# FINANCIAL MATHEMATICS 

## I-Liang Chern

Department of Mathematics
National Taiwan University

## Contents

1 Introduction ..... 1
1.1 Financial Markets ..... 1
1.2 Financial Derivatives ..... 1
1.3 Examples ..... 2
1.4 Payoff functions ..... 3
1.5 Other kinds of options ..... 5
1.6 Types of traders ..... 5
1.7 Basic assumption ..... 6
2 Asset Price Model ..... 7
2.1 Efficient market hypothesis ..... 7
2.2 The asset price model ..... 7
2.2.1 The discrete asset price model ..... 7
2.2.2 The continuous asset price model ..... 8
2.3 Random walk ..... 8
2.4 The solution of the discrete asset price model ..... 10
2.5 The Brownian motion ..... 10
2.5.1 The definition of a Brownian motion ..... 10
2.5.2 The Brownian as a limit of random walk ..... 11
2.5.3 Properties of Brownian motion ..... 12
2.6 Itô's formula ..... 12
2.7 The solution of the continuous asset price model ..... 13
2.8 Continuous model as a limit of the discrete model ..... 14
2.9 Simulation of asset price model ..... 16
3 Black-Scholes Analysis ..... 19
3.1 The hypothesis of no-arbitrage-opportunities ..... 19
3.2 Basic properties of option prices ..... 20
3.2.1 The relation between payoff and options ..... 20
3.2.2 European options ..... 20
3.2.3 Basic properties of American options ..... 22
3.2.4 Dividend Case ..... 24
3.3 The Black-Scholes Equation ..... 26
3.3.1 Black-Scholes Equation ..... 26
3.3.2 Boundary and Final condition for European options ..... 27
3.4 Exact solution for the B-S equation for European options ..... 28
3.4.1 Reduction to parabolic equation with constant coefficients ..... 28
3.4.2 Further reduction ..... 29
3.4.3 Black-Scholes formula ..... 30
3.4.4 Special cases ..... 31
3.5 Risk Neutrality ..... 33
3.6 The delta hedging ..... 33
3.6.1 Time-Dependent $r, \sigma, \mu$ ..... 35
3.7 Trading strategy involving options ..... 36
3.7.1 Strategies involving a single option and stock ..... 37
3.7.2 Bull spreads ..... 38
3.7.3 Bear spreads ..... 39
3.7.4 Butterfly spread ..... 39
4 Variations on Black-Scholes models ..... 41
4.1 Options on dividend-paying assets ..... 41
4.1.1 Constant dividend yield ..... 41
4.1.2 Discrete dividend payments ..... 43
4.2 Warrants ..... 43
4.3 Futures and futures options ..... 44
4.3.1 Forward contracts ..... 44
4.3.2 Futures ..... 45
4.3.3 Futures options ..... 46
4.3.4 Black-Scholes analysis on futures options ..... 47
5 Numerical Methods ..... 51
5.1 Monte Carlo method ..... 51
5.2 Binomial Methods ..... 51
5.2.1 Binomial method for asset price model ..... 52
5.2.2 Binomial method for option ..... 52
5.3 Finite difference methods (for the modified B-S eq.) ..... 53
5.3.1 Discretization methods ..... 54
5.3.2 Binomial method is a forward Euler finite difference method ..... 55
5.3.3 Stability ..... 55
5.3.4 Convergence ..... 60
5.3.5 Boundary condition ..... 61
5.4 Converting the B-S equation to finite domain ..... 62
5.5 Fast algorithms for solving linear systems ..... 63
5.5.1 Direct methods ..... 64
5.5.2 Iterative methods ..... 66
6 American Option ..... 69
6.1 Introduction ..... 69
6.2 American options as a free boundary value problem ..... 70
6.2.1 American put option ..... 70
6.2.2 American call option on a dividend-paying asset ..... 72
6.3 American option as a linear complementary problem ..... 72
6.4 Numerical Methods ..... 74
6.4.1 Projective method for American put ..... 74
6.4.2 Projective method for American call ..... 75
6.4.3 Implicit method ..... 76
6.5 Converting American option to a fixed domain problem ..... 76
6.5.1 American call option with dividend paying asset ..... 76
6.5.2 American put option ..... 78
7 Exotic Options ..... 79
7.1 Binaries ..... 79
7.2 Compounds ..... 80
7.3 Chooser options ..... 80
7.4 Barrier option ..... 81
7.4.1 down-and-out call(knockout) ..... 81
7.4.2 down-and-in(knock-in) option ..... 82
7.5 Asian options and lookback options ..... 83
8 Path-Dependent Options ..... 85
8.1 Introduction ..... 85
8.2 General Method ..... 85
8.3 Average strike options ..... 86
8.3.1 European calls ..... 86
8.3.2 American call options ..... 87
8.3.3 Put-call parity for average strike option ..... 87
8.4 Lookback Option ..... 88
8.4.1 A lookback put with European exercise feature ..... 89
8.4.2 Lookback put option with American exercise feature ..... 90
9 Bonds and Interest Rate Derivatives ..... 91
9.1 Bond Models ..... 91
9.1.1 Deterministic bond model ..... 91
9.1.2 Stochastic bond model ..... 91
9.2 Interest models ..... 93
9.2.1 A functional approach for interest rate model ..... 93
9.3 Convertible Bonds ..... 94
A Basic theory of stochastic calculus ..... 97
A. 1 Brownian motion ..... 97
A. 2 Stochastic integral ..... 99
A. 3 Stochastic differential equation ..... 100
A. 4 Diffusion process ..... 101

