PART • I

Pricing Theory and Risk Management

CHAPTER • 1

Pricing Theory

Pricing theory for derivative securities is a highly technical topic in finance; its foundations rest on trading practices and its theory relies on advanced methods from stochastic calculus and numerical analysis. This chapter summarizes the main concepts while presenting the essential theory and basic mathematical tools for which the modeling and pricing of financial derivatives can be achieved.

Financial assets are subdivided into several classes, some being quite basic while others are structured as complex contracts referring to more elementary assets. Examples of *elementary asset classes* include *stocks*, which are ownership rights to a corporate entity; *bonds*, which are promises by one party to make cash payments to another in the future; *commodities*, which are assets, such as wheat, metals, and oil that can be consumed; and *real estate* assets, which have a convenience yield deriving from their use. A more general example of an asset is that of a contractual *contingent claim* associated with the obligation of one party to enter a stream of more elementary financial transactions, such as cash payments or deliveries of shares, with another party at future dates. The value of an individual transaction is called a *pay-off* or *payout*. Mathematically, a pay-off can be modeled by means of a *payoff function* in terms of the prices of other, more elementary assets.

There are numerous examples of contingent claims. *Insurance policies*, for instance, are structured as contracts that envision a payment by the insurer to the insured in case a specific event happens, such as a car accident or an illness, and whose pay-off is typically linked to the damage suffered by the insured party. *Derivative assets* are claims that distinguish themselves by the property that the *payoff function* is expressed in terms of the price of an *underlying asset*. In finance jargon, one often refers to underlying assets simply as *underlyings*. To some extent, there is an overlap between insurance policies and derivative assets, except the nomenclature differs because the first are marketed by insurance companies while the latter are traded by banks.

A *trading strategy* consists of a set of rules indicating what positions to take in response to changing market conditions. For instance, a rule could say that one has to adjust the position in a given stock or bond on a daily basis to a level given by evaluating a certain function. The implementation of a trading strategy results in pay-offs that are typically random. A major difference that distinguishes derivative instruments from insurance contracts is that most traded derivatives are structured in such a way that it is possible to implement trading strategies in the underlying assets that generate streams of pay-offs that *replicate* the pay-offs of the derivative claim. In this sense, trading strategies are substitutes for derivative claims. One of the driving forces behind derivatives markets is that some market participants, such as market makers, have a competitive advantage in implementing replication strategies, while their clients are interested in taking certain complex risk exposures synthetically by entering into a single contract.

A key property of replicable derivatives is that the corresponding payoff functions depend only on prices of tradable assets, such as stocks and bonds, and are not affected by events, such as car accidents or individual health conditions that are not directly linked to an asset price. In the latter case, risk can be reduced only by diversification and reinsurance. A related concept is that of *portfolio immunization*, which is defined as a trade intended to offset the risk of a portfolio over at least a short time horizon. A perfect replication strategy for a given claim is one for which a position in the strategy combined with an offsetting position in the claim are perfectly immunized, i.e., risk free. The position in an asset that immunizes a given portfolio against a certain risk is traditionally called *hedge ratio*.[1] An immunizing trade is called a *hedge*. One distinguishes between *static* and *dynamic hedging*, depending on whether the hedge trades can be executed only once or instead are carried over time while making adjustments to respond to new information.

The assets traded to execute a replication strategy are called *hedging instruments*. A set of hedging instruments in a financial model is *complete* if all derivative assets can be replicated by means of a trading strategy involving only positions in that set. In the following, we shall define the mathematical notion of financial models by listing a set of hedging instruments and assuming that there are *no redundancies*, in the sense that no hedging instrument can be replicated by means of a strategy in the other ones. Another very common expression is that of *risk factor*: The risk factors underlying a given financial model with a complete basis of hedging instruments are given by the prices of the hedging instruments themselves or functions thereof; as these prices change, risk factor values also change and the prices of all other derivative assets change accordingly. The statistical analysis of risk factors allows one to assess the risk of financial holdings.

Transaction costs are impediments to the execution of replication strategies and correspond to costs associated with adjusting a position in the hedging instruments. The market for a given asset is *perfectly liquid* if unlimited amounts of the asset can be traded without affecting the asset price. An important notion in finance is that of *arbitrage*: If an asset is replicable by a trading strategy and if the price of the asset is different from that of the replicating strategy, the opportunity for riskless gains/profits arises. Practical limitations to the size of possible gains are, however, placed by the inaccuracy of replication strategies due to either market incompleteness or lack of liquidity. In such situations, either riskless replication strategies are not possible or prices move in response to posting large trades. For these reasons, arbitrage opportunities are typically short lived in real markets.

Most financial models in pricing theory account for finite liquidity indirectly, by postulating that prices are *arbitrage free*. Also, market incompleteness is accounted for indirectly and is reflected in corrections to the probability distributions in the price processes. In this stylized mathematical framework, each asset has a unique *price*.[2]

[1] Notice that the term *hedge ratio* is part of the finance jargon. As we shall see, in certain situations hedge ratios are computed as mathematical ratios or limits thereof, such as derivatives. In other cases, expressions are more complicated.

[2] To avoid the perception of a linguistic ambiguity, when in the following we state that a given asset is *worth* a certain amount, we mean that amount is the asset price.

Most financial models are built upon the *perfect-markets hypothesis*, according to which:

- There are no trading impediments such as transaction costs.
- The set of basic hedging instruments is complete.
- Liquidity is infinite.
- No arbitrage opportunities are present.

These hypotheses are robust in several ways. If liquidity is not perfect, then arbitrage opportunities are short lived because of the actions of arbitrageurs. The lack of completeness and the presence of transaction costs impacts prices in a way that is uniform across classes of derivative assets and can safely be accounted for implicitly by adjusting the process probabilities.

The existence of replication strategies, combined with the perfect-markets hypothesis, makes it possible to apply more sophisticated pricing methodologies to financial derivatives than is generally possible to devise for insurance claims and more basic assets, such as stocks. The key to finding derivative prices is to construct mathematical models for the underlying asset price processes and the replication strategies. Other sources of information, such as a country's domestic product or a takeover announcement, although possibly relevant to the underlying prices, affect derivative prices only indirectly.

This first chapter introduces the reader to the mathematical framework of pricing theory in parallel with the relevant notions of probability, stochastic calculus, and stochastic control theory. The dynamic evolution of the risk factors underlying derivative prices is *random*, i.e., not deterministic, and is subject to uncertainty. Mathematically, one uses *stochastic processes*, defined as random variables with probability distributions on sets of paths. Replicating and hedging strategies are formulated as sets of rules to be followed in response to changing price levels. The key principle of pricing theory is that if a given payoff stream can be replicated by means of a dynamic trading strategy, then the cost of executing the strategy must equal the price of a contractual claim to the payoff stream itself. Otherwise, arbitrage opportunities would ensue. Hence pricing can be reduced to a mathematical optimization problem: to replicate a certain payoff function while minimizing at the same time replication costs and replication risks. In perfect markets one can show that one can achieve perfect replication at a finite cost, while if there are imperfections one will have to find the right trade-off between risk and cost. The *fundamental theorem of asset pricing* is a far-reaching mathematical result that states;

- The solution of this optimization problem can be expressed in terms of a *discounted expectation of future pay-offs* under a pricing (or probability) measure.
- This representation is unique (with respect to a given discounting) as long as markets are complete.

Discounting can be achieved in various ways: using a bond, using the money market account, or in general using a reference *numeraire asset* whose price is positive. This is because pricing assets is a relative, as opposed to an absolute, concept: One values an asset by computing its worth as compared to that of another asset. A key point is that expectations used in pricing theory are computed under a probability measure tailored to the numeraire asset.

In this chapter, we start the discussion with a simple single-period model, where trades can be carried out only at one point in time and gains or losses are observed at a later time, a fixed date in the future. In this context, we discuss static hedging strategies. We then briefly review some of the relevant and most basic elements of probability theory in the context of multivariate continuous random variables. Brownian motion and martingales are then discussed as an introduction to stochastic processes. We then move on to further discuss continuous-time stochastic processes and review the basic framework of stochastic (Itô) calculus. Geometric Brownian motion is then presented, with some preliminary derivations of Black–Scholes formulas for single-asset and multiasset price models. We then proceed to introduce a more general mathematical framework for dynamic hedging and derive the fundamental theorem of asset pricing (FTAP) for continuous-state-space and continuous-time-diffusion processes. We then apply the FTAP to European-style options. Namely, by the use of change of numeraire and stochastic calculus techniques, we show how exact pricing formulas based on geometric Brownian motions for the underlying assets are obtained for a variety of situations, ranging from elementary stock options to foreign exchange and quanto options. The partial differential equation approach for option pricing is then presented. We then discuss pricing theory for early-exercise or American-style options.

1.1 Single-Period Finite Financial Models

The simplest framework in pricing theory is given by *single-period financial models*, in which calendar time t is restricted to take only two values, current time $t = 0$ and a future date $t = T > 0$. Such models are appropriate for analyzing situations where trades can be made *only* at current time $t = 0$. Revenues (i.e., profits or losses) can be realized only at the later date T, while trades at intermediate times are not allowed.

In this section, we focus on the particular case in which only a finite number of *scenarios* $\omega_1, \ldots, \omega_m$ can occur. *Scenario* is a common term for an outcome or event. The *scenario set* $\Omega = \{\omega_1, \ldots, \omega_m\}$ is also called the *probability space*. A *probability measure* P is given by a set of numbers p_i, $i = 1, \ldots, m$, in the interval $[0, 1]$ that sum up to 1; i.e.,

$$\sum_{i=1}^{m} p_i = 1, \qquad 0 \le p_i \le 1. \tag{1.1}$$

p_i is the *probability* that scenario (event) ω_i occurs, i.e., that the ith state is attained. Scenario ω_i is *possible* if it can occur with strictly positive probability $p_i > 0$. Neglecting scenarios that cannot possibly occur, the probabilities p_i will henceforth be assumed to be strictly positive; i.e., $p_i > 0$. A *random variable* is a function on the scenario set, $f : \Omega \to \mathbb{R}$, whose values $f(\omega_i)$ represent observables. As we discuss later in more detail, examples of random variables one encounters in finance include the price of an asset or an interest rate at some point in the future or the pay-off of a derivative contract. The *expectation* of the random variable f is defined as the sum

$$E^P[f] = \sum_{i=1}^{m} p_i f(\omega_i). \tag{1.2}$$

Asset prices and other financial observables, such as interest rates, are modeled by *stochastic processes*. In a single-period model, a *stochastic process* is given by a value f_0 at current time $t = 0$ and by a random variable f_T that models possible values at time T. In finance, probabilities are obtained with two basically different procedures: They can either be *inferred* from historical data by *estimating* a statistical model, or they can be *implied* from current asset valuations by *calibrating* a pricing model. The former are called *historical*, *statistical*, or, better, *real-world* probabilities. The latter are called *implied probabilities*. The calibration procedure involves using the fundamental theorem of asset pricing to represent prices as discounted expectations of future pay-offs and represents one of the central topics to be discussed in the rest of this chapter.

Definition 1.1. Financial Model *A finite, single-period financial model $\mathcal{M} = (\Omega, \mathcal{A})$ is given by a finite scenario set $\Omega = \{\omega_1, \ldots, \omega_m\}$ and n basic asset price processes for hedging instruments:*

$$\mathcal{A} = \{A_t^1, \ldots, A_t^n; t = 0, T\}. \tag{1.3}$$

Here, A_0^i models the current price of the ith asset at current (or initial) time $t = 0$ and A_T^i is a random variable such that the price at time $T > 0$ of the ith asset in case scenario ω_j occurs is given by $A_T^i(\omega_j)$. The basic asset prices A_t^i, $i = 1, \ldots, n$, are assumed real and positive.

Definition 1.2. Portfolio and Asset *Let $\mathcal{M} = (\Omega, \mathcal{A})$ be a financial model. A portfolio $\boldsymbol{\pi}$ is given by a vector with components $\pi_i \in \mathbb{R}, i = 1, \ldots, n$, representing the positions or holdings in the the family of basic assets with prices $A_t^1, \ldots, A_t^n$. The worth of the portfolio at terminal time T is given by $\sum_{i=1}^n \pi_i A_T^i(\omega)$ given the state or scenario ω, whereas the current price is $\sum_{i=1}^n \pi_i A_0^i$. A portfolio is nonnegative if it gives rise to nonnegative pay-offs under all scenarios, i.e., $\sum_{i=1}^n \pi_i A_T^i(\omega_j) \geq 0, \ \forall j = 1, \ldots, m$. An asset price process $A_t = A_t(\omega)$ (a generic one, not necessarily that of a hedging instrument) is a process of the form*

$$A_t = \sum_{i=1}^n \pi_i A_t^i \tag{1.4}$$

for some portfolio $\boldsymbol{\pi} \in \mathbb{R}^n$.

The modeling assumption behind this definition is that market liquidity is infinite, meaning that asset prices don't vary as a consequence of agents trading them. As we discussed at the start of this chapter, this hypothesis is valid only in case trades are relatively small, for large trades cause market prices to change. In addition, a financial model with infinite liquidity is mathematically consistent only if there are no arbitrage opportunities.

Definition 1.3. Arbitrage: Single-Period Discrete Case *An arbitrage opportunity or arbitrage portfolio is a portfolio $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)$ such that either of the following conditions holds:*
A1. *The current price of $\boldsymbol{\pi}$ is negative, $\sum_{i=1}^n \pi_i A_0^i < 0$, and the pay-off at terminal time T is nonnegative, i.e., $\sum_{i=1}^n \pi_i A_T^i(\omega_j) \geq 0$ for all j states.*
A2. *The current price of $\boldsymbol{\pi}$ is zero, i.e., $\sum_{i=1}^n \pi_i A_0^i = 0$, and the pay-off at terminal time T in at least one scenario ω_j is positive, i.e., $\sum_{i=1}^n \pi_i A_T^i(\omega_j) > 0$ for some jth state, and the pay-off at terminal time T is nonnegative.*

Definition 1.4. Market Completeness *The financial model $\mathcal{M} = (\Omega, \mathcal{A})$ is complete if for all random variables $f_t : \Omega \to \mathbb{R}$, where f_t is a bounded payoff function, there exists an asset price process or portfolio A_t in the basic assets contained in $\mathcal{A}$ such that $A_T(\omega) = f_T(\omega)$ for all scenarios $\omega \in \Omega$.*

This definition essentially states that any pay-off (or state-contingent claim) can be replicated, i.e., is attainable by means of a portfolio consisting of positions in the set of basic assets. If an arbitrage portfolio exists, one says there is *arbitrage*. The first form of arbitrage occurs whenever there exists a trade of negative initial cost at time $t = 0$ by means of which one can form a portfolio that under all scenarios at future time $t = T$ has a nonnegative pay-off. The second form of arbitrage occurs whenever one can perform a trade at zero cost at an initial time $t = 0$ and then be assured of a strictly positive payout at future time T under at least one possible scenario, with no possible downside. In reality, in either case investors would want to perform arbitrage trades and take arbitrarily large positions in the arbitrage portfolios. The existence of these trades, however, infringes on the modeling assumption of infinite liquidity, because market prices would shift as a consequence of these large trades having been placed.

Let's start by considering the simplest case of a single-period economy consisting of only two hedging instruments (i.e., $n = 2$ basic assets) with price processes $A_t^1 = B_t$ and $A_t^2 = S_t$. The scenario set, or sample space, is assumed to consist of only two possible states of the world: $\Omega = \{\omega_+, \omega_-\}$. S_t is the price of a risky asset, which can be thought of as a stock price. The riskless asset is a *zero-coupon bond*, defined as a process B_t that is known to be worth the so-called *nominal* amount $B_T = N$ at time T while at time $t = 0$ has worth

$$B_0 = (1 + rT)^{-1} N. \tag{1.5}$$

Here $r > 0$ is called the *interest rate*. As is discussed in more detail in Chapter 2, interest rates can be defined with a number of different *compounding rules*; the definition chosen here for r corresponds to selecting T itself as the compounding interval, with simple (or discrete) compounding assumed. At current time $t = 0$, the stock has known worth S_0. At a later time $t = T$, two scenarios are possible for the stock. If the scenario ω_+ occurs, then there is an upward move and $S_T = S_T(\omega_+) \equiv S_+$; if the scenario ω_- occurs, there is a downward move and $S_T = S_T(\omega_-) \equiv S_-$, where $S_+ > S_-$. Since the bond is riskless we have $B_T(\omega_+) = B_T(\omega_-) = B_T$. Assume that the real-world probabilities that these events will occur are $p_+ = p \in (0, 1)$ and $p_- = (1 - p)$, respectively.

Figure 1.1 illustrates this simple economy. In this situation, the hypothesis of arbitrage freedom demands that the following strict inequality be satisfied:

$$\frac{S_-}{1 + rT} < S_0 < \frac{S_+}{1 + rT}. \tag{1.6}$$

In fact, if, for instance, one had $S_0 < \frac{S_-}{1+rT}$, then one could make unbounded riskless profits by initially borrowing an arbitrary amount of money and buying an arbitrary number of shares in the stock at price S_0 at time $t = 0$, followed by selling the stock at time $t = T$ at a higher return level than r. Inequality (1.6) is an example of a restriction resulting from the *condition of absence of arbitrage*, which is defined in more detail later.

A *derivative* asset, of worth A_t at time t, is a claim whose pay-off is contingent on future values of risky underlying assets. In this simple economy the underlying asset is the stock. An example is a derivative that pays f_+ dollars if the stock is worth S_+, and f_- otherwise, at final time T: $A_T = A_T(\omega_+) = f_+$ if $S_T = S_+$ and $A_T = A_T(\omega_-) = f_-$ if $S_T = S_-$. Assuming one can take fractional positions, this payout can be statically replicated by means of a portfolio

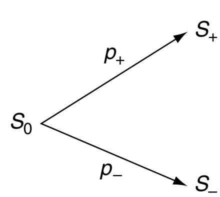

FIGURE 1.1 A single-period model with two possible future prices for an asset S.

consisting of a shares of the stock and b bonds such that the following replication conditions under the two scenarios are satisfied:

$$aS_- + bN = f_-, \tag{1.7}$$

$$aS_+ + bN = f_+. \tag{1.8}$$

The solution to this system is

$$a = \frac{f_+ - f_-}{S_+ - S_-}, \qquad b = \frac{f_- S_+ - f_+ S_-}{N(S_+ - S_-)}. \tag{1.9}$$

The price of the replicating portfolio, with pay-off identical to that of the derivative, must be the price of the derivative asset; otherwise there would be an arbitrage opportunity. That is, one could make unlimited riskless profits by buying (or selling) the derivative asset and, at the same time, taking a short (or long) position in the portfolio at time $t = 0$. At time $t = 0$, the arbitrage-free price of the derivative asset, A_0, is then

$$\begin{aligned} A_0 &= aS_0 + b(1 + rT)^{-1}N \\ &= \left(\frac{S_0 - (1 + rT)^{-1}S_-}{S_+ - S_-}\right) f_+ + \left(\frac{(1 + rT)^{-1}S_+ - S_0}{S_+ - S_-}\right) f_-. \end{aligned} \tag{1.10}$$

Dimensional considerations are often useful to understand the structure of pricing formulas and detect errors. It is important to remember that prices at different moments in calendar time are not equivalent and that they are related by discount factors. The hedge ratios a and b in equation (1.9) are dimensionless because they are expressed in terms of ratios of prices at time T. In equation (1.10) the variables $f_\pm$ and $S_+ - S_-$ are measured in dollars at time T, so their ratio is dimensionless. Both S_0 and the discounted prices $(1 + rT)^{-1}S_\pm$ are measured in dollars at time 0, as is also the derivative price A_0.

Rewriting this last equation as

$$A_0 = (1 + rT)^{-1}\left[\left(\frac{(1 + rT)S_0 - S_-}{S_+ - S_-}\right) f_+ + \left(\frac{S_+ - (1 + rT)S_0}{S_+ - S_-}\right) f_-\right] \tag{1.11}$$

shows that price A_0 can be interpreted as the discounted expected pay-off. However, the probability measure is *not* the real-world one (i.e., not the physical measure P) with probabilities $p_\pm$ for up and down moves in the stock price. Rather, current price A_0 is the discounted expectation of future prices A_T, in the following sense:

$$A_0 = (1 + rT)^{-1}\, E^Q[A_T] = (1 + rT)^{-1}[q_+ A_T(\omega_+) + q_- A_T(\omega_-)] \tag{1.12}$$

under the measure Q with probabilities (strictly between 0 and 1)

$$q_+ = \frac{(1 + rT)S_0 - S_-}{S_+ - S_-}, \qquad q_- = \frac{S_+ - (1 + rT)S_0}{S_+ - S_-}, \tag{1.13}$$

$q_+ + q_- = 1$. The measure Q is called the *pricing measure*. Pricing measures also have other, more specific names. In the particular case at hand, since we are discounting with a constant interest rate within the time interval $[0, T]$, Q is commonly named the *risk-neutral* or *risk-adjusted probability measure*, where $q_\pm$ are so-called risk-neutral (or risk-adjusted) probabilities. Later we shall see that this measure is also the *forward measure*, where the bond price B_t is used as numeraire asset. In particular, by expressing all asset prices relative to (i.e., in units of) the bond price A_t^i/B_t, with $B_T = N$, regardless of the scenario and $B_0/B_T = (1 + rT)^{-1}$, we can hence recast the foregoing expectation as: $A_0 = B_0\, E^Q[A_T/B_T]$. Hence Q corresponds to the forward measure. We can also use as numeraire a discretely compounded *money-market account* having value $(1 + rt)$ (or $(1 + rt)N$). By expressing all asset prices relative to this quantity, it is trivially seen that the corresponding measure is the same as the forward measure in this simple model. As discussed later, the name *risk-neutral measure* shall, however, refer to the case in which the money-market account (to be defined more generally later in this chapter) is used as numeraire, and this measure generally differs from the forward measure for more complex financial models.

Later in this chapter, when we cover pricing in *continuous time*, we will be more specific in defining the terminology needed for pricing under general choices of numeraire asset. We will also see that what we just unveiled in this particularly simple case is a general and far-reaching property: Arbitrage-free prices can be expressed as discounted expectations of future pay-offs. More generally, we will demonstrate that asset prices can be expressed in terms of expectations of *relative asset price processes*. A pricing measure is then a *martingale measure*, under which all relative asset price processes (i.e., relative to a given choice of numeraire asset) are so-called *martingales*. Since our primary focus is on continuous-time pricing models, as introduced later in this chapter, we shall begin to explicitly cover some of the essential elements of martingales in the context of stochastic calculus and continuous-time pricing. For a more complete and elaborate mathematical construction of the martingale framework in the case of discrete-time finite financial models, however, we refer the reader to other literature (for example, see [Pli97, MM03]).

We now extend the pricing formula of equation (1.12) to the case of n assets and m possible scenarios.

Definition 1.5. Pricing Measure *A probability measure $Q = (q_1, \ldots, q_m)$, $0 < q_j < 1$, for the scenario set $\Omega = \{\omega_1, \ldots, \omega_m\}$ is a pricing measure if asset prices can be expressed as follows:*

$$A_0^i \;=\; \alpha\, E^Q[A_T^i] = \alpha \sum_{j=1}^{m} q_j A_T^i(\omega_j) \tag{1.14}$$

for all $i = 1, \ldots, n$ and some real number $\alpha > 0$. The constant α is called the discount factor.

Theorem 1.1. Fundamental Theorem of Asset Pricing (Discrete, single-period case) *Assume that all scenarios in Ω are possible. Then the following statements hold true:*

- *There is no arbitrage if and only if there is a pricing measure for which all scenarios are possible.*
- *The financial model is complete, with no arbitrage if and only if the pricing measure is unique.*

Proof. First, we prove that if a pricing measure $Q = (q_1, \ldots, q_m)$ exists and prices $A_0^i = \alpha\, E^Q[A_T^i]$ for all $i = 1, \ldots, n$, then there is no arbitrage. If $\sum_i \pi_i A_T^i(\omega_j) \geq 0$, for all $\omega_j \in \Omega$, then from equation (1.14) we must have $\sum_i \pi_i A_0^i \geq 0$. If $\sum_i \pi_i A_0^i = 0$, then from equation (1.14) we cannot satisfy the payoff conditions in (A2) of Definition 1.3. Hence there is no arbitrage, for any choice of portfolio $\boldsymbol{\pi} \in \mathbb{R}^n$.

On the other hand, assume that there is no arbitrage. The possible price-payoff $(m + 1)$-tuples

$$\mathcal{P} = \left\{ \left(\sum_{i=1}^{n} \pi_i A_0^i, \sum_{i=1}^{n} \pi_i A_T^i(\omega_1), \ldots, \sum_{i=1}^{n} \pi_i A_T^i(\omega_m) \right), \qquad \boldsymbol{\pi} \in \mathbb{R}^n \right\} \tag{1.15}$$

make up a plane in $\mathbb{R} \times \mathbb{R}^m$. Since there is no arbitrage, the plane $\mathcal{P}$ intersects the octant $\mathbb{R}_+ \times \mathbb{R}_+^m$ made up of vectors of nonnegative coordinates only in the origin. Let $\mathcal{N}$ be the set of all vectors $(-\beta, \gamma_1, \ldots, \gamma_m)$ normal to the plane $\mathcal{P}$ and normalized so that $\beta > 0$. Vectors in $\mathcal{N}$ satisfy the normality condition

$$-\beta\left(\sum_{i=1}^{n} \pi_i A_0^i\right) + \sum_{j=1}^{m} \gamma_j \left(\sum_{i=1}^{n} \pi_i A_T^i(\omega_j)\right) = 0 \tag{1.16}$$

for all portfolios $\boldsymbol{\pi}$.

Next we obtain two Lemmas to complete the proof.

Lemma 1.1. *Suppose the financial model on the scenario set Ω and with instruments $(A^1, \ldots, A^n)$ is arbitrage free and let m be the dimension of the linear space $\mathcal{P}$. If the matrix rank* $\dim \mathcal{P} < m$*, then one can define $l = (m - \dim \mathcal{P})$ price-payoff tuples $(-B_0^k, B_T^k(\omega))$, $k = 1, \ldots, l$, so that the extended financial model with basic assets $(A^1, \ldots, A^n, B^1, \ldots, B^l)$ and scenario set Ω is complete and arbitrage free.*

Proof. The price-payoff tuples $(-B_0^k, B_T^k(\omega_1), \ldots, B_T^k(\omega_l))$ can be found iteratively. Suppose that $l = m - \dim \mathcal{P} > 0$. Then the complement to the linear space $\mathcal{P}$ has dimension $l + 1 \geq 2$. Let $X = (-X_0^k, X_T^k(\omega))$ and $Y = (-Y_0^k, Y_T^k(\omega))$ be two vectors orthogonal to each other and orthogonal to $\mathcal{P}$. Then there is an angle θ such that the vector $B^1 = \cos\theta X + \sin\theta Y$ has at least one strictly positive coordinate and one strictly negative coordinate, i.e., $B^1 \notin \mathbb{R} \times \mathbb{R}_+$. Hence the financial model with instruments $(A^1, \ldots, A^n, B^1)$ is arbitrage free. Iterating the argument, one can complete the market while retaining arbitrage freedom. □

Lemma 1.2. *If markets are complete, the space $\mathcal{N}$ orthogonal to $\mathcal{P}$ is spanned by a vector $(\beta, \gamma_1, \ldots, \gamma_m)$ lying in the main octant $\mathcal{B} = \mathbb{R}_+ \times \mathbb{R}_+^m$ of vectors with strictly positive coordinates.*

Proof. In fact if $\beta = 0$, then $\mathcal{P}$ contains the line $(x, 0, \ldots, 0)$ and all positive payouts would be possible, even for an empty portfolio, which is absurd. It is also absurd that $\gamma_j = 0$, $\forall j$. In fact, in this case, since markets are complete, there is an instrument paying one dollar in case the scenario ω_j occurs and zero otherwise, and since $\gamma_j = 0$, the price of this instrument at time $t = 0$ is zero, which is absurd. □

If markets are not complete, one can still conclude that the set $\mathcal{N}$ contains a vector $(\beta, \gamma_1, \ldots, \gamma_m)$ with strictly positive coordinates. In fact, thanks to Lemma 1.1, one can complete it while preserving arbitrage freedom by introducing auxiliary assets and the normal vector can be chosen to have positive coordinates. Hence, in all cases of π_i values, according to equation (1.16) we have

$$A_0^i = \alpha\, E^Q[A_T^i] = \alpha \sum_{j=1}^{m} q_j A_T^i(\omega_j), \tag{1.17}$$

where Q is the measure with probabilities

$$q_j = \frac{\gamma_j}{\sum_{j=1}^{m} \gamma_j} \tag{1.18}$$

and discount factor

$$\alpha = \beta^{-1} \sum_{j=1}^{m} \gamma_j. \tag{1.19}$$

□

The first project of Part II of this book is a study on single-period arbitrage. We refer the interested reader to that project for a more detailed and practical exposition of the foregoing theory. In particular, the project provides an explicit discussion of a numerical linear algebra implementation for detecting arbitrage in single-period, finite financial models.

Problems

Problem 1. Consider the simple example in Figure 1.1 and assume the interest rate is r. Under what condition is there no arbitrage in the model?

Problem 2. Compute $E^Q[S_T]$ within the single-period two-state model. Explain your result.

Problem 3. Let p_i^0 denote the current price A_0^i of the ith security and denote by $D_{ij} = A_T^i(\omega_j)$ the matrix elements of the $n \times m$ dividend matrix with $i = 1, \ldots, n,\ j = 1, \ldots, m$. Using equation (1.14) with $\alpha = (1 + rT)^{-1}$ show that the risk-neutral expected return on any security A^i is given by the risk-free interest rate

$$E^Q\left[\frac{A_T^i - A_0^i}{A_0^i}\right] = \sum_{j=1}^{m} q_j \left(\frac{D_{ij}}{p_i^0} - 1\right) = rT, \tag{1.20}$$

where q_j are the risk-neutral probabilities.

Problem 4. State the explicit matrix condition for market completeness in the single-period two-state model with the two basic assets as the riskless bond and the stock. Under what condition is this market complete?

Problem 5. *Arrow–Debreu securities* are claims with unit pay-offs in only one state of the world. Assuming a single-period two-state economy, these claims are denoted by $E_\pm$ and defined by

$$E_+(\omega) = \begin{cases} 1, & \text{if } \omega = \omega_+ \\ 0, & \text{if } \omega = \omega_- \end{cases}, \qquad E_-(\omega) = \begin{cases} 0, & \text{if } \omega = \omega_+ \\ 1, & \text{if } \omega = \omega_- \end{cases}.$$

(a) Find exact replicating portfolios $\boldsymbol{\pi}_+ = (a_+, b_+)$ and $\boldsymbol{\pi}_- = (a_-, b_-)$ for E_+ and E_-, respectively. The coefficients a and b are positions in the stock and the riskless bond, respectively.

(b) Letting F_T represent an arbitrary pay-off, find the unique portfolio of Arrow–Debreu securities that replicates F_T.

1.2 Continuous State Spaces

This section, together with the next section, presents a review of basic elements of probability theory for random variables that can take on a continuum of values while emphasizing some of the financial interpretation of mathematical concepts.

Modern probability theory is based on measure theory. Referring the reader to textbook literature for more detailed and exhaustive formal treatments, we will just simply recall here that measure theory deals with the definition of *measurable sets* D, *probability measures* μ, and *integrable functions* $f : D \to \mathbb{R}$ for which one can evaluate expectations as integrals

$$E[f] = \int_D f(\mathbf{x})\mu(d\mathbf{x}). \tag{1.21}$$

In finance, one typically deals with situations where the measurable set $D \subset \mathbb{R}^d$, with integer $d \geq 1$. Realizations of the vector variable $\mathbf{x} \in D$ correspond to scenarios for the risk factors or random variables in a financial model.

Future asset prices are real-valued functions of underlying risk factors $f(\mathbf{x})$ defined for $\mathbf{x} \in D$ and hence themselves define random variables. Probability measures $\mu(d\mathbf{x})$ are often defined as $\mu(d\mathbf{x}) = p(\mathbf{x})d\mathbf{x}$, where $p(\mathbf{x})$ is a real-valued continuous *probability distribution function* that is nonnegative and integrates to 1; i.e.,

$$p(\mathbf{x}) \geq 0, \qquad \int_D p(\mathbf{x})d\mathbf{x} = 1. \tag{1.22}$$

The expectation $E^P\big[f\big]$ of f under the probability measure with p as density is defined by the d-dimensional integral

$$E^P\big[f\big] = \int_D f(\mathbf{x})p(\mathbf{x})d\mathbf{x}. \tag{1.23}$$

The pair $(D, \mu(d\mathbf{x}))$ is called a *probability space*.

In particular, this formalism can also allow for the case of a finite scenario set of vectors $D = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}\}$, as was considered in the previous section. In this case the probability distribution is a sum of Dirac *delta functions*,

$$p(\mathbf{x}) = \sum_{i=1}^{N} p_i \delta(\mathbf{x} - \mathbf{x}^{(i)}). \tag{1.24}$$

As further discussed shortly, a delta function can be thought of as a singular function that is positive, integrates to 1 over all space, and corresponds to the infinite limiting case of a sequence of integrable functions with support only at the origin. Probabilistically, a distribution, such as equation (1.24), which is a sum of delta functions, corresponds to a situation where only the scenarios $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}$ can possibly occur, and they do with probabilities $p_1, \ldots, p_N$. These probabilities must be positive and add up to 1; i.e.,

$$\sum_{i=1}^{N} p_i = 1. \tag{1.25}$$

In the case of a finite scenario set (i.e., a finite set of possible events with finite integer N), the random variable $f = f(\mathbf{x})$ is a function defined on the set of scenarios D, and its expectation under the measure with p as density is given by the finite sum

$$E^P[f] = \sum_{i=1}^{N} p_i f(\mathbf{x}^{(i)}). \tag{1.26}$$

For an infinitely countable set of scenarios, then, the preceding expressions must be considered in the limit $N \to \infty$. Hence in the case of a discrete set of scenarios (as opposed to a continuum) the probability density function collapses into the usual *probability mass function*, as occurs in standard probability theory of discrete-valued random variables.

The Dirac delta function is not an ordinary function in $\mathbb{R}^d$ but, rather, a so-called *distribution*. Mathematically, a distribution is defined through its value when integrated against a smooth function. One can regard $\delta(\mathbf{x} - \mathbf{x}')$, $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^d$, as the limit of an infinitesimally narrow d-dimensional normal distribution:

$$\int_{\mathbb{R}^d} f(\mathbf{x})\delta(\mathbf{x} - \mathbf{x}')d\mathbf{x} = \lim_{\sigma \to 0} \frac{1}{(\sigma\sqrt{2\pi})^d} \int_{\mathbb{R}^d} f(\mathbf{x}) \exp\left(-\frac{|\mathbf{x} - \mathbf{x}'|^2}{2\sigma^2}\right)d\mathbf{x} = f(\mathbf{x}'). \tag{1.27}$$

For example, in one dimension a representation of the delta function is

$$\delta(x-x') = \lim_{\sigma\to 0} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-x')^2/2\sigma^2}. \tag{1.28}$$

Events are modeled as subsets $G \subset D$ for which one can compute the integral that gives the expectation $E^P[\mathbf{1}_G]$. The function $\mathbf{1}_G(\mathbf{x})$ denotes the random variable equal to 1 for $\mathbf{x} \in G$ and to zero otherwise; $\mathbf{1}_G(\mathbf{x})$ is called the *indicator function* of the set G. This expectation is interpreted as the *probability P(G)* that event $G \subset D$ will occur; i.e.,

$$P(G) = E^P[\mathbf{1}_G] = \int_{\mathbb{R}^d} \mathbf{1}_G(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} = \int_G p(\mathbf{x}) d\mathbf{x}. \tag{1.29}$$

Examples of events are subsets, e.g., such as

$$G = \{\mathbf{x} \in D: \ a < f(\mathbf{x}) < b\}, \tag{1.30}$$

with $b > a$ and where f is some function. An important concept associated with events is that of *conditional expectation*. Given a random variable f, the expectation of f conditioned to knowing that event G will occur is

$$E^P\big[f|G\big] = \frac{E^P\big[f \cdot \mathbf{1}_G\big]}{P(G)}. \tag{1.31}$$

Two probability measures $\tilde{\mu}(d\mathbf{x}) = \tilde{p}(\mathbf{x}) d\mathbf{x}$ and $\mu(d\mathbf{x}) = p(\mathbf{x}) d\mathbf{x}$ are said to be equivalent (or absolutely continuous with respect to one another) if they share the same sets of null probability; i.e., $\tilde{\mu} \sim \mu$ if the probability condition $P(G) > 0$ implies $\tilde{P}(G) > 0$, where

$$\tilde{P}(G) = E^{\tilde{P}}[\mathbf{1}_G] = \int_{\mathbb{R}^d} \mathbf{1}_G(\mathbf{x}) \tilde{p}(\mathbf{x}) d\mathbf{x} = \int_G \tilde{p}(\mathbf{x}) d\mathbf{x}, \tag{1.32}$$

with $E^{\tilde{P}}[\,]$ denoting the expectation with respect to the measure $\tilde{\mu}$. When computing the expectation of a real-valued random variable, say, of the general form of a function of a random vector (such functions are further defined in the next section), $f = f(\mathbf{X}) : \mathbb{R}^d \to \mathbb{R}$, it is sometimes useful to switch from one choice of probability measure to another, equivalent one. One can use the following *change of measure* (known as the *Radon–Nikodym theorem*) for computing expectations:

$$E^P[f] = \int_D f(\mathbf{x}) \mu(d\mathbf{x}) = \int_D f(\mathbf{x}) \frac{d\mu}{d\tilde{\mu}}(\mathbf{x}) \tilde{\mu}(d\mathbf{x}) = E^{\tilde{P}}\left[f \frac{d\mu}{d\tilde{\mu}}\right]. \tag{1.33}$$

The nonnegative random variable denoted by $\frac{d\mu}{d\tilde{\mu}}$ is called the *Radon–Nikodym derivative* of μ with respect to $\tilde{\mu}$ (or P w.r.t. $\tilde{P}$). From this result it also follows that $\frac{d\mu}{d\tilde{\mu}} = \left(\frac{d\tilde{\mu}}{d\mu}\right)^{-1}$ and $E^{\tilde{P}}[\frac{d\mu}{d\tilde{\mu}}] = 1$. As will be seen later in the chapter, a more general adaptation of this result for computing certain types of conditional expectations involving martingales will turn out to form one of the basic tools for pricing financial derivatives using changes of numeraire. Another particular example of the use of this change-of-measure technique is in the Monte Carlo estimation of integrals by so-called importance-sampling methods, as described in Chapter 4.

Just as integrals are approximated with arbitrary accuracy by finite integral sums, continuous probability distributions can be approximated by discrete ones. For instance, let $D \subset \mathbb{R}^d$ be a bounded domain and $p(\mathbf{x})$ be a continuous probability density on D and let $\{G_1, \ldots, G_m\}$

be a *partition of D* made up of a family of nonintersecting events $G_i \subset D$ whose union covers the entire state space D and that have the shape of hypercubes. Let p_i be the probability of event G_i under the probability measure with density *p(x)*. Then an approximation for $p(\mathbf{x})$ is

$$p(\mathbf{x}) = \sum_{i=1}^{m} p_i \delta(\mathbf{x} - \mathbf{x}_i), \tag{1.34}$$

where $\mathbf{x}_i$ is the center of the hypercube corresponding to event G_i. Let δ be the volume of the largest hypercube among the cubes in the partition $\{G_1, \ldots, G_m\}$ and let $f(\mathbf{x})$ be a random variable on D. In the limit $\delta \to 0$, as the partition becomes finer and finer, the number of events $m(\delta)$ will diverge to ∞. In this limit, we find

$$E^P[f] = \lim_{\delta \to 0} \sum_{i=1}^{m(\delta)} p_i f(\mathbf{x}_i). \tag{1.35}$$

By using sums as approximations to expectations, which are essentially multidimensional Riemann integrals, one can extend the theorem in the previous section to the case of continuous probability distributions. Consider a single-period financial model with current (i.e., initial) time $t = 0$ and time horizon $t = T$ and with n basic assets whose current prices are A_0^i, $i = 1, \ldots, n$. The prices of these basic assets at time T are indexed by a continuous state space represented by the domain $\Omega \subset \mathbb{R}^d$, and the values of the basic assets are random variables $A_T^i(\mathbf{x})$, with $\mathbf{x} \in \Omega$. That is, the asset prices A_t^i are random variables assumed to take on real positive values, i.e., $A_t^i : \Omega \to \mathbb{R}_+$. Let's denote by $p(\mathbf{x})d\mathbf{x}$ the real-world probability measure in Ω and assume that the measure of all open subsets of Ω is strictly positive. A portfolio is modeled by a vector $\boldsymbol{\pi}$ whose components denote positions or holdings π_i, $i = 1, \ldots, n$, in the basic assets. The definition of arbitrage extends as follows.

Definition 1.6. Nonnegative Portfolio *A portfolio is nonnegative if it gives rise to nonnegative expected pay-offs under almost all events $G \subset \Omega$ of nonzero probability, i.e., such that*

$$E^P\left[\sum_{i=1}^{n} \pi_i \, A_T^i(\mathbf{x}) \middle| \mathbf{x} \in G\right] \geq 0. \tag{1.36}$$

Definition 1.7. Arbitrage: Single-Period Continuous Case *The market admits arbitrage if either of the following conditions holds:*
A1. *There is a nonnegative portfolio $\boldsymbol{\pi}$ of negative initial price $\sum_{i=1}^{n} \pi_i A_0^i < 0$.*
A2. *There is a nonnegative portfolio of zero initial cost, $\sum_{i=1}^{n} \pi_i A_0^i = 0$, for which the expected payoff is strictly positive, i.e., $E^P\left[\sum_{i=1}^{n} \pi_i A_T^i\right] > 0$.*

Definition 1.8. Pricing Measure: Single-Period Continuous Case[3] *A probability measure Q of density $q(\mathbf{x})d\mathbf{x}$ on D is a pricing measure if all asset prices at current time $t = 0$ can be expressed as follows:*

$$A_0^i = \alpha E^Q[f_i] = \alpha \int_{\Omega} f_i(\mathbf{x}) q(\mathbf{x}) d\mathbf{x} \tag{1.37}$$

for some real number $\alpha > 0$. The constant α is called the discount factor. The functions $f_i(\boldsymbol{x}) = A_T^i(\boldsymbol{x})$ are payoff functions for a given state or scenario $\boldsymbol{x}$.

[3] Later we relate such pricing measures to the case of arbitrary choices of numeraire asset wherein the pricing formula involves an expectation of asset prices relative to the chosen numeraire asset price. Changes in numeraire correspond to changes in the probability measure.

Market completeness is defined in a manner similar to that in the single-period discrete case of the previous section. From the foregoing definitions of arbitrage and pricing measure we then have the following result, whose proof is left as an exercise.

Theorem 1.2. Fundamental Theorem of Asset Pricing (Continuous Single-Period Case) *Assume that all scenarios in Ω are possible. Then the following statements hold true.*

- *There is no arbitrage if and only if there is a pricing measure for which all scenarios are possible.*
- *If the linear span of the set of basic instruments A_T^i, $i = 1, \ldots, n$, is complete and there is no arbitrage, then there is a unique pricing measure Q consistent with the prices A_0^i of the reference assets at current time $t = 0$.*

The single-period pricing formalism can also be extended to the case of a multiperiod discrete-time financial model, where trading is allowed to take place at a finite number of intermediate dates. This feature gives rise to dynamic trading strategies, with portfolios in the basic assets being rebalanced at discrete points in time. The foregoing definitions and notions of arbitrage and asset pricing must then be modified and extended substantially. Rather than present the theory for such discrete-time models, we shall instead introduce more important theoretical tools in the following sections that will allow us ultimately to consider continuous-time financial models. Multiperiod discrete-time (continuous-state-space) models can then be obtained, if desired, as special cases of the continuous models via a discretization of time. A further discretization of the state space leads to discrete-time multiperiod finite financial models.

1.3 Multivariate Continuous Distributions: Basic Tools

Marginal probability distributions arise, for instance, when one is computing expectations on some reduced subspace of random variables. Consider, for example, a set of continuous random variables that can be separated or grouped into two random vector spaces $\mathbf{X} = (X_1, \ldots, X_m)$ and $\mathbf{Y} = (Y_1, \ldots, Y_{n-m})$ that can take on values $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$ and $\mathbf{y} = (y_1, \ldots, y_{n-m}) \in \mathbb{R}^{n-m}$, respectively, with $1 \leq m < n$, $n \geq 2$. The function $p(\mathbf{x}, \mathbf{y})$ is the joint probability density or *probability distribution function (pdf)* in the product space $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$. The integral

$$p_y(\mathbf{y}) \equiv \int_{\mathbb{R}^m} p(\mathbf{x}, \mathbf{y}) d\mathbf{x} \tag{1.38}$$

defines a marginal density $p_y(\mathbf{y})$. This function describes a probability density in the subspace of random vectors $\mathbf{Y} \in \mathbb{R}^{n-m}$ and integrates to unity over $\mathbb{R}^{n-m}$. The *conditional density function*, denoted by $p(\mathbf{x}|\mathbf{Y} = \mathbf{y}) \equiv p(\mathbf{x}|\mathbf{y})$ for the random vector $\mathbf{X}$, is defined on the subspace of $\mathbb{R}^m$ (for a given vector value $\mathbf{Y} = \mathbf{y}$) and is defined by the ratio of the joint probability density function and the marginal density function for the random vector $\mathbf{Y}$ evaluated at $\mathbf{y}$:

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x}, \mathbf{y})}{p_y(\mathbf{y})}, \tag{1.39}$$

assuming $p_y(\mathbf{y}) \neq 0$. From the foregoing two relations it is simple to see that, for any given $\mathbf{y}$, the conditional density also integrates to unity over $\mathbf{x} \in \mathbb{R}^m$.

