^2 dB_t + dB_t dW_t^2 \\
&= B_t(2W_t dW_t + dt) + W_t^2 sign(W_t)dW_t + sign(W_t)dW_t(2W_t dW_t + dt) \\
&= 2B_t W_t dW_t + B_t dt + sign(W_t)W_t^2 dW_t + 2sign(W_t)W_t dt.
\end{aligned}$$

So

$$\begin{aligned}
E[B_tW_t^2] &= E[\int_0^t B_s ds] + 2E[\int_0^t sign(W_s)W_s ds] \\
&= \int_0^t E[B_s]ds + 2\int_0^t E[sign(W_s)W_s]ds \\
&= 2\int_0^t (E[W_s 1_{\{W_s\geq 0\}}] - E[W_s 1_{\{W_s<0\}}])ds \\
&= 4\int_0^t \int_0^\infty x\frac{e^{-\frac{x^2}{2s}}}{\sqrt{2\pi s}}dxds \\
&= 4\int_0^t \sqrt{\frac{s}{2\pi}}ds \\
&\neq 0 = E[B_t]\cdot E[W_t^2].
\end{aligned}$$

Since $E[B_tW_t^2] \neq E[B_t]\cdot E[W_t^2]$, B_t and W_t are not independent. □

4.20. (i)

Proof. $f(x) = \begin{cases} x-K, & \text{if } x \geq K \\ 0, & \text{if } x < K. \end{cases}$ So $f'(x) = \begin{cases} 1, & \text{if } x > K \\ \text{undefined}, & \text{if } x = K \\ 0, & \text{if } x < K \end{cases}$ and $f''(x) = \begin{cases} 0, & \text{if } x \neq K \\ \text{undefined}, & \text{if } x = K. \end{cases}$ □

(ii)

Proof. $E[f(W_T)] = \int_K^\infty (x-K)\frac{e^{-\frac{x^2}{2T}}}{\sqrt{2\pi T}}dx = \sqrt{\frac{T}{2\pi}}e^{-\frac{K^2}{2T}} - K\Phi(-\frac{K}{\sqrt{T}})$ where Φ is the distribution function of standard normal random variable. If we suppose $\int_0^T f''(W_t)dt = 0$, the expectation of RHS of (4.10.42) is equal to 0. So (4.10.42) cannot hold. □

(iii)

Proof. This is trivial to check. □

(iv)

Proof. If $x = K$, $\lim_{n\to\infty} f_n(x) = \frac{1}{8n} = 0$; if $x > K$, for n large enough, $x \geq K + \frac{1}{2n}$, so $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty}(x-K) = x-K$; if $x < K$, for n large enough, $x \leq K - \frac{1}{2n}$, so $lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} 0 = 0$. In summary, $lim_{n\to\infty} f_n(x) = (x-K)^+$. Similarly, we can show

$$\lim_{n\to\infty} f'_n(x) = \begin{cases} 0, & \text{if } x < K \\ \frac{1}{2}, & \text{if } x = K \\ 1, & \text{if } x > K. \end{cases} \tag{5}$$

□

(v)

Proof. Fix ω, so that $W_t(\omega) < K$ for any $t \in [0,T]$. Since $W_t(\omega)$ can obtain its maximum on $[0,T]$, there exists n_0, so that for any $n \geq n_0$, $\max_{0\leq t\leq T} W_t(\omega) < K - \frac{1}{2n}$. So

$$L_K(T)(\omega) = \lim_{n\to\infty} n\int_0^T 1_{(K-\frac{1}{2n}, K+\frac{1}{2n})}(W_t(\omega))dt = 0.$$

□

(vi)

Proof. Take expectation on both sides of the formula (4.10.45), we have

$$E[L_K(T)] = E[(W_T - K)^+] > 0.$$

So we cannot have $L_K(T) = 0$ a.s.. $\square$

Remark 1. *Cf. Kallenberg Theorem 22.5, or the paper by Alsmeyer and Jaeger (2005).*

4.21. (i)

Proof. There are two problems. First, the transaction cost could be big due to active trading; second, the purchases and sales cannot be made at exactly the same price K. For more details, see Hull [6]. $\square$

(ii)

Proof. No. The RHS of (4.10.26) is a martingale, so its expectation is 0. But $E[(S_T - K)^+] > 0$. So $X_T \neq (S_T - K)^+$. $\square$

5 Risk-Neutral Pricing

★ **Comments**:

1) *Heuristics to memorize the conditional Bayes formula (Lemma 5.2.2 on page 212 of the textbook)*

$$\widetilde{\mathbb{E}}[Y|\mathcal{F}(s)] = \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)].$$

We recall an example in Durrett [4, page 223] (Example 5.1.3): suppose Ω_1, Ω_2, $\cdots$ is a finite or infinite partition of Ω into disjoint sets, each of which has positive probability, and let $\mathcal{F} = \sigma(\Omega_1, \Omega_2, \cdots)$ be the σ-field generated by these sets. Then

$$\mathbb{E}[X|\mathcal{F}(s)] = \sum_i 1_{\Omega_i} \frac{\mathbb{E}[X; \Omega_i]}{\mathbb{P}(\Omega_i)}.$$

Therefore, if $\mathcal{F}(s) = \sigma(\Omega_1, \Omega_2, \cdots)$, we have

$$\widetilde{\mathbb{E}}[Y|\mathcal{F}(s)] = \sum_i 1_{\Omega_i} \frac{\widetilde{\mathbb{E}}[Y; \Omega_i]}{\widetilde{\mathbb{P}}(\Omega_i)} = \sum_i 1_{\Omega_i} \frac{\mathbb{E}[YZ(t); \Omega_i]}{\mathbb{E}[Z(s); \Omega_i]}.$$

Consequently, on any Ω_i,

$$\widetilde{\mathbb{E}}[Y|\mathcal{F}(s)] = \frac{\mathbb{E}[YZ(t); \Omega_i]/\mathbb{P}(\Omega_i)}{\mathbb{E}[Z(s); \Omega_i]/\mathbb{P}(\Omega_i)} = \frac{\mathbb{E}[YZ(t)|\mathcal{F}(s)]}{\mathbb{E}[Z(s)|\mathcal{F}(s)]} = \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)].$$

2) *Heuristics to memorize the one-dimensional Girsanov's Theorem (Theorem 5.2.3 on page 212 of the textbook): suppose under a change of measure with density process* $(Z(t))_{t\geq 0}$, $\widetilde{W}(t) = W(t) + \int_0^t \Theta(u)du$ *becomes a martingale, then we must have*

$$Z(t) = \exp\left\{-\int_0^t \Theta(u)dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)du\right\}.$$

Recall Z is a positive martingale under the original probability measure $\mathbb{P}$, so by martingale representation theorem under the Brownian filtration, Z must satisfy the SDE

$$dZ(t) = Z(t)X(t)dW(t)$$

for some adapted process $X(t)$.[7] Since an adapted process M is a $\widetilde{\mathbb{P}}$-martingale if and only if MZ is a $\mathbb{P}$-martingale,[8] we conclude $\widetilde{W}$ is a $\widetilde{\mathbb{P}}$-martingale if and only if $\widetilde{W}Z$ is a $\mathbb{P}$-martingale. Since

$$\begin{aligned} d[\widetilde{W}(t)Z(t)] &= Z(t)[dW(t)+\Theta(t)dt]+\widetilde{W}(t)Z(t)X(t)dW(t)+[dW(t)+\Theta(t)dt]Z(t)X(t)dW(t) \\ &= (\cdots)dW(t)+Z(t)[\Theta(t)+X(t)]dt, \end{aligned}$$

we must require $X(t)=-\Theta(t)$. So $Z(t)=\exp\left\{-\int_0^t\Theta(u)dW(u)-\frac{1}{2}\int_0^t\Theta^2(u)du\right\}$.

3) *Main idea of Example 5.4.4.* Combining formula (5.4.15) and (5.4.22), we have

$$d[D(t)X(t)]=\sum_{i=1}^m\Delta_i(t)d[D(t)S_i(t)]=\sum_{i=1}^m\Delta_i(t)D(t)S_i(t)[(\alpha_i(t)-R(t))dt+\sigma_i(t)dB_i(t)].$$

To create arbitrage, we want $D(t)X(t)$ to have a deterministic and positive return rate, that is,

$$\begin{cases}\sum_{i=1}^m\Delta_i(t)S_i(t)\sigma_i(t)=0\\ \sum_{i=1}^m\Delta_i(t)S_i(t)[\alpha_i(t)-R(t)]>0.\end{cases}$$

In Example 5.4.4, $i=2$, $\Delta_1(t)=\frac{1}{S_1(t)\sigma_1}$, $\Delta_2(t)=-\frac{1}{S_2(t)\sigma_2}$ and

$$\frac{\alpha_1-r}{\sigma_1}-\frac{\alpha-r}{\sigma_2}>0.$$

5.1. (i)

Proof.

$$\begin{aligned} df(X_t) &= f'(X_t)dt+\frac{1}{2}f''(X_t)d\langle X\rangle_t \\ &= f(X_t)(dX_t+\frac{1}{2}d\langle X\rangle_t) \\ &= f(X_t)\left[\sigma_tdW_t+(\alpha_t-R_t-\frac{1}{2}\sigma_t^2)dt+\frac{1}{2}\sigma_t^2dt\right] \\ &= f(X_t)(\alpha_t-R_t)dt+f(X_t)\sigma_tdW_t. \end{aligned}$$

This is formula (5.2.20). □

(ii)

Proof. $d(D_tS_t)=S_tdD_t+D_tdS_t+dD_tdS_t=-S_tR_tD_tdt+D_t\alpha_tS_tdt+D_t\sigma_tS_tdW_t=D_tS_t(\alpha_t-R_t)dt+D_tS_t\sigma_tdW_t$. This is formula (5.2.20). □

5.2.

Proof. By Lemma 5.2.2., $\widetilde{E}[D_TV_T|\mathcal{F}_t]=E\left[\frac{D_TV_TZ_T}{Z_t}|\mathcal{F}_t\right]$. Therefore (5.2.30) is equivalent to $D_tV_tZ_t=E[D_TV_TZ_T|\mathcal{F}_t]$. □

5.3. (i)

[7] The existence of X is easy: if $dZ(t)=\xi(t)dW(t)$, then $X(t)=\frac{\xi(t)}{Z(t)}$.

[8] To see this, note $\widetilde{\mathbb{E}}[M(t)|\mathcal{F}(s)]=\mathbb{E}[M(t)Z(t)|\mathcal{F}(s)]/Z(s)$.

Proof.

$$
\begin{aligned}
c_x(0,x) &= \frac{d}{dx}\widetilde{E}[e^{-rT}(xe^{\sigma\widetilde{W}_T+(r-\frac{1}{2}\sigma^2)T}-K)^+] \\
&= \widetilde{E}\left[e^{-rT}\frac{d}{dx}h(xe^{\sigma\widetilde{W}_T+(r-\frac{1}{2}\sigma^2)T})\right] \\
&= \widetilde{E}\left[e^{-rT}e^{\sigma\widetilde{W}_T+(r-\frac{1}{2}\sigma^2)T}1_{\{xe^{\sigma\widetilde{W}_T+(r-\frac{1}{2}\sigma^2)T}>K\}}\right] \\
&= e^{-\frac{1}{2}\sigma^2T}\widetilde{E}\left[e^{\sigma\widetilde{W}_T}1_{\{\widetilde{W}_T>\frac{1}{\sigma}(\ln\frac{K}{x}-(r-\frac{1}{2}\sigma^2)T)\}}\right] \\
&= e^{-\frac{1}{2}\sigma^2T}\widetilde{E}\left[e^{\sigma\sqrt{T}\frac{\widetilde{W}_T}{\sqrt{T}}}1_{\{\frac{\widetilde{W}_T}{\sqrt{T}}-\sigma\sqrt{T}>\frac{1}{\sigma\sqrt{T}}(\ln\frac{K}{x}-(r-\frac{1}{2}\sigma^2)T)-\sigma\sqrt{T}\}}\right] \\
&= e^{-\frac{1}{2}\sigma^2T}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}e^{\sigma\sqrt{T}z}1_{\{z-\sigma\sqrt{T}>-d_+(T,x)\}}dz \\
&= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{(z-\sigma\sqrt{T})^2}{2}}1_{\{z-\sigma\sqrt{T}>-d_+(T,x)\}}dz \\
&= N(d_+(T,x)).
\end{aligned}
$$

□

(ii)

Proof. If we set $\widehat{Z}_T = e^{\sigma\widetilde{W}_T-\frac{1}{2}\sigma^2T}$ and $\widehat{Z}_t = \widetilde{E}[\widehat{Z}_T|\mathcal{F}_t]$, then $\widehat{Z}$ is a $\widetilde{P}$-martingale, $\widehat{Z}_t > 0$ and $E[\widehat{Z}_T] = \widetilde{E}[e^{\sigma\widetilde{W}_T-\frac{1}{2}\sigma^2T}] = 1$. So if we define $\widehat{P}$ by $d\widehat{P} = Z_Td\widetilde{P}$ on $\mathcal{F}_T$, then $\widehat{P}$ is a probability measure equivalent to $\widetilde{P}$, and

$$c_x(0,x) = \widetilde{E}[\widehat{Z}_T1_{\{S_T>K\}}] = \widehat{P}(S_T>K).$$

Moreover, by Girsanov's Theorem, $\widehat{W}_t = \widetilde{W}_t + \int_0^t(-\sigma)du = \widetilde{W}_t - \sigma t$ is a $\widehat{P}$-Brownian motion (set $\Theta = -\sigma$ in Theorem 5.4.1.) □

(iii)

Proof. $S_T = xe^{\sigma\widetilde{W}_T+(r-\frac{1}{2}\sigma^2)T} = xe^{\sigma\widehat{W}_T+(r+\frac{1}{2}\sigma^2)T}$. So

$$\widehat{P}(S_T>K) = \widehat{P}(xe^{\sigma\widehat{W}_T+(r+\frac{1}{2}\sigma^2)T}>K) = \widehat{P}\left(\frac{\widehat{W}_T}{\sqrt{T}} > -d_+(T,x)\right) = N(d_+(T,x)).$$

□

5.4. First, a few typos. In the SDE for S, "$\sigma(t)d\widetilde{W}(t)$" $\to$ "$\sigma(t)S(t)d\widetilde{W}(t)$". In the first equation for $c(0,S(0))$, $E \to \widetilde{E}$. In the second equation for $c(0,S(0))$, the variable for BSM should be

$$BSM\left(T,S(0);K,\frac{1}{T}\int_0^T r(t)dt,\sqrt{\frac{1}{T}\int_0^T\sigma^2(t)dt}\right).$$

(i)

Proof. $d\ln S_t = \frac{dS_t}{S_t} - \frac{1}{2S_t^2}d\langle S\rangle_t = r_tdt + \sigma_td\widetilde{W}_t - \frac{1}{2}\sigma_t^2dt$. So $S_T = S_0\exp\{\int_0^T(r_t-\frac{1}{2}\sigma_t^2)dt + \int_0^T\sigma_td\widetilde{W}_t\}$. Let $X = \int_0^T(r_t-\frac{1}{2}\sigma_t^2)dt + \int_0^T\sigma_td\widetilde{W}_t$. The first term in the expression of X is a number and the second term is a Gaussian random variable $N(0,\int_0^T\sigma_t^2dt)$, since both r and σ ar deterministic. Therefore, $S_T = S_0e^X$, with $X \sim N(\int_0^T(r_t-\frac{1}{2}\sigma_t^2)dt, \int_0^T\sigma_t^2dt)$,. □

(ii)

Proof. For the standard BSM model with constant volatility Σ and interest rate R, under the risk-neutral measure, we have $S_T = S_0 e^Y$, where $Y = (R - \frac{1}{2}\Sigma^2)T + \Sigma\widetilde{W}_T \sim N((R-\frac{1}{2}\Sigma^2)T, \Sigma^2 T)$, and $\widetilde{E}[(S_0e^Y - K)^+] = e^{RT}BSM(T, S_0; K, R, \Sigma)$. Note $R = \frac{1}{T}(E[Y] + \frac{1}{2}Var(Y))$ and $\Sigma = \sqrt{\frac{1}{T}Var(Y)}$, we can get

$$\widetilde{E}[(S_0e^Y - K)^+] = e^{E[Y]+\frac{1}{2}Var(Y)}BSM\left(T, S_0; K, \frac{1}{T}\left(E[Y] + \frac{1}{2}Var(Y)\right), \sqrt{\frac{1}{T}Var(Y)}\right).$$

So for the model in this problem,

$$\begin{aligned} c(0,S_0) &= e^{-\int_0^T r_t dt}\widetilde{E}[(S_0e^X - K)^+] \\ &= e^{-\int_0^T r_t dt}e^{E[X]+\frac{1}{2}Var(X)}BSM\left(T, S_0; K, \frac{1}{T}\left(E[X]+\frac{1}{2}Var(X)\right), \sqrt{\frac{1}{T}Var(X)}\right) \\ &= BSM\left(T, S_0; K, \frac{1}{T}\int_0^T r_t dt, \sqrt{\frac{1}{T}\int_0^T \sigma_t^2 dt}\right). \end{aligned}$$

□

5.5. (i)

Proof. Let $f(x) = \frac{1}{x}$, then $f'(x) = -\frac{1}{x^2}$ and $f''(x) = \frac{2}{x^3}$. Note $dZ_t = -Z_t\Theta_t dW_t$, so

$$d\left(\frac{1}{Z_t}\right) = f'(Z_t)dZ_t + \frac{1}{2}f''(Z_t)dZ_t dZ_t = -\frac{1}{Z_t^2}(-Z_t)\Theta_t dW_t + \frac{1}{2}\frac{2}{Z_t^3}Z_t^2\Theta_t^2 dt = \frac{\Theta_t}{Z_t}dW_t + \frac{\Theta_t^2}{Z_t}dt.$$

□

(ii)

Proof. By Lemma 5.2.2., for $s, t \geq 0$ with $s < t$, $\widetilde{M}_s = \widetilde{E}[\widetilde{M}_t|\mathcal{F}_s] = E\left[\frac{Z_t\widetilde{M}_t}{Z_s}|\mathcal{F}_s\right]$. That is, $E[Z_t\widetilde{M}_t|\mathcal{F}_s] = Z_s\widetilde{M}_s$. So $M = Z\widetilde{M}$ is a P-martingale. □

(iii)

Proof.

$$d\widetilde{M}_t = d\left(M_t \cdot \frac{1}{Z_t}\right) = \frac{1}{Z_t}dM_t + M_t d\frac{1}{Z_t} + dM_t d\frac{1}{Z_t} = \frac{\Gamma_t}{Z_t}dW_t + \frac{M_t\Theta_t}{Z_t}dW_t + \frac{M_t\Theta_t^2}{Z_t}dt + \frac{\Gamma_t\Theta_t}{Z_t}dt.$$

□

(iv)

Proof. In part (iii), we have

$$d\widetilde{M}_t = \frac{\Gamma_t}{Z_t}dW_t + \frac{M_t\Theta_t}{Z_t}dW_t + \frac{M_t\Theta_t^2}{Z_t}dt + \frac{\Gamma_t\Theta_t}{Z_t}dt = \frac{\Gamma_t}{Z_t}(dW_t + \Theta_t dt) + \frac{M_t\Theta_t}{Z_t}(dW_t + \Theta_t dt).$$

Let $\widetilde{\Gamma}_t = \frac{\Gamma_t + M_t\Theta_t}{Z_t}$, then $d\widetilde{M}_t = \widetilde{\Gamma}_t d\widetilde{W}_t$. This proves Corollary 5.3.2. □

5.6.

Proof. By Theorem 4.6.5, it suffices to show $\widetilde{W}_i(t)$ is an $\mathcal{F}_t$-martingale under $\widetilde{P}$ and $[\widetilde{W}_i, \widetilde{W}_j](t) = t\delta_{ij}$ $(i, j = 1, 2)$. Indeed, for $i = 1, 2$, $\widetilde{W}_i(t)$ is an $\mathcal{F}_t$-martingale under $\widetilde{P}$ if and only if $\widetilde{W}_i(t)Z_t$ is an $\mathcal{F}_t$-martingale under P, since

$$\widetilde{E}[\widetilde{W}_i(t)|\mathcal{F}_s] = E\left[\frac{\widetilde{W}_i(t)Z_t}{Z_s}|\mathcal{F}_s\right].$$

By Itô's product formula, we have

$$\begin{aligned} d(\widetilde{W}_i(t)Z_t) &= \widetilde{W}_i(t)dZ_t + Z_t d\widetilde{W}_i(t) + dZ_t d\widetilde{W}_i(t) \\ &= \widetilde{W}_i(t)(-Z_t)\Theta(t) \cdot dW_t + Z_t(dW_i(t) + \Theta_i(t)dt) + (-Z_t\Theta_t \cdot dW_t)(dW_i(t) + \Theta_i(t)dt) \\ &= \widetilde{W}_i(t)(-Z_t)\sum_{j=1}^{d}\Theta_j(t)dW_j(t) + Z_t(dW_i(t) + \Theta_i(t)dt) - Z_t\Theta_i(t)dt \\ &= \widetilde{W}_i(t)(-Z_t)\sum_{j=1}^{d}\Theta_j(t)dW_j(t) + Z_t dW_i(t) \end{aligned}$$

This shows $\widetilde{W}_i(t)Z_t$ is an $\mathcal{F}_t$-martingale under P. So $\widetilde{W}_i(t)$ is an $\mathcal{F}_t$-martingale under $\widetilde{P}$. Moreover,

$$[\widetilde{W}_i, \widetilde{W}_j](t) = \left[W_i + \int_0^\cdot \Theta_i(s)ds, W_j + \int_0^\cdot \Theta_j(s)ds\right](t) = [W_i, W_j](t) = t\delta_{ij}.$$

Combined, this proves the two-dimensional Girsanov's Theorem. □

5.7. (i)

Proof. Let a be any strictly positive number. We define $X_2(t) = (a + X_1(t))D(t)^{-1}$. Then

$$P\left(X_2(T) \geq \frac{X_2(0)}{D(T)}\right) = P(a + X_1(T) \geq a) = P(X_1(T) \geq 0) = 1,$$

and $P\left(X_2(T) > \frac{X_2(0)}{D(T)}\right) = P(X_1(T) > 0) > 0$, since a is arbitrary, we have proved the claim of this problem. □

Remark 2. *The intuition is that we invest the positive starting fund a into the money market account, and construct portfolio X_1 from zero cost. Their sum should be able to beat the return of money market account.*

(ii)

Proof. We define $X_1(t) = X_2(t)D(t) - X_2(0)$. Then $X_1(0) = 0$,

$$P(X_1(T) \geq 0) = P\left(X_2(T) \geq \frac{X_2(0)}{D(T)}\right) = 1,\ P(X_1(T) > 0) = P\left(X_2(T) > \frac{X_2(0)}{D(T)}\right) > 0.$$

□

5.8. The basic idea is that for any positive $\widetilde{P}$-martingale M, $dM_t = M_t \cdot \frac{1}{M_t}dM_t$. By Martingale Representation Theorem, $dM_t = \widetilde{\Gamma}_t d\widetilde{W}_t$ for some adapted process $\widetilde{\Gamma}_t$. So $dM_t = M_t(\frac{\widetilde{\Gamma}_t}{M_t})d\widetilde{W}_t$, i.e. any positive martingale must be the exponential of an integral w.r.t. Brownian motion. Taking into account discounting factor and apply Itô's product rule, we can show every strictly positive asset is a generalized geometric Brownian motion.

(i)

Proof. $V_t D_t = \widetilde{E}[e^{-\int_0^T R_u du}V_T|\mathcal{F}_t] = \widetilde{E}[D_T V_T|\mathcal{F}_t]$. So $(D_t V_t)_{t\geq 0}$ is a $\widetilde{P}$-martingale. By Martingale Representation Theorem, there exists an adapted process $\widetilde{\Gamma}_t$, $0 \leq t \leq T$, such that $D_t V_t = \int_0^t \widetilde{\Gamma}_s d\widetilde{W}_s$, or equivalently, $V_t = D_t^{-1}\int_0^t \widetilde{\Gamma}_s d\widetilde{W}_s$. Differentiate both sides of the equation, we get $dV_t = R_t D_t^{-1}\int_0^t \widetilde{\Gamma}_s d\widetilde{W}_s dt + D_t^{-1}\widetilde{\Gamma}_t d\widetilde{W}_t$, i.e. $dV_t = R_t V_t dt + \frac{\widetilde{\Gamma}_t}{D_t}dW_t$. □

(ii)

Proof. We prove the following more general lemma.

Lemma 5.1. *Let X be an almost surely positive random variable (i.e. $X>0$ a.s.) defined on the probability space $(\Omega,\mathcal{G},P)$. Let $\mathcal{F}$ be a sub σ-algebra of $\mathcal{G}$, then $Y=E[X|\mathcal{F}]>0$ a.s.*

Proof. By the property of conditional expectation $Y_t\geq 0$ a.s. Let $A=\{Y=0\}$, we shall show $P(A)=0$. Indeed, note $A\in\mathcal{F}$, $0=E[YI_A]=E[E[X|\mathcal{F}]I_A]=E[XI_A]=E[X1_{A\cap\{X\geq 1\}}]+\sum_{n=1}^\infty E[X1_{A\cap\{\frac{1}{n}>X\geq\frac{1}{n+1}\}}]\geq P(A\cap\{X\geq 1\})+\sum_{n=1}^\infty\frac{1}{n+1}P(A\cap\{\frac{1}{n}>X\geq\frac{1}{n+1}\})$. So $P(A\cap\{X\geq 1\})=0$ and $P(A\cap\{\frac{1}{n}>X\geq\frac{1}{n+1}\})=0$, $\forall n\geq 1$. This in turn implies $P(A)=P(A\cap\{X>0\})=P(A\cap\{X\geq 1\})+\sum_{n=1}^\infty P(A\cap\{\frac{1}{n}>X\geq\frac{1}{n+1}\})=0$. ☐

By the above lemma, it is clear that for each $t\in[0,T]$, $V_t=\widetilde{E}[e^{-\int_t^T R_u du}V_T|\mathcal{F}_t]>0$ a.s.. Moreover, by a classical result of martingale theory (Revuz and Yor [11], Chapter II, Proposition (3.4)), we have the following stronger result: for a.s. ω, $V_t(\omega)>0$ for any $t\in[0,T]$. ☐

(iii)

Proof. By (ii), $V>0$ a.s., so $dV_t=V_t\frac{1}{V_t}dV_t=V_t\frac{1}{V_t}\left(R_tV_tdt+\frac{\widetilde{\Gamma}_t}{D_t}d\widetilde{W}_t\right)=V_tR_tdt+V_t\frac{\widetilde{\Gamma}_t}{V_tD_t}d\widetilde{W}_t=R_tV_tdt+\sigma_tV_td\widetilde{W}_t$, where $\sigma_t=\frac{\widetilde{\Gamma}_t}{V_tD_t}$. This shows V follows a generalized geometric Brownian motion. ☐

5.9.

Proof. $c(0,T,x,K)=xN(d_+)-Ke^{-rT}N(d_-)$ with $d_\pm=\frac{1}{\sigma\sqrt{T}}(\ln\frac{x}{K}+(r\pm\frac{1}{2}\sigma^2)T)$. Let $f(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$, then $f'(y)=-yf(y)$,

$$
\begin{aligned}
c_K(0,T,x,K) &= xf(d_+)\frac{\partial d_+}{\partial y}-e^{-rT}N(d_-)-Ke^{-rT}f(d_-)\frac{\partial d_-}{\partial y}\\
&= xf(d_+)\frac{-1}{\sigma\sqrt{T}K}-e^{-rT}N(d_-)+e^{-rT}f(d_-)\frac{1}{\sigma\sqrt{T}},
\end{aligned}
$$

and

$$
\begin{aligned}
&c_{KK}(0,T,x,K)\\
=\;& xf(d_+)\frac{1}{\sigma\sqrt{T}K^2}-\frac{x}{\sigma\sqrt{T}K}f(d_+)(-d_+)\frac{\partial d_+}{\partial y}-e^{-rT}f(d_-)\frac{\partial d_-}{\partial y}+\frac{e^{-rT}}{\sigma\sqrt{T}}(-d_-)f(d_-)\frac{d_-}{\partial y}\\
=\;& \frac{x}{\sigma\sqrt{T}K^2}f(d_+)+\frac{xd_+}{\sigma\sqrt{T}K}f(d_+)\frac{-1}{K\sigma\sqrt{T}}-e^{-rT}f(d_-)\frac{-1}{K\sigma\sqrt{T}}-\frac{e^{-rT}d_-}{\sigma\sqrt{T}}f(d_-)\frac{-1}{K\sigma\sqrt{T}}\\
=\;& f(d_+)\frac{x}{K^2\sigma\sqrt{T}}[1-\frac{d_+}{\sigma\sqrt{T}}]+\frac{e^{-rT}f(d_-)}{K\sigma\sqrt{T}}[1+\frac{d_-}{\sigma\sqrt{T}}]\\
=\;& \frac{e^{-rT}}{K\sigma^2T}f(d_-)d_+-\frac{x}{K^2\sigma^2T}f(d_+)d_-.
\end{aligned}
$$

☐

5.10. (i)

Proof. At time t_0, the value of the chooser option is $V(t_0)=\max\{C(t_0),P(t_0)\}=\max\{C(t_0),C(t_0)-F(t_0)\}=C(t_0)+\max\{0,-F(t_0)\}=C(t_0)+(e^{-r(T-t_0)}K-S(t_0))^+$. ☐

(ii)

Proof. By the risk-neutral pricing formula, $V(0) = \widetilde{E}[e^{-rt_0}V(t_0)] = \widetilde{E}[e^{-rt_0}C(t_0) + (e^{-rT}K - e^{-rt_0}S(t_0)^+] = C(0) + \widetilde{E}[e^{-rt_0}(e^{-r(T-t_0)}K - S(t_0))^+]$. The first term is the value of a call expiring at time T with strike price K and the second term is the value of a put expiring at time t_0 with strike price $e^{-r(T-t_0)}K$. □

5.11.

Proof. We first make an analysis which leads to the hint, then we give a formal proof.

(Analysis) If we want to construct a portfolio X that exactly replicates the cash flow, we must find a solution to the backward SDE

$$\begin{cases} dX_t = \Delta_t dS_t + R_t(X_t - \Delta_t S_t)dt - C_t dt \\ X_T = 0. \end{cases}$$

Multiply D_t on both sides of the first equation and apply Itô's product rule, we get $d(D_tX_t) = \Delta_t d(D_tS_t) - C_tD_tdt$. Integrate from 0 to T, we have $D_TX_T - D_0X_0 = \int_0^T \Delta_t d(D_tS_t) - \int_0^T C_tD_tdt$. By the terminal condition, we get $X_0 = D_0^{-1}(\int_0^T C_tD_tdt - \int_0^T \Delta_t d(D_tS_t))$. X_0 is the theoretical, no-arbitrage price of the cash flow, provided we can find a trading strategy Δ that solves the BSDE. Note the SDE for S gives $d(D_tS_t) = (D_tS_t)\sigma_t(\theta_t dt + dW_t)$, where $\theta_t = \frac{\alpha_t - R_t}{\sigma_t}$. Take the proper change of measure so that $\widetilde{W}_t = \int_0^t \theta_s ds + W_t$ is a Brownian motion under the new measure $\widetilde{P}$, we get

$$\int_0^T C_tD_tdt = D_0X_0 + \int_0^T \Delta_t d(D_tS_t) = D_0X_0 + \int_0^T \Delta_t(D_tS_t)\sigma_t d\widetilde{W}_t.$$

This says the random variable $\int_0^T C_tD_tdt$ has a stochastic integral representation $D_0X_0 + \int_0^T \Delta_tD_tS_t\sigma_t d\widetilde{W}_t$. This inspires us to consider the martingale generated by $\int_0^T C_tD_tdt$, so that we can apply Martingale Representation Theorem and get a formula for Δ by comparison of the integrands.