## Chapter 1

## Introduction

### 1.1 Financial Markets

A society improves its welfare through investment. The financial market provides a link between saving and investment. Savers can earn high returns from their saving and borrowers can execute their investment plans to earn future profits. In financial markets, assets are traded in. There are many kinds of financial markets:

- stock markets,
- bond markets,
- currency markets, foreign exchange markets,
- commodity markets (oil, wheat, gold),
- futures and options markets.

In futures or options, more complex contracts than simple buy/sell trades have been introduced. These are called financial derivatives.

### 1.2 Financial Derivatives

1. Forwards contract: A forward contract is an agreement which allows the holder of the contract to buy or sell a certain asset at or by a certain day at a certain price. Here,

- the certain day-maturity or expiration date,
- the certain price-delivery price,
- the person who write the contract (has the asset) is called in short position,
- the person who holds the contract is called in long position.

2. Futures (futures contracts): A future contract, like a forward contract, except,

- it is normally traded in an exchange;
- it has standard features (including contract size, quality, delivery arrangement, price quotes, daily price movement, position limit, etc.);
- it is a margin trading (certain minimal amount of money should be maintained in a margin account);
- clearinghouse.

3. Options: There are two kinds of options - call options and put options. A call (put) option is a contract between two parties, in which the holder has the right to buy (sell) and the writer has the obligation to sell (buy) an asset at certain time in the future at a certain price. The price is called the exercise price (or strike price). The holder is called in long position, while the writer is called in short position. The underlying assets of an option can be commodity, stocks, stock indices, foreign currencies, or future contracts.

There are two kinds of exercise features:

- European options : Options can only be exercised at the maturity date.
- American options: Options can be exercised any time up to the maturity date.


### 1.3 Examples

## Notation

- $t$ current time
- $T$ maturity date
- $S$ current asset price
- $S_{T}$ asset price at time T
- $E$ strike price
- $c$ premium, the price of call option
- $r$ bank interest rate

1. An investor buys 100 European call options on IBM stock with strike price $\$ 140$. Suppose

$$
\begin{aligned}
E & =140 \\
S_{t} & =138 \\
T & =2 \text { months } \\
c & =5 \text { (the price of one call option). }
\end{aligned}
$$

If at time $T, S_{T}>E$, then he should exercise this option. The payoff is $100 \times\left(S_{T}-\right.$ $E)=100 \times(146-140)=600$, The premium is $5 \times 100=500$. Hence, he earns $\$ 100$. If $S_{T} \leq T$, then he should not exercise his call contracts. The payoff is 0 .