Conditional distributions play an important role in finance and pricing theory. As we will see later, derivative instruments can be priced by computing conditional expectations. Assuming a conditional distribution, the *conditional expectation* of a continuous random variable $g = g(\mathbf{X}, \mathbf{Y})$, given $\mathbf{Y} = \mathbf{y}$, is defined by

$$E[g|\mathbf{Y} = \mathbf{y}] = \int_{\mathbb{R}^m} g(\mathbf{x}, \mathbf{y}) p(\mathbf{x}|\mathbf{y}) d\mathbf{x}. \tag{1.40}$$

Given any two continuous random variables X and Y, then $E[X|Y = y]$ is a number while $E[X|Y]$ is itself a random variable as Y is random, i.e., has not been fixed. We then have the following property that relates unconditional and conditional expectations:

$$E[X] = E\big[E[X|Y]\big] = \int_{-\infty}^{\infty} E[X|Y = y] p_y(y) dy. \tag{1.41}$$

This property is useful for computing expectations by conditioning. More generally, for a random variable given by the function $g = g(\mathbf{X}, \mathbf{Y})$ we have the property

$$\begin{aligned} E[g] &= \int_{\mathbb{R}^{n-m}} \int_{\mathbb{R}^m} g(\mathbf{x}, \mathbf{y}) p(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \int_{\mathbb{R}^{n-m}} \left[\int_{\mathbb{R}^m} g(\mathbf{x}, \mathbf{y}) p(\mathbf{x}|\mathbf{y}) d\mathbf{x} \right] p_y(\mathbf{y}) d\mathbf{y} \\ &= \int_{\mathbb{R}^{n-m}} E[g|\mathbf{Y} = \mathbf{y}] p_y(\mathbf{y}) d\mathbf{y} = E\big[E[g|\mathbf{Y}]\big]. \end{aligned} \tag{1.42}$$

Functions of random variables, such as $g(\mathbf{X}, \mathbf{Y})$, are of course also random variables. In general, the pdf of a random variable given by a mapping $f = f(\mathbf{X}) : \mathbb{R}^n \to \mathbb{R}$ is the function $p_f : \mathbb{R} \to \mathbb{R}$,

$$p_f(\xi) = \lim_{\delta\xi \to 0} \frac{\mathrm{P}\big(f(\mathbf{X}) \in [\xi, \xi + \delta\xi)\big)}{(\delta\xi)}, \tag{1.43}$$

defined on some open or closed interval between a and b. This interval may be finite or infinite; some examples are $\xi \in [0, 1]$, $[0, \infty)$, and $(-\infty, \infty)$. The *cumulative distribution function (cdf)* C_f for the random variable f is defined as

$$C_f(z) = \int_a^z p_f(\xi) d\xi \tag{1.44}$$

and gives the probability $P(a \le f \le z)$, with $C_f(b) = 1$. Let us consider another independent real-valued random variable $g \in (c, d)$, where *(c,d)* is generally any other interval. We recall that any two random variables f and g are independent if the joint pdf (or cdf) of f and g is given by the product of the respective marginal pdfs (or cdfs). The sum of two independent random variables f and g is again a random variable $h = f + g$. The cumulative distribution function, denoted by C_h, for the random variable h is given by the convolution integral

$$\begin{aligned} C_h(\zeta) &= \iint\limits_{\xi+\eta\le\zeta} p_f(\xi) p_g(\eta) d\xi d\eta \\ &= \int_a^b p_f(\xi) C_g(\zeta - \xi) d\xi = \int_c^d p_g(\eta) C_f(\zeta - \eta) d\eta, \end{aligned} \tag{1.45}$$

where p_g and C_g are the density and cumulative distribution functions, respectively, for the random variable g. By differentiating the cumulative distribution function we find the density function for the variable h:

$$p_h(\zeta) = \int_a^b p_f(\xi) p_g(\zeta - \xi) d\xi = \int_c^d p_g(\eta) p_f(\zeta - \eta) d\eta. \tag{1.46}$$

The preceding formulas are sometimes useful because they provide the cumulative (or density) functions for a sum of two independent random variables as convolution integrals of the separate density and cumulative functions.

The definition for cumulative distribution functions extends into the multivariate case in the obvious manner. Given a pdf $p : \mathbb{R}^n \to \mathbb{R}$ for $\mathbb{R}^n$-valued random vectors $\mathbf{X} = (X_1, \ldots, X_n)$, the corresponding cdf is the function $C_p : \mathbb{R}^n \to \mathbb{R}$ defined by the joint probability

$$C_p(\mathbf{x}) = P(X_1 \leq x_1, \ldots, X_n \leq x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} p(\mathbf{x}') d\mathbf{x}'. \tag{1.47}$$

We recall that any two random variables X_i and X_j $(i \neq j)$ are *independent* if the joint probability $P(X_i \leq a, X_j \leq b) = P(X_i \leq a) P(X_j \leq b)$ for all $a, b \in \mathbb{R}$, i.e., if the events $\{X_i \leq a\}$ and $\{X_j \leq b\}$ are independent. Hence, for two independent random variables the joint cdf and joint pdf are equal to the product of the marginal cdf and marginal pdf, respectively: $p(x_i, x_j) = p_i(x_i) p_j(x_j)$ and $C_p(x_i, x_j) = C_i(x_i) C_j(x_j)$.

Another useful formula for multivariate distributions is the relationship between probability densities (within the same probability measure, say, $\mu(d\mathbf{x})$) expressed on different variable spaces or coordinate variables. That is, if $p(\mathbf{x})$ and $p_{\tilde{\mathbf{X}}}(\tilde{\mathbf{x}})$ represent probability densities on n-dimensional real-valued vector spaces $\mathbf{x}$ and $\tilde{\mathbf{x}}$, respectively and the two spaces are related by a one-to-one continuously differentiable mapping $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}(\mathbf{x})$, then

$$p(\mathbf{x}) = p_{\tilde{\mathbf{X}}}(\tilde{\mathbf{x}}) \left| \frac{d\tilde{\mathbf{x}}}{d\mathbf{x}} \right|, \tag{1.48}$$

where $\frac{d\tilde{\mathbf{x}}}{d\mathbf{x}}$ is the Jacobian matrix of the invertible transformation $\mathbf{x} \to \tilde{\mathbf{x}}$. The notation $|\mathbf{M}|$ refers to the determinant of a matrix $\mathbf{M}$.

A probability distribution that plays a distinguished role is the n-dimensional *Gaussian* (or *normal*) distribution, with *mean* (or average) vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$, defined on $\mathbf{x} \in \mathbb{R}^n$ as follows:

$$p(\mathbf{x}; \boldsymbol{\mu}, \mathbf{C}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{C}|}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}) \cdot \mathbf{C}^{-1} \cdot (\mathbf{x} - \boldsymbol{\mu}) \right). \tag{1.49}$$

The shorthand notation $\mathbf{x} \sim N_n(\boldsymbol{\mu}, \mathbf{C})$ is also used to denote the values of an n-dimensional random vector with components $x_1, \ldots, x_n$ that are obtained by sampling with distribution $p(\mathbf{x}; \boldsymbol{\mu}, \mathbf{C})$. $\mathbf{C} = (C_{ij})$ is called *covariance matrix* and enjoys the property of being *positive definite*, i.e., is such that the inner product $(\mathbf{x}, \mathbf{C}\mathbf{x}) \equiv \mathbf{x} \cdot (\mathbf{C}\mathbf{x}) > 0$ for all real vectors $\mathbf{x}$, and $C_{ij} = C_{ji}$. It follows that the cdf of the n-dimensional multivariate normal random vector is defined by the n-dimensional Gaussian integral

$$\Phi_n(\mathbf{x}; \boldsymbol{\mu}, \mathbf{C}) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} p(\mathbf{x}'; \boldsymbol{\mu}, \mathbf{C}) d\mathbf{x}'. \tag{1.50}$$

A particularly important special case of equation (1.50) for $n = 1$ is the univariate standard normal cdf (i.e., $\Phi_1(x; 0, 1)$), defined by

$$N(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy. \tag{1.51}$$

The mean of a random vector $\mathbf{X}$ with given pdf $p(\mathbf{x})$, is defined by the components

$$\mu_i = E[X_i] = \int_{\mathbb{R}^n} x_i p(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}} x p_i(x) dx, \tag{1.52}$$

and the covariance matrix elements are defined by the expectations

$$C_{ij} \equiv \mathrm{Cov}(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)] = \int_{\mathbb{R}^n} (x_i - \mu_i)(x_j - \mu_j) p(\mathbf{x}) d\mathbf{x}, \tag{1.53}$$

for all $i, j = 1, \ldots, n$. The *standard deviation* of the random variable X_i is defined as the square root of the variance:

$$\sigma_i \equiv \sqrt{\mathrm{Var}(X_i)} = \sqrt{E[(X_i - \mu_i)^2]}, \tag{1.54}$$

and the *correlation* between two random variables X_i and X_j is defined as follows:

$$\rho_{ij} \equiv \mathrm{Corr}(X_i, X_j) = \frac{C_{ij}}{\sigma_i \sigma_j}. \tag{1.55}$$

Since $\sqrt{C_{ii}} = \sigma_i$, the correlation matrix has a unit diagonal, i.e., $\rho_{ii} = 1$. As well, they obey the inequality $|\rho_{ij}| \le 1$ (see Problem 1 of this section). For random variables that may be positively or negatively correlated (e.g., as is the case for different stock returns) it follows that

$$-1 \le \rho_{ij} \le 1. \tag{1.56}$$

In the particular case of a multivariate normal distribution with positive definite covariance matrix as in equation (1.49), the strict inequalities $-1 < \rho_{ij} < 1$ hold.

The main property of normal distributions is that the convolution of two normal distributions is also normal. A random variable that is a sum of random normal variables is, therefore, also normally distributed (see Problem 2). Because of this property, multivariate normal distributions can be regarded as affine transformations of standard normal distributions with $\boldsymbol{\mu} = \mathbf{0}_{n\times 1}$ and $\mathbf{C} = \mathbf{I}_{n\times n}$ (the identity matrix). Consider the vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ of independent standard normal variables with zero mean and unit covariance, i.e., with probability density

$$p(\boldsymbol{\xi}) = \prod_{i=1}^{n} \frac{e^{-\xi_i^2/2}}{\sqrt{2\pi}}. \tag{1.57}$$

If $\mathbf{L} = (L_{ij})$, is an n-dimensional matrix, then the random vector $\mathbf{X} = \boldsymbol{\mu} + \mathbf{L}\boldsymbol{\xi}$ is normally distributed with mean $\boldsymbol{\mu}$ and covariance $\mathbf{C} = \mathbf{L}\mathbf{L}^\dagger$, $\dagger \equiv$ matrix transpose. Indeed, taking expectations over the components gives

$$E[X_i] = E\left[\mu_i + \sum_{j=1}^{n} L_{ij}\xi_j\right] = \mu_i, \tag{1.58}$$

and

$$E[(X_i - \mu_i)(X_j - \mu_j)] = E\left[\left(\sum_{k=1}^{n} L_{ik}\xi_k\right)\left(\sum_{l=1}^{n} L_{jl}\xi_l\right)\right] = \sum_{k=1}^{n} L_{ik}L_{jk} = C_{ij}. \tag{1.59}$$

Here we have used $E[\xi_i] = 0$ and $E[\xi_i \xi_j] = \delta_{ij}$, where δ_{ij} is Kronecker's delta, with value 1 if $i = j$ and zero otherwise.

Conversely, given a positive definite matrix $\mathbf{C}$, one can show that there is a lower triangular matrix $\mathbf{L} = (L_{ij})$ with $L_{ij} = 0$ if $j > i$, such that $\mathbf{C} = \mathbf{L}\mathbf{L}^\dagger$. The matrix $\mathbf{L}$ can be evaluated with a procedure known as *Cholesky factorization*. As discussed later in the book, this algorithm is at the basis of Monte Carlo methods for generating scenarios obeying a multivariate normal distribution with a given covariance matrix.

A special case of a multivariate normal is the bivariate distribution defined for $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$:

$$p(x_1, x_2; \mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \frac{e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - 2\rho\frac{(x_1-\mu_1)}{\sigma_1}\frac{(x_2-\mu_2)}{\sigma_2}\right]}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}.$$

The parameters μ_i and $\sigma_i > 0$ are the mean and the standard deviation of X_i, $i = 1, 2$, respectively, and ρ $(-1 < \rho < 1)$ is the correlation between X_1 and X_2, i.e., $\rho = \rho_{12} = C_{12}/\sigma_1\sigma_2$. In this case the covariance matrix is

$$\mathbf{C} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \tag{1.60}$$

and the lower Cholesky factorization of $\mathbf{C}$ is given by

$$\mathbf{L} = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{pmatrix}. \tag{1.61}$$

The correlation matrix is simply

$$\boldsymbol{\rho} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \tag{1.62}$$

with Cholesky factorization $\boldsymbol{\rho} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\dagger$,

$$\boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}. \tag{1.63}$$

The covariance matrix has inverse

$$\mathbf{C}^{-1} = \frac{1}{(1-\rho^2)} \begin{pmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{pmatrix}. \tag{1.64}$$

Conditional and marginal densities of the bivariate distribution are readily obtained by integrating over one of the variables in the foregoing joint density (see Problem 3).

For multivariate normal distributions one has the following general result, which we state without proof.

Proposition. *Consider the random vector* $\mathbf{X} \in \mathbb{R}^n$ *with partition* $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$, $\mathbf{X}_1 \in \mathbb{R}^m$, $\mathbf{X}_2 \in \mathbb{R}^{n-m}$ *with* $1 \le m < n$, $n \ge 2$. *Let* $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{C})$ *with mean* $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$ *and* $n \times n$ *covariance*

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix}$$

with nonzero determinant $|\mathbf{C}_{22}| \neq 0$, *where* $\mathbf{C}_{11}$ *and* $\mathbf{C}_{22}$ *are* $m \times m$ *and* $(n-m) \times (n-m)$ *covariance matrices of* $\mathbf{X}_1$ *and* $\mathbf{X}_2$, *respectively, and* $\mathbf{C}_{12} = \mathbf{C}_{21}^{\dagger}$ *is the* $m \times (n-m)$ *cross-covariance matrix of the two subspace vectors. The conditional distribution of* $\mathbf{X}_1$, *given* $\mathbf{X}_2 = \mathbf{x}_2$, *is the m-dimensional normal density with mean* $\tilde{\boldsymbol{\mu}} = \boldsymbol{\mu}_1 + \mathbf{C}_{12}\mathbf{C}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$ *and covariance* $\tilde{\mathbf{C}} = \mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_{21}$, *i.e.,* $\mathbf{x}_1 \sim N_m(\tilde{\boldsymbol{\mu}}, \tilde{\mathbf{C}})$ *conditional on* $\mathbf{X}_2 = \mathbf{x}_2$.

A relatively simple proof of this result follows by application of known identities for partitioned matrices. This result is useful in manipulating multidimensional integrals involving normal distributions.

In deriving analytical properties associated with expectations or conditional expectations of random variables, the concept of a *characteristic function* is useful. Given a pdf $p : \mathbb{R}^n \to \mathbb{R}$ for a continuous random vector $\mathbf{X} = (X_1, \ldots, X_n)$, the (joint) characteristic function is the function $\phi_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\phi_{\mathbf{X}}(\mathbf{u}) = E[e^{i\mathbf{u}\cdot\mathbf{X}}] = \int_{\mathbb{R}^n} e^{i\mathbf{u}\cdot\mathbf{x}} p(\mathbf{x}) d\mathbf{x}, \tag{1.65}$$

where $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$, $i \equiv \sqrt{-1}$. Since $\phi_{\mathbf{X}}$ is the Fourier transform of p, then from the theory of Fourier integral transforms we know that the characteristic function gives a complete characterization of the probabilitic laws of $\mathbf{X}$, equivalently as p does. That is, any two random variables having the same characteristic function are identically distributed; i.e., the characteristic function uniquely determines the distribution. From the definition we observe that $\phi_{\mathbf{X}}$ is always a well-defined continuous function, given that p is a bonafide distribution. Evaluating at the origin gives $\phi_{\mathbf{X}}(\mathbf{0}) = E[1] = 1$. The existence of derivatives $\partial^k \phi_{\mathbf{X}}(\mathbf{0})/\partial u_i^k$, $k \geq 1$ is dependent upon the existence of the respective *moments* of the random variables X_i. The kth moment of a single random variable $X \in \mathbb{R}$ is defined by

$$m_k = E[X^k] = \int_{-\infty}^{\infty} x^k p(x) dx, \tag{1.66}$$

while the kth centered moment is defined by

$$\mu^{(k)} = E[(X-\mu)^k] = \int_{-\infty}^{\infty} (x-\mu)^k p(x) dx, \tag{1.67}$$

$\mu = E[X]$, $k \geq 1$. [Note: for $X = X_i$ then $p \to p_i$ is the ith marginal pdf, $\mu \to \mu_i = E[X_i]$, $\mu^{(k)} \to \mu_i^{(k)} = E[(X_i - \mu_i)^k]$, etc.] From these integrals we thus see that the existence of the moments depends on the decay behavior of p at the limits $x \to \pm\infty$. For instance, a distribution that exhibits asymptotic decay at least as fast as a decaying exponential has finite moments to all orders. Obvious examples of these include the distributions of normal, exponential, and uniform random variables. In contrast, distributions that decay as some polynomial to a negative power may, at most, only possess a number of finite moments. A classic case is the Student t distribution with integer d degrees of freedom, which can be shown to possess only moments up to order d. This distribution is discussed in Chapter 4 with respect to modeling risk-factor return distributions.

The moments can be obtained from the derivatives of $\phi_{\mathbf{X}}$ at the origin. However, it is a little more convenient to work directly with the *moment-generating function* (mgf). The (joint) moment-generating function is given by

$$M_{\mathbf{X}}(\mathbf{u}) = E[e^{\mathbf{u}\cdot\mathbf{X}}] = \int_{\mathbb{R}^n} e^{\mathbf{u}\cdot\mathbf{x}} p(\mathbf{x}) d\mathbf{x}. \tag{1.68}$$

If the mgf exists (which is not always true), then it is related to the characteristic function: $M_{\mathbf{X}}(\mathbf{u}) = \phi_{\mathbf{X}}(-i\mathbf{u})$. It can be shown that if $E[|X|^r] < \infty$, then $M_X(u)$ (and $\phi_X(u)$) has continuous rth derivative at $u = 0$ with moments given by

$$m_k = E[X^k] = \frac{d^k M_X(0)}{du^k} = (-i)^k \frac{d^k \phi_X(0)}{du^k}, \qquad k = 1, \dots, r. \tag{1.69}$$

Hence, a random variable X has finite moments of all orders when $M_X(u)$ (or $\phi_X(u)$) is continuously differentiable to any order with $m_k = M_X^{(k)}(0) = (-i)^k \phi_X^{(k)}(0)$, $k = 1, \dots$.

Given two independent random variables X and Y, the characteristic function of the sum $X + Y$ simplifies into a product of functions: $\phi_{X+Y}(u) = E[e^{iu(X+Y)}] = E[e^{iuX}]E[e^{iuY}] = \phi_X(u)\phi_Y(u)$. Hence for $Z = \sum_{i=1}^n X_i$ we have $\phi_Z(u) = \Pi_{i=1}^n \phi_{X_i}(u)$ if all X_i are independent. Characteristic functions or mgfs can be obtained in analytically closed form for various common distributions.

Problems

Problem 1. Make use of equations (1.53) and (1.54) and the Schwarz inequality,

$$\left(\int_{\mathbb{R}^n} f(\mathbf{x}) g(\mathbf{x}) d\mathbf{x} \right)^2 \leq \left(\int_{\mathbb{R}^n} (f(\mathbf{x}))^2 d\mathbf{x} \right) \left(\int_{\mathbb{R}^n} (g(\mathbf{x}))^2 d\mathbf{x} \right), \qquad \mathbf{x} \in \mathbb{R}^n, \tag{1.70}$$

to demonstrate the inequality $|C_{ij}| \leq \sigma_i \sigma_j$, hence $|\rho_{ij}| \leq 1$.

Problem 2. Consider two independent normal random variables X and Y with probability distributions

$$p_x(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-(x-\mu_x)^2/2\sigma_x^2} \qquad \text{and} \qquad p_y(y) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-(y-\mu_y)^2/2\sigma_y^2}, \tag{1.71}$$

respectively. Use convolution to show that $Z = X + Y$ is a normal random variable with probability distribution

$$p_z(z) = \frac{1}{\sigma_z \sqrt{2\pi}} e^{-(z-\mu_z)^2/2\sigma_z^2}, \tag{1.72}$$

where $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$ and $\mu_z = \mu_x + \mu_y$.

Problem 3. Show that the joint density function for the bivariate normal has the form

$$\begin{aligned} p(x, y; \mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = {} & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-(y-\mu_2)^2/2\sigma_2^2} \\ & \exp\left[-\frac{1}{2\sigma_1^2(1-\rho^2)} \left(x - \mu_1 - \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2) \right)^2 \right], \end{aligned} \tag{1.73}$$

and thereby obtain the marginal and conditional distributions:

$$p_Y(Y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-(Y-\mu_2)^2/2\sigma_2^2}, \tag{1.74}$$

$$p(x|Y) = \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} \exp\left[-\frac{1}{2(1-\rho^2)\sigma_1^2} \left[x - \mu_1 - \rho\frac{\sigma_1}{\sigma_2}(Y - \mu_2) \right]^2 \right]. \tag{1.75}$$

Verify that this same result follows as a special case of the foregoing proposition.

Problem 4. Find the moment-generating function for the following distributions:

(a) The uniform distribution on the interval *(a,b)* with pdf: $p(x) = (b-a)^{-1}\mathbf{1}_{x\in(a,b)}$.
(b) The exponential distribution with parameter $\lambda > 0$ and pdf: $p(x) = \lambda e^{-\lambda x}\mathbf{1}_{x\geq 0}$.
(c) The gamma distribution with parameters (n, λ), $n = 1, 2, \ldots, \lambda > 0$, and pdf: $p(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{n-1}}{(n-1)!}\mathbf{1}_{x\geq 0}$.

By differentiating the mgf, obtain the mean and variance of the random variable X for each distribution (a)–(c).

Problem 5. Obtain the moment-generating function for:

(a) The multivariate normal with density given by equation (1.49).
(b) The chi-squared random variable with n degrees of freedom: $Y = \sum_{i=1}^{n} Z_i^2$, where $Z_i \sim N(0, 1)$.

Problem 6. Rederive the result in problem 2 using an argument based solely on moment-generating functions.

Problem 7. Consider two independent exponential random variables X_1 and X_2 with respective parameters λ_1 and λ_2, $\lambda_1 \neq \lambda_2$. Find the pdf for $X_1 + X_2$ and the probability $P(X_1 < X_2)$. Hint: Use convolution and conditioning, respectively.

1.4 Brownian Motion, Martingales, and Stochastic Integrals

A particularly important example of a multivariate normal distribution is provided by a random path evaluated at a sequence of dates in the future. Consider a time interval $[0, t] = [t_0 = 0, t_1, \ldots, t_N = t]$, and subdivide it into $N \geq 1$ subintervals $[t_i, t_{i+1}]$ of length $\delta t_i = t_{i+1} - t_i$, $i = 0, \ldots, N-1$. The path points (t, x_t) are defined for all $t = t_i$ by means of the recurrence relation

$$x_{t_{i+1}} = x_{t_i} + \mu(t_i)\delta t_i + \sigma(t_i)\delta W_{t_i}, \tag{1.76}$$

where the *increments* $\delta W_{t_i} = W_{t_i+1} - W_{t_i}$ are assumed uncorrelated (independent) normal random variables with probability density at $\delta W_{t_i} = \delta w_i$:

$$p_i(\delta w_i) = \frac{1}{\sqrt{2\pi\delta t_i}} e^{-(\delta w_i)^2/2\delta t_i}. \tag{1.77}$$

Since the increments are assumed independent, the joint pdf for all increments is

$$p(\delta w_0, \ldots, \delta w_{N-1}) = \prod_{i=0}^{N-1} p_i(\delta w_i). \tag{1.78}$$

This gives rise to two important unconditional expectations:

$$E\big[\delta W_{t_i}\delta W_{t_j}\big] = \delta_{ij}\delta t_i \qquad E\big[\delta W_{t_i}\big] = 0. \tag{1.79}$$

By usual convention we fix $W_0 = 0$. The joint pdf for the random variables $W_{t_1}, \ldots, W_{t_N}$ representing the probability density at the path points $W_{t_i} = w_i$ ($w_0 = 0$) is then also a multivariate Gaussian function, which is obtained by simply setting $\delta w_i = w_{i+1} - w_i$ in equation (1.78). The set of real-valued random variables $(W_{t_i})_{i=0,\dots,N}$ therefore represents the time-discretized *standard Brownian motion* (or *Wiener process*) at arbitrary discrete points in time. Iterating equation (1.76) gives

$$x_t = x_0 + \sum_{j=0}^{N-1} \left[\mu(t_j)\delta t_j + \sigma(t_j)\delta W_{t_j}\right], \tag{1.80}$$

where $x_{t_N} = x_t$ and $x_{t_0} = x_0$. The random variable x_t is normal with mean

$$E_0[x_t] = x_0 + \sum_{i=0}^{N-1} \mu(t_i)\delta t_i \tag{1.81}$$

and variance

$$E_0[(x_t - E_0[x_t])^2] = E_0\left[\left(\sum_{i=0}^{N-1} \sigma(t_i)\delta W_{t_i}\right)^2\right] = \sum_{i=0}^{N-1} \sigma(t_i)^2 \delta t_i. \tag{1.82}$$

Note: We use $E_0[\,]$ to denote the expectation conditional only on the value of paths being fixed at initial time; i.e., $x_{t_0} = x_0 =$ fixed value. This is hence an unconditional expectation with respect to path values at any later time $t > 0$. Later, we will at times simply use the unconditional expectation $E[\,]$ to denote $E_0[\,]$. Sample paths of a process with zero mean and constant volatility are displayed in Figure 1.2.

Typical stochastic processes in finance are meaningful if time is discretized. The choice of the elementary unit of time is part of the modeling assumptions and depends on the applications at hand. In pricing theory, the natural elementary unit is often one day but can also be one week, one month as well as five minutes or one tick, depending on the objective. The mathematical theory, however, simplifies in the *continuous-time limit*, where the elementary time is infinitesimal with respect to the other time units in the problem, such as

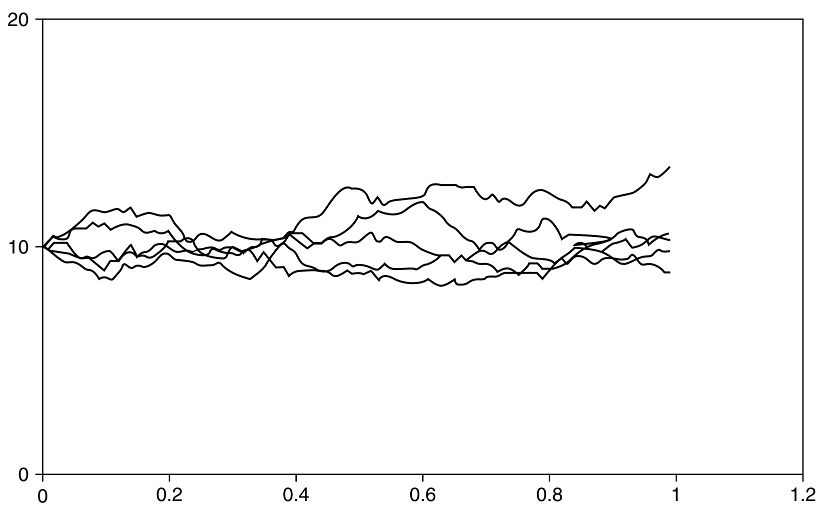

FIGURE 1.2 A simulation of five stochastic paths using equation (1.76), with $x_0 = 10$, constant $\mu(t) = 0.1$, $\sigma(t) = 0.2$, $N = 100$, and time steps $\delta t_i = 0.01$.

option maturities and cash flow periods. Mathematically, one can construct continuous-time processes by starting from a sequence of approximating processes defined for discrete-time values $i\delta t$, $i = 0, \ldots, N$, and then pass to the limit as $\delta t \to 0$. More precisely, one can define a continuous-time process in an interval $[t_0, t_N]$ by subdividing it into N subintervals of equal length, defining a discrete time process $x_t^N \equiv x_{t_N}$ and then compute the limit

$$x_t = \lim_{N\to\infty} x_t^N \tag{1.83}$$

by assuming that the discrete-time process x_t^N is constant over the partition subintervals. The elementary increments $\delta x_t = x_{t+\delta t} - x_t$ are random variables that obviously tend to zero as $\delta t \to 0$, but which are still meaningful in this case. The convention is to denote these increments as dx in the limit $\delta t \to 0$ and to consider the straight d as a reminder that, at the end of the calculations, one is ultimately interested in the limit as $\delta t \to 0$.[4]

The continuous-time limit is obtained by holding the terminal time $t = t_N$ fixed and letting $N \to \infty$, i.e.,

$$E_0[x_t] \equiv \lim_{N\to\infty} E_0[x_t^N] = x_0 + \int_0^t \mu(\tau)d\tau \equiv x_0 + \bar{\mu}(t)t, \tag{1.84}$$

and

$$E_0[(x_t - E_0[x_t])^2] \equiv \lim_{N\to\infty} E_0[(x_t^N - E_0[x_t^N])^2] = \int_0^t \sigma^2(\tau)d\tau \equiv \bar{\sigma}(t)^2 t, \tag{1.85}$$

where we introduced the time-averaged drift $\bar{\mu} = \bar{\mu}(t)$ and volatility $\bar{\sigma} = \bar{\sigma}(t)$ over the time period $[0, t]$, $t \in \mathbb{R}_+$. Since x_t is normally distributed, we finally arrive at the *transition probability density* for a stochastic path to attain value x_t at time t, given an initially known value x_0 at time $t = 0$:

$$p(x_t, x_0; t) = \frac{1}{\bar{\sigma}\sqrt{2\pi t}} \exp\left(- \frac{(x_t - (x_0 + \bar{\mu}t))^2}{2\bar{\sigma}^2 t} \right). \tag{1.86}$$

This density, therefore, gives the distribution (conditional on a starting value x_0) for a process with continuous motion on the entire real line $x_t \in (-\infty, \infty)$ with constant drift and volatility. [Note: x_0, x_t are real numbers (not random) in equation (1.86).]

A *Markov chain* is a discrete-time stochastic process such that for all times $t \in \mathcal{T}$ the increments $x_{t+\delta t} - x_t$ are random variables independent of x_t. A *Markov process* is the continuous-time limit of a Markov chain. The process just introduced provides an example of a Markov chain because the increments are independent.

The probability space for a general discrete-time stochastic process where calendar time can take on values $t_0 < t_1 < \cdots < t_N$ is the space of vectors $\mathbf{x} \in \mathbb{R}^N$ with an appropriate multivariate measure, such as $P(d\mathbf{x}) = p(x_1, \ldots x_N)d\mathbf{x}$, where p is a probability density. By considering a process x_t only up to an intermediate time t_i, $i < N$, we are essentially restricting the information set of possible events or probability space of paths. The family $(\mathcal{F}_t)_{t\geq 0}$ of all reduced (or filtered) probability spaces $\mathcal{F}_t$ up to time t, for all times $t \geq 0$, is called *filtration*. One can think of $\mathcal{F}_t$ as the set of all paths up to time t. A pay-off of a derivative contract

[4] These definitions are admittedly NOT entirely rigorous, but they are meant to allow the reader to quickly develop an intuition in case she doesn't have a formal probability education. In keeping with the purpose of this book, our objective is to have the reader learn how to master the essential techniques in stochastic calculus that are useful in finance without assuming that she first learn the formal mathematical theory.

occuring at time t is a well-defined (measurable) random variable on all the spaces $\mathcal{F}_{t'}$ with $t' \geq t$ but not on the spaces with $t' < t$. Filtrations are essentially hierarchies of probability spaces (or information sets) through which more and more information is revealed to us as time progresses; i.e., $\mathcal{F}_{t'} \subset \mathcal{F}_t$ if $t' < t$ so that given a time partition $t_0 < t_1 < \cdots < t_N$, $\mathcal{F}_{t_0} \subset \mathcal{F}_{t_1} \subset \cdots \subset \mathcal{F}_{t_N}$. We say that a random variable or process is $\mathcal{F}_t$-measurable if its value is revealed at time t. Such a random variable or process is also said to be *nonanticipative* with respect to the filtration or *$\mathcal{F}_t$-adapted* (see later for a definition of nonanticipative functions, while a definition of an adapted process is also provided in Section 1.9 in the context of continuous-time asset pricing). *Conditional expectations with respect to a filtration* $\mathcal{F}_t$ represent expectations conditioned on knowing all of the information about the process only up to time t. It is customary to use the following shorthand notation for conditional probabilities:

$$E_t[\cdot] = E[\cdot\,|\mathcal{F}_t]. \tag{1.87}$$

Definition 1.9. Martingale *A real-valued $\mathcal{F}_t$-adapted continuous-time process $(x_t)_{t\geq 0}$ is said to be a P-martingale if the boundedness condition $E[|x_t|] < \infty$ holds for all $t \geq 0$ and*

$$x_t = E_t[x_T], \tag{1.88}$$

for $0 \leq t < T < \infty$

This definition implies that the conditional expectation for the value of a martingale process at a future time T, given all previous history up to the current time t (i.e., adapted to a filtration $\mathcal{F}_t$), is its current time t value. Our best prediction of future values of such a process is therefore just the presently observed value. [Note: Although we have used the same notation, i.e., x_t, this definition generally applies to arbitrary continuous-time processes that satisfy the required conditions; the pure Wiener process or standard Brownian motion is just a special case.] We remark that the expectation $E[\,] \equiv E^P[\,]$ and conditional expectation $E_t[\,] \equiv E_t^P[\,]$ are assumed here to be taken with respect to a given probability measure P. For ease of notation in what follows we drop the explicit use of the superscript P unless the probability measure must be made explicit. If one changes filtration or the probability space associated with the process, then the same process may not be a martingale with respect to the new probability measure and filtration. However, the reverse also applies, in the sense that a process may be converted into a martingale by modifying the probability measure.

A more general property satisfied by a stochastic process $(x_t)_{t\geq 0}$ (regardless of whether the process is a martingale or not) is the so-called *tower property* for $s < t < T$:

$$E_s[E_t[x_T]] = E_s[x_T]. \tag{1.89}$$

This follows from the basic property of conditional expectations: The expectation of a future expectation must be equal to the present expectation or presently forecasted value. Another way to see this is that a recursive application of conditional expectations always gives the conditional expectation with respect to the smallest information set. In this case $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}_T$. A martingale process $f_t = f(x_t, t)$ can also be specified by considering a conditional expectation over some (payoff) function ϕ of an underlying process. In particular, consider an underlying process x_t starting at time t_0 with some value x_0 and the conditional expectation

$$f_t = f(x_t, t) = E_t[\phi(x_T, T)], \tag{1.90}$$

for any $t_0 \le s \le t \le T$, then f_t satisfies the martingale property. In fact

$$E_s\big[f(x_t, t)\big] = E_s\Big[E_t\big[\phi(x_T, T)\big]\Big] = E_s\big[\phi(x_T, T)\big] = f(x_s, s). \tag{1.91}$$

The process introduced in equation (1.76) is a martingale in case the drift function $\mu(t)$ is identically zero. In fact, in this case if $t_i < t_j$, we have

$$E_{t_i}[x_{t_j}] = E_{t_i}\big[\cdots E_{t_{j-1}}\big[E_{t_j}\big[x_{t_j}\big]\big]\cdots\big] \tag{1.92}$$

$$= E_{t_i}\big[\cdots E_{t_{j-1}}\big[x_{t_j}\big]\cdots\big] = x_{t_i}. \tag{1.93}$$

Bachelier was one of the pioneers of stochastic calculus, and he proposed to use a process similar to x_t as defined by equation (1.76) in the continuous-time limit to model stock price processes.[5] A difficulty with the Bachelier model was that stock prices can attain negative values. The problem can be corrected by regarding x_t to be the natural logarithm of stock prices; this conditional density turns out to be related to (although not equivalent to) the risk-neutral density used for pricing derivatives within the Black–Scholes formulation, as is seen in Section 1.6, where we take a close look at *geometric Brownian motion*. The density in equation (1.86) leads to Bachelier's formula for the expectation of the random variable $(x_t - K)_+$, with constant $K > 0$, where $(x)_+ \equiv x$ if $x > 0$, $(x)_+ \equiv 0$ if $x \le 0$ (see Problem 9). In passing to the continuous-time limit, we have, based on equation (1.86), arrived at an expression for the random variable x_t in terms of the random variable W_t for the standard Brownian motion (or Wiener process):

$$x_t = x_0 + \bar{\mu}t + \bar{\sigma}W_t. \tag{1.94}$$

The distribution for the zero-drift random variable $(W_t)_{t\ge 0}$, representing the real-valued standard Brownian motion (Wiener process) at time t with $W_{t=0} \equiv W_0 = 0$, is given by

$$p_W(w, t) = \frac{1}{\sqrt{2\pi t}} e^{-w^2/2t} \tag{1.95}$$

at $W_t = w$. Note that this is also entirely consistent with the marginal density obtained by integrating out all intermediate variables $w_1, \ldots, w_{N-1}$ in the joint pdf of the discretized process $(W_{t_i})_{i=0,\ldots,N}$ with $w = w_N$, $t = t_N$.

According to the distributions given by equations (1.77) and (1.95), one concludes that standard Brownian motion (or the Wiener process) is a martingale process characterized by independent Gaussian (normal) increments with trajectories [i.e., path points (t, x_t)] that are continuous in time $t \ge 0$: $\delta W_t = W_{t+\delta t} - W_t \sim N(0, \delta t)$ (i.e., normally distributed with mean zero and variance δt) and $W_{t+\delta t} - W_t$ is independent of W_s for $\delta t > 0$, $0 \le s \le t$, $0 \le t < \infty$. Moreover, specializing to the case of zero drift and $\bar{\sigma} = 1$ and putting $t_0 = s$, the corresponding

[5] The date March 29, 1900, should be considered as the birth date of mathematical finance. On that day, Louis Bachelier successfully defended at the Sorbonne his thesis *Théorie de la Spéculation*. As a work of exceptional merit, strongly supported by Henri Poincaré, Bachelier's supervisor, it was published in *Annales Scientifiques de l' Ecole Normale Supérieure*, one of the most influential French scientific journals. This model was a breakthrough that motivated much of the future work by Kolmogorov and others on the foundations of modern stochastic calculus. The stochastic process proposed by Bachelier was independently analyzed by Einstein (1905) and is referred to as *Brownian motion* in the physics literature. It is also referred to as the *Wiener–Bachelier process* in a book by Feller, *An Introduction to Probability Theory and Its Applications* [Fel71]. However, this terminology didn't affirm itself, and now the process is commonly called the *Wiener process*.

probability distribution given by equation (1.86) with shifted time $t \to (t-s)$ then gives the well-known property: $W_t - W_s \sim N(0, t-s)$, $W_t \sim N(0, t)$. In fact we have the homogeneity property for the increments: $W_{t+s} - W_s \sim W_t - W_0 = W_t \sim N(0, t)$. In particular, $E[W_t] = 0$ and $E[W_t^2] = t$. An additional property is $E[W_s W_t] = \min(s, t)$. This last identity obtains from the independence of the increments [i.e., equation (1.79)]. Indeed consider any $t_i < t_j$, $0 \le i < j \le N$, then:

$$\begin{aligned} E\big[W_{t_i} W_{t_j}\big] &= E\big[(W_{t_i} - W_0)((W_{t_j} - W_{t_i}) + (W_{t_i} - W_0))\big] \\ &= E\big[(W_{t_i} - W_0)^2\big] = E\big[W_{t_i}^2\big] = t_i. \end{aligned} \tag{1.96}$$

A similar argument with $t_j < t_i$ gives t_j, while for $t_i = t_j$ we obviously obtain t_i. All of these properties also follow by taking expectations with respect to the joint pdf for the Wiener paths.

An important aspect of martingales is whether or not their trajectories or paths are continuous in time. Consider any real-valued martingale x_t, then $\delta x_t = x_{t+\delta t} - x_t$ is a process corresponding to the change in a path over an arbitrary time difference $\delta t > 0$. From equation (1.88), $E_t\big[\delta x_t\big] = 0$, so, not surprisingly, the increments of a martingale path are unpredictable (irregular), even in the infinitesimal limit $\delta t \to 0$. However, the irregularity of paths can be either continuous or discontinuous. An example of a martingale with discontinuous paths is a *jump process*, where paths are generally right continuous at every point in time as a consequence of incorporating jump discontinuities in the process at a random yet countable number of points within a time period. We refer the interested reader to recent works on the growing subject of financial modeling with jump processes (see, for example, [CT04]). Here and throughout, we focus on continuous diffusion models for asset pricing; hence our discussion is centered on continuous martingales (i.e., martingales with continuous paths). Let $f(t) = x_t(\omega)$, $t \ge 0$, represent a particular realized path indexed by the scenario ω, then continuity in the usual sense implies that the graph of *f(t)* against time is continuous for all $t \ge 0$. Denoting the left and right limits at t by $f(t-) = \lim_{s \to t^-} f(s)$ and $f(t+) = \lim_{s \to t^+} f(s)$, then $f(t) = f(t-) = f(t+)$. Every Brownian path or any path of a stochastic process generated by an underlying Brownian motion displays this property, as can be observed, for example, in Figure 1.2. [In contrast, a path of a jump diffusion process would display a similar continuity in piecewise time intervals but with the additional feature of vertical jump discontinuities at random points in time at which only right continuity holds. If $\bar{t}$ is a jump time, then $f(\bar{t}-)$, $f(\bar{t}+)$ both exist, yet $f(\bar{t}-) \ne f(\bar{t}+)$ with $f(\bar{t}) = f(\bar{t}+)$, where $f(\bar{t}) - f(\bar{t}-)$ is the size of the jump at time $\bar{t}$.]

Stochastic continuity refers to continiuty of sample paths of a process $(x_t)_{t \ge 0}$ in the probabilitistic sense as defined by

$$\lim_{s \to t} P\big(|x_s - x_t| > \epsilon\big) = 0, \quad s, t > 0 \tag{1.97}$$

for any $\epsilon > 0$. This is readily seen to hold for Brownian motion and for continuous martingales. The class of continuous-time martingales that are of interest are so-called *continuous square integrable martingales*, i.e., martingales with finite unconditional variance or finite second moment: $E[x_t^2] < \infty$ for $t \ge 0$. Such processes are closely related to Brownian motion and include Brownian motion itself. Further important properties of the paths of a continuous square integrable martingale (e.g., Brownian motion) then also follow. Consider again the time discretization $[0, t] = [t_0 = 0, t_1, \ldots, t_N = t]$ with subintervals $[t_i, t_{i+1}]$ and path points (t_i, x_{t_i}). The *variation* and *quadratic variation* of the path are, respectively, defined as:

$$V_1 = \lim_{N \to \infty} V_1^N \equiv \lim_{N \to \infty} \sum_{i=0}^{N-1} |\delta x_{t_i}| \tag{1.98}$$

and

$$V_2 = \lim_{N\to\infty} V_2^N \equiv \lim_{N\to\infty} \sum_{i=0}^{N-1} (\delta x_{t_i})^2, \tag{1.99}$$

$\delta x_{t_i} = x_{t_{i+1}} - x_{t_i}$. The properties of V_1 and V_2 provide two differing measures of how paths behave over time and give rise to important implications for stochastic calculus. Since the process is generally of nonzero variance, then $P(V_2^N > 0) = 1$ and $P(V_2 > 0) = 1$. In particular, if we let $\delta t_i = \delta t = t/N$ and consider the case of Brownian motion $x_t = W_t$, then by rewriting V_2 we have with probability 1:

$$V_2 = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=0}^{N-1} (\delta x_{t_i})^2 \right) N = \lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=0}^{N-1} (\delta W_{t_i})^2 \right) N = t. \tag{1.100}$$

Here we used the Strong law of large numbers and the fact that the $(\delta W_{t_i})^2$ are identically and independently distributed random variables with common mean of δt. Based on this important property of nonzero quadratic variation, Brownian paths, although continuous, are not differentiable. For finite N the variation V_1^N is finite. As the number N of increments goes to infinity, $\delta t_i \to 0$ and, from property (1.97), we see that the size of the increments approaches zero. The question that arises then is whether V_1 exists or not. Except for the trivial case of a constant martingale, the result is that $V_1^N \to \infty$ as $N \to \infty$; i.e., the variation V_1 is in fact infinite. Without trying to provide any rigorous proof of this here, we simply state the usual heuristic and somewhat instructive argument for this fact based on the following observation:

$$V_2^N = \sum_{i=0}^{N-1} |\delta x_{t_i}|^2 \le \left[\max_{0\le i\le N} \{|\delta x_{t_i}|\} \right] \sum_{i=0}^{N-1} |\delta x_{t_i}| = \left[\max_{0\le i\le N} \{|\delta x_{t_i}|\} \right] V_1^N. \tag{1.101}$$

Since the quadratic variation V_2 is greater than zero, taking the limit $N \to \infty$ on both sides of the inequality shows that the right-hand side must have a nonzero limit. Yet from equation (1.97) we have $\max\{|\delta x_{t_i}|\} \to 0$ as $N \to \infty$. Hence we must have that the right-hand side is a limit of an indeterminate form (of type $0 \cdot \infty$); that is, $V_1 = \lim_{N\to\infty} V_1^N = \infty$, which is what we wanted to show.