(Formal proof) Let $M_T = \int_0^T C_tD_tdt$, and $M_t = \widetilde{E}[M_T|\mathcal{F}_t]$. Then by Martingale Representation Theorem, we can find an adapted process $\widetilde{\Gamma}_t$, so that $M_t = M_0 + \int_0^t \widetilde{\Gamma}_t d\widetilde{W}_t$. If we set $\Delta_t = \frac{\widetilde{\Gamma}_t}{D_tS_t\sigma_t}$, we can check $X_t = D_t^{-1}(D_0X_0 + \int_0^t \Delta_u d(D_uS_u) - \int_0^t C_uD_udu)$, with $X_0 = M_0 = \widetilde{E}[\int_0^T C_tD_tdt]$ solves the SDE

$$\begin{cases} dX_t = \Delta_t dS_t + R_t(X_t - \Delta_t S_t)dt - C_t dt \\ X_T = 0. \end{cases}$$

Indeed, it is easy to see that X satisfies the first equation. To check the terminal condition, we note $X_TD_T = D_0X_0 + \int_0^T \Delta_tD_tS_t\sigma_t d\widetilde{W}_t - \int_0^T C_tD_tdt = M_0 + \int_0^T \widetilde{\Gamma}_t d\widetilde{W}_t - M_T = 0$. So $X_T = 0$. Thus, we have found a trading strategy Δ, so that the corresponding portfolio X replicates the cash flow and has zero terminal value. So $X_0 = \widetilde{E}[\int_0^T C_tD_tdt]$ is the no-arbitrage price of the cash flow at time zero. □

Remark 3. *As shown in the analysis, $d(D_tX_t) = \Delta_t d(D_tS_t) - C_tD_tdt$. Integrate from t to T, we get $0 - D_tX_t = \int_t^T \Delta_u d(D_uS_u) - \int_t^T C_uD_udu$. Take conditional expectation w.r.t. $\mathcal{F}_t$ on both sides, we get $-D_tX_t = -\widetilde{E}[\int_t^T C_uD_udu|\mathcal{F}_t]$. So $X_t = D_t^{-1}\widetilde{E}[\int_t^T C_uD_udu|\mathcal{F}_t]$. This is the no-arbitrage price of the cash flow at time t, and we have justified formula (5.6.10) in the textbook.*

5.12. (i)

Proof. $d\widetilde{B}_i(t) = dB_i(t) + \gamma_i(t)dt = \sum_{j=1}^d \frac{\sigma_{ij}(t)}{\sigma_i(t)}dW_j(t) + \sum_{j=1}^d \frac{\sigma_{ij}(t)}{\sigma_i(t)}\Theta_j(t)dt = \sum_{j=1}^d \frac{\sigma_{ij}(t)}{\sigma_i(t)}d\widetilde{W}_j(t)$. So B_i is a martingale. Since $d\widetilde{B}_i(t)d\widetilde{B}_i(t) = \sum_{j=1}^d \frac{\sigma_{ij}(t)^2}{\sigma_i(t)^2}dt = dt$, by Lévy's Theorem, $\widetilde{B}_i$ is a Brownian motion under $\widetilde{P}$. □

(ii)

Proof.

$$\begin{aligned}
dS_i(t) &= R(t)S_i(t)dt + \sigma_i(t)S_i(t)d\widetilde{B}_i(t) + (\alpha_i(t) - R(t))S_i(t)dt - \sigma_i(t)S_i(t)\gamma_i(t)dt \\
&= R(t)S_i(t)dt + \sigma_i(t)S_i(t)d\widetilde{B}_i(t) + \sum_{j=1}^{d}\sigma_{ij}(t)\Theta_j(t)S_i(t)dt - S_i(t)\sum_{j=1}^{d}\sigma_{ij}(t)\Theta_j(t)dt \\
&= R(t)S_i(t)dt + \sigma_i(t)S_i(t)d\widetilde{B}_i(t).
\end{aligned}$$

□

(iii)

Proof. $d\widetilde{B}_i(t)d\widetilde{B}_k(t) = (dB_i(t) + \gamma_i(t)dt)(dB_j(t) + \gamma_j(t)dt) = dB_i(t)dB_j(t) = \rho_{ik}(t)dt.$ □

(iv)

Proof. By Itô's product rule and martingale property,

$$\begin{aligned}
E[B_i(t)B_k(t)] &= E[\int_0^t B_i(s)dB_k(s)] + E[\int_0^t B_k(s)dB_i(s)] + E[\int_0^t dB_i(s)dB_k(s)] \\
&= E[\int_0^t \rho_{ik}(s)ds] = \int_0^t \rho_{ik}(s)ds.
\end{aligned}$$

Similarly, by part (iii), we can show $\widetilde{E}[\widetilde{B}_i(t)\widetilde{B}_k(t)] = \int_0^t \rho_{ik}(s)ds$. □

(v)

Proof. By Itô's product formula,

$$E[B_1(t)B_2(t)] = E[\int_0^t sign(W_1(u))du] = \int_0^t [P(W_1(u) \geq 0) - P(W_1(u) < 0)]du = 0.$$

Meanwhile,

$$\begin{aligned}
\widetilde{E}[\widetilde{B}_1(t)\widetilde{B}_2(t)] &= \widetilde{E}[\int_0^t sign(W_1(u))du \\
&= \int_0^t [\widetilde{P}(W_1(u) \geq 0) - \widetilde{P}(W_1(u) < 0)]du \\
&= \int_0^t [\widetilde{P}(\widetilde{W}_1(u) \geq u) - \widetilde{P}(\widetilde{W}_1(u) < u)]du \\
&= \int_0^t 2\left(\frac{1}{2} - \widetilde{P}(\widetilde{W}_1(u) < u)\right)du \\
&< 0,
\end{aligned}$$

for any $t > 0$. So $E[B_1(t)B_2(t)] = \widetilde{E}[\widetilde{B}_1(t)\widetilde{B}_2(t)]$ for all $t > 0$. □

5.13. (i)

Proof. $\widetilde{E}[W_1(t)] = \widetilde{E}[\widetilde{W}_1(t)] = 0$ and $\widetilde{E}[W_2(t)] = \widetilde{E}[\widetilde{W}_2(t) - \int_0^t \widetilde{W}_1(u)du] = 0$, for all $t \in [0, T]$. □

(ii)

Proof.

$$\begin{aligned}
\widetilde{Cov}[W_1(T), W_2(T)] &= \widetilde{E}[W_1(T)W_2(T)] \\
&= \widetilde{E}\left[\int_0^T W_1(t)dW_2(t) + \int_0^T W_2(t)dW_1(t)\right] \\
&= \widetilde{E}\left[\int_0^T \widetilde{W}_1(t)(d\widetilde{W}_2(t) - \widetilde{W}_1(t)dt)\right] + \widetilde{E}\left[\int_0^T W_2(t)d\widetilde{W}_1(t)\right] \\
&= -\widetilde{E}\left[\int_0^T \widetilde{W}_1(t)^2 dt\right] \\
&= -\int_0^T t dt \\
&= -\frac{1}{2}T^2.
\end{aligned}$$

□

5.14. Equation (5.9.6) can be transformed into $d(e^{-rt}X_t) = \Delta_t[d(e^{-rt}S_t) - ae^{-rt}dt] = \Delta_t e^{-rt}[dS_t - rS_t dt - adt]$. So, to make the discounted portfolio value $e^{-rt}X_t$ a martingale, we are motivated to change the measure in such a way that $S_t - r\int_0^t S_u du - at$ is a martingale under the new measure. To do this, we note the SDE for S is $dS_t = \alpha_t S_t dt + \sigma S_t dW_t$. Hence $dS_t - rS_t dt - adt = [(\alpha_t - r)S_t - a]dt + \sigma S_t dW_t = \sigma S_t\left[\frac{(\alpha_t - r)S_t - a}{\sigma S_t}dt + dW_t\right]$. Set $\theta_t = \frac{(\alpha_t - r)S_t - a}{\sigma S_t}$ and $\widetilde{W}_t = \int_0^t \theta_s ds + W_t$, we can find an equivalent probability measure $\widetilde{P}$, under which S satisfies the SDE $dS_t = rS_t dt + \sigma S_t d\widetilde{W}_t + adt$ and $\widetilde{W}_t$ is a BM. This is the rational for formula (5.9.7).

This is a good place to pause and think about the meaning of "martingale measure." *What is to be a martingale?* The new measure $\widetilde{P}$ should be such that *the discounted value process of the replicating portfolio* is a martingale, *not the discounted price process of the underlying.* First, we want $D_t X_t$ to be a martingale under $\widetilde{P}$ because we suppose that X is able to replicate the derivative payoff at terminal time, $X_T = V_T$. In order to avoid arbitrage, we must have $X_t = V_t$ for any $t \in [0, T]$. The difficulty is how to calculate X_t and the magic is brought by the martingale measure in the following line of reasoning: $V_t = X_t = D_t^{-1}\widetilde{E}[D_T X_T|\mathcal{F}_t] = D_t^{-1}\widetilde{E}[D_T V_T|\mathcal{F}_t]$. You can think of martingale measure as a calculational convenience. That is *all* about martingale measure! *Risk neutral* is just a *perception*, referring to the actual effect of constructing a hedging portfolio! Second, we note when the portfolio is self-financing, the discounted price process of the underlying is a martingale under $\widetilde{P}$, as in the classical Black-Scholes-Merton model without dividends or cost of carry. This is not a coincidence. Indeed, we have in this case the relation $d(D_t X_t) = \Delta_t d(D_t S_t)$. So $D_t X_t$ being a martingale under $\widetilde{P}$ is more or less equivalent to $D_t S_t$ being a martingale under $\widetilde{P}$. However, when the underlying pays dividends, or there is cost of carry, $d(D_t X_t) = \Delta_t d(D_t S_t)$ no longer holds, as shown in formula (5.9.6). The portfolio is no longer *self-financing*, but *self-financing with consumption*. What we still want to retain is the martingale property of $D_t X_t$, not that of $D_t S_t$. This is how we choose martingale measure in the above paragraph.

Let V_T be a payoff at time T, then for the martingale $M_t = \widetilde{E}[e^{-rT}V_T|\mathcal{F}_t]$, by Martingale Representation Theorem, we can find an adapted process $\widetilde{\Gamma}_t$, so that $M_t = M_0 + \int_0^t \widetilde{\Gamma}_s d\widetilde{W}_s$. If we let $\Delta_t = \frac{\widetilde{\Gamma}_t e^{rt}}{\sigma S_t}$, then the value of the corresponding portfolio X satisfies $d(e^{-rt}X_t) = \widetilde{\Gamma}_t d\widetilde{W}_t$. So by setting $X_0 = M_0 = \widetilde{E}[e^{-rT}V_T]$, we must have $e^{-rt}X_t = M_t$, for all $t \in [0, T]$. In particular, $X_T = V_T$. Thus the portfolio perfectly hedges V_T. This justifies the risk-neutral pricing of European-type contingent claims in the model where cost of carry exists. Also note the risk-neutral measure is different from the one in case of no cost of carry.

Another perspective for perfect replication is the following. We need to solve the backward SDE

$$\begin{cases} dX_t = \Delta_t dS_t - a\Delta_t dt + r(X_t - \Delta_t S_t)dt \\ X_T = V_T \end{cases}$$

for two unknowns, X and Δ. To do so, we find a probability measure $\widetilde{P}$, under which $e^{-rt}X_t$ is a martingale, then $e^{-rt}X_t = \widetilde{E}[e^{-rT}V_T|\mathcal{F}_t] := M_t$. Martingale Representation Theorem gives $M_t = M_0 + \int_0^t \widetilde{\Gamma}_u d\widetilde{W}_u$ for some adapted process $\widetilde{\Gamma}$. This would give us a theoretical representation of Δ by comparison of integrands, hence a perfect replication of V_T.

(i)

Proof. As indicated in the above analysis, if we have (5.9.7) under $\widetilde{P}$, then $d(e^{-rt}X_t) = \Delta_t[d(e^{-rt}S_t) - ae^{-rt}dt] = \Delta_t e^{-rt}\sigma S_t d\widetilde{W}_t$. So $(e^{-rt}X_t)_{t\geq 0}$, where X is given by (5.9.6), is a $\widetilde{P}$-martingale. □

(ii)

Proof. By Itô's formula, $dY_t = Y_t[\sigma d\widetilde{W}_t + (r - \frac{1}{2}\sigma^2)dt] + \frac{1}{2}Y_t\sigma^2 dt = Y_t(\sigma d\widetilde{W}_t + rdt)$. So $d(e^{-rt}Y_t) = \sigma e^{-rt}Y_t d\widetilde{W}_t$ and $e^{-rt}Y_t$ is a $\widetilde{P}$-martingale. Moreover, if $S_t = S_0 Y_t + Y_t \int_0^t \frac{a}{Y_s}ds$, then

$$dS_t = S_0 dY_t + \int_0^t \frac{a}{Y_s}ds dY_t + adt = \left(S_0 + \int_0^t \frac{a}{Y_s}ds\right) Y_t(\sigma d\widetilde{W}_t + rdt) + adt = S_t(\sigma d\widetilde{W}_t + rdt) + adt.$$

This shows S satisfies (5.9.7). □

Remark 4. *To obtain this formula for S, we first set $U_t = e^{-rt}S_t$ to remove the $rS_t dt$ term. The SDE for U is $dU_t = \sigma U_t d\widetilde{W}_t + ae^{-rt}dt$. Just like solving linear ODE, to remove U in the $d\widetilde{W}_t$ term, we consider $V_t = U_t e^{-\sigma \widetilde{W}_t}$. Itô's product formula yields*

$$\begin{aligned}
dV_t &= e^{-\sigma\widetilde{W}_t}dU_t + U_t e^{-\sigma\widetilde{W}_t}\left((-\sigma)d\widetilde{W}_t + \frac{1}{2}\sigma^2 dt\right) + dU_t \cdot e^{-\sigma\widetilde{W}_t}\left((-\sigma)d\widetilde{W}_t + \frac{1}{2}\sigma^2 dt\right) \\
&= e^{-\sigma\widetilde{W}_t}ae^{-rt}dt - \frac{1}{2}\sigma^2 V_t dt.
\end{aligned}$$

Note V appears only in the dt term, so multiply the integration factor $e^{\frac{1}{2}\sigma^2 t}$ on both sides of the equation, we get

$$d(e^{\frac{1}{2}\sigma^2 t}V_t) = ae^{-rt-\sigma\widetilde{W}_t + \frac{1}{2}\sigma^2 t}dt.$$

Set $Y_t = e^{\sigma\widetilde{W}_t + (r-\frac{1}{2}\sigma^2)t}$, we have $d(S_t/Y_t) = adt/Y_t$. So $S_t = Y_t(S_0 + \int_0^t \frac{ads}{Y_s})$.

(iii)

Proof.

$$\begin{aligned}
\widetilde{E}[S_T|\mathcal{F}_t] &= S_0\widetilde{E}[Y_T|\mathcal{F}_t] + \widetilde{E}\left[Y_T\int_0^t \frac{a}{Y_s}ds + Y_T\int_t^T \frac{a}{Y_s}ds|\mathcal{F}_t\right] \\
&= S_0\widetilde{E}[Y_T|\mathcal{F}_t] + \int_0^t \frac{a}{Y_s}ds\widetilde{E}[Y_T|\mathcal{F}_t] + a\int_t^T \widetilde{E}\left[\frac{Y_T}{Y_s}|\mathcal{F}_t\right]ds \\
&= S_0 Y_t\widetilde{E}[Y_{T-t}] + \int_0^t \frac{a}{Y_s}dsY_t\widetilde{E}[Y_{T-t}] + a\int_t^T \widetilde{E}[Y_{T-s}]ds \\
&= \left(S_0 + \int_0^t \frac{a}{Y_s}ds\right)Y_t e^{r(T-t)} + a\int_t^T e^{r(T-s)}ds \\
&= \left(S_0 + \int_0^t \frac{ads}{Y_s}\right)Y_t e^{r(T-t)} - \frac{a}{r}(1 - e^{r(T-t)}).
\end{aligned}$$

In particular, $\widetilde{E}[S_T] = S_0 e^{rT} - \frac{a}{r}(1 - e^{rT})$. □

(iv)

Proof.

$$\begin{aligned}
d\widetilde{E}[S_T|\mathcal{F}_t] &= ae^{r(T-t)}dt + \left(S_0 + \int_0^t \frac{ads}{Y_s}\right)(e^{r(T-t)}dY_t - rY_t e^{r(T-t)}dt) + \frac{a}{r}e^{r(T-t)}(-r)dt \\
&= \left(S_0 + \int_0^t \frac{ads}{Y_s}\right) e^{r(T-t)}\sigma Y_t d\widetilde{W}_t.
\end{aligned}$$

So $\widetilde{E}[S_T|\mathcal{F}_t]$ is a $\widetilde{P}$-martingale. As we have argued at the beginning of the solution, risk-neutral pricing is valid even in the presence of cost of carry. So by an argument similar to that of §5.6.2, the process $\widetilde{E}[S_T|\mathcal{F}_t]$ is the futures price process for the commodity. ☐

(v)

Proof. We solve the equation $\widetilde{E}[e^{-r(T-t)}(S_T - K)|\mathcal{F}_t] = 0$ for K, and get $K = \widetilde{E}[S_T|\mathcal{F}_t]$. So $For_S(t,T) = Fut_S(t,T)$. ☐

(vi)

Proof. We follow the hint. First, we solve the SDE

$$\begin{cases} dX_t = dS_t - adt + r(X_t - S_t)dt \\ X_0 = 0. \end{cases}$$

By our analysis in part (i), $d(e^{-rt}X_t) = d(e^{-rt}S_t) - ae^{-rt}dt$. Integrate from 0 to t on both sides, we get $X_t = S_t - S_0e^{rt} + \frac{a}{r}(1 - e^{rt}) = S_t - S_0e^{rt} - \frac{a}{r}(e^{rt} - 1)$. In particular, $X_T = S_T - S_0e^{rT} - \frac{a}{r}(e^{rT} - 1)$. Meanwhile, $For_S(t,T) = Fut_s(t,T) = \widetilde{E}[S_T|\mathcal{F}_t] = \left(S_0 + \int_0^t \frac{ads}{Y_s}\right) Y_t e^{r(T-t)} - \frac{a}{r}(1 - e^{r(T-t)})$. So $For_S(0,T) = S_0e^{rT} - \frac{a}{r}(1 - e^{rT})$ and hence $X_T = S_T - For_S(0,T)$. After the agent delivers the commodity, whose value is S_T, and receives the forward price $For_S(0,T)$, the portfolio has exactly zero value. ☐

6 Connections with Partial Differential Equations

★ **Comments**:

1) A rigorous presentation of (strong) Markov property can be found in Revuz and Yor [11], Chapter III.

2) A rigorous presentation of the Feymann-Kac formula can be found in Øksendal [9, page 143] (also see 钱敏平等 [10, page 236]). To reconcile the version presented in this book and the one in Øksendal [9] (Theorem 8.2.1), we note for the undiscounted version, in this book

$$g(t, X(t)) = \mathbb{E}[h(X(T))|\mathcal{F}(t)] = \mathbb{E}^{X(t)}[h(X(T-t))]$$

while in Øksendal [9]

$$v(t,x) = E^x[h(X_t)].$$

So $v(t,x) = g(T-t,x)$. The discounted version can connected similarly.

3) *Hedging equation.* Recall the SDE satisfied by the value process $X(t)$ of a self-financing portfolio is (see formula (5.4.21) and (5.4.22) on page 230)

$$\begin{aligned}
d[D(t)X(t)] &= D(t)[dX(t) - R(t)X(t)dt] \\
&= D(t)\left\{\sum_{i=1}^m \Delta_i(t)dS_i(t) + R(t)\left[X(t) - \sum_{i=1}^m \Delta_i(t)S_i(t)\right]dt - R(t)X(t)dt\right\} \\
&= D(t)\left[\sum_{i=1}^m \Delta_i(t)dS_i(t) - R(t)\sum_{i=1}^m \Delta_i(t)S_i(t)dt\right] \\
&= \sum_{i=1}^m \Delta_i(t)d[D(t)S_i(t)].
\end{aligned}$$

Under the multidimensional market model (formula (5.4.6) on page 226)

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

and assuming the existence of the risk-neutral measure (that is, the *market price of risk equations* has a solution (formula (5.4.18) on page 228): $\alpha_i(t) - R(t) = \sum_{j=1}^{d}\sigma_{ij}(t)\Theta_j(t),\ i = 1, \cdots, m$), $D(t)S_i(t)$ is a martingale under the risk-neutral measure (formula (5.4.17) on page 228) satisfying the SDE

$$d[D(t)S_i(t)] = D(t)S_i(t)\sum_{j=1}^{d}\sigma_{ij}(t)d\widetilde{W}_j(t),\ dS_i(t) = R(t)S_i(t)dt + S_i(t)\sum_{j=1}^{d}\sigma_{ij}(t)d\widetilde{W}_j(t),\ i = 1, \cdots, m.$$

Consequently, $d[D(t)X(t)] = \sum_{i=1}^{m}\Delta_i(t)d[D(t)S_i(t)]$ becomes

$$d[D(t)X(t)] = D(t)\sum_{i=1}^{m}\left[\Delta_i(t)S_i(t)\sum_{j=1}^{d}\sigma_{ij}(t)d\widetilde{W}_j(t)\right].$$

If we further assume $X(t)$ has the form $v(t, S(t))$, then the above equation becomes

$$\begin{aligned}
& D(t)\left[v_t(t,S(t))dt + \sum_{i=1}^{m}v_{x_i}(t,S(t))dS_i(t) + \frac{1}{2}\sum_{k,l=1}^{m}v_{x_kx_l}(t,S(t))dS_k(t)dS_l(t)\right] - R(t)v(t,X(t))dt \\
= \ & (\cdots)dt + D(t)\sum_{i=1}^{m}v_{x_i}(t,S(t))\left[S_i(t)\sum_{j=1}^{d}\sigma_{ij}(t)d\widetilde{W}_j(t)\right] \\
= \ & D(t)\sum_{i=1}^{m}\left[\Delta_i(t)S_i(t)\sum_{j=1}^{d}\sigma_{ij}(t)d\widetilde{W}_j(t)\right].
\end{aligned}$$

Equating the coefficient of each $d\widetilde{W}_j(t)$, we have the *hedging equation*

$$\sum_{i=1}^{m}v_{x_i}(t,S(t))S_i(t)\sigma_{ij}(t) = \sum_{i=1}^{m}\Delta_i(t)S_i(t)\sigma_{ij}(t),\ j = 1, \cdots, d.$$

One solution of the hedging equation is

$$\Delta_i(t) = v_{x_i}(t, S(t)),\ i = 1, \cdots, m.$$

► **Exercise 6.1.** Consider the stochastic differential equation

$$dX(u) =$$

(i)

Proof. $Z_t = 1$ is obvious. Note the form of Z is similar to that of a geometric Brownian motion. So by Itô's formula, it is easy to obtain $dZ_u = b_uZ_udu + \sigma_uZ_udW_u$, $u \geq t$. □

(ii)

Proof. If $X_u = Y_u Z_u$ ($u \geq t$), then $X_t = Y_t Z_t = x \cdot 1 = x$ and

$$\begin{aligned}
dX_u &= Y_u dZ_u + Z_u dY_u + dY_u Z_u \\
&= Y_u(b_u Z_u du + \sigma_u Z_u dW_u) + Z_u\left(\frac{a_u - \sigma_u\gamma_u}{Z_u}du + \frac{\gamma_u}{Z_u}dW_u\right) + \sigma_u Z_u \frac{\gamma_u}{Z_u}du \\
&= [Y_u b_u Z_u + (a_u - \sigma_u\gamma_u) + \sigma_u\gamma_u]du + (\sigma_u Z_u Y_u + \gamma_u)dW_u \\
&= (b_u X_u + a_u)du + (\sigma_u X_u + \gamma_u)dW_u.
\end{aligned}$$

□

Remark 5. *To see how to find the above solution, we manipulate the equation (6.2.4) as follows. First, to remove the term $b_u X_u du$, we multiply on both sides of (6.2.4) the integrating factor $e^{-\int_t^u b_v dv}$. Then*

$$d(X_u e^{-\int_t^u b_v dv}) = e^{-\int_t^u b_v dv}(a_u du + (\gamma_u + \sigma_u X_u)dW_u).$$

Let $\bar{X}_u = e^{-\int_t^u b_v dv}X_u$, $\bar{a}_u = e^{-\int_t^u b_v dv}a_u$ and $\bar{\gamma}_u = e^{-\int_t^u b_v dv}\gamma_u$, then $\bar{X}$ satisfies the SDE

$$d\bar{X}_u = \bar{a}_u du + (\bar{\gamma}_u + \sigma_u \bar{X}_u)dW_u = (\bar{a}_u du + \bar{\gamma}_u dW_u) + \sigma_u \bar{X}_u dW_u.$$

To deal with the term $\sigma_u \bar{X}_u dW_u$, we consider $\hat{X}_u = \bar{X}_u e^{-\int_t^u \sigma_v dW_v}$. Then

$$\begin{aligned}
d\hat{X}_u &= e^{-\int_t^u \sigma_v dW_v}[(\bar{a}_u du + \bar{\gamma}_u dW_u) + \sigma_u \bar{X}_u dW_u] + \bar{X}_u\left(e^{-\int_t^u \sigma_v dW_v}(-\sigma_u)dW_u + \frac{1}{2}e^{-\int_t^u \sigma_v dW_v}\sigma_u^2 du\right) \\
&\quad +(\bar{\gamma}_u + \sigma_u \bar{X}_u)(-\sigma_u)e^{-\int_t^u \sigma_v dW_v}du \\
&= \hat{a}_u du + \hat{\gamma}_u dW_u + \sigma_u \hat{X}_u dW_u - \sigma_u \hat{X}_u dW_u + \frac{1}{2}\hat{X}_u \sigma_u^2 du - \sigma_u(\hat{\gamma}_u + \sigma_u \hat{X}_u)du \\
&= (\hat{a}_u - \sigma_u\hat{\gamma}_u - \frac{1}{2}\hat{X}_u\sigma_u^2)du + \hat{\gamma}_u dW_u,
\end{aligned}$$

where $\hat{a}_u = \bar{a}_u e^{-\int_t^u \sigma_v dW_v}$ and $\hat{\gamma}_u = \bar{\gamma}_u e^{-\int_t^u \sigma_v dW_v}$. Finally, use the integrating factor $e^{\int_t^u \frac{1}{2}\sigma_v^2 dv}$, we have

$$d\left(\hat{X}_u e^{\frac{1}{2}\int_t^u \sigma_v^2 dv}\right) = e^{\frac{1}{2}\int_t^u \sigma_v^2 dv}(d\hat{X}_u + \hat{X}_u \cdot \frac{1}{2}\sigma_u^2 du) = e^{\frac{1}{2}\int_t^u \sigma_v^2 dv}[(\hat{a}_u - \sigma_u\hat{\gamma}_u)du + \hat{\gamma}_u dW_u].$$

Write everything back into the original X, a and γ, we get

$$d\left(X_u e^{-\int_t^u b_v dv - \int_t^u \sigma_v dW_v + \frac{1}{2}\int_t^u \sigma_v^2 dv}\right) = e^{\frac{1}{2}\int_t^u \sigma_v^2 dv - \int_t^u \sigma_v dW_v - \int_t^u b_v dv}[(a_u - \sigma_u\gamma_u)du + \gamma_u dW_u],$$

i.e.

$$d\left(\frac{X_u}{Z_u}\right) = \frac{1}{Z_u}[(a_u - \sigma_u\gamma_u)du + \gamma_u dW_u] = dY_u.$$

This inspired us to try $X_u = Y_u Z_u$.

6.2. (i)

Proof. The portfolio is self-financing, so for any $t \leq T_1$, we have

$$dX_t = \Delta_1(t)df(t, R_t, T_1) + \Delta_2(t)df(t, R_t, T_2) + R_t(X_t - \Delta_1(t)f(t, R_t, T_1) - \Delta_2(t)f(t, R_t, T_2))dt,$$

and

$$
\begin{aligned}
& d(D_tX_t) \\
= & -R_tD_tX_tdt + D_tdX_t \\
= & D_t[\Delta_1(t)df(t,R_t,T_1) + \Delta_2(t)df(t,R_t,T_2) - R_t(\Delta_1(t)f(t,R_t,T_1) + \Delta_2(t)f(t,R_t,T_2))dt] \\
= & D_t[\Delta_1(t)\left(f_t(t,R_t,T_1)dt + f_r(t,R_t,T_1)dR_t + \frac{1}{2}f_{rr}(t,R_t,T_1)\gamma^2(t,R_t)dt\right) \\
& +\Delta_2(t)\left(f_t(t,R_t,T_2)dt + f_r(t,R_t,T_2)dR_t + \frac{1}{2}f_{rr}(t,R_t,T_2)\gamma^2(t,R_t)dt\right) \\
& -R_t(\Delta_1(t)f(t,R_t,T_1) + \Delta_2(t)f(t,R_t,T_2))dt] \\
= & \Delta_1(t)D_t[-R_tf(t,R_t,T_1) + f_t(t,R_t,T_1) + \alpha(t,R_t)f_r(t,R_t,T_1) + \frac{1}{2}\gamma^2(t,R_t)f_{rr}(t,R_t,T_1)]dt \\
& +\Delta_2(t)D_t[-R_tf(t,R_t,T_2) + f_t(t,R_t,T_2) + \alpha(t,R_t)f_r(t,R_t,T_2) + \frac{1}{2}\gamma^2(t,R_t)f_{rr}(t,R_t,T_2)]dt \\
& +D_t\gamma(t,R_t)[D_t\gamma(t,R_t)[\Delta_1(t)f_r(t,R_t,T_1) + \Delta_2(t)f_r(t,R_t,T_2)]]dW_t \\
= & \Delta_1(t)D_t[\alpha(t,R_t) - \beta(t,R_t,T_1)]f_r(t,R_t,T_1)dt + \Delta_2(t)D_t[\alpha(t,R_t) - \beta(t,R_t,T_2)]f_r(t,R_t,T_2)dt \\
& +D_t\gamma(t,R_t)[\Delta_1(t)f_r(t,R_t,T_1) + \Delta_2(t)f_r(t,R_t,T_2)]dW_t.
\end{aligned}
$$

□

(ii)

Proof. Let $\Delta_1(t) = S_tf_r(t,R_t,T_2)$ and $\Delta_2(t) = -S_tf_r(t,R_t,T_1)$, then

$$
\begin{aligned}
d(D_tX_t) &= D_tS_t[\beta(t,R_t,T_2) - \beta(t,R_t,T_1)]f_r(t,R_t,T_1)f_r(t,R_t,T_2)dt \\
&= D_t|[\beta(t,R_t,T_1) - \beta(t,R_t,T_2)]f_r(t,R_t,T_1)f_r(t,R_t,T_2)|dt.
\end{aligned}
$$

Integrate from 0 to T on both sides of the above equation, we get

$$
D_TX_T - D_0X_0 = \int_0^T D_t|[\beta(t,R_t,T_1) - \beta(t,R_t,T_2)]f_r(t,R_t,T_1)f_r(t,R_t,T_2)|dt.
$$

If $\beta(t,R_t,T_1) \neq \beta(t,R_t,T_2)$ for some $t \in [0,T]$, under the assumption that $f_r(t,r,T) \neq 0$ for all values of r and $0 \leq t \leq T$, $D_TX_T - D_0X_0 > 0$. To avoid arbitrage (see, for example, Exercise 5.7), we must have for a.s. ω, $\beta(t,R_t,T_1) = \beta(t,R_t,T_2)$, $\forall t \in [0,T]$. This implies $\beta(t,r,T)$ does not depend on T. □

(iii)

Proof. In (6.9.4), let $\Delta_1(t) = \Delta(t)$, $T_1 = T$ and $\Delta_2(t) = 0$, we get

$$
\begin{aligned}
d(D_tX_t) &= \Delta(t)D_t\left[-R_tf(t,R_t,T) + f_t(t,R_t,T) + \alpha(t,R_t)f_r(t,R_t,T) + \frac{1}{2}\gamma^2(t,R_t)f_{rr}(t,R_t,T)\right]dt \\
&\quad +D_t\gamma(t,R_t)\Delta(t)f_r(t,R_t,T)dW_t.
\end{aligned}
$$

This is formula (6.9.5).