The payoff function for a call option is $\Lambda=\max \left\{S_{T}-E, 0\right\}$. One needs to pay premium $\left(c_{t}\right)$ to buy the options. Thus the net profit from buying this call is

$$
\Lambda-c_{t} e^{r(T-t)}
$$

2. Suppose

$$
\begin{aligned}
\text { today is } t & =8 / 22 / 95, \\
\text { expiration is } T & =4 / 14 / 96, \\
\text { the strike price } E & =250
\end{aligned}
$$

for some stock. If $S_{T}=270$ at expiration, which is smaller than the strike price, we should exercise this call option, then buy the share for 250 , and sell it in the market immediately for 270 . The payoff $\Lambda=270-250=20$. If $S_{T}=230$, we should give up our option, and the payoff is 0 . Suppose the share take 230 or 270 with equal probability. Then the expected profit is

$$
\frac{1}{2} \times 0+\frac{1}{2} \times 20=10
$$

Ignoring the interest of bank, then a reasonable price for this call option should be 10. If $S_{T}=270$, then the net profit $=20-10=10$. This means that the profits is $100 \%$ (He paid 10 for the option). If $S_{T}=230$ the loss is 10 for the premium. The loss is also $100 \%$. On the other hand, if the investor had instead purchased the share for 250 at $t$, then the corresponding profit or loss at $T$ is $\pm 20$. Which is only $\pm 8 \%$ of the original investment. Thus, option is of high risk and with high return.

### 1.4 Payoff functions

At the expiration day, the payoff of a future or an option is the follows.

1. The payoff function of a future is

$$
\Lambda=S_{T}-E
$$




Payoff of a future, long position (left) and short position (right)
2. The payoff function of a call option is

$$
\Lambda=\max \left\{S_{T}-E, 0\right\}
$$



Payoff of a call, long position (left) and short position (right)
3. The payoff function of a put option is

$$
\Lambda=\max \left\{E-S_{T}, 0\right\}
$$




Payoff of a put, long position (left) and short position (right)
4. Below is a portion of a call option copied from the Financial Times.

$$
\begin{aligned}
\text { the current time } t & =\text { Feb } 3 \\
\text { the expiration } T & =\text { end of Feb, } \\
T-t & \approx 10 \text { days } \\
S_{t} & =2872
\end{aligned}
$$



The FT-SE index call option values versus exercise price.

### 1.5 Other kinds of options

- Barrier option: The option only exists when the underlying asset price is in some prescribed value before expiry.
- Asian option: It is a contract giving the holder the right to buy or sell an asset for its average price over some prescribed period.
- Look-back option: The payoff depends not only on the asset price at expiry but also its maximum or minimum over some period price to expiry. For example, $\Lambda=$ $\max \left\{J-S_{0}, 0\right\}, J=\max _{0 \leq \tau \leq T} S(\tau)$.


### 1.6 Types of traders

1. Speculators (high risk, high rewards)
2. Hedgers (to make the outcomes more certain)
3. Arbitrageurs (Working on more than one markets, p12, p13, p14, Hull).

### 1.7 Basic assumption

Arbitrage opportunities cannot last for long. Only small arbitrage opportunities are observed in financial markets. Our arguments concerning future prices and option prices will be based on the assumption that "there is no arbitrage opportunities".

## Chapter 2

## Asset Price Model

### 2.1 Efficient market hypothesis

The asset prices move randomly because of the following efficient market hypothesis:

1. The past history is fully reflected in the present price, which does not hold any future information. This means the future price of the asset only depends on its current value and does not depends on its value one month ago, or one year ago. If this were not true, technical analysis could make above-average return by interpreting chart of the past history of the asset price. This contradicts to the hypothesis of no arbitrage opportunities. In fact, there is very little evidence that they are able to do so.
2. Market reponds immediately to any new information about an asset.

### 2.2 The asset price model

We shall introduce a discrete model and a continuous model. We will show that the continuous model is the continuous limit of the discrete model.

### 2.2.1 The discrete asset price model

The time is discrete in this model. The time sequence is $n \Delta t, n \in \mathbb{N}$. Let us denote the asset price at time step $n$ by $S_{n}$. We model the asset price by

$$
\frac{S_{n+1}}{S_{n}}= \begin{cases}u & \text { with probability } p  \tag{2.1}\\ d & \text { with probability } 1-p .\end{cases}
$$

Here, $0<d<1<u$. The information we are looking for is the following transition probability $P\left(S_{n}=S \mid S_{0}\right)$, the probability that the asset price is $S$ at time step $n$ with initial price $S_{0}$. We shall find this transition probability later.