Once we are equipped with a standard Brownian motion and a filtered probability space, then the notion of stochastic integration arises by considering the concept of a nonanticipative function. Essentially, a (random) function f_t is said to be nonanticipative w.r.t. a Brownian motion or process W_t if its value at any time $t > 0$ is independent of future information. That is, f_t is possibly only a function of the history of paths up to time t and time t itself: $f_t = f(\{(W_s)_{0\le s\le t}\}, t)$. The value of this function at time t for a particular realization or scenario ω may be denoted by $f_t(\omega)$. Nonanticipative functions therefore include all deterministic (i.e., nonrandom) functions as a special case. Given a continuous nonanticipative function f_t that satisfies the "nonexplosive"condition

$$E\left[\int_0^t f_s^2 ds \right] < \infty, \tag{1.102}$$

the Itô (stochastic) integral is the random variable denoted by

$$I_t(f) = \int_0^t f_s dW_s < \infty \tag{1.103}$$

and is defined by the limit

$$I_t(f) = \lim_{N\to\infty} \sum_{i=0}^{N-1} f_{t_i} \delta W_{t_i} = \lim_{N\to\infty} \sum_{i=0}^{N-1} f_{t_i} [W_{t_{i+1}} - W_{t_i}]. \tag{1.104}$$

It can be shown that this limit exists for any choice of time partitioning of the interval $[0, t]$; e.g., we can choose $\delta t_i = \delta t = t/N$. Each term in the sum is given by a random number f_{t_i} [but fixed over the next time increment (t_i, t_{i+1})] times a random Gaussian variable δW_{t_i}. Because of this, the Itô integral can be thought of as a random walk on increments with randomly varying amplitudes. Since f_t is nonanticipative, then for each ith step we have the conditional expectation for each increment in the sum: $E_{t_i}[f_{t_i}\delta W_{t_i}] = f_{t_i} E_{t_i}[\delta W_{t_i}] = 0$. Given nonanticipative functions f_t and g_t, the following formulas provide us with the first and second moments as well as the variance-covariance properties of Itô integrals:

$$(i)\ E[I_t(f)] = E\left[\int_0^t f_s dW_s\right] = 0, \tag{1.105}$$

$$(ii)\ E[(I_t(f))^2] = E\left[\left(\int_0^t f_s dW_s\right)^2\right] = E\left[\int_0^t f_s^2 ds\right], \tag{1.106}$$

$$(iii)\ E[I_t(f)I_t(g)] = E\left[\left(\int_0^t f_s dW_s\right)\left(\int_0^t g_s dW_s\right)\right] = E\left[\int_0^t f_s g_s ds\right]. \tag{1.107}$$

Based on the definition of $I_t(f)$ and the properties of Brownian increments, it is not difficult to obtain these relations. We leave this as an exercise for the reader. Of interest in finance are nonanticipative functions of the form $f_t = f(x_t, t)$, where x_t is generally a continuous stochastic (price) process $(x_t)_{t\geq 0}$. The Itô integral is then of the form

$$I_t(f) = \int_0^t f(x_s, s) dW_s, \tag{1.108}$$

and, assuming that condition (1.102) holds, then properties (i)–(iii) also apply. Another notable property is that the Itô integral is a martingale, since $E_t[I_u(f)] = I_t(f)$, for $0 < t < u$.

The Itô integral leads us into important types of processes and the concept of a *stochastic differential equation* (SDE). In fact the general class of stochastic processes that take the form of sums of stochastic integrals are (not surprisingly) known as *Itô processes*. It is of interest to consider nonanticipative processes of the type $a_t = a(x_t, t)$ and $b_t = b(x_t, t)$, $t \geq 0$, where $(x_t)_{t\geq 0}$ is a random process. A stochastic process $(x_t)_{t\geq 0}$ is then an Itô process if there exist two nonanticipative processes $(a_t)_{t\geq 0}$ and $(b_t)_{t\geq 0}$ such that the conditions

$$P\left(\int_0^t |a_s| ds < \infty\right) = 1 \qquad \text{and} \qquad P\left(\int_0^t b_s^2 ds < \infty\right) = 1$$

are satisfied, and

$$x_t = x_0 + \int_0^t a(x_s, s) ds + \int_0^t b(x_s, s) dW_s, \tag{1.109}$$

for $t > 0$. These probability conditions are commonly imposed smoothness conditions on the drift and volatility functions. This stochastic integral equation is conveniently and formally abbreviated by simply writing it in SDE form:

$$dx_t = a(x_t, t)dt + b(x_t, t)dW_t. \tag{1.110}$$

We shall use SDE notation in most of our future discussions of Itô processes.

Itô integrals give rise to an important property, known as *Doob–Meyer decomposition*. In particular, it can be shown that if $(M_s)_{0\le s\le t}$ is a square integrable martingale process, then there exists a (nonanticipative) process $(f_s)_{0\le s\le t}$ that satisfies equation (1.102) such that

$$M_t = M_0 + \int_0^t f_s dW_s. \tag{1.111}$$

From this we observe that an Itô process x_t as given by equation (1.109) is divisible into a sum of a martingale component and a (generally random) drift component.

Problems

Problem 1. Show that the finite difference $\frac{x_{t_{i+1}} - x_{t_i}}{\delta t_i}$ of the Brownian motion in equation (1.76) is a normally distributed random variable with mean $\mu(t_i)$ and volatility $\sigma(t_i)/\sqrt{\delta t_i}$. Hint: Use equation (1.76) and take expectations while using equation (1.79).

Problem 2. Show that the random variable

$$\xi = \sum_{i=0}^{N-1} a(t_i)\delta x_{t_i}, \tag{1.112}$$

where $\delta x_{t_i} = x_{t_{i+1}} - x_{t_i}$, and x_{t_i} defined by equation (1.76), is a normal random variable. Compute its mean and variance. Hint: Take appropriate expectations while using equation (1.79).

Problem 3. Suppose that the time intervals are given by $\delta t_i = t/N$, where t is any finite time value and N is an integer. Show that equations (1.84) and (1.85) follow in the continuous-time limit as $N \to \infty$ for fixed t.

Problem 4. Show that the random variable $\xi = \sum_{i=1}^{N} a(t_i)(\delta W_{t_i})^2$ has mean and variance given by

$$E[\xi] = \sum_{i=1}^{N} a(t_i)\delta t_i, \qquad E[(\xi - E[\xi])^2] = 2\sum_{i=1}^{N} a(t_i)^2(\delta t_i)^2 \tag{1.113}$$

Hint: Since $\delta W_{t_i} \sim N(0, \delta t_i)$ independently for each i, one can use the identity in Problem 2 of Section 1.6. That is, by considering $E[\exp(\alpha\delta W_{t_i})]$ for nonzero parameter α and applying a Taylor expansion of the exponential and matching terms in the power series in α^n, one obtains $E[(\delta W_{t_i})^n]$ for any $n \ge 0$. For this problem you only need terms up to $n = 4$.

Problem 5. Show that the distribution $p(x, x_0; t)$ in equation (1.86) approaches the one-dimensional Dirac delta function $\delta(x - x_0)$ in the limit $t \to 0$.

Problem 6. (i) Obtain the joint marginal pdf of the random variables W_s and W_t, $s \ne t$. Evaluate $E[(W_t - W_s)^2]$ for all $s, t \ge 0$. (ii) Compute $E_t[W_s^3]$ for $s > t$.

Problem 7. Let the processes $(x_t)_{t\ge0}$ and $(y_t)_{t\ge0}$ be given by $x_t = x_0 + \mu_x t + \sigma_x W_t$ and $y_t = y_0 + \mu_y t + \sigma_y W_t$, where μ_x, μ_y, σ_x, σ_y are constants. Find:

(i) the means $E[x_t]$, $E[y_t]$;
(ii) the unconditional variances $\mathrm{Var}(x_t)$, $\mathrm{Var}(y_t)$;
(iii) the unconditional covariances $\mathrm{Cov}(x_t, y_t)$ and $\mathrm{Cov}(x_s, y_t)$ for all $s, t \ge 0$.

Problem 8. Obtain $E[X_t]$, $\mathrm{Var}(X_t)$, and $\mathrm{Cov}(X_s, X_t)$ for the processes

$$(a) X_t = X_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dW_s, \quad t \ge 0, \tag{1.114}$$

$$(b) X_t = \alpha(1 - t/T) + \beta(t/T) + (T - t)\int_0^t \frac{dW_s}{T - s}, \quad 0 \le t \le T, \tag{1.115}$$

where α, β, σ are constant parameters and time T is fixed in (b). The process in (a) describes the so-called *Ornstein–Uhlenbeck process*, while (b) describes a *Brownian bridge*, whereby the process is Brownian in nature, yet it is also exactly pinned down at initial time and final time T, i.e., $X_0 = \alpha$, $X_T = \beta$. For (a) assume X_0 is a constant.

Problem 9. Assume that x_t is described by a random process given by equation (1.94), or equivalently by the conditional density in equation (1.86). Show that the conditional expectation at time $t = 0$ defined by

$$C(t, K) = E_0[(x_t - K)_+], \tag{1.116}$$

where $(x)_+ = x$ if $x > 0$ and zero otherwise gives the formula

$$C(t, K) = (x_0 + \bar{\mu}t - K)N\left(\frac{x_0 + \bar{\mu}t - K}{\bar{\sigma}\sqrt{t}}\right) + \bar{\sigma}\sqrt{t}\varphi\left(\frac{x_0 + \bar{\mu}t - K}{\bar{\sigma}\sqrt{t}}\right), \tag{1.117}$$

where $N(\cdot)$ is the standard cumulative normal distribution function and

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{1.118}$$

By further restricting the drift, $\mu = 0$ gives Bachelier's formula. This corresponds (from the viewpoint of pricing theory) to the fair price of a standard call option struck at K, and maturing in time t, assuming a zero interest rate and *simple* Brownian motion for the underlying "stock" level x_t at time t. Hint: One way to obtain equation (1.117) is by direct integration over all x_t of the product of the density p [of equation (1.86)] and the payoff function $(x_t - K)_+$. Use appropriate changes of integration variables and the property $1 - N(x) = N(-x)$ to arrive at the final expression.

1.5 Stochastic Differential Equations and Itô's Formula

For purposes of describing asset price processes it is of interest to consider SDEs for diffusion processes x_t that are defined in terms of a lognormal drift function $\mu(x, t)$ and a lognormal volatility function $\sigma(x, t)$ and are written as follows:[6]

$$dx_t = \mu(x_t, t)x_t dt + \sigma(x_t, t)x_t dW_t. \tag{1.119}$$

Assuming the drift and volatility are smooth functions, the discretization process in the previous section extends to this case and produces a solution to equation (1.119) as the limit as $N \to \infty$ of the Markov chain $x_{t_0}, \ldots, x_{t_N}$ defined by means of the recurrence relations

$$x_{t_{i+1}} = x_{t_i} + \mu(x_{t_i}, t_i)x_{t_i}\delta t_i + \sigma(x_{t_i}, t_i)x_{t_i}\delta W_{t_i}. \tag{1.120}$$

[6] When the drift and volatility (or diffusion) terms in the SDE are written in the form given by equation (1.119) it is common to refer to μ and σ as the lognormal drift and volatility, respectively. The reason for using this terminology stems from the fact that in the special case that μ and σ are at most only functions of time t (i.e., not dependent on x_t), the SDE leads to geometric Brownian motion, and, in particular, the conditional transition density is exactly given by a lognormal distribution, as discussed in the next section.

From this discrete form of equation (1.119) we observe that $x_{t+\delta t} - x_t = \delta x_t = \mu(x_t, t)x_t\delta t + \sigma(x_t, t)x_t\delta W_t$. Alternatively, the solution to equation (1.119) can be characterized as the process x_t such that

$$\mu(x_t, t) = \lim_{\delta t \to 0} \frac{E_t[x_{t+\delta t} - x_t]}{x_t \delta t}, \qquad \sigma(x_t, t)^2 = \lim_{\delta t \to 0} \frac{E_t[(x_{t+\delta t} - x_t)^2]}{x_t^2 \delta t}. \tag{1.121}$$

These expectations follow from the properties $E_t[\delta W_t] = 0$ and $E_t[(\delta W_t)^2] = \delta t$. Notice that, although an SDE defines a stochastic process in a fairly constructive way, conditional distribution probabilities, such as the one for the Wiener process in equation (1.86), can be computed in analytically closed form only in some particular cases. Advanced methods for obtaining closed-form conditional (transition) probability densities for certain families of drift and volatility functions are discussed in Chapter 3, where the corresponding Kolmogorov (or Fokker–Planck) partial differential equation approach is presented in detail.

A method for constructing stochastic processes is by means of nonlinear transformations. The stochastic differential equation satisfied by a nonlinear transformation as a function of another diffusion process is given by Itô's lemma:

Lemma 1.3. Itô's Lemma *If the function $f_t = f(x_t, t)$ is smooth with continuous derivatives $\partial f/\partial t$, $\partial f/\partial x$, and $\partial^2 f/\partial x^2$ and x_t satisfies the stochastic differential*

$$dx_t = a(x_t, t)dt + b(x_t, t)dW_t, \tag{1.122}$$

where $a(x, t)$ and $b(x, t)$ are smooth functions of x and t, then the stochastic differential of f_t is given by

$$\begin{aligned} df_t &= \left(\frac{\partial f}{\partial t} + a(x_t, t)\frac{\partial f}{\partial x} + \frac{b(x_t, t)^2}{2}\frac{\partial^2 f}{\partial x^2} \right) dt + b(x_t, t)\frac{\partial f}{\partial x} dW_t \\ &\equiv A(x_t, t)dt + B(x_t, t)dW_t. \end{aligned} \tag{1.123}$$

In stochastic integral form:

$$f_t = f_0 + \int_0^t A(x_s, s)ds + \int_0^t B(x_s, s)dW_s. \tag{1.124}$$

A nonrigorous, yet instructive, "proof" is as follows.[7]

Proof. Using a Taylor expansion we find

$$\begin{aligned} \delta f_t &= f(x_t + \delta x_t, t + \delta t) - f(x_t, t) \\ &= \frac{\partial f}{\partial t}(x_t, t)\delta t + \frac{\partial f}{\partial x}(x_t, t)\delta x_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_t, t)(\delta x_t)^2 + O((\delta t)^{\frac{3}{2}}), \end{aligned} \tag{1.125}$$

where the remainder has an expectation and variance converging to zero as fast as $(\delta t)^2$ in the limit $\delta t \to 0$. Inserting the finite differential form of equation (1.122) into equation (1.125) while replacing $(\delta W_t)^2 \to \delta t$ and retaining only terms up to $O(\delta t)$ gives

$$\begin{aligned} \delta f_t = &\left(\frac{\partial f}{\partial t}(x_t, t) + a(x_t, t)\frac{\partial f}{\partial x}(x_t, t) + \frac{b(x_t, t)^2}{2}\frac{\partial^2 f}{\partial x^2}(x_t, t) \right) \delta t \\ &+ b(x_t, t)\frac{\partial f}{\partial x}(x_t, t)\delta W_t + O((\delta t)^{\frac{3}{2}}). \end{aligned} \tag{1.126}$$

[7] For more formal rigorous treatments and proofs see, for example, [IW89, Øks00, JS87].

Taking the limit $N \to \infty$ ($\delta t \to 0$), the finite difference δt is the infinitesimal differential dt, δW_t is the stochastic differential dW_t, the remainder term drops out, and we finally obtain equation (1.123). Alternatively, with the use of equation (1.125) we can obtain the drift function of the f_t process:

$$
\begin{aligned}
A(x_t, t) &= \lim_{\delta t \to 0} \frac{E_t[\delta f_t]}{\delta t} \\
&= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \lim_{\delta t \to 0} \frac{E_t[\delta x_t]}{\delta t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} \lim_{\delta t \to 0} \frac{E_t[(\delta x_t)^2]}{\delta t} \\
&= \frac{\partial f}{\partial t}(x_t, t) + a(x_t, t)\frac{\partial f}{\partial x}(x_t, t) + \frac{b(x_t, t)^2}{2}\frac{\partial^2 f}{\partial x^2}(x_t, t);
\end{aligned}
$$

and the volatility function of the f_t process:

$$
\begin{aligned}
B(x_t, t)^2 &= \lim_{\delta t \to 0} \frac{E_t[(\delta f_t)^2]}{\delta t} \\
&= \left(\frac{\partial f}{\partial x}\right)^2 \lim_{\delta t \to 0} \frac{E_t[(\delta x_t)^2]}{\delta t} = b(x_t, t)^2 \left(\frac{\partial f}{\partial x}(x_t, t)\right)^2.
\end{aligned}
$$

The drift and volatility functions therefore give equation (1.123), as required. Here we have made use of the expectations

$$
a(x_t, t) = \lim_{\delta t \to 0} \frac{E_t[\delta x_t]}{\delta t}, \qquad b(x_t, t)^2 = \lim_{\delta t \to 0} \frac{E_t[(\delta x_t)^2]}{\delta t}
$$

following from the finite differential form of equation (1.122). □

Note: Itô's formula is rather simple to remember if one just takes the Taylor expansion of the infinitesimal change df up to second order in dx and up to first order in the time increment dt and then inserts the stochastic expression for dx and replaces $(dx)^2$ by $b(x, t)^2 dt$.

As we will later see, in most pricing applications, x_t represents some asset price process, and therefore it proves convenient to consider Itô's lemma applied to the SDE of equation (1.119); i.e., $a(x, t) = x\mu(x, t)$, $b(x, t) = x\sigma(x, t)$, written in terms involving the lognormal drift and volatility functions for the random variable x. Equation (1.123) then gives

$$
\begin{aligned}
df_t &= \left(\frac{\partial f}{\partial t} + x\mu\frac{\partial f}{\partial x} + \frac{x^2\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}\right)dt + x\sigma\frac{\partial f}{\partial x}dW_t && (1.127) \\
&\equiv \mu_f f_t dt + \sigma_f f_t dW_t && (1.128)
\end{aligned}
$$

From this form of the SDE we identify the corresponding lognormal drift $\mu_f = \mu_f(x, t)$ and volatility $\sigma_f = \sigma_f(x, t)$ for the process f_t.

The foregoing derivation of Itô's lemma for one underlying random variable can be extended to the general case of a function $f(x_1, \ldots, x_n, t)$ depending on n random variables $x = (x_1, \ldots, x_n)$ and time t. [Note: To simplify notation, we shall avoid the use of subscript t in the variables, i.e., $x_{1,t} = x_1$, etc.] We can readily derive Itô's formula by assuming that the x_i, $i = 1, \ldots, n$, satisfy the stochastic differential equations

$$
dx_i = a_i dt + b_i \sum_{j=1}^{n} \Lambda_{ij} dW_t^j. \qquad (1.129)
$$

Here the coefficients $a_i = a_i(x_1, \ldots, x_n, t)$ and $b_i = b_i(x_1, \ldots, x_n, t)$ are any smooth functions of the arguments. Furthermore we assume that the Wiener processes W_t^j are mutually independent, i.e.,

$$E[dW_t^i dW_t^j] = \delta_{ij}\ dt. \tag{1.130}$$

The constants $\rho_{ij} = \rho_{ji}$ (with $\rho_{ii} = 1$) are correlation matrix elements and are convenient for introducing correlations among the increments (e.g. see equation (1.176) of Section 1.6):

$$\begin{aligned} E[(dx_i)(dx_j)] &= b_i b_j \sum_{k=1}^{n}\sum_{l=1}^{n} \Lambda_{ik}\Lambda_{jl} E[dW_t^k dW_t^l] \\ &= b_i b_j \sum_{k=1}^{n} \Lambda_{ik}\Lambda_{jk} dt \equiv b_i b_j \rho_{ij} dt. \end{aligned} \tag{1.131}$$

When $i = j$ this gives $E[(dx_i)^2] = b_i^2 dt$. Taylor expanding *df* up to second order in the dx_i increments and to first order in *dt* we have

$$df = \frac{\partial f}{\partial t}dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}dx_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(dx_i)(dx_j) \tag{1.132}$$

Now replacing $(dx_i)(dx_j)$ by the right-hand side of equation (1.131) while substituting the above expression for dx_i and collecting terms in *dt* and the dW_t^i gives the final expression:

$$\begin{aligned} df = &\left(\frac{\partial f}{\partial t} + \sum_{i=1}^{n}\left[a_i \frac{\partial f}{\partial x_i} + \frac{b_i^2}{2}\frac{\partial^2 f}{\partial x_i^2}\right] + \sum_{i<j=1}^{n} b_i b_j \rho_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}\right)dt \\ &+ \sum_{j=1}^{n}\left(\sum_{i=1}^{n} \rho_{ij} b_i \frac{\partial f}{\partial x_i}\right)dW_t^j. \end{aligned} \tag{1.133}$$

This procedure can be straightforwardly applied or extended to stochastic differentials of various processes that are dependent on groups of underlying random variables.

As we shall see in the coming sections, where we cover derivatives pricing in continuous time, it is important to work out the stochastic differential of the quotient of two processes, namely; $f_t \equiv g_t/h_t$, where

$$\frac{dg_t}{g_t} = \mu_g dt + \sum_{i=1}^{n} \sigma_g^i dW_t^i, \qquad \frac{dh_t}{h_t} = \mu_h dt + \sum_{i=1}^{n} \sigma_h^i dW_t^i \tag{1.134}$$

are stochastic differential equations assumed satisfied by g_t and h_t, respectively. Note that the drift and volatility functions[8] are generally considered functions of time and of the underlying processes, $\mu_g = \mu_g(g_t, h_t, t)$, $\mu_h = \mu_h(g_t, h_t, t)$, $\sigma_g^i = \sigma_g^i(g_t, h_t, t)$, $\sigma_h^i = \sigma_h^i(g_t, h_t, t)$. The function σ_g^i is the volatility of the process g_t with respect to the *i*th independent Wiener process (or *i*th risk factor).[9] The stochastic differential of the ratio $f_t = g_t/h_t$ can be obtained via the Taylor expansion of the differential *df* up to first order in *dt* and up to second order

[8] Here and throughout the rest of the book we shall sometimes take the liberty to refer to the lognormal drift and volatility functions simply as the drift and volatility so as to avoid excessive use of such terminology.

[9] In what follows we shall at times also refer to independent Brownian motions as risk factors.

in the dg and dh terms. Hence considering f as function of g, h, and t and taking appropriate partial derivatives gives

$$df = \frac{1}{h}dg - \frac{g}{h^2}dh - \frac{1}{h^2}(dg)(dh) + \frac{g}{h^3}(dh)^2. \tag{1.135}$$

Here $\frac{\partial f}{\partial t} = 0$, since there is no explicit time dependence. Moreover, since $\frac{\partial^2 f}{\partial g^2} = 0$, the $(dg)^2$ term is absent. This last SDE takes on a particularly simple form when we divide through by f:

$$\frac{df}{f} = \frac{dg}{g} - \frac{dh}{h} - \frac{dg}{g}\frac{dh}{h} + \left(\frac{dh}{h}\right)^2 = \left(\frac{dg}{g} - \frac{dh}{h}\right)\left(1 - \frac{dh}{h}\right) \tag{1.136}$$

Substituting equations (1.134), expanding out, and setting to zero any term containing $(dW_t^i)(dt)$ or $(dt)^2$ [i.e., terms of $O((dt)^{3/2})$ or higher] then gives

$$\frac{df}{f} = \left[\mu_g - \mu_h - \sum_{i=1}^{n} \sigma_h^i(\sigma_g^i - \sigma_h^i)\right]dt + \sum_{i=1}^{n}(\sigma_g^i - \sigma_h^i)dW_t^i. \tag{1.137}$$

Here we have also made use of the replacement $dW_t^i\, dW_t^j = \delta_{ij}\, dt$. This gives the stochastic differential of $f_t = g_t/h_t$. Note that this equation in compact form reads

$$\frac{df}{f} = \mu_f dt + \sum_{i=1}^{n} \sigma_f^i dW_t^i, \tag{1.138}$$

where the drift of f is $\mu_f = \mu_g - \mu_h - \sum_{i=1}^{n} \sigma_h^i(\sigma_g^i - \sigma_h^i)$ and the volatility is given by $\sigma_f^i = \sigma_g^i - \sigma_h^i$. It is important to note that pricing formulas ultimately involve the absolute value or square of the volatilities, i.e., $\sigma_f^i = \left|\sigma_g^i - \sigma_h^i\right| = \sqrt{(\sigma_g^i)^2 + (\sigma_h^i)^2 - 2\sigma_g^i\sigma_h^i}$. This will become clear in the sections that follow. Namely, a rigorous justification of this arises from consideration of the partial differential equation (i.e., the forward or backward Kolmogorov equation) satisfied by the corresponding transition probability density function, which explicitly involves only terms in the square of the volatilities. Finally, note that for the case of only one risk factor, i.e., $n = 1$, we have equation (1.138) with $\mu_f = \mu_g - \mu_h - \sigma_h(\sigma_g - \sigma_h)$ and $\sigma_f = \sigma_g - \sigma_h$. For general n, using vector notation ($\boldsymbol{\sigma}_f = \boldsymbol{\sigma}_g - \boldsymbol{\sigma}_h$, $\mu_f = \mu_g - \mu_h - \boldsymbol{\sigma}_h \cdot (\boldsymbol{\sigma}_g - \boldsymbol{\sigma}_h)$) and equation (1.138) takes the form:

$$\frac{df}{f} = \mu_f dt + \boldsymbol{\sigma}_f \cdot d\mathbf{W}_t. \tag{1.139}$$

Recall that a martingale process, which we shall here simply denote by f_t, is a stochastic process for which $E_t^P[f_T] = f_t$, $t \le T$, under a given probability measure P. Recall that this is a driftless process, in the sense that its expected value, under P, is constant over all future times. We have already encountered a simple example of such a process, namely, the standard Brownian motion, or Wiener process W_t. Equation (1.90) provides a method of generating a martingale process. Based on Itô's Lemma we now have the following result.

Theorem. (Feynman–Kac) *If $f(x,t)$ is the function given by the conditional expectation*

$$f(x, t) = E_t[\phi(x_T)], \tag{1.140}$$

at time $t \le T$, with $x_t = x$ and underlying process obeying equation (1.122), then $f(x,t)$ satisfies the partial differential equation

$$\frac{\partial f(x,t)}{\partial t} + a(x,t)\frac{\partial f(x,t)}{\partial x} + \frac{b(x,t)^2}{2}\frac{\partial^2 f(x,t)}{\partial x^2} = 0, \tag{1.141}$$

with terminal time condition $f(x, T) = \phi(x)$.

Proof. The proof follows by considering the conditional expectation of equation (1.126) at time t, which leaves us with only the drift term in δt (to order δt), since the Wiener term is Markovian. On the other hand,

$$E_t[\delta f_t] = E_t[f_{t+\delta t}] - f_t = 0. \tag{1.142}$$

The last equality is due to the martingale property of f_t. In the limit of infinitesimal time step we are left with the infinitesimal drift term, which vanishes identically only if equation (1.141) is satisfied. The terminal condition follows simply because $f(x, t = T) = E_T[\phi(x_T)] = \phi(x)$, with $x_T = x$ imposed when $t = T$. □

The Black–Scholes partial differential equation discussed in Section 1.13 is a special case of the Feynman–Kac result. The generalization of equation (1.141) to n dimensions is also readily obtained by using Itô's lemma in n dimensions.

Problems

Problem 1. Consider the stochastic processes g_t and h_t defined earlier. Further assume that the volatilities of the two processes are identical with respect to all Brownian increments, i.e., $\sigma_g^i = \sigma_h^i$ for all i. Show that the process $f_t = g_t/h_t$ is deterministic with solution

$$f_T = f_t \exp\left(\int_t^T \left(\mu_g(g_s, s) - \mu_h(h_s, s)\right) ds\right). \tag{1.143}$$

Problem 2. Consider two processes defined by $g_t = g_0 e^{\sigma_g W_t + \mu_g t}$ and $h_t = h_0 e^{\sigma_h W_t + \mu_h t}$, where W_t is a standard Wiener process and μ_g, μ_h, σ_g, σ_h, g_0, and h_0 are constants. Use Itô's lemma to show that

$$\frac{dg_t}{g_t} = \left(\mu_g + \frac{\sigma_g^2}{2}\right) dt + \sigma_g dW_t, \qquad \frac{dh_t}{h_t} = \left(\mu_h + \frac{\sigma_h^2}{2}\right) dt + \sigma_h dW_t. \tag{1.144}$$

Then assume $df_t/f_t = \mu_f dt + \sigma_f dW_t$. Find these drift and volatility coefficients in terms of μ_g, μ_h, σ_g, and σ_h, for the cases $f_t = g_t/h_t$ and $f_t = g_t h_t$.

Problem 3. Obtain the stochastic differential equations satisfied by the Ornstein–Uhlenbeck and Brownian bridge processes in Problem 8 of Section 1.4.

1.6 Geometric Brownian Motion

Univariate geometric Brownian motion with time-dependent coefficients is characterized by the SDE of the form

$$dS_t = \mu(t) S_t \, dt + \sigma(t) S_t \, dW_t, \tag{1.145}$$

with initial condition S_0, where $\mu = \mu(t)$ and $\sigma = \sigma(t)$ are deterministic functions of time t. This equation can be solved by means of the change of variable

$$x_t = \log \frac{S_t}{S_0}. \tag{1.146}$$

The transformed equation is obtained using Itô's lemma,

$$dx_t = \left(\mu(t) - \frac{\sigma(t)^2}{2}\right) dt + \sigma(t) dW_t, \tag{1.147}$$

and is to be solved with initial condition $x_0 = 0$. Following the procedure in Section 1.4 we discretize this equation in the time interval $[0,T]$ using a partition in N subintervals of length $\delta t = \frac{T}{N}$:

$$x_{t_{i+1}} = x_{t_i} + \left(\mu(t_i) - \frac{\sigma(t_i)^2}{2}\right) \delta t + \sigma(t_i)\ \delta W_{t_i}. \tag{1.148}$$

By iterating the recurrence relations up to time T, we find

$$x_T = \sum_{i=0}^{N-1} \left[\left(\mu(t_i) - \frac{\sigma(t_i)^2}{2}\right) \delta t + \sigma(t_i)\ \delta W_{t_i}\right]. \tag{1.149}$$

Hence x_T is a normal random variable for all $N > 1$. In the limit as $N \to \infty$, the mean of x_T is given by

$$E_0[x_T] = \lim_{N\to\infty} \sum_{i=0}^{N-1} \left(\mu(t_i) - \frac{\sigma(t_i)^2}{2}\right) \delta t = \int_0^T \left(\mu(t) - \frac{\sigma(t)^2}{2}\right) dt \tag{1.150}$$

and the variance is given by

$$E_0[x_T^2] - (E_0[x_T])^2 = \lim_{N\to\infty} \sum_{i=0}^{N-1} \sigma(t_i)^2\ \delta t = \int_0^T \sigma(t)^2\ dt. \tag{1.151}$$

Introducing the time-averaged drift and volatility

$$\bar{\mu}(T) \equiv \frac{1}{T} \int_0^T \mu(t) dt \tag{1.152}$$

and

$$\bar{\sigma}(T) \equiv \sqrt{\frac{1}{T} \int_0^T \sigma(t)^2\ dt}, \tag{1.153}$$

we conclude that $x_T = \log \frac{S_T}{S_0} \sim N\left(\left(\bar{\mu}(T) - \frac{\bar{\sigma}^2(T)}{2}\right) T, \bar{\sigma}^2(T)T \right)$. This result is also easily verified by directly applying properties (1.105) and (1.106) to the integrated form of equation (1.147).

The solution to stochastic differential equation (1.145) for all $t \geq 0$ is hence

$$S_t = S_0 \exp\left(\left(\bar{\mu}(t) - \frac{\bar{\sigma}^2(t)}{2}\right)t + \bar{\sigma}(t) W_t\right), \tag{1.154}$$

where $\bar{\mu}(t)$ and $\sigma(t)$ are given by equations (1.152) and (1.153), respectively. This solution (which is actually a strong solution) can also be verified by a direct application of Itô's lemma (see Problem 1). Note that this represents a solution, in the sense that the random variable denoted by S_t and parameterized by time t is expressed in terms of the underlying random variable, W_t, for the pure Wiener process.

This solution gives a closed-form expression for generating sample paths for geometric Brownian motion. Equation (1.154) provides a general expression for the case of time-dependent drift and volatility. It is very instructive at this point to compute expectations of functions of S_t. Let us consider the process in equation (1.145) and proceed now to compute the expectations $E_0[S_t]$ and $E_0[(S_t - K)_+]$, for some constant $K \geq 0$, where $(x)_+ \equiv \max(x, 0) = x$ if $x > 0$ and zero if $x \leq 0$. Using the solution in equation (1.154), the expectation of S_t under the density of equation (1.95) (i.e., conditional on $S_{t=0} = S_0$, hence we write $E_0[\,]$) is

$$\begin{aligned} E_0[S_t] &= S_0 e^{(\bar{\mu} - \bar{\sigma}^2/2)t} E_0[e^{\bar{\sigma} W_t}] \\ &= S_0 e^{(\bar{\mu} - \bar{\sigma}^2/2)t} e^{\bar{\sigma}^2 t/2} = S_0 e^{\bar{\mu} t}. \end{aligned} \tag{1.155}$$

To compact notation we denote $\bar{\mu} \equiv \bar{\mu}(t)$, $\bar{\sigma} \equiv \bar{\sigma}(t)$. In the last step we have used an important identity derived in Problem 2 of this section. This result shows that the stock price is expected to grow exponentially at a rate of $\bar{\mu}$.

Using equation (1.154), the expectation $E_0[(S_t - K)_+]$ is given by

$$\begin{aligned} E_0[(S_t - K)_+] &= \int_{-\infty}^{\infty} p(y, t)\big(S_0 e^{(\bar{\mu} - \bar{\sigma}^2/2)t} e^{\bar{\sigma} y} - K\big)_+ dy \\ &= \frac{S_0 e^{(\bar{\mu} - \bar{\sigma}^2/2)t}}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-y^2/2t} \left(e^{\bar{\sigma} y} - \frac{K}{S_0} e^{-(\bar{\mu} - \bar{\sigma}^2/2)t} \right)_+ dy \end{aligned} \tag{1.156}$$

The last step obtains from the identity $(ax - b)_+ = a(x - b/a)_+$, for $a > 0$. Changing integration variable $y = \sqrt{t}x$ while employing this identity again gives

$$E_0[(S_t - K)_+] = \frac{S_0 e^{(\bar{\mu} - \frac{\bar{\sigma}^2}{2})t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2 + \bar{\sigma}\sqrt{t}x} \left(1 - \frac{K}{S_0} e^{-[(\bar{\mu} - \frac{\bar{\sigma}^2}{2})t + \bar{\sigma}\sqrt{t}x]} \right)_+ dx \tag{1.157}$$

Since $e^{-\bar{\sigma}\sqrt{t}x}$ is a monotonically decreasing function of x, there is a value x_K such that

$$\left(1 - \frac{K}{S_0} e^{-[(\bar{\mu} - \bar{\sigma}^2/2)t + \bar{\sigma}\sqrt{t}x]} \right)_+ = \begin{cases} 1 - \dfrac{K}{S_0} e^{-[(\bar{\mu} - \bar{\sigma}^2/2)t + \bar{\sigma}\sqrt{t}x]}, & x > x_K \\ \\ 0, & x \leq x_K \end{cases} \tag{1.158}$$

where

$$x_K = -\frac{\log(S_0/K) + (\bar{\mu} - \bar{\sigma}^2/2)t}{\bar{\sigma}\sqrt{t}}. \tag{1.159}$$

Hence, the integral in equation (1.157) becomes a sum of two parts in the region $x \in (x_K, \infty)$:

$$E_0[(S_t - K)_+] = \frac{S_0 e^{(\bar{\mu} - \frac{\bar{\sigma}^2}{2})t}}{\sqrt{2\pi}} \int_{x_K}^{\infty} e^{-x^2/2 + \bar{\sigma}\sqrt{t}x}\, dx - \frac{K}{\sqrt{2\pi}} \int_{x_K}^{\infty} e^{-x^2/2}\, dx. \tag{1.160}$$

Completing the square in the first integration gives

$$\begin{aligned} E_0[(S_t - K)_+] &= S_0 e^{\bar{\mu} t}(1 - N(x_K - \bar{\sigma}\sqrt{t})) - KN(-x_K) \\ &= S_0 e^{\bar{\mu} t} N(\bar{\sigma}\sqrt{t} - x_K) - KN(-x_K) \\ &= S_0 e^{\bar{\mu} t} N(d_+) - KN(d_-), \end{aligned} \tag{1.161}$$

where $N(\cdot)$ is the standard cumulative normal distribution function and

$$d_{\pm} = \frac{\log(S_0/K) + (\bar{\mu} \pm \bar{\sigma}^2/2)t}{\bar{\sigma}\sqrt{t}}. \tag{1.162}$$

Note that here we have used the property $N(-x) = 1 - N(x)$.

The Black–Scholes pricing formula for a plain European call option follows automatically. In particular, assuming a risk-neutral pricing measure, the drift is given by the instantaneous risk-free rate $\mu(t) = r(t)$. Hence, the price of a call at current time ($t = 0$) with current stock level (or spot) S_0, strike K, and maturing in time t is given by the discounted expectation

$$C_0(S_0, K, t) = e^{-\bar{r}t} E_0[(S_t - K)_+] = S_0 N(d_+) - e^{-\bar{r}t} K N(d_-), \tag{1.163}$$

where $\bar{r}$ is the time-averaged continuously compounded risk-free interest rate

$$\bar{r} = \bar{r}(t) \equiv \frac{1}{t}\int_0^t r(\tau)d\tau, \tag{1.164}$$

and $d_{\pm}$ is given by equation (1.162) with $\bar{\mu} = \bar{r}$. It is instructive to note the inherent difference between the Black–Scholes pricing formula in equation (1.163) and Bachelier's formula in equation (1.117). Bachelier's formula is a result of assuming a standard Brownian motion for the underlying stock price process [i.e., equation (1.94)]. In contrast, formulas of the Black–Scholes type are equivalent to the assumption of geometric Brownian motion for the underlying price process. Using equation (1.154) as defining a change of probability variables $W_t \to S_t$, the one-dimensional analogue of equation (1.48) together with equation (1.95) gives

$$p(S_t, S_0; t) = \frac{1}{S_t \bar{\sigma}\sqrt{2\pi t}} e^{-[\log(S_t/S_0) - (\bar{\mu} - \bar{\sigma}^2/2)t]^2 / 2\bar{\sigma}^2 t}. \tag{1.165}$$

This is the *lognormal distribution* function defined on positive stock price space $S_t \in (0, \infty)$. The log-returns $\log(S_t/S_0)$ are distributed normally with mean $(\bar{\mu} - \bar{\sigma}^2/2)t$ and variance $\bar{\sigma}^2 t$. Setting $\bar{\mu} = \bar{r}$ gives the *risk-neutral conditional probability density* for a stock attaining a value S_t at time $t > 0$ given an initial value S_0 at time $t = 0$. Hence, the Black–Scholes pricing formula for European options can also be obtained by taking discounted expectations of payoff functions with respect to this risk-neutral density. In particular, a European-style claim having pay-off $\Lambda(S_T)$ as a function of the terminal stock level S_T, where $T > 0$ is a maturity time, has arbitrage-free price $f_0(S_0, T)$ at time $t = 0$ expressible as

$$f_0(S_0, T) = e^{-\bar{r}(T)T} E_0^Q[\Lambda(S_T)] = e^{-\bar{r}(T)T} \int_0^\infty p(S_T, S_0; T)\Lambda(S_T)dS_T. \tag{1.166}$$

Here the superscript Q is used to denote an expectation with respect to the risk-neutral density given by equation (1.165) with drift $\bar{\mu} = \bar{r}(T)$. Note that within this probability measure, equation (1.155) shows that stock prices drift at the time-averaged risk-free rate *r(t)* at time *t*. As will become apparent in the following sections, this must be the case in order to ensure arbitrage-free pricing.

For pricing applications, discussed in greater length in later sections of this chapter, it is useful to consider a slight extension of the foregoing closed-form solutions to geometric Brownian motion. Namely, we can extend equation (1.154) by a simple shift in time variables as follows:

$$S_T = S_t \exp\left(\left(\bar{\mu}(t, T) - \frac{\bar{\sigma}^2(t, T)}{2}\right)(T - t) + \bar{\sigma}(t, T) W_{T-t}\right), \tag{1.167}$$

with time-averaged drift and volatility over the period *[t,T]*

$$\bar{\mu}(t,T) \equiv \frac{1}{T-t}\int_t^T \mu(\tau)d\tau, \qquad \bar{\sigma}^2(t,T) \equiv \frac{1}{T-t}\int_t^T \sigma^2(\tau)d\tau. \tag{1.168}$$

Here $W_{T-t} = W_T - W_t$ is the Wiener normal random variable with mean zero and variance $T-t$; i.e., $W_{T-t} \sim \sqrt{T-t}x$, $x \sim N(0,1)$. For constant drift and volatility this solution simplifies in the obvious manner. The formula for the conditional expectation now extends to give

$$E_t[(S_T-K)_+] = e^{\bar{\mu}(T-t)}S_tN(d_+) - KN(d_-), \tag{1.169}$$

with

$$d_\pm = \frac{\log(S_t/K) + (\bar{\mu} \pm \bar{\sigma}^2/2)(T-t)}{\bar{\sigma}\sqrt{T-t}} \tag{1.170}$$

and $\bar{\mu} = \bar{\mu}(t,T)$, $\bar{\sigma} = \bar{\sigma}(t,T)$. A related expectation that is useful for pricing purposes is (see Problem 3)

$$E_t[(K-S_T)_+] = KN(-d_-) - e^{\bar{\mu}(T-t)}S_tN(-d_+). \tag{1.171}$$

Within the risk-neutral probability measure, $\bar{\mu} = \bar{r}$. Hence discounting this expectation by $e^{-\bar{r}(T-t)}$ gives the analogue of equation (1.163) for the Black–Scholes price of a put option at calendar time t, spot S_t, and maturing at time T with strike K:

$$P_t(S_t,K,T) = e^{-\bar{r}(T-t)}KN(-d_-) - S_tN(-d_+), \tag{1.172}$$

where $d_\pm$ is given by equation (1.170) with $\bar{\mu} = \bar{r} \equiv r(t,T)$.

In closing this section, we consider the more general multidimensional case of geometric Brownian motion. Multivariate geometric Brownian motions describe n-dimensional state spaces of vector valued processes $S_t^1, \ldots, S_t^n$ and can be described with two different but equivalent sets of notations. Let's consider n uncorrelated standard Wiener processes

$$W_t^1, \ldots, W_t^n, \text{with } \; E_t[dW_t^i dW_t^j] = \delta_{ij}\; dt. \tag{1.173}$$

A simple way to introduce correlations among the price processes is to allow for correlated Wiener processes by defining a new set of n processes $W_t^{S^i}$ as

$$dW_t^{S^i} = \sum_{j=1}^n \Lambda_{ij}\; dW_t^j, \tag{1.174}$$

or, in matrix-vector notation,

$$d\mathbf{W}_t^S = \mathbf{\Lambda}\cdot d\mathbf{W}_t. \tag{1.175}$$

Using equation (1.174) we have

$$E_t\left[dW_t^{S^i}dW_t^{S^j}\right] = \sum_{k,l=1}^n \Lambda_{ik}\Lambda_{jl}\;\delta_{kl}\; dt = \sum_{k=1}^n \Lambda_{ik}\Lambda_{jk}\; dt \equiv \rho_{ij}\; dt, \tag{1.176}$$

where the last relation defines a correlation matrix $\boldsymbol{\rho}$, with elements ρ_{ij}, and lower Cholesky decomposition given by

$$\boldsymbol{\rho} = \mathbf{\Lambda}\mathbf{\Lambda}^\dagger. \tag{1.177}$$

Throughout this section, superscript † denotes matrix transpose.

Stochastic differential equations for the stock price processes can be written as follows:

$$\frac{dS_t^i}{S_t^i} = \mu_i \; dt + \sigma_i \; dW_t^{S^i} \tag{1.178}$$

$$= \mu_i \; dt + \sigma_i \sum_{j=1}^{n} \Lambda_{ij} \; dW_t^j \equiv \mu_i \; dt + \sum_{j=1}^{n} L_{ij} \; dW_t^j \tag{1.179}$$

where the last expression defines the matrix $\mathbf{L}$, $L_{ij} = \sigma_i \Lambda_{ij}$. Note that the lognormal drifts μ_i and volatilities σ_i can generally depend on time, although to simplify notation we have chosen not to denote this explicitly. The last relation in equation (1.179) defines a lower Cholesky factorization of the covariance matrix

$$\mathbf{C} = \mathbf{L}\mathbf{L}^{\dagger} = \mathbf{\Sigma}\mathbf{\Lambda}\mathbf{\Lambda}^{\dagger}\mathbf{\Sigma} = \mathbf{\Sigma}\boldsymbol{\rho}\mathbf{\Sigma}. \tag{1.180}$$

Here $\mathbf{\Sigma}$ is the diagonal matrix of lognormal volatilities with *(ij)*-elements given by $\delta_{ij} \; \sigma_i$, $\mathbf{L} = \mathbf{\Sigma}\mathbf{\Lambda}$ and $\mathbf{\Sigma} = \mathbf{\Sigma}^{\dagger}$. In vector notation we can write equations (1.179) in a compact form as

$$\frac{dS_t^i}{S_t^i} = \mu_i \; dt + \boldsymbol{\sigma}_i \cdot d\mathbf{W}_t, \tag{1.181}$$

where $\boldsymbol{\sigma}_i = (\sigma_{i1}, \ldots, \sigma_{in})$ is the volatility vector for the ith stock, whose jth component $\sigma_{ij} = L_{ij}$ gives the lognormal volatility with respect to the jth risk factor.