If $f_r(t,r,T) = 0$, then $d(D_tX_t) = \Delta(t)D_t\left[-R_tf(t,R_t,T) + f_t(t,R_t,T) + \frac{1}{2}\gamma^2(t,R_t)f_{rr}(t,R_t,T)\right]dt$. We choose $\Delta(t) = \text{sign}\left\{\left[-R_tf(t,R_t,T) + f_t(t,R_t,T) + \frac{1}{2}\gamma^2(t,R_t)f_{rr}(t,R_t,T)\right]\right\}$. To avoid arbitrage in this case, we must have $f_t(t,R_t,T) + \frac{1}{2}\gamma^2(t,R_t)f_{rr}(t,R_t,T) = R_tf(t,R_t,T)$, or equivalently, for any r in the range of R_t, $f_t(t,r,T) + \frac{1}{2}\gamma^2(t,r)f_{rr}(t,r,T) = rf(t,r,T)$. □

6.3.

Proof. We note

$$\frac{d}{ds}\left[e^{-\int_0^s b_v dv}C(s,T)\right] = e^{-\int_0^s b_v dv}[C(s,T)(-b_s) + b_s C(s,T) - 1] = -e^{-\int_0^s b_v dv}.$$

So integrate on both sides of the equation from t to T, we obtain

$$e^{-\int_0^T b_v dv}C(T,T) - e^{-\int_0^t b_v dv}C(t,T) = -\int_t^T e^{-\int_0^s b_v dv}ds.$$

Since $C(T,T) = 0$, we have $C(t,T) = e^{\int_0^t b_v dv}\int_t^T e^{-\int_0^s b_v dv}ds = \int_t^T e^{\int_s^t b_v dv}ds$. Finally, by $A'(s,T) = -a(s)C(s,T) + \frac{1}{2}\sigma^2(s)C^2(s,T)$, we get

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

Since $A(T,T) = 0$, we have $A(t,T) = \int_t^T (a(s)C(s,T) - \frac{1}{2}\sigma^2(s)C^2(s,T))ds$. □

6.4. (i)

Proof. By the definition of φ, we have

$$\varphi'(t) = e^{\frac{1}{2}\sigma^2\int_t^T C(u,T)du}\frac{1}{2}\sigma^2(-1)C(t,T) = -\frac{1}{2}\varphi(t)\sigma^2 C(t,T).$$

So $C(t,T) = -\frac{2\varphi'(t)}{\phi(t)\sigma^2}$. Differentiate both sides of the equation $\varphi'(t) = -\frac{1}{2}\varphi(t)\sigma^2 C(t,T)$, we get

$$\begin{aligned}\varphi''(t) &= -\frac{1}{2}\sigma^2[\varphi'(t)C(t,T) + \varphi(t)C'(t,T)] \\ &= -\frac{1}{2}\sigma^2[-\frac{1}{2}\varphi(t)\sigma^2 C^2(t,T) + \varphi(t)C'(t,T)] \\ &= \frac{1}{4}\sigma^4\varphi(t)C^2(t,T) - \frac{1}{2}\sigma^2\varphi(t)C'(t,T).\end{aligned}$$

So $C'(t,T) = \left[\frac{1}{4}\sigma^4\varphi(t)C^2(t,T) - \varphi''(t)\right]/\frac{1}{2}\varphi(t)\sigma^2 = \frac{1}{2}\sigma^2 C^2(t,T) - \frac{2\varphi''(t)}{\sigma^2\varphi(t)}$. □

(ii)

Proof. Plug formulas (6.9.8) and (6.9.9) into (6.5.14), we get

$$-\frac{2\varphi''(t)}{\sigma^2\varphi(t)} + \frac{1}{2}\sigma^2 C^2(t,T) = b(-1)\frac{2\varphi'(t)}{\sigma^2\varphi(t)} + \frac{1}{2}\sigma^2 C^2(t,T) - 1,$$

i.e. $\varphi''(t) - b\varphi'(t) - \frac{1}{2}\sigma^2\varphi(t) = 0$. □

(iii)

Proof. The characteristic equation of $\varphi''(t) - b\varphi'(t) - \frac{1}{2}\sigma^2\varphi(t) = 0$ is $\lambda^2 - b\lambda - \frac{1}{2}\sigma^2 = 0$, which gives two roots $\frac{1}{2}(b \pm \sqrt{b^2 + 2\sigma^2}) = \frac{1}{2}b \pm \gamma$ with $\gamma = \frac{1}{2}\sqrt{b^2 + 2\sigma^2}$. Therefore by standard theory of ordinary differential equations, a general solution of φ is $\varphi(t) = e^{\frac{1}{2}bt}(a_1 e^{\gamma t} + a_2 e^{-\gamma t})$ for some constants a_1 and a_2. It is then easy to see that we can choose appropriate constants c_1 and c_2 so that

$$\varphi(t) = \frac{c_1}{\frac{1}{2}b + \gamma}e^{-(\frac{1}{2}b+\gamma)(T-t)} - \frac{c_2}{\frac{1}{2}b - \gamma}e^{-(\frac{1}{2}b-\gamma)(T-t)}.$$

□

(iv)

Proof. From part (iii), it is easy to see $\varphi'(t) = c_1 e^{-(\frac{1}{2}b+\gamma)(T-t)} - c_2 e^{-(\frac{1}{2}b-\gamma)(T-t)}$. In particular,

$$0 = C(T,T) = -\frac{2\varphi'(T)}{\sigma^2\varphi(T)} = -\frac{2(c_1-c_2)}{\sigma^2\varphi(T)}.$$

So $c_1 = c_2$. $\square$

(v)

Proof. We first recall the definitions and properties of sinh and cosh:

$$\sinh z = \frac{e^z - e^{-z}}{2},\ \cosh z = \frac{e^z + e^{-z}}{2},\ (\sinh z)' = \cosh z,\ \text{and}\ (\cosh z)' = \sinh z.$$

Therefore

$$\begin{aligned}
\varphi(t) &= c_1 e^{-\frac{1}{2}b(T-t)}\left[\frac{e^{-\gamma(T-t)}}{\frac{1}{2}b+\gamma} - \frac{e^{\gamma(T-t)}}{\frac{1}{2}b-\gamma}\right] \\
&= c_1 e^{-\frac{1}{2}b(T-t)}\left[\frac{\frac{1}{2}b-\gamma}{\frac{1}{4}b^2-\gamma^2}e^{-\gamma(T-t)} - \frac{\frac{1}{2}b+\gamma}{\frac{1}{4}b^2-\gamma^2}e^{\gamma(T-t)}\right] \\
&= \frac{2c_1}{\sigma^2}e^{-\frac{1}{2}b(T-t)}\left[-(\frac{1}{2}b-\gamma)e^{-\gamma(T-t)} + (\frac{1}{2}b+\gamma)e^{\gamma(T-t)}\right] \\
&= \frac{2c_1}{\sigma^2}e^{-\frac{1}{2}b(T-t)}[b\sinh(\gamma(T-t)) + 2\gamma\cosh(\gamma(T-t))].
\end{aligned}$$

and

$$\begin{aligned}
\varphi'(t) &= \frac{1}{2}b\cdot\frac{2c_1}{\sigma^2}e^{-\frac{1}{2}b(T-t)}[b\sinh(\gamma(T-t)) + 2\gamma\cosh(\gamma(T-t))] \\
&\quad + \frac{2c_1}{\sigma^2}e^{-\frac{1}{2}b(T-t)}[-\gamma b\cosh(\gamma(T-t)) - 2\gamma^2\sinh(\gamma(T-t))] \\
&= 2c_1 e^{-\frac{1}{2}b(T-t)}\left[\frac{b^2}{2\sigma^2}\sinh(\gamma(T-t)) + \frac{b\gamma}{\sigma^2}\cosh(\gamma(T-t)) - \frac{b\gamma}{\sigma^2}\cosh(\gamma(T-t))\right. \\
&\quad \left. -\frac{2\gamma^2}{\sigma^2}\sinh(\gamma(T-t))\right] \\
&= 2c_1 e^{-\frac{1}{2}b(T-t)}\frac{b^2-4\gamma^2}{2\sigma^2}\sinh(\gamma(T-t)) \\
&= -2c_1 e^{-\frac{1}{2}b(T-t)}\sinh(\gamma(T-t)).
\end{aligned}$$

This implies

$$C(t,T) = -\frac{2\varphi'(t)}{\sigma^2\varphi(t)} = \frac{\sinh(\gamma(T-t))}{\gamma\cosh(\gamma(T-t)) + \frac{1}{2}b\sinh(\gamma(T-t))}.$$

$\square$

(vi)

Proof. By (6.5.15) and (6.9.8), $A'(t,T) = \frac{2a\varphi'(t)}{\sigma^2\varphi(t)}$. Hence

$$A(T,T) - A(t,T) = \int_t^T \frac{2a\varphi'(s)}{\sigma^2\varphi(s)}ds = \frac{2a}{\sigma^2}\ln\frac{\varphi(T)}{\varphi(t)},$$

and

$$A(t,T) = -\frac{2a}{\sigma^2}\ln\frac{\varphi(T)}{\varphi(t)} = -\frac{2a}{\sigma^2}\ln\left[\frac{\gamma e^{\frac{1}{2}b(T-t)}}{\gamma\cosh(\gamma(T-t)) + \frac{1}{2}b\sinh(\gamma(T-t))}\right].$$

$\square$

6.5. (i)

Proof. Since $g(t, X_1(t), X_2(t)) = E[h(X_1(T), X_2(T))|\mathcal{F}_t]$ and

$$e^{-rt}f(t, X_1(t), X_2(t)) = E[e^{-rT}h(X_1(T), X_2(T))|\mathcal{F}_t],$$

iterated conditioning argument shows $g(t, X_1(t), X_2(t))$ and $e^{-rt}f(t, X_1(t), X_2(t))$ ar both martingales. □

(ii) and (iii)

Proof. We note

$$\begin{aligned}
& dg(t, X_1(t), X_2(t)) \\
= \;& g_t dt + g_{x_1} dX_1(t) + g_{x_2} dX_2(t) + \frac{1}{2} g_{x_1x_2} dX_1(t)dX_1(t) + \frac{1}{2} g_{x_2x_2} dX_2(t)dX_2(t) + g_{x_1x_2} dX_1(t)dX_2(t) \\
= \;& \Big[g_t + g_{x_1}\beta_1 + g_{x_2}\beta_2 + \frac{1}{2} g_{x_1x_1}(\gamma_{11}^2 + \gamma_{12}^2 + 2\rho\gamma_{11}\gamma_{12}) + g_{x_1x_2}(\gamma_{11}\gamma_{21} + \rho\gamma_{11}\gamma_{22} + \rho\gamma_{12}\gamma_{21} + \gamma_{12}\gamma_{22}) \\
& + \frac{1}{2} g_{x_2x_2}(\gamma_{21}^2 + \gamma_{22}^2 + 2\rho\gamma_{21}\gamma_{22}) \Big] dt + \text{martingale part.}
\end{aligned}$$

So we must have

$$\begin{aligned}
& g_t + g_{x_1}\beta_1 + g_{x_2}\beta_2 + \frac{1}{2} g_{x_1x_1}(\gamma_{11}^2 + \gamma_{12}^2 + 2\rho\gamma_{11}\gamma_{12}) + g_{x_1x_2}(\gamma_{11}\gamma_{21} + \rho\gamma_{11}\gamma_{22} + \rho\gamma_{12}\gamma_{21} + \gamma_{12}\gamma_{22}) \\
& + \frac{1}{2} g_{x_2x_2}(\gamma_{21}^2 + \gamma_{22}^2 + 2\rho\gamma_{21}\gamma_{22}) = 0.
\end{aligned}$$

Taking $\rho = 0$ will give part (ii) as a special case. The PDE for f can be similarly obtained. □

6.6. (i)

Proof. Multiply $e^{\frac{1}{2}bt}$ on both sides of (6.9.15), we get

$$d(e^{\frac{1}{2}bt}X_j(t)) = e^{\frac{1}{2}bt}\left(X_j(t)\frac{1}{2}bdt + (-\frac{b}{2}X_j(t)dt + \frac{1}{2}\sigma dW_j(t)\right) = e^{\frac{1}{2}bt}\frac{1}{2}\sigma dW_j(t).$$

So $e^{\frac{1}{2}bt}X_j(t) - X_j(0) = \frac{1}{2}\sigma\int_0^t e^{\frac{1}{2}bu}dW_j(u)$ and $X_j(t) = e^{-\frac{1}{2}bt}\left(X_j(0) + \frac{1}{2}\sigma\int_0^t e^{\frac{1}{2}bu}dW_j(u)\right)$. By Theorem 4.4.9, $X_j(t)$ is normally distributed with mean $X_j(0)e^{-\frac{1}{2}bt}$ and variance $\frac{e^{-bt}}{4}\sigma^2\int_0^t e^{bu}du = \frac{\sigma^2}{4b}(1 - e^{-bt})$. □

(ii)

Proof. Suppose $R(t) = \sum_{j=1}^d X_j^2(t)$, then

$$\begin{aligned}
dR(t) &= \sum_{j=1}^d (2X_j(t)dX_j(t) + dX_j(t)dX_j(t)) \\
&= \sum_{j=1}^d \left(2X_j(t)dX_j(t) + \frac{1}{4}\sigma^2 dt\right) \\
&= \sum_{j=1}^d \left(-bX_j^2(t)dt + \sigma X_j(t)dW_j(t) + \frac{1}{4}\sigma^2 dt\right) \\
&= \left(\frac{d}{4}\sigma^2 - bR(t)\right)dt + \sigma\sqrt{R(t)}\sum_{j=1}^d \frac{X_j(t)}{\sqrt{R(t)}}dW_j(t).
\end{aligned}$$

Let $B(t) = \sum_{j=1}^d \int_0^t \frac{X_j(s)}{\sqrt{R(s)}}dW_j(s)$, then B is a local martingale with $dB(t)dB(t) = \sum_{j=1}^d \frac{X_j^2(t)}{R(t)}dt = dt$. So by Lévy's Theorem, B is a Brownian motion. Therefore $dR(t) = (a - bR(t))dt + \sigma\sqrt{R(t)}dB(t)$ ($a := \frac{d}{4}\sigma^2$) and R is a CIR interest rate process. □

(iii)

Proof. By (6.9.16), $X_j(t)$ is dependent on W_j only and is normally distributed with mean $e^{-\frac{1}{2}bt}X_j(0)$ and variance $\frac{\sigma^2}{4b}[1-e^{-bt}]$. So $X_1(t)$, $\cdots$, $X_d(t)$ are i.i.d. normal with the same mean $\mu(t)$ and variance $v(t)$. □

(iv)

Proof.

$$\begin{aligned}
E\left[e^{uX_j^2(t)}\right] &= \int_{-\infty}^{\infty} e^{ux^2}\frac{e^{-\frac{(x-\mu(t))^2}{2v(t)}}dx}{\sqrt{2\pi v(t)}} \\
&= \int_{-\infty}^{\infty} \frac{e^{-\frac{(1-2uv(t))x^2-2\mu(t)x+\mu^2(t)}{2v(t)}}}{\sqrt{2\pi v(t)}}dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi v(t)}}e^{-\frac{\left(x-\frac{\mu(t)}{1-2uv(t)}\right)^2+\frac{\mu^2(t)}{1-2uv(t)}-\frac{\mu^2(t)}{(1-2uv(t))^2}}{2v(t)/(1-2uv(t))}}dx \\
&= \int_{-\infty}^{\infty} \frac{\sqrt{1-2uv(t)}}{\sqrt{2\pi v(t)}}e^{-\frac{\left(x-\frac{\mu(t)}{1-2uv(t)}\right)^2}{2v(t)/(1-2uv(t))}}dx\cdot\frac{e^{-\frac{\mu^2(t)(1-2uv(t))-\mu^2(t)}{2v(t)(1-2uv(t))}}}{\sqrt{1-2uv(t)}} \\
&= \frac{e^{-\frac{u\mu^2(t)}{1-2uv(t)}}}{\sqrt{1-2uv(t)}}.
\end{aligned}$$

□

(v)

Proof. By $R(t)=\sum_{j=1}^d X_j^2(t)$ and the fact $X_1(t)$, $\cdots$, $X_d(t)$ are i.i.d.,

$$E[e^{uR(t)}] = (E[e^{uX_1^2(t)}])^d = (1-2uv(t))^{-\frac{d}{2}}e^{\frac{ud\mu^2(t)}{1-2uv(t)}} = (1-2uv(t))^{-\frac{2a}{\sigma^2}}e^{-\frac{e^{-bt}uR(0)}{1-2uv(t)}}.$$

□

6.7. (i)

Proof. $e^{-rt}c(t,S_t,V_t)=\widetilde{E}[e^{-rT}(S_T-K)^+|\mathcal{F}_t]$ is a martingale by iterated conditioning argument. Since

$$\begin{aligned}
& d(e^{-rt}c(t,S_t,V_t)) \\
= \;& e^{-rt}\Bigg[c(t,S_t,V_t)(-r)+c_t(t,S_t,V_t)+c_s(t,S_t,V_t)rS_t+c_v(t,S_t,V_t)(a-bV_t)+\frac{1}{2}c_{ss}(t,S_t,V_t)V_tS_t^2+ \\
& \frac{1}{2}c_{vv}(t,S_t,V_t)\sigma^2V_t+c_{sv}(t,S_t,V_t)\sigma V_tS_t\rho\Bigg]dt+\text{martingale part},
\end{aligned}$$

we conclude $rc=c_t+rsc_s+c_v(a-bv)+\frac{1}{2}c_{ss}vs^2+\frac{1}{2}c_{vv}\sigma^2v+c_{sv}\sigma sv\rho$. This is equation (6.9.26). □

(ii)

Proof. Suppose $c(t,s,v) = sf(t,\log s,v) - e^{-r(T-t)}Kg(t,\log s,v)$, then

$$\begin{aligned}
c_t &= sf_t(t,\log s,v) - re^{-r(T-t)}Kg(t,\log s,v) - e^{-r(T-t)}Kg_t(t,\log s,v),\\
c_s &= f(t,\log s,v) + sf_s(t,\log s,v)\frac{1}{s} - e^{-r(T-t)}Kg_s(t,\log s,v)\frac{1}{s},\\
c_v &= sf_v(t,\log s,v) - e^{-r(T-t)}Kg_v(t,\log s,v),\\
c_{ss} &= f_s(t,\log s,v)\frac{1}{s} + f_{ss}(t,\log s,v)\frac{1}{s} - e^{-r(T-t)}Kg_{ss}(t,\log s,v)\frac{1}{s^2} + e^{-r(T-t)}Kg_s(t,\log s,v)\frac{1}{s^2},\\
c_{sv} &= f_v(t,\log s,v) + f_{sv}(t,\log s,v) - e^{-r(T-t)}\frac{K}{s}g_{sv}(t,\log s,v),\\
c_{vv} &= sf_{vv}(t,\log s,v) - e^{-r(T-t)}Kg_{vv}(t,\log s,v).
\end{aligned}$$

So

$$\begin{aligned}
& c_t + rsc_s + (a-bv)c_v + \frac{1}{2}s^2vc_{ss} + \rho\sigma svc_{sv} + \frac{1}{2}\sigma^2vc_{vv}\\
= \; & sf_t - re^{-r(T-t)}Kg - e^{-r(T-t)}Kg_t + rsf + rsf_s - rKe^{-r(T-t)}g_s + (a-bv)(sf_v - e^{-r(T-t)}Kg_v)\\
& +\frac{1}{2}s^2v\left[-\frac{1}{s}f_s + \frac{1}{s}f_{ss} - e^{-r(T-t)}\frac{K}{s^2}g_{ss} + e^{-r(T-t)}K\frac{g_s}{s^2}\right] + \rho\sigma sv\left(f_v + f_{sv} - e^{-r(T-t)}\frac{K}{s}g_{sv}\right)\\
& +\frac{1}{2}\sigma^2v(sf_{vv} - e^{-r(T-t)}Kg_{vv})\\
= \; & s\left[f_t + (r+\frac{1}{2}v)f_s + (a-bv+\rho\sigma v)f_v + \frac{1}{2}vf_{ss} + \rho\sigma vf_{sv} + \frac{1}{2}\sigma^2vf_{vv}\right] - Ke^{-r(T-t)}\left[g_t + (r-\frac{1}{2}v)g_s\right.\\
& \left. +(a-bv)g_v + \frac{1}{2}vg_{ss} + \rho\sigma vg_{sv} + \frac{1}{2}\sigma^2vg_{vv}\right] + rsf - re^{-r(T-t)}Kg\\
= \; & rc.
\end{aligned}$$

That is, c satisfies the PDE (6.9.26). □

(iii)

Proof. First, by Markov property, $f(t,X_t,V_t) = E[1_{\{X_T\geq\log K\}}|\mathcal{F}_t]$. So $f(T,X_t,V_t) = 1_{\{X_T\geq\log K\}}$, which implies $f(T,x,v) = 1_{\{x\geq\log K\}}$ for all $x\in\mathbb{R}$, $v\geq 0$. Second, $f(t,X_t,V_t)$ is a martingale, so by differentiating f and setting the dt term as zero, we have the PDE (6.9.32) for f. Indeed,

$$\begin{aligned}
df(t,X_t,V_t) \;=\; & \Big[f_t(t,X_t,V_t) + f_x(t,X_t,V_t)(r+\frac{1}{2}V_t) + f_v(t,X_t,V_t)(a-bv_t+\rho\sigma V_t) + \frac{1}{2}f_{xx}(t,X_t,V_t)V_t\\
& +\frac{1}{2}f_{vv}(t,X_t,V_t)\sigma^2V_t + f_{xv}(t,X_t,V_t)\sigma V_t\rho\Big]dt + \text{martingale part.}
\end{aligned}$$

So we must have $f_t + (r+\frac{1}{2}v)f_x + (a-bv+\rho\sigma v)f_v + \frac{1}{2}f_{xx}v + \frac{1}{2}f_{vv}\sigma^2v + \sigma v\rho f_{xv} = 0$. This is (6.9.32). □

(iv)

Proof. Similar to (iii). □

(v)

Proof. $c(T,s,v) = sf(T,\log s,v) - e^{-r(T-t)}Kg(T,\log s,v) = s1_{\{\log s\geq\log K\}} - K1_{\{\log s\geq\log K\}} = 1_{\{s\geq K\}}(s-K) = (s-K)^+$. □

6.8.

Proof. We follow the hint. Suppose h is smooth and compactly supported, then it is legitimate to exchange integration and differentiation:

$$
\begin{aligned}
g_t(t,x) &= \frac{\partial}{\partial t}\int_0^\infty h(y)p(t,T,x,y)dy = \int_0^\infty h(y)p_t(t,T,x,y)dy,\\
g_x(t,x) &= \int_0^\infty h(y)p_x(t,T,x,y)dy,\\
g_{xx}(t,x) &= \int_0^\infty h(y)p_{xx}(t,T,x,y)dy.
\end{aligned}
$$

So (6.9.45) implies $\int_0^\infty h(y)\left[p_t(t,T,x,y)+\beta(t,x)p_x(t,T,x,y)+\frac{1}{2}\gamma^2(t,x)p_{xx}(t,T,x,y)\right]dy = 0$. By the arbitrariness of h and assuming β, p_t, p_x, v, p_{xx} are all continuous, we have

$$
p_t(t,T,x,y)+\beta(t,x)p_x(t,T,x,y)+\frac{1}{2}\gamma^2(t,x)p_{xx}(t,T,x,y)=0.
$$

This is (6.9.43). □

6.9.

Proof. We first note

$$
\begin{aligned}
dh_b(X_u) &= h_b'(X_u)dX_u + \frac{1}{2}h_b''(X_u)dX_u dX_u\\
&= \left[h_b'(X_u)\beta(u,X_u)+\frac{1}{2}\gamma^2(u,X_u)h_b''(X_u)\right]du + h_b'(X_u)\gamma(u,X_u)dW_u.
\end{aligned}
$$

Integrate on both sides of the equation, we have

$$
h_b(X_T)-h_b(X_t) = \int_t^T \left[h_b'(X_u)\beta(u,X_u)+\frac{1}{2}\gamma^2(u,X_u)h_b''(X_u)\right]du + \text{martingale part}.
$$

Take expectation on both sides, we get

$$
\begin{aligned}
E^{t,x}[h_b(X_T)-h_b(X_t)] &= \int_{-\infty}^{\infty} h_b(y)p(t,T,x,y)dy - h(x)\\
&= \int_t^T E^{t,x}[h_b'(X_u)\beta(u,X_u)+\frac{1}{2}\gamma^2(u,X_u)h_b''(X_u)]du\\
&= \int_t^T\int_{-\infty}^{\infty}\left[h_b'(y)\beta(u,y)+\frac{1}{2}\gamma^2(u,y)h_b''(y)\right]p(t,u,x,y)dydu.
\end{aligned}
$$

Since h_b vanishes outside $(0,b)$, the integration range can be changed from $(-\infty,\infty)$ to $(0,b)$, which gives (6.9.48).

By integration-by-parts formula, we have

$$
\begin{aligned}
\int_0^b \beta(u,y)p(t,u,x,y)h_b'(y)dy &= h_b(y)\beta(u,y)p(t,u,x,y)|_0^b - \int_0^b h_b(y)\frac{\partial}{\partial y}(\beta(u,y)p(t,u,x,y))dy\\
&= -\int_0^b h_b(y)\frac{\partial}{\partial y}(\beta(u,y)p(t,u,x,y))dy,
\end{aligned}
$$

and

$$
\begin{aligned}
\int_0^b \gamma^2(u,y)p(t,u,x,y)h_b''(y)dy &= -\int_0^b \frac{\partial}{\partial y}(\gamma^2(u,y)p(t,u,x,y))h_b'(y)dy\\
&= \int_0^b \frac{\partial^2}{\partial y}(\gamma^2(u,y)p(t,u,x,y))h_b(y)dy.
\end{aligned}
$$

Plug these formulas into (6.9.48), we get (6.9.49).

Differentiate w.r.t. T on both sides of (6.9.49), we have

$$\int_0^b h_b(y)\frac{\partial}{\partial T}p(t,T,x,y)dy = -\int_0^b \frac{\partial}{\partial y}[\beta(T,y)p(t,T,x,y)]h_b(y)dy + \frac{1}{2}\int_0^b \frac{\partial^2}{\partial y^2}[\gamma^2(T,y)p(t,T,x,y)]h_b(y)dy,$$

that is,

$$\int_0^b h_b(y)\left[\frac{\partial}{\partial T}p(t,T,x,y) + \frac{\partial}{\partial y}(\beta(T,y)p(t,T,x,y)) - \frac{1}{2}\frac{\partial^2}{\partial y^2}(\gamma^2(T,y)p(t,T,x,y))\right]dy = 0.$$

This is (6.9.50).

By (6.9.50) and the arbitrariness of h_b, we conclude for any $y \in (0,\infty)$,

$$\frac{\partial}{\partial T}p(t,T,x,y) + \frac{\partial}{\partial y}(\beta(T,y)p(t,T,x,y)) - \frac{1}{2}\frac{\partial^2}{\partial y^2}(\gamma^2(T,y)p(t,T,x,y)) = 0.$$

□

6.10.

Proof. Under the assumption that $\lim_{y\to\infty}(y-K)ry\widetilde{p}(0,T,x,y) = 0$, we have

$$\begin{aligned}
-\int_K^\infty (y-K)\frac{\partial}{\partial y}(ry\widetilde{p}(0,T,x,y))dy &= -(y-K)ry\widetilde{p}(0,T,x,y)|_K^\infty + \int_K^\infty ry\widetilde{p}(0,T,x,y)dy \\
&= \int_K^\infty ry\widetilde{p}(0,T,x,y)dy.
\end{aligned}$$

If we further assume (6.9.57) and (6.9.58), then use integration-by-parts formula twice, we have

$$\begin{aligned}
&\frac{1}{2}\int_K^\infty (y-K)\frac{\partial^2}{\partial y^2}(\sigma^2(T,y)y^2\widetilde{p}(0,T,x,y))dy \\
= &\frac{1}{2}\left[(y-K)\frac{\partial}{\partial y}(\sigma^2(T,y)y^2\widetilde{p}(0,T,x,y))|_K^\infty - \int_K^\infty \frac{\partial}{\partial y}(\sigma^2(T,y)y^2\widetilde{p}(0,T,x,y))dy\right] \\
= &-\frac{1}{2}(\sigma^2(T,y)y^2\widetilde{p}(0,T,x,y)|_K^\infty) \\
= &\frac{1}{2}\sigma^2(T,K)K^2\widetilde{p}(0,T,x,K).
\end{aligned}$$

Therefore,

$$\begin{aligned}
c_T(0,T,x,K) &= -rc(0,T,x,K) + e^{-rT}\int_K^\infty (y-K)\widetilde{p}_T(0,T,x,y)dy \\
&= -re^{-rT}\int_K^\infty (y-K)\widetilde{p}(0,T,x,y)dy + e^{-rT}\int_K^\infty (y-K)\widetilde{p}_T(0,T,x,y)dy \\
&= -re^{-rT}\int_K^\infty (y-K)\widetilde{p}(0,T,x,y)dy - e^{-rT}\int_K^\infty (y-K)\frac{\partial}{\partial y}(ry\widetilde{p}(t,T,x,y))dy \\
&\quad +e^{-rT}\int_K^\infty (y-K)\frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(T,y)y^2\widetilde{p}(t,T,x,y))dy \\
&= -re^{-rT}\int_K^\infty (y-K)\widetilde{p}(0,T,x,y)dy + e^{-rT}\int_K^\infty ry\widetilde{p}(0,T,x,y)dy \\
&\quad +e^{-rT}\frac{1}{2}\sigma^2(T,K)K^2\widetilde{p}(0,T,x,K) \\
&= re^{-rT}K\int_K^\infty \widetilde{p}(0,T,x,y)dy + \frac{1}{2}e^{-rT}\sigma^2(T,K)K^2\widetilde{p}(0,T,x,K) \\
&= -rKc_K(0,T,x,K) + \frac{1}{2}\sigma^2(T,K)K^2c_{KK}(0,T,x,K).
\end{aligned}$$

□

7 Exotic Options

★ **Comments**:

On the PDE approach to pricing knock-out barrier options. We give some clarification to the explanation below Theorem 7.3.1 (page 301-302), where the key is that $V(t)$ and $v(t,x)$ differ by an indicator function and all the hassles come from this difference.