### 2.2.2 The continuous asset price model

Let us denote the asset price at time $t$ by $S(t)$. The meaningful quantity for the change of an asset price is its relative change

$$
\frac{d S}{S}
$$

which is called the return. The change $\frac{d S}{S}$ can be decomposed into two parts: one is deterministic, the other is random.

- Deterministic part: This can be modeled by

$$
\frac{d S}{S}=\mu d t
$$

Here, $\mu$ is a measure of the growth rate of the asset. We may think $\mu$ is a constant during the life of an option.

- Random part: this part is a random change in response to external effects, such as unexpected news. It is modeled by a Brownian motion

$$
\sigma d z
$$

the $\sigma$ is the order of fluctuations or the variance of the return and is called the volatility. The quantity $d z$ is sampled from a normal distribution which we shall discuss below.

The overall asset price model is then given by

$$
\begin{equation*}
\frac{d S}{S}=\mu d t+\sigma d z \tag{2.2}
\end{equation*}
$$

We shall look for the transition probability density function $\mathcal{P}\left(S(t)=S \mid S(0)=S_{0}\right)$. Or equivalently, the integral

$$
\int_{a}^{b} \mathcal{P}\left(S(t)=S \mid S(0)=S_{0}\right) d S
$$

is the probability that the asset price $S(t)$ lies in $(a, b)$ at time $t$ and is $S_{0}$ initially.

### 2.3 Random walk

To study the discrete asset price model, we study a simple model-the random walk in one dimension-first. Consider a particle moving randomly on a uniformly distributed grid points on the real lines. Suppose the grid points are located at $m \Delta x, m \in \mathbb{Z}$. In each time step, the particle moves to its left adjacent grid point or right adjacent grid point with equal probability. Suppose the particle is located at 0 initially. Let $Z_{n}$ denote the location of this particle at time step $n$. Let $w(m \Delta x, n \Delta t)$ denotes for the probability that the particle is
located at the $m \Delta x$ cell at the time $n \Delta t$. That is, $w(m \Delta x, n \Delta t)=P\left(Z_{n}=m \Delta x \mid Z_{0}=0\right)$. By our rule,

$$
Z_{n+1}-Z_{n}= \begin{cases}\Delta x & \text { with probability } \frac{1}{2} \\ -\Delta x & \text { with probability } \frac{1}{2}\end{cases}
$$

and

$$
\begin{equation*}
w(m \Delta x,(n+1) \Delta t)=\frac{1}{2} w((m-1) \Delta, n \Delta t)+\frac{1}{2} w((m+1) \Delta, n \Delta t) . \tag{2.3}
\end{equation*}
$$

Suppose in $n$ times, the particle moves $p$ times toward right and $n-p$ time toward left. Then

$$
m=p-(n-p)=2 p-n \text { or } p=\frac{1}{2}(n+m) .
$$

Notice that $m$ is even(odd), when $n$ is even(odd). There is a one-to-one correspondence between $\{p \mid 0 \leq p \leq n\}$ and $\{m \mid-n \leq m \leq n, m+n$ is even $\}$. Notice also that the number of choices in $n$ steps that the particle moves $p$ times toward right is $\binom{n}{p}:=\frac{n!}{(n-p)!p p}$. When $p=\frac{1}{2}(n+m)$, we have

$$
w(m \Delta x, n \Delta t)= \begin{cases}0, & \text { if } m+n \text { is odd } \\ \binom{n}{p}\left(\frac{1}{2}\right)^{n}, & \text { if } m+n \text { is even } .\end{cases}
$$

We may check that $w(m \Delta x, n \Delta t)$ is a probability density function. Namely,

1. $w(m \Delta x, n \Delta t) \geq 0$.
2. $\sum_{m} w(m \Delta x, n \Delta t)=1$.

Given any function $f(m)$, we define its expectation value at $n \Delta t$ by

$$
<f(m)>:=\sum_{m} f(m) w(m \Delta x, n \Delta t)
$$

The moments $<m^{k}>, k \in \mathbb{N}$ are particularly important. The first moment $<m>$ is called the mean, while the second moment of the variation from mean $<(m-<m>)^{2}>$ is called the variance. They can be found by computing $\left\langle p^{k}\right\rangle$, which in turn can be computed through the help of the following generating function:

$$
\begin{aligned}
G(u) & :=\sum_{p} u^{p}\left(\frac{1}{2}\right)^{n}\binom{n}{p} \\
& =\left(\frac{1+u}{2}\right)^{n} .
\end{aligned}
$$