Equation (1.61) in Section 1.2 gives $\mathbf{L}$ for the case $n = 2$. In particular, in the case of two stocks we can introduce a correlation ρ, where equations (1.179) now take the specific form

$$\frac{dS_t^1}{S_t^1} = \mu_1 \; dt + \sigma_1 \; dW_t^1, \tag{1.182}$$

$$\frac{dS_t^2}{S_t^2} = \mu_2 \; dt + \rho\sigma_2 \; dW_t^1 + \sqrt{1-\rho^2}\sigma_2 \; dW_t^2, \tag{1.183}$$

with infinitesimal variances and covariances

$$E_t\left[\left(\frac{dS_t^1}{S_t^1}\right)^2\right] = \sigma_1^2 \; dt, E_t\left[\left(\frac{dS_t^2}{S_t^2}\right)^2\right] = \sigma_2^2 \; dt, E_t\left[\frac{dS_t^1}{S_t^1}\frac{dS_t^2}{S_t^2}\right] = \rho\sigma_1\sigma_2 \; dt. \tag{1.184}$$

For this case the volatility vectors are given by $\boldsymbol{\sigma}_1 = (\sigma_1, 0)$ and $\boldsymbol{\sigma}_2 = \left(\rho\sigma_2, \sigma_2\sqrt{1-\rho^2}\right)$ for stock prices S_t^1 and S_t^2, respectively.

More generally, equations (1.179) [or (1.181)] describe geometric Brownian motion for an arbitrary number of n stocks with infinitesimal correlations and variances:

$$E_t\left[\frac{dS_t^i}{S_t^i}\frac{dS_t^j}{S_t^j}\right] = C_{ij} \; dt, \qquad E_t\left[\left(\frac{dS_t^i}{S_t^i}\right)^2\right] = \sigma_i^2 \; dt. \tag{1.185}$$

The vectors $\boldsymbol{\sigma}_i$ are seen to be given by the ith rows of matrix $\mathbf{L}$, i.e., the matrix of the lower Cholesky factorization of the covariance matrix.

A solution to the system of stochastic differential equations (1.179) [or (1.181)] is readily obtained by employing a simple change-of-variable approach (see Problem 4). In particular,

$$S_T^i = S_t^i \exp\left(\left(\mu_i - \frac{\sigma_i^2}{2}\right)(T-t) + \sigma_i \sum_{j=1}^{n} \Lambda_{ij} W_{T-t}^j\right); \qquad i = 1, \ldots, n, \tag{1.186}$$

where we denote $W^j_{T-t} = W^j_T - W^j_t$, for each jth independent Wiener normal random variable with mean zero and variance $T-t$; i.e., $W^j_{T-t} = \sqrt{T-t}x_j$, $x_j \sim N(0,1)$ independently for all $j = 1, \ldots, n$. From this result one readily obtains the multivariate lognormal distribution function $p(\mathbf{S}_T, \mathbf{S}_t; T-t)$, i.e., the analogue of equation (1.165) [see equation (1.198) in Problem 5]. The pricing of European-style options whose pay-offs depend on a group of n stocks, i.e., European basket options, can then proceed by computing expectations of such pay-offs over this density, where the drifts are set by risk neutrality. That is, let's assume a money-market account $B_t = e^{rt}$ with constant risk-free rate r, then within the risk-neutral measure the stock prices must all drift at the same rate, giving $\mu_i = r$.[10] Let V_t denote the option price at time t for a European-style contract with payoff function at maturity time T given by $V_T = \Pi(\mathbf{S}_T)$, $\mathbf{S}_T = (S^1_T, \ldots, S^n_T)$. The arbitrage-free price is then given by the expectation

$$\begin{aligned} V_t &= e^{-r(T-t)} E_t^{Q(B)}\big[\Pi(\mathbf{S}_T)\big] \\ &= e^{-r(T-t)} \int_{\mathbb{R}^n_+} p(\mathbf{S}_T, \mathbf{S}_t; T-t)\Pi(\mathbf{S}_T)d\mathbf{S}_\mathbf{T} \\ &= \frac{e^{-r(T-t)}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{1}{2}|\mathbf{x}|^2}\Pi(\mathbf{S}_T(\mathbf{x}))d\mathbf{x}, \end{aligned} \tag{1.187}$$

where $\mathbf{S}_T(\mathbf{x})$ has components $S^i_T(\mathbf{x})$ given by equation (1.186), $\mathbf{x} = (x_1, \ldots, x_n)$. The price hence involves an n-dimensional integral over a multivariate normal times some payoff function. Exact analytical expressions for basket options are generally difficult to obtain, depending on the type of payoff function as well as the number of dimensions n. Numerical integration methods can be used in general. Monte Carlo simulation methods are very useful for this purpose. The reader interested in gaining insight into the numerical implementation of standard Monte Carlo methods for pricing such options is referred to Project 8 on Monte Carlo pricing of basket options in Part II of this book.

Exact analytical pricing formulas for certain types of elementary basket options, however, can be obtained, as demonstrated in the following worked-out example.

Example. Chooser basket options on two stocks.

Consider a basket of two stocks with prices S^1_t (for stock 1) and S^2_t (for stock 2) modeled as before with constants μ_1, μ_2, ρ, σ_1, σ_2. Specifically, the risk-neutral geometric Brownian motions of the two stocks are given by

$$S^1_T = S^1_T(x_1, x_2) = S^1_0 e^{(r-\frac{\sigma_1^2}{2})T + \sigma_1\sqrt{T}x_1}, \tag{1.188}$$

$$S^2_T = S^2_T(x_1, x_2) = S^2_0 e^{(r-\frac{\sigma_2^2}{2})T + \sigma_2\sqrt{T}(\rho x_1 + \sqrt{1-\rho^2}x_2)}, \tag{1.189}$$

where S^1_0, S^2_0 are initially known stock prices at current time $t = 0$. The earlier pricing formula gives

$$V_0 = \frac{e^{-rT}}{2\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x_1^2+x_2^2)}\Pi(S^1_T(x_1,x_2), S^2_T(x_1,x_2))dx_1\, dx_2 \tag{1.190}$$

[10] This drift restriction is further clarified later in the chapter where we discuss the asset-pricing theorem in continuous time.

for the general payoff function. A *simple chooser* option is a European contract defined by the payoff $\max(S_T^1, S_T^2)$. This pay-off has a simple relation to other elementary pay-offs; i.e., $\max(S_T^1, S_T^2) = (S_T^2 - S_T^1)_+ + S_T^1 = (S_T^1 - S_T^2)_+ + S_T^2$. The current price V_0 of the simple chooser is hence given by $V_0 = C_0 + S_0^1$, where C_0 denotes the price of the contract with payoff $(S_T^2 - S_T^1)_+$. This follows since an expectation of a sum is the sum of expectations and from the fact that the stock prices drift at rate r; i.e., $e^{-rT} E_0^{Q(B)}[S_T^i] = S_0^i$. The problem remains to find the price C_0 given by the integral

$$C_0 = \frac{e^{-rT}}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x_1^2+x_2^2)} \big(S_T^2(x_1, x_2) - S_T^1(x_1, x_2)\big)_+ \, dx_1 \, dx_2. \tag{1.191}$$

The integrand is nonzero on the domain $\{(x_1, x_2) \in \mathbb{R}^2; S_T^2(x_1, x_2) > S_T^1(x_1, x_2)\}$. From equations (1.188) and (1.189) we find the domain is $\{(x_1, x_2) \in \mathbb{R}^2; x_1 < ax_2 + b\}$, where

$$a \equiv \frac{\sigma_2\sqrt{1-\rho^2}}{(\sigma_1 - \rho\sigma_2)}, \qquad b \equiv \frac{\log(S_0^2/S_0^1) + \frac{1}{2}(\sigma_1^2 - \sigma_2^2)T}{(\sigma_1 - \rho\sigma_2)\sqrt{T}}.$$

Here we assume $\sigma_1 - \rho\sigma_2 > 0$ and leave it to the reader to verify that a similar derivation of the same price given next also follows for the case $\sigma_1 - \rho\sigma_2 \leq 0$. Using this integration domain and inserting expressions (1.188) and (1.189) into the last integral gives

$$\begin{aligned} C_0 = {} & \frac{S_0^2 e^{-\frac{1}{2}\sigma_2^2 T}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x_2^2 + \sqrt{1-\rho^2}\sigma_2\sqrt{T}x_2} \left[\int_{-\infty}^{ax_2+b} e^{-\frac{1}{2}x_1^2 + \rho\sigma_2\sqrt{T}x_1} dx_1 \right] dx_2 \\ & - \frac{S_0^1 e^{-\frac{1}{2}\sigma_1^2 T}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x_2^2} \left[\int_{-\infty}^{ax_2+b} e^{-\frac{1}{2}x_1^2 + \sigma_1\sqrt{T}x_1} dx_1 \right] dx_2 \end{aligned}$$

By completing the square in the exponents, the integrals in x_1 give cumulative normal functions $N(\cdot)$. In particular,

$$\begin{aligned} C_0 = {} & \frac{S_0^2 e^{-\frac{1}{2}(1-\rho^2)\sigma_2^2 T}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x_2^2 + \sqrt{1-\rho^2}\sigma_2\sqrt{T}x_2} N(ax_2 + b - \rho\sigma_2\sqrt{T}) dx_2 \\ & - \frac{S_0^1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x_2^2} N(ax_2 + b - \sigma_1\sqrt{T}) dx_2. \end{aligned}$$

At this point we make use of the integral identity (see Problem 6),

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2 + Cx} N(Ax + B) dx = e^{\frac{1}{2}C^2} N\left(\frac{AC + B}{\sqrt{1 + A^2}}\right), \tag{1.192}$$

for any constants A, B, and C, giving

$$C_0 = S_0^2 N\left(\frac{(a\sqrt{1-\rho^2} - \rho)\sigma_2\sqrt{T} + b}{\sqrt{1+a^2}}\right) - S_0^1 N\left(\frac{b - \sigma_1\sqrt{T}}{\sqrt{1+a^2}}\right).$$

After a bit of algebra, using a and b just defined, we finally obtain the exact expression for the price in terms of the initial stock prices and the effective volatility ν as

$$C_0 = S_0^2 N(d_+) - S_0^1 N(d_-), \tag{1.193}$$

with

$$d_\pm = \frac{\log(S_0^2/S_0^1) \pm \frac{1}{2}\nu^2 T}{\nu\sqrt{T}}, \tag{1.194}$$

$\nu^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2.$

Changes of numeraire methods for obtaining exact analytical solutions in the form of Black–Scholes–type formulas for basket options on two stocks, as well as other options involving two correlated underlying random variables, are discussed later in this chapter.

Problems

Problem 1. Use Itô's lemma to verify that equation (1.154) provides a solution to equation (1.145).

Problem 2. Consider an exponential function of a normal random variable X, e^{aX} for any parameter a, where $X \in (-\infty, \infty)$ has probability density at $X = x$ given by

$$p(x, t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}, \qquad (t > 0).$$

Show that

$$E[e^{aX}] = \exp\big(a^2 t/2\big).$$

Hint: make use of the integral identity

$$\int_{-\infty}^{\infty} e^{-ax^2+bx}\, dx = \sqrt{\frac{\pi}{a}} e^{b^2/4a},$$

where $a > 0$ and b are constants.

Problem 3. Derive the expectation in equation (1.171) by making use of the identity $(a-b)_+ = (b-a)_+ + a - b$.

Problem 4. Consider the general correlated n-dimensional geometric Brownian process discussed in this section. Use Itô's lemma to show that the processes $Y_t^i \equiv \log S_t^i$ obey

$$dY_t^i = (\mu_i - \sigma_i^2/2)dt + \sigma_i \sum_{j=1}^{n} \lambda_{ij} dW_t^j. \tag{1.195}$$

Assuming all volatilities are nonzero, the correlation matrix is positive definite. Hence, $\boldsymbol{\lambda}$ has an inverse $\boldsymbol{\lambda}^{-1}$. Define new random variables $X_t^j \equiv \sum_{i=1}^{n} \sigma_i^{-1} \lambda_{ji}^{-1} Y_t^i$ and show that

$$dX_t^j = \tilde{\mu}_j\, dt + dW_t^j, \tag{1.196}$$

with $\tilde{\mu}_j \equiv \sum_{i=1}^{n} \sigma_i^{-1} \lambda_{ji}^{-1} (\mu_i - \frac{1}{2}\sigma_i^2)$, has solution

$$X_T^j = X_t^j + \tilde{\mu}_j (T - t) + W_T^j - W_t^j, \qquad j = 1, \ldots, n. \tag{1.197}$$

Invert this solution back into the old random variables, hence obtaining equation (1.186).

Problem 5. Treat W_{T-t}^j and $\log(S_T^i/S_t^i)$ as two sets of n independent variables in equation (1.186) and thereby compute the Jacobian of the transformation among the variables. Then invert equation (1.186) and use the identity in equation (1.48) with the distribution function for the n independent uncorrelated Wiener processes to show that the analytical formula for the transition probability density for geometric Brownian motion is given by

$$p(\mathbf{S}_T, \mathbf{S}_t; T - t) = (2\pi(T-t))^{-\frac{n}{2}} |\mathbf{C}|^{-\frac{1}{2}} \exp\big(-\tfrac{1}{2}\mathbf{z} \cdot \mathbf{C}^{-1} \cdot \mathbf{z}\big), \tag{1.198}$$

where the n-dimensional vector $\mathbf{z}$ has components

$$z_i \equiv \frac{\log(S^i_T/S^i_t) - (\mu_i - \frac{1}{2}\sigma_i^2)(T-t)}{\sqrt{T-t}}. \tag{1.199}$$

Problem 6. Using the definition of the cumulative normal function, write

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2+Cx} N(Ax+B)dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2+Cx}\left[\int_{-\infty}^{Ax+B} e^{-\frac{1}{2}y^2}dy\right]dx. \tag{1.200}$$

Introduce a change of variables $(\eta, \xi) \equiv (y - Ax, y + Ax)$ and integrate while completing squares to obtain equation (1.192).

1.7 Forwards and European Calls and Puts

Consider a situation with a stock price that at current time $t = 0$ has price S_0 while at time $T > 0$ in the future is described by a certain random variable S_T. Suppose that there is also a *zero-coupon bond* maturing at time T, i.e., a riskless claim to one unit of account at time T. Let

$$Z_t(T) = e^{-r(T-t)} \tag{1.201}$$

be its price at time t. Here r is the yield up to time T. Unlike the rate introduced in equation (1.5), in this case r is defined with the continuously compounded rule; we refer again to Chapter 2 for a more systematic discussion of fixed-income terminology.

Let's consider a situation where S_t is contained in the half-line of positive real numbers $\mathbb{R}_+$. Let P be the real-world measure with density *p(S)*; P is inferred through statistical estimations based on historical data. Pricing measures, instead, are evaluated as the result of a calibration procedure starting from option prices. Also, as discussed in detail later in this chapter, pricing measures depend on the choice of a *numeraire asset*. In our framework, a numeraire asset is given by an asset price process, g_t, that is strictly positive at initial time $t = 0$ and any other future time t, $t \leq T$. The corresponding pricing measure is denoted by *Q(g)*, specifying the fact that the asset price g_t is the chosen numeraire. A possible choice of numeraire is given by the bond $g_t = Z_t(T)$; this choice corresponds to the pricing measure denoted by *Q(Z(T))*, which is called the *forward measure*. Note that since r is constant, this also coincides with the risk-neutral measure. Technically speaking the name for the risk-neutral measure corresponds to using the continuously compounded money-market account $B_t = e^{rt}$ (i.e., the continuously compounded value of one unit of account deposited at time $t = 0$ earning interest rate r) as numeraire.[11] For constant interest rate, the two measures are then easily shown to be equivalent since $Z_t(T) = B_t/B_T$. This point is further clarified in Chapter 2. Other choices of numeraire asset are also possible; for example, $g_t = S_t$ corresponds to using the stock price as numeraire. As mentioned earlier and also described in detail later in the chapter, expectations taken based on the information available up to current time t with respect to the pricing measure *Q(g)*, with g_t as numeraire asset price, are denoted by $E_t^{Q(g)}[\,]$. In this section, note that (without loss in generality) we are simply setting $t = 0$ as current time and allowing T to be any future time.

[11] Note that we previously used the symbol B_t to denote the bond price. However, here we instead use B_t to denote the value of the money-market account.

By applying risk-neutral valuation to the zero-coupon bond, we find that

$$Z_0(T) = e^{-rT} = \alpha E_0^{Q(Z(T))}[Z_T(T)] = \alpha E_0^{Q(Z(T))}[1] = \alpha, \tag{1.202}$$

where $Z_T(T) = 1$. Hence, the discount factor α can be interpreted as the initial price of the zero-coupon bond. Although we have not yet formally introduced continuous-time financial models at this point in the chapter, the arguments presented in this section are generally valid if we assume dynamic trading is allowed in continuous time.

Risky assets are modeled by a function $\phi : \mathbb{R}_+ \to \mathbb{R}$ of the stock price at time T. Let $(A_t)_{0\le t\le T}$ be a price process such that $A_T = \phi(S_T)$; such an asset is called a *European-style option on the stock S with maturity T and payoff function* $\phi(S_T)$. Applying the asset-pricing theorem, the arbitrage-free price A_0 at time $t = 0$ of this option can be written as a discounted expectation under a pricing measure *Q(Z(T))*,

$$A_0 = e^{-rT} E_0^{Q(Z(T))}\left[\phi(S_T)\right]. \tag{1.203}$$

An alternative and instructive way of writing this equation is

$$\frac{A_0}{Z_0(T)} = E_0^{Q(Z(T))}\left[\frac{A_T}{Z_T(T)}\right]. \tag{1.204}$$

Although the numeraire asset in equation (1.204) is the riskless bond $Z_t(T)$, the pricing formula can be extended to the case of a generic numeraire asset g. Let's denote *Q(g)* as the probability measure, with g_t as numeraire asset price at time t, and defined so that

$$\frac{A_0}{g_0} = E_0^{Q(g)}\left[\frac{A_T}{g_T}\right] \tag{1.205}$$

for all random variables $A_T = \phi(S_T)$ and for all $T > 0$. Assuming the price is unique, equating the price A_0 in equation (1.204) with that in this last equation gives a relationship for the equivalence of the two pricing (or probability) measures:

$$g_0 E_0^{Q(g)}\left[\frac{\phi(S_T)}{g_T}\right] = Z_0(T) E_0^{Q(Z(T))}\left[\frac{\phi(S_T)}{Z_T(T)}\right]. \tag{1.206}$$

A variety of numeraire assets can be chosen for derivative pricing. Depending on the pay-off, one choice over another may be more convenient for evaluating the expectation and hence obtaining the derivative price, as seen in detail in the examples of pricing derivations in Section 1.12.

A *forward contract* on an underlying stock S stipulated at initial time $t = 0$ and with maturity time $t = T$ is a European-style claim with payoff $S_T - F_0$ at time T. Here F_0 is the *forward price* at time $t = 0$. Forward contracts are entered at the equilibrium forward price F_0, for which their present value is zero. A simple arbitrage argument gives a (*model-independent*) forward price F_0 as

$$F_0 = Z_0(T)^{-1} S_0. \tag{1.207}$$

Indeed, to replicate the pay-off of a forward contract one can buy the underlying stock at price S_0 and carry it to maturity while funding the purchase with a loan to be returned also at maturity. The nominal of the loan to be paid back at time T is then $Z_0(T)^{-1}S_0$ (e.g., this equals $e^{rT}S_0$ if we assume a constant interest rate).

Since the forward contract is initially worthless, the valuation formula yields

$$0 = E_0^{Q(Z(T))}\big[S_T - F_0\big]. \tag{1.208}$$

Since F_0 is constant, we have that

$$E_0^{Q(Z(T))}\big[S_T\big] = F_0 = Z_0(T)^{-1}S_0 = e^{rT}S_0. \tag{1.209}$$

The interpretation of this formula is that, under the pricing measure *Q(Z(T))*, the expected return on a stock is the risk-free yield r over the maturity T. The argument just outlined is model independent and can be shown to extend to all assets with no intermediate cash flows, thus no carry costs, before maturity time T. The expected return on any asset under the pricing measure $Q(Z(T))$ is the risk-free rate, no matter how volatile they are. Also notice that the expected return with respect to the real-world measure is quite different.

The popular geometric Brownian motion model, also called the *Black–Scholes model*, gives a lognormal risk-neutral probability density for the stock price process. As derived in Section 1.6, the stock price at time T is a lognormal random variable,

$$S_T = S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}x\right), \tag{1.210}$$

where $x \sim N(0, 1)$ and $\sigma > 0$ is the model volatility parameter. As we have seen, the risk-neutral distribution for S_T is defined in such a way as to satisfy the growth condition in equation (1.209)

$$E_0^{Q(Z(T))}[S_T] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}x\right)e^{-\frac{x^2}{2}}\,dx = S_0e^{rT}. \tag{1.211}$$

Two important examples of European-style securities are the *call option struck at K and of maturity T* with price process C_t and payoff function

$$C_T \equiv (S_T - K)_+ \tag{1.212}$$

and the *put option struck at K and of maturity T* with price process P_t and payoff function

$$P_T \equiv (K - S_T)_+. \tag{1.213}$$

Theorem 1.3. (Put-Call parity). *If $C_0(S_0, K, T)$ and $P_0(S_0, K, T)$ denote the prices at time $t = 0$ of a plain European call and a plain European put, respectively, both maturing at a later time T and both struck at K, then we have the put-call parity relationship, namely,*

$$C_0(S_0, K, T) - P_0(S_0, K, T) = S_0 - KZ_0(T). \tag{1.214}$$

The proof of the put-call parity relationship descends from the fact that a portfolio with a long position in a call struck at K and maturing at T and a short position in a put struck at K and maturing at T has the same pay-off as a forward contract stipulated at the forward price K. (See Section 1.8.)

In contrast to the put-call parity relationship in equation (1.214), the evaluation of the price of a call or put option requires making an assumption on the measure *Q(Z(T))* and the stock price process. Under the Black–Scholes model, where the stock at time T is given by equation (1.210), the expectation $E_0^{Q(Z(T))}[(S_T - K)_+]$ can be reduced to a simple integral. As shown in a detailed derivation in Section 1.6,

$$E_0^{Q(Z(T))}[(S_T - K)_+] = S_0 e^{rT} N(d_+) - KN(d_-), \tag{1.215}$$

where $N(\cdot)$ is the standard cumulative normal distribution function,

$$d_\pm = \frac{\log(S_0/K) + (r \pm \sigma^2/2)T}{\sigma\sqrt{T}}, \tag{1.216}$$

and the pricing formula for a plain European call option (with constant interest rate) in the Black–Scholes model is

$$\begin{aligned} C_{BS}(S_0, K, T, \sigma, r) &= e^{-rT} E_0^{Q(Z(T))}[(S_T - K)_+] \\ &= S_0 N(d_+) - Ke^{-rT} N(d_-). \end{aligned} \tag{1.217}$$

European put options are priced analytically in similar fashion by computing the expectation $e^{-rT} E_0^{Q(Z(T))}[(K - S_T)_+]$, as seen in the derivation of equation (1.172) of Section 1.6. From this formula, or by applying the put-call parity relation (1.214) using equation (1.217), we have the equivalent formulas for the put option price:

$$\begin{aligned} P_{BS}(S_0, K, T, \sigma, r) &= e^{-rT} E_0^{Q(Z(T))}[(K - S_T)_+] \\ &= S_0 N(d_+) - Ke^{-rT} N(d_-) - S_0 + Ke^{-rT} \\ &= Ke^{-rT} N(-d_-) - S_0 N(-d_+). \end{aligned} \tag{1.218}$$

A direct calculation shows that the functions C_{BS} and P_{BS} satisfy the *Black–Scholes partial differential equation* (BSPDE). Analytical and numerical methods for solving this equation are discussed at length throughout later sections and chapters of this book. The numerical projects in Part II provide implementation details for finite-difference lattice approaches to option pricing. A derivation of the BSPDE based on a dynamic replication strategy is provided in Section 1.9 (and a general derivation is given in Section 1.13), but here we simply quote it for the purposes of the present discussion. In terms of the partial derivatives with respect to the time to maturity T and current stock price S_0 (with r and σ constants) this equation can be rewritten in the form

$$\frac{\partial V}{\partial T} = \frac{\sigma^2 S_0^2}{2} \frac{\partial^2 V}{\partial S_0^2} + rS_0 \frac{\partial V}{\partial S_0} - rV, \tag{1.219}$$

where the option value $V = V(S_0, T)$. The original Black–Scholes equation is really a backward-time equation involving $\partial V/\partial t$ in calendar time t, where the price V is expressed in terms of t and equals the pay-off at maturity (or expiry) $t = T$. That is, if we were to express the option value explicitly in terms of such a function of calendar time t, then, for example, for the case of a call struck at K, $C(S, t = T) = (S - K)_+$. Note that in the present context, however, since we are expressing the option value with respect to the time to maturity, denoted here by the variable T, the option price equals the pay-off when $T = 0$ (i.e., at zero time to expiry): $C_{BS}(S, K, T = 0) = (S - K)_+$ and $P_{BS}(S, K, T = 0) = (K - S)_+$, as is easily verified via equations (1.217) and (1.218) in the limit $T \to 0$. Since the Black–Scholes equation is time homogeneous for time-independent interest rate and volatility, option prices are generally functions of $T - t$ (where t and $T \geq t$ represent actual calendar times), so $\partial/\partial t = -\partial/\partial T$

in the original Black–Scholes equation. By replacing $T - t \to T$ (without loss in generality this corresponds to setting current time $t = 0$), we further simplify all expressions, wherein T now represents the time to maturity. The form in equation (1.219) is convenient for the following discussion.

Whether the pricing measure *Q(Z(T))* is unique or not depends on the choice of hedging instruments. The asset-pricing theorem (in the single-period setting as stated earlier and in the continuous-time case discussed later in this chapter) only implies that — assuming absence of arbitrage — there exists such a measure and that this measure prices all pay-offs. Indeterminacies in *Q(Z(T))* arise in case there is no perfect replication strategy for the given pay-off, which can be priced independently. The Black–Scholes model provides the most basic pricing model that captures option prices through the single volatility parameter σ. Since in finance there is no fundamental theory ruling asset price processes, all models are inaccurate to some degree. The Black–Scholes model is perhaps the most inaccurate among all those used, but also the most basic because of its simplicity. Inaccuracies in the Black–Scholes model are captured by the *implied volatility surface*, defined as the function $\sigma_{BS}(K, T)$ such that

$$C_{BS}(S_0, K, T, \sigma_{BS}(K, T), r) = C_0(K, T), \tag{1.220}$$

where $C_0(K, T)$ is the *observed market price* of the call option struck at K and maturing at time T. This describes a surface $\sigma_I = \sigma_{BS}(K, T)$ in which the implied volatility σ_I is graphed as a function of two variables K, T, i.e., across a range of strikes K and time to maturity values T. For any fixed pair of values *(K,T)* (and assumed fixed S_0, r), the function C_{BS} is monotonically increasing in σ [see equation (1.222)], hence the preceding equation can be uniquely inverted to give a value for the so-called *Black–Scholes implied volatility* σ_I for any observed market price of a call. If the Black–Scholes (i.e., lognormal) model were accurate, the implied volatility surface would be flat and constant, for one single volatility parameter would price all options. Empirical evidence shows that implied volatility surfaces are instead curved (not flat!).

A practical and widely used approach to risk management involving the Black–Scholes pricing formulas is based on the calculation of portfolio sensitivities. Sensitivities of option prices in the Black–Scholes model with respect to changes in the underlying parameters r, T, S, σ are of importance to hedging and computing risk for nonlinear portfolios. Within the Black–Scholes formulation, these sensitivities are easily obtained analytically by taking the respective partial derivatives of the European-style option price V for a given pay-off. The list of sensitivities (also known as the *Greeks*) are defined as follows, where we specialize to provide the exact expressions for the case of a plain-vanilla call under the Black–Scholes model:

- The *delta*, denoted by Δ, is defined as the derivative

$$\Delta = \frac{\partial V}{\partial S_0} = \frac{\partial C_{BS}}{\partial S_0} = N(d_+). \tag{1.221}$$

- The *vega*, denoted by Λ, is defined as the derivative

$$\Lambda = \frac{\partial V}{\partial \sigma} = \frac{\partial C_{BS}}{\partial \sigma} = S_0 \sqrt{T} \frac{e^{-d_+^2/2}}{\sqrt{2\pi}}. \tag{1.222}$$

- The *gamma*, denoted by Γ, is defined as the second derivative

$$\Gamma = \frac{\partial^2 V}{\partial S_0^2} = \frac{\partial^2 C_{BS}}{\partial S_0^2} = \frac{e^{-d_+^2/2}}{\sigma S_0 \sqrt{2\pi T}}, \tag{1.223}$$

- The *rho*, denoted by ρ, is defined as the derivative

$$\rho = \frac{\partial V}{\partial r} = \frac{\partial C_{BS}}{\partial r} = KTe^{-rT} N(d_-), \tag{1.224}$$

- The *theta*, denoted by Θ, is defined as the derivative[12]

$$\Theta = \frac{\partial V}{\partial T} = \frac{\partial C_{BS}}{\partial T} = (\sigma^2 S_0^2/2)\Gamma + r(S_0\Delta - C_{BS}). \tag{1.225}$$

The numerical project called "The Black–Scholes Model" in Part II provides the interested reader with an in-depth implementation of such formulas for calls as well as for puts and so-called butterfly spread options. The corresponding spreadsheet is then useful for numerically graphing and analyzing the dependence of the various option prices and their sensitivities as functions of either r, σ, S_0, K, or T.

Given the sensitivities, one can approximate the change in price δC of a call option due to small changes $T \to T + \delta T$, $S_0 \to S_0 + \delta S_0$, $\sigma \to \sigma + \delta\sigma$, $r \to r + \delta r$ by means of a truncated Taylor expansion,

$$\delta C \cong \Delta(\delta S_0) + \Lambda(\delta\sigma(K, T)) + \frac{1}{2}\Gamma(\delta S_0)^2 + \rho(\delta r) + \Theta(\delta T). \tag{1.226}$$

Here, δS_0, δr, $\delta\sigma(K, T)$, and δT are small changes in the stock price, the interest rate, the implied Black–Scholes volatility $\sigma = \sigma(K, T)$, and the time to maturity T of the option at hand. In the Black–Scholes model, $\sigma(K, T)$ does not depend on the two arguments and these parameters are constant, so the only source of randomness is the price of the underlying. However, in practice one observes that implied volatilities and interest rates also change over time and affect option values.

As we discuss in more detail in Chapter 4, the risk of option positions is hedged on a portfolio basis and risk-reducing trades are placed in such a way as to decrease portfolio sensitivities to the underlyings. In particular:

- The delta can be reduced by taking a position in the stock or, more commonly, in a forward or futures contract on the stock.
- The vega and gamma can be reduced by taking a position in another option.
- The rho can be reduced by taking a position in a zero-coupon bond of maturity T.

Problems

Problem 1. Derive the formulas in equations (1.221)–(1.225).

Problem 2. Obtain formulas analoguous to equations (1.221)–(1.225) for the corresponding put option with value P_{BS}.

[12] In other literature this is sometimes defined as $-\partial V/\partial T$.

Problem 3. Consider a portfolio with positions θ_i in N securities, each with price f_i, $i = 1, \ldots, N$, respectively. Assume the security prices are functions of the same spot S_0 at current time t_0 and that each price function $f_i = f_i(S_0, T_i - t_0)$ satisfies the time-homogeneous BSPDE with constant interest rate and volatility. The contract maturity dates T_i are allowed to be distinct. Find the relation between the Θ, Δ, and Γ of the portfolio.

1.8 Static Hedging and Replication of Exotic Pay-Offs

Options other than the calls and puts considered in the previous section are often called *exotic*. In this section, we consider the replication of arbitrary pay-offs via portfolios made up of standard instruments (i.e., consisting of calls, puts, underlying stock, and cash). In finance, such replicating portfolios are useful for the static hedging of European-style options.

A *butterfly spread* option maturing in time T is a portfolio of three calls with current value

$$B_0(S_0, K, T, \epsilon) = \frac{1}{\epsilon^2}(C_0(S_0, K-\epsilon, T) + C_0(S_0, K+\epsilon, T) - 2C_0(S_0, K, T)), \qquad (1.227)$$

for some $\epsilon > 0$, where $C_0(S_0, K, T)$ represents the (model-independent) price of a European call with current stock price S_0, strike K, and time to maturity T. We observe that (apart from the normalization constant) this option consists of a long position in a call struck at $K+\epsilon$, a long position in a call struck at $K-\epsilon$, and two short positions in a call struck at K, with all calls maturing at the same time. At expiry $T \to 0$ we simply have the payoff function for the butterfly spread:

$$\begin{aligned}\delta_\epsilon(S_T - K) &= \frac{1}{\epsilon^2}(C_T(S_T, K-\epsilon) + C_T(S_T, K+\epsilon) - 2C_T(S_T, K)) \\ &= \frac{1}{\epsilon^2}\begin{cases}(S_T - (K-\epsilon))_+, & S_T \le K \\ ((K+\epsilon) - S_T)_+, & S_T > K.\end{cases}\end{aligned} \qquad (1.228)$$

Here we have used $C_T(S_T, K) \equiv (S_T - K)_+$ for the pay-off of a call. The normalization factor hence ensures that the area under the graph of the pay-off (as function of S_T) is unity, for all choices of ϵ (see Figure 1.3). In the limit $\epsilon \to 0$, the function $\delta_\epsilon(S_T - K)$ converges to the Dirac delta function $\delta(S_T - K)$ (see Problem 1).

From the one-dimensional version of equation (1.27), we have

$$\lim_{\epsilon \to 0}\int_0^\infty \delta_\epsilon(S_T - K) f(K)dK = \int_0^\infty \delta(S_T - K) f(K)dK = f(S_T), \qquad (1.229)$$

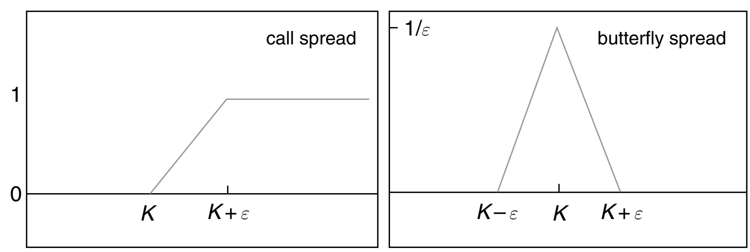

FIGURE 1.3 Payoff functions for a call spread and a corresponding unit butterfly spread struck at K, where 2ϵ is the width of the butterfly spread.

for any $S_T > 0$ and any continuous function f. From the linearity property of expectations and risk-neutral pricing we must have

$$B_0(S_0, K, T, \epsilon) = e^{-rT} E_0^Q[\delta_\epsilon(S_T - K)]. \tag{1.230}$$

In particular, we find that in the limit $\epsilon \to 0$,

$$\begin{aligned}
\lim_{\epsilon\to 0} B_0(S_0, K, T, \epsilon) &= \lim_{\epsilon\to 0} e^{-rT} E_0^Q\left[\delta_\epsilon(S_T - K)\right] \\
&= e^{-rT} \lim_{\epsilon\to 0} \int_0^\infty p(S_0, 0; S_T, T)\delta_\epsilon(S_T - K)dS_T \\
&= e^{-rT} \int_0^\infty p(S_0, 0; S_T, T)\delta(S_T - K)dS_T \\
&= e^{-rT} p(S_0, 0; K, T),
\end{aligned} \tag{1.231}$$

where $p(S_0, 0; K, T)$ is the risk-neutral probability density that the stock price S_T equals K at time $t = T$, conditional to its equaling S_0 at initial time $t = 0$. This result basically tells us that the price of an infinitely narrow butterfly spread is the price of a so-called Arrow–Debreu security, i.e., the value of a security that pays one unit of account if the stock price (i.e., the state) $S_T = K$ is attained at maturity. One concludes that knowledge of the prices of European calls at all strikes is equivalent to the knowledge of the risk-neutral transition probability density $p(S_0, 0; S_T, T)$ for all S_T. Notice, though, that this does not uniquely identify the price process under the risk-neutral measure because all possible transition probabilities $p(S_t, t; K, T)$ for any $t > 0$ are not uniquely determined.[13] By recognizing that equation (1.227) is in fact a representation of the finite difference for the second derivative, we obtain from the last equation

$$\frac{\partial^2 C_0(S_0, K, T)}{\partial K^2} = e^{-rT} p(S_0, 0; K, T). \tag{1.232}$$

We will arrive at this equation again in Section 1.13 when we discuss the Black–Scholes partial differential equation and its dual equation.

Other common portfolios of trades include the following.

- *Covered calls* consist of a long position in the underlying and a short position in a call, typically struck above the spot at the contract inception. This position is meant to trade potential returns above the strike at future time for the option price. The pay-off at the option maturity is

$$S_T - (S_T - K)_+. \tag{1.233}$$

- *Bull spreads* are option spread positions consisting of one long call struck at K_1 and one short call struck at K_2 with payoff function

$$(S_T - K_1)_+ - (S_T - K_2)_+, \tag{1.234}$$

  $K_1 < K_2$. This portfolio is designed to profit from a rally in the price of the underlying security.

[13] There are in general a variety of models involving jumps, stochastic or state-dependent volatility, or a combination of all that result in the same prices for European options but yield different valuations for path-dependent pay-offs.

- *Bear spreads* are option spread positions in one short put struck at K_1 and one long put struck at K_2 with payoff function

$$-(K_1 - S_T)_+ + (K_2 - S_T)_+, \tag{1.235}$$

$K_1 < K_2$. This portfolio profits from a decline in price of the underlying security.
- *Digitals* obtain in the limit that $(K_2 - K_1) \to 0$ in a spread option with positions scaled by the strike spread $(K_2 - K_1)^{-1}$. A digital is also called a *binary*. For instance, the pay-off of a bull digital (or digital call) is a unit step function obtained when such a limit is taken in a bull spread with $(K_2 - K_1)^{-1}$ long positions in a call struck at K_1 and $(K_2 - K_1)^{-1}$ short positions in a call struck at K_2, with $K_1 < K_2$:

$$\theta(S_T - K) = \begin{cases} 1 & \text{if } S_T \geq K \\ 0 & \text{otherwise} \end{cases}. \tag{1.236}$$

The bear digital (or digital put) obtains similarly by considering the limiting case of the bear spread, and the pay-off is $\theta(K - S_T) = 1 - \theta(S_T - K)$, giving 1 if $S_T < K$ and zero otherwise.
- *Wingspreads (also called Condors)* consist of two long and two short positions in calls. These are similar to butterfly spreads, except the body of the payoff function has a flat maximum instead of a vertex; in formulas, the payoff function is

$$(S_T - K_1)_+ - (S_T - K_2)_+ - (S_T - K_3)_+ + (S_T - K_4)_+, \tag{1.237}$$

with $K_1 < K_2 < K_3 < K_4$ and $K_2 - K_1 = K_4 - K_3$.
- *Straddles* involve the simultaneous purchase or sale of an equivalent number of calls and puts on the same underlying with the same strike and same expiration. The straddle buyer speculates that the realized volatility up to the option's maturity will be large and cause large deviations for the price of the underlying asset. The pay-off is

$$(S_T - K)_+ + (K - S_T)_+. \tag{1.238}$$

- *Strangles* are similar to straddles, except the call is struck at a different level than the put; i.e.,

$$(S_T - K_1)_+ + (K_2 - S_T)_+, \tag{1.239}$$

with $K_1 > K_2$ or $K_1 < K_2$. The case $K_1 < K_2$ is an in-the-money strangle, and $K_1 > K_2$ is an out-of-the-money strangle, since the minimum payoff values attained are $K_2 - K_1$ and zero, respectively.
- *Calendar spreads* are spread options where the expiration dates are different and the strike prices are the same, for example:

$$(S_{T_1} - K)_+ - (S_{T_2} - K)_+, \tag{1.240}$$

with $T_1 \neq T_2$. This option strategy is added here for completeness, although it differs from all of the foregoing because the portfolio involves options of varying expiry dates.

Consider the problem of replicating a generic payoff function $\phi(S)$, $0 < S < \infty$, assumed throughout to be twice differentiable. By virtue of equation (1.229), one can achieve replication by means of positions in infinitely narrow butterfly spreads of all possible strikes.

A perhaps more instructive replication strategy involves positions in the underlying stock, a zero-coupon bond and European call options, of all possible strikes and fixed expiration time T. Assuming $\phi(0)$ $\phi'(0)$ exist, the formula is

$$\phi(S) = \phi(0) + \phi'(0)S + \int_0^\infty n(K)C_T(S, K)dK. \tag{1.241}$$

$n(K)dK$ represents the size of the position in the call of strike K. The function *n(S)* is related to the payoff function and can be evaluated by differentiating equation (1.241) twice:

$$\phi''(S) = \int_0^\infty n(K)\delta(S-K)dK = n(S). \tag{1.242}$$

Here we make use of the identity

$$\frac{\partial^2}{\partial S^2}(S-K)_+ = \frac{\partial^2}{\partial K^2}(S-K)_+ = \delta(S-K). \tag{1.243}$$

As shown in Problem 3 of this section, equation (1.241) can be derived via an integration-by-parts procedure. The conclusion we can draw is that if calls of all strikes are available, the arbitrage-free price $f_0 = f_0(S_0, T)$ at time $t = 0$ of a contingent European claim with payoff $\phi(S_T)$ at maturity $t = T$ is

$$f_0 = \phi(0)Z_0(T) + \phi'(0)S_0 + \int_0^\infty \phi''(K)C_0(S_0, K, T)dK. \tag{1.244}$$

Besides the basic assumption that asset prices satisfy equation (1.205), it is crucial to point out that the foregoing replication formulas follow without any assumption on the model of the underlying stock motion; i.e., the replication equations are also true by assuming a stochastic process of a more general form that includes the lognormal model as a special case. Moreover, these equations can be extended to apply to a payoff $\phi(S)$ defined on a region $S \in [S_0, S_1]$, where S_0, S_1 may be taken as either finite or infinite. Specifically, let us consider the space $[S_0, S_1]$, then, using the delta function integration property[14] and assuming $\phi(S_0)$, $\phi'(S_0)$ exist, one can derive

$$\phi(S) = \phi(S_0) + \phi'(S_0)(S-S_0) + \int_{S_0}^{S_1} \phi''(K)(S-K)_+ \, dK. \tag{1.245}$$

The discretized form of this formula reads

$$\phi(S) \approx \phi(S_0) + \phi'(S_0)(S-S_0) + \sum_{i=1}^{N}(\Delta K_i)\phi''(K_i)(S-K_i)_+, \tag{1.246}$$

where K_i are chosen as $S_0 < K_1 < K_2 < \cdots < K_N < S_1$. Let us assume that the strikes are chosen as equally spaced, $\Delta K_i = K_i - K_{i-1} = \Delta K$. Hence, the replication consists of a cash position of size $\phi(S_0) - \phi'(S_0)S_0$, a stock position of size $\phi'(S_0)$, and N call positions of size $(\Delta K_i)\phi''(K_i)$ in calls struck at K_i. In most practical cases, this formula actually offers a more accurate discrete representation than the analogous form obtained from discretizing the integral in equation (1.241). This is especially the case when considering a pay-off whose nonzero values are localized to a region $[S_0, S_1]$ for finite S_1 or to a region $[S_0, \infty)$, with $S_0 > 0$.

[14] Here one uses the general property $\int_{S-\beta}^{S+\eta} \delta(S-K)\phi(K)dK = \phi(S)$ for any real constants $\beta, \eta > 0$.

This is the situation for pay-offs of the general form $\Lambda(S, X)\mathbf{1}_{\mathcal{A}}$, for some function $\Lambda(S, X)$ with strike $X > 0$. Here $\mathbf{1}_{\mathcal{A}}$ is the indicator function having nonzero value only if condition $\mathcal{A}$ is satisfied. If $\mathcal{A}$ is chosen as the condition $S > X$, then $\mathbf{1}_{S>X} = \theta(S - X)$. The plain European call pay-off obtains with the obvious choice $\Lambda(S, X) = S - X$. It should also be noted that an alternate replication formula involving puts at various strikes (instead of calls) is readily obtained in a manner similar as before or by a simple application of put-call parity (see Problem 6), giving

$$\phi(S) = \phi(S_1) + \phi'(S_1)(S - S_1) + \int_{S_0}^{S_1} \phi''(K)(K - S)_+ \, dK, \tag{1.247}$$

assuming that $\phi(S_1)$, $\phi'(S_1)$ exist.