More precisely, we define the first passage time ρ by following the notation of the textbook:

$$\rho = \inf\{t > 0 : S(t) = B\}.$$

Then risk-neutral pricing gives the time-t price of the knock-out barrier option as

$$V(t) = \widetilde{\mathbb{E}}\left[e^{-r(T-t)}V(T)\middle|\mathcal{F}(t)\right]$$

where $V(T) = (S(T)-K)^+1_{\{\rho \geq T\}} = (S(T)-K)^+1_{\{\rho > T\}}$, since the event $\{\rho = T\} \subset \{S(T) = B\}$ has zero probability.

Therefore, the discounted value process $e^{-rt}V(t) = \widetilde{\mathbb{E}}\left[e^{-rT}V(T)\middle|\mathcal{F}(t)\right]$ is a martingale under the risk-neutral measure $\widetilde{\mathbb{P}}$, but we cannot say $V(t)$ is solely a function of t and $S(t)$, since it depends on the path property before t as well. To see this analytically, we note

$$\{\rho > T\} = \{\rho > t, \rho \circ \theta_t > T - t\},$$

where θ_t is the *shift operator* used for sample paths (i.e. $\theta_t(\omega_\cdot) = \omega_{t+\cdot}$. See Øksendal [9, page 119] for details). Then

$$V(t) = \widetilde{\mathbb{E}}\left[e^{-r(T-t)}V(T)\middle|\mathcal{F}(t)\right] = \widetilde{\mathbb{E}}\left[e^{-r(T-t)}(S(T)-K)^+1_{\{\rho\circ\theta_t > T-t\}}\middle|\mathcal{F}(t)\right]1_{\{\rho > t\}}.$$

Note $\rho \circ \theta_t$ is solely dependent on the behavior of sample paths between time t and T. Therefore, we can apply Markov property

$$V(t) = e^{-r(T-t)}\widetilde{\mathbb{E}}^{S(t)}[(S(T-t)-K)^+1_{\{\rho > T-t\}}]1_{\{\rho > t\}}$$

Define $v(t,x) = e^{-r(T-t)}\widetilde{\mathbb{E}}^x[(S(T-t)-K)^+1_{\{\rho > T-t\}}]$, then $V(t) = v(t,x)1_{\{\rho > t\}}$ and $v(t,x)$ satisfies the conditions (7.3.4)-(7.3.7) listed in Theorem 7.3.1. Indeed, it is easy to see $v(t,x)$ satisfies all the boundary conditions (7.3.5)-(7.3.7), by the arguments in the paragraph immediately after Theorem 7.3.1. For the Black-Scholes-Merton PDE (7.3.4), we note

$$d[e^{-rt}V(t)] \quad = \quad d[e^{-rt}v(t,S(t))1_{\{\rho > t\}}] = 1_{\{\rho > t\}}d[e^{-rt}v(t,S(t))] + e^{-rt}v(t,S(t))d1_{\{\rho > t\}}.$$

Following the computation in equation (7.3.13), We can see the Black-Scholes-Merton equation (7.3.4) must hold for $\{(t,x) : 0 \leq t < T, 0 \leq x < B\}$.

7.1. (i)

Proof. Since $\delta_\pm(\tau,s) = \frac{1}{\sigma\sqrt{\tau}}[\log s + (r \pm \frac{1}{2}\sigma^2)\tau] = \frac{\log s}{\sigma}\tau^{-\frac{1}{2}} + \frac{r \pm \frac{1}{2}\sigma^2}{\sigma}\sqrt{\tau}$,

$$\begin{aligned}
\frac{\partial}{\partial t}\delta_\pm(\tau,s) &= \frac{\log s}{\sigma}(-\frac{1}{2})\tau^{-\frac{3}{2}}\frac{\partial\tau}{\partial t} + \frac{r \pm \frac{1}{2}\sigma^2}{\sigma}\frac{1}{2}\tau^{-\frac{1}{2}}\frac{\partial\tau}{\partial t} \\
&= -\frac{1}{2\tau}\left[\frac{\log s}{\sigma}\frac{1}{\sqrt{\tau}}(-1) - \frac{r \pm \frac{1}{2}\sigma^2}{\sigma}\sqrt{\tau}(-1)\right] \\
&= -\frac{1}{2\tau}\cdot\frac{1}{\sigma\sqrt{\tau}}\left[-\log ss + (r \pm \frac{1}{2}\sigma^2)\tau)\right] \\
&= -\frac{1}{2\tau}\delta_\pm(\tau,\frac{1}{s}).
\end{aligned}$$

□

(ii)

Proof.
$$\frac{\partial}{\partial x}\delta_{\pm}(\tau,\frac{x}{c}) = \frac{\partial}{\partial x}\left(\frac{1}{\sigma\sqrt{\tau}}\left[\log\frac{x}{c}+(r\pm\frac{1}{2}\sigma^2)\tau\right]\right) = \frac{1}{x\sigma\sqrt{\tau}},$$
$$\frac{\partial}{\partial x}\delta_{\pm}(\tau,\frac{c}{x}) = \frac{\partial}{\partial x}\left(\frac{1}{\sigma\sqrt{\tau}}\left[\log\frac{c}{x}+(r\pm\frac{1}{2}\sigma^2)\tau\right]\right) = -\frac{1}{x\sigma\sqrt{\tau}}.$$
□

(iii)

Proof.
$$N'(\delta_{\pm}(\tau,s)) = \frac{1}{\sqrt{2\pi}}e^{-\frac{\delta_{\pm}(\tau,s)}{2}} = \frac{1}{\sqrt{2\pi}}e^{-\frac{(\log s+r\tau)^2\pm\sigma^2\tau(\log s+r\tau)+\frac{1}{4}\sigma^4\tau^2}{2\sigma^2\tau}}.$$
Therefore
$$\frac{N'(\delta_+(\tau,s))}{N'(\delta_-(\tau,s))} = e^{-\frac{2\sigma^2\tau(\log s+r\tau)}{2\sigma^2\tau}} = \frac{e^{-r\tau}}{s}$$
and $e^{-r\tau}N'(\delta_-(\tau,s)) = sN'(\delta_+(\tau,s))$. □

(iv)

Proof.
$$\frac{N'(\delta_{\pm}(\tau,s))}{N'(\delta_{\pm}(\tau,s^{-1}))} = e^{-\frac{\left[(\log s+r\tau)^2-(\log\frac{1}{s}+r\tau)^2\right]\pm\sigma^2\tau(\log s-\log\frac{1}{s})}{2\sigma^2\tau}} = e^{-\frac{4r\tau\log s\pm 2\sigma^2\tau\log s}{2\sigma^2\tau}} = e^{-(\frac{2r}{\sigma^2}\pm 1)\log s} = s^{-(\frac{2r}{\sigma^2}\pm 1)}.$$
So $N'(\delta_{\pm}(\tau,s^{-1})) = s^{(\frac{2r}{\sigma^2}\pm 1)}N'(\delta_{\pm}(\tau,s))$. □

(v)

Proof. $\delta_+(\tau,s)-\delta_-(\tau,s) = \frac{1}{\sigma\sqrt{\tau}}\left[\log s+(r+\frac{1}{2}\sigma^2)\tau\right] - \frac{1}{\sigma\sqrt{\tau}}\left[\log s+(r-\frac{1}{2}\sigma^2)\tau\right] = \frac{1}{\sigma\sqrt{\tau}}\sigma^2\tau = \sigma\sqrt{\tau}$. □

(vi)

Proof. $\delta_{\pm}(\tau,s)-\delta_{\pm}(\tau,s^{-1}) = \frac{1}{\sigma\sqrt{\tau}}\left[\log s+(r\pm\frac{1}{2}\sigma^2)\tau\right] - \frac{1}{\sigma\sqrt{\tau}}\left[\log s^{-1}+(r\pm\frac{1}{2}\sigma^2)\tau\right] = \frac{2\log s}{\sigma\sqrt{\tau}}$. □

(vii)

Proof. $N'(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$, so $N''(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}(-\frac{y^2}{2})' = -yN'(y)$. □

To be continued ...

7.3.

Proof. We note $S_T = S_0e^{\sigma\widehat{W}_T} = S_te^{\sigma(\widehat{W}_T-\widehat{W}_t)}$, $\widehat{W}_T-\widehat{W}_t = (\widetilde{W}_T-\widetilde{W}_t)+\alpha(T-t)$ is independent of $\mathcal{F}_t$, $\sup_{t\le u\le T}(\widehat{W}_u-\widehat{W}_t)$ is independent of $\mathcal{F}_t$, and
$$\begin{aligned}
Y_T &= S_0e^{\sigma\widehat{M}_T}\\
&= S_0e^{\sigma\sup_{t\le u\le T}\widehat{W}_u}1_{\{\widehat{M}_t\le\sup_{t\le u\le T}\widehat{W}_t\}} + S_0e^{\sigma\widehat{M}_t}1_{\{\widehat{M}_t>\sup_{t\le u\le T}\widehat{W}_u\}}\\
&= S_te^{\sigma\sup_{t\le u\le T}(\widehat{W}_u-\widehat{W}_t)}1_{\{\frac{Y_t}{S_t}\le e^{\sigma\sup_{t\le u\le T}(\widehat{W}_u-\widehat{W}_t)}\}} + Y_t1_{\{\frac{Y_t}{S_t}\le e^{\sigma\sup_{t\le u\le T}(\widehat{W}_u-\widehat{W}_t)}\}}.
\end{aligned}$$
So $E[f(S_T,Y_T)|\mathcal{F}_t] = E[f(x\frac{S_{T-t}}{S_0}, x\frac{Y_{T-t}}{S_0}1_{\{\frac{y}{x}\le\frac{Y_{T-t}}{S_0}\}}+y1_{\{\frac{y}{x}\le\frac{Y_{T-t}}{S_0}\}})]$, where $x=S_t$, $y=Y_t$. Therefore $E[f(S_T,Y_T)|\mathcal{F}_t]$ is a Borel function of (S_t,Y_t).

□

7.4.

Proof. By Cauchy's inequality and the monotonicity of Y, we have

$$\begin{aligned} |\sum_{j=1}^{m}(Y_{t_j}-Y_{t_{j-1}})(S_{t_j}-S_{t_{j-1}})| &\leq \sum_{j=1}^{m}|Y_{t_j}-Y_{t_{j-1}}||S_{t_j}-S_{t_{j-1}}| \\ &\leq \sqrt{\sum_{j=1}^{m}(Y_{t_j}-Y_{t_{j-1}})^2}\sqrt{\sum_{j=1}^{m}(S_{t_j}-S_{t_{j-1}})^2} \\ &\leq \sqrt{\max_{1\leq j\leq m}|Y_{t_j}-Y_{t_{j-1}}|(Y_T-Y_0)}\sqrt{\sum_{j=1}^{m}(S_{t_j}-S_{t_{j-1}})^2}. \end{aligned}$$

If we increase the number of partition points to infinity and let the length of the longest subinterval $\max_{1\leq j\leq m}|t_j-t_{j-1}|$ approach zero, then $\sqrt{\sum_{j=1}^{m}(S_{t_j}-S_{t_{j-1}})^2} \to \sqrt{[S]_T-[S]_0}<\infty$ and $\max_{1\leq j\leq m}|Y_{t_j}-Y_{t_{j-1}}|\to 0$ a.s. by the continuity of Y. This implies $\sum_{j=1}^{m}(Y_{t_j}-Y_{t_{j-1}})(S_{t_j}-S_{t_{j-1}})\to 0$. □

8 American Derivative Securities

★ **Comments**:

Justification of Definition 8.3.1. For any given stopping time τ, let $C_t=\int_0^t(K-S(u))d1_{\{\tau\leq u\}}$, then

$$e^{-r\tau}(K-S(\tau))1_{\{\tau<\infty\}}=\int_0^\infty e^{-rt}dC_t.$$

Regarding C as a cumulative cash flow, valuation of put option via (8.3.2) is justified by the risk-neutral valuation of a cash flow via (5.6.10).

8.1.

Proof. $v_L'(L+)=(K-L)(-\frac{2r}{\sigma^2})(\frac{x}{L})^{-\frac{2r}{\sigma^2}-1}\frac{1}{L}\Big|_{x=L}=-\frac{2r}{\sigma^2L}(K-L)$. So $v_L'(L+)=v_L'(L-)$ if and only if $-\frac{2r}{\sigma^2L}(K-L)=-1$. Solve for L, we get $L=\frac{2rK}{2r+\sigma^2}$. □

8.2.

Proof. By the calculation in Section 8.3.3, we can see $v_2(x)\geq(K_2-x)^+\geq(K_1-x)^+$, $rv_2(x)-rxv_2'(x)-\frac{1}{2}\sigma^2x^2v_2''(x)\geq 0$ for all $x\geq 0$, and for $0\leq x<L_{1*}<L_{2*}$,

$$rv_2(x)-rxv_2'(x)-\frac{1}{2}\sigma^2x^2v_2''(x)=rK_2>rK_1>0.$$

So the linear complementarity conditions for v_2 imply $v_2(x)=(K_2-x)^+=K_2-x>K_1-x=(K_1-x)^+$ on $[0,L_{1*}]$. Hence $v_2(x)$ does not satisfy the third linear complementarity condition for v_1: for each $x\geq 0$, equality holds in either (8.8.1) or (8.8.2) or both. □

8.3. (i)

Proof. Suppose x takes its values in a domain bounded away from 0. By the general theory of linear differential equations, if we can find two linearly independent solutions $v_1(x)$, $v_2(x)$ of (8.8.4), then any solution of (8.8.4) can be represented in the form of $C_1v_1+C_2v_2$ where C_1 and C_2 are constants. So it suffices to find two linearly independent special solutions of (8.8.4). Assume $v(x)=x^p$ for some constant p to be determined, (8.8.4) yields $x^p(r-pr-\frac{1}{2}\sigma^2p(p-1))=0$. Solve the quadratic equation $0=r-pr-\frac{1}{2}\sigma^2p(p-1)=(-\frac{1}{2}\sigma^2p-r)(p-1)$, we get $p=1$ or $-\frac{2r}{\sigma^2}$. So a general solution of (8.8.4) has the form $C_1x+C_2x^{-\frac{2r}{\sigma^2}}$. □

(ii)

Proof. Assume there is an interval $[x_1, x_2]$ where $0 < x_1 < x_2 < \infty$, such that $v(x) \not\equiv 0$ satisfies (8.3.19) with equality on $[x_1, x_2]$ and satisfies (8.3.18) with equality for x at and immediately to the left of x_1 and for x at and immediately to the right of x_2, then we can find some C_1 and C_2, so that $v(x) = C_1x + C_2x^{-\frac{2r}{\sigma^2}}$ on $[x_1, x_2]$. If for some $x_0 \in [x_1, x_2]$, $v(x_0) = v'(x_0) = 0$, by the uniqueness of the solution of (8.8.4), we would conclude $v \equiv 0$. This is a contradiction. So such an x_0 cannot exist. This implies $0 < x_1 < x_2 < K$ (if $K \leq x_2$, $v(x_2) = (K - x_2)^+ = 0$ and $v'(x_2)$=the right derivative of $(K - x)^+$ at x_2, which is 0). [9] Thus we have four equations for C_1 and C_2:

$$\begin{cases} C_1x_1 + C_2x_1^{-\frac{2r}{\sigma^2}} = K - x_1 \\ C_1x_2 + C_2x_2^{-\frac{2r}{\sigma^2}} = K - x_2 \\ C_1 - \frac{2r}{\sigma^2}C_2x_1^{-\frac{2r}{\sigma^2}-1} = -1 \\ C_1 - \frac{2r}{\sigma^2}C_2x_2^{-\frac{2r}{\sigma^2}-1} = -1. \end{cases}$$

Since $x_1 \neq x_2$, the last two equations imply $C_2 = 0$. Plug $C_2 = 0$ into the first two equations, we have $C_1 = \frac{K-x_1}{x_1} = \frac{K-x_2}{x_2}$; plug $C_2 = 0$ into the last two equations, we have $C_1 = -1$. Combined, we would have $x_1 = x_2$. Contradiction. Therefore our initial assumption is incorrect, and the only solution v that satisfies the specified conditions in the problem is the zero solution. ☐

(iii)

Proof. If in a right neighborhood of 0, v satisfies (8.3.19) with equality, then part (i) implies $v(x) = C_1x + C_2x^{-\frac{2r}{\sigma^2}}$ for some constants C_1 and C_2. Then $v(0) = \lim_{x\downarrow 0} v(x) = 0 < (K - 0)^+$, i.e. (8.3.18) will be violated. So we must have $rv - rxv' - \frac{1}{2}\sigma^2x^2v'' > 0$ in a right neighborhood of 0. According to (8.3.20), $v(x) = (K - x)^+$ near o. So $v(0) = K$. We have thus concluded simultaneously that v cannot satisfy (8.3.19) with equality near 0 and $v(0) = K$, starting from first principles (8.3.18)-(8.3.20). ☐

(iv)

Proof. This is already shown in our solution of part (iii): near 0, v cannot satisfy (8.3.19) with equality. ☐

(v)

Proof. If v satisfy $(K - x)^+$ with equality for all $x \geq 0$, then v cannot have a continuous derivative as stated in the problem. This is a contradiction. ☐

(vi)

Proof. By the result of part (i), we can start with $v(x) = (K - x)^+$ on $[0, x_1]$ and $v(x) = C_1x + C_2x^{-\frac{2r}{\sigma^2}}$ on $[x_1, \infty)$. By the assumption of the problem, both v and v' are continuous. Since $(K - x)^+$ is not differentiable at K, we must have $x_1 \leq K$.This gives us the equations

$$\begin{cases} K - x_1 = (K - x_1)^+ = C_1x_1 + C_2x_1^{-\frac{2r}{\sigma^2}} \\ -1 = C_1 - \frac{2r}{\sigma^2}C_2x_1^{-\frac{2r}{\sigma^2}-1}. \end{cases}$$

Because v is assumed to be bounded, we must have $C_1 = 0$ and the above equations only have two unknowns: C_2 and x_1. Solve them for C_2 and x_1, we are done. ☐

8.4. (i)

Proof. This is already shown in part (i) of Exercise 8.3. ☐

[9] Note we have interpreted the condition "$v(x)$ satisfies (8.3.18) with equality for x at and immediately to the right of x_2" as "$v(x_2) = (K - x_2)^+$ and $v'(x_2)$ =the right derivative of $(K - x)^+$ at x_2." This is weaker than "$v(x) = (K - x)$ in a right neighborhood of x_2."

(ii)

Proof. We solve for A, B the equations

$$\begin{cases} AL^{-\frac{2r}{\sigma^2}} + BL = K - L \\ -\frac{2r}{\sigma^2}AL^{-\frac{2r}{\sigma^2}-1} + B = -1, \end{cases}$$

and we obtain $A = \frac{\sigma^2 K L^{\frac{2r}{\sigma^2}}}{\sigma^2+2r}$, $B = \frac{2rK}{L(\sigma^2+2r)} - 1$. □

(iii)

Proof. By (8.8.5), $B > 0$. So for $x \geq K$, $f(x) \geq BK > 0 = (K-x)^+$. If $L \leq x < K$,

$$\begin{aligned} f(x) - (K-x)^+ &= \frac{\sigma^2 K L^{\frac{2r}{\sigma^2}}}{\sigma^2+2r} x^{-\frac{2r}{\sigma^2}} + \frac{2rKx}{L(\sigma^2+2r)} - K \\ &= x^{-\frac{2r}{\sigma^2}} \frac{KL^{\frac{2r}{\sigma^2}}\left[\sigma^2 + 2r(\frac{x}{L})^{\frac{2r}{\sigma^2}+1} - (\sigma^2+2r)(\frac{x}{L})^{\frac{2r}{\sigma^2}}\right]}{(\sigma^2+2r)L}. \end{aligned}$$

Let $g(\theta) = \sigma^2 + 2r\theta^{\frac{2r}{\sigma^2}+1} - (\sigma^2+2r)\theta^{\frac{2r}{\sigma^2}}$ with $\theta \geq 1$. Then $g(1) = 0$ and $g'(\theta) = 2r(\frac{2r}{\sigma^2}+1)\theta^{\frac{2r}{\sigma^2}} - (\sigma^2+2r)\frac{2r}{\sigma^2}\theta^{\frac{2r}{\sigma^2}-1} = \frac{2r}{\sigma^2}(\sigma^2+2r)\theta^{\frac{2r}{\sigma^2}-1}(\theta-1) \geq 0$. So $g(\theta) \geq 0$ for any $\theta \geq 1$. This shows $f(x) \geq (K-x)^+$ for $L \leq x < K$. Combined, we get $f(x) \geq (K-x)^+$ for all $x \geq L$. □

(iv)

Proof. Since $\lim_{x\to\infty} v(x) = \lim_{x\to\infty} f(x) = \infty$ and $\lim_{x\to\infty} v_{L_*}(x) = \lim_{x\to\infty}(K-L_*)(\frac{x}{L_*})^{-\frac{2r}{\sigma^2}} = 0$, $v(x)$ and $v_{L_*}(x)$ are different. By part (iii), $v(x) \geq (K-x)^+$. So v satisfies (8.3.18). For $x \geq L$, $rv - rxv' - \frac{1}{2}\sigma^2x^2v'' = rf - rxf - \frac{1}{2}\sigma^2x^2f'' = 0$. For $0 \leq x \leq L$, $rv - rxv' - \frac{1}{2}\sigma^2x^2v'' = r(K-x) + rx = rK$. Combined, $rv - rxv' - \frac{1}{2}\sigma^2x^2v'' \geq 0$ for $x \geq 0$. So v satisfies (8.3.19). Along the way, we also showed v satisfies (8.3.20). In summary, v satisfies the linear complementarity condition (8.3.18)-(8.3.20), but v is not the function v_{L_*} given by (8.3.13). □

(v)

Proof. By part (ii), $B = 0$ if and only if $\frac{2rK}{L(\sigma^2+2r)} - 1 = 0$, i.e. $L = \frac{2rK}{2r+\sigma^2}$. In this case, $v(x) = Ax^{-\frac{2r}{\sigma^2}} = \frac{\sigma^2 K}{\sigma^2+2r}(\frac{x}{L})^{-\frac{2r}{\sigma^2}} = (K-L)(\frac{x}{L})^{-\frac{2r}{\sigma^2}} = v_{L_*}(x)$, on the interval $[L, \infty)$. □

8.5. The difficulty of the dividend-paying case is that from Lemma 8.3.4, we can only obtain $\widetilde{E}[e^{-(r-a)\tau_L}]$, not $\widetilde{E}[e^{-r\tau_L}]$. So we have to start from Theorem 8.3.2.

(i)

Proof. By (8.8.9), $S_t = S_0 e^{\sigma\widetilde{W}_t + (r-a-\frac{1}{2}\sigma^2)t}$. Assume $S_0 = x$, then $S_t = L$ if and only if $-\widetilde{W}_t - \frac{1}{\sigma}(r-a-\frac{1}{2}\sigma^2)t = \frac{1}{\sigma}\log\frac{x}{L}$. By Theorem 8.3.2,

$$\widetilde{E}[e^{-r\tau_L}] = e^{-\frac{1}{\sigma}\log\frac{x}{L}\left[\frac{1}{\sigma}(r-a-\frac{1}{2}\sigma^2) + \sqrt{\frac{1}{\sigma^2}(r-a-\frac{1}{2}\sigma^2)^2 + 2r}\right]}.$$

If we set $\gamma = \frac{1}{\sigma^2}(r-a-\frac{1}{2}\sigma^2) + \frac{1}{\sigma}\sqrt{\frac{1}{\sigma^2}(r-a-\frac{1}{\sigma^2})^2 + 2r}$, we can write $\widetilde{E}[e^{-r\tau_L}]$ as $e^{-\gamma\log\frac{x}{L}} = (\frac{x}{L})^{-\gamma}$. So the risk-neutral expected discounted pay off of this strategy is

$$v_L(x) = \begin{cases} K - x, & 0 \leq x \leq L \\ (K-L)(\frac{x}{L})^{-\gamma}, & x > L. \end{cases}$$

□

(ii)

Proof. $\frac{\partial}{\partial L}v_L(x) = -(\frac{x}{L})^{-\gamma}(1-\frac{\gamma(K-L)}{L})$. Set $\frac{\partial}{\partial L}v_L(x)=0$ and solve for L_*, we have $L_*=\frac{\gamma K}{\gamma+1}$. □

(iii)

Proof. By Itô's formula, we have

$$d\left[e^{-rt}v_{L_*}(S_t)\right] = e^{-rt}\left[-rv_{L_*}(S_t)+v'_{L_*}(S_t)(r-a)S_t+\frac{1}{2}v''_{L_*}(S_t)\sigma^2S_t^2\right]dt+e^{-rt}v'_{L_*}(S_t)\sigma S_t d\widetilde{W}_t.$$

If $x>L_*$,

$$\begin{aligned}
& -rv_{L_*}(x)+v'_{L_*}(x)(r-a)x+\frac{1}{2}v''_{L_*}(x)\sigma^2x^2 \\
= & -r(K-L_*)\left(\frac{x}{L_*}\right)^{-\gamma}+(r-a)x(K-L_*)(-\gamma)\frac{x^{-\gamma-1}}{L_*^{-\gamma}}+\frac{1}{2}\sigma^2x^2(-\gamma)(-\gamma-1)(K-L_*)\frac{x^{-\gamma-2}}{L_*^{-\gamma}} \\
= & (K-L_*)\left(\frac{x}{L_*}\right)^{-\gamma}\left[-r-(r-a)\gamma+\frac{1}{2}\sigma^2\gamma(\gamma+1)\right].
\end{aligned}$$

By the definition of γ, if we define $u=r-a-\frac{1}{2}\sigma^2$, we have

$$\begin{aligned}
& r+(r-a)\gamma-\frac{1}{2}\sigma^2\gamma(\gamma+1) \\
= & r-\frac{1}{2}\sigma^2\gamma^2+\gamma(r-a-\frac{1}{2}\sigma^2) \\
= & r-\frac{1}{2}\sigma^2\left(\frac{u}{\sigma^2}+\frac{1}{\sigma}\sqrt{\frac{u^2}{\sigma^2}+2r}\right)^2+\left(\frac{u}{\sigma^2}+\frac{1}{\sigma}\sqrt{\frac{u^2}{\sigma^2}+2r}\right)u \\
= & r-\frac{1}{2}\sigma^2\left(\frac{u^2}{\sigma^4}+\frac{2u}{\sigma^3}\sqrt{\frac{u^2}{\sigma^2}+2r}+\frac{1}{\sigma^2}\left(\frac{u^2}{\sigma^2}+2r\right)\right)+\frac{u^2}{\sigma^2}+\frac{u}{\sigma}\sqrt{\frac{u^2}{\sigma^2}+2r} \\
= & r-\frac{u^2}{2\sigma^2}-\frac{u}{\sigma}\sqrt{\frac{u^2}{\sigma^2}+2r}-\frac{1}{2}\left(\frac{u^2}{\sigma^2}+2r\right)+\frac{u^2}{\sigma^2}+\frac{u}{\sigma}\sqrt{\frac{u^2}{\sigma^2}+2r} \\
= & 0.
\end{aligned}$$

If $x<L_*$, $-rv_{L_*}(x)+v'_{L_*}(x)(r-a)x+\frac{1}{2}v''_{L_*}(x)\sigma^2x^2=-r(K-x)+(-1)(r-a)x=-rK+ax$. Combined, we get

$$d\left[e^{-rt}v_{L_*}(S_t)\right]=-e^{-rt}1_{\{S_t<L_*\}}(rK-aS_t)dt+e^{-rt}v'_{L_*}(S_t)\sigma S_t d\widetilde{W}_t.$$

Following the reasoning in the proof of Theorem 8.3.5, we only need to show $1_{\{x<L_*\}}(rK-ax)\geq 0$ to finish the solution. This is further equivalent to proving $rK-aL_*\geq 0$. Plug $L_*=\frac{\gamma K}{\gamma+1}$ into the expression and note $\gamma\geq\frac{1}{\sigma}\sqrt{\frac{1}{\sigma^2}(r-a-\frac{1}{2}\sigma^2)^2}+\frac{1}{\sigma^2}(r-a-\frac{1}{2}\sigma^2)\geq 0$, the inequality is further reduced to $r(\gamma+1)-a\gamma\geq 0$. We prove this inequality as follows.

Assume for some K, r, a and σ (K and σ are assumed to be strictly positive, r and a are assumed to be non-negative), $rK-aL_*<0$, then necessarily $r<a$, since $L_*=\frac{\gamma K}{\gamma+1}\leq K$. As shown before, this means $r(\gamma+1)-a\gamma<0$. Define $\theta=\frac{r-a}{\sigma}$, then $\theta<0$ and $\gamma=\frac{1}{\sigma^2}(r-a-\frac{1}{2}\sigma^2)+\frac{1}{\sigma}\sqrt{\frac{1}{\sigma^2}(r-a-\frac{1}{2}\sigma^2)^2+2r}=$

$\frac{1}{\sigma}(\theta - \frac{1}{2}\sigma) + \frac{1}{\sigma}\sqrt{(\theta - \frac{1}{2}\sigma)^2 + 2r}$. We have

$$\begin{aligned}
r(\gamma+1) - a\gamma < 0 &\iff (r-a)\gamma + r < 0 \\
&\iff (r-a)\left[\frac{1}{\sigma}(\theta - \frac{1}{2}\sigma) + \frac{1}{\sigma}\sqrt{(\theta - \frac{1}{2}\sigma)^2 + 2r}\right] + r < 0 \\
&\iff \theta(\theta - \frac{1}{2}\sigma) + \theta\sqrt{(\theta - \frac{1}{2}\sigma)^2 + 2r} + r < 0 \\
&\iff \theta\sqrt{(\theta - \frac{1}{2}\sigma)^2 + 2r} < -r - \theta(\theta - \frac{1}{2}\sigma)(< 0) \\
&\iff \theta^2[(\theta - \frac{1}{2}\sigma)^2 + 2r] > r^2 + \theta^2(\theta - \frac{1}{2}\sigma)^2 + 2\theta r(\theta - \frac{1}{2}\sigma^2) \\
&\iff 0 > r^2 - \theta r\sigma^2 \\
&\iff 0 > r - \theta\sigma^2.
\end{aligned}$$

Since $\theta\sigma^2 < 0$, we have obtained a contradiction. So our initial assumption is incorrect, and $rK - aL_* \geq 0$ must be true. □

(iv)

Proof. The proof is similar to that of Corollary 8.3.6. Note the only properties used in the proof of Corollary 8.3.6 are that $e^{-rt}v_{L_*}(S_t)$ is a supermartingale, $e^{-rt\wedge\tau_{L_*}}v_{L_*}(S_t \wedge \tau_{L_*})$ is a martingale, and $v_{L_*}(x) \geq (K-x)^+$. Part (iii) already proved the supermartingale-martingale property, so it suffices to show $v_{L_*}(x) \geq (K-x)^+$ in our problem. Indeed, by $\gamma \geq 0$, $L_* = \frac{\gamma K}{\gamma+1} < K$. For $x \geq K > L_*$, $v_{L_*}(x) > 0 = (K-x)^+$; for $0 \leq x < L_*$, $v_{L_*}(x) = K - x = (K-x)^+$; finally, for $L_* \leq x \leq K$,

$$\frac{d}{dx}(v_{L_*}(x) - (K-x)) = -\gamma(K - L_*)\frac{x^{-\gamma-1}}{L_*^{-\gamma}} + 1 \geq -\gamma(K - L_*)\frac{L_*^{-\gamma-1}}{L_*^{-\gamma}} + 1 = -\gamma(K - \frac{\gamma K}{\gamma+1})\frac{1}{\frac{\gamma K}{\gamma+1}} + 1 = 0.$$

and $(v_{L_*}(x) - (K-x))|_{x=L_*} = 0$. So for $L_* \leq x \leq K$, $v_{L_*}(x) - (K-x)^+ \geq 0$. Combined, we have $v_{L_*}(x) \geq (K-x)^+ \geq 0$ for all $x \geq 0$. □

8.6.