Note that these formulas assume that the payoff function is well behaved at either the lower endpoint or the upper endpoint. A formula that is valid irrespective of whether the payoff function is singular at either endpoint can be obtained by subdividing the interval $[S_0, S_1]$ into two regions: a lower region $[S_0, \bar{S}]$ and an upper region $[\bar{S}, S_1]$ for any $\bar{S}$ with $S_0 < \bar{S} < S_1$. In the lower region we use puts, while calls are used for the upper region. In particular, via a straightforward integration-by-parts procedure one can derive (see Problem 7)

$$\phi(S) = \phi(\bar{S}) + \phi'(\bar{S})(S - \bar{S}) + \int_{S_0}^{\bar{S}} \phi''(K)(K - S)_+ \, dK + \int_{\bar{S}}^{S_1} \phi''(K)(S - K)_+ \, dK. \tag{1.248}$$

One is then at liberty to choose $\bar{S}$, which acts as a kind of separation boundary for whether calls or puts are used. Note that in the limit $\bar{S} \to S_0$ the formula reduces to that in equation (1.245), with only calls being used, while the opposing limit $\bar{S} \to S_1$ gives equation (1.247), with only puts used for replication. A similar approximate discretization scheme as discussed earlier may be used for these integrals, giving rise to a replication in terms of a finite number of calls and puts at appropriate strikes. This last formula may hence prove advantageous in practice when liquidity issues are present. In particular, this replication can be exploited to better balance the use of available market contracts that are either in-the-money or out-of-the-money puts or calls.

We now give some examples of applications of the foregoing replication theory.

Example 1. Exponential Pay-Off.

As a first example, let

$$\phi(S) = (e^{S-X} - 1)_+ = [e^{S-X} - 1]\theta(S - X) = \begin{cases} e^{S-X} - 1, & S \geq X \\ 0, & S < X. \end{cases} \tag{1.249}$$

One can readily verify that this payoff function can be exactly replicated using the right-hand side of either equation (1.241) or equation (1.245) with $S_1 = \infty$. Using $\phi(X) = 0$, $\phi'(K) = \phi''(K) = e^{K-X}$ (for $K > X$), and adopting the replication formula in equation (1.246) with $S_0 = X$ and any $S_1 > X$ gives

$$\phi(S) \approx S - X + \sum_{i=1}^{N} w_i (S - K_i)_+, \tag{1.250}$$

with call positions (i.e., weights) $w_i = (\Delta K)e^{K_i - X}$ and strikes $K_i = X + i\,\Delta K$. Note that one may also use slightly different subdivisions, all of which converge to the same result in the

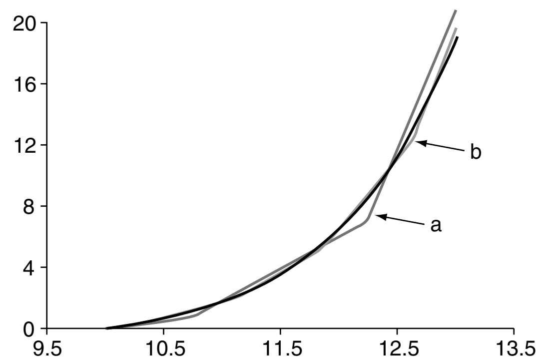

FIGURE 1.4 Rapid convergence of the static replication of the exponential pay-off defined in equation (1.249) (in the region $[X, X+L]$ with $X = 10$, $L = 3$) using equation (1.250) with a sum of (a) two calls with $K_1 = 10.75$, $K_2 = 12.25$ versus (b) four calls with $K_1 = 10.375$, $K_2 = 11.125$, $K_3 = 11.875$, $K_4 = 12.625$.

limit of infinitesimal spacing $\Delta K \to 0$. Figure 1.4 partly shows the result of this replication strategy in practice. Nearly exact replication is already achieved with only eight strikes.

Example 2. Sinusoidal Pay-Off.

Consider the sinusoidal pay-off

$$\phi(S) = \sin\left(\frac{\pi(S-X)}{L}\right)\mathbf{1}_{X \le S \le X+L}, \qquad X, L > 0. \tag{1.251}$$

The choice of strikes $K_i = X + iL/N$, $i = 1, \ldots, N$, with $S_0 = X$ and $S_1 = X + L$, within equation (1.246) gives

$$\phi(S) \approx \frac{\pi}{L}(S - X) + \sum_{i=1}^{N} w_i (S - K_i)_+, \tag{1.252}$$

where $w_i = -(\pi^2/NL)\sin(i\pi/N)$. Figure 1.5 shows the convergence using this replication strategy.

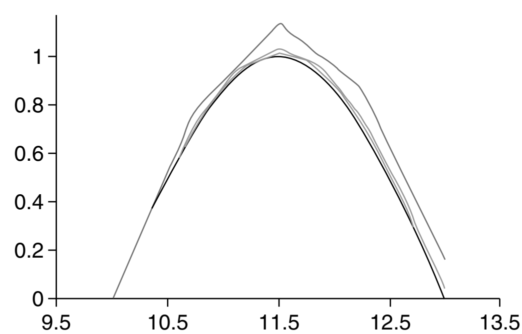

FIGURE 1.5 A comparison of three replication curves and the exact sine pay-off defined in equation (1.251) (in the region $[X, X+L]$ with $X = 10$, $L = 3$) with $N = 4$, $N = 8$, and $N = 12$ short calls, a long position in the stock, and a short cash position using equation (1.252). With $N = 12$ the replication is already very accurate.

Example 3. Finite Number of Market Strikes.

In realistic applications there typically is only a select number of strikes available in the market, so the trader has no control over the values of K_i to be used in the replication strategy. In this situation the set of calls (puts) with strikes K_i, $i = 1, \ldots, N$, is already given (i.e., preassigned) for some fixed N, and the spacing between strikes is not necessarily uniform. A solution to this problem is to consider a slight variation to equation (1.246) and write the finite expansion

$$\phi(S) \approx w_{-1} + w_0 S + \sum_{i=1}^{N} w_i (S - K_i)_+. \tag{1.253}$$

The coefficient w_{-1} gives the cash position, while the weight w_0 gives the stock position, and the weights w_i give the positions in the calls struck at values K_i. The goal is to find the positions w_i providing the best fit, in the linear least squares sense, as follows. By subdividing the stock price space $[S_0, S_1]$ into M interval slices $S^{(j)}$, with $S^{(j)} < S^{(j+1)}$, $j = 1, \ldots, M$, the $N+2$ positions w_i can be determined by matching the approximate payoff function on the right-hand side of equation (1.253) to the value of the exact payoff function $\phi(S^{(j)})$ at these M stock points. This leads to a linear system of M equations in the $N+2$ unknown weights w_i:

$$\phi(S^{(j)}) = w_{-1} + w_0 S^{(j)} + \sum_{i=1}^{N} w_i (S^{(j)} - K_i)_+, \qquad j = 1, \ldots, M. \tag{1.254}$$

One can always make the choice $M \geq N+2$ so that there are at least as many equations as unknown weights. A solution to this system can be found within the linear least squares sense, giving the w_i. This technique is fairly robust and also offers a rapidly convergent replication. The reader interested in gaining further experience with the actual numerical implementation of this procedure as applied to logarithmic pay-offs is referred to the numerical project in Part II of this book dealing specifically with the replication of the static component of variance swap contracts.

Problems

Problem 1. A particular representation of the Dirac delta function $\delta(x)$ is given by the limit $\epsilon \to 0$ of the sequence of functions $f_\epsilon(x) = (1/\epsilon^2)(\epsilon - |x|)_+$. Using this fact, demonstrate that the butterfly spread pay-off defined in equation (1.228) gives the Dirac delta function $\delta(S_T - K)$ in the limit $\epsilon \to 0$.

Problem 2. Consider the bull spread portfolio with maximum pay-off normalized to unity:

$$\frac{C_T(S, K+\epsilon) - C_T(S, K)}{\epsilon}, \tag{1.255}$$

$C_T(S, K) = (S-K)_+$. Compute the limit $\epsilon \to 0$ and thereby obtain the pay-off of a bull digital.

Problem 3. Show that under suitable assumptions on the function ϕ [i.e., $\phi(0)$ and $\phi'(0)$ exist] we have

$$\int_0^\infty \phi''(K)(S-K)_+ \, dK = \phi(S) - \phi'(0)S - \phi(0), \tag{1.256}$$

hence verifying equation (1.241). For this purpose use integration by parts twice, together with the property in equation (1.243) as well as the identity

$$\frac{\partial}{\partial S}(S-K)_+ = \theta(S-K), \tag{1.257}$$

where $\theta(x)$ is the Heaviside unit step function having value 1, or 0 for $x \geq 0$, or $x < 0$, respectively. Note that the derivative of this function gives the Dirac delta function.

Problem 4. Demonstrate explicitly that the pay-offs of Examples 1 and 2 of this section satisfy equation (1.245) with $S_0 = X$, $S_1 = X + L$, $L > 0$.

Problem 5. Assume that calls of all strikes are available for trade and have a known price. Express the present value of the log payoff $\phi(S_T) = \log \frac{S_T + a}{S_0}$, with constant $a > 0$, in terms of call option prices of all strikes $K > 0$. Find a similar expression in terms of put option prices.

Problem 6. Apply equation (1.241) to a call payoff $\phi(S) = (S - X)_+$, with constant X, to obtain the put-call parity relation

$$(S - X)_+ = S - X + (X - S)_+, \tag{1.258}$$

for all $S > 0$. In deriving this result, the property in equation (1.243) is useful. Now make use of the right-hand side of this put-call parity formula into equation (1.245) and integrate by parts to arrive at equation (1.247).

Problem 7. Consider the interval $S \in [S_0, S_1]$. Integrate by parts twice while using the general properties stated earlier for the functions $\theta(x)$, $(x)_+$, and the delta function $\delta(x)$ to arrive at the identities

$$\int_{S_0}^{\bar{S}} \phi''(K)(K - S)_+ dK = \phi(S)\mathbf{1}_{S_0 < S < \bar{S}} - \phi(\bar{S})\theta(\bar{S} - S) + \phi'(\bar{S})(\bar{S} - S)_+ \tag{1.259}$$

and

$$\int_{\bar{S}}^{S_1} \phi''(K)(S - K)_+ dK = \phi(S)\mathbf{1}_{\bar{S} \leq S < S_1} - \phi(\bar{S})\theta(S - \bar{S}) - \phi'(\bar{S})(S - \bar{S})_+ \tag{1.260}$$

where $\mathbf{1}_{\mathcal{D}}$ is the indicator function having unit value for the domain $\mathcal{D}$ and zero otherwise. Add these two expressions to finally obtain equation (1.248).

Problem 8. Using risk-neutral valuation, i.e., equation (1.166), derive the Black–Scholes pricing formula for the price of a European digital call and that of a digital put struck at K with time to maturity T. For simplicity assume geometric Brownian motion with constant interest rate and volatility. Interpret the meaning of the digital option prices in terms of the price of a standard call. Hint: The derivation of the European digital call boils down to computing the risk-neutral probability $P(S_T \geq K)$, where the algebraic steps are similar to what is used to derive a standard call price.

Problem 9. Derive the Greeks Δ, Γ, and vega for a European digital call.

1.9 Continuous-Time Financial Models

In this section, we introduce the basic concepts in continuous-time finance. Derivative claims are structured as contracts written on underlying assets that can be used as hedging instruments. An elegant mathematical structure underlying these financial concepts is reviewed in this section.

In perfect-markets models, a basic asset price process is given by a *money-market account* on which we can deposit and out of which we can borrow without limits. The value at time t of one dollar deposited in a money-market account at initial time $t = 0$ with continuously compounded interest up to time t, is denoted by B_t.

Definition 1.10. Money-Market Account. *Assuming continuous compounding, a money-market account is an asset price process B_t that is monotonically increasing in time, has zero volatility, and follows an equation of the form*

$$dB_t = r_t B_t \, dt, \tag{1.261}$$

where r_t is a stochastic process that is positive at all times.[15] *By integrating equation (1.261) we find the stochastic integral representation*

$$B_t = e^{\int_0^t r_s ds}. \tag{1.262}$$

The *instantaneous rate* (or *short rate*) r_t is assumed positive at all times. This is a way to implicitly account for an important restriction: If interest rates were negative, an arbitrage strategy would be to borrow money at negative interest and hold the cash in a safety deposit instead of in an interest-bearing account. Assuming that security costs to store money in a safety deposit are negligible, the existence of such a strategy constrains interest rates to stay positive.

Definition 1.11. Financial Model: Continuous Time. *A continuous-time financial model $\mathcal{M} = (\mathcal{F}_t, A_t^1, \ldots, A_t^n)$ is given by a filtration $\mathcal{F}_t$ and n price processes as basic hedging instruments:*

$$(A_t^1, \ldots, A_t^n), \quad t \in \mathbb{R}_+. \tag{1.263}$$

The value A_0^i can be used to model the current (or spot) price of the ith asset if current time is set as $t = 0$ and the random variable A_t^i models the price of the ith asset at any time $t > 0$.

Definition 1.12. Diffusion Pricing Model. *In a diffusion model the price processes of all hedging instruments (or securities) obey stochastic differential equations of the form*

$$\frac{dA_t^i}{A_t^i} = \mu_t^{A^i} dt + \sum_{\alpha=1}^{M} \sigma_{\alpha,t}^{A^i} \, dW_t^\alpha. \tag{1.264}$$

Here, the dW_t^α, $\alpha = 1, \ldots, M$, are independent Brownian motions (or Wiener processes) with $E[dW_t^\alpha] = 0$ and $E[dW_t^\alpha \; dW_t^\beta] = \delta_{\alpha\beta} \, dt$. The functions $\sigma_{\alpha,t}^{A^i}$ are so-called lognormal volatilities of the ith asset price process $(A_t^i)_{t\geq 0}$ with respect to the αth Brownian motion (i.e., with respect to the αth risk factor), and the functions $\mu_t^{A^i}$ are lognormal drifts of the ith asset price process. These are generally functions of the asset values $A_t^1, \ldots, A_t^n$ and time t.

Note: We can assume further that one of the assets, e.g., A_t^1, is the money-market account, which is the only asset characterized by having zero volatility; in this case $\sigma_{\alpha,t}^{A^1} = 0$ for all $\alpha = 1, \ldots, M$.

Definition 1.13. Adapted Process. *A stochastic process ξ_t is adapted to the filtration $\mathcal{F}_t$ if ξ_t is a random variable in the probability space generated by $\mathcal{F}_t$. In other words, the value of ξ_t depends only on the values taken by the paths $(A_s^1, \ldots, A_s^n)$ for $0 \leq s \leq t$, as they were realized up to time t, i.e., ξ_t is $\mathcal{F}_t$-measurable.*

[15] Technically, B_t is of zero quadratic variation because the differential contains no term with dW_t; however, r_t can generally be stochastic.

Definition 1.14. Stopping Time. *A stopping time* $\tau \in (0, T]$, *for any finite time T, is an* $\mathcal{F}_t$*-measurable positive random variable such that the time event* $\{t = \tau\}$, *with probability* $P(\tau < \infty) = 1$, *corresponds to a decision to stop and is determined entirely by the information set* $\mathcal{F}_t$ *up to time* $t = \tau$. *That is, given the filtration* $\mathcal{F}_t$ *we know whether or not* $\tau \leq t$.

Note that for asset-pricing purposes the information set $\mathcal{F}_t$ basically derives from the set of all asset price paths $(A_t^1, \ldots, A_t^n)$, $0 \leq t \leq \tau$. This rather technical definition and abstract concept of a stopping time is best illustrated with examples. For instance, let x_t be some real-valued diffusion process (e.g., a Wiener process) and let $[a, b] \subset \mathbb{R}$ be a given fixed finite interval. Assume initially $x_0 \notin [a, b]$ at time $t = 0$ and allow the process to evolve in time $t > 0$ up to time T. The random variable defined by

$$\tau = \begin{cases} \min\{t;\ \text{such that } x_t \in [a, b]\}, & \text{if } 0 < t < T \\ T, & \text{otherwise} \end{cases} \tag{1.265}$$

is then a stopping time and corresponds to the first entry time $t < T$ of the process x_t into the interval *[a,b]*. Some basic useful properties of stopping times follow readily, such as additivity: If τ_1 and τ_2 are two stopping times in a given time interval, then $\tau = \tau_1 + \tau_2$ is also a stopping time and, moreover, $\min(\tau_1, \tau_2)$ and $\max(\tau_1, \tau_2)$ are also stopping times. In the pricing of European-style options the expiration time is an example of a stopping time that is actually known at contract inception. In contrast, for American-style options the expiration period (or lifetime of the contract) is still finite, yet there is the added freedom of early exercise. As we shall see in Section 1.14, the early-exercise time is actually an example of an optimal stopping time that is (dynamically) determined by the level of the asset or stock price at the time of early exercise. Other examples of stopping times and derivative instruments are given by *barrier contracts*, for which the pay-off depends on whether or not a certain price process crosses a given barrier in the future. Suppose H is a fixed number, and define τ as the time $t = \tau$ at which $A_t = H$ for the first time, subject to the initial condition A_0. Then τ is a stopping time. Cash flows for barrier options can occur at the time the barrier is crossed or at maturity. A counterexample to a stopping time is the time τ', defined as the last time before a given maturity date T for which $A_{\tau'} = H$. τ' is not a stopping time because knowledge about when τ' occurs requires information on the full path x_t for all $t \in [0, T]$ and in particular for times after τ' itself.

Definition 1.15. Derivative instrument.[16] *A derivative instrument, or contingent claim, is a contractual agreement between two parties who agree to exchange a cash flow stream in the future, where the cash flow amounts are adapted processes and the timings are stopping times in the given financial model. A discrete cash flow stream is modeled by a sequence of pairs* (τ_j, c_j), $j = 1, \ldots, m$, *where the* τ_j *are stopping times and the* c_j *are cash flow amounts depending on the price processes* $(A_t^1, \ldots, A_t^n)$ *up to time* τ_j. *Continuous cash flow streams are modeled by more general adapted processes* γ_t *such that* $d\gamma_t$ *is the cash flow occurring in the time interval* $[t, t + dt)$. *In the particular case of a discrete cash flow stream* (τ_j, c_j), $\tau_j = \tau_1, \ldots, \tau_m < t$, *the continuous-time representation* c_t *is given by*

$$\int_0^t d\gamma_t = \sum_{j=1}^{m} c_j. \tag{1.266}$$

[16] It should be clearly understood that we are throughout assuming all claims or assets are nondefaultable; e.g., the money-market account is assumed nondefaultable. The definition must be modified in the case of defaultable (credit) derivatives, where pricing depends on time of default and recovery, quantities not directly observable from market-traded instruments.

An example of a continuous cash flow stream is given by exchange-traded futures and options contracts. These contracts have the same final pay-off as forward and ordinary option contracts. However, to reduce credit risk to a minimum, exchanges ask investors to hold a margin account and mark-to-market gains and losses on a daily basis based on realized prices or to unwind the position. This results in a daily stream of cash flows that can be modeled as continuous.

Definition 1.16. Self-Financing Trading Strategy. *A self-financing trading strategy in the hedging instruments $A_t^1, \ldots, A_t^n$ is a zero cash flow–replicating strategy for all time $t \in [0, T]$. That is, this strategy consists of a portfolio of positions ξ_t^i in the assets A_t^i, with value $V_t = \sum_{i=1}^n \xi_t^i A_t^i$, where the ξ_t^i, $i = 1, \ldots, n$, are adapted processes such that at all times $t \in [0, T]$ we have*

$$\sum_{i=1}^{n} (A_t^i + dA_t^i) d\xi_t^i = 0. \tag{1.267}$$

The meaning of the self-financing condition is that the cash flow $d\gamma_t$ resulting at time $t + dt$ are reinvested in the underlying assets by adjusting the positions ξ_{t+dt}^i by purchasing or selling the corresponding hedging instruments at the prices $A_t^i + dA_t^i$ at an infinitesimally later time $t + dt$ (i.e., positions are readjusted only after the prices have changed during time dt). In this sense the positions are adapted, i.e., nonanticipative with respect to the stochastic changes in the asset prices. The infinitesimal change in the portfolio value V_t of a self-financing strategy is only due to changes in the prices of the underlying instruments since there are no allowed additional cash inflows or outflows after initial time; hence,[17]

$$dV_t = \sum_{i=1}^{n} \xi_t^i dA_t^i. \tag{1.268}$$

In integral form this is written as

$$V_t = V_0 + \sum_{i=1}^{n} \int_0^t \xi_s^i dA_s^i. \tag{1.269}$$

Using Itô's lemma, the change in portfolio value, $dV_t = V_{t+dt} - V_t$, must also satisfy

$$dV_t = \sum_{i=1}^{n} \left[\xi_t^i dA_t^i + A_t^i d\xi_t^i + (d\xi_t^i)(dA_t^i) \right]. \tag{1.270}$$

Equating these two expressions then gives the self-financing condition rewritten in the form contained in equation (1.267).

Definition 1.17. Self-Financing Replicating Strategy. *A self-financing replicating strategy (or perfect hedge) in the hedging instruments $A_t^1, \ldots, A_t^n$ that replicates a given cash flow stream $d\gamma_t$, where γ_t is a given contingent claim at time t in some time interval $t \in [0, T]$, is defined as a family of adapted processes ξ_t^i, $i = 1, \ldots, n$, such that at all times $t \in [0, T]$ we have*

$$\gamma_t = \gamma_0 + \sum_{i=1}^{n} \int_0^t \xi_s^i \, dA_s^i, \tag{1.271}$$

[17] Note: We assume throughout that the assets do not pay dividends, although in the case of dividends the appropriate formulas extend in a simple manner.

or, equivalently in differential form,

$$d\gamma_t = \sum_{i=1}^{n} \xi_t^i \, dA_t^i. \tag{1.272}$$

In the case of a European-style option with payoff $\phi(S_T)$ at time T, where S_t is the underlying stock price process, a self-financing replication strategy in the stock and the money-market account, with value $\xi_t^1 B_t + \xi_t^2 S_t$ at time t, would satisfy

$$B_t \, d\xi_t^1 + (S_t + dS_t) d\xi_t^2 = 0 \tag{1.273}$$

for all times $t \in [0, T)$. [Note that the term $dB_t = r_t B_t \, dt$ vanishes since it gives rise to a term of $O((dt)d\xi_t^1)$, i.e., of order greater than dt.] At time T, the position is unwound so that the payout $\phi(S_T)$ [i.e., $\gamma_T = \phi(S_T)$ in this case] is generated; i.e., the portfolio has terminal value

$$\xi_T^1 B_T + \xi_T^2 S_T = \phi(S_T). \tag{1.274}$$

In the case of a barrier or American option, where the payout occurs at a stopping time $0 \le \tau \le T$, the equation (1.273) is valid until time τ, at which point we have

$$B_\tau \xi_\tau^1 + S_\tau \xi_\tau^2 = \phi(S_\tau). \tag{1.275}$$

One of the main problems in pricing theory is whether or not the cash flow streams associated with a contingent claim can be replicated by means of a self-financing trading strategy. If a self-financing trading strategy exists and reproduces all the cash flows of a given contingent claim, then the present value of the cash flow stream can (uniquely in case of no arbitrage) be identified as the cost of setting up the self-financing trading strategy. The question of whether such a self-financing strategy exists relates to attainability and market completeness.

The practical implementation of trading strategies is limited by the existence of transaction costs, by liquidity effects, which pose restrictions on the amounts of a given instrument that can be traded at the posted price, and by the delays with which information reaches market participants. To a first approximation, these effects can be taken into account implicitly by assuming that there are no imperfections. A key role is played by the condition of *absence of arbitrage*, which is stated next and which implies that all portfolios with the same payoff structure have the same price. Asking for absence of arbitrage is a way of accounting for finite market liquidity since, in fact, if an asset had two different prices, trades to exploit the opportunity would cause the prices to realign.

Definition 1.18. Arbitrage: Continuous Time. *The self-financing trading strategy* $(\xi_t^1, \ldots, \xi_t^n)$, $0 \le t \le T$, *in the hedging assets* $(A_t^1, \ldots, A_t^n)$ *is an arbitrage strategy if either of the following two conditions holds.*
A1. *The portfolio value process*

$$V_t = \sum_{i=1}^{n} \xi_t^i A_t^i \tag{1.276}$$

is such that $V_0 < 0$ *and with probability* $P(V_T \ge 0) = 1$.
A2. *The value process* V_t *is such that* $V_0 = 0$ *and* $P(V_T > 0) > 0$ *with* $P(V_t \ge 0) = 1$ *for all* $t \in [0, T]$.

In plain language, condition A2 says that an arbitrage opportunity is a self-financed strategy that can generate a profit at zero cost and with no possibility of a loss at any time during the strategy.

Typically, when solving the replication problem for a cash flow stream, the current price of the stream is not known, a priori. Knowledge of the cash flow stream, however, is sufficient because if a trading strategy replicates the cash flows, in virtue of the hypothesis of absence of arbitrage, the value of this strategy at all times yields the price or value process V_t. Next we consider a couple of examples of replication (or hedging) strategies. One is static in time; the other is dynamic.

Example 1. Perpetual Double Barrier Option.

Suppose there are no carry costs such as interest rates or dividends for holding a position in the stock. Consider a perpetual option with two barriers: a lower barrier at stock value L and an upper barrier at H, with $L < H$. If the stock price touches the lower barrier before it touches the upper barrier, the holder receives R_L dollars and the contract terminates. Otherwise, whenever the upper barrier is hit first, the holder receives R_H dollars and the contract terminates. The problem is to find the price and a hedging strategy for this contract.

To solve this problem, let τ_L be the stopping time for hitting the lower barrier and τ_H be the stopping time for hitting the higher barrier. The stopping time τ at which the option expires is the minimum of these times,

$$\tau = \min(\tau_L, \tau_H). \tag{1.277}$$

If one considers a replicating portfolio $f_t = aS_t + b$ at any time t, then the barrier levels give rise to two equations:

$$aH + b = R_H, \qquad aL + b = R_L, \tag{1.278}$$

corresponding to the portfolio value (i.e., payout) for hitting either barrier. The value f_τ of the perpetual double barrier contract evaluated at the stopping time $t = \tau$ is

$$f_\tau \ = \ aS_\tau + b. \tag{1.279}$$

Solving the system in equation (1.278) for the portfolio weights a and b, we find that

$$a = \frac{R_H - R_L}{H - L}, \qquad b = R_H - aH. \tag{1.280}$$

Absence of arbitrage therefore implies that the price process followed by f_t is given by the value of the portfolio $aS_t + b$ that replicates the cash flows.

Example 2. Dynamic Hedging in the Black–Scholes Model

Consider the Black–Scholes model with a stock price following geometric Brownian motion,

$$\frac{dS_t}{S_t} = \mu \ dt + \sigma \ dW_t. \tag{1.281}$$

In this model, the price at time t of a call struck at K and maturity at calendar time $T > t$ is given by the function $C_{BS}(S_t, K, T - t, \sigma, r)$ in equation (1.217). Let's assume that in this economy interest rates are constant and equal to r.

One can show that the pay-off of the call can be replicated by means of a self-financing trading strategy that costs $C_t = C_{BS}(S_t, K, T - t, \sigma, r)$ to set up at calendar time t. This strategy involves two adapted processes a_t and b_t for the hedge ratios that give the positions at calendar time t in two assets: the stock of price S_t and a zero-coupon bond maturing at time T of price $Z_t(T) = e^{-r(T-t)}$. Namely,

$$C_t = a_t S_t + b_t Z_t(T). \tag{1.282}$$

To show this, we need to find the two processes a_t and b_t. Let us note that self-financing condition (1.267) in this case reads

$$(S_t + dS_t)da_t + Z_t(T)db_t = 0. \tag{1.283}$$

By the differential of equation (1.282) and using the self-financing condition we find

$$dC_t = a_t \, dS_t + rb_t Z_t(T)dt. \tag{1.284}$$

On the other hand, applying Itô's lemma (in one dimension) to the price process C_t (considered as function of t and S_t) we find

$$dC_t = \left(\frac{\partial C_{BS}}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 C_{BS}}{\partial S^2}\right)dt + \frac{\partial C_{BS}}{\partial S} dS_t,$$

where $S = S_t$. By equating coefficients in dt and dS_t with the previous equation we find

$$a_t = \frac{\partial C_{BS}}{\partial S} \tag{1.285}$$

and

$$rb_t Z_t(T) = \frac{\partial C_{BS}}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 C_{BS}}{\partial S^2}. \tag{1.286}$$

Solving for b_t from replication equation (1.282) gives

$$b_t = Z_t(T)^{-1}(C_t - a_t S_t). \tag{1.287}$$

Substituting b_t as given by equation (1.287), as well as a_t from equation (1.285) into equation (1.286), we arrive at the Black–Scholes partial differential equation in current time t and spot price $S = S_t$:

$$\frac{\partial C_{BS}}{\partial t} + rS\frac{\partial C_{BS}}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 C_{BS}}{\partial S^2} - rC_{BS} = 0 \tag{1.288}$$

This is precisely the equation satisfied by the function $C_{BS}(S_t, K, T-t, \sigma, r)$ given by equation (1.217) with $T \to T-t$.

Notice that the parameter μ in the equation for the stock price process (1.281) appears in neither the Black–Scholes formula, the Black–Scholes equation, nor the hedge ratios a_t and b_t. Section 1.10 provides a more general explanation of this very notable simplification.

1.10 Dynamic Hedging and Derivative Asset Pricing in Continuous Time

In this section, we present the main theorem for pricing derivative assets within the continuous-time framework.

Theorem 1.4. Fundamental Theorem of Asset Pricing (Continuous-Time Case). Part I. *Consider a diffusion continuous-time financial model $\mathcal{M} = (\mathcal{F}_t, A_t^1, \ldots, A_t^n)$, where the hedging instruments are assumed to satisfy a diffusion equation of the form (1.264), i.e.,*

$$\frac{dA_t^i}{A_t^i} = \mu_t^{A^i} dt + \sum_{\alpha=1}^{M} \sigma_{\alpha,t}^{A^i} \, dW_t^\alpha, \qquad i = 1, \ldots, n, \tag{1.289}$$

where dW_t^α are understood to be standard Brownian increments with respect to a specified probability measure. Also, suppose there exists a money-market account B_t with

$$dB_t = r_t B_t \, dt. \tag{1.290}$$

Finally, suppose there are no arbitrage opportunities. Then:

(i) Under all equivalent probability measures, there exists a family of adapted processes $q_{\alpha,t}$, $\alpha = 1, \ldots, M$ (one for each risk factor), such that, for any asset price process A_t obeying an equation similar to equation (1.289) with drift μ_t^A and volatilities $\sigma_{\alpha,t}^A$, the drift term is linked to the corresponding volatilities by the equation

$$\mu_t^A = r_t + \sum_{\alpha=1}^{M} q_{\alpha,t} \sigma_{\alpha,t}^A, \tag{1.291}$$

where $q_{\alpha,t}$ are independent of the asset A in question.

In finance parlance, the adapted processes $q_{\alpha,t}$ are known as the *price of risk* for the αth risk factor (or αth Brownian motion). Note that this result applies to any asset obeying a diffusion process: In particular, the drifts $\mu_t^{A^i}$ and volatilities $\sigma_{\alpha,t}^{A^i}$ of the base asset prices A_t^i are themselves also linked by an equation similar to equation (1.291), with $q_{\alpha,t}$ independent of the prices A_t^i.

Definition 1.19. Numeraire Asset. *Any asset g_t whose price process is positive, in the sense that $g_t > 0$ for all t, is chosen as the numeraire for pricing. That is, g_t is an asset price relative to which the value of all other assets A_t are expressed using the ratio $\frac{A_t}{g_t}$.*

Theorem 1.5. Fundamental Theorem of Asset Pricing (Continuous-Time Case). Part II. *Under the hypotheses in Part I of the theorem, we have the following: (ii) If g_t is a numeraire asset, then there exists a probability measure Q(g) for which the price A_t at time t of any attainable instrument without cash flows up to a stopping time $\tau > t$ is given by the martingale condition*

$$\frac{A_t}{g_t} = E_t^{Q(g)}\left[\frac{A_\tau}{g_\tau}\right]. \tag{1.292}$$

Under the measure Q(g) the prices of risk in equation (1.291) for the αth factors are given by the volatilities of g_t for the corresponding αth factors:

$$q_{\alpha,t}^g = \sigma_{\alpha,t}^g. \tag{1.293}$$

Note that we are throughout assuming that the contingent claim or derivative instrument to be priced is *attainable*, meaning that one can find a self-financing replicating strategy that exactly replicates the cash flows of the claim. If one also assumes that the financial model satisfies market completeness, then every contingent claim or cash flow stream is assumed attainable.

Definition 1.20. Pricing Measure: Continuous Time. *Given a numeraire asset price process g_t, the pricing measure associated with g is the martingale measure $Q(g)$ for which pricing formula (1.292) holds for any asset price process A_t.*

Definition 1.21. Risk-Neutral Measure. *Assuming continuous compounding, the risk-neutral measure $Q(B)$ is the martingale measure with the money-market account as numeraire asset $g_t = B_t = e^{\int_0^t r_s ds}$.*

Theorem 1.6. Fundamental Theorem of Asset Pricing (Continuous-Time Case). Part III. *Under the hypotheses in Part I of the theorem, we have the following: (iii) Under the risk-neutral measure $Q(B)$ all the components of the price-of-risk vector, $q^g_{\alpha,t}$, $\alpha = 1, \dots, M$, vanish, and the drift μ^A_t of any asset price A_t at time t is equal to the riskless rate r_t. The price process for any attainable instrument without cash flows up to any stopping time $\tau > t$ is given by the expectation at time t:*

$$A_t = E_t^{Q(B)}\left[e^{-\int_t^\tau r_u du} A_\tau\right]. \tag{1.294}$$

(iv) Any attainable price process A_t can be replicated by means of a self-financing trading strategy with portfolio value $V_t = \zeta_t^{(0)} B_t + \sum_{i=1}^n \zeta_t^{(i)} A_t^i$ in the base assets A_t^i and in the money-market account B_t:

$$dA_t = dV_t = \zeta_t^{(0)} r_t B_t \; dt + \sum_{i=1}^n \zeta_t^{(i)} \; dA_t^i, \tag{1.295}$$

where the positions $\zeta_t^{(i)}$ satisfy the self-financing condition

$$B_t \; d\zeta_t^{(0)} + \sum_{i=1}^n (A_t^i + dA_t^i) d\zeta_t^{(i)} = 0. \tag{1.296}$$

Proof.
(i). Assume no arbitrage and consider a self-financing trading strategy, with components $\zeta_t^{(1)}, \dots, \zeta_t^{(n)}$ as adapted positions in the family of base assets $A_t^1, \dots, A_t^n$. Then

$$\sum_{i=1}^n (A_t^i + dA_t^i) d\zeta_t^{(i)} = 0 \tag{1.297}$$

holds. This strategy has portfolio value at time t given by

$$\Pi_t \;=\; \sum_{i=1}^n \zeta_t^{(i)} A_t^i. \tag{1.298}$$

This strategy is instantaneously riskless if the stochastic component is zero, i.e., $d\Pi_t = r_t \Pi_t \; dt$. Given our assumptions, a riskless strategy exists and can be explicitly constructed as follows. Using the self-financing condition in equation (1.297) and Itô's lemma for the stochastic differential $d\Pi_t$ we obtain the infinitesimal change in portfolio value in time $[t, t+dt)$:

$$d\Pi_t = \sum_{i=1}^n \left[(A_t^i + dA_t^i) d\zeta_t^{(i)} + \zeta_t^{(i)} dA_t^i\right] = \sum_{i=1}^n \zeta_t^{(i)} dA_t^i. \tag{1.299}$$

Due to the assumption of no arbitrage, the rate of return on this portfolio over the period $[t, t+dt)$ must equal the riskless rate of return on the money-market account, i.e., $d\Pi_t = r_t \Pi_t \, dt$.[18] Substituting equation (1.289) into the foregoing stochastic differential and setting the coefficients in all the stochastic terms dW_t^α to zero gives

$$\sum_{i=1}^{n} \sigma_{\alpha,t}^{A^i} \zeta_t^{(i)} A_t^i = 0, \tag{1.300}$$

for all $\alpha = 1, \ldots, M$. Here the functions $\sigma_{\alpha,t}^{A^i}$ are volatilities in the αth factor for each asset A^i. This equation states that the $\mathbb{R}^n$-dimensional vector of components $\zeta_t^{(i)} A_t^i$ is orthogonal to the subspace of M vectors (labeled by $\alpha = 1, \ldots, M$) in $\mathbb{R}^n$ having components $\sigma_{\alpha,t}^{A^i}$, $i = 1, \ldots, n$.

Absence of arbitrage also implies that the portfolio earns a risk-free rate, $d\Pi_t = r_t \Pi_t dt$; hence, setting the drift coefficient in the stochastic differential $d\Pi_t$ to $r_t \Pi_t$ while using equation (1.298) gives this additional condition:

$$\sum_{i=1}^{n} (\mu_t^{A^i} - r_t) \zeta_t^{(i)} A_t^i = 0. \tag{1.301}$$

Here, the quantities $\mu_t^{A^i}$ are drifts for each ith asset. Hence equation (1.300) must imply equation (1.301) for all arbitrage-free strategies satisfying the self-financing condition. Equation (1.301) states that the $\mathbb{R}^n$-dimensional vector of components $\zeta_t^{(i)} A_t^i$ must be orthogonal to the $\mathbb{R}^n$-dimensional vector with components $(\mu_t^{A^i} - r_t)$. This means that if the vector with components $\zeta_t^{(i)} A_t^i$ is orthogonal to the M vectors of components $\sigma_{\alpha,t}^{A^i}$, then it is also orthogonal to the vector of components $(\mu_{\alpha,t}^{A^i} - r_t)$. From linear algebra we know that this is possible if and only if the vector of components $(\mu_t^{A^i} - r_t)$ is a linear combination of the M vectors of components $\sigma_{\alpha,t}^{A^i}$ (i.e., is contained in the linear subspace spanned by the M vectors). Hence for any given time t, we have

$$\mu_t^{A^i} = r_t + \sum_{\alpha=1}^{M} q_{\alpha,t} \sigma_{\alpha,t}^{A^i}, \tag{1.302}$$

with coefficients $q_{\alpha,t}$ *independent of the asset* A^i, for all $i = 1, \ldots, n$. Since this is true for all self-financing strategies and choices of base assets, this implies that the same relation must follow for any asset A_t; namely, equation (1.291) obtains.

(ii) Let g be a numeraire asset. The measure *Q(g)* is specified by the condition in equation (1.292). At this point we make use of a previously derived result contained in equation (1.138). Applying that formula now to the quotient A_t/g_t, where A_t satisfies an equation of the form (1.264) (with A^i replaced by A) and the numeraire asset g_t satisfies a similar equation,

$$\frac{dg_t}{g_t} = \mu_t^g \, dt + \sum_{\alpha=1}^{M} \sigma_{\alpha,t}^g \, dW_t^\alpha, \tag{1.303}$$

[18] A simple argument shows that if the portfolio return is greater than r_t, then an arbitrage strategy exists by borrowing money at the lower rate r_t at time t and investing in the portfolio until time $t+dt$. On the other hand, if the portfolio return is less than r_t, then an arbitrage strategy also exists by short-selling the portfolio at time t and investing the earnings in the money-market account. Both strategies yield a zero-cost profit.

immediately gives the drift component:

$$E_t\left[d\frac{A_t}{g_t}\right] = \frac{A_t}{g_t}\left[\mu_t^A - \mu_t^g - \sum_{\alpha=1}^{M} \sigma_{\alpha,t}^g(\sigma_{\alpha,t}^A - \sigma_{\alpha,t}^g)\right]dt \tag{1.304}$$

$$= \frac{A_t}{g_t}\sum_{\alpha=1}^{M}(q_{\alpha,t} - \sigma_{\alpha,t}^g)(\sigma_{\alpha,t}^A - \sigma_{\alpha,t}^g)dt. \tag{1.305}$$

In the last equation we have used equation (1.291) for both g_t and A_t. In order for the ratio A_t/g_t to be a martingale process for all (arbitrary) choices of the asset A_t, this expectation must be zero. This is the case if and only if the process for the price of risk q_t^α is related to the numeraire asset g_t, $q_{\alpha,t} = q_{\alpha,t}^g$, as follows:

$$q_{\alpha,t}^g = \sigma_{\alpha,t}^g, \qquad \alpha = 1, \ldots, M. \tag{1.306}$$

That is, the prices of risk q_t^α are equal to the volatilities of the numeraire asset for each respective risk factor.

(iii) This is a particular case of (ii) and follows when money-market account B_t is chosen as numeraire asset. Since $dB_t = r_t B_t\, dt$, the prices of risk in this case are all zero, i.e., $q_{\alpha,t}^B = 0$, and therefore $\mu_t^A = r_t$ for all asset price processes A_t. In particular, we have that

$$A_t = E_t^{Q(B)}\left[A_\tau \frac{B_t}{B_\tau}\right] = E_t^{Q(B)}\left[A_\tau e^{-\int_t^\tau r_s ds}\right], \tag{1.307}$$

giving the result. Here we have used the fact that B_t at time t is a known (i.e., nonstochastic) quantity that can be taken inside the expectation.

(iv). Consider the trading strategy with positions $\zeta_t^{(i)}$ in the base assets A_t^i. A long position in this trading strategy and a short position in the generic asset A_t is a riskless combination that accrues at the risk-free rate. By adjusting the position in the money-market account $\zeta_0^{(0)}$ so that the trading strategy has the same value of asset A_0 at initial time $t = 0$, the resulting trading strategy will track the price process A_t for all times. This trading strategy is also self-financing. In fact

$$dA_t = d\left(\zeta_t^{(0)} B_t + \sum_{i=1}^{n} \zeta_t^{(i)} A_t^i\right)$$

$$= \zeta_t^{(0)} r_t B_t\, dt + B_t\, d\zeta_t^{(0)} + \sum_{i=1}^{n}[(A_t^i + dA_t^i)d\zeta_t^{(i)} + \zeta_t^{(i)} dA_t^i]. \tag{1.308}$$

Hence equation (1.295) obtains from equation (1.296). □

In summary, we observe that the asset pricing theorem is connected to the evaluation of conditional expectations of martingales (i.e., relative asset price processes) within a filtered probability space and under a choice of an equivalent probability measure (also called an *equivalent martingale measure*). A measure is specified by the chosen numeraire asset g obeying a stochastic price process of its own, given by equation (1.303). Given a numeraire g, the relative asset price process A_t/g_t, for a generic asset price A_t, is a martingale under the corresponding measure $Q(g)$. Equivalent martingale measures then arise by considering different choices of numeraire assets. In particular, consider another numeraire asset, denoted by $\tilde{g}$, with price process $\tilde{g}_t$, and suppose that measure $Q(\tilde{g})$ is equivalent to $Q(g)$, then prices computed under any two equivalent measures must be equal:

$$A_t = g_t E_t^{Q(g)}\left[\frac{A_T}{g_T}\right] = \tilde{g}_t E_t^{Q(\tilde{g})}\left[\frac{A_T}{\tilde{g}_T}\right]. \tag{1.309}$$

Rearranging terms gives

$$E_t^{Q(g)}\left[\frac{A_T}{g_T}\right] = \frac{\tilde{g}_t}{g_t} E_t^{Q(\tilde{g})}\left[\frac{A_T}{\tilde{g}_T}\right]$$
$$= E_t^{Q(\tilde{g})}\left[\frac{g_T/\tilde{g}_T}{g_t/\tilde{g}_t}\frac{A_T}{g_T}\right]. \tag{1.310}$$

Note that this holds true for an arbitrary random variable $X_T = A_T/g_T$. We hence obtain the general property under two equivalent measures:

$$E_t^{Q(g)}\big[X_T\big] = \rho_t^{-1} E_t^{Q(\tilde{g})}\big[X_T \rho_T\big], \tag{1.311}$$

where $\rho_t = g_t/\tilde{g}_t \equiv \left(\frac{dQ(g)}{dQ(\tilde{g})}\right)_t$, $t \in [0, T]$, is a Radon–Nikodym derivative of *Q(g)* with respect to $Q(\tilde{g})$ (with both measures being restricted to the filtration $\mathcal{F}_t$). For $t = T$ we write $\left(\frac{dQ(g)}{dQ(\tilde{g})}\right)_T = \frac{dQ(g)}{dQ(\tilde{g})}$. Choosing $X_t = 1$ in the foregoing equation shows that ρ_t is also a martingale with respect to $Q(\tilde{g})$.

Let's now fix our choice for one of the numeraires; i.e., let $\tilde{g}_t = B_t$ be the value process of the money-market account so that $Q(\tilde{g}) = Q(B)$ is the risk-neutral measure. Taking the stochastic differential of the quotient process $\rho_t = g_t/B_t$ gives

$$\frac{d\rho_t}{\rho_t} = (\mu_t^g - r_t)dt + \sum_{\alpha=1}^{M} \sigma_{\alpha,t}^g \, dW_t^\alpha. \tag{1.312}$$

Under the risk-neutral measure with dW_t^α as Brownian increments under *Q(B)*, this process must be driftless so that we have $\mu_t^g = r_t$. In particular, this martingale takes the form of an *exponential martingale*,

$$\rho_t = \frac{g_t}{B_t} = \exp\left(-\frac{1}{2}\int_0^t ||\boldsymbol{\sigma}_s^g||^2 ds + \int_0^t \boldsymbol{\sigma}_s^g \cdot d\mathbf{W}_s\right), \tag{1.313}$$

where $||\boldsymbol{\sigma}_s^g||^2 = \boldsymbol{\sigma}_s^g \cdot \boldsymbol{\sigma}_s^g = \sum_{\alpha=1}^M (\sigma_{\alpha,s}^g)^2$ and $\boldsymbol{\sigma}_s^g \cdot d\mathbf{W}_s = \sum_{\alpha=1}^M \sigma_{\alpha,s}^g \, dW_s^\alpha$. At this point we can implement the Girsanov theorem for exponential martingales, which tells us that the $\mathbb{R}^M$-valued vector increment defined by

$$d\mathbf{W}_t^g = -\boldsymbol{\sigma}_t^g \, dt + d\mathbf{W}_t \tag{1.314}$$

is a standard Brownian vector increment under the measure *Q(g)*. In the risk-neutral measure the base assets must all drift at the same risk-free rate,

$$\frac{dA_t^i}{A_t^i} = r_t \, dt + \sum_{\alpha=1}^{M} \sigma_{\alpha,t}^{A^i} \, dW_t^\alpha, \qquad i = 1, \ldots n. \tag{1.315}$$

Substituting for $d\mathbf{W}_t$ using equation (1.314) into this equation and compacting to vector notation gives

$$\frac{dA_t^i}{A_t^i} = (r_t + \boldsymbol{\sigma}_t^g \cdot \boldsymbol{\sigma}_t^{A^i})dt + \boldsymbol{\sigma}_t^{A^i} \cdot d\mathbf{W}_t^g \,, \qquad i = 1, \ldots, n. \tag{1.316}$$

This last equation is therefore entirely consistent with the formulation presented earlier in terms of the prices of risk. In particular, equation (1.316) is precisely equation (1.289), wherein the Brownian increments are understood to be w.r.t. $Q(g)$, with g_t as an arbitrary choice of numeraire asset-price process. From equation (1.316) we again see that the vector of the prices of risk is $\mathbf{q}_t = \boldsymbol{\sigma}_t^g$. In financial terms, each component of $\mathbf{q}_t$ essentially represents the excess return on the risk-free rate (per unit of risk or volatility for the component risk factor) required by investors in a fair market.

Example 1. Perpetual Double Barrier Option — Risk-Neutral Measures.

Reconsider the case of the perpetual double barrier option with zero interest rates discussed previously. The pricing formula for f_t is independent of the real-world stock price drift, although this drift does in fact affect the real-world probability of hitting one barrier before the other. Since interest rates vanish, no discounting is required, and the price process f_t has the following representation under the risk-neutral measure $Q = Q(B)$:

$$f_t = E_t^Q[f_\tau]. \tag{1.317}$$

In this case, the price process f_t is a martingale under the risk-neutral measure because interest rates are zero for all time and the value of the money-market account is constant, i.e., unity. Hence the martingale property gives

$$f_t = R_H \text{ Prob}^Q[S_\tau = H|S_t] + R_L \text{ Prob}^Q[S_\tau = L|S_t], \tag{1.318}$$

where the probabilities are conditional on the current stock price's value S_t. These probabilities of hitting either barrier must also sum to unity,

$$\text{Prob}^Q[S_\tau = H|S_t] + \text{Prob}^Q[S_\tau = L|S_t] = 1. \tag{1.319}$$

Note that $f_t = aS_t + b$, where a and b are given by equations (1.280). Hence, the probability of hitting either barrier under the risk-neutral measure can be found by solving equations (1.318) and (1.319). Notice that these probabilities do not depend on the drift of the stock price under the real-world measure.

Problems

Problem 1. Find explicit expressions for the preceding risk-neutral probabilities $P_L = \text{Prob}^Q[S_\tau = L|S_t]$ and $P_H = \text{Prob}^Q[S_\tau = H|S_t]$. Find the limiting expressions for the case that $H >> L$ (i.e., $H \to \infty$ for fixed L). What is the price of the perpetual double barrier for this case?

1.11 Hedging with Forwards and Futures

Let A_t be an asset price process for the asset A. A *forward contract*, with value V_t at time t, on the underlying asset A (e.g., a stock) is a contingent claim with maturity T and pay-off at time T equal to

$$V_T = A_T - F, \tag{1.320}$$

where F is a fixed amount. According to the fundamental theorem of asset pricing (FTAP), the price of this contract at time $t < T$ prior to maturity is equal to $A_t - FZ_t(T)$, where $Z_t(T)$ is the value at calendar time t of a zero-coupon (discount) bond maturing at time T. This can be seen in several ways. The first is the following. The payout A_T can be replicated by holding a position in the asset A at all times, while the cash payment F at time T is equivalent to holding a zero-coupon bond of nominal F and maturing at time T. Alternatively, to assess the current price V_t of the forward contract using FTAP of Section 1.10, we can evaluate the following expectation at time t of the pay-off under the forward measure with $g_t = Z_t(T)$ as numeraire, giving

$$V_t = Z_t(T) E_t^{Q(Z(T))}[A_T - F] = A_t - F Z_t(T). \tag{1.321}$$

Here we used the facts that at maturity $Z_T(T) = 1$ and that $E_t^{Q(Z(T))}[A_T] = A_t / Z_t(T)$, $E_t^{Q(Z(T))}[F] = F$. The *equilibrium forward price* (at time t), denoted by $F_t(A, T)$, is the so-called *forward price* such that the value V_t of the forward contract at time t is zero. Setting $V_t = 0$ in equation (1.321), we find

$$F_t(A, T) = \frac{A_t}{Z_t(T)}. \tag{1.322}$$

Let's assume stochastic interest rates, i.e., a diffusion process for the zero-coupon bond [satisfying equation (1.349) of Problem 1], as well as diffusion processes for the asset A_t [satisfying equation (1.348) of Problem 1] and the equilibrium forward price satisfying

$$\frac{dF_t(A, T)}{F_t(A, T)} = \mu_t^{F(A,T)} dt + \sigma_t^{F(A,T)} dW_t. \tag{1.323}$$

Then a relatively straightforward calculation using Itô's lemma yields the following form for the lognormal volatility of the forward price (see Problem 1 of this section):

$$\sigma_t^{F(A,T)} = \sigma_t^A - \sigma_t^{Z(T)}, \tag{1.324}$$

and its drift

$$\mu_t^{F(A,T)} = \mu_t^A - \mu_t^{Z(T)} - \sigma_t^{Z(T)}(\sigma_t^A - \sigma_t^{Z(T)}), \tag{1.325}$$

where $\sigma_t^{Z(T)}$ is the lognormal volatility of the zero-coupon bond price and σ_t^A that of the asset. We note that the foregoing drift and volatility functions are generally functions of the underlying asset price A_t, calendar time t, and maturity T. Moreover, these relationships hold for any choice of numeraire asset g_t. As part of Problem 1 of this section, the reader is also asked to derive more explicit expressions for the drifts and volatilities of the forward price under various choices of numeraire.

Definition 1.22. Futures Contract. *Futures contracts are characterized by an underlying asset of price process A_t and a maturity T. Let us partition the lifetime interval [0,T] in N subintervals of length $\delta t = \frac{T}{N}$. Let $t_i = i \cdot \delta t$ be the endpoints of the intervals. The futures contract with reset period δt is characterized by a futures price $F_{t_i}^*(A, T)$ for all $i = 0, \ldots, N$, and at all times t_i the following cash flow occurs at time t_{i+1}:*

$$c_{t_{i+1}} = F_{t_{i+1}}^*(A, T) - F_{t_i}^*(A, T). \tag{1.326}$$

Furthermore, the futures price at time $t_N = T$ equals the asset price $F_T^*(A, T) = A_T$, while at previous times the futures price is set in such a way that the present value of the futures contract is zero.

Recall that under the risk-neutral measure *Q(B)*, the price of risk is zero (i.e., the numeraire g_t is the money-market account B_t with zero volatility with respect to all risk factors — $\sigma^g_{\alpha,t} = 0$). Hence, according to equation (1.291) of the asset pricing theorem, all asset prices A_t drift at the riskless rate $\mu^A_t = r_t$ under *Q(B)*:

$$\frac{dA_t}{A_t} = r_t\ dt + \sum_{\alpha=1}^{M} \sigma^A_{\alpha,t}\ dW^\alpha_t, \tag{1.327}$$

where we have assumed *M* risk factors or, in the case of one risk factor, we simply have

$$\frac{dA_t}{A_t} = r_t\ dt + \sigma^A_t\ dW_t. \tag{1.328}$$

Proposition. *In the limit as $\delta t \to 0$, futures prices behave as (zero-drift) martingales under the risk-neutral measure.*

Proof. By definition, the futures price is such that the present value of a futures contract is zero at all reset times *t*, and the cash flows at the subsequent times $t + \delta t$ are given by the random variable $\delta F^*_t(A, T) = F^*_{t+\delta t}(A, T) - F^*_t(A, T)$, so the following condition holds under the risk-neutral measure:

$$E^{Q(B)}_t \left[\frac{\delta F^*_t(A, T)}{B_{t+\delta t}} \right] = 0, \tag{1.329}$$

where we discount at times $t + \delta t$. Taking the limit $\delta t \to 0$, gives $B^{-1}_t E^{Q(B)}_t [dF^*_t(A, T)] = 0$. Since $B_t \neq 0$, the stochastic differential $dF^*_t(A, T)$ has zero-drift terms for all *t*; i.e., $F^*_t(A, T)$ is a martingale under the measure *Q(B)*, with $E^{Q(B)}_t [dF^*_t(A, T)] = 0$. □

The price spread between futures and forwards is given by

$$F^*_t(A, T) - F_t(A, T) = E^{Q(B)}_t [A_T] - \frac{A_t}{Z_t(T)}, \tag{1.330}$$

with $F^*_T(A, T) = F_T(A, T)$ (i.e., at maturity the two prices are the same). In Chapter 2 we shall derive a formula for this spread based on a simple diffusion model for the asset and discount bond. The topic of stochastic interest rates and bond pricing will be covered in Chapter 2. However, we note here that when interest rates are deterministic (nonstochastic), where r_t is a known ordinary function of *t*, then the discount bond price is simply given by a time integral: $Z_t(T) = \exp(-\int_t^T r_s\ ds) = B_t/B_T$. When interest rates are stochastic (i.e., nondeterministic), as is more generally the case, then we can use equation (1.294) of the asset-pricing theorem, for the case $Z_t(T)$ as asset, to express the discount bond price as an expectation of the payoff $Z_T(T) = 1$ (i.e., the payout of exactly one dollar for certain at maturity) under the measure with the money-market account as numeraire:

$$Z_t(T) = E^{Q(B)}_t [e^{-\int_t^T r_s ds} Z_T(T)] = E^{Q(B)}_t [e^{-\int_t^T r_s ds}]. \tag{1.331}$$

[This expectation is not a simple integral (as arises in the pricing of European options) and can in fact generally be expressed as a multidimensional path integral. See, for example, the project on interest rate trees in Part II.] In the case that the interest rate process is a deterministic function of time or, more generally, when the underlying asset price process A_t is statistically independent of the interest rate process (where both processes may be nondeterminsitic), then forward and future prices coincide and the spread vanishes. In fact, in this case

$$E_t^{Q(B)}[A_T] = \frac{E_t^{Q(B)}[e^{-\int_t^T r_s ds}]}{Z_t(T)} E_t^{Q(B)}[A_T] = \frac{E_t^{Q(B)}[e^{-\int_t^T r_s ds} A_T]}{Z_t(T)} = \frac{A_t}{Z_t(T)}, \tag{1.332}$$

where we have used equations (1.331) and (1.294).

Definition 1.23. European-Style Futures Options. *European-style futures options are contracts with a payoff function $\phi(A_T)$ at maturity T. They are similar to the regular earlier European-style option contracts, except those are written on the underlying and traded over the counter with upfront payment, while futures options are traded using a margin account mechanism similar to that of futures contracts. Namely, futures options are traded in terms of a futures option price A_t^* that equals $\phi(A_T)$ at maturity $t = T$, while the associated cash flow stream to the holder's margin account is given by*

$$c_t = A_t^* - A_{t-dt}^*. \tag{1.333}$$

Notice that, similar to an ordinary futures contract, futures option prices A_t^* follow martingale processes under the risk-neutral measure.

Example 1. European Futures Options.

The futures option price V_t^* for a European-style option with payoff function $\phi(A_T)$ is thus given by the martingale condition

$$V_t^* = E_t^{Q(B)}\left[\phi(A_T)\right]. \tag{1.334}$$

The analogue of the Black–Scholes (i.e., lognormal) model can be written as follows under the risk-neutral measure

$$dA_t^* = \sigma A_t^* \, dW_t, \tag{1.335}$$

where the drift is zero because of the martingale property. We remark here that, in case interest rates are stochastic, the implied Black–Scholes volatility on the futures option does not necessarily coincide with the implied Black–Scholes volatility for plain vanilla equivalents.

Let $A_t^* = F_t^*(A, T)$ be the futures price on the asset. At maturity we have $A_T = F_T^*(A, T) = F_T(A, T)$. The pricing formula for a futures call option struck at the futures price K is given by

$$C_t^*(K, T) = E_t^{Q(B)}[(A_T - K)_+] = F_t^*(A, T)N(d_+) - KN(d_-), \tag{1.336}$$

where

$$d_\pm = \frac{\log(F_t^*(A, T)/K) \pm (\sigma^2/2)(T - t)}{\sigma\sqrt{T - t}} \tag{1.337}$$

and we have used the standard expectation formula in equation (1.169) for the case of zero drift, and where the underlying variable S_t in that formula is now replaced by $F_t^*(A, T)$. Notice that this formula carries no explicit dependence on interest rates.

Example 2. Variance Swaps.

An example of a dynamic trading strategy involving futures contracts and the static hedging strategies discussed in Section 1.8 is provided by variance swaps. Variance swaps are defined as contracts yielding the pay-off at maturity time T:

$$\mathcal{N}\left[\frac{1}{T}\int_0^T \sigma_t^2\ dt - \Sigma^2\right], \tag{1.338}$$

where $\mathcal{N}$ is a fixed notional amount in dollars per annualized variance. Assuming that technical upfront fees are negligible, variance swaps are priced by specifying the variance Σ^2, which, as we show, is computed in such a way that the value of the variance swap contract is zero at contract inception ($t = 0$); i.e., since this is structured as a forward contract, it must have zero initial cost. Computing the expectation of the pay-off at initial time $t = 0$ and setting this to zero therefore gives the fair value of this variance in terms of a stochastic integral:

$$\Sigma^2 = \frac{1}{T}E_0^{Q(B)}\left[\int_0^T \sigma_t^2\ dt\right]. \tag{1.339}$$

We shall compute this expectation by recasting the integrand as follows. Assuming a diffusion process for futures prices and assuming that European call and put options of all strikes and maturity T are available, such a contract can be replicated exactly.[19]

More precisely, assume that futures prices $F_t^* \equiv F_t^*(A, T)$ on a contract maturing at time T with underlying asset price A_t (e.g., a stock price) at time t obeys the following zero-drift process under the risk-neutral measure *Q(B)*:

$$\frac{dF_t^*}{F_t^*} = \sigma_t\ dW_t, \tag{1.340}$$

where the volatility σ_t is a random process that can generally depend on time as well as on other stochastic variables.

Then consider the dynamic trading strategy, whereby at time t one holds $\frac{1}{F_t^*}$ futures contracts. If one starts implementing the strategy at time $t = 0$ and accumulates all the gains and losses from the futures position into a money-market account, then the worth Π_T accumulated at time T is

$$\Pi_T = \int_0^T \frac{dF_t^*}{F_t^*} = \int_0^T \sigma_t\ dW_t. \tag{1.341}$$

Due to Itô's lemma we have

$$d\log F_t^* = \frac{dF_t^*}{F_t^*} - \frac{1}{2}\left(\frac{dF_t^*}{F_t^*}\right)^2 = \frac{dF_t^*}{F_t^*} - \frac{\sigma_t^2}{2}dt, \tag{1.342}$$

and integrating from time $t = 0$ to T we find

$$\log F_T^* - \log F_0^* = \int_0^T \sigma_t\ dW_t - \frac{1}{2}\int_0^T \sigma_t^2\ dt = \Pi_T - \frac{1}{2}\int_0^T \sigma_t^2\ dt, \tag{1.343}$$

[19] We point out that in actuality the price of a variance swap is largely model independent. That is, it is possible to replicate the cash flows as long as the trader can set up a static hedge and trade futures on the underlying.

where equation (1.340) has been used. Rearranging this equation gives the integrand in equation (1.339) as

$$\frac{1}{T}\int_0^T \sigma_t^2 \, dt = \frac{2}{T}\Pi_T - \frac{2}{T}\log\frac{F_T^*}{F_0^*}. \tag{1.344}$$

This last expression demonstrates the precise nature of the replication. This contains (i) a static part given by the logarithmic payoff function and (ii) a dynamic part given by the stochastic time integral Π_T. Substituting this last expression into equation (1.339) and using the fact that Π_t is a martingale,[20] i.e., $E_0^{Q(B)}[\Pi_T] = 0$, we obtain

$$\Sigma^2 = -\frac{2}{T}E_0^{Q(B)}\left[\log\frac{F_T^*}{F_0^*}\right]. \tag{1.345}$$

Replicating the logarithmic payoff function in terms of standard call and/or put pay-offs of various strikes using the replication schemes described in Section 1.8 then gives a formula for Σ^2 in terms of futures calls and/or puts. In particular, by applying replication equation (1.248) on the domain $F_T^* \in (0, \infty)$ and taking expectations, equation (1.345) takes the form (see Problem 2)

$$\Sigma^2 = \frac{2}{T}\left[1 - \frac{F_0^*}{\bar{F}} - \log\frac{\bar{F}}{F_0^*} + \int_0^{\bar{F}} P_0^*(K,T)\frac{dK}{K^2} + \int_{\bar{F}}^{\infty} C_0^*(K,T)\frac{dK}{K^2}\right], \tag{1.346}$$

with any choice of nonzero parameter $\bar{F} \in (0, \infty)$, and where $C_0^*(K, T)$ and $P_0^*(K, T)$ represent the current $t = 0$ prices of a futures call and put option, respectively, at strike K and maturity T. Note that this formula holds irrespective of what particular assumed form for the volatility σ_t. In the cases of analytically solvable diffusion models, such as some classes of state-dependent models studied in Chapter 3, the call and put options can be expressed in closed analytical form. Of course, if $\sigma_t = \sigma(t)$, i.e., a deterministic function of only time, then the futures price obeys a geometric Brownian motion, and in this case, according to our previous analysis, we have simple analytical expressions of the Black–Scholes type, with $C_t^*(K, T)$ given by equation (1.336), and

$$P_t^*(K,T) = E_t^{Q(B)}[(K - F_T^*)_+] = KN(-d_-) - F_t^*(A,T)N(-d_+), \tag{1.347}$$

with $d_\pm$ given by equation (1.337), wherein $\sigma \to \bar{\sigma} \equiv \sqrt{(T-t)^{-1}\int_t^T \sigma^2(s)ds}$. For a numerical implementation of the efficient replication of logarithmic pay-offs for variance swaps in cases where only a select number of market call contracts is assumed available, the reader is encouraged to complete the project on variance swaps in Part II.

Problems

Problem 1. Derive the equations for the drift and volatility of the forward price as discussed in this section. For the domestic asset assume the process

$$\frac{dA_t}{A_t} = \mu_t^A dt + \sigma_t^A \, dW_t. \tag{1.348}$$

[20] Here we recall the property for the first moment $E_0[\int_0^t f_s \, dW_s] = 0$, which is valid under a suitable measure and conditions on the adapted process f_t.

Let $Z_t(T)$ be the price process of a domestic discount bond of maturity T. For any fixed maturity $T > t$, the discount bond price process is assumed to obey a stochastic differential equation of the form

$$\frac{dZ_t(T)}{Z_t(T)} = \mu^Z dt + \sigma^Z dW_t, \tag{1.349}$$

where shorthand notation is used ($\mu^Z \equiv \mu_t^{Z(T)}$, $\sigma^Z \equiv \sigma_t^{Z(T)}$) to denote the lognormal drift and volatility functions of the discount bond. Find the drift of the forward price process $F_t(A, T)$, defined by equation (1.322), within the following three different choices of numeraire asset g_t: (i) the money-market account: $g_t = B_t = e^{\int_0^t r_s ds}$, where r_t is the domestic short rate at time t, (ii) the discount bond: $g_t = Z_t(T)$, and (iii) the asset : $g_t = A_t$. Hint: Make use of the formula for the stochastic differential of a quotient of two processes that was derived in a previous section.

Problem 2. Use equation (1.248) with payoff function $\phi(F) = -\log \frac{F}{F_0^*}$, $F \equiv F_T^*$, $\bar{S} = \bar{F}$, $S_0 = 0$, $S_1 = \infty$, with $0 < \bar{F} < \infty$, to show

$$\phi(F) = 1 - \frac{F}{\bar{F}} - \log \frac{\bar{F}}{F_0^*} + \int_0^{\bar{F}} (K - F)_+ \frac{dK}{K^2} + \int_{\bar{F}}^{\infty} (F - K)_+ \frac{dK}{K^2}. \tag{1.350}$$

Now, arrive at the formula in equation (1.346) by taking the expectation of this pay-off at $t = 0$ under the measure *Q(B)* while making use of the fact that an expectation can be taken inside any integral over K and the fact that $E_t^{Q(B)}[F_T^*] = F_t^*$, i.e., that F_t^* is a martingale within this measure.

1.12 Pricing Formulas of the Black–Scholes Type

In this section we apply the fundamental theorem of asset pricing of Section 1.10 to derive a few exact pricing formulas. The worked-out examples are meant to demonstrate the use of different numeraire assets for option pricing.

Example 1. Plain European Call Option.

As a first example, let's revisit the problem of pricing the plain European call. Consider the Black–Scholes model (i.e., geometric Brownian motion) for a stock of constant volatility σ and in an economy with a constant interest rate r. Under the risk-neutral measure with money-market account $g_t = B_t = e^{rt}$ as numeraire, the expected return on the stock is just the risk-free rate r; hence,

$$dS_t = rS_t \; dt + \sigma S_t \; dW_t. \tag{1.351}$$

The stock price process is given in terms of a standard normal random variable [i.e., equation (1.154)]: $S_T = S_t e^{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\sqrt{T-t}x}$, $x \sim N(0, 1)$. Using equation (1.292), the arbitrage-free price at time t of a European call option struck at $K > 0$ with maturity $T > t$ is hence the discounted expectation under the risk-neutral measure *Q(B)*:

$$\begin{aligned} C_t(S_t, K, T) &= e^{-r(T-t)} E_t^{Q(B)}\left[(S_T - K)_+\right] \\ &= \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \left(S_t e^{\left(r - \frac{\sigma^2}{2}\right)(T-t) + \sigma\sqrt{T-t}x} - K \right)_+ dx \\ &= S_t N(d_+) - K e^{-r(T-t)} N(d_-), \end{aligned} \tag{1.352}$$

where

$$d_{\pm} = \frac{\log(S_t/K) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}. \tag{1.353}$$

Note that the details of this integral expectation were presented in Section 1.6.

This Black–Scholes pricing formula plays a particularly important role because it is the prototype for a large number of pricing formulas. As we shall see in a number of examples in this and the following chapters, analytically solvable pricing problems for European-style options often lead to pricing formulas of a similar structure. In the case that the underlying asset pays *continuous dividends*, the foregoing pricing formula for a European call (and the corresponding put) must be slightly modified. A similar derivation procedure also applies, as shown at the end of this section.

If the drift and the volatility are deterministic functions of time, $r = r(t)$ and $\sigma = \sigma(t)$, the Black–Scholes formula extends thanks to the formula in equation (1.167) of Section 1.6. Using again the money-market account $g_t = B_t = \exp(\int_0^t r(s)ds)$ as numeraire asset and setting

$$\bar{r}(t,T) = \frac{1}{(T-t)}\int_t^T r(u)du$$

gives $B_t/B_T = e^{-\bar{r}(t,T)(T-t)}$, and we find

$$\begin{aligned} C_t(S_t, K, T) &= e^{-\bar{r}(t,T)(T-t)} E_t^{Q(B)}\left[(S_T - K)_+\right] \\ &= \frac{e^{-\bar{r}(t,T)(T-t)}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\left(S_t e^{\left(\bar{r}(t,T)-\frac{\bar{\sigma}(t,T)^2}{2}\right)(T-t)+\bar{\sigma}(t,T)\sqrt{T-t}x} - K\right)_+ dx \\ &= S_t N(d_+) - K e^{-\bar{r}(t,T)(T-t)} N(d_-), \end{aligned} \tag{1.354}$$

where

$$\bar{\sigma}(t,T)^2 = \frac{1}{(T-t)}\int_t^T \sigma(t)^2\, dt, \tag{1.355}$$

and

$$d_{\pm} = \frac{\log(S_t/K) + \left(\bar{r}(t,T) \pm \frac{1}{2}\bar{\sigma}(t,T)^2\right)(T-t)}{\bar{\sigma}(t,T)\sqrt{T-t}}. \tag{1.356}$$

Note that, in agreement with the results obtained in Section 1.6, the Black–Scholes pricing formula now involves the time-averaged interest rate and volatility over the maturity time $T - t$.

Example 2. A Currency Option.

Let

$$dX_t = \mu_X X_t\, dt + \sigma_X X_t\, dW_t \tag{1.357}$$

be a model for the foreign exchange rate X_t at time t, assuming that the lognormal volatility σ_X of the exchange rate and drift μ_X are constants. Suppose that the domestic risk-free interest rate r^d and the foreign interest rate r^f are both constant, and let $B_t^d = e^{r^d t}$ and $B_t^f = e^{r^f t}$ be the worth of the two money-market accounts, respectively. The drift μ_X can be computed as follows. First we note that the foreign currency money-market account, after conversion into domestic currency, is a domestic asset and therefore must obey a price process of the form

$$d(X_t B_t^f) = (r^d + \sigma_g \sigma_{XB^f})(X_t B_t^f)dt + \sigma_{XB^f}(X_t B_t^f)dW_t, \tag{1.358}$$

where σ_g and σ_{XB^f} are lognormal volatilities of the numeraire g_t and $X_t B_t^f$, respectively. We shall choose $g_t = B_t^d$ (i.e., the domestic risk-neutral measure) giving $\sigma_g = 0$. By direct application of Itô's lemma for the product of two processes we also have the stochastic differential

$$d(X_t B_t^f) = X_t\, dB_t^f + B_t^f\, dX_t + (dX_t)(dB_t^f) = X_t\, dB_t^f + B_t^f\, dX_t, \tag{1.359}$$

where the third term in the middle expression is of order $dt\ dW_t$ and hence set to zero. This follows since both domestic and foreign money-market accounts satisfy a deterministic differential equation, in particular,

$$dB_t^f = r^f B_t^f dt. \tag{1.360}$$

Plugging this and equation (1.357) into equation (1.359) gives

$$d(X_t B_t^f) = (r^f + \mu_X)(X_t B_t^f)dt + \sigma_X (X_t B_t^f)dW_t. \tag{1.361}$$

Hence, comparing equations (1.358) and (1.361) gives $\mu_X = r^d - r^f$. The foreign exchange rate therefore follows a geometric Brownian motion with this constant drift and constant volatility σ_X. The pricing formula for a foreign exchange call option struck at exchange rate K is then

$$\begin{aligned} C_t(X_t, K, T) &= e^{-r^d(T-t)} E_t^{Q(B^d)}\big[(X_T - K)_+\big] \\ &= \frac{e^{-r^d(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \left(X_t e^{((r^d - r^f) - \frac{\sigma_X^2}{2})(T-t) + \sigma_X \sqrt{T-t}x} - K \right)_+ dx \\ &= e^{-r^d(T-t)}\big[e^{(r^d - r^f)(T-t)} X_t N(d_+) - K N(d_-)\big], \\ &= e^{-r^f(T-t)} X_t N(d_+) - K e^{-r^d(T-t)} N(d_-), \end{aligned} \tag{1.362}$$

where

$$d_{\pm} = \frac{\log(X_t/K) + (r^d - r^f \pm \frac{1}{2}\sigma_X^2)(T-t)}{\sigma_X \sqrt{T-t}}. \tag{1.363}$$

Example 3. A Quanto Option.

Consider the case of a quanto option, in which we have a stock denominated in a foreign currency with geometric Brownian process

$$dS_t^f = \mu S_t^f\ dt + \sigma_S S_t^f\ dW_t^S, \tag{1.364}$$

and the foreign exchange process is also a geometric Brownian motion, with

$$dX_t = (r^d - r^f) X_t\ dt + \sigma_X X_t\ dW_t^X, \tag{1.365}$$

under the risk-neutral measure with numeraire $g_t = B_t^d$. Note that the drift rate $\mu_X = r^d - r^f$ was derived in the previous example. The constants σ_S and σ_X are the lognormal volatilities of the stock and foreign exchange rate, respectively. These Brownian increments are not independent; however, the foregoing equations can also be written equivalently in terms of two independent Brownian increments dW_t^1, dW_t^2, where

$$dW_t^X = \rho\ dW_t^1 + \sqrt{1-\rho^2}\ dW_t^2, \qquad dW_t^S = dW_t^1.$$

Here ρ is a correlation between the stock price and the foreign exchange rate at time t, with

$$dW^S\ dW^X = \rho\ dt. \tag{1.366}$$

In vector notation, $d\mathbf{W}_t = (dW_t^1, dW_t^2)$ and

$$\frac{dX_t}{X_t} = (r^d - r^f)dt + \boldsymbol{\sigma}_X \cdot d\mathbf{W}_t, \tag{1.367}$$

$$\frac{dS_t^f}{S_t^f} = \mu dt + \boldsymbol{\sigma}_S \cdot d\mathbf{W}_t, \tag{1.368}$$

where $\boldsymbol{\sigma}_X = (\rho\sigma_X, \sigma_X\sqrt{1-\rho^2})$, $\boldsymbol{\sigma}_S = (\sigma_S, 0)$. Suppose one wants to price a call option on the stock S_t^f struck at K and then to convert this into domestic currency at a preassigned fixed rate $\bar{X}$. Since $g_t = B_t^d$, the prices of all domestic assets (as well as the prices of foreign assets denominated in domestic currency) drift at the domestic risk-free rate. Hence the return on the price process $X_t S_t^f$ must be r^d. This also follows because the price of risk $q^g = q^{B^d} = \sigma_{B^d} = 0$. By direct application of Itô's lemma we also have

$$\frac{d(X_t S_t^f)}{X_t S_t^f} = \frac{dS_t^f}{S_t^f} + \frac{dX_t}{X_t} + \frac{dS_t^f}{S_t^f}\frac{dX_t}{X_t}. \tag{1.369}$$

Plugging the preceding expressions into this equation gives

$$\begin{aligned}\frac{d(X_t S_t^f)}{X_t S_t^f} &= (\mu + r^d - r^f + \boldsymbol{\sigma}_X \cdot \boldsymbol{\sigma}_S)dt + (\boldsymbol{\sigma}_X + \boldsymbol{\sigma}_S) \cdot d\mathbf{W}_t \\ &= (\mu + r^d - r^f + \rho\sigma_X\sigma_S)dt + \sigma_S dW_t^S + \sigma_X dW_t^X\end{aligned} \tag{1.370}$$

Since the drift must equal r^d,

$$\mu = r^f - \rho\sigma_S\sigma_X \tag{1.371}$$

is the constant drift of S_t^f in equation (1.364). The arbitrage-free price of a quanto call option struck at foreign price K is then given by

$$\begin{aligned}C_t(S_t^f, K, T) &= \bar{X}e^{-r^d(T-t)}E_t^{Q(B^d)}\left[\left(S_T^f - K\right)_+\right] \\ &= \bar{X}e^{-r^d(T-t)}\left[e^{\mu(T-t)}S_t^f N(d_+) - KN(d_-)\right] \\ &= \bar{X}\left[e^{-(r^d - r^f + \rho\sigma_S\sigma_X)(T-t)}S_t^f N(d_+) - e^{-r^d(T-t)}KN(d_-)\right],\end{aligned} \tag{1.372}$$

where

$$d_\pm = \frac{\log(S_t^f/K) + \left(r^f - \rho\sigma_S\sigma_X \pm \frac{1}{2}\sigma_S^2\right)(T-t)}{\sigma_S\sqrt{T-t}}. \tag{1.373}$$

Example 4. Elf-X Option (Equity-Linked Foreign Exchange Option).

Assume equation (1.364), as in the previous example, and now write

$$dX_t = \mu_X X_t \, dt + \sigma_X X_t \, dW_t^X \tag{1.374}$$

for the foreign exchange process, with μ_X dependent on the choice of numeraire. Consider the case where the pay-off is

$$C_T = (X_T - K)_+ S_T^f. \tag{1.375}$$

The foreign asset price S_t^f cannot be used as a domestic numeraire asset, but the converted process $g_t = X_t S_t^f$ can. Indeed this is a positive price process denominated in domestic currency. Under the measure with g_t as numeraire we first need to compute the drift μ_X explicitly. This is done by considering the process $X_t B_t^f$, which must drift at the domestic risk-free rate plus a price-of-risk component

$$\frac{d(X_t B_t^f)}{X_t B_t^f} = (r^d + \boldsymbol{\sigma}_{XS^f} \cdot \boldsymbol{\sigma}_{XB^f})dt + \boldsymbol{\sigma}_{XB^f} \cdot d\mathbf{W}_t, \tag{1.376}$$

where $\boldsymbol{\sigma}_{XS^f}$ and $\boldsymbol{\sigma}_{XB^f}$ are volatility vectors of the price processes $X_t S_t^f$ and $X_t B_t^f$, respectively. These are expressible in the basis of either (dW_t^1, dW_t^2) or (dW_t^S, dW_t^X), as described in the previous example. [Note also that the Brownian increments, written still as $d\mathbf{W}_t$ in the SDE are actually w.r.t. the measure $Q(XS^f)$.] From equation (1.370) we have $\boldsymbol{\sigma}_{XS^f} = \boldsymbol{\sigma}_X + \boldsymbol{\sigma}_S$. From a direct application of Itô's lemma we also have

$$\frac{d(X_t B_t^f)}{X_t B_t^f} = (r^f + \mu_X)dt + \boldsymbol{\sigma}_X \cdot d\mathbf{W}_t. \tag{1.377}$$

By equating drifts and the volatility vectors in these two expressions we find $\boldsymbol{\sigma}_{XB^f} = \boldsymbol{\sigma}_X$ and

$$r^d + (\boldsymbol{\sigma}_S + \boldsymbol{\sigma}_X) \cdot \boldsymbol{\sigma}_X = r^f + \mu_X. \tag{1.378}$$

Hence,

$$\mu_X = r^d - r^f + \boldsymbol{\sigma}_S \cdot \boldsymbol{\sigma}_X + ||\boldsymbol{\sigma}_X||^2.$$

The drift $\mu_{X^{-1}}$ and volatility of the inverse exchange rate X_t^{-1} under the same measure are computed using Itô's lemma [i.e., apply equation (1.138) with numerator = 1 and denominator = X_t]:

$$\frac{dX_t^{-1}}{X_t^{-1}} = (-\mu_X + \sigma_X^2)dt - \sigma_X \, dW_t^X.$$

Hence,

$$\mu_{X^{-1}} = -\mu_X + \sigma_X^2 = r^f - r^d - \boldsymbol{\sigma}_S \cdot \boldsymbol{\sigma}_X = r^f - r^d - \rho\sigma_X\sigma_S,$$

where the square of the volatility is the same as that of X_t, namely σ_X^2. Using the measure $Q(XS^f)$, we therefore have the arbitrage-free price:

$$\begin{aligned} C_t(S_t^f, X_t, K, T) &= (X_t S_t^f) E_t^{Q(XS^f)}\left[\frac{S_T^f (X_T - K)_+}{X_T S_T^f}\right] \\ &= K X_t S_t^f E_t^{Q(XS^f)}[(K^{-1} - X_T^{-1})_+] \\ &= K X_t S_t^f \left[K^{-1} N(-d_-) - e^{\mu_{X^{-1}}(T-t)} X_t^{-1} N(-d_+)\right] \\ &= S_t^f \left[X_t N(-d_-) - e^{-(r^d - r^f + \rho\sigma_X\sigma_S)(T-t)} K N(-d_+)\right], \end{aligned} \tag{1.379}$$

where

$$d_{\pm} = \frac{\log(K/X_t) + (r^f - r^d - \rho\sigma_X\sigma_S \pm \frac{1}{2}\sigma_X^2)(T-t)}{\sigma_X\sqrt{T-t}}. \tag{1.380}$$

Let us now consider Black–Scholes pricing formulas as well as symmetry relations for European calls and puts under an economy whereby the underlying asset pays continuous dividends. This will be useful for the discussion on American options in Section 1.14. In particular, let us assume that the asset price S_t follows geometric Brownian motion, as in Example 1, but with an additional drift term due to a constant dividend yield q:

$$dS_t = (r-q)S_t\, dt + \sigma S_t\, dW_t. \tag{1.381}$$

Note that from equation (1.165) we readily have the risk-neutral lognormal transition density for this asset price process,

$$p(S_T, S_t; \tau) = \frac{1}{\sigma S_T\sqrt{2\pi\tau}} e^{-[\log(S_T/S_t)-(r-q-\frac{\sigma^2}{2})\tau]^2/2\sigma^2\tau}, \tag{1.382}$$

$\tau = T - t$. We follow Example 1 and choose $B_t = e^{rt}$ as numeraire. Then, using equation (1.169) with drift $(r-q)$ as given by equation (1.381), the price of a European call struck at K with underlying asset paying continuous dividend q is

$$\begin{aligned} C_t(S_t, K, T) &= e^{-r(T-t)}E_t^{Q(B)}\big[(S_T - K)_+\big] \\ &= e^{-r(T-t)}\big[e^{(r-q)(T-t)}S_tN(d_+) - KN(d_-)\big] \\ &= e^{-q(T-t)}S_tN(d_+) - Ke^{-r(T-t)}N(d_-), \end{aligned} \tag{1.383}$$

with

$$d_{\pm} = \frac{\log(S_t/K) + \left(r - q \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}. \tag{1.384}$$

The corresponding European put price is easily derived in similar fashion, giving

$$P_t(S_t, K, T) = Ke^{-r(T-t)}N(-d_-) - S_te^{-q(T-t)}N(-d_+). \tag{1.385}$$

The previous put-call parity relation for plain European calls and puts, i.e., equation (1.214), is now modified to read

$$C_t(S_t, K, T) - P_t(S_t, K, T) = e^{-q(T-t)}S_t - Ke^{-r(T-t)} \tag{1.386}$$

for generally nonzero q.

This put-call parity is a rather general property that obtains whenever relative asset prices are martingales. Within the geometric Brownian motion model, we can further establish *another* special symmetry property that relates a call price to its corresponding put price. In particular, explicitly denoting the dependence on the interest rate r and dividend yield q, we have

$$C_t(S, K, T; r, q) = P_t(K, S, T; q, r). \tag{1.387}$$

This relation states that the Black–Scholes pricing formula for a call, with spot $S_t = S$, strike K, interest rate r, and dividend q, is the same as the Black–Scholes pricing formula for a put where one inputs the strike as S, spot as $S_t = K$, interest rate as q, and dividend as r. That is, by interchanging r and q and interchanging S and K, the call and put pricing formulas give the same price. For this reason we can also refer to identity (1.387) as a *put-call reversal symmetry*. This result can be established by relating expectations under different numeraires as follows. Consider the modified asset price process defined by $\tilde{S}_t \equiv e^{qt} S_t$, then Itô's lemma gives

$$d\tilde{S}_t = r\tilde{S}_t \, dt + \sigma \tilde{S}_t \, dW_t \tag{1.388}$$

within the risk-neutral measure. By alternatively choosing $\tilde{g}_t = \tilde{S}_t$ as numeraire, equation (1.292) gives the arbitrage-free price of the call as

$$C_t(S, K, T; r, q) = SKe^{-q(T-t)} E_t^{Q(\tilde{g})} \left[(K^{-1} - X_T)_+ \right] \tag{1.389}$$

where we have used the spot value $S_t = S$ and defined the process $X_t \equiv S_t^{-1}$. From equation (1.388), we see that the lognormal volatility of $\tilde{g}_t$ (or the price of risk) is σ; therefore, under the new measure, $Q(\tilde{g})$, equation (1.381) becomes

$$dS_t = (r - q + \sigma^2) S_t \, dt + \sigma S_t \, d\tilde{W}_t, \tag{1.390}$$

where $d\tilde{W}_t$ denotes the Brownian increment under measure $Q(\tilde{g})$. Using this equation and applying Itô's lemma to $X_t = S_t^{-1}$ gives

$$dX_t = (q - r) X_t \, dt - \sigma X_t \, d\tilde{W}_t \tag{1.391}$$

Under $Q(\tilde{g})$, the transition density $\tilde{p}$ for the process X_t is hence given by equation (1.382) with r and q interchanged and the replacement $S_t \to X_t$, $S_T \to X_T$:

$$\tilde{p}(X_T, X_t; \tau) = \frac{1}{\sigma X_T \sqrt{2\pi\tau}} e^{-[\log(X_T/X_t) - (q - r - \frac{\sigma^2}{2})\tau]^2 / 2\sigma^2\tau}. \tag{1.392}$$

Under $Q(\tilde{g})$, the drift of the lognormal diffusion X_t is $q - r$. Using equations (1.171) and (1.392) with spot $X_t = 1/S_t = 1/S$ at current time t, the expectation in equation (1.389) is evaluated to give

$$C_t(S, K, T; r, q) = SKP_t(1/S, 1/K, T; q, r). \tag{1.393}$$

This establishes the identity, which is actually equivalent to equation (1.387), as can be verified using equation (1.385). Finally, note that equation (1.387) is also verified by directly manipulating equation (1.385) or (1.383).

A class of slightly more sophisticated options that can also be valued analytically within the Black–Scholes model are European-style *compound options*. Such contracts are options on an option. Examples are a *call-on-a-call* and a *call-on-a-put*. Such compound options are hence characterized by two expiration dates, T_1 and T_2, and two strike values. Let us specifically consider a call-on-a-call option. This contract gives the holder the right (not the obligation) to purchase an underlying call option for a fixed strike price K_1 at calendar time T_1. The underlying call is a call option on an asset or stock with strike K_2 and expiring at a later calendar time $T_2 > T_1$ — we denote its value by $C_{T_1}(S_{T_1}, K_2, T_2)$, where S_{T_1} denotes the stock price at T_1. Hence at time T_1 this underlying call will be purchased (i.e., the compound call-on-a-call will be exercised at time T_1) only if $C_{T_1}(S_{T_1}, K_2, T_2) > K_1$. Let t denote current calendar time, $t < T_1 < T_2$, then the pay-off of the call-on-a-call at T_1 is $\left(C_{T_1}(S_{T_1}, K_2, T_2) - K_1 \right)_+$.

Since C_{T_1} is a monotonically increasing function of S_{T_1}, this pay-off is nonzero only for values of S_{T_1} above a (critical) value S_1^* defined as the unique solution to the (nonlinear) equation $C_{T_1}(S_1^*, K_2, T_2) = K_1$. Hence $\left(C_{T_1}(S_{T_1}, K_2, T_2) - K_1\right)_+ = C_{T_1}(S_{T_1}, K_2, T_2) - K_1$ for $S_{T_1} > S_1^*$ and zero otherwise.

Denoting the value of the call-on-a-call option by $V^{cc}(S, t)$, where $S_t = S$ is the spot at time t, and assuming a constant interest rate with $g_t = e^{rt}$ as numeraire asset price, we generally have

$$V^{cc}(S,t) = e^{-r(T_1-t)} E_t^{Q(B)}\left[\left(C_{T_1}(S_{T_1}, K_2, T_2) - K_1\right)_+\right]. \tag{1.394}$$

Specializing to the case where the stock price process obeys equation (1.381) within the risk-neutral measure *Q(B)*, this expectation is readily evaluated in terms of the standard univariate cumulative normal and bivariate cumulative normal functions. Inserting the price of the call from equation (1.383) gives

$$V^{cc}(S,t) = e^{-r\tau_1} \int_{S_1^*}^{\infty} \left[e^{-q(T_2-T_1)} S_1 N(d_+) - K_2 e^{-r(T_2-T_1)} N(d_-) - K_1\right] p(S_1, S; \tau_1) dS_1, \tag{1.395}$$

$\tau_1 = T_1 - t$, where p is the transition density function defined in equation (1.382) and

$$d_\pm = \frac{\log(S_1/K_2) + \left(r - q \pm \frac{1}{2}\sigma^2\right)(T_2 - T_1)}{\sigma\sqrt{T_2 - T_1}}. \tag{1.396}$$

Equation (1.395) is a sum of three integrals. The third integral term involves the risk-neutral probability that the stock price is above S_1^* after a time τ_1 and having initiated at S. This integral is reduced to a standard cumulative normal function by changing the integration variable to $x = \log S_1$:

$$\int_{S_1^*}^{\infty} p(S_1, S; \tau_1) dS_1 = N(a_-), \tag{1.397}$$

where we define

$$a_\pm = \frac{\log(S/S_1^*) + \left(r - q \pm \frac{1}{2}\sigma^2\right)\tau_1}{\sigma\sqrt{\tau_1}}. \tag{1.398}$$

The second integral term in equation (1.395) can be rewritten using

$$N(d_-) = \int_{K_2}^{\infty} p(S_2, S_1; T_2 - T_1) dS_2, \tag{1.399}$$

giving

$$\int_{S_1^*}^{\infty} N(d_-) p(S_1, S; \tau_1) dS_1 = \int_{S_1^*}^{\infty} \int_{K_2}^{\infty} p(S_2, S_1; T_2 - T_1) p(S_1, S; \tau_1) dS_2\, dS_1. \tag{1.400}$$

This double integral can be recast in terms of a standard bivariate cumulative normal function

$$N_2(a, b; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{a} \int_{-\infty}^{b} \exp\left[-\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho^2)}\right] dy\, dx, \tag{1.401}$$

where ρ is a correlation coefficient. For this purpose it proves useful to define

$$\tau_2 = T_2 - t \qquad \text{and} \qquad \rho = \sqrt{\tau_1/\tau_2}, \tag{1.402}$$

hence $T_2 - T_1 = \tau_2 - \tau_1$. Introducing the change of variables

$$-x = \frac{\log(S_1/S) - \left(r - q - \frac{1}{2}\sigma^2\right)\tau_1}{\sigma\sqrt{\tau_1}}, -y = \frac{\log(S_2/S) - \left(r - q - \frac{1}{2}\sigma^2\right)\tau_2}{\sigma\sqrt{\tau_2}}$$

Equation (1.400) then becomes (after some algebraic manipulation)

$$\begin{aligned}\int_{S_1^*}^{\infty} N(d_-)p(S_1, S; \tau_1)dS_1 &= \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{a_-}\int_{-\infty}^{b_-} \exp\left[-\frac{x^2}{2} - \frac{(y-\rho x)^2}{2(1-\rho^2)}\right]dy\,dx,\\ &= N_2(a_-, b_-; \rho),\end{aligned} \tag{1.403}$$

where

$$b_{\pm} = \frac{\log(S/K_2) + \left(r - q \pm \frac{1}{2}\sigma^2\right)\tau_2}{\sigma\sqrt{\tau_2}}. \tag{1.404}$$

We leave it to the reader to verify that the first integral term in equation (1.395) can be reduced, using similar manipulations as earlier, to give

$$\int_{S_1^*}^{\infty} S_1 N(d_+)p(S_1, S; \tau_1)dS_1 = Se^{(r-q)\tau_1}N_2(a_+, b_+; \rho). \tag{1.405}$$

Combining the three integrals in equation (1.395) finally gives

$$V^{cc}(S, t) = Se^{-q\tau_2}N_2(a_+, b_+; \rho) - K_2 e^{-r\tau_2}N_2(a_-, b_-; \rho) - K_1 e^{-r\tau_1}N(a_-). \tag{1.406}$$

Derivations of similar pricing formulas for related types of compound options are left to the interested reader (see Problem 10).