Proof. By Lemma 8.5.1, $X_t = e^{-rt}(S_t - K)^+$ is a submartingale. For any $\tau \in \Gamma_{0,T}$, Theorem 8.8.1 implies

$$\widetilde{E}[e^{-rT}(S_T - K)^+] \geq \widetilde{E}[e^{-r\tau\wedge T}(S_{\tau\wedge T} - K)^+] \geq E[e^{-r\tau}(S_\tau - K)^+ 1_{\{\tau<\infty\}}] = E[e^{-r\tau}(S_\tau - K)^+],$$

where we take the convention that $e^{-r\tau}(S_\tau - K)^+ = 0$ when $\tau = \infty$. Since τ is arbitrarily chosen, $\widetilde{E}[e^{-rT}(S_T - K)^+] \geq \max_{\tau\in\Gamma_{0,T}} \widetilde{E}[e^{-r\tau}(S_\tau - K)^+]$. The other direction "$\leq$" is trivial since $T \in \Gamma_{0,T}$. □

8.7.

Proof. Suppose $\lambda \in [0,1]$ and $0 \leq x_1 \leq x_2$, we have $f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2) \leq (1-\lambda)h(x_1) + \lambda h(x_2)$. Similarly, $g((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)h(x_1) + \lambda h(x_2)$. So

$$h((1-\lambda)x_1 + \lambda x_2) = \max\{f((1-\lambda)x_1 + \lambda x_2), g((1-\lambda)x_1 + \lambda x_2)\} \leq (1-\lambda)h(x_1) + \lambda h(x_2).$$

That is, h is also convex. □

9 Change of Numéraire

★ **Comments**:

1) To provide an intuition for change of numéraire, we give *a summary of results for change of numéraire in discrete case*. This summary is based on Shiryaev [13].

Consider a model of financial market $(\widetilde{B}, \bar{B}, \mathbf{S})$ as in Delbaen and Schachermayer [2] Definition 2.1.1 or Shiryaev [13, page 383]. Here $\widetilde{B}$ and $\bar{B}$ are both one-dimensional while $\mathbf{S}$ could be a vector price process. Suppose $\widetilde{B}$ and $\bar{B}$ are both strictly positive, then both of them can be chosen as numéaire.

Several results hold under this model. First, no-arbitrage and completeness properties of market are independent of the choice of numéraire (see, for example, Shiryaev [13, page 413, 481]).

Second, if the market is arbitrage-free, then corresponding to $\widetilde{B}$ (resp. $\bar{B}$), there is an equivalent probability measure $\widetilde{\mathbb{P}}$ (resp. $\bar{\mathbb{P}}$), such that $\left(\frac{\bar{B}}{\widetilde{B}}, \frac{\mathbf{S}}{\widetilde{B}}\right)$ (resp. $\left(\frac{\widetilde{B}}{\bar{B}}, \frac{\mathbf{S}}{\bar{B}}\right)$) is a martingale under $\widetilde{\mathbb{P}}$ (resp. $\bar{\mathbb{P}}$).

Third, if the market is both arbitrage-free and complete, we have the relation

$$d\bar{\mathbb{P}} = \frac{\bar{B}_T}{\widetilde{B}_T} \frac{1}{\mathbb{E}\left[\frac{\bar{B}_0}{\widetilde{B}_0}\right]} d\widetilde{\mathbb{P}}. \tag{6}$$

See Shiryaev [13, page 510], formula (12).

Finally, if f_T is a European contingent claim with maturity N and the market is both arbitrage-free and complete, then [10]

$$\bar{B}_t \bar{\mathbb{E}}\left[\left.\frac{f_T}{\bar{B}_T}\right|\mathcal{F}_t\right] = \widetilde{B}_t \widetilde{\mathbb{E}}\left[\left.\frac{f_T}{\widetilde{B}_T}\right|\mathcal{F}_t\right].$$

That is, the risk-neutral price of f_T is independent of the choice of numéraire. See Shiryaev [13], Chapter VI, §1b.2.[11]

2) The above theoretical results can be applied to market involving foreign money market account. We consider the following market: a domestic money market account M $(M_0 = 1)$, a foreign money market account M^f $(M_0^f = 1)$, a (vector) asset price process $\mathbf{S}$ denominated in domestic currency called stock. Suppose the domestic vs. foreign currency exchange rate is Q. Note Q is not a traded asset. Denominated by domestic currency, the traded assets are $(M, M^f Q, \mathbf{S})$, where $M^f Q$ can be seen as the price process of one unit foreign currency. Domestic risk-neutral measure $\widetilde{\mathbb{P}}$ is such that $\left(\frac{M^f Q}{M}, \frac{\mathbf{S}}{M}\right)$ is a $\widetilde{\mathbb{P}}$-martingale. Denominated by foreign currency, the traded assets are $\left(M^f, \frac{M}{Q}, \frac{\mathbf{S}}{Q}\right)$. Foreign risk-neutral measure $\widetilde{\mathbb{P}}^f$ is such that $\left(\frac{M}{QM^f}, \frac{\mathbf{S}}{QM^f}\right)$ is a $\widetilde{\mathbb{P}}^f$-martingale. This is a change of numéraire in the market denominated by domestic currency, from M to $M^f Q$. If we assume the market is arbitrage-free and complete, the foreign

[10] If the market is incomplete but the contingent claim is still replicable, this result still holds for $t = 0$. Indeed, in Shiryaev [13], Chapter V §1c.2, it is easy to generalize formula (12) to the case of $N \geq 1$ with $x^* = \sup_{\widetilde{\mathbb{P}} \in \mathcal{P}(\mathbb{P})} \widetilde{\mathbb{E}}\left[\frac{f_N}{B_N}\right] B_0$, $x_* = \inf_{\widetilde{\mathbb{P}} \in \mathcal{P}(\mathbb{P})} \widetilde{\mathbb{E}}\left[\frac{f_N}{B_N}\right] B_0$, $C^*(\mathbb{P}) = \inf\{x \geq 0 : \exists \pi \in SF, x_0^\pi = x, x_N^\pi \geq f_N\}$, and $C_*(\mathbb{P}) = \sup\{x \geq 0 : \exists \pi \in SF, x_0^\pi = x, x_N^\pi \leq f_N\}$. No-arbitrage implies $C_* \leq C^*$. Since

$$\frac{X_N^\pi}{B_N} = \frac{X_0^\pi}{B_0} + \sum_{k=1}^N \gamma_k \Delta\left(\frac{S_k}{B_k}\right),$$

we must have $C_* \leq x_* \leq x^* \leq C^*$. The replicability of f_N implies $C^* \leq C_*$. So, if the market is incomplete but the contingent claim is still replicable, we still have $C_* = x_* = x^* = C^*$, i.e. the risk-neutral pricing formula still holds for $t = 0$.

[11] The invariance of risk-neutral price gives us a mnemonics to memorize formula (6): take $f_T = 1_A \bar{B}_T$ with $A \in \mathcal{F}_T$ and set $t = 0$, we have

$$\bar{B}_0 \bar{\mathbb{P}}(A) = \widetilde{B}_0 \widetilde{\mathbb{E}}\left[\frac{\bar{B}_T}{\widetilde{B}_T}; A\right].$$

So the Radon-Nikodým derivative is $d\bar{\mathbb{P}}/d\widetilde{\mathbb{P}} = \frac{\bar{B}_T/\bar{B}_0}{\widetilde{B}_T/\widetilde{B}_0}$. This leads to formula (9.2.6) in Shreve [14, page 378]. The textbook's chapter summary also provides a good way to memorize: the Radon-Nikodým derivative "is the numéraire itself, discounted in order to be a martingale and normalized by its initial condition in order to have expected value 1".

risk-neutral measure is given by

$$d\widetilde{\mathbb{P}}^f = \frac{Q_T M_T^f}{M_T \mathbb{E}\left[\frac{Q_0 M_0^f}{M_0}\right]} d\widetilde{\mathbb{P}} = \frac{Q_T D_T M_T^f}{Q_0} d\widetilde{\mathbb{P}}$$

on $\mathcal{F}_T$. For a European contingent claim f_T denominated in domestic currency, its payoff in foreign currency is f_T/Q_T. Therefore its foreign price is $\widetilde{\mathbb{E}}^f\left[\frac{D_T^f f_T}{D_t^f Q_T}\middle|\mathcal{F}_t\right]$. Convert this price into domestic currency, we have $Q_t \widetilde{\mathbb{E}}^f\left[\frac{D_T^f f_T}{D_t^f Q_T}\middle|\mathcal{F}_t\right]$. Use the relation between $\widetilde{\mathbb{P}}^f$ and $\widetilde{\mathbb{P}}$ on $\mathcal{F}_T$ and the Bayes formula, we get

$$Q_t \widetilde{\mathbb{E}}^f\left[\frac{D_T^f f_T}{D_t^f Q_T}\middle|\mathcal{F}_t\right] = \widetilde{\mathbb{E}}\left[\frac{D_T f_T}{D_t}\middle|\mathcal{F}_t\right].$$

The RHS is exactly the price of f_T in domestic market if we apply risk-neutral pricing.

3) *Alternative proof of Theorem 9.4.2 (Black-Scholes-Merton option pricing with random interest rate).* By (9.4.7), $V(t) = B(t,T)\widetilde{\mathbb{E}}^T[V(T)|\mathcal{F}(t)]$, where $V(T) = (S(T)-K)^+$ with

$$S(T) = \text{For}_S(T,T) = \text{For}_S(t,T)\exp\left\{\sigma[\widetilde{W}^T(T) - \widetilde{W}^T(t)] - \frac{1}{2}\sigma^2(T-t)\right\}.$$

Then the valuation of $\widetilde{\mathbb{E}}^T[V(T)|\mathcal{F}(t)]$ is like the Black-Scholes-Merton model for stock options, where interest rate is 0 and the underlying has price $\text{For}_S(t,T)$ at time t. So its value is

$$\widetilde{\mathbb{E}}^T[V(T)|\mathcal{F}(t)] = \text{For}_S(t,T)N(d_+(t)) - KN(d_-(t))$$

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

$$V(t) = B(t,T)[\text{For}_S(t,T)N(d_+) - KN(d_-)] = S(t)N(d_+(t)) - KB(t,T)N(d_-(t)).$$

9.1. (i)

Proof. For any $0 \le t \le T$, by Lemma 5.5.2,

$$E^{(M_2)}\left[\frac{M_1(T)}{M_2(T)}\middle|\mathcal{F}_t\right] = E\left[\frac{M_2(T)}{M_2(t)}\frac{M_1(T)}{M_2(T)}\middle|\mathcal{F}_t\right] = \frac{E[M_1(T)|\mathcal{F}_t]}{M_2(t)} = \frac{M_1(t)}{M_2(t)}.$$

So $\frac{M_1(t)}{M_2(t)}$ is a martingale under P^{M_2}. □

(ii)

Proof. Let $M_1(t) = D_t S_t$ and $M_2(t) = D_t N_t / N_0$. Then $\widetilde{P}^{(N)}$ as defined in (9.2.6) is $P^{(M_2)}$ as defined in Remark 9.2.5. Hence $\frac{M_1(t)}{M_2(t)} = \frac{S_t}{N_t}N_0$ is a martingale under $\widetilde{P}^{(N)}$, which implies $S_t^{(N)} = \frac{S_t}{N_t}$ is a martingale under $\widetilde{P}^{(N)}$. □

9.2. (i)

Proof. Since $N_t^{-1} = N_0^{-1}e^{-\nu\widetilde{W}_t - (r-\frac{1}{2}\nu^2)t}$, we have

$$d(N_t^{-1}) = N_0^{-1}e^{-\nu\widetilde{W}_t-(r-\frac{1}{2}\nu^2)t}[-\nu d\widetilde{W}_t - (r-\frac{1}{2}\nu^2)dt + \frac{1}{2}\nu^2 dt] = N_t^{-1}(-\nu d\widehat{W}_t - rdt).$$

□

(ii)

Proof.

$$d\widehat{M}_t = M_t d\left(\frac{1}{N_t}\right) + \frac{1}{N_t} dM_t + d\left(\frac{1}{N_t}\right) dM_t = \widehat{M}_t(-\nu d\widehat{W}_t - rdt) + r\widehat{M}_t dt = -\nu \widehat{M}_t d\widehat{W}_t.$$

Remark: This can also be obtained directly from Theorem 9.2.2. □

(iii)

Proof.

$$\begin{aligned} d\widehat{X}_t &= d\left(\frac{X_t}{N_t}\right) = X_t d\left(\frac{1}{N_t}\right) + \frac{1}{N_t} dX_t + d\left(\frac{1}{N_t}\right) dX_t \\ &= (\Delta_t S_t + \Gamma_t M_t) d\left(\frac{1}{N_t}\right) + \frac{1}{N_t}(\Delta_t dS_t + \Gamma_t dM_t) + d\left(\frac{1}{N_t}\right)(\Delta_t dS_t + \Gamma_t dM_t) \\ &= \Delta_t \left[S_t d\left(\frac{1}{N_t}\right) + \frac{1}{N_t} dS_t + d\left(\frac{1}{N_t}\right) dS_t\right] + \Gamma_t \left[M_t d\left(\frac{1}{N_t}\right) + \frac{1}{N_t} dM_t + d\left(\frac{1}{N_t}\right) dM_t\right] \\ &= \Delta_t d\widehat{S}_t + \Gamma_t d\widehat{M}_t. \end{aligned}$$

□

9.3. To avoid singular cases, we need to assume $-1 < \rho < 1$.

(i)

Proof. $N_t = N_0 e^{\nu \widetilde{W}_3(t) + (r - \frac{1}{2}\nu^2)t}$. So

$$\begin{aligned} dN_t^{-1} &= d(N_0^{-1} e^{-\nu \widetilde{W}_3(t) - (r - \frac{1}{2}\nu^2)t}) \\ &= N_0^{-1} e^{-\nu \widetilde{W}_3(t) - (r - \frac{1}{2}\nu^2)t}\left[-\nu d\widetilde{W}_3(t) - (r - \frac{1}{2}\nu^2)dt + \frac{1}{2}\nu^2 dt\right] \\ &= N_t^{-1}[-\nu d\widetilde{W}_3(t) - (r - \nu^2)dt], \end{aligned}$$

and

$$\begin{aligned} dS_t^{(N)} &= N_t^{-1} dS_t + S_t dN_t^{-1} + dS_t dN_t^{-1} \\ &= N_t^{-1}(rS_t dt + \sigma S_t d\widetilde{W}_1(t)) + S_t N_t^{-1}[-\nu d\widetilde{W}_3(t) - (r - \nu^2)dt] \\ &= S_t^{(N)}(rdt + \sigma d\widetilde{W}_1(t)) + S_t^{(N)}[-\nu d\widetilde{W}_3(t) - (r - \nu^2)dt] - \sigma S_t^{(N)} \rho dt \\ &= S_t^{(N)}(\nu^2 - \sigma\rho)dt + S_t^{(N)}(\sigma d\widetilde{W}_1(t) - \nu d\widetilde{W}_3(t)). \end{aligned}$$

Define $\gamma = \sqrt{\sigma^2 - 2\rho\sigma\nu + \nu^2}$ and $\widetilde{W}_4(t) = \frac{\sigma}{\gamma}\widetilde{W}_1(t) - \frac{\nu}{\gamma}\widetilde{W}_3(t)$, then $\widetilde{W}_4$ is a martingale with quadratic variation

$$[\widetilde{W}_4]_t = \frac{\sigma^2}{\gamma^2} t - 2\frac{\sigma\nu}{\gamma^2}\rho t + \frac{\nu^2}{r^2} t = t.$$

By Lévy's Theorem, $\widetilde{W}_4$ is a BM and therefore, $S_t^{(N)}$ has volatility $\gamma = \sqrt{\sigma^2 - 2\rho\sigma\nu + \nu^2}$. □

(ii)

Proof. This problem is the same as Exercise 4.13, we define $\widetilde{W}_2(t) = \frac{-\rho}{\sqrt{1-\rho^2}}\widetilde{W}_1(t) + \frac{1}{\sqrt{1-\rho^2}}\widetilde{W}_3(t)$, then $\widetilde{W}_2$ is a martingale, with

$$(d\widetilde{W}_2(t))^2 = \left(-\frac{\rho}{\sqrt{1-\rho^2}} d\widetilde{W}_1(t) + \frac{1}{\sqrt{1-\rho^2}} d\widetilde{W}_3(t)\right)^2 = \left(\frac{\rho^2}{1-\rho^2} + \frac{1}{1-\rho^2} - \frac{2\rho^2}{1-\rho^2}\right) dt = dt,$$

and $d\widetilde{W}_2(t) d\widetilde{W}_1(t) = -\frac{\rho}{\sqrt{1-\rho^2}} dt + \frac{\rho}{\sqrt{1-\rho^2}} dt = 0$. So $\widetilde{W}_2$ is a BM independent of $\widetilde{W}_1$, and $dN_t = rN_t dt + \nu N_t d\widetilde{W}_3(t) = rN_t dt + \nu N_t[\rho d\widetilde{W}_1(t) + \sqrt{1-\rho^2} d\widetilde{W}_2(t)]$. □

(iii)

Proof. Under $\widetilde{P}$, $(\widetilde{W}_1, \widetilde{W}_2)$ is a two-dimensional BM, and

$$\begin{cases} dS_t = rS_t dt + \sigma S_t d\widetilde{W}_1(t) = rS_t dt + S_t(\sigma, 0) \cdot \begin{pmatrix} d\widetilde{W}_1(t) \\ d\widetilde{W}_2(t) \end{pmatrix} \\ dN_t = rN_t dt + \nu N_t d\widetilde{W}_3(t) = rN_t dt + N_t(\nu\rho, \nu\sqrt{1-\rho^2}) \cdot \begin{pmatrix} d\widetilde{W}_1(t) \\ d\widetilde{W}_2(t) \end{pmatrix}. \end{cases}$$

So under $\widetilde{P}$, the volatility vector for S is $(\sigma, 0)$, and the volatility vector for N is $(\nu\rho, \nu\sqrt{1-\rho^2})$. By Theorem 9.2.2, under the measure $\widetilde{P}^{(N)}$, the volatility vector for $S^{(N)}$ is $(v_1, v_2) = (\sigma - \nu\rho, -\nu\sqrt{1-\rho^2}$. In particular, the volatility of $S^{(N)}$ is

$$\sqrt{v_1^2 + v_2^2} = \sqrt{(\sigma - \nu\rho)^2 + (-\nu\sqrt{1-\rho^2})^2} = \sqrt{\sigma^2 - 2\nu\rho\sigma + \nu^2},$$

consistent with the result of part (i). □

9.4.

Proof. From (9.3.15), we have $M_t^f Q_t = M_0^f Q_0 e^{\int_0^t \sigma_2(s) d\widetilde{W}_3(s) + \int_0^t (R_s - \frac{1}{2}\sigma_2^2(s)) ds}$. So

$$\frac{D_t^f}{Q_t} = D_0^f Q_0^{-1} e^{-\int_0^t \sigma_2(s) d\widetilde{W}_3(s) - \int_0^t (R_s - \frac{1}{2}\sigma_2^2(s)) ds}$$

and

$$d\left(\frac{D_t^f}{Q_t}\right) = \frac{D_t^f}{Q_t}[-\sigma_2(t) d\widetilde{W}_3(t) - (R_t - \frac{1}{2}\sigma_2^2(t))dt + \frac{1}{2}\sigma_2^2(t)dt] = \frac{D_t^f}{Q_t}[-\sigma_2(t) d\widetilde{W}_3(t) - (R_t - \sigma_2^2(t))dt].$$

To get (9.3.22), we note

$$\begin{aligned} d\left(\frac{M_t D_t^f}{Q_t}\right) &= M_t d\left(\frac{D_t^f}{Q_t}\right) + \frac{D_t^f}{Q_t} dM_t + dM_t d\left(\frac{D_t^f}{Q_t}\right) \\ &= \frac{M_t D_t^f}{Q_t}[-\sigma_2(t) d\widetilde{W}_3(t) - (R_t - \sigma_2^2(t))dt] + \frac{R_t M_t D_t^f}{Q_t} dt \\ &= -\frac{M_t D_t^f}{Q_t}(\sigma_2(t) d\widetilde{W}_3(t) - \sigma_2^2(t)dt) \\ &= -\frac{M_t D_t^f}{Q_t}\sigma_2(t) d\widetilde{W}_3^f(t). \end{aligned}$$

To get (9.3.23), we note

$$\begin{aligned} d\left(\frac{D_t^f S_t}{Q_t}\right) &= \frac{D_t^f}{Q_t} dS_t + S_t d\left(\frac{D_t^f}{Q_t}\right) + dS_t d\left(\frac{D_t^f}{Q_t}\right) \\ &= \frac{D_t^f}{Q_t} S_t(R_t dt + \sigma_1(t) d\widetilde{W}_1(t)) + \frac{S_t D_t^f}{Q_t}[-\sigma_2(t) d\widetilde{W}_3(t) - (R_t - \sigma_2^2(t))dt] \\ &\quad + S_t \sigma_1(t) d\widetilde{W}_1(t) \frac{D_t^f}{Q_t}(-\sigma_2(t)) d\widetilde{W}_3(t) \\ &= \frac{D_t^f S_t}{Q_t}[\sigma_1(t) d\widetilde{W}_1(t) - \sigma_2(t) d\widetilde{W}_3(t) + \sigma_2^2(t)dt - \sigma_1(t)\sigma_2(t)\rho_t dt] \\ &= \frac{D_t^f S_t}{Q_t}[\sigma_1(t) d\widetilde{W}_1^f(t) - \sigma_2 d\widetilde{W}_3^f(t)]. \end{aligned}$$

□

9.5.

Proof. We combine the solutions of all the sub-problems into a single solution as follows. The payoff of a quanto call is $(\frac{S_T}{Q_T}-K)^+$ units of domestic currency at time T. By risk-neutral pricing formula, its price at time t is $\widetilde{E}[e^{-r(T-t)}(\frac{S_T}{Q_T}-K)^+|\mathcal{F}_t]$. So we need to find the SDE for $\frac{S_t}{Q_t}$ under risk-neutral measure $\widetilde{P}$. By formula (9.3.14) and (9.3.16), we have $S_t = S_0 e^{\sigma_1 \widetilde{W}_1(t)+(r-\frac{1}{2}\sigma_1^2)t}$ and

$$Q_t = Q_0 e^{\sigma_2 \widetilde{W}_3(t)+(r-r^f-\frac{1}{2}\sigma_2^2)t} = Q_0 e^{\sigma_2\rho\widetilde{W}_1(t)+\sigma_2\sqrt{1-\rho^2}\widetilde{W}_2(t)+(r-r^f-\frac{1}{2}\sigma_2^2)t}.$$

So $\frac{S_t}{Q_t} = \frac{S_0}{Q_0} e^{(\sigma_1-\sigma_2\rho)\widetilde{W}_1(t)-\sigma_2\sqrt{1-\rho^2}\widetilde{W}_2(t)+(r^f+\frac{1}{2}\sigma_2^2-\frac{1}{2}\sigma_1^2)t}$. Define

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

Then $\widetilde{W}_4$ is a martingale with $[\widetilde{W}_4]_t = \frac{(\sigma_1-\sigma_2\rho)^2}{\sigma_4^2}t + \frac{\sigma_2(1-\rho^2)}{\sigma_4^2}t + t$. So $\widetilde{W}_4$ is a Brownian motion under $\widetilde{P}$. So if we set $a = r - r^f + \rho\sigma_1\sigma_2 - \sigma_2^2$, we have

$$\frac{S_t}{Q_t} = \frac{S_0}{Q_0} e^{\sigma_4\widetilde{W}_4(t)+(r-a-\frac{1}{2}\sigma_4^2)t} \text{ and } d\left(\frac{S_t}{Q_t}\right) = \frac{S_t}{Q_t}[\sigma_4 d\widetilde{W}_4(t)+(r-a)dt].$$

Therefore, under $\widetilde{P}$, $\frac{S_t}{Q_t}$ behaves like dividend-paying stock and the price of the quanto call option is like the price of a call option on a dividend-paying stock. Thus formula (5.5.12) gives us the desired price formula for quanto call option. □

9.6. (i)

Proof. $d_+(t) - d_-(t) = \frac{1}{\sigma\sqrt{T-t}}\sigma^2(T-t) = \sigma\sqrt{T-t}$. So $d_-(t) = d_+(t) - \sigma\sqrt{T-t}$. □

(ii)

Proof. $d_+(t) + d_-(t) = \frac{2}{\sigma\sqrt{T-t}} \log \frac{\text{For}_S(t,T)}{K}$. So

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

□

(iii)

Proof.

$$\begin{aligned}
\text{For}_S(t,T)e^{-d_+^2(t)/2} - Ke^{-d_-^2(t)} &= e^{-d_+^2(t)/2}[\text{For}_S(t,T) - Ke^{d_+^2(t)/2-d_-^2(t)/2}] \\
&= e^{-d_+^2(t)/2}[\text{For}_S(t,T) - Ke^{\log\frac{\text{For}_S(t,T)}{K}}] \\
&= 0.
\end{aligned}$$

□

(iv)

Proof.

$$
\begin{aligned}
& dd_+(t) \\
= \;& \frac{1}{2}\sqrt{1}\sigma\sqrt{(T-t)^3}[\log\frac{\mathrm{For}_S(t,T)}{K}+\frac{1}{2}\sigma^2(T-t)]dt+\frac{1}{\sigma\sqrt{T-t}}\left[\frac{d\mathrm{For}_S(t,T)}{\mathrm{For}_S(t,T)}-\frac{(d\mathrm{For}_S(t,T))^2}{2\mathrm{For}_S(t,T)^2}-\frac{1}{2}\sigma dt\right] \\
= \;& \frac{1}{2\sigma\sqrt{(T-t)^3}}\log\frac{\mathrm{For}_S(t,T)}{K}dt+\frac{\sigma}{4\sqrt{T-t}}dt+\frac{1}{\sigma\sqrt{T-t}}(\sigma d\widetilde{W}^T(t)-\frac{1}{2}\sigma^2 dt-\frac{1}{2}\sigma^2 dt) \\
= \;& \frac{1}{2\sigma(T-t)^{3/2}}\log\frac{\mathrm{For}_S(t,T)}{K}dt-\frac{3\sigma}{4\sqrt{T-t}}dt+\frac{d\widetilde{W}^T(t)}{\sqrt{T-t}}.
\end{aligned}
$$

□

(v)

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

(vi)

Proof. By (iv) and (v), $(dd_-(t))^2=(dd_+(t))^2=\frac{dt}{T-t}$. □

(vii)

Proof.

$$
\begin{aligned}
dN(d_+(t)) &= N'(d_+(t))dd_+(t)+\frac{1}{2}N''(d_+(t))(dd_+(t))^2 \\
&= \frac{1}{\sqrt{2\pi}}e^{-\frac{d_+^2(t)}{2}}dd_+(t)+\frac{1}{2}\frac{1}{\sqrt{2\pi}}e^{-\frac{d_+^2(t)}{2}}(-d_+(t))\frac{dt}{T-t}.
\end{aligned}
$$

□

(viii)

Proof.

$$
\begin{aligned}
dN(d_-(t)) &= N'(d_-(t))dd_-(t)+\frac{1}{2}N''(d_-(t))(dd_-(t))^2 \\
&= \frac{1}{\sqrt{2\pi}}e^{-\frac{d_-^2(t)}{2}}\left(dd_+(t)+\frac{\sigma dt}{2\sqrt{T-t}}\right)+\frac{1}{2}\frac{e^{-\frac{d_-^2(t)}{2}}}{\sqrt{2\pi}}(-d_-(t))\frac{dt}{T-t} \\
&= \frac{1}{\sqrt{2\pi}}e^{-d_-^2(t)/2}dd_+(t)+\frac{\sigma e^{-d_-^2(t)/2}}{2\sqrt{2\pi(T-t)}}dt+\frac{e^{-\frac{d_-^2(t)(\sigma\sqrt{T-t}-d_+(t))}{2}}}{2(T-t)\sqrt{2\pi}}dt \\
&= \frac{1}{\sqrt{2\pi}}e^{-d_-^2(t)/2}dd_+(t)+\frac{\sigma e^{-d_-^2(t)/2}}{\sqrt{2\pi(T-t)}}dt-\frac{d_+(t)e^{-\frac{d_-^2(t)}{2}}}{2(T-t)\sqrt{2\pi}}dt.
\end{aligned}
$$

□

(ix)

Proof.

$$
d\mathrm{For}_S(t,T)dN(d_+(t))=\sigma\mathrm{For}_S(t,T)d\widetilde{W}^T(t)\frac{e^{-d_+^2(t)/2}}{\sqrt{2\pi}}\frac{1}{\sqrt{T-t}}d\widehat{W}^T(t)=\frac{\sigma\mathrm{For}_S(t,T)e^{-d_+^2(t)/2}}{\sqrt{2\pi(T-t)}}dt.
$$

□

(x)

Proof.

$$\begin{aligned}
& \mathrm{For}_S(t,T)dN(d_+(t)) + d\mathrm{For}_S(t,T)dN(d_+(t)) - KdN(d_-(t)) \\
= \; & \mathrm{For}_S(t,T)\left[\frac{1}{\sqrt{2\pi}}e^{-d_+^2(t)/2}dd_+(t) - \frac{d_+(t)}{2(T-t)\sqrt{2\pi}}e^{-d_+^2(t)/2}dt\right] + \frac{\sigma\mathrm{For}_S(t,T)e^{-d_+^2(t)/2}}{\sqrt{2\pi(T-t)}}dt \\
& -K\left[\frac{e^{-d_-^2(t)/2}}{\sqrt{2\pi}}dd_+(t) + \frac{\sigma}{\sqrt{2\pi(T-t)}}e^{-d_-^2(t)/2}dt - \frac{d_+(t)}{2(T-t)\sqrt{2\pi}}e^{-d_-^2(t)/2}dt\right] \\
= \; & \left[\frac{\mathrm{For}_S(t,T)d_+(t)}{2(T-t)\sqrt{2\pi}}e^{-d_+^2(t)/2} + \frac{\sigma\mathrm{For}_S(t,T)e^{-d_+^2(t)/2}}{\sqrt{2\pi(T-t)}} - \frac{K\sigma e^{-d_-^2(t)/2}}{\sqrt{2\pi(T-t)}} - \frac{Kd_+(t)}{2(T-t)\sqrt{2\pi}}e^{-d_-^2(t)/2}\right]dt \\
& +\frac{1}{\sqrt{2\pi}}\left(\mathrm{For}_S(t,T)e^{-d_+^2(t)/2} - Ke^{-d_-^2(t)/2}\right)dd_+(t) \\
= \; & 0.
\end{aligned}$$

The last "=" comes from (iii), which implies $e^{-d_-^2(t)/2} = \frac{\mathrm{For}_S(t,T)}{K}e^{-d_+^2(t)/2}$. □

10 Term-Structure Models

★ **Comments**:

1) *Computation of $e^{\Lambda t}$ (Lemma 10.2.3).* For a systematic treatment of the computation of matrix exponential function $e^{\Lambda t}$, see 丁同仁等 [3, page 175] or Arnold et al. [1, page 221]. For the sake of Lemma 10.2.3, a direct computation to find (10.2.35) goes as follows.

i) The case of $\lambda_1 = \lambda_2$. In this case, set $\Gamma_1 = \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$ and $\Gamma_2 = \begin{bmatrix}0 & 0\\ \lambda_{21} & 0\end{bmatrix}$. Then $\Lambda = \Gamma_1 + \Gamma_2$, $\Gamma_2^n = 0_{2\times 2}$ for $n \geq 2$, and $\Gamma_1^n = \begin{bmatrix}\lambda_1^n & 0\\ 0 & \lambda_2^n\end{bmatrix}$. Note Γ_1 and Γ_2 commute, we therefore have

$$e^{\Lambda t} = e^{\Gamma_1 t}\cdot e^{\Gamma_2 t} = \begin{bmatrix}e^{\lambda_1 t} & 0\\ 0 & e^{\lambda_2 t}\end{bmatrix}\left(I_{2\times 2} + \begin{bmatrix}0 & 0\\ \lambda_{21}t & 0\end{bmatrix}\right) = \begin{bmatrix}e^{\lambda_1 t} & 0\\ \lambda_{21}te^{\lambda_2 t} & e^{\lambda_2 t}\end{bmatrix} = \begin{bmatrix}e^{\lambda_1 t} & 0\\ \lambda_{21}te^{\lambda_1 t} & e^{\lambda_2 t}\end{bmatrix}.$$

ii) The case of $\lambda_1 \neq \lambda_2$. In this case, Λ has two distinct eigenvalues λ_1 and λ_2. So it can be diagonalized, that is, we can find a matrix P such that

$$P^{-1}\Lambda P = J = \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}.$$

Solving the matrix equation $P\Lambda = PJ$, we have

$$P = \begin{bmatrix}1 & 0\\ \frac{\lambda_{21}}{\lambda_1-\lambda_2} & 1\end{bmatrix},$$

and consequently

$$P^{-1} = \begin{bmatrix}1 & 0\\ -\frac{\lambda_{21}}{\lambda_1-\lambda_2} & 1\end{bmatrix}.$$

Therefore

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

2) Intuition of Theorem 10.4.1 (Price of backset LIBOR). The intuition here is the same as that of a discrete-time model: the payoff $\delta L(T,T) = (1+\delta L(T,T)) - 1$ at time $T+\delta$ is equivalent to an investment of 1 at time T (which gives the payoff $1+\delta L(T,T)$ at time $T+\delta$), subtracting a payoff of 1 at time $T+\delta$. The first part has time t price $B(t,T)$ while the second part has time t price $B(t,T+\delta)$.