Problems

Within the problems involving a single underlying asset or stock, assume we are in a Black–Scholes world where the asset price process obeys geometric Brownian motion of the form

$$dA_t = (r + q \cdot \sigma_A)A_t\, dt + \sigma_A A_t\, dW_t, \tag{1.407}$$

where dW_t is the assumed Brownian increment under the given measure, the interest rate r (in the appropriate economy) and volatility σ_A are constants, and q is a market price of risk.

Problem 1. Find the price of a call option on foreign stock struck in foreign currency, i.e., of the contract with payoff

$$C_T = X_T(S_T^f - K)_+. \tag{1.408}$$

Problem 2. Find the price of a call option on foreign stock struck in domestic currency with payoff

$$C_T = (X_T S_T^f - K)_+, \tag{1.409}$$

where X_t is the exchange rate at time t.

Problem 3. Consider again the example of the quanto option in Example 3. Compute the coefficient α in such a way that the price process

$$g_t \equiv \bar{X} e^{\alpha t} S_t^f \tag{1.410}$$

is a domestic asset price process. Further, price the quanto option in Example 3 using g_t as a numeraire asset. Describe the replication strategy for the numeraire asset g_t.

Problem 4. Derive the price of an Elf-X option from the point of view of the foreign investor taking as payoff

$$C_T = (S_T^f - K S_T^f Y_T)_+, \tag{1.411}$$

where $Y_T = 1/X_T$.

Problem 5. A forward starting call on a stock S is structured as follows. The holder will receive at a preassigned future time T_1 a call struck at $K = \alpha S_{T_1}$ and maturing at time $T_2 > T_1$. Here, α is a positive preassigned constant and S_{T_1} is the stock price realized at time T_1.

Find (i) the present time $t = t_0 \leq T_1$ price of the forward starting call prior to maturity T_1 and (ii) a static hedging strategy that applies up to time T_1. Using the result in (i), show that the price of the contract simplifies to that of a standard call struck at $K = \alpha S_0$ with time to maturity $T_2 - t_0$ in the limiting case that $T_1 \to t_0$ (with t_0, T_2 held fixed). On the other hand, show that in the limit $T_2 \to T_1$ (with t_0, T_1 held fixed) the contract price is simply given by $S_0(1-\alpha)_+$. This last result is consistent with the price of a standard call with maturity $t = T_1$ and strike $K = \alpha S_{T_1}$.

Problem 6. Consider two stocks S^1 and S^2 described by correlated geometric Brownian motion with constant volatilities σ_1 and σ_2 and with correlation ρ. As seen in Section 1.6, a simple chooser option yields the pay-off as the maximum of the two stock levels,

$$\max(S_T^1, S_T^2), \tag{1.412}$$

at the maturity date T. Find the price of this instrument at time $t < T$. Find the relationship between the price of this chooser option and that of the chooser with payoff $(S_T^2 - S_T^1)_+$.

One Solution: To solve for either option price, pick the price of stock 1 as numeraire, $g_t = S_t^1$. So, for instance, to price the latter option, show that the price C_t is given by an expectation

$$C_t = S_t^1 E_t^p\left[(f_T - 1)_+\right], \tag{1.413}$$

where the random variable $f_t = S_t^2/S_t^1$ obeys

$$\frac{df_t}{f_t} = (\rho\sigma_2 - \sigma_1)dW_t^1 + \sigma_2\sqrt{1-\rho^2}\; dW_t^2. \tag{1.414}$$

From this, show that we have

$$\log\frac{f_T}{f_t} = -\frac{\nu^2}{2}(T-t) + (\rho\sigma_2 - \sigma_1)(W_T^1 - W_t^1) + \sigma_2\sqrt{1-\rho^2}(W_T^2 - W_t^2), \tag{1.415}$$

where W_t^i are independent Wiener processes at time t and

$$\nu \equiv \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} = |\boldsymbol{\sigma}_f|.$$

Since $\log(f_T/f_t)$ is normally distributed, find its mean and variance and thereby obtain the lognormal drift and volatility of f_t, i.e., the lognormal density $p = p(f_T, f_t; T-t)$, giving the price

$$C_t = S_t^2 N(d_+) - S_t^1 N(d_-), \tag{1.416}$$

where

$$d_{\pm} \equiv \frac{\log(S_t^2/S_t^1) \pm \frac{1}{2}\nu^2(T-t)}{\nu\sqrt{T-t}}. \tag{1.417}$$

Problem 7. Derive the standard call option-pricing formula of Example 1 of this section, but this time use the stock price as numeraire, i.e., $g_t = S_t$. In particular, show that with this choice of numeraire,

$$\frac{d(1/S_t)}{(1/S_t)} = -r\ dt - \sigma\ dW_t, \tag{1.418}$$

where dW_t stands for the Brownian increment under the measure *Q(S)* with S_t as numeraire. Then show that this leads to

$$C_t(S_t, K, T) = KS_t E_t^{Q(S)}\big[(1/K - 1/S_T)_+\big]. \tag{1.419}$$

Note: This is related to the price of an European put contract where the random variable is now the inverse of the stock price struck at the inverse of the strike, i.e., $1/K$, and with drift $= -r$. Compute this expectation to obtain the final expression.

Problem 8. Consider a foreign money-market account $B_t^f = e^{\int_0^t r_s^f\,ds}$ (with interest rate in foreign currency given by r_t^f at time t), a domestic asset with price A_t^d, and a foreign asset with price A_t^f. Let X_t be the exchange rate process in converting foreign currency into domestic. Suppose we choose $g_t = A_t^d$ as our numeraire asset. Compute the drift of the following processes: X_t, B_t^f, and A_t^f, within the $Q(g)$ measure.

Problem 9. Consider a domestic asset with price A_t^d and a foreign asset with price A_t^f. Let the constant κ be the conversion factor

$$\kappa = \frac{A_0^d}{A_0^f}. \tag{1.420}$$

[Note that this is given in terms of the asset prices at some current time $t = 0$.]

(i) Find a pricing formula for the contract at current time $t = 0$ with payoff function

$$\max(A_T^d, \kappa A_T^f) \tag{1.421}$$

at maturity $t = T$. Assume that all relevant lognormal volatilities and correlations are constant.

(ii) How can one hedge this contract? Is it necessary to trade the foreign currency dynamically?

Problem 10. Derive pricing formulas analogous to equation (1.406) for (i) a call-on-a-put, (ii) a put-on-a-put, and (iii) a put-on-a-call.

1.13 Partial Differential Equations for Pricing Functions and Kernels

Consider the continuous-time model with *state-dependent volatility*

$$\frac{dS_t}{S_t} = (r(t) + q\sigma(S_t, t))dt + \sigma(S_t, t)dW_t, \tag{1.422}$$

where q is the price of risk (also equal to the volatility of the numeraire asset). Here, *r(t)* is a deterministic, time-dependent short rate consistent with the term structure of interest rates. The state-dependent volatility $\sigma(S, t)$ is sometimes called the *local volatility*.

The asset price process A_t of an European-style option contingent on the asset S in the model described by equation (1.422) is given by a *pricing function* *A(S,t)* through a formula of the form

$$A_t = A(S_t, t). \tag{1.423}$$

The existence of a pricing function is an expression of the fact that the current price of an European option depends only on current calendar time and on the current (i.e., spot) price $S = S_t$ for the underlying asset (assuming all other contract parameters are held fixed as the maturity time, etc.).

Theorem. (Black–Scholes Equation) *The pricing function A(S, t) of a European claim contingent on the asset S in equation (1.422) satisfies the Black–Scholes equation*

$$\frac{\partial A}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 A}{\partial S^2} + rS\frac{\partial A}{\partial S} - rA = 0, \tag{1.424}$$

where $r = r(t)$, $\sigma = \sigma(S, t)$.

This is a backward time parabolic partial differential equation related closely to the backward Kolmogorov equation, as we shall see later.

Proof. Choosing as numeraire asset the money-market account $B_t = e^{\int_0^t r(s)ds}$, the price of risk $q = 0$ and the risk-neutral pricing formula yields

$$E^{Q(B)}\left[dA_t\right] = r(t)A_t\ dt \tag{1.425}$$

Equation (1.424) follows by applying Itô's lemma to the calculation of $dA_t = dA(S_t, t)$. Namely,

$$\frac{\partial A}{\partial t} + rS\frac{\partial A}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial A}{\partial S^2} = rA, \tag{1.426}$$

$r = r(t)$, $\sigma = \sigma(S, t)$. Lastly, note that this follows simply from the Feynman–Kac theorem. □

A second important partial differential equation concerns the probability density function *P(S, t)* under the risk-neutral measure for the stock price values S at time t, given an initial Dirac delta function distribution at time $t = t_0$:

$$P(S, t = t_0) = \delta(S - S_0). \tag{1.427}$$

More explicitly, this function is given by $P(S, t) \equiv p(S, t; S_0, t_0)$; i.e., this is the risk-neutral transition probability density for the price of the underlying asset to begin at value S_0 at initial time t_0 and end with value $S_t = S$ at time t. The function $p(S, t; S_0, t_0)$ is also commonly referred to as a *pricing kernel*. We have already seen a specific example of this as the lognormal transition density for geometric Brownian motion. In general, the resulting equation, called the *Fokker–Planck* (or *forward Kolmogorov*) equation, is contained in the following statement.

Theorem 1.7. (Fokker–Planck Equation) *The probability density function P(S, t) under the risk-neutral measure for the stock price values S at time t satisfying initial condition (1.427) obeys the following equation:*

$$\frac{\partial P}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial S^2}(\sigma^2 S^2 P) - r\frac{\partial}{\partial S}(SP), \tag{1.428}$$

where $r = r(t)$, $\sigma = \sigma(S, t)$.

Proof. This result can be derived as a consequence of the Black–Scholes equation. Consider a generic asset with pricing function *A(S, t)* defined in the interval $t \in (t_0, T)$, we then have from risk-neutral valuation that at any time t,

$$A(S_0, t_0) = e^{-\int_{t_0}^{t} r(s)ds}\int_0^\infty P(S, t)A(S, t)dS. \tag{1.429}$$

Note here that we assume that the range of solution is $S \in (0, \infty)$, although the derivation can be extended to cases with different ranges. Taking the partial derivative with respect to calendar time t on both sides of this equation, we find

$$\int_0^\infty \left[-rPA + A\frac{\partial P}{\partial t} + P\left(rA - rS\frac{\partial A}{\partial S} - \frac{S^2\sigma^2}{2}\frac{\partial^2 A}{\partial S^2}\right)\right]dS = 0,$$

where $r = r(t)$, $\sigma = \sigma(S, t)$, and the Black–Scholes equation (1.424) has been used for $\frac{\partial A}{\partial t}$. Integrating the last two terms by parts we obtain:

$$-\int_0^\infty PS\frac{\partial A}{\partial S}dS = -(PSA)\Big|_0^\infty + \int_0^\infty A\frac{\partial}{\partial S}(SP)dS = \int_0^\infty A\frac{\partial}{\partial S}(SP)dS,$$

and

$$\begin{aligned}
-\int_0^\infty \frac{\sigma^2 S^2}{2}P\frac{\partial^2 A}{\partial S^2}dS &= -\frac{1}{2}\frac{\partial}{\partial S}(\sigma^2 S^2 P)\frac{\partial A}{\partial S}\Big|_0^\infty + \int_0^\infty \frac{\partial A}{\partial S}\frac{\partial}{\partial S}\left(\frac{\sigma^2 S^2 P}{2}\right)dS \\
&= -\frac{1}{2}(\frac{\partial A}{\partial S} - A)\frac{\partial}{\partial S}(\sigma^2 S^2 P)\Big|_0^\infty - \frac{1}{2}\int_0^\infty A\frac{\partial^2}{\partial S^2}(\sigma^2 S^2 P)dS \\
&= -\frac{1}{2}\int_0^\infty A\frac{\partial^2}{\partial S^2}(\sigma^2 S^2 P)dS.
\end{aligned}$$

In the last equation we have integrated by parts twice. Notice that the nonintegral terms all vanish, due to the boundary conditions on the probability density function P, namely, that the function P and the first and second partial derivatives with respect to S, are assumed to be rapidly decaying functions of S as $S \to 0$ and $S \to \infty$. Collecting terms gives that for *any* derivative pricing function *A(S,t)*,

$$\int_0^\infty A(S, t)\left[\frac{\partial P}{\partial t} + r\frac{\partial}{\partial S}(SP) - \frac{1}{2}\frac{\partial^2}{\partial S^2}(\sigma^2 S^2 P)\right]dS = 0.$$

This can only occur if the integrand term in brackets is identically zero; hence equation (1.428) is fulfilled. □

The corresponding backward Kolmogorov equation for the density is given by the so-called *Lagrange adjoint* of equation (1.428). By combining equations (1.429) [with $P(S, t) \equiv P \equiv p(S, t; S_0, t_0)$] and (1.424), we readily see that $e^{-\int_{t_0}^{t} r(s)ds} P$ must satisfy the same equation as $A(S_0, t_0)$ for all initial times $t_0 < t$. Simplifying the equation in terms of P only, we find the backward Kolmogorov equation:

$$\frac{\partial P}{\partial t_0} + \frac{1}{2}\sigma^2(S_0, t_0)S_0^2 \frac{\partial^2 P}{\partial S_0^2} + r(t_0)S_0 \frac{\partial P}{\partial S_0} = 0. \tag{1.430}$$

This is a backward-time parabolic partial differential equation of the form of the Black–Scholes equation [i.e., replacing *(S,t)* by (S_0, t_0) in equation (1.424)]. The only term missing is the compounding term $r(t_0)A$. However, as just mentioned, the function $e^{-\int_{t_0}^{t} r(s)ds} P$ does exactly satisfy the Black–Scholes equation. This is, not surprisingly, consistent with our discussion in Section 1.8, where we showed [see equation (1.231)] that the discounted transition density gives the current price of a European butterfly option with inifinitely narrow spread (i.e., the price of an Arrow–Debreu security).

A partial differential equation satisfied by the pricing function of European-style call options *C(S, t; K, T)* regarded explicitly as functions of the strike and maturity time arguments *(K,T)* [instead of functions of the arguments *(S, t)*, which are held fixed] can now be derived as follows.

Theorem 1.8. (Dual Black–Scholes Equation) *The pricing function for a European call option C(S, t; K, T) satisfies the following equation:*

$$\frac{\partial C}{\partial T} = -r(T)K \frac{\partial C}{\partial K} + \frac{1}{2}K^2\sigma^2(K, T)\frac{\partial^2 C}{\partial K^2}. \tag{1.431}$$

Proof. European-style call prices admit the following representation in terms of the risk-neutral transition probability density [i.e., the density for the risk-neutral measure *Q(B)*]:

$$C(K, T) = Z_0(T)E_0^{Q(B)}[(S-K)_+] = Z_0(T)\int_0^\infty P(S, T)(S-K)_+ \, dS, \tag{1.432}$$

where $Z_0(T) = e^{-\int_0^T r(s)ds}$. Without loss of generality we simply set current time $t = 0$. Using the property $\partial(S-K)_+/\partial K = -\theta(S-K)$, where $\theta(x)$ is the Heaviside step function with value 1 for $x \geq 0$ and value 0 for $x < 0$, the first and second derivatives of equation (1.432) with respect to the strike K give

$$\frac{\partial C}{\partial K} = -Z_0(T)\int_K^\infty P(S, T)dS, \tag{1.433}$$

and

$$\frac{\partial^2 C}{\partial K^2} = Z_0(T)P(K, T). \tag{1.434}$$

The derivative with respect to maturity is given by

$$\begin{aligned}\frac{\partial C}{\partial T} &= -rZ_0(T)\int_0^\infty (S-K)_+ P \, dS + Z_0(T)\int_0^\infty \frac{\partial P}{\partial T}\,(S-K)_+ \, dS \\ &= -rC + Z_0(T)\int_0^\infty \left[-r\frac{\partial}{\partial S}(SP) + \frac{1}{2}\frac{\partial^2}{\partial S^2}(\sigma^2 S^2 P)\right](S-K)_+ \, dS,\end{aligned}$$

where $r = r(T)$, $\sigma = \sigma(S, T)$. Note that we have used equation (1.428) with $t = T$. The integral containing the first derivative with respect to S can be evaluated by parts as follows:

$$\begin{aligned}\int_0^\infty (S-K)_+ \frac{\partial}{\partial S}(SP)dS &= -\int_K^\infty SP\; dS \\ &= -\int_0^\infty (S-K)_+ P\; dS - K\int_K^\infty P\; dS \\ &= \left[-C + K\frac{\partial C}{\partial K}\right] Z_0(T)^{-1},\end{aligned}$$

where we used the identity $S = (S-K)_+ + K$, for $S \in [K, \infty)$, and equations (1.432) and (1.433). The integral containing the second derivative can again be evaluated by parts:

$$\begin{aligned}\int_0^\infty (S-K)_+ \frac{\partial^2}{\partial S^2}(\sigma^2 S^2 P)dS &= -\int_K^\infty \frac{\partial}{\partial S}(\sigma^2 S^2 P)dS \\ &= \sigma^2(K, T)K^2 P(K, T) \\ &= Z_0(T)^{-1}\sigma^2(K, T)K^2\frac{\partial^2 C}{\partial K^2}.\end{aligned}$$

Collecting the intermediate results obtained so far, we arrive at the following dual Black–Scholes equation:

$$\begin{aligned}\frac{\partial C}{\partial T} &= -rC + rC - rK\frac{\partial C}{\partial K} + \frac{1}{2}K^2\sigma^2(K, T)\frac{\partial^2 C}{\partial K^2} \\ &= -rK\frac{\partial C}{\partial K} + \frac{1}{2}K^2\sigma^2(K, T)\frac{\partial^2 C}{\partial K^2}.\end{aligned}$$

□

A consequence of this result is the following, which may be used in practice to calibrate a local volatility surface $\sigma_l = \sigma(K, T)$ via market European call option prices across a range of maturities and strikes.

Theorem 1.9. (Derman–Kani) *If a local volatility function exists, then it is unique and it can be expressed in analytical closed form as follows in terms of call option prices:*

$$\sigma^2(K, T) = \frac{2}{K^2}\frac{\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K}}{\frac{\partial^2 C}{\partial K^2}}. \tag{1.435}$$

This PDE pricing formalism extends readily into arbitrary dimensions. A general connection between a system of SDEs and the corresponding forward (backward) Kolmogorov PDEs that govern the transition probability density is as follows. Consider a diffusion model with n correlated random processes $\mathbf{x}_t = (x_t^1, \ldots, x_t^n) \in \mathbb{R}^n$ satisfying the system of SDEs:

$$\frac{dx_t^i}{x_t^i} = \mu_i(\mathbf{x}_t, t)dt + \sum_{\alpha=1}^M \sigma_{i,\alpha}(\mathbf{x}_t, t)dW_t^\alpha; \qquad i = 1, \ldots, n, \tag{1.436}$$

with $M \geq 1$ independent Brownian motions, $dW_t^\alpha\, dW_t^\beta = \delta_{\alpha,\beta}\, dt$, and where the drifts and volatilites are generally functions of time t and $\mathbf{x}_t$. Let us define the differential operator $\mathcal{L}$ by

$$\mathcal{L}_{\mathbf{x},t} f \equiv \sum_{i=1}^n x_i\mu_i(\mathbf{x}, t)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n x_i x_j \nu_{i,j}(\mathbf{x}, t)\frac{\partial^2 f}{\partial x_i \partial x_j}, \tag{1.437}$$

with Lagrange adjoint operator $\mathcal{L}^\dagger$ given by

$$\mathcal{L}^\dagger_{\mathbf{x},t} f \equiv -\sum_{i=1}^{n} \frac{\partial\left[x_i \mu_i(\mathbf{x},t) f\right]}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2\left[x_i x_j \nu_{i,j}(\mathbf{x},t) f\right]}{\partial x_i \, \partial x_j}, \tag{1.438}$$

where the functions $\nu_{i,j}$, $i, j = 1, \ldots, n$, are defined by

$$\nu_{i,j}(\mathbf{x},t) = \sum_{\alpha=1}^{M} \sigma_{i,\alpha}(\mathbf{x},t)\sigma_{j,\alpha}(\mathbf{x},t). \tag{1.439}$$

These operators act on any sufficiently differentiable function $f = f(\mathbf{x}, t)$. The transition probability density $p = p(\mathbf{x}, t; \mathbf{x}_0, t_0)$ associated with the foregoing diffusion process then satisfies the forward (Fokker–Planck) Kolmogorov PDE,

$$\frac{\partial p}{\partial t} = \mathcal{L}^\dagger_{\mathbf{x},t} p \tag{1.440}$$

as well as the corresponding backward PDE,

$$\frac{\partial p}{\partial t_0} + \mathcal{L}_{\mathbf{x}_0,t_0} p = 0, \tag{1.441}$$

for all $t_0 < t$, with initial (or final) time condition

$$p(\mathbf{x}, t = t_0; \mathbf{x}_0, t_0) = p(\mathbf{x}, t; \mathbf{x}_0, t_0 = t) = \delta(\mathbf{x} - \mathbf{x}_0).$$

Assuming that a diffusion path starting at some point $\mathbf{x}_0$ at time t_0 and ending at a point $\mathbf{x}$ at time t must be at all possible points $\bar{\mathbf{x}}$ at any intermediate time $\bar{t}$, $t_0 \le \bar{t} \le t$, then a consistency requirement in the theory is the so-called *Chapman–Kolmogorov integral equation*:

$$p(\mathbf{x}, t; \mathbf{x}_0, t_0) = \int_{\mathbb{R}^n} p(\mathbf{x}, t; \bar{\mathbf{x}}, \bar{t}) p(\bar{\mathbf{x}}, \bar{t}; \mathbf{x}_0, t_0) d\bar{\mathbf{x}}. \tag{1.442}$$

Prices of European-style contingent claims can then be computed by taking integrals over an appropriate pricing kernel as follows. Suppose we are within a certain measure *Q(g)* where underlying assets depend on random variables x^i_t that have appropriate drift and volatilites in accordance with equation (1.436). Assuming the existence of a martingale measure where the numeraire is, for example, of the form $e^{\int_0^t \lambda(\mathbf{x}_s, s) ds}$ (i.e., with λ as a discounting function), then according to the asset pricing theorem of the previous section, the price of a contingent claim $A(\mathbf{x}, t)$ with payoff $\phi(\mathbf{x})$ is given by the expectation

$$A(\mathbf{x}, t) = E^Q_t\left[e^{-\int_t^T \lambda(\mathbf{x}_s, s) ds} \phi(\mathbf{x})\right]. \tag{1.443}$$

Then due to the Feynman–Kac formula (in n dimensions) we have the corresponding Black–Scholes PDE:

$$\frac{\partial A(\mathbf{x}, t)}{\partial t} + \mathcal{L}_{\mathbf{x},t} A(\mathbf{x}, t) - \lambda(\mathbf{x}, t) A(\mathbf{x}, t) = 0, \tag{1.444}$$

$t < T$, with terminal condition $A(\mathbf{x}, T) = \phi(\mathbf{x})$, as required. From this analysis we see that the price of the contingent claim satisfying this Black–Scholes type of PDE can in fact be expressed as an integral over the set of diffusion paths. With the particular choice

$\lambda(\mathbf{x}_t, t) = r(t)$ (the risk-free rate), then, the density p is the risk-neutral density expressed in the $\mathbf{x}$-space variables. The claim's price is then simply given by an integral in $\mathbb{R}^n$:

$$A(\mathbf{x}, t) = e^{-\int_t^T r(s)ds} \int p(\mathbf{x}_T, T; \mathbf{x}, t)\phi(\mathbf{x}_T)d\mathbf{x}_T. \tag{1.445}$$

This is a multidimensional extension of equation (1.429). Note also that here, variables $\mathbf{x}$ do not necessarily represent prices. In general, asset prices are functions of $\mathbf{x}$ and time t. A nice feature of such integral equations, among others, is the fact that they provide a solution whereby the kernel p and hence the expected values can be propagated forward in the time variable T, starting from $T = t$, where the delta function condition is employed.

Problems

Problem 1. Consider the one-dimensional lognormal density $p(S, S_0; t - t_0)$ given by equation (1.165). Show that it satisfies forward and backward equations of the form (1.428) and (1.430) as well as the Chapman–Kolmogorov equation,

$$\int_0^\infty p(S, \bar{S}; t - \bar{t})p(\bar{S}, S_0; \bar{t} - t_0)d\bar{S} = p(S, S_0; t - t_0), \tag{1.446}$$

$t_0 \le \bar{t} \le t$.

Problem 2. Consider the n-dimensional lognormal density given by equation (1.198). Verify that this density satisfies the appropriate Kolmogorov equations.

1.14 American Options

In this section we briefly present the theory for pricing American, or *early-exercise*, options. The distinction between an American-style option and its European counterpart is that the holder of the American option has the additional freedom or right to exercise the option at *any* date from contract inception until expiration. This additional *time optionality* generally gives rise to an additional worth, appropriately also referred to as the *early-exercise premium*. We mainly focus our discussion on calls and puts, although the theory is also useful for treating other types of pay-offs. Throughout this section, we shall assume that we are within a Black–Scholes world with only one underlying asset. Although the formal theory readily extends into the multiasset case, the practical implementation and analysis issues are nontrivial and not within the scope of our present discussion. The development of numerical methods for pricing multiasset American options remains a topic of active research (see, for example, [BD96, BG97b, BG97a, BKT01, Gla04]).

1.14.1 *Arbitrage-Free Pricing and Optimal Stopping Time Formulation*

To begin our discussion, we consider the case where the underlying asset (or stock) price process $(S_t)_{t\ge 0}$ follows the geometric Brownian motion model as given by equation (1.381) in the risk-neutral measure, where r is the risk-free interest rate and q is a continuous dividend yield. We therefore assume that $r \ge 0$, $q \ge 0$, σ are constants (i.e., state and time independent), although the formalism (i.e., the governing equations) readily extends to the case of state-dependent drift and volatility functions. Let t_0 be the present time (i.e., contract inception). An American call (or put) option struck at K with expiration at time T is a claim to a payoff $(S_t - K)_+$ (or $(K - S_t)_+$) that the holder can exercise at any intermediate time

t prior to maturity, i.e., $t_0 \le t \le T$. The time at which the option is exercised is a stopping time. Recall the simpler situation in which the stopping time is initially known (i.e., as in the case of a European-style claim), then from the theorem of asset pricing the arbitrage-free price of a claim with a given pay-off occurring at time t is simply given by the discounted expectation via equation (1.294). In particular, the value at present time t_0 of a cash flow $(S_t - K)_+$ delivered at a later time t is given by

$$e^{-r(t-t_0)} E_0\big[(S_t - K)_+\big],$$

where $E_0[\cdot] = E^Q[\cdot|\mathcal{F}_{t_0}] = E^Q[\cdot|S_{t_0} = S_0]$ is used as a simplified notation to denote the expectation at time t_0 within the risk-neutral measure *Q(B)*, with $B_t = e^{rt}$ as numeraire, conditional on $S_{t=t_0} = S_0$. This expectation gives us the fair value of the cash flow as long as the delivery time t is a given stopping time, which may either be deterministic or random. For the case in which the stopping time is given by the maturity, e.g., $t = T$, the foregoing expectation obviously corresponds to the price of an European call [as given by equation (1.383), with t, S_t replaced by t_0, S_0].

For American contracts the holder has the freedom to exercise at any time within the continuous set of values $\mathcal{T} = \{t : t_0 \le t \le T\}$, giving rise to an *optimal stopping time* (i.e., *early-exercise time*) at which the holder should exercise the option for maximal gain. In particular, we shall see that an *early-exercise boundary* arises on the (t, S_t)-plane (i.e., time-spot plane) that separates the domain $[t_0, T] \times \mathbb{R}_+$ into two subdomains. These consist of a so-called *continuation* domain, for which the option is not yet exercised, and a *stopping* domain, whereby the option is exercised early. Hence, a main distinction from the European case is that the exercise time is *not known prematurely* and must be optimally determined as part of the solution to the pricing problem. As observed later, the basic financial reasoning for the emergence of an early-exercise boundary is that the holder can either claim a profit from the underlying dividend income by opting to purchase the asset (e.g., for the case of a call) or profit from the interest that arises from selling the underlying asset and investing the proceeds in a money-market account (e.g., for a put).

More generally, let us consider a nonnegative payoff function $\phi(S)$, $S \in \mathbb{R}_+$. The values of the European and corresponding American claim to such a pay-off are given, respectively, by

$$V_E(S_0, T - t_0) = E_0\big[e^{-r(T-t_0)}\phi(S_T)\big] \tag{1.447}$$

and

$$V(S_0, T - t_0) = \sup_{t \in \mathcal{T}} E_0\big[e^{-r(t-t_0)}\phi(S_t)\big]. \tag{1.448}$$

Throughout this section we use V_E to distinguish the European price from its American counterpart. In equation (1.448) the supremum is taken over all possible stopping times in the set $\mathcal{T}$. Note that both pricing functions are functions of the current time to maturity $T - t_0$, as is generally true when the drift and volatility terms have no explicit time dependence. We remark that although various theoretical frameworks exist for the determination of optimal stopping times, exact analytical formulas for such quantities as well as for American option values in terms of known transcendental functions have not been found to date. This is the case for the geometric Brownian motion model and, of course, for the more complex state-dependent models. In Section 1.14.4 we develop an integral-equations approach for computing the early-exercise boundary and the American option value, whereas in this section we provide a discrete-time backward induction formulation, which is useful for approximating the continuous-time quantities.

Formally, the optimal stopping time, denoted by t^*, is given by the infimum over the set $\mathcal{T}$ such that the value of the American option is equal to its *intrinsic value* (or *face value*) as given by the pay-off at the observed asset price:

$$t^* = \inf\{t \in \mathcal{T}, V(S_t, T-t) = \phi(S_t)\}. \tag{1.449}$$

The *stopping domain*, corresponding to spot and time values for which it is optimal to exercise prematurely, consists of the set of points

$$\mathcal{D} = \{(t, S) : t \in \mathcal{T}, V(S, T-t) = \phi(S)\}, \tag{1.450}$$

while the *continuation domain*, corresponding to spot and time values for which the option is not exercised prematurely, is the set of points

$$\mathcal{C} = \{(t, S) : t \in \mathcal{T}, V(S, T-t) > \phi(S)\}. \tag{1.451}$$

Assuming there exists an optimal stopping time t^*, then from asset-pricing theory this time is given implicitly by

$$E_0\left[e^{-r(t^*-t_0)}\phi(S_{t^*})\right] = V(S_0, T-t_0). \tag{1.452}$$

This is a result that is not practical as it stands since the equation involves the American option value on the right-hand side, which is itself not yet known and dependent upon the stopping domain. This is a common feature among optimal stopping problems for Markov processes in continuous time, because they are essentially *free-boundary* value problems as shown shortly.

The structure of the stopping domains may be quite complicated for certain classes of payoff functions and diffusion models. However, for standard piecewise call/put types of pay-offs considered here, the domains turn out to be simply connected. In particular, the boundary of $\mathcal{D}$ is an early-exercise boundary curve given by

$$\partial\mathcal{D} = \{(\tau, S) : 0 \le \tau \le T - t_0, S = S^*(\tau)\}, \tag{1.453}$$

with $S^*(\tau)$ given by a smooth curve

$$S^*(\tau) = \min\{S > 0 : V(S, \tau) = (S-K)_+\} \tag{1.454}$$

for a call and

$$S^*(\tau) = \max\{S > 0 : V(S, \tau) = (K-S)_+\} \tag{1.455}$$

for a put struck at K. Here the function $V(S, \tau)$ represents the value of the American call $C(S, K, \tau)$ or put $P(S, K, \tau)$, respectively, where S is the value of the underlying spot. From equation (1.451) it is obvious that the continuation domain is the set of all points (τ, S) such that $V(S, \tau)$ is greater than the respective payoff function at S. As we will see, the subscript $+$ signs are actually redundant in equations (1.454) and (1.455). Note that here we have simply expressed the boundary and the option price in terms of the time-to-maturity variable $\tau = T - t \in [0, T - t_0]$ rather than the calendar time $t \in [t_0, T]$. This is convenient for what follows since the diffusion models are assumed to be time homogeneous. The optimal-exercise decision for the holder therefore depends on the observed spot (or stock price level) and the time to maturity (or calendar time) of the observation. In this sense, Amercian options can be characterized as having a kind of path dependence.

Before any further analysis, we make note of one very basic and important property of the early-exercise premium (or value): *The European option value V_E satisfies the condition (i) $V_E(S, \tau) \geq \phi(S)$ for all (S, τ) if and only if the corresponding American option value V satisfies (ii) $V(S, \tau) = V_E(S, \tau)$ for all (S, τ).* That is, if the corresponding European price is always above its intrinsic value during the contract lifetime, then it is never optimal to exercise the American option at any time earlier than expiry; i.e., there is no early-exercise premium and $V = V_E$. To show this, note that equation (1.448) implies $V(S, \tau) \geq V_E(S, \tau)$. Hence condition (i) gives $V(S, \tau) \geq \phi(S)$, so the American option is always above the intrinsic value, implying that the holder would not exercise earlier for a lower value. The optimal exercise (stopping) time is therefore at expiry T; hence (i) implies (ii). To prove the converse, observe that since the American option value must satisfy $V(S, \tau) \geq \phi(S)$ for all (S, τ), condition (ii) implies (i). This result is essentially a statement of the fact that an early-exercise boundary (and premium) arises only if the corresponding European option value falls below the intrinsic (payoff function) value. Because of this we have the following rather well-known result.

Proposition.
(i) An Amercian call has a nonzero early-exercise premium if and only if $q > 0$.
(ii) An Amercian put has a nonzero early-exercise premium if and only if $r > 0$.

This result will be seen to follow explicitly from the early-exercise boundary properties and the formulas for the early-exercise premiums developed in the following subsections. However, a simple and instructive proof goes as follows.

Proof. The put-call parity relation for European calls and puts gives

$$C_E(S, K, \tau) - P_E(S, K, \tau) = e^{-q\tau}S - e^{-r\tau}K. \tag{1.456}$$

Rewriting this we have

$$C_E(S, K, \tau) = S - K + P_E(S, K, \tau) + [(e^{-q\tau} - 1)S - (e^{-r\tau} - 1)K]. \tag{1.457}$$

Since $P_E(S, K, \tau) > 0$, then for $q = 0$ either of these expressions gives $C_E(S, K, \tau) > S - e^{-r\tau}K \geq S - K$. Hence C_E is always above its intrinsic value, and from the previous property we conclude that the European call value is equal to the American call value, $C_E(S, K, \tau) = C(S, K, \tau)$, so the early-exercise premium is zero. For the case $q > 0$, we use equation (1.457) and note that since the European put is a decreasing function of S, there exist large enough values of $S > K$ such that $P_E(S, K, \tau) + [(e^{-q\tau} - 1)S - (e^{-r\tau} - 1)K] < 0$, i.e., $C_E(S, K, \tau) < S - K$ for some $S > K$. From the previous result we therefore have $C(S, K, \tau) \neq C_E(S, K, \tau)$ and hence conclude that the early-exercise premium is nonzero for $q > 0$. This proves (i), while statement (ii) is proved in a similar fashion by reversing the roles of S,q with K,r and is left as an exercise. □

An obvious consequence of this proposition is that: (i) for an American call on a non-dividend-paying stock the exercise boundary is trivial (i.e., it is never optimal to exercise early), and (ii) for an American put on a nondividend-paying stock the exercise boundary is nontrivial (i.e., there is an optimal early-exercise time) if the interest rate is positive. In what follows (and also from the framework of Section 1.14.4) we will be able to further assess such properties.

Pricing by Recurrence: Dynamic Programming Approach

We now consider specifically the recursive formulation for pricing American options. This involves an iteration method that goes backward in calendar time (or forward in time to maturity). Formally, the American option price is given by equation (1.448). In order to actually implement this formula in a practical manner, we subdivide the time interval $[t_0, T] = [t_0, t_1, \ldots, t_N = T]$ into $N \geq 1$ subintervals $[t_i, t_{i+1}]$, $\delta t_i = t_{i+1} - t_i > 0$, $i = 0, \ldots, N-1$. For notational purposes it is useful to introduce the price function $V_t(S)$. For the case of time-homogeneous diffusions we have

$$V_t(S) \equiv V(S, T-t) = V(S, \tau), \tag{1.458}$$

with $\tau = T - t$ being the time remaining to maturity. We therefore assume that exercise can only occur at a fixed set of (intermediate stopping) times given by $\{t_i : i = 0, \ldots, N\}$. Equation (1.448) can then be approximated by

$$V_0(S_0) = \sup_{t \in \{t_i : i=0, \ldots, N\}} E_0\big[e^{-r(t-t_0)}\phi(S_t)\big], \tag{1.459}$$

$V_0(S_0) \equiv V_{t_0}(S_0) = V(S_0, T-t_0)$. For small δt_i values we expect equation (1.459) to be a good approximation to equation (1.448). From the theory of optimal stopping rules, one can show that in the limit $\delta t_i \to 0$ ($N \to \infty$) this approximation approaches the exact American option value in equation (1.448), which allows for continuous-time exercise. We remark that equation (1.459) actually gives the exact price of a *Bermudan* option with payoff function ϕ. Bermudans are bonafide contracts that essentially lie in between European and American contracts and are in reality structured specifically with only a fixed set of allowable exercise dates. Moreover, in any realistic trading strategy it is interesting to note that the actual information on asset price levels can only be accessible to the trader at intermittent times (i.e., at best one obtains "tick-by-tick"data). Hence, for the holder of an American option the exercise decision times, although approaching the continuum limit, essentially occur at discretely spaced points in time.

By discretizing time, the underlying asset price process with values $S_{t_i} \in \mathbb{R}_+$, $i = 0, \ldots, N$, is then a Markov chain. Iterating backward in calendar time starting from maturity, equation (1.459) is readily shown to imply that the option price at any intermediate time satisfies the recurrence relation

$$V_{t_i}(S) = \max\big\{\phi(S), E_{t_i}\big[e^{-r\delta t_i} V_{t_{i+1}}(S_{t_{i+1}}) | S_{t_i} = S\big]\big\}, \tag{1.460}$$

$i = N-1, \ldots, 0$, where $V_T(S) = \phi(S)$. This result states that the option price at each date t_i is given by the maximum of the pay-off (or the immediate-exercise value) and the discounted expected value of continuing without early exercise at time t_i. Note that at each ith step the expectation is conditional on $S_{t_i} = S$. [Remark: Equation (1.460) can also be rewritten as a forward recurrence relation in terms of a discretized time to maturity variable $\tau_i = T - t_i$ using equation (1.458)]. This formulation can be applied to asset prices that obey diffusion processes with generally state- and time-dependent drift and volatility functions. Here and in the following subsections, however, we are assuming time-homogeneous solutions; i.e., the drift and volatility functions of the asset price process are only allowed to be explicitly state dependent. Assuming a generally state-dependent Markov diffusion process $(S_t)_{t \geq 0}$, $S_t \in \mathbb{R}_+$ with assumed risk-neutral transition probability density function $p(S', S; \tau)$, the earlier expectation then gives

$$V_{t_i}(S) = \max\big\{\phi(S), \tilde{V}_{t_i}(S)\big\}, \tag{1.461}$$

where

$$\tilde{V}_{t_i}(S) = e^{-r\delta t_i} \int_0^\infty p(S', S; \delta t_i) V_{t_{i+1}}(S') dS' \tag{1.462}$$

represents the *continuation value* of the option at time t_i. For the particular process, of equation (1.381), p is specifically the lognormal density function given by equation (1.382). In this iteration approach, the American (or Bermudan) option prices are obtained without necessarily computing the early-exercise boundary. However, this can also be obtained simultaneously. From equation (1.461) we see that equations (1.449), (1.450), and (1.451) give the stopping rule

$$t^* = \min\{t_i; i = 0, \ldots, N : \phi(S_{t_i}) = \tilde{V}_{t_i}(S_{t_i})\}, \tag{1.463}$$

the early-exercise (stopping) domain as the union of line segments

$$\mathcal{D} = \cup_{i=0,\ldots,N}\{(t_i, S) : \phi(S) \geq \tilde{V}_{t_i}(S)\}, \tag{1.464}$$

and the continuation domain

$$\mathcal{C} = \cup_{i=0,\ldots,N}\{(t_i, S) : \phi(S) < \tilde{V}_{t_i}(S)\}. \tag{1.465}$$

Relation to Lattice (Tree) Methods

The dynamic programming approach provides a basis for implementing a number of different numerical methods for computing option prices using either Monte Carlo simulations, quadrature rules of integration, lattice methods, or a combination of such methods. In particular, the dynamic programming formulation can be directly related to the simplest of the lattice models — the binomial and trinomial lattices. For a detailed exposition on the implementations of lattice methods for pricing American options (as well as their European counterparts) the reader is urged to take a close look at the relevant numerical projects in Part II. The intricate details as well as the relevant equations and algorithms are explicitly described in those projects — the reader is also given the opportunity to numerically program the option-pricing applications. Here we shall simply give a very brief and generic discussion, meant only to emphasize the basic connection between the dynamic programming formulation and the lattice pricing models without having to repeat the underlying details.

Lattice methods can be viewed as either: (i) approximate solutions to recurrence relation (1.460) (or alternatively as approximate solutions to the equivalent option-pricing PDE by way of finite differences) or (ii) option-pricing models in their own right. Lattice models can accommodate time-inhomogeneous processes, as is the case for time-dependent drift and/or volatility functions. However, let's assume time-homogeneous models, where the underlying asset or stock price process is essentially modeled as a Markov chain on a discrete set of possible states. Generally, one assumes that the stock price can only move on a set of *nodes*, each denoted by a pair of integers *(i,j)* corresponding to a stock price value S_j^i. The lattice is a mesh or grid made up of all such nodes, where the integer j is an index for the spatial position of the stock price on the lattice at time t_i, $i = 0, \ldots, N$. Lattice models allow for the implementation of time steps of fixed or variable size, but for the sake of simplicity let's assume a fixed time step of size $\delta t = (T - t_0)/N$. In fact, most implementations are based on equal-size time steps. Then conditional on $S_{t_i} = S_j^i$, the probability of a movement of the stock price within a single time step δt from a node (i, j) into a successor node $(i+1, j')$, with value $S_{t_{i+1}} = S_{j'}^{i+1}$, is given by the transition probability value $P(S_{t_{i+1}} = S_{j'}^{i+1} | S_{t_i} = S_j^i) \equiv p_{j \to j'} > 0$.

Although not critical to the present discussion, we note that for the binomial model there are only two successor nodes with $j' = j, j+1$, whereas the trinomial model has three successor nodes with $j' = j-1, j, j+1$, and so on.

The positive quantities $p_{j\to j'}$ are risk-adjusted probabilities and must obviously obey probability conservation,

$$\sum_{j'} p_{j\to j'} = 1, \qquad \text{for all } j, \tag{1.466}$$

where the sum is over all successor nodes in the model. Assuming the risk-neutral measure with money market as numeraire, the expected rate of return of the stock must equal the risk-free rate; i.e., $E_t[S_{t+\delta t}] = S_t e^{r\delta t}$. This is the risk-neutrality or no-arbitrage condition. For the lattice model it takes the form

$$\sum_{j'} p_{j\to j'} S_{j'}^{i+1} = e^{\mu\delta t} S_j^i, \tag{1.467}$$

for all (i, j) nodes, where $\mu = r$ or $\mu = r - q$ for nondividend- or dividend-paying stock. In order to capture the variance in the asset price returns, the lattice model is also built to take into account the asset price volatilty. For instance, one can relate the variation either of stock prices or of the log-returns that are computed separately using the diffusion model and the lattice model. If the variation or second moment of the log-returns are considered, then we have $E_t[(\delta \log S_t)^2] = (\sigma(S_t))^2\, \delta t$ within order δt, where $\sigma(S_t)$ is the *local* volatility function for the general case of a state-dependent diffusion model of the form $\delta S_t = \mu(S_t)S_t\, \delta t + \sigma(S_t)S_t\, \delta W_t$. Applying this same expectation at each node within the lattice model and equating the two expectations gives

$$(\sigma_j^i)^2 \delta t = \sum_{j'} p_{j\to j'} \log^2(S_{j'}^{i+1}/S_j^i), \tag{1.468}$$

where $\sigma_j^i = \sigma(S_j^i)$ forms a set of volatility parameters. This is just one possible way of introducing lattice volatility parameters into the model. Equations (1.466), (1.467), and (1.468) are therefore collective constraints on the lattice geometry and the nodal transition probabilities. These form an integral part of the construction of the lattice model and its parameters — this is part of the model *calibration* procedure. Further steps in the calibration can also be undertaken by fitting the lattice parameters so that certain computed option prices exactly match the corresponding market prices. In most applications the number of adjustable lattice parameters is greatly reduced. In particular, for geometric Brownian motion there is only one volatility parameter, i.e., $\sigma_j^i \to \sigma$. Moreover, most lattice models are simplified by assuming that the nodal transitions are independent of the starting node, as is the case for constant local volatilities, i.e., $p_{j\to j'} \to p_{j'}$. For specific details on the contruction of lattices and on implementing various calibration schemes for American and European option pricing within the binomial and trinomial models, we again refer the reader to the relevant projects in Part II.