Pricing under $(T+\delta)$*-forward measure*: Using the $(T+\delta)$-forward measure, the time-t no-arbitrage price of a payoff $\delta L(T,T)$ at time $(T+\delta)$ is

$$B(t,T+\delta)\widetilde{\mathbb{E}}^{T+\delta}[\delta L(T,T)] = B(t,T+\delta)\cdot \delta L(t,T),$$

where we have used the observation that the forward rate process $L(t,T) = \frac{1}{\delta}\left(\frac{B(t,T)}{B(t,T+\delta)} - 1\right)$ $(0\le t\le T)$ is a martingale under the $(T+\delta)$-forward measure.

► **Exercise 10.1 (Statistics in the two-factor Vasicek model).** According to Example 4.7.3, $Y_1(t)$ and $Y_2(t)$ in (10.2.43)-(10.2.46) are Gaussian processes.

(i) Show that

$$\widetilde{\mathbb{E}}Y_1(t) = e^{-\lambda_1 t}Y_1(0), \tag{10.7.1}$$

that when $\lambda_1 \neq \lambda_2$, then

$$\widetilde{\mathbb{E}}Y_2(t) = \frac{\lambda_{21}}{\lambda_1-\lambda_2}\left(e^{-\lambda_1 t} - e^{-\lambda_2 t}\right)Y_1(0) + e^{-\lambda_2 t}Y_2(0), \tag{10.7.2}$$

and when $\lambda_1 = \lambda_2$, then

$$\widetilde{\mathbb{E}}Y_2(t) = -\lambda_{21}te^{-\lambda_1 t}Y_1(0) + e^{-\lambda_1 t}Y_2(0). \tag{10.7.3}$$

We can write

$$Y_1(t) - \widetilde{\mathbb{E}}Y_1(t) = e^{-\lambda_1 t}I_1(t),$$

when $\lambda_1 \neq \lambda_2$,

$$Y_2(t) - \widetilde{\mathbb{E}}Y_2(t) = \frac{\lambda_{21}}{\lambda_1-\lambda_2}\left(e^{-\lambda_1 t}I_1(t) - e^{-\lambda_2 t}I_2(t)\right) - e^{-\lambda_2 t}I_3(t),$$

and when $\lambda_1 = \lambda_2$,

$$Y_2(t) - \widetilde{\mathbb{E}}Y_2(t) = -\lambda_{21}te^{-\lambda_1 t}I_1(t) + \lambda_{21}e^{-\lambda_1 t}I_4(t) + e^{-\lambda_1 t}I_3(t),$$

where the Itô integrals

$$I_1(t) = \int_0^t e^{\lambda_1 u}d\widetilde{W}_1(u), \qquad I_2(t) = \int_0^t e^{\lambda_2 u}d\widetilde{W}_1(u),$$
$$I_3(t) = \int_0^t e^{\lambda_2 u}d\widetilde{W}_2(u), \qquad I_4(t) = \int_0^t ue^{\lambda_1 u}d\widetilde{W}_1(u),$$

all have expectation zero under the risk-neutral measure $\widetilde{\mathbb{P}}$. Consequently, we can determine the variances of $Y_1(t)$ and $Y_2(t)$ and the covariance of $Y_1(t)$ and $Y_2(t)$ under the risk-neutral measure from the variances and covariances of $I_j(t)$ and $I_k(t)$. For example, if $\lambda_1 = \lambda_2$, then

$$\begin{aligned}
\mathrm{Var}(Y_1(t)) &= e^{-2\lambda_1 t}\widetilde{\mathbb{E}}I_1^2(t),\\
\mathrm{Var}(Y_2(t)) &= \lambda_{21}^2t^2e^{-2\lambda_1 t}\widetilde{\mathbb{E}}I_1^2(t) + \lambda_{21}^2e^{-2\lambda_1 t}\widetilde{\mathbb{E}}I_4^2(t) + e^{-2\lambda_1 t}\widetilde{\mathbb{E}}I_3^2(t)\\
&\quad -2\lambda_{21}^2te^{-2\lambda_1 t}\widetilde{\mathbb{E}}[I_1(t)I_4(t)] - 2\lambda_{21}te^{-2\lambda_1 t}\widetilde{\mathbb{E}}[I_1(t)I_3(t)]\\
&\quad +2\lambda_{21}e^{-2\lambda_1 t}\widetilde{\mathbb{E}}[I_4(t)I_3(t)],\\
\mathrm{Cov}(Y_1(t),Y_2(t)) &= -\lambda_{21}te^{-2\lambda_1 t}\widetilde{\mathbb{E}}I_1^2(t) + \lambda_{21}e^{-2\lambda_1 t}\widetilde{\mathbb{E}}[I_1(t)I_4(t)] + e^{-2\lambda_1 t}\widetilde{\mathbb{E}}[I_1(t)I_3(t)],
\end{aligned}$$

where the variances and covariance above are under the risk-neutral measure $\widetilde{\mathbb{P}}$.

Proof. Using the notation $I_1(t)$, $I_2(t)$, $I_3(t)$ and $I_4(t)$ introduced in the problem, we can write $Y_1(t)$ and $Y_2(t)$ as

$$Y_1(t) = e^{-\lambda_1 t}Y_1(0) + e^{-\lambda_1 t}I_1(t)$$

and

$$\begin{aligned} & Y_2(t) \\ = & \begin{cases} \frac{\lambda_{21}}{\lambda_1-\lambda_2}(e^{-\lambda_1 t} - e^{-\lambda_2 t})Y_1(0) + e^{-\lambda_2 t}Y_2(0) + \frac{\lambda_{21}}{\lambda_1-\lambda_2}\left[e^{-\lambda_1 t}I_1(t) - e^{-\lambda_2 t}I_2(t)\right] - e^{-\lambda_2 t}I_3(t), & \lambda_1 \neq \lambda_2; \\ -\lambda_{21}te^{-\lambda_1 t}Y_1(0) + e^{-\lambda_1 t}Y_2(0) - \lambda_{21}\left[te^{-\lambda_1 t}I_1(t) - e^{-\lambda_1 t}I_4(t)\right] + e^{-\lambda_1 t}I_3(t), & \lambda_1 = \lambda_2. \end{cases} \end{aligned}$$

Since all the $I_k(t)$'s ($k = 1, \cdots, 4$) are normally distributed with zero mean, we can conclude

$$\widetilde{\mathbb{E}}[Y_1(t)] = e^{-\lambda_1 t}Y_1(0)$$

and

$$\widetilde{\mathbb{E}}[Y_2(t)] = \begin{cases} \frac{\lambda_{21}}{\lambda_1-\lambda_2}(e^{-\lambda_1 t} - e^{-\lambda_2 t})Y_1(0) + e^{-\lambda_2 t}Y_2(0), & \text{if } \lambda_1 \neq \lambda_2; \\ -\lambda_{21}te^{-\lambda_1 t}Y_1(0) + e^{-\lambda_1 t}Y_2(0), & \text{if } \lambda_1 = \lambda_2. \end{cases}$$

□

(ii) Compute the five terms

$$\widetilde{\mathbb{E}}I_1^2(t),\ \widetilde{\mathbb{E}}[I_1(t)I_2(t)],\ \widetilde{\mathbb{E}}[I_1(t)I_3(t)],\ \widetilde{\mathbb{E}}[I_1(t)I_4(t)],\ \widetilde{\mathbb{E}}[I_4^2(t)].$$

The five other terms, which you are not being asked to compute, are

$$\begin{aligned} \mathbb{E}I_2^2(t) &= \frac{1}{2\lambda_2}\left(e^{2\lambda_2 t} - 1\right), \\ \mathbb{E}[I_2(t)I_3(t)] &= 0, \\ \mathbb{E}[I_2(t)I_4(t)] &= \frac{t}{\lambda_1+\lambda_2}e^{(\lambda_1+\lambda_2)t} + \frac{1}{(\lambda_1+\lambda_2)^2}\left(1 - e^{(\lambda_1+\lambda_2)t}\right), \\ \mathbb{E}I_3^2(t) &= \frac{1}{\lambda_2}\left(e^{2\lambda_2 t} - 1\right), \\ \mathbb{E}[I_3(t)I_4(t)] &= 0. \end{aligned}$$

Solution. The calculation relies on the following fact: if X_t and Y_t are both martingales, then $X_tY_t - [X,Y]_t$ is also a martingale. In particular, $\widetilde{\mathbb{E}}[X_tY_t] = \widetilde{\mathbb{E}}\{[X,Y]_t\}$. Thus

$$\widetilde{\mathbb{E}}[I_1^2(t)] = \int_0^t e^{2\lambda_1 u}du = \frac{e^{2\lambda_1 t} - 1}{2\lambda_1},\ \widetilde{\mathbb{E}}[I_1(t)I_2(t)] = \int_0^t e^{(\lambda_1+\lambda_2)u}du = \frac{e^{(\lambda_1+\lambda_2)t} - 1}{\lambda_1+\lambda_2},$$

$$\widetilde{\mathbb{E}}[I_1(t)I_3(t)] = 0,\ \widetilde{\mathbb{E}}[I_1(t)I_4(t)] = \int_0^t ue^{2\lambda_1 u}du = \frac{1}{2\lambda_1}\left[te^{2\lambda_1 t} - \frac{e^{2\lambda_1 t} - 1}{2\lambda_1}\right]$$

and

$$\widetilde{\mathbb{E}}[I_4^2(t)] = \int_0^t u^2e^{2\lambda_1 u}du = \frac{t^2e^{2\lambda_1 t}}{2\lambda_1} - \frac{te^{2\lambda_1 t}}{2\lambda_1^2} + \frac{e^{2\lambda_1 t} - 1}{4\lambda_1^3}.$$

□

(iii) Some derivative securities involve *time spread* (i.e., they depend on the interest rate at two different times). In such cases, we are interested in the joint statistics of the factor processes at different times. These are still jointly normal and depend on the statistics of the Itô integral I_j at different times. Compute $\widetilde{\mathbb{E}}[I_1(s)I_2(t)]$, where $0 \leq s < t$. (Hint: Fix $s \geq 0$ and define

$$J_1(t) = \int_0^t e^{\lambda_1 u}1_{\{u \leq s\}}d\widetilde{W}_1(u),$$

where $1_{\{u\leq s\}}$ is the function of u that is 1 if $u \leq s$ and 0 if $u > s$. Note that $J_1(t) = I_1(s)$ when $t \geq s$.)

Solution. Following the hint, we have

$$\widetilde{\mathbb{E}}[I_1(s)I_2(t)] = \widetilde{\mathbb{E}}[J_1(t)I_2(t)] = \int_0^t e^{(\lambda_1+\lambda_2)u} 1_{\{u\leq s\}} du = \frac{e^{(\lambda_1+\lambda_2)s} - 1}{\lambda_1 + \lambda_2}.$$

□

▶ **Exercise 10.2 (Ordinary differential equations for the mixed affine-yield model).** In the mixed model of Subsection 10.2.3, as in the two-factor Cox-Ingersoll-Ross model, zero-coupon bond prices have the affine-yield form

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

where $C_1(0) = C_2(0) = A(0) = 0$.

(i) Find the partial differential equation satisfied by $f(t, y_1, y_2)$.

Solution. Assume $B(t,T) = \widetilde{\mathbb{E}}\left[e^{-\int_t^T R_s ds}\middle|\mathcal{F}_t\right] = f(t, Y_1(t), Y_2(t))$. Then

$$d(D(t)B(t,T)) = D(t)[-R(t)f(t, Y_1(t), Y_2(t))dt + df(t, Y_1(t), Y_2(t))].$$

By Itô's formula,

$$\begin{aligned} df(t, Y_1(t), Y_2(t)) \quad &= \quad [f_t(t, Y_1(t), Y_2(t)) + f_{y_1}(t, Y_1(t), Y_2(t))(\mu - \lambda_1 Y_1(t)) + f_{y_2}(t, Y_1(t), Y_2(t))(-\lambda_2)Y_2(t) \\ &\quad + f_{y_1 y_2}(t, Y_1(t), Y_2(t))\sigma_{21} Y_1(t) + \frac{1}{2} f_{y_1 y_1}(t, Y_1(t), Y_2(t)) Y_1(t) \\ &\quad \left. + \frac{1}{2} f_{y_2 y_2}(t, Y_1(t), Y_2(t))(\sigma_{21}^2 Y_1(t) + \alpha + \beta Y_1(t))\right] dt + \text{martingale part}. \end{aligned}$$

Since $D(t)B(t,T)$ is a martingale, we must have

$$\begin{aligned} &\left[-(\delta_0 + \delta_1 y_1 + \delta_2 y_2) + \frac{\partial}{\partial t} + (\mu - \lambda_1 y_1)\frac{\partial}{\partial y_1} - \lambda_2 y_2 \frac{\partial}{\partial y_2}\right] f \\ &+ \left[\frac{1}{2}\left(2\sigma_{21} y_1 \frac{\partial^2}{\partial y_1 \partial y_2} + y_1 \frac{\partial^2}{\partial y_1^2} + (\sigma_{21}^2 y_1 + \alpha + \beta y_1)\frac{\partial^2}{\partial y_2^2}\right)\right] f \\ = \quad & 0. \end{aligned}$$

□

(ii) Show that C_1, C_2, and A satisfy the system of ordinary differential equations[12]

$$C_1' = -\lambda_1 C_1 - \frac{1}{2}C_1^2 - \sigma_{21} C_1 C_2 - \frac{1}{2}(\sigma_{21}^2 + \beta)C_2^2 + \delta_1, \tag{10.7.4}$$

$$C_2' = -\lambda_2 C_2 + \delta_2, \tag{10.7.5}$$

$$A' = \mu C_1 - \frac{1}{2}\alpha C_2^2 + \delta_0. \tag{10.7.6}$$

[12] The textbook has a typo in formula (10.7.4): the coefficient of C_2^2 should be $-\frac{1}{2}(\sigma_{21}^2 + \beta)$ instead of $-(1+\beta)$. See http://www.math.cmu.edu/~shreve/ for details.

Proof. If we suppose $f(t, y_1, y_2) = e^{-y_1 C_1(T-t) - y_2 C_2(T-t) - A(T-t)}$, then

$$\begin{aligned}
\frac{\partial f}{\partial t} &= [y_1 C_1'(T-t) + y_2 C_2'(T-t) + A'(T-t)]f \\
\frac{\partial f}{\partial y_1} &= -C_1(T-t)f \\
\frac{\partial f}{\partial y_2} &= -C_2(T-t)f \\
\frac{\partial^2 f}{\partial y_1 \partial y_2} &= C_1(T-t)C_2(T-t)f \\
\frac{\partial^2 f}{\partial y_1^2} &= C_1^2(T-t)f \\
\frac{\partial^2 f}{\partial y_2^2} &= C_2^2(T-t)f.
\end{aligned}$$

So the PDE in part (i) becomes

$$\begin{aligned}
& -(\delta_0 + \delta_1 y_1 + \delta_2 y_2) + y_1 C_1' + y_2 C_2' + A' - (\mu - \lambda_1 y_1)C_1 + \lambda_2 y_2 C_2 \\
& + \frac{1}{2}\left[2\sigma_{21} y_1 C_1 C_2 + y_1 C_1^2 + (\sigma_{21}^2 y_1 + \alpha + \beta y_1)C_2^2\right] \\
= \;& 0.
\end{aligned}$$

Sorting out the LHS according to the independent variables y_1 and y_2, we get

$$\begin{cases}
-\delta_1 + C_1' + \lambda_1 C_1 + \sigma_{21} C_1 C_2 + \frac{1}{2}C_1^2 + \frac{1}{2}(\sigma_{21}^2 + \beta)C_2^2 = 0 \\
-\delta_2 + C_2' + \lambda_2 C_2 = 0 \\
-\delta_0 + A' - \mu C_1 + \frac{1}{2}\alpha C_2^2 = 0.
\end{cases}$$

In other words, we can obtain the ODEs for C_1, C_2 and A as follows

$$\begin{cases}
C_1' = -\lambda_1 C_1 - \sigma_{21} C_1 C_2 - \frac{1}{2}C_1^2 - \frac{1}{2}(\sigma_{21}^2 + \beta)C_2^2 + \delta_1 \\
C_2' = -\lambda_2 C_2 + \delta_2 \\
A' = \mu C_1 - \frac{1}{2}\alpha C_2^2 + \delta_0.
\end{cases}$$

□

▶ **Exericse 10.3 (Calibration of the two-factor Vasicek model).** Consider the canonical two-factor Vasicek model (10.2.4), (10.2.5), but replace the interest rate equation (10.2.6) by

$$R(t) = \delta_0(t) + \delta_1 Y_1(t) + \delta_2 Y_2(t), \tag{10.7.7}$$

where δ_1 and δ_2 are constant but $\delta_0(t)$ is a nonrandom function of time. Assume that for each T there is a zero-coupon bond maturing at time T. The price of this bond at time $t \in [0, T]$ is

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

Because the pair of processes $(Y_1(t), Y_2(t))$ is Markov, there must exist some function $f(t, T, y_1, y_2)$ such that $B(t, T) = f(t, T, Y_1(t), Y_2(t))$. (We indicate the dependence of f on the maturity T because, unlike in Subsection 10.2.1, here we shall consider more than one value of T.)

(i) The function $f(t, T, y_1, y_2)$ is of the affine-yield form

$$f(t, T, y_1, y_2) = e^{-y_1 C_1(t,T) - y_2 C_2(t,T) - A(t,T)}. \tag{10.7.8}$$

Holding T fixed, derive a system of ordinary differential equations for $\frac{d}{dt}C_1(t, T)$, $\frac{d}{dt}C_2(t, T)$, and $\frac{d}{dt}A(t, T)$.

Solution. We have $d(D_tB(t,T)) = D_t[-R_tf(t,T,Y_1(t),Y_2(t))dt + df(t,T,Y_1(t),Y_2(t))]$ and

$$\begin{aligned}
& df(t,T,Y_1(t),Y_2(t)) \\
= \ & [f_t(t,T,Y_1(t),Y_2(t)) + f_{y_1}(t,T,Y_1(t),Y_2(t))(-\lambda_1Y_1(t)) + f_{y_2}(t,T,Y_1(t),Y_2(t))(-\lambda_{21}Y_1(t) - \lambda_2Y_2(t)) \\
& + \frac{1}{2}f_{y_1y_1}(t,T,Y_1(t),Y_2(t)) + \frac{1}{2}f_{y_2y_2}(t,T,Y_1(t),Y_2(t))]dt + \text{martingale part.}
\end{aligned}$$

Since $D_tB(t,T)$ is a martingale under risk-neutral measure, we have the following PDE:

$$\left[-(\delta_0(t) + \delta_1y_1 + \delta_2y_2) + \frac{\partial}{\partial t} - \lambda_1y_1\frac{\partial}{\partial y_1} - (\lambda_{21}y_1 + \lambda_2y_2)\frac{\partial}{\partial y_2} + \frac{1}{2}\frac{\partial^2}{\partial y_1^2} + \frac{1}{2}\frac{\partial}{\partial y_2^2}\right] f(t,T,y_1,y_2) = 0.$$

Suppose $f(t,T,y_1,y_2) = e^{-y_1C_1(t,T) - y_2C_2(t,T) - A(t,T)}$, then

$$\begin{cases}
f_t(t,T,y_1,y_2) = \left[-y_1\frac{d}{dt}C_1(t,T) - y_2\frac{d}{dt}C_2(t,T) - \frac{d}{dt}A(t,T)\right] f(t,T,y_1,y_2), \\
f_{y_1}(t,T,y_1,y_2) = -C_1(t,T)f(t,T,y_1,y_2), \\
f_{y_2}(t,T,y_1,y_2) = -C_2(t,T)f(t,T,y_1,y_2), \\
f_{y_1y_2}(t,T,y_1,y_2) = C_1(t,T)C_2(t,T)f(t,T,y_1,y_2), \\
f_{y_1y_1}(t,T,y_1,y_2) = C_1^2(t,T)f(t,T,y_1,y_2), \\
f_{y_2y_2}(t,T,y_1,y_2) = C_2^2(t,T)f(t,T,y_1,y_2).
\end{cases}$$

So the PDE becomes

$$\begin{aligned}
& -(\delta_0(t) + \delta_1y_1 + \delta_2y_2) + \left(-y_1\frac{d}{dt}C_1(t,T) - y_2\frac{d}{dt}C_2(t,T) - \frac{d}{dt}A(t,T)\right) + \lambda_1y_1C_1(t,T) \\
& + (\lambda_{21}y_1 + \lambda_2y_2)C_2(t,T) + \frac{1}{2}C_1^2(t,T) + \frac{1}{2}C_2^2(t,T) = 0.
\end{aligned}$$

Sorting out the terms according to independent variables y_1 and y_2, we get

$$\begin{cases}
-\delta_0(t) - \frac{d}{dt}A(t,T) + \frac{1}{2}C_1^2(t,T) + \frac{1}{2}C_2^2(t,T) = 0 \\
-\delta_1 - \frac{d}{dt}C_1(t,T) + \lambda_1C_1(t,T) + \lambda_{21}C_2(t,T) = 0 \\
-\delta_2 - \frac{d}{dt}C_2(t,T) + \lambda_2C_2(t,T) = 0.
\end{cases}$$

That is

$$\begin{cases}
\frac{d}{dt}C_1(t,T) = \lambda_1C_1(t,T) + \lambda_{21}C_2(t,T) - \delta_1 \\
\frac{d}{dt}C_2(t,T) = \lambda_2C_2(t,T) - \delta_2 \\
\frac{d}{dt}A(t,T) = \frac{1}{2}C_1^2(t,T) + \frac{1}{2}C_2^2(t,T) - \delta_0(t).
\end{cases}$$

□

(ii) Using the terminal conditions $C_1(T,T) = C_2(T,T) = 0$, solve the equations in (i) for $C_1(t,T)$ and $C_2(t,T)$. (As in Subsection 10.2.1, the functions C_1 and C_2 depend on t and T only through the difference $\tau = T - t$; however, the function A discussed in part (iii) below depends on t and T separately.)

Solution. For C_2, we note $\frac{d}{dt}[e^{-\lambda_2t}C_2(t,T)] = -e^{-\lambda_2t}\delta_2$ from the ODE in (i). Integrate from t to T, we have $0 - e^{-\lambda_2t}C_2(t,T) = -\delta_2\int_t^T e^{-\lambda_2s}ds = \frac{\delta_2}{\lambda_2}(e^{-\lambda_2T} - e^{-\lambda_2t})$. So $C_2(t,T) = \frac{\delta_2}{\lambda_2}\left(1 - e^{-\lambda_2(T-t)}\right)$. For C_1, we note

$$\frac{d}{dt}(e^{-\lambda_1t}C_1(t,T)) = (\lambda_{21}C_2(t,T) - \delta_1)e^{-\lambda_1t} = \frac{\lambda_{21}\delta_2}{\lambda_2}(e^{-\lambda_1t} - e^{-\lambda_2T + (\lambda_2-\lambda_1)t}) - \delta_1e^{-\lambda_1t}.$$

Integrate from t to T, we get

$$\begin{aligned}
& -e^{-\lambda_1t}C_1(t,T) \\
= \ & \begin{cases}
-\frac{\lambda_{21}\delta_2}{\lambda_2\lambda_1}(e^{-\lambda_1T} - e^{-\lambda_1t}) - \frac{\lambda_{21}\delta_2}{\lambda_2}e^{-\lambda_2T}\frac{e^{(\lambda_2-\lambda_1)T} - e^{(\lambda_2-\lambda_1)t}}{\lambda_2-\lambda_1} + \frac{\delta_1}{\lambda_1}(e^{-\lambda_1T} - e^{-\lambda_1T}) & \text{if } \lambda_1 \neq \lambda_2 \\
-\frac{\lambda_{21}\delta_2}{\lambda_2\lambda_1}(e^{-\lambda_1T} - e^{-\lambda_1t}) - \frac{\lambda_{21}\delta_2}{\lambda_2}e^{-\lambda_2T}(T-t) + \frac{\delta_1}{\lambda_1}(e^{-\lambda_1T} - e^{-\lambda_1T}) & \text{if } \lambda_1 = \lambda_2.
\end{cases}
\end{aligned}$$

So

$$C_1(t,T) = \begin{cases} \frac{\lambda_{21}\delta_2}{\lambda_2\lambda_1}(e^{-\lambda_1(T-t)}-1) + \frac{\lambda_{21}\delta_2}{\lambda_2}\frac{e^{-\lambda_1(T-t)}-e^{-\lambda_2(T-t)}}{\lambda_2-\lambda_1} - \frac{\delta_1}{\lambda_1}(e^{-\lambda_1(T-t)}-1) & \text{if } \lambda_1 \neq \lambda_2 \\ \frac{\lambda_{21}\delta_2}{\lambda_2\lambda_1}(e^{-\lambda_1(T-t)}-1) + \frac{\lambda_{21}\delta_2}{\lambda_2}e^{-\lambda_2 T+\lambda_1 t}(T-t) - \frac{\delta_1}{\lambda_1}(e^{-\lambda_1(T-t)}-1) & \text{if } \lambda_1 = \lambda_2. \end{cases}.$$

□

(iii) Using the terminal condition $A(T,T)=0$, write a formula for $A(t,T)$ as an integral involving $C_1(u,T)$, $C_2(u,T)$, and $\delta_0(u)$. You do not need to evaluate this integral.

Solution. From the ODE $\frac{d}{dt}A(t,T) = \frac{1}{2}(C_1^2(t,T)+C_2^2(t,T)) - \delta_0(t)$, we get

$$A(t,T) = \int_t^T \left[\delta_0(s) - \frac{1}{2}(C_1^2(s,T)+C_2^2(s,T))\right] ds.$$

□

(iv) Assume that the model parameters $\lambda_1 > 0$ $\lambda_2 > 0$, λ_{21}, δ_1, and δ_2 and the initial conditions $Y_1(0)$ and $Y_2(0)$ are given. We wish to choose a *function* δ_0 so that the zero-coupon bond prices given by the model match the bond prices given by the market at the initial time zero. In other words, we want to choose a function $\delta(T)$, $T \geq 0$, so that

$$f(0,T,Y_1(0),Y_2(0)) = B(0,T),\ T \geq 0.$$

In this part of the exercise, we regard both t and T as variables and uses the notation $\frac{\partial}{\partial t}$ to indicate the derivative with respect to t when T is held fixed and the notation $\frac{\partial}{\partial T}$ to indicate the derivative with respect to T when t is held fixed. Give a formula for $\delta_0(T)$ in terms of $\frac{\partial}{\partial T}\log B(0,T)$ and the model parameters. (Hint: Compute $\frac{\partial}{\partial T}A(0,T)$ in two ways, using (10.7.8) and also using the formula obtained in (iii). Because $C_i(t,T)$ depends only on t and T through $\tau = T-t$, there are functions $\overline{C}_i(\tau)$ such that $\overline{C}_i(\tau) = \overline{C}_i(T-t) = C_i(t,T)$, $i=1,2$. Then

$$\frac{\partial}{\partial t}C_i(t,T) = -\overline{C}_i'(\tau),\ \frac{\partial}{\partial T}C_i(t,T) = \overline{C}_i'(\tau),$$

where $'$ denotes differentiation with respect to τ. This shows that

$$\frac{\partial}{\partial T}C_i(t,T) = -\frac{\partial}{\partial t}C_i(t,T),\ i=1,2, \tag{10.7.9}$$

a fact that you will need.)