Once the lattice geometry and transition probabilities are determined, i.e., the lattice is calibrated, the option prices at each node in the lattice, $V_j^i = V_{t_i}(S_j^i)$, can be determined by recurrence:

$$V_j^i = \max\left\{\phi(S_j^i),\, e^{-r\delta t} \sum_{j'} p_{j\to j'} V_{j'}^{i+1}\right\}. \tag{1.469}$$

The current option price $V_0^0 = V_0(S_0)$ at spot $S_0^0 \equiv S_0$ is obtained by simply iterating over N time steps, starting from the known payoff $V_j^N = \phi(S_j^N)$ at the terminal node values S_j^N.

Equation (1.469) also divides up the lattice into two groups of nodes: (i) a stopping domain as the set $\{(i, j) : V_j^i = \phi(S_j^i)\}$ and (ii) a continuation domain as the set $\{(i, j) : V_j^i > \phi(S_j^i)\}$. This second set gives the times t_i and spot values S_j^i for which the option should not be exercised early. According to equation (1.463), the optimal stopping time is

$$t^* = \min\{t_i = i\delta t : V_j^i = \phi(S_j^i)\}. \tag{1.470}$$

The early-exercise boundary is then also readily obtained. For instance, for a call this is the set of points $(i\delta t, S_*^i)$, $i = 0, \ldots, N$, where $S_*^i = \max\{S_j^i : V_j^i > S_j^i - K\}$; for a put, $S_*^i = \min\{S_j^i : V_j^i > K - S_j^i\}$. This offers a simple approach for approximating the early-exercise boundary curve in the continuous diffusion model corresponding to the limit $\delta t \to 0$. However, the resulting curve will not be smooth, even for relatively small time steps. More accurate calculations are afforded by applying more advanced techniques, such as the integral-equation approach discussed in Section 1.14.4. For the case of a trinomial lattice, equation (1.469) is related to the *explicit finite-difference* scheme for solving the Black–Scholes PDE. Alternative PDE solvers are based on *implicit finite-difference schemes*. Implicit schemes require the solution of a linear system of equations (or matrix inversion) for each time step in the propagation, yet they may offer more flexibility in the allowable range of lattice parameters for achieving accuracy and numerical stability. We refer the reader to the "Crank–Nicolson Option Pricer" project in Part II, which discusses a special type of implementation of the Crank–Nicolson implicit scheme for calibration and option pricing on a mesh.

The Smooth Pasting Condition and PDE Approach

Although the free-boundary curve is not analytically computable as a function of time, one can generally establish the *smooth pasting condition*. This property guarantees that the price function for an American option has a continuous derivative at the exercise boundary and that the derivative is equal to the derivative of the payoff function at the exercise boundary. The following proposition summarizes this result.

Proposition. *Let $\mathcal{D}_\tau$, with time to expiry $\tau = T - t > 0$, be the early-exercise domain for which $V_t(S) \equiv V(S, \tau) = \phi(S)$ when $S \in \mathcal{D}_\tau$, where ϕ is any differentiable payoff function. Then the American option price function V satisfies the smooth pasting condition at the boundary denoted by $S^*(\tau) \equiv S_t^*$:*

$$\left.\frac{\partial V(S, \tau)}{\partial S}\right|_{S=S^*(\tau)} = \phi'(S^*(\tau)), \tag{1.471}$$

and the zero-time-decay condition obtains on the early-exercise domain,

$$\frac{\partial V(S, \tau)}{\partial \tau} = 0, \qquad \textit{for } S \in \mathcal{D}_\tau. \tag{1.472}$$

Remark: The condition in equation (1.471) is also obviously valid for $S \in \mathcal{D}_\tau$ (excluding the boundary) since $V(S, \tau) = \phi(S)$ on that domain. What is important to emphasize here is that the derivative is continuous at the boundary of the stopping and continuation domains. These properties are valid under general proper Itô diffusion models. For a call (or put), then, equation (1.471) simply gives $\frac{\partial V(S^*(\tau),\tau)}{\partial S} = 1$ (or -1). This is illustrated in Figure 1.6. Although this proposition can be formally proven from the PDE approach, we shall instead demonstrate how it arises based on a dynamic hedging strategy argument, which turns out to be financially more insightful. First we note that the graph of the American option value

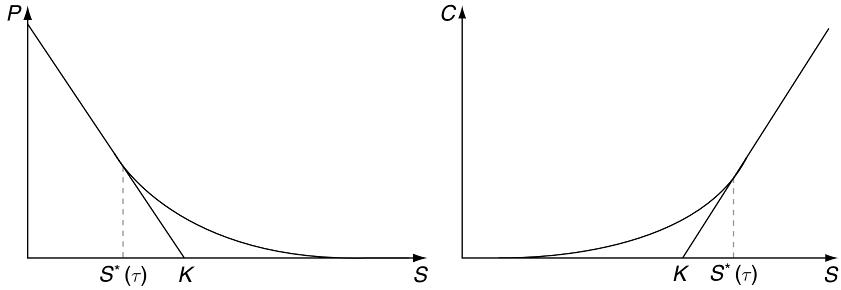

FIGURE 1.6 The pricing functions for an American put (left) and an American call (right) with continuous dividend yield satisfy the smooth pasting condition with slope equal to -1 and 1, respectively, at the optimal exercise boundary $S^*(\tau)$ for given time to expiry $\tau > 0$.

is never below that of the payoff function. Moreover, for given calendar time t (or time to maturity τ), the slope of the graph of $V_t(S) = V(S, \tau)$ at the exercise boundary point $S = S^*(\tau) \equiv S_t^*$ must be less (greater) than or equal to that of the payoff function if the latter is an increasing (decreasing) function at the boundary. That is: (i) $\frac{\partial V_t(S)}{\partial S}|_{S=S_t^{*(-)}} \leq \phi'(S_t^*)$ for the case $\phi'(S_t^*) \geq 0$ or (ii) $\frac{\partial V_t(S)}{\partial S}|_{S=S_t^{*(+)}} \geq \phi'(S_t^*)$ for the case $\phi'(S_t^*) \leq 0$. Here we use $S_t^{*(\pm)}$ to denote the limiting values from the right (+) or left ($-$) of S_t^*. Our objective is to show that these inequalities in the slopes are actually strict equalities. We now show this for case (i) as the argument follows in identical fashion for case (ii). In particular, let us assume that the asset or stock price at calendar time t is at the boundary; i.e., let $S_t = S_t^*$. After an infinitesimally small time lapse δt, the stock price can move either up into the exercise domain $\mathcal{D}_\tau$ or down into the (no-exercise) domain of continuation. If the stock price moves upward, then its change is $\delta S_t = S_{t+\delta t} - S_t^* > 0$, so $S_{t+\delta t} > S_t^*$ and it remains in the exercise domain. In this case, $V_{t+\delta t}(S_{t+\delta t}) = \phi(S_{t+\delta t})$ and the option value changes by an amount $\delta V_t = \phi(S_{t+\delta t}) - \phi(S_t^*) = \phi'(S_t^*)\delta S_t$, to leading order in δt. So to achieve a delta hedge for an upward tick over time δt, the option writer has to buy $\Delta_t = \phi'(S_t^*)$ shares of the stock. The writer's delta-hedge portfolio at time t consists of one short position in the option and Δ_t shares in the stock. Hence for an upward tick the hedge portfolio has value $\pi_t = -V_t(S_t^*) + \Delta_t S_t^* = -V_t(S_t^*) + \phi'(S_t^*)S_t^*$, and the change in portfolio value is $\delta\pi_t = -\delta V_t + \phi'(S_t^*)\delta S_t = 0$, to leading order in δt. On the other hand, if at time t the stock ticks down, then $\delta S_t < 0$, $S_{t+\delta t} < S_t^*$; hence the stock price falls into the domain of continuation. Now assume the SDE in equation (1.381) holds. [Note: The same argument also readily follows if we assume a more general Itô diffusion with state- and time-dependent drift and volatility.] To leading order, then,

$$\delta S_t = \sigma S_t^* \delta W_t = -\sigma S_t^* \sqrt{\delta t}|z|, \tag{1.473}$$

where $z \sim N(0, 1)$, since $\delta W_t < 0$ for a downward tick. Now, $\delta V_t = \frac{\partial V_t(S)}{\partial S}|_{S=S_t^{*(-)}} \delta S_t$ and, using the foregoing expression, the hedge portfolio changes by

$$\begin{aligned}\delta\pi_t &= -\delta V_t + \phi'(S_t^*)\delta S_t \\ &= \left[\frac{\partial V_t(S)}{\partial S}\bigg|_{S=S_t^{*(-)}} - \phi'(S_t^*) \right] \sigma S_t^* \sqrt{\delta t}|z|. \end{aligned} \tag{1.474}$$

Taking expectations and using $E[|z|] = \sqrt{2/\pi}$ gives the expected change in the hedge portfolio:

$$E[\delta\pi_t] = \sqrt{\frac{2}{\pi}}\left[\frac{\partial V_t(S)}{\partial S}\bigg|_{S=S_t^{*(-)}} - \phi'(S_t^*)\right]\sigma S_t^* \sqrt{\delta t}. \tag{1.475}$$

We hence conclude that the writer cannot exactly set up a delta hedge portfolio and in particular is expected to suffer a loss every time the underlying stock is in the vicinity of the boundary *unless* $\frac{\partial V_t(S_t^{*(-)})}{\partial S} = \phi'(S_t^*)$. Since $\frac{\partial V_t(S_t^{*(+)})}{\partial S} = \phi'(S_t^*)$, the function $\frac{\partial V_t(S)}{\partial S} \equiv \frac{\partial V(S,\tau)}{\partial S}$ is continuous at the boundary and we have established equation (1.471).

The zero-time-decay condition is shown by simply considering the total change in the American option value *along* the boundary $S = S^*(\tau)$ as the calendar time (or time to maturity) changes and the boundary point moves accordingly. Along the boundary we have $V(S^*(\tau), \tau) = \phi(S^*(\tau))$, and differentiating both sides of this relation w.r.t. τ gives (Note: The analysis in terms of t is the same):

$$\frac{\partial V(S^*(\tau), \tau)}{\partial S}\frac{dS^*(\tau)}{d\tau} + \frac{\partial V(S^*(\tau), \tau)}{\partial \tau} = \phi'(S^*(\tau))\frac{dS^*(\tau)}{d\tau}. \tag{1.476}$$

Hence, using equation (1.471) gives $\frac{\partial V(S^*(\tau),\tau)}{\partial \tau} = 0$, and since the option is given by the time-independent payoff function everywhere else on the stopping domain, we have equation (1.472).

Delta hedging and continuous-time replication arguments apply to American options in the same way they apply to European options. Within the (no-exercise) continuation domain we therefore expect and require that the option price function satisfy the Black–Scholes PDE. The connection between the optimal stopping time formulation and the PDE approach can be shown as follows. Consider recurrence relation (1.460) with time step $\delta t > 0$ for any calendar time $t < T$,

$$V_t(S) = \max\left\{\phi(S), e^{-r\delta t}E_t\left[V_{t+\delta t}(S_{t+\delta t})|S_t = S\right]\right\}. \tag{1.477}$$

Assuming $V_t(S)$ is sufficiently smooth with continuous derivatives then to leading order $O(\delta t)$, we can Taylor-expand $V_{t+\delta t}(S_{t+\delta t})$ while using Itô's lemma. For a generally state- and time-dependent process obeying $\delta S_t = \mu(S_t, t)\delta t + \sigma(S_t, t)\delta W_t$, we have

$$\begin{aligned} V_t(S) = \max &\left\{\phi(S), (1 - r\delta t)E_t\left[V_t(S_t) + \left(\frac{\partial V_t(S_t)}{\partial t} + \mu(S_t, t)\frac{\partial V_t(S_t)}{\partial S_t}\right.\right.\right. \\ &\left.\left.\left. + \frac{1}{2}\sigma^2(S_t, t)\frac{\partial^2 V_t(S_t)}{\partial S_t^2}\right)\delta t + \sigma(S_t, t)\frac{\partial V_t(S_t)}{\partial S_t}\delta W_t \bigg| S_t = S\right]\right\} + O((\delta t)^2) \\ = \max &\left\{\phi(S), V_t(S) + \left[\frac{\partial V_t(S)}{\partial t} + \mathcal{L}_{BS}V_t(S)\right]\delta t\right\} + O((\delta t)^2). \end{aligned} \tag{1.478}$$

The second equation obtains by evaluating the conditional expectation (which sets $S_t = S$ and eliminates the δW_t term) and then collecting terms up to $O(\delta t)$. This expression has been written more compactly using the Black–Scholes differential operator (for general drift and volatility functions) defined by

$$\mathcal{L}_{BS}V \equiv \frac{1}{2}\sigma^2(S, t)\frac{\partial^2 V}{\partial S^2} + \mu(S, t)\frac{\partial V}{\partial S} - rV \equiv (\mathcal{L}_{S,t} - r)V. \tag{1.479}$$

For values of S in the continuation domain, the inequality $V_t(S) > \phi(S)$ is satisfied, and hence, from equation (1.478) we must have the Black–Scholes PDE:

$$\frac{\partial V_t(S)}{\partial t} + \mathcal{L}_{BS} V_t(S) = 0, \qquad \text{for all } S \notin \mathcal{D}_\tau. \tag{1.480}$$

By specializing to the geometric Brownian motion model, then, $\mu(S, t) = (r - q)S$, $\sigma(S, t) = \sigma S$ and the Black–Scholes PDE is

$$\frac{\partial V}{\partial \tau} = \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} + (r - q)S\frac{\partial V}{\partial S} - rV \equiv \mathcal{L}_{BS} V, \qquad \text{for all } S \notin \mathcal{D}_\tau. \tag{1.481}$$

Thanks to the time-homogeneous property of the solution in this case, we have a PDE in terms of the time-to-maturity variable, $V = V(S, \tau)$, which will be convenient in subsequent discussions.

1.14.2 Perpetual American Options

An option with infinite time to maturity is called a *perpetual option*. Here we consider perpetual American calls and puts. These options are instructive since simple analytic solutions exist. Moreover, since the exercise boundary $S^*(\tau)$ is a monotonic function of time to maturity τ (i.e., increasing for a dividend-paying American call and decreasing for an American put), the perpetual option price provides us with the asymptotic limit $\lim_{\tau\to\infty} S^*(\tau) \equiv S^*$ of the exercise boundary for times infinitely far from maturity. We again consider an asset price process S_t following geometric Brownian motion with constant interest rate r and continuous dividend yield at constant rate q. Since a perpetual option has infinite time to maturity, its value does not depend on the passage of time; i.e., the price function is independent of time. Hence the time derivative of the price function is zero and the Black–Scholes partial differential equation (1.481) for the price of a perpetual option reduces to a time-independent ordinary differential equation (ODE).

We first consider the case of a perpetual put struck at K. The price function denoted by *P(S)* must satisfy the ODE

$$\frac{1}{2}\sigma^2 S^2 \frac{d^2 P}{dS^2} + (r - q)S\frac{dP}{dS} - rP = 0 \tag{1.482}$$

for values away from the exercise boundary, $S^* < S < \infty$. The optimal exercise price S^* is therefore the asset price at which the perpetual American put should be exercised. Since the value of the perpetual put must be equal to the intrinsic value at all values of $S \le S^*$ and $S^* < K$, (see Figure 1.6) the boundary conditions on *P(S)* are

$$\lim_{S\to\infty} P(S) = 0, \quad P(S^*) = K - S^*. \tag{1.483}$$

S^* is yet unknown but uniquely determined once *P(S)* is obtained in terms of S^* as described just next. Equation (1.482) is an ODE of the Cauchy–Euler (equidimensional) type and therefore has the general solution

$$P(S) = a_+ S^{\gamma_+} + a_- S^{\gamma_-}, \tag{1.484}$$

where $a_\pm$ are arbitrary constants and $\gamma_\pm$ are roots of the auxiliary quadratic equation

$$\tfrac{\sigma^2}{2}\gamma^2 + (r - q - \tfrac{\sigma^2}{2})\gamma - r = 0. \tag{1.485}$$

Solving for the roots gives

$$\gamma_{\pm} = \frac{-(r-q-\frac{\sigma^2}{2}) \pm \sqrt{(r-q-\frac{\sigma^2}{2})^2 + 2\sigma^2 r}}{\sigma^2}. \tag{1.486}$$

Assuming positive interest rate r, then γ_- and γ_+ are negative and positive roots, respectively. To satisfy the first condition at infinity in equation (1.483) we must have $a_+ = 0$. By satisfying the second boundary condition in equation (1.483), $a_- = (K - S^*)/(S^*)^{\gamma_-}$, we obtain the price function in the form

$$P(S) = (K - S^*)\left(\frac{S}{S^*}\right)^{\gamma_-}, \qquad S \geq S^*. \tag{1.487}$$

The exercise boundary value S^* can now be determined as the optimal value that maximizes the price *P(S)* for all possible choices of S^*. The derivative w.r.t. the parameter S^* of this price function gives

$$\frac{\partial P}{\partial S^*} = -\left(\frac{S}{S^*}\right)^{\gamma_-}\left(1 + \frac{K - S^*}{S^*}\gamma_-\right). \tag{1.488}$$

Setting this derivative to zero yields the extremum

$$S^* = \frac{K\gamma_-}{\gamma_- - 1}. \tag{1.489}$$

Computing the second derivative at this extremum gives $\frac{\partial^2 P}{\partial S^{*2}} = \frac{K\gamma_-}{(S^*)^2}(\frac{S}{S^*})^{\gamma_-} < 0$. Hence S^* in equation (1.489) is a maximum, and inserting its value into equation (1.487) gives the price of the perpetual American put in the equivalent forms

$$\begin{aligned} P(S) &= \frac{K}{1-\gamma_-}\left(\frac{\gamma_- - 1}{\gamma_-}\right)^{\gamma_-}\left(\frac{S}{K}\right)^{\gamma_-} \\ &= -\frac{S^*}{\gamma_-}\left(\frac{S}{S^*}\right)^{\gamma_-}, \end{aligned} \tag{1.490}$$

for $S \geq S^*$. This solution is easily shown to satisfy the required smooth pasting condition

$$\left.\frac{dP}{dS}\right|_{S=S^*} = -1. \tag{1.491}$$

Next we consider the perpetual American call struck at K. As in the case of the put, the price function now denoted by *C(S)* also satisfies equation (1.482), but for values $0 < S < S^*$. The optimal value S^* is therefore the asset price at which the call should be exercised. The value *C(S)* must be given by the intrinsic value of the call pay-off for values on the boundary $S \geq S^*$, where $S^* > K$; hence the boundary conditions are

$$\lim_{S\to 0} C(S) = 0, \quad C(S^*) = S^* - K. \tag{1.492}$$

The general solution is again given by equations (1.484) and (1.486). However, by satisfying the boundary conditions in equation (1.492) we now instead have $a_- = 0$ and $a_+ = (S^* - K)/(S^*)^{\gamma_+}$, giving

$$C(S) = (S^* - K)\left(\frac{S}{S^*}\right)^{\gamma_+}, \qquad 0 < S < S^*. \tag{1.493}$$

Using the same procedure as for the put, the optimal exercise boundary is determined by finding the maximum of *C(S)* w.r.t. S^*, giving

$$S^* = \frac{K\gamma_+}{\gamma_+ - 1}. \tag{1.494}$$

Using S^* from equation (1.494) in equation (1.493) gives the price of the perpetual American call, written equivalently in terms of K or S^*:

$$\begin{aligned} C(S) &= \frac{K}{\gamma_+ - 1}\left(\frac{\gamma_+ - 1}{\gamma_+}\right)^{\gamma_+}\left(\frac{S}{K}\right)^{\gamma_+} \\ &= \frac{S^*}{\gamma_+}\left(\frac{S}{S^*}\right)^{\gamma_+}. \end{aligned} \tag{1.495}$$

This satisfies the required smooth pasting condition

$$\left.\frac{dC}{dS}\right|_{S=S^*} = 1. \tag{1.496}$$

It is instructive to examine what happens to the exercise boundary in the two separate limiting cases: (i) zero interest rate $r = 0$ and (ii) zero dividend yield $q = 0$. In case (i) we have from equation (1.486) that $\gamma_- = 0$ (assuming $q \geq -\sigma^2/2$, which is the case if $q \geq 0$). From equation (1.489) we see that $S^* = 0$; hence, for zero interest rate the perpetual put is never exercised early. This is consistent with the property of an American put for r = 0 and for any finite time to maturity, as shown in the next section. From a financial standpoint, there is no time value gained from an early pay-off with zero interest. For case (ii): Equation (1.486) gives $\gamma_+ = 1$ (assuming $r \geq -\sigma^2/2$, which is the case for $r \geq 0$). Moreover, $\gamma_+ \to 1^+$ as $q \to 0^+$ and from equation (1.494) we have $S^* \to \infty$. Hence in the limit of zero dividend yield the perpetual call is never exercised early, irrespective of the interest rate. This feature is also consistent with the plain American call of finite maturity, as shown in the next section.

1.14.3 *Properties of the Early-Exercise Boundary*

The perpetual American option formulas of the previous section already allowed us to determine the precise behavior of the optimal exercise boundary in the asymptotic limit of infinite time to expiry, i.e., as $\tau \to \infty$. To further complete the analysis of the boundary we now consider the opposite limit, of infinitesimally small positive time to maturity $\tau \to 0^+$. In particular, let us consider the case of the Amercian call struck at K with continuous dividend yield q and price function denoted by $C(S, K, \tau)$ at spot S. Since $C(S, K, \tau)$ is an increasing function of τ, for $\tau > 0$, the graph of the American call price (plotted as a function of S) with greater time to maturity τ_2 must lie above the graph of the price function for the corresponding call with time to maturity $\tau_1 < \tau_2$. Furthermore, the smooth pasting condition guarantees that the price functions join the intrinsic line at levels $S^*(\tau_1) - K$ and $S^*(\tau_2) - K$, respectively, giving $S^*(\tau_1) < S^*(\tau_2)$. Hence, we conclude that $S^*(\tau)$ is a continuously increasing function of positive τ. To put this in financial terms, an American call with greater time to maturity should be exercised deeper in the money to account for the loss of time value on the strike K. Due to the fact that one would never prematurely exercise at a spot value below the strike level (i.e., exercising for a nonpositive pay-off), the early-exercise boundary for an Amercian call must, in addition, satisfy the property $S^*(\tau) > K$ for all $\tau > 0$.

To determine the boundary in the limit $\tau \to 0^+$, note that the option value approaches the intrinsic value; i.e., at expiry it is exactly given by the payoff function $C(S, K, \tau = 0) = S - K$ for values on the exercise boundary. Inserting this function into the right-hand side of equation (1.481) and taking derivatives gives

$$\frac{\partial C(S, K, 0^+)}{\partial \tau} = rK - qS \tag{1.497}$$

for $S > K$. Since the condition $\partial C(S, K, 0^+)/\partial\tau > 0$ ensures that the option is still alive (i.e., not yet exercised), the spot value S at which $\partial C(S, K, 0^+)/\partial\tau$ becomes negative and hence for which the call is exercised at an instant just before expiry is given by $S = \frac{r}{q}K$. This is the case, however, if the value $\frac{r}{q}K$ is in the interval $S > K$, that is, if $r > q > 0$. In this case, just prior to expiry the call is not yet exercised if the spot is in the region $K < S < \frac{r}{q}K$ but would be exercised if $S \geq \frac{r}{q}K$. Hence, $S^*(0^+) = \frac{r}{q}K$ for $r > q > 0$. In the other case, $r \leq q$, so $\frac{r}{q}K \leq K$. Yet $S > K$, so $S^*(0^+) = K$ for $r \leq q$. Note that the condition $S^*(0^+) > K$ is not possible in this case because this leads to a suboptimal early exercise, since the loss in dividends would have greater value than the interest earned over the infinitesimal time interval until expiry. Combining these arguments we arrive at the general limiting condition for the exercise boundary of an American call just prior to expiry:

$$\lim_{\tau \to 0^+} S^*(\tau) = \max(K, \frac{r}{q}K). \tag{1.498}$$

From this property we see that $S^*(0^+) \to \infty$ as $q \to 0$. Hence, for zero dividend yield the American call is never exercised early, which is consistent with the fact that the plain (nondividend) American call has exactly the same worth as the plain European call.

Similar arguments can also be employed in the case of the Amercian put struck at K with continuous dividend yield q. At expiry the put has value $P(S, K, \tau = 0) = K - S$ for values on the exercise boundary. We leave it as an exercise for the reader to show that the exercise boundary of an American put just prior to expiry is given by

$$\lim_{\tau \to 0^+} S^*(\tau) = \min(K, \frac{r}{q}K). \tag{1.499}$$

For $r = 0$ we therefore have $S^*(0^+) = 0$, irrespective of the value of q. Since $S^*(\tau)$ is a decreasing function of τ, we conclude that the early-exercise boundary is always at zero, meaning that the American put with zero interest rate is never exercised before maturity. This is consistent with the conclusion we arrived at earlier, where we considered the perpetual American put. For $q \leq r$ we observe that the early-exercise boundary just before expiry is at the strike, $S^*(0^+) = K$. A special case of this is the vanilla American put, i.e., when $r > 0$ and $q = 0$. Figure 1.7 gives an illustration of typical early-exercise boundaries for a call and put. Given a time to maturity of T at contract inception, we see that the American call with nonzero dividend is not yet exercised (i.e., is still alive) on the domain of points (S, τ) below the exercise curve: $S \in [0, S^*(\tau))$ and $\tau \in (0, T]$. In contrast, the American put is kept alive above the exercise curve: $S \in (S^*(\tau), \infty)$ and $\tau \in (0, T]$.

1.14.4 The Partial Differential Equation and Integral Equation Formulation

The problem of pricing an American option can be formulated as an initial-value partial differential equation (PDE) with a time-dependent free boundary. The early-exercise boundary is an unknown function of time, which must also be determined as part of the solution. In particular, let $V(S, \tau)$ represent the pricing function of an American option with spot S

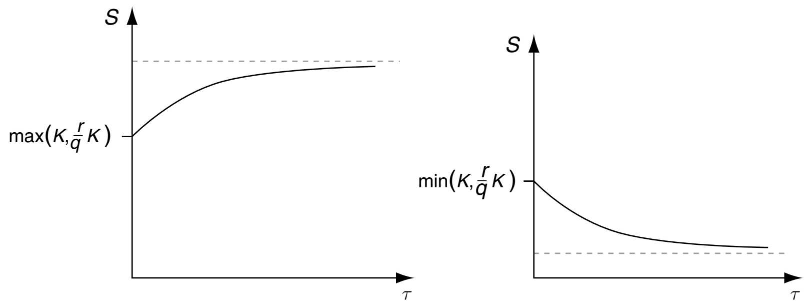

FIGURE 1.7 Early-exercise smooth boundary curves $S = S^*(\tau)$ for the American call (left), with $q > 0$, and put (right), with values depicted just before expiry $\tau \to 0^+$. In the limit of infinite time to expiry, the curves approach the horizontal asymptotes at $S = S^*$, where S^* is given by equation (1.494) or equation (1.489) for the call or put, respectively.

and time to maturity τ, $0 \le \tau \le T$, and having payoff or intrinsic function $V(S, 0) = \phi(S)$. Here we assume the pay-off is time independent, although the formulation also extends to the case of a known time-dependent payoff function. For given τ, the solution domain is divisible into a union of two regions: (1) a continuation region $(S, \tau) \in \mathcal{D}'_\tau \times [0, T]$, for which the option is still alive or not exercised, and (2) a stopping region $(S, \tau) \in \mathcal{D}_\tau \times [0, T]$, where $\mathcal{D}_\tau$ is the complement of $\mathcal{D}'_\tau$ within $\mathbb{R}_+$, for which the American option is already exercised. The domains depend on τ. As seen in the previous section, in the case of the American call, $\phi(S) = S - K$ on $\mathcal{D}_\tau = [S^*(\tau), \infty)$ (and $\mathcal{D}'_\tau = (0, S^*(\tau))$, while for the put, $\phi(S) = K - S$ on $\mathcal{D}_\tau = (0, S^*(\tau)]$ (and $\mathcal{D}'_\tau = (S^*(\tau), \infty)$. Assuming the underlying asset follows equation (1.381), equation (1.481) holds for $S \in \mathcal{D}'_\tau$. In contrast, the homogeneous Black–Scholes PDE does not hold on the domain of the early-exercise boundary, where the American option is given by the time-independent payoff function $V(S, \tau) = \phi(S)$. Since $\frac{\partial \phi(S)}{\partial \tau} = 0$, the solution on $\mathcal{D}_\tau$ satisfies $\frac{\partial V}{\partial \tau} = 0$. Combining regions and assuming the pay-off is twice differentiable gives a *nonhomogeneous* Black–Scholes PDE:

$$\frac{\partial V(S, \tau)}{\partial \tau} = \mathcal{L}_{BS} V(S, \tau) + f(S, \tau), \tag{1.500}$$

with (source) function

$$f(S, \tau) = \begin{cases} 0, & S \in \mathcal{D}'_\tau \\ \\ -\mathcal{L}_{BS}\phi(S), & S \in \mathcal{D}_\tau, \end{cases} \tag{1.501}$$

where $\mathcal{L}_{BS}$ is the Black–Scholes differential operator. For geometric Brownian motion, $\mathcal{L}_{BS}$ is defined by equation (1.481). Given the function $f(S, \tau)$, whose time dependence is determined in terms of the free boundary, the solution to equation (1.500), subject to the initial condition $V(S, \tau = 0) = \phi(S)$ and boundary conditions $V(S = 0, \tau) = \phi(0)$, $V(S = \infty, \tau) = \phi(\infty)$, can be obtained in terms of the solution to the corresponding homogeneous Black–Scholes PDE. Recall from previous discussions that the transition probability density function $p(S', S; \tau)$ solves the forward Kolmogorov PDE in the S' variable and the backward PDE in the spot variable S with zero boundary conditions at $S = 0, \infty$ for all $\tau > 0$. As already mentioned, for process (1.381) p is just the lognormal density given by equation (1.382). We also know that $e^{-r\tau} p$ solves the homogeneous Black–Scholes PDE. Combining these facts and applying

Laplace transforms, one arrives at the well-known Duhamel's solution to equation (1.500) in the form

$$
\begin{aligned}
V(S,\tau) &= e^{-r\tau}\int_0^{\infty} p(S',S;\tau)\phi(S')dS' \\
&\quad + \int_0^{\tau} e^{-r\tau'}\left[\int_0^{\infty} p(S',S;\tau')f(S',\tau-\tau')dS'\right]d\tau' \\
&\equiv V_E(S,\tau)+V^e(S,\tau). \qquad (1.502)
\end{aligned}
$$

One can readily verify that this solves equation (1.500), even for the more general case of state-dependent models (see Problem 1). An important aspect of this result is that the American option value $V(S,\tau)$ is expressible as a sum of two components. The first term is simply the European option value V_E, as given by the discounted risk-neutral expectation of the pay-off. Hence the second term, denoted by $V^e(S,\tau)$, must represent the early-exercise premium, which gives the holder the additional liberty of early exercise.

Assuming geometric Brownian motion for the underlying asset, equations (1.500) and (1.501) for the American call and put specialize to

$$
\frac{\partial C}{\partial \tau}-\frac{\sigma^2S^2}{2}\frac{\partial^2 C}{\partial S^2}-(r-q)S\frac{\partial C}{\partial S}+rC=\begin{cases}0, & S<S^*(\tau)\\ qS-rK, & S\geq S^*(\tau)\end{cases} \qquad (1.503)
$$

and

$$
\frac{\partial P}{\partial \tau}-\frac{\sigma^2S^2}{2}\frac{\partial^2 P}{\partial S^2}-(r-q)S\frac{\partial P}{\partial S}+rP=\begin{cases}rK-qS, & S\leq S^*(\tau)\\ 0, & S>S^*(\tau)\end{cases}, \qquad (1.504)
$$

respectively. Here we used $\mathcal{L}_{BS}(S-K)=rK-qS$, and $S^*(\tau)$ denotes the early-exercise boundary for the respective call and put with strike K. The right-hand sides of these nonhomogeneous PDEs are nonzero only within the respective stopping regions. Using equation (1.502), the solutions to equations (1.503) and (1.504) for the American call and put price are given by

$$
C(S,K,\tau)=C_E(S,K,\tau)+C^e(S,K,\tau) \qquad (1.505)
$$

and

$$
P(S,K,\tau)=P_E(S,K,\tau)+P^e(S,K,\tau), \qquad (1.506)
$$

where the respective early-exercise premiums take on the integral forms

$$
C^e(S,K,\tau)=\int_0^{\tau} e^{-r\tau'}\left[\int_{S^*(\tau-\tau')}^{\infty} p(S',S;\tau')(qS'-rK)dS'\right]d\tau' \qquad (1.507)
$$

and

$$
P^e(S,K,\tau)=\int_0^{\tau} e^{-r\tau'}\left[\int_0^{S^*(\tau-\tau')} p(S',S;\tau')(rK-qS')dS'\right]d\tau'. \qquad (1.508)
$$

These premiums can also be recast as

$$
C^e(S,K,\tau)=\int_0^{\tau} e^{-r\tau'}E_0\left[(qS_{\tau'}-rK)\mathbf{1}_{\{S_{\tau'}\geq S^*(\tau-\tau')\}}\right]d\tau' \qquad (1.509)
$$

and

$$P^e(S,K,\tau) = \int_0^{\tau} e^{-r\tau'} E_0\big[(rK - qS_{\tau'})\mathbf{1}_{\{S_{\tau'} \le S^*(\tau-\tau')\}}\big]d\tau', \tag{1.510}$$

where E_0 denotes the current-time expectation, conditional on asset paths starting at $S_0 = S$ under the risk-neutral measure with density $p(S_{\tau'}, S; \tau')$. The time integral is over all intermediate times to maturity, and the indicator functions ensure that all asset paths fall within the early-exercise region. The properties of the early-exercise boundaries established in the previous section guarantee that the early-exercise premiums are nonnegative. For a dividend-paying call, equation (1.498), together with the indicator function condition, leads to $S_{\tau'} \ge \max(\frac{r}{q}K, K) \ge \frac{r}{q}K$; hence $qS_{\tau'} - rK \ge 0$ and C^e is positive. A similar analysis follows for the put premium. The exercise premiums hence involve a continuous stream of discounted expected cash flows, beginning from contract inception until maturity. This lends itself to an interesting financial interpretation, as follows. Consider the case of the American put (a similar argument applies to the dividend-paying call) and an infinitesimal intermediate time interval $[\tau', \tau' + d\tau']$. Then from the holder's perspective the option should be optimally exercised if the asset price, given by $S_{\tau'}$ at time τ', attains the stopping region (i.e., reaches the early-exercise boundary with $S_{\tau'} \le S^*(\tau - \tau')$ and $\tau - \tau'$ as the remaining time to maturity). Assuming that the holder is instead forced to keep the American put alive until expiry, the holder would have to be fairly compensated for the loss due to the delay in exercising during the time interval $d\tau'$. The value of this compensation is the difference between the interest on K dollars and the dividend earned on the asset value $S_{\tau'}$, continuously compounded over time $d\tau'$. This cash flow is an amount $(rK - qS_{\tau'})d\tau'$, and corresponds to the early-exercise gain if the holder in fact had the privilege to optimally exercise. Allowing for all possible asset price scenarios from S to $S_{\tau'}$ that attain the boundary gives rise to the expectation integral under the risk-neutral density for all intermediate times $0 \le \tau' \le \tau$. Summing up all of these infinitesimal cash flows and discounting their values to present time by an amount $e^{-r\tau'}$ gives the time integral, as in equation (1.508) or (1.510). We conclude that the early-exercise premium has an equivalent and alternative interpretation as a delay-exercise compensation.

The foregoing integral representations for the American call and put price can also be applied to cases where the volatility of the asset price process S_t is considered generally state dependent. In order to implement the integral formulas, we need to be able to compute the transition density function p, either analytically or numerically. Moreover, the integrals can only be computed after having determined the early-exercise boundary $S^*(\tau')$ for $0 \le \tau' \le \tau$. For the geometric Brownian motion model (with constants r,q,σ), p is given by the lognormal density, and the foregoing double integrals readily simplify to single time integrals in terms of standard cumulative normal functions. In particular, one readily derives explicit integral representations for the price of the American call and put (see Problem 2):

$$\begin{aligned} C(S,K,\tau) &= Se^{-q\tau}N(d_+) - Ke^{-r\tau}N(d_-) \\ &\quad + \int_0^{\tau} \big[qSe^{-q(\tau-\tau')}N(d_+^*(\tau')) - rKe^{-r(\tau-\tau')}N(d_-^*(\tau'))\big]d\tau', \end{aligned} \tag{1.511}$$

$$\begin{aligned} P(S,K,\tau) &= Ke^{-r\tau}N(-d_-) - Se^{-q\tau}N(-d_+) \\ &\quad + \int_0^{\tau} \big[rKe^{-r(\tau-\tau')}N(-d_-^*(\tau')) - qSe^{-q(\tau-\tau')}N(-d_+^*(\tau'))\big]d\tau', \end{aligned} \tag{1.512}$$

where

$$d_{\pm} = \frac{\log \frac{S}{K} + \left(r - q \pm \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}, \tag{1.513}$$

$$d^*_{\pm}(\tau') = \frac{\log \frac{S}{S^*(\tau')} + \left(r - q \pm \frac{1}{2}\sigma^2\right)(\tau - \tau')}{\sigma\sqrt{\tau - \tau'}}. \tag{1.514}$$

These integral representations are valid for $S \in (0, \infty)$, $\tau \geq 0$. By setting $S = S^*(\tau)$ and applying the respective boundary conditions, $C(S^*(\tau), K, \tau) = S^*(\tau) - K$ for the call and $P(S^*(\tau), K, \tau) = K - S^*(\tau)$ for the put, equations (1.511) and (1.512) give rise to integral equations for the early-exercise boundary. For the call,

$$\begin{aligned} S^*(\tau) - K &= Se^{-q\tau}N(\tilde{d}_+) - Ke^{-r\tau}N(\tilde{d}_-) \\ &\quad + \int_0^\tau \left[qSe^{-q(\tau-\tau')}N(\tilde{d}^*_+(\tau')) - rKe^{-r(\tau-\tau')}N(\tilde{d}^*_-(\tau'))\right]d\tau'. \end{aligned} \tag{1.515}$$

and separately for the put,

$$\begin{aligned} K - S^*(\tau) &= Ke^{-r\tau}N(-\tilde{d}_-) - Se^{-q\tau}N(-\tilde{d}_+) \\ &\quad + \int_0^\tau \left[rKe^{-r(\tau-\tau')}N(-\tilde{d}^*_-(\tau')) - qSe^{-q(\tau-\tau')}N(-\tilde{d}^*_+(\tau'))\right]d\tau', \end{aligned} \tag{1.516}$$

where

$$\tilde{d}_{\pm} = \frac{\log \frac{S^*(\tau)}{K} + \left(r - q \pm \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}, \tag{1.517}$$

$$\tilde{d}^*_{\pm}(\tau') = \frac{\log \frac{S^*(\tau)}{S^*(\tau')} + \left(r - q \pm \frac{1}{2}\sigma^2\right)(\tau - \tau')}{\sigma\sqrt{\tau - \tau'}}. \tag{1.518}$$

Note that equations (1.515) and (1.516) involve a variable upper integration limit and the integrands are nonlinear functions of $S^*(\tau)$, $S^*(\tau')$, τ and τ'. From the theory of integral equations, equations (1.515) and (1.516) are known as *nonlinear Volterra integral equations*. Note that the solution $S^*(\tau)$, at time to maturity τ, is dependent on the solution $S^*(\tau')$ from zero time to maturity $\tau' = 0$ up to $\tau' = \tau$. Although equations (1.515) and (1.516) are not analytically tractable, simple and efficient algorithms can be employed to solve for $S^*(\tau)$ numerically. For detailed descriptions on various numerical algorithms for solving these types of integral equations, see, for example, [DM88]. A typical procedure divides the solution domain into a regular mesh: $\tau_0 = 0$, $\tau_i = ih$, $i = 1, \ldots, n$, with n steps spaced as $h = \tau/n$. By approximating the time integral via a quadrature rule (e.g., the trapezoidal rule), one obtains a system of algebraic equations in the values $S^*(\tau_i)$, which can be iteratively solved starting from the known value $S^*(\tau_0) = S^*(\tau = 0^+)$ at zero time to maturity. Alternatively, popular Runge–Kutta methods usually used for solving initial-value nonlinear ODEs can be also adapted to these integral equations. Once the early-exercise boundary is determined, the integral in equation (1.511) or (1.512) for the respective call or put can be computed. In particular, a quadrature rule that makes use of the computed points $S^*(\tau_i)$ can be implemented. Accurate approximations to the early-exercise boundary are obtained by choosing the number n of points to be sufficiently large.

Problems

Problem 1. Consider the state-dependent model $dS_t = \mu(S_t)dt + \sigma(S_t)dW_t$. Assuming $f(S, \tau)$ is differentiable w.r.t. τ, show that equation (1.502) satisfies equation (1.500) for the appropriate operator $\mathcal{L}_{BS}$. Hint: Since V_E satisfies the homogeneous Black–Scholes PDE, from superposition one need only show that V^e satisfies equation (1.500). Use the property of interchanging order of differentiation and integration, integration by parts, and the fact that $e^{-r\tau}p$ satisfies the homogeneous Black–Scholes PDE with initial condition $p(S', S; 0) = \delta(S' - S)$. Provide an extension to equation (1.502), if possible, for the more general case of explicitly time-dependent drift and volatility.

Problem 2. (a) By employing similar manipulations as were used to obtain the standard Black–Scholes formulas in Section 1.6, derive equations (1.511) and (1.512) from equations (1.507) and (1.508). (b) Show that the pricing formulas for the American call and put in equations (1.511) and (1.512) satisfy the required boundary conditions at $S = 0$ and $S = \infty$.

Problem 3. Find an analytical formula for the price as well as the early-exercise boundaries of a perpetual American butterfly option with payoff function $\delta_\epsilon(S - K)$ given by equation (1.228) of Section 1.8. Assume $K - \epsilon > 0$ and that the underlying asset price obeys geometric Brownian motion with constant interest rate r and continuous dividend yield q.

Problem 4. Using equations (1.511) and (1.512), derive integral representations for the delta, gamma, and vega sensitivities of the American call and put.

Problem 5. Let $V(S, \tau)$ and $V_E(S, \tau)$ denote the American and European option values, respectively, with spot S, time to maturity τ, and payoff function $\phi(S)$. Assume a constant interest rate r and continuous dividend yield q under the geometric Brownian motion model for the process S_t. Prove the equivalence of these two statements:

(i) $V(S, \tau) > V_E(S, \tau)$ for all $S > 0$, $\tau > 0$.
(ii) $\phi(S) > e^{-r\tau}\phi(e^{(r-q)\tau}S)$ for some point (S, τ). Explain why American options on futures have a nonzero early-exercise premium.

Problem 6. Consider a Bermudan put option with strike K at maturity T with only a single intermediate early-exercise date $T_1 \in [0, T]$. Assume the underlying stock price obeys equation (1.381) within the risk-neutral measure, and let $P(S_t, K, T - t)$ denote the option value at calendar time t with spot S_t. Find an analytically closed-form expression for the present-time $t = 0$ price $P(S_0, K, T)$. Hint: This problem is very closely related to the valuation of a compound option discussed at the end of Section 1.12. In particular, proceed as follows. From backward recurrence show that

$$P(S_0, K, T) = e^{-rT_1} E_0\big[P(S_{T_1}, K, T - T_1)\big], \tag{1.519}$$

with

$$P(S_{T_1}, K, T - T_1) = \begin{cases} P_E(S_{T_1}, K, T - T_1), & S_{T_1} > S^*_{T_1} \\ \\ K - S_{T_1}, & S_{T_1} \le S^*_{T_1}, \end{cases} \tag{1.520}$$

where P_E is the European put price function, $E_0[\,]$ is the risk-neutral expectation at time 0, and the critical value $S^*_{T_1}$ for the early-exercise boundary at calendar time T_1 solves

$$P_E(S^*_{T_1}, K, T - T_1) = K - S^*_{T_1}.$$

Compute this expectation as a sum of two integrals, one over the domain $S_{T_1} > S^*_{T_1}$ and the other over $0 < S_{T_1} \leq S^*_{T_1}$ while using equations (1.382) and (1.385) to finally arrive at the expression for $P(S_0, K, T)$ in terms of univariate and bivariate cumulative normal functions. Show whether $S^*_{T_1}$ is a strictly increasing or decreasing function of the volatility σ, and explain your answer. What is this functional dependency for the case of a Bermudan call? Explain.