Solution. We want to find δ_0 so that $f(0,T,Y_1(0),Y_2(0)) = e^{-Y_1(0)C_1(0,T)-Y_2(0)C_2(0,T)-A(0,T)} = B(0,T)$ for all $T>0$. Take logarithm on both sides and plug in the expression of $A(t,T)$, we get

$$\log B(0,T) = -Y_1(0)C_1(0,T) - Y_2(0)C_2(0,T) + \int_0^T \left[\frac{1}{2}(C_1^2(s,T)+C_2^2(s,T)) - \delta_0(s)\right] ds.$$

Taking derivative w.r.t. T, we have

$$\frac{\partial}{\partial T}\log B(0,T) = -Y_1(0)\frac{\partial}{\partial T}C_1(0,T) - Y_2(0)\frac{\partial}{\partial T}C_2(0,T) + \frac{1}{2}C_1^2(T,T) + \frac{1}{2}C_2^2(T,T) - \delta_0(T).$$

Therefore

$$\begin{aligned} \delta_0(T) &= -Y_1(0)\frac{\partial}{\partial T}C_1(0,T) - Y_2(0)\frac{\partial}{\partial T}C_2(0,T) - \frac{\partial}{\partial T}\log B(0,T) \\ &= \begin{cases} -Y_1(0)\left[\delta_1 e^{-\lambda_1 T} - \frac{\lambda_{21}\delta_2}{\lambda_2}e^{-\lambda_2 T}\right] - Y_2(0)\delta_2 e^{-\lambda_2 T} - \frac{\partial}{\partial T}\log B(0,T) & \text{if } \lambda_1 \neq \lambda_2 \\ -Y_1(0)\left[\delta_1 e^{-\lambda_1 T} - \lambda_{21}\delta_2 e^{-\lambda_2 T}T\right] - Y_2(0)\delta_2 e^{-\lambda_2 T} - \frac{\partial}{\partial T}\log B(0,T) & \text{if } \lambda_1 = \lambda_2. \end{cases} \end{aligned}$$

□

▶ **Exercise 10.4.** Hull and White [89] propose the two-factor model

$$dU(t) = -\lambda_1 U(t)dt + \sigma_1 d\widetilde{B}_2(t), \tag{10.7.10}$$

$$dR(t) = [\theta(t) + U(t) - \lambda_2 R(t)]dt + \sigma_2 d\widetilde{B}_1(t), \tag{10.7.11}$$

where λ_1, λ_2, σ_1, and σ_2 are positive constants, $\theta(t)$ is a nonrandom function, and $\widetilde{B}_1(t)$ and $\widetilde{B}_2(t)$ are correlated Brownian motions with $d\widetilde{B}_1(t)d\widetilde{B}_2(t) = \rho dt$ for some $\rho \in (-1, 1)$. In this exercise, we discuss how to reduce this to the two-factor Vasicek model of Subsection 10.2.1, except that, instead of (10.2.6), the interest rate is given by (10.7.7), in which $\delta_0(t)$ is a nonrandom function of time.

(i) Define

$$X(t) = \begin{bmatrix} U(t) \\ R(t) \end{bmatrix},\ K = \begin{bmatrix} \lambda_1 & 0 \\ -1 & \lambda_2 \end{bmatrix},\ \Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

$$\Theta(t) = \begin{bmatrix} 0 \\ \theta(t) \end{bmatrix},\ \widetilde{B}(t) = \begin{bmatrix} \widetilde{B}_1(t) \\ \widetilde{B}_2(t) \end{bmatrix},$$

so that (10.7.10) and (10.7.11) can be written in vector notation as

$$dX(t) = \Theta(t)dt - KX(t)dt + \Sigma d\widetilde{B}(t). \tag{10.7.12}$$

Now set

$$\widehat{X}(t) = X(t) - e^{-Kt}\int_0^t e^{Ku}\Theta(u)du.$$

Show that

$$d\widehat{X}(t) = -K\widehat{X}(t)dt + \Sigma d\widetilde{B}(t). \tag{10.7.13}$$

Proof.

$$\begin{aligned} d\widehat{X}(t) &= dX(t) + Ke^{-Kt}\int_0^t e^{Ku}\Theta(u)dudt - \Theta(t)dt \\ &= -KX(t)dt + \Sigma d\widetilde{B}(t) + Ke^{-Kt}\int_0^t e^{Ku}\Theta(u)dudt \\ &= -K\widehat{X}(t)dt + \Sigma d\widetilde{B}(t). \end{aligned}$$

□

Remark 6. *The definition of $\widehat{X}$ is motivated by the observation that Y has homogeneous dt term in (10.2.4)-(10.2.5) and e^{Kt} is the integrating factor for X: $d\left(e^{Kt}X(t)\right) = e^{Kt}\left(dX(t) - KX(t)dt\right)$.*

(ii) With

$$C = \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{bmatrix},$$

define $Y(t) = C\widehat{X}(t)$, $\widetilde{W}(t) = C\Sigma\widetilde{B}(t)$. Show that the components of $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ are independent Brownian motions and

$$dY(t) = -\Lambda Y(t)dt + d\widetilde{W}(t), \tag{10.7.14}$$

where

$$\Lambda = CKC^{-1} = \begin{bmatrix} \lambda_1 & 0 \\ \frac{\rho\sigma_2(\lambda_2-\lambda_1)-\sigma_1}{\sigma_2\sqrt{1-\rho^2}} & \lambda_2 \end{bmatrix}.$$

Equation (10.7.14) is the vector form of the canonical two-factor Vasicek equations (10.2.4) and (10.2.5).

Proof.

$$\widetilde{W}(t) = C\Sigma\widetilde{B}(t) = \begin{pmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \widetilde{B}(t) = \begin{pmatrix} 1 & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}} & \frac{1}{\sqrt{1-\rho^2}} \end{pmatrix} \widetilde{B}(t).$$

So $\widetilde{W}$ is a martingale with $\langle \widetilde{W}^1 \rangle(t) = \langle \widetilde{B}^1 \rangle(t) = t$,

$$\langle \widetilde{W}^2 \rangle(t) = \left\langle -\frac{\rho}{\sqrt{1-\rho^2}}\widetilde{B}^1 + \frac{1}{\sqrt{1-\rho^2}}\widetilde{B}^2 \right\rangle(t) = \frac{\rho^2 t}{1-\rho^2} + \frac{t}{1-\rho^2} - 2\frac{\rho}{1-\rho^2}\rho t = \frac{\rho^2+1-2\rho^2}{1-\rho^2}t = t,$$

and

$$\langle \widetilde{W}^1, \widetilde{W}^2 \rangle(t) = \left\langle \widetilde{B}^1, -\frac{\rho}{\sqrt{1-\rho^2}}\widetilde{B}^1 + \frac{1}{\sqrt{1-\rho^2}}\widetilde{B}^2 \right\rangle(t) = -\frac{\rho t}{\sqrt{1-\rho^2}} + \frac{\rho t}{\sqrt{1-\rho^2}} = 0.$$

Therefore $\widetilde{W}$ is a two-dimensional BM. Moreover, $dY_t = Cd\widehat{X}_t = -CK\widehat{X}_t dt + C\Sigma d\widetilde{B}_t = -CKC^{-1}Y_t dt + d\widetilde{W}_t = -\Lambda Y_t dt + d\widetilde{W}_t$, where

$$\begin{aligned}
\Lambda &= CKC^{-1} = \begin{pmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ -1 & \lambda_2 \end{pmatrix} \cdot \frac{1}{|C|} \begin{pmatrix} \frac{1}{\sigma_2\sqrt{1-\rho^2}} & 0 \\ \frac{\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_1} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\lambda_1}{\sigma_1} & 0 \\ -\frac{\rho\lambda_1}{\sigma_1\sqrt{1-\rho^2}} - \frac{1}{\sigma_2\sqrt{1-\rho^2}} & \frac{\lambda_2}{\sigma_2\sqrt{1-\rho^2}} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & 0 \\ \frac{\rho\sigma_2(\lambda_2-\lambda_1)-\sigma_1}{\sigma_2\sqrt{1-\rho^2}} & \lambda_2 \end{pmatrix}.
\end{aligned}$$

□

(iii) Obtain a formula for $R(t)$ of the form (10.7.7). What are $\delta_0(t)$, δ_1, and δ_2?

Solution.

$$\begin{aligned}
X(t) &= \widehat{X}(t) + e^{-Kt}\int_0^t e^{Ku}\Theta(u)du = C^{-1}Y_t + e^{-Kt}\int_0^t e^{Ku}\Theta(u)du \\
&= \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} Y_1(t) \\ Y_2(t) \end{pmatrix} + e^{-Kt}\int_0^t e^{Ku}\Theta(u)du \\
&= \begin{pmatrix} \sigma_1 Y_1(t) \\ \rho\sigma_2 Y_1(t) + \sigma_2\sqrt{1-\rho^2}Y_2(t) \end{pmatrix} + e^{-Kt}\int_0^t e^{Ku}\Theta(u)du.
\end{aligned}$$

So

$$R(t) = X_2(t) = \rho\sigma_2 Y_1(t) + \sigma_2\sqrt{1-\rho^2}Y_2(t) + \delta_0(t),$$

where $\delta_0(t)$ is the second coordinate of $e^{-Kt}\int_0^t e^{Ku}\Theta(u)du$ and can be derived explicitly by Lemma 10.2.3. Note

$$\begin{aligned}
e^{-Kt}\int_0^t e^{Ku}\Theta(u)du &= \int_0^t \begin{bmatrix} e^{\lambda_1(u-t)} & 0 \\ * & e^{\lambda_2(u-t)} \end{bmatrix} \begin{bmatrix} 0 \\ \theta(u) \end{bmatrix} du \\
&= \int_0^t \begin{bmatrix} 0 \\ \theta(u)e^{\lambda_2(u-t)} \end{bmatrix} du
\end{aligned}$$

Then $\delta_0(t) = e^{-\lambda_2 t}\int_0^t e^{\lambda_2 u}\theta(u)du$, $\delta_1 = \rho\sigma_2$, and $\delta_2 = \sigma_2\sqrt{1-\rho^2}$. □

▶ **Exercise 10.5 (Correlation between long rate and short rate in the one-factor Vasicek model).** The one-factor Vasicek model is the one-factor Hull-White model of Example 6.5.1 with constant parameters,

$$dR(t) = (a - bR(t))dt + \sigma d\widetilde{W}(t), \tag{10.7.15}$$

where a, b, and σ are positive constants and $\widetilde{W}(t)$ is a one-dimensional Brownian motion. In this model, the price at time $t \in [0,T]$ of the zero-coupon bond maturing at time T is

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

where $C(t,T)$ and $A(t,T)$ are given by (6.5.10) and (6.5.11):

$$\begin{aligned} C(t,T) &= \int_t^T e^{-\int_t^s b dv} ds = \frac{1}{b}\left(1 - e^{-b(T-t)}\right), \\ A(t,T) &= \int_t^T \left(aC(s,t) - \frac{1}{2}\sigma^2 C^2(s,T)\right) ds \\ &= \frac{2ab - \sigma^2}{2b^2}(T-t) + \frac{\sigma^2 - ab}{b^3}\left(1 - e^{-b(T-t)}\right) - \frac{\sigma^2}{4b^3}\left(1 - e^{-2b(T-t)}\right). \end{aligned}$$

In the spirit of the discussion of the short rate and the long rate in Subsection 10.2.1, we fix a positive relative maturity $\bar{\tau}$ and define the long rate $L(t)$ at time t by (10.2.30):

$$L(t) = -\frac{1}{\bar{\tau}} \log B(t, t+\bar{\tau}).$$

Show that changes in $L(t)$ and $R(t)$ are perfectly correlated (i.e., for any $0 \leq t_1 < t_2$, the correlation coefficient between $L(t_2) - L(t_1)$ and $R(t_2) - R(t_1)$ is one). This characteristic of one-factor models caused the development of models with more than one factor.

Proof. We note $C(t,T)$ and $A(t,T)$ are dependent only on $T-t$. So $C(t,t+\bar{\tau})$ and $A(t,t+\bar{\tau})$ are constants when $\bar{\tau}$ is fixed. So

$$\begin{aligned} \frac{d}{dt}L(t) &= -\frac{B(t,t+\bar{\tau})[-C(t,t+\bar{\tau})R'(t) - A(t,t+\bar{\tau})]}{\bar{\tau}B(t,t+\bar{\tau})} \\ &= \frac{1}{\bar{\tau}}[C(t,t+\bar{\tau})R'(t) + A(t,t+\bar{\tau})] \\ &= \frac{1}{\bar{\tau}}[C(0,\bar{\tau})R'(t) + A(0,\bar{\tau})]. \end{aligned}$$

Integrating from t_1 to t_2 on both sides, we have $L(t_2) - L(t_1) = \frac{1}{\bar{\tau}}C(0,\bar{\tau})[R(t_2) - R(t_1)] + \frac{1}{\bar{\tau}}A(0,\bar{\tau})(t_2 - t_1)$. Since $L(t_2) - L(t_1)$ is a linear transformation of $R(t_2) - R(t_1)$, their correlation is 1. □

▶ **Exercise 10.6 (Degenerate two-factor Vasicek model).** In the discussion of short rates and long rates in the two-factor Vasicek model of Subsection 10.2.1, we made the assumptions that $\delta_2 \neq 0$ and $(\lambda_1 - \lambda_2)\delta_1 + \lambda_{21}\delta_2 \neq 0$ (see Lemma 10.2.2). In this exercise, we show that if either of these conditions is violated, the two-factor Vasicek model reduces to a one-factor model, for which long rates and short rates are perfectly correlated (see Exercise 10.5).

(i) Show that if $\delta_2 = 0$ (and $\delta_0 > 0$, $\delta_1 > 0$), then the short rate $R(t)$ given by the system of equations (10.2.4)-(10.2.6) satisfies the one-dimensional stochastic differential equation

$$dR(t) = (a - bR(t))dt + d\widetilde{W}_1(t). \tag{10.7.16}$$

Define a and b in terms of the parameters in (10.2.4)-(10.2.6).

Proof. If $\delta_2 = 0$, then

$$\begin{aligned} dR(t) &= \delta_1 \left(-\lambda_1 Y_1(t)dt + d\widetilde{W}_1(t)\right) \\ &= \delta_1 \left[\left(\frac{\delta_0}{\delta_1} - \frac{R(t)}{\delta_1}\right)\lambda_1 dt + d\widetilde{W}_1(t)\right] \\ &= (\delta_0\lambda_1 - \lambda_1 R(t))dt + \delta_1 d\widetilde{W}_1(t). \end{aligned}$$

So $a = \delta_0\lambda_1$ and $b = \lambda_1$. □

(ii) Show that if $(\lambda_1 - \lambda_2)\delta_1 + \lambda_{21}\delta_2 = 0$ (and $\delta_0 > 0$, $\delta_1^2 + \delta_2^2 \neq 0$), then the short rate $R(t)$ given by the system of equations (10.2.4)-(10.2.6) satisfies the one-dimensional stochastic differential equation

$$dR(t) = (a - bR(t))dt + \sigma d\widetilde{B}(t). \tag{10.7.17}$$

Define a and b in terms of the parameters in (10.2.4)-(10.2.6) and define the Brownian motion $\widetilde{B}(t)$ in terms of the independent Brownian motions $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ in (10.2.4) and (10.2.5).

Proof.

$$\begin{aligned} dR(t) &= \delta_1 dY_1(t) + \delta_2 dY_2(t) \\ &= -\delta_1\lambda_1 Y_1(t)dt + \lambda_1 d\widetilde{W}_1(t) - \delta_2\lambda_{21}Y_1(t)dt - \delta_2\lambda_2 Y_2(t)dt + \delta_2 d\widetilde{W}_2(t) \\ &= -Y_1(t)(\delta_1\lambda_1 + \delta_2\lambda_{21})dt - \delta_2\lambda_2 Y_2(t)dt + \delta_1 d\widetilde{W}_1(t) + \delta_2 d\widetilde{W}_2(t) \\ &= -Y_1(t)\lambda_2\delta_1 dt - \delta_2\lambda_2 Y_2(t)dt + \delta_1 d\widetilde{W}_1(t) + \delta_2 d\widetilde{W}_2(t) \\ &= -\lambda_2(Y_1(t)\delta_1 + Y_2(t)\delta_2)dt + \delta_1 d\widetilde{W}_1(t) + \delta_2 d\widetilde{W}_2(t) \\ &= -\lambda_2(R_t - \delta_0)dt + \sqrt{\delta_1^2 + \delta_2^2}\left[\frac{\delta_1}{\sqrt{\delta_1^2+\delta_2^2}}d\widetilde{W}_1(t) + \frac{\delta_2}{\sqrt{\delta_1^2+\delta_2^2}}d\widetilde{W}_2(t)\right]. \end{aligned}$$

So $a = \lambda_2\delta_0$, $b = \lambda_2$, $\sigma = \sqrt{\delta_1^2 + \delta_2^2}$ and $\widetilde{B}(t) = \frac{\delta_1}{\sqrt{\delta_1^2+\delta_2^2}}\widetilde{W}_1(t) + \frac{\delta_2}{\sqrt{\delta_1^2+\delta_2^2}}\widetilde{W}_2(t)$. □

► **Exercise 10.7 (Forward measure in the two-factor Vasicek model).** Fix a maturity $T > 0$. In the two-factor Vasicek model of Subsection 10.2.1, consider the T-forward measure $\widetilde{\mathbb{P}}^T$ of Definition 9.4.1:

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

(i) Show that the two-dimensional $\widetilde{\mathbb{P}}^T$-Brownian motion $\widetilde{W}_1^T(t)$, $\widetilde{W}_2^T(t)$ of (9.2.5) are

$$\widetilde{W}_j^T(t) = \int_0^t C_j(T-u)du + \widetilde{W}_j(t),\ j = 1, 2, \tag{10.7.18}$$

where $C_1(\tau)$ and $C_2(\tau)$ are given by (10.2.26)-(10.2.28).

Proof. We use the canonical form of the model as in formulas (10.2.4)-(10.2.6). By (10.2.20),

$$\begin{aligned} dB(t,T) &= df(t, Y_1(t), Y_2(t)) \\ &= de^{-Y_1(t)C_1(T-t) - Y_2(t)C_2(T-t) - A(T-t)} \\ &= (\cdots)dt + B(t,T)[-C_1(T-t)d\widetilde{W}_1(t) - C_2(T-t)d\widetilde{W}_2(t)] \\ &= (\cdots)dt + B(t,T)(-C_1(T-t), -C_2(T-t))\begin{pmatrix} d\widetilde{W}_1(t) \\ d\widetilde{W}_2(t)\end{pmatrix}. \end{aligned}$$

So the volatility vector of $B(t,T)$ under $\widetilde{\mathbb{P}}$ is $(-C_1(T-t), -C_2(T-t))$. By (9.2.5), $\widetilde{W}_j^T(t) = \int_0^t C_j(T-u)du + \widetilde{W}_j(t)$ $(j = 1, 2)$ form a two-dimensional $\widetilde{\mathbb{P}}^T-$BM. □

(ii) Consider a call option on a bond maturing at time $\bar{T} > T$. The call expires at time T and has strike price K. Show that at time zero the risk-neutral price of this option is

$$B(0,T)\widetilde{\mathbb{E}}^T\left[\left(e^{-C_1(\bar{T}-T)Y_1(T)-C_2(\bar{T}-T)Y_2(T)-A(\bar{T}-T)} - K\right)^+\right]. \tag{10.7.19}$$

Proof. Under the T-forward measure $\widetilde{\mathbb{P}}^T$, the numeraire is $B(t,T)$. By risk-neutral pricing, at time zero the risk-neutral price $V(0)$ of the option satisfies

$$\frac{V(0)}{B(0,T)} = \widetilde{\mathbb{E}}^T\left[\frac{1}{B(T,T)}\left(e^{-C_1(\bar{T}-T)Y_1(T)-C_2(\bar{T}-T)Y_2(T)-A(\bar{T}-T)} - K\right)^+\right].$$

Note $B(T,T) = 1$, we get (10.7.19). □

(iii) Show that, under the T-forward measure $\widetilde{\mathbb{P}}^T$, the term

$$X = -C_1(\bar{T}-T)Y_1(T) - C_2(\bar{T}-T)Y_2(T) - A(\bar{T}-T)$$

appearing in the exponent in (10.7.19) is normally distributed.

Proof. Using (10.7.18), we can rewrite (10.2.4) and (10.2.5) as

$$\begin{cases} dY_1(t) = -\lambda_1 Y_1(t)dt + d\widetilde{W}_1^T(t) - C_1(T-t)dt \\ dY_2(t) = -\lambda_{21}Y_1(t)dt - \lambda_2 Y_2(t)dt + d\widetilde{W}_2^T(t) - C_2(T-t)dt. \end{cases}$$

Then

$$\begin{cases} Y_1(t) = Y_1(0)e^{-\lambda_1 t} + \int_0^t e^{\lambda_1(s-t)}d\widetilde{W}_1^T(s) - \int_0^t C_1(T-s)e^{\lambda_1(s-t)}ds \\ Y_2(t) = Y_0 e^{-\lambda_2 t} - \lambda_{21}\int_0^t Y_1(s)e^{\lambda_2(s-t)}ds + \int_0^t e^{\lambda_2(s-t)}d\widetilde{W}_2(s) - \int_0^t C_2(T-s)e^{\lambda_2(s-t)}ds. \end{cases}$$

Since C_1 is deterministic, Y_1 has Gaussian distribution. As the consequence, the second term in the expression of Y_2, $\int_0^t Y_1(s)e^{\lambda_2(s-t)}ds$ also has Gaussian distribution and is uncorrelated to $\int_0^t e^{\lambda_2(s-t)}d\widetilde{W}_2(s)$, since $\widetilde{W}_1^T$ and $\widetilde{W}_2^T$ are uncorrelated. Therefore, (Y_1, Y_2) is jointly Gaussian and X as a linear combination of them is also Gaussian. □

(iv) It is a straightforward but lengthy computation, like the computations in Exercise 10.1, to determine the mean and variance of the term X. Let us call its variance σ^2 and its mean $\mu - \frac{1}{2}\sigma^2$, so that we can write X as

$$X = \mu - \frac{1}{2}\sigma^2 - \sigma Z,$$

where Z is a standard normal random variable under $\widetilde{\mathbb{P}}^T$. Show that the call option price in (10.7.19) is

$$B(0,T)\left(e^\mu N(d_+) - KN(d_-)\right),$$

where

$$d_\pm = \frac{1}{\sigma}\left(\mu - \log K \pm \frac{1}{2}\sigma^2\right).$$

Proof. The call option price in (10.7.19) is

$$B(0,T)\widetilde{\mathbb{E}}^T\left[\left(e^X - K\right)^+\right]$$

with $\widetilde{\mathbb{E}}^T[X] = \mu - \frac{1}{2}\sigma^2$ and $\mathrm{Var}(X) = \sigma^2$. Comparing with the Black-Scholes formula for call options: if $dS_t = rS_t dt + \sigma S_t d\widetilde{W}_t$, then

$$\widetilde{\mathbb{E}}\left[e^{-rT}\left(S_0 e^{\sigma W_T + (r-\frac{1}{2}\sigma^2)T} - K\right)^+\right] = S_0 N(d_+) - Ke^{-rT}N(d_-)$$

with $d_\pm = \frac{1}{\sigma\sqrt{T}}(\log\frac{S_0}{K} + (r \pm \frac{1}{2}\sigma^2)T)$, we can set in the Black-Scholes formula $r = \mu$, $T = 1$ and $S_0 = 1$, then

$$B(0,T)\widetilde{\mathbb{E}}^T\left[\left(e^X - K\right)^+\right] = B(0,T)e^{\mu}\widetilde{\mathbb{E}}^T\left[e^{-\mu}\left(e^X - K\right)^+\right] = B(0,T)\left(e^{\mu}N(d_+) - KN(d_-)\right)$$

where $d_\pm = \frac{1}{\sigma}(-\log K + (\mu \pm \frac{1}{2}\sigma^2))$. $\square$

▶ **Exercise 10.8 (Reversal of order of integration in forward rates).** The forward rate formula (10.3.5) with v replacing T states that

$$f(t,v) = f(0,v) + \int_0^t \alpha(u,v)du + \int_0^t \sigma(u,v)dW(u).$$

Therefore

$$-\int_t^T f(t,v)dv = -\int_t^T \left[f(0,v) + \int_0^t \alpha(u,v)du + \int_0^t \sigma(u,v)dW(u)\right]dv. \tag{10.7.20}$$

(i) Define

$$\widehat{\alpha}(u,t,T) = \int_t^T \alpha(u,v)dv,\ \widehat{\sigma}(u,t,T) = \int_t^T \sigma(u,v)dv.$$

Show that if we reverse the order of integration in (10.7.20), we obtain the equation

$$-\int_t^T f(t,v)dv = -\int_t^T f(0,v)dv - \int_0^t \widehat{(}u,t,T)du - \int_0^t \widehat{\sigma}(u,t,T)dW(u). \tag{10.7.21}$$

(In one case, this is a reversal of the order of two Riemann integrals, a step that uses only the theory of ordinary calculus. In the other case, the order of a Riemann and an Itô integral are being reversed. This step is justified in the appendix of [83]. You may assume without proof that this step is legitimate.)

Proof. Starting from (10.7.20), we have

$$\begin{aligned}
-\int_t^T f(t,v)dv &= -\int_t^T f(0,v)dv - \int_{\{t\le v\le T, 0\le u\le t\}} \alpha(u,v)dudv - \int_{\{t\le v\le T, 0\le u\le t\}} \sigma(u,v)dW(u)dv \\
&= -\int_t^T f(0,v)dv - \int_0^t du\int_t^T \alpha(u,v)dv - \int_0^t dW(u)\int_t^T \sigma(u,v)dv \\
&= -\int_t^T f(0,v)dv - \int_0^t \widehat{\alpha}(u,t,T)du - \int_0^t \widehat{\sigma}(u,t,T)dW(u).
\end{aligned}$$

$\square$

(ii) Take the differential with respect to t in (10.7.21), remembering to get two terms from each of the integrals $\int_0^t \widehat{\alpha}(u,t,T)du$ and $\int_0^t \widehat{\sigma}(u,t,T)dW(u)$ because one must differentiate with respect to each of the two ts appearing in these integrals.

Solution. Differentiating both sides of (10.7.21), we have

$$\begin{aligned}
& d\left(-\int_t^T f(t,v)dv\right) \\
=\ & f(0,t)dt - \widehat{\alpha}(t,t,T)dt - \int_0^t \frac{\partial}{\partial t}\widehat{\alpha}(u,t,T)du - \widehat{\sigma}(t,t,T)dW(t) - \int_0^t \frac{\partial}{\partial t}\widehat{\sigma}(u,t,T)dW(u) \\
=\ & f(0,t)dt - \int_t^T \alpha(t,v)dvdt + \int_0^t \alpha(u,t)dudt - \int_t^T \sigma(t,v)dvdW(t) + \int_0^t \sigma(u,t)dW(u)dt.
\end{aligned}$$

$\square$

(iii) Check that your formula in (ii) agrees with (10.3.10).

Proof. We note $R(t) = f(t,t) = f(0,t) + \int_0^t \alpha(u,v)du + \int_0^t \sigma(u,v)dW(u)$. From (ii), we therefore have

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

which is (10.3.10). □

▶ **Exercise 10.9 (Multifactor HJM model).** Suppose the Heath-Jarrow-Morton model is driven by a d-dimensional Brownian motion, so that $\sigma(t,T)$ is also a d-dimensional vector and the forward rate dynamics are given by

$$df(t,T) = \alpha(t,T)dt + \sum_{j=1}^{d} \sigma_j(t,T)dW_j(t).$$

(i) Show that (10.3.16) becomes

$$\alpha(t,T) = \sum_{j=1}^{d} \sigma_j(t,T)[\sigma_j^*(t,T) + \Theta_j(t)].$$

Proof. We first derive the SDE for the discounted bond price $D(t)B(t,T)$. We first note

$$\begin{aligned} d\left(-\int_t^T f(t,v)dv\right) &= f(t,t)dt - \int_t^T df(t,v)dv = R(t)dt - \int_t^T \left[\alpha(t,v)dt + \sum_{j=1}^{d} \sigma_j(t,v)dW_j(t)\right] dv \\ &= R(t)dt - \alpha^*(t,T)dt - \sum_{j=1}^{d} \sigma_j^*(t,T)dW_j(t). \end{aligned}$$

The applying Itô's formula, we have

$$\begin{aligned} dB(t,T) &= de^{-\int_t^T f(t,v)dv} = B(t,T)\left[d\left(-\int_t^T f(t,v)dv\right) + \frac{1}{2}\sum_{j=1}^{d}(\sigma_j^*(t,T))^2 dt\right] \\ &= B(t,T)\left[R(t) - \alpha^*(t,T) + \frac{1}{2}\sum_{j=1}^{d}(\sigma_j^*(t,T))^2\right] dt - B(t,T)\sum_{j=1}^{d}\sigma_j^*(t,T)dW_j(t). \end{aligned}$$

Using integration-by-parts formula, we have

$$\begin{aligned} d(D(t)B(t,T)) &= D(t)[-B(t,T)R(t)dt + dB(t,T)] \\ &= D(t)B(t,T)\left[\left(-\alpha^*(t,T) + \frac{1}{2}\sum_{j=1}^{d}(\sigma_j^*(t,T))^2\right) dt - \sum_{j=1}^{d}\sigma_j^*(t,T)dW_j(t)\right] \end{aligned}$$

If no-arbitrage condition holds, we can find a risk-neutral measure $\widetilde{\mathbb{P}}$ under which

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

is a d-dimensional Brownian motion for some d-dimensional adapted process $\Theta(t)$. This implies

$$-\alpha^*(t,T) + \frac{1}{2}\sum_{j=1}^{d}(\sigma_j^*(t,T))^2 + \sum_{j=1}^{d}\sigma_j^*(t,T)\Theta_j(t) = 0$$

Differentiating both sides w.r.t. T, we obtain

$$-\alpha(t,T)+\sum_{j=1}^{d}\sigma_j^*(t,T)\sigma_j(t,T)+\sum_{j=1}^{d}\sigma_j(t,T)\Theta_j(t)=0.$$

Or equivalently,

$$\alpha(t,T)=\sum_{j=1}^{d}\sigma_j(t,T)[\sigma_j^*(t,T)+\Theta_j(t)].$$

□

(ii) Suppose there is an adapted, d-dimensional process

$$\Theta(t)=(\Theta_1(t),\cdots,\Theta_d(t))$$

satisfying this equation for all $0\leq t\leq T\leq \overline{T}$. Show that if there are maturities T_1, $\cdots$, T_d such that the $d\times d$ matrix $(\sigma_j(t,T_i))_{i,j}$ is nonsingular, then $\Theta(t)$ is unique.

Proof. Let $\sigma(t,T)=[\sigma_1(t,T),\cdots,\sigma_d(t,T)]^{tr}$ and $\sigma^*(t,T)=[\sigma_1^*(t,T),\cdots,\sigma_d^*(t,T)]^{tr}$. Then the no-arbitrage condition becomes

$$\alpha(t,T)-(\sigma^*(t,T))^{tr}\sigma(t,T)=(\sigma(t,T))^{tr}\Theta(t).$$

Let T iterate T_1, $\cdots$, T_d, we have a matrix equation

$$\begin{bmatrix}\alpha(t,T_1)-(\sigma^*(t,T_1))^{tr}\sigma(t,T_1)\\ \vdots\\ \alpha(t,T_d)-(\sigma^*(t,T_d))^{tr}\sigma(t,T_d)\end{bmatrix}=\begin{bmatrix}\sigma_1(t,T_1) & \ldots & \sigma_d(t,T_1)\\ \vdots & \ddots & \vdots\\ \sigma_1(t,T_d) & \ldots & \sigma_d(t,T_d)\end{bmatrix}\begin{bmatrix}\Theta_1(t)\\ \vdots\\ \Theta_d(t)\end{bmatrix}$$

Therefore, if $(\sigma_j(t,T_i))_{i,j}$ is nonsingular, $\Theta(t)$ can be uniquely solved from the above matrix equation. □

► **Exercise 10.10.** (i) Use the ordinary differential equations (6.5.8) and (6.5.9) satisfied by the functions $A(t,T)$ and $C(t,T)$ in the one-factor Hull-White model to show that this model satisfies the HJM no-arbitrage condition (10.3.27)

Proof. Recall (6.5.8) and (6.5.9) are

$$\begin{cases}C'(t,T)=b(t)C(t,T)-1\\ A'(t,T)=-a(t)C(t,T)+\frac{1}{2}\sigma^2(t)C^2(t,T).\end{cases}$$

Then by $\beta(t,r)=a(t)-b(t)r$ and $\gamma(t,r)=\sigma(t)$ (p. 430), we have

$$\begin{aligned}&\frac{\partial}{\partial T}C(t,T)\beta(t,R(t))+R(t)\frac{\partial}{\partial T}C'(t,T)+\frac{\partial}{\partial T}A'(t,T)\\ =\;&\frac{\partial}{\partial T}C(t,T)\beta(t,R(t))+R(t)b(t)\frac{\partial}{\partial T}C(t,T)+\left[-a(t)\frac{\partial}{\partial T}C(t,T)+\sigma^2(t)C(t,T)\frac{\partial}{\partial T}C(t,T)\right]\\ =\;&\frac{\partial}{\partial T}C(t,T)\left[\beta(t,R(t))+R(t)b(t)-a(t)+\sigma^2(t)C(t,T)\right]\\ =\;&\left(\frac{\partial}{\partial T}C(t,T)\right)C(t,T)\gamma^2(t,R(t)).\end{aligned}$$

□

(ii) Use the ordinary differential equations (6.5.14) and (6.5.15) satisfied by the functions $A(t,T)$ and $C(t,T)$ in the one-factor Cox-Ingersoll-Ross model to show that this model satisfies the HJM no-arbitrage condition (10.3.27).

Proof. Recall (6.5.14) and (6.5.15) are

$$\begin{cases} C'(t,T) = bC(t,T) + \frac{1}{2}\sigma^2 C^2(t,T) - 1 \\ A'(t,T) = -aC(t,T). \end{cases}$$

Then by $\beta(t,r) = a - br$ and $\gamma(t,r) = \sigma\sqrt{r}$ (p. 430), we have

$$\begin{aligned} & \frac{\partial}{\partial T}C(t,T)\beta(t,R(t)) + R(t)\frac{\partial}{\partial T}C'(t,T) + \frac{\partial}{\partial T}A'(t,T) \\ = \; & \frac{\partial}{\partial T}C(t,T)\beta(t,R(t)) + R(t)\left[b\frac{\partial}{\partial T}C(t,T) + \sigma^2 C(t,T)\frac{\partial}{\partial T}C(t,T)\right] - a\frac{\partial}{\partial T}C(t,T) \\ = \; & \frac{\partial}{\partial T}C(t,T)\left[\beta(t,R(t)) + R(t)b + R(t)\sigma^2 C(t,T) - a\right] \\ = \; & \left(\frac{\partial}{\partial T}C(t,T)\right)C(t,T)\gamma^2(t,R(t)). \end{aligned}$$

□

► **Exercise 10.11.** Let $\delta > 0$ be given. Consider an interest rate swap paying a fixed interest rate K and receiving backset LIBOR $L(T_{j-1}, T_{j-1})$ on a principal of 1 at each of the payment dates $T_j = \delta j$, $j = 1, 2, \cdots, n+1$. Show that the value of the swap is

$$\delta K \sum_{j=1}^{n+1} B(0,T_j) - \delta \sum_{j=1}^{n+1} B(0,T_j)L(0,T_{j-1}). \tag{10.7.22}$$

Remark 10.7.1 The *swap rate* is defined to be the value of K that makes the initial value of the swap equal to zero. Thus, the swap rate is

$$K = \frac{\sum_{j=1}^{n+1} B(0,T_j)L(0,T_{j-1})}{\sum_{j=1}^{n+1} B(0,T_j)}. \tag{10.7.23}$$

Proof. On each payment date T_j, the payoff of this swap contract is $\delta(K - L(T_{j-1}, T_{j-1}))$. Its no-arbitrage price at time 0 is $\delta(KB(0,T_j) - B(0,T_j)L(0,T_{j-1}))$ by Theorem 10.4. So the value of the swap is

$$\sum_{j=1}^{n+1} \delta[KB(0,T_j) - B(0,T_j)L(0,T_{j-1})] = \delta K \sum_{j=1}^{n+1} B(0,T_j) - \delta \sum_{j=1}^{n+1} B(0,T_j)L(0,T_{j-1}).$$

□

► **Exercise 10.12.** In the proof of Theorem 10.4.1, we showed by an arbitrage argument that the value at time 0 of a payment of backset LIBOR $L(T,T)$ at time $T+\delta$ is $B(0,T+\delta)L(0,T)$. The risk-neutral price of this payment, computed at time zero, is

$$\widetilde{\mathbb{E}}[D(T+\delta)L(T,T)].$$

Use the definitions

$$\begin{aligned} L(T,T) &= \frac{1 - B(T,T+\delta)}{\delta B(T,T+\delta)}, \\ B(0,T+\delta) &= \widetilde{\mathbb{E}}[D(T+\delta)], \end{aligned}$$

and the properties of conditional expectations to show that

$$\widetilde{\mathbb{E}}[D(T+\delta)L(T,T)] = B(0,T+\delta)L(0,T).$$

Proof. Since $L(T,T) = \frac{1-B(T,T+\delta)}{\delta B(T,T+\delta)} \in \mathcal{F}_T$, we have

$$\begin{aligned}
\widetilde{\mathbb{E}}[D(T+\delta)L(T,T)] &= \widetilde{\mathbb{E}}[\widetilde{\mathbb{E}}[D(T+\delta)L(T,T)|\mathcal{F}_T]] \\
&= \widetilde{\mathbb{E}}\left[\frac{1-B(T,T+\delta)}{\delta B(T,T+\delta)}\widetilde{\mathbb{E}}[D(T+\delta)|\mathcal{F}_T]\right] \\
&= \widetilde{\mathbb{E}}\left[\frac{1-B(T,T+\delta)}{\delta B(T,T+\delta)}D(T)B(T,T+\delta)\right] \\
&= \widetilde{\mathbb{E}}\left[\frac{D(T)-D(T)B(T,T+\delta)}{\delta}\right] \\
&= \frac{B(0,T)-B(0,T+\delta)}{\delta} \\
&= B(0,T+\delta)L(0,T).
\end{aligned}$$

□

Remark 7. *An alternative proof is to use the $(T+\delta)$-forward measure $\widetilde{\mathbb{P}}^{T+\delta}$: the time-$t$ no-arbitrage price of a payoff $\delta L(T,T)$ at time $(T+\delta)$ is*

$$B(t,T+\delta)\widetilde{\mathbb{E}}^{T+\delta}[\delta L(T,T)] = B(t,T+\delta)\cdot\delta L(t,T),$$

where we have used the observation that the forward rate process $L(t,T) = \frac{1}{\delta}\left(\frac{B(t,T)}{B(t,T+\delta)} - 1\right)$ $(0 \le t \le T)$ is a martingale under the $(T+\delta)$-forward measure.

11 Introduction to Jump Processes

★ **Comments**:

1) A mathematically rigorous presentation of semimartingale theory and stochastic calculus can be found in He et al. [5] and Kallenberg [7].

2) *Girsanov's theorem.* The most general version of Girsanov's theorem for local martingales can be found in He et al. [5, page 340] (Theorem 12.13) or Kallenberg [7, page 523] (Theorem 26.9). It has the following form

Theorem 1 (Girsanov's theorem for local martingales)**.** *Let $Q = Z_t \cdot P$ on $\mathcal{F}_t$ for all $t \ge 0$, and consider a local P-martingale M such that the process $[M,Z]$ has locally integrable variation and P-compensator $\langle M,Z\rangle$. Then $\bar{M} = M - Z_-^{-1}\cdot\langle M,Z\rangle$ is a local Q-martingale.*[13]

Applying Girsanov's theorem to Lemma 11.6.1, in order to change the intensity of a Poisson process via change of measure, we need to find a $\mathbb{P}$-martingale L such that the Radon-Nikodým derivative $Z_\cdot = d\widetilde{\mathbb{P}}/d\mathbb{P}$ satisfies the SDE

$$dZ(t) = Z(t-)dL(t)$$

and

$$N(t) - \lambda t - \langle N(t) - \lambda t, L(t)\rangle = N(t) - \widetilde{\lambda}t$$

is a martingale under $\widetilde{\mathbb{P}}$. This implies we should find L such that

$$\langle N(t) - \lambda t, L(t)\rangle = (\widetilde{\lambda} - \lambda)t.$$

[13]For the difference between the *predictable quadratic variation* $\langle\cdot\rangle$ (or *angle bracket process*) and the *quadratic variation* (or *square bracket process*), see He et al. [5, page 185-187]. Basically, we have $[M,N]_t = M_0N_0 + \langle M^c,N^c\rangle_t + \sum_{s\le t}\Delta M_s\Delta N_s$ and $\langle M,N\rangle$ is the dual predictable projection of $[M,N]$ (i.e. $[M,N]_t - \langle M,N\rangle_t$ is a martingale).

Assuming martingale representation theorem, we suppose $L(t) = L(0) + \int_0^t H(s)d(N(s) - \lambda s)$, then

$$[N(t) - \lambda t, L(t)] = \int_0^t H(s)d[N(s) - \lambda s, N(s) - \lambda s] = \sum_{0<s\leq t} H(s)(\Delta N(s))^2 = \int_0^t H(s)dN(s).$$

So $\langle N(t) - \lambda t, L(t)\rangle = \int_0^t H(s)\lambda ds$. Solving the equation

$$\int_0^t H(s)\lambda ds = (\widetilde{\lambda} - \lambda)t,$$

we get $H(s) = \frac{\widetilde{\lambda}-\lambda}{\lambda}$. Combined, we conclude Z should be determined by the equation (11.6.2)

$$dZ(t) = \frac{\widetilde{\lambda} - \lambda}{\lambda} Z(t-)dM(t).$$

► **Exercise 11.1.** Let $M(t)$ be the compensated Poisson process of Theorem 11.2.4.
(i) Show that $M^2(t)$ is a submartingale.

Proof. First, $M^2(t) = N^2(t) - 2\lambda tN(t) + \lambda^2 t^2$. So $\mathbb{E}[M^2(t)] < \infty$. $\varphi(x) = x^2$ is a convex function, by the conditional Jensen's inequality (Shreve [14, page 70], Theorem 2.3.2 (v)),

$$\mathbb{E}[\varphi(M(t))|\mathcal{F}(s)] \geq \varphi(\mathbb{E}[M(t)|\mathcal{F}(s)]) = \varphi(M(s)),\ \forall s \leq t.$$

So $M^2(t)$ is a submartingale. □

(ii) Show that $M^2(t) - \lambda t$ is a martingale.

Proof. We note $M(t) = N(t) - \lambda t$ has independent and stationary increment. So $\forall s \leq t$,

$$\begin{aligned}
& \mathbb{E}[M^2(t) - M^2(s)|\mathcal{F}(s)] \\
= \ & \mathbb{E}[(M(t) - M(s))^2|\mathcal{F}(s)] + \mathbb{E}[(M(t) - M(s)) \cdot 2M(s)|\mathcal{F}(s)] \\
= \ & \mathbb{E}[M^2(t-s)] + 2M(s)\mathbb{E}[M(t-s)] \\
= \ & \mathrm{Var}(N(t-s)) + 0 \\
= \ & \lambda(t-s).
\end{aligned}$$

That is, $\mathbb{E}[M^2(t) - \lambda t|\mathcal{F}(s)] = M^2(s) - \lambda s$. □

► **Exercise 11.2.** Suppose we have observed a Poisson process up to time s, have seen that $N(s) = k$, and are interested in the values of $N(s+t)$ for small positive t. Show that

$$\begin{aligned}
\mathbb{P}\{N(s+t) = k|N(s) = k\} &= 1 - \lambda t + O(t^2), \\
\mathbb{P}\{N(s+t) = k+1|N(s) = k\} &= \lambda t + O(t^2), \\
\mathbb{P}\{N(s+t) \geq k+2|N(s) = k\} &= O(t^2),
\end{aligned}$$

where $O(t^2)$ is used to denote terms involving t^2 and higher powers of t.

Proof. We note

$$\mathbb{P}(N(s+t) = k|N(s) = k) = \mathbb{P}(N(s+t) - N(s) = 0|N(s) = k) = \mathbb{P}(N(t) = 0) = e^{-\lambda t} = 1 - \lambda t + O(t^2).$$

Similarly, we have

$$\mathbb{P}(N(s+t) = k+1|N(s) = k) = \mathbb{P}(N(t) = 1) = \frac{(\lambda t)^1}{1!}e^{-\lambda t} = \lambda t(1 - \lambda t + O(t^2)) = \lambda t + O(t^2),$$

and

$$\mathbb{P}(N(s+t) \geq k+2|N(t) = k) = \mathbb{P}(N(t) \geq 2) = \sum_{k=2}^{\infty} \frac{(\lambda t)^k}{k!}e^{-\lambda t} = O(t^2).$$

□

▶ **Exercise 11.3 (Geometric Poisson process).** Let $N(t)$ be a Poisson process with intensity $\lambda > 0$, and let $S(0) > 0$ and $\sigma > -1$ be given. Using Theorem 11.2.3 rather than the Itô-Doeblin formula for jump processes, show that

$$S(t) = \exp\{N(t)\log(\sigma+1) - \lambda\sigma t\} = (\sigma+1)^{N(t)}e^{-\lambda\sigma t}$$

is a martingale.

Proof. For any $t \leq u$, we have

$$\begin{aligned}
\mathbb{E}\left[\frac{S(u)}{S(t)}\middle|\mathcal{F}(t)\right] &= \mathbb{E}\left[(\sigma+1)^{N(t)-N(u)}e^{-\lambda\sigma(t-u)}|\mathcal{F}(t)\right] \\
&= e^{-\lambda\sigma(t-u)}\mathbb{E}\left[(\sigma+1)^{N(t-u)}\right] \\
&= e^{-\lambda\sigma(t-u)}\mathbb{E}[e^{N(t-u)\log(\sigma+1)}] \\
&= e^{-\lambda\sigma(t-u)}\exp\left\{\lambda(t-u)\left(e^{\log(\sigma+1)}-1\right)\right\} \quad \text{(by (11.3.4))} \\
&= e^{-\lambda\sigma(t-u)}e^{\lambda\sigma(t-u)} \\
&= 1.
\end{aligned}$$

So $S(t) = \mathbb{E}[S(u)|\mathcal{F}(t)]$ and S is a martingale. □

▶ **Exercise 11.4.** Suppose $N_1(t)$ and $N_2(t)$ are Poisson processes with intensities λ_1 and λ_2, respectively, both defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and relative to the same filtration $\mathcal{F}(t)$, $t \geq 0$. Show that almost surely $N_1(t)$ and $N_2(t)$ can have no simultaneous jump. (Hint: Define the compensated Poisson processes $M_1(t) = N_1(t) - \lambda t$ and $M_2(t) = N_2(t) - \lambda_2 t$, which like N_1 and N_2 are independent. Use Itô product rule for jump processes to compute $M_1(t)M_2(t)$ and take expectations.)

Proof. The problem is ambiguous in that the relation between N_1 and N_2 is not clearly stated. According to page 524, paragraph 2, we would guess the condition should be that N_1 and N_2 are independent.

Suppose N_1 and N_2 are independent. Define $M_1(t) = N_1(t) - \lambda_1 t$ and $M_2(t) = N_2(t) - \lambda_2 t$. Then by independence $\mathbb{E}[M_1(t)M_2(t)] = \mathbb{E}[M_1(t)]\mathbb{E}[M_2(t)] = 0$. Meanwhile, by Itô's product formula (Corollary 11.5.5),

$$M_1(t)M_2(t) = \int_0^t M_1(s-)dM_2(s) + \int_0^t M_2(s-)dM_1(s) + [M_1, M_2](t).$$

Both $\int_0^t M_1(s-)dM_2(s)$ and $\int_0^t M_2(s-)dM_1(s)$ are martingales. So taking expectation on both sides, we get

$$0 = 0 + \mathbb{E}\{[M_1, M_2](t)\} = \mathbb{E}\left[\sum_{0<s\leq t}\Delta N_1(s)\Delta N_2(s)\right].$$

Since $\sum_{0<s\leq t}\Delta N_1(s)\Delta N_2(s) \geq 0$ a.s., we conclude $\sum_{0<s\leq t}\Delta N_1(s)\Delta N_2(s) = 0$ a.s. By letting $t = 1, 2, \cdots$, we can find a set Ω_0 of probability 1, so that $\forall\omega \in \Omega_0$, $\sum_{0<s\leq t}\Delta N_1(\omega; s)\Delta N_2(\omega; s) = 0$ for all $t > 0$. Therefore N_1 and N_2 can have no simultaneous jump. □

▶ **Exercise 11.5.** Suppose $N_1(t)$ and $N_2(t)$ are Poisson processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ relative to the same filtration $\mathcal{F}(t)$, $t \geq 0$. Assume that almost surely $N_1(t)$ and $N_2(t)$ have no simultaneous jump. Show that, for each fixed t, the random variable $N_1(t)$ and $N_2(t)$ are independent. (Hint: Adapt the proof of Corollary 11.5.3.) (In fact, the whole path of N_1 is independent of the whole path of N_2, although you are not being asked to prove this stronger statement.)

Proof. We shall prove the whole path of N_1 is independent of the whole path of N_2, following the scheme suggested by page 489, paragraph 1.

Fix $s \geq 0$, we consider $X_t = u_1(N_1(t) - N_1(s)) + u_2(N_2(t) - N_2(s)) - \lambda_1(e^{u_1} - 1)(t-s) - \lambda_2(e^{u_2} - 1)(t-s)$, $t > s$. Then by Itô's formula for jump process, we have

$$\begin{aligned} e^{X_t} - e^{X_s} &= \int_s^t e^{X_u} dX_u^c + \frac{1}{2}\int_s^t e^{X_u} dX_u^c dX_u^c + \sum_{s<u\leq t}(e^{X_u} - e^{X_{u-}}) \\ &= \int_s^t e^{X_u}[-\lambda_1(e^{u_1} - 1) - \lambda_2(e^{u_2} - 1)]du + \sum_{0<u\leq t}(e^{X_u} - e^{X_{u-}}). \end{aligned}$$

Since $\Delta X_t = u_1 \Delta N_1(t) + u_2 \Delta N_2(t)$ and N_1, N_2 have no simultaneous jump,

$$e^{X_u} - e^{X_{u-}} = e^{X_{u-}}(e^{\Delta X_u} - 1) = e^{X_{u-}}[(e^{u_1} - 1)\Delta N_1(u) + (e^{u_2} - 1)\Delta N_2(u)].$$

Note $\int_s^t e^{X_u} du = \int_s^t e^{X_{u-}} du$[14], we have

$$\begin{aligned} & e^{X_t} - e^{X_s} \\ = \ & \int_s^t e^{X_{u-}}[-\lambda_1(e^{u_1} - 1) - \lambda_2(e^{u_2} - 1)]du + \sum_{s<u\leq t} e^{X_{u-}}[(e^{u_1} - 1)\Delta N_1(u) + (e^{u_2} - 1)\Delta N_2(u)] \\ = \ & \int_s^t e^{X_{u-}} \left[(e^{u_1} - 1)d(N_1(u) - \lambda_1 u) - (e^{u_2} - 1)d(N_2(u) - \lambda_2 u)\right]. \end{aligned}$$

Therefore, $\mathbb{E}[e^{X_t}] = \mathbb{E}[e^{X_s}] = 1$, which implies

$$\begin{aligned} \mathbb{E}\left[e^{u_1(N_1(t)-N_1(s))+u_2(N_2(t)-N_2(s))}\right] &= e^{\lambda_1(e^{u_1}-1)(t-s)} e^{\lambda_2(e^{u_2}-1)(t-s)} \\ &= \mathbb{E}\left[e^{u_1(N_1(t)-N_1(s))}\right] \mathbb{E}\left[e^{u_2(N_2(t)-N_2(s))}\right]. \end{aligned}$$

This shows $N_1(t) - N_1(s)$ is independent of $N_2(t) - N_2(s)$.

Now, suppose we have $0 \leq t_1 < t_2 < t_3 < \cdots < t_n$, then the vector $(N_1(t_1), \cdots, N_1(t_n))$ is independent of $(N_2(t_1), \cdots, N_2(t_n))$ if and only if $(N_1(t_1), N_1(t_2) - N_1(t_1), \cdots, N_1(t_n) - N_1(t_{n-1}))$ is independent of $(N_2(t_1), N_2(t_2) - N_2(t_1), \cdots, N_2(t_n) - N_2(t_{n-1}))$. Let $t_0 = 0$, then

$$\begin{aligned} & \mathbb{E}\left[e^{\sum_{i=1}^n u_i(N_1(t_i)-N_1(t_{i-1}))+\sum_{j=1}^n v_j(N_2(t_j)-N_2(t_{j-1}))}\right] \\ = \ & \mathbb{E}\left[e^{\sum_{i=1}^{n-1} u_i(N_1(t_i)-N_1(t_{i-1}))+\sum_{j=1}^{n-1} v_j(N_2(t_j)-N_2(t_{j-1}))}\cdot\right. \\ & \left.\mathbb{E}\left[e^{u_n(N_1(t_n)-N_1(t_{n-1}))+v_n(N_2(t_n)-N_2(t_{n-1}))}\Big|\mathcal{F}(t_{n-1})\right]\right] \\ = \ & \mathbb{E}\left[e^{\sum_{i=1}^{n-1} u_i(N_1(t_i)-N_1(t_{i-1}))+\sum_{j=1}^{n-1} v_j(N_2(t_j)-N_2(t_{j-1}))}\right]\mathbb{E}\left[e^{u_n(N_1(t_n)-N_1(t_{n-1}))+v_n(N_2(t_n)-N_2(t_{n-1}))}\right] \\ = \ & \mathbb{E}\left[e^{\sum_{i=1}^{n-1} u_i(N_1(t_i)-N_1(t_{i-1}))+\sum_{j=1}^{n-1} v_j(N_2(t_j)-N_2(t_{j-1}))}\right]\mathbb{E}\left[e^{u_n(N_1(t_n)-N_1(t_{n-1}))}\right]\mathbb{E}\left[e^{v_n(N_2(t_n)-N_2(t_{n-1}))}\right], \end{aligned}$$

where the second equality comes from the independence of $N_i(t_n) - N_i(t_{n-1})$ $(i = 1, 2)$ relative to $\mathcal{F}(t_{n-1})$ and the third equality comes from the result obtained just before. Working by induction, we have

$$\begin{aligned} & \mathbb{E}\left[e^{\sum_{i=1}^n u_i(N_1(t_i)-N_1(t_{i-1}))+\sum_{j=1}^n v_j(N_2(t_j)-N_2(t_{j-1}))}\right] \\ = \ & \prod_{i=1}^n \mathbb{E}\left[e^{u_i(N_1(t_i)-N_1(t_{i-1}))}\right]\cdot\prod_{j=1}^n \mathbb{E}\left[e^{v_j(N_2(t_j)-N_2(t_{j-1}))}\right] \\ = \ & \mathbb{E}\left[e^{\sum_{i=1}^n u_i(N_1(t_i)-N_1(t_{i-1}))}\right]\cdot\mathbb{E}\left[e^{\sum_{j=1}^n v_j(N_2(t_j)-N_2(t_{j-1}))}\right]. \end{aligned}$$

This shows the whole path of N_1 is independent of the whole path of N_2. $\square$

[14]On the interval $[s, t]$, the sample path $X_\cdot(\omega)$ has only finitely many jumps for each ω, so the Riemann integrals of $X_\cdot$ and $X_{\cdot-}$ should agree with each other.

► **Exercise 11.6.** Let $W(t)$ be a Brownian motion and let $Q(t)$ be a compound Poisson process, both defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and relative to the same filtration $\mathcal{F}(t)$, $t \geq 0$. Show that, for each t, the random variables $W(t)$ and $Q(t)$ are independent. (In fact, the whole path of W is independent of the whole path of Q, although you are not being asked to prove this stronger statement.)

Proof. Let $X_t = u_1 W(t) - \frac{1}{2}u_1^2 t + u_2 Q(t) - \lambda t(\varphi(u_2) - 1)$ where φ is the moment generating function of the jump size Y. Itô's formula for jump process yields

$$e^{X_t} - 1 = \int_0^t e^{X_s}\left(u_1 dW(s) - \frac{1}{2}u_1^2 ds - \lambda(\varphi(u_2)-1)ds\right) + \frac{1}{2}\int_0^t e^{X_s}u_1^2 ds + \sum_{0<s\leq t}(e^{X_s} - e^{X_{s-}}).$$

Note $\Delta X_t = u_2 \Delta Q(t) = u_2 Y_{N(t)}\Delta N(t)$, where $N(t)$ is the Poisson process associated with $Q(t)$. So

$$e^{X_t} - e^{X_{t-}} = e^{X_{t-}}(e^{\Delta X_t} - 1) = e^{X_{t-}}(e^{u_2 Y_{N(t)}} - 1)\Delta N(t).$$

Consider the compound Poisson process $H_t = \sum_{i=1}^{N(t)}(e^{u_2 Y_i} - 1)$, then

$$H_t - \lambda \mathbb{E}\left[e^{u_2 Y_1} - 1\right] t = H_t - \lambda(\varphi(u_2) - 1)t$$

is a martingale, $e^{X_t} - e^{X_{t-}} = e^{X_{t-}}\Delta H_t$ and

$$\begin{aligned} e^{X_t} - 1 &= \int_0^t e^{X_s}\left(u_1 dW(s) - \frac{1}{2}u_1^2 ds - \lambda(\varphi(u_2)-1)ds\right) + \frac{1}{2}\int_0^t e^{X_s}u_1^2 ds + \int_0^t e^{X_{s-}} dH_s \\ &= \int_0^t e^{X_s} u_1 dW(s) + \int_0^t e^{X_{s-}} d(H_s - \lambda(\varphi(u_2)-1)s). \end{aligned}$$

This shows e^{X_t} is a martingale and $\mathbb{E}\left[e^{X_t}\right] \equiv 1$. So

$$\mathbb{E}\left[e^{u_1 W(t) + u_2 Q(t)}\right] = e^{\frac{1}{2}u_1 t} e^{\lambda t(\varphi(u_2)-1)t} = \mathbb{E}\left[e^{u_1 W(t)}\right]\mathbb{E}\left[e^{u_2 Q(t)}\right].$$

This shows $W(t)$ and $Q(t)$ are independent. □

Remark 8. *It is easy to see that, if we follow the steps of solution to Exercise 11.5, firstly proving the independence of $W(t) - W(s)$ and $Q(t) - Q(s)$, then proving the independence of $(W(t_1), W(t_2) - W(t_1), \cdots, W(t_n) - W(t_{n-1}))$ and $(Q(t_1), Q(t_2) - Q(t_1), \cdots, Q(t_n) - Q(t_{n-1}))$ ($0 \leq t_1 < t_2 < \cdots < t_n$), then we can show the whole path of W is independent of the whole path of Q.*

► **Exercise 11.7.** Use Theorem 11.3.2 to prove that a compound Poisson process is Markov. In other words, show that, whenever we are given two times $0 \leq t \leq T$ and a function $h(x)$, there is another function $g(t, x)$ such that

$$\mathbb{E}[h(Q(T))|\mathcal{F}(t)] = g(t, Q(t)).$$

Proof. By Independence Lemma, we have

$$\mathbb{E}[h(Q(T))|\mathcal{F}(t)] = \mathbb{E}[h(Q(T) - Q(t) + Q(t))|\mathcal{F}(t)] = \mathbb{E}[h(Q(T-t) + x)]|_{x=Q(t)} = g(t, Q(t)),$$

where $g(t, x) = \mathbb{E}[h(Q(T-t) + x)]$. □

References

[1] Vladimir I. Arnold and Roger Cooke. *Ordinary differential equations*, 3rd edition, Springer, 1992. 78

[2] Freddy Delbaen and Walter Schachermayer. *The mathematics of arbitrage*. Springer, 2006. 72

[3] 丁同仁、李承治：《常微分方程教程》，北京：高等教育出版社，1991. 78

[4] Richard Durrett. *Probability: Theory and examples*, 4th edition. Cambridge University Press, New York, 2010. 2, 44

[5] Sheng-wu He, Jia-gang Wang, and Jia-an Yan. *Semimartingale theory and stochastic calculus*. Science Press & CRC Press Inc, 1992. 94

[6] John Hull. *Options, futures, and other derivatives*, 4th edition. Prentice-Hall International Inc., New Jersey, 2000. 44

[7] Olav Kallenberg. *Foundations of modern probability*, 2nd edition. Springer, 2002. 94

[8] J. David Logan. *A first course in differential equations*, 2nd edition. Springer, New York, 2010. 27

[9] B. Øksendal. *Stochastic differential equations: An introduction with applications*, 6th edition. Springer-Verlag, Berlin, 2003. 54, 65

[10] 钱敏平、龚光鲁：《随机过程论（第二版）》，北京：北京大学出版社，1997.10。54

[11] Daniel Revuz and Marc Yor. *Continous martingales and Brownian motion*, 3rd edition. Springer-Verlag, Berlin, 1998. 49, 54

[12] A. N. Shiryaev. *Probability*, 2nd edition. Springer, 1995. 23

[13] A. N. Shiryaev. *Essentials of stochastic finance: facts, models, theory*. World Scientific, Singapore, 1999. 72

[14] Steven Shreve. *Stochastic calculus for finance II: Continuous-time models*. Springer-Verlag, New York, 2004. 1, 72, 95

[15] Paul Wilmott. *The mathematics of financial derivatives: A student introduction*. Cambridge University Press, 1995.

27, 38