Hence

$$
<p>=G^{\prime}(1)=\sum_{p} p\left(\frac{1}{2}\right)^{n}\binom{n}{p}=\frac{n}{2} .
$$

From $m=2 p-n$, we have

$$
<m\rangle=2\langle p\rangle-n=0 .
$$

To compute the second moment $<m^{2}>$, from $m=2 p-n$, we have

$$
<m^{2}>=4<p^{2}>-4 n<p>+n^{2} .
$$

With the help of the generating function,

$$
\begin{aligned}
G^{\prime \prime}(1) & =\sum_{p=0}^{n} p(p-1)\binom{n}{p}\left(\frac{1}{2}\right)^{n} \\
& =<p^{2}>-<p> \\
& =<p^{2}>-\frac{n}{2}
\end{aligned}
$$

On the other hand, from $G(u)=\left(\frac{1+u}{2}\right)^{n}$, we obtain $G^{\prime \prime}(1)=\frac{n(n-1)}{4}$. Hence, $\left\langle p^{2}\right\rangle=$ $\frac{n^{2}}{4}+\frac{n}{4}$ and

$$
<m^{2}>=4<p^{2}>-4 n<p>+n^{2}=n .
$$

The mean of this random walk is $<m>=0$, while its variance is $<(m-<m>)^{2}>=n$.

## Exercise

1. Find the transition probability, mean and variance for the case

$$
Z_{n+1}-Z_{n}= \begin{cases}\Delta x & \text { with probability } p \\ -\Delta x & \text { with probability } 1-p\end{cases}
$$

2. One can also find the transition probability $w$ by solving the difference equation (2.3).

### 2.4 The solution of the discrete asset price model

Let us consider the case

$$
\frac{S_{n+1}}{S_{n}}= \begin{cases}u & \text { with probability } \frac{1}{2} \\ d & \text { with probability } \frac{1}{2}\end{cases}
$$

for simplicity. In $n$ movements of the asset price, if the price goes up $p$ times, then the price at time step $n \Delta t$ is $S_{n}=S_{0} u^{p} d^{n-p}$. Since there are $\binom{n}{p}$ such choices, we then obtain the transition probability of the asset:

$$
P\left(S_{n}=S \mid S_{0}\right)= \begin{cases}\binom{n}{p}\left(\frac{1}{2}\right)^{n} & \text { if } S=S_{0} u^{p} d^{n-p}  \tag{2.4}\\ 0 & \text { otherwise }\end{cases}
$$

### 2.5 The Brownian motion

### 2.5.1 The definition of a Brownian motion

The definition of the (standard) Brownian motion $z(t)$ is the following:

1. $\forall t, z(t)$ is a random variable.
2. The increment $z(t+s)-z(t), z(t)-z(t-u), u>0, s>0$ are independent.
3. $z(t)$ is continuous in $t$.
4. $\forall s>0, z_{t+s}-z_{t}$ is normally distributed with mean zero and variance $s$, i.e., its probability density is $\mathcal{N}(0, s)$ (i.e., $\frac{1}{\sqrt{2 \pi s}} e^{\frac{-x^{2}}{2 s}}$ ).

### 2.5.2 The Brownian motion as a limit of random walk

We may realize the Brownian motion as the limit of the random walk in the previous section. Namely, $Z_{n} \rightarrow z(t)$ as $n \rightarrow \infty$ with $m \Delta x \rightarrow x, n \Delta t \rightarrow t$ and $\frac{(\Delta x)^{2}}{\Delta t}=\sigma$ fixed. This can be proved by the Stirling formula:

$$
n!\approx \sqrt{2 \pi} n^{n+\frac{1}{2}} e^{-n}
$$

Recall that the probability

$$
P\left(Z_{n}=m \Delta x \mid Z_{0}=0\right)=\binom{n}{\frac{1}{2}(m+n)}\left(\frac{1}{2}\right)^{n} .
$$

Using the Stirling formula, we have for $n, p, n-p \gg 1$,

$$
\begin{aligned}
\binom{n}{\frac{1}{2}(m+n)}\left(\frac{1}{2}\right)^{n} & =\left(\frac{1}{2}\right)^{n} \frac{n!}{\left(\frac{1}{2}(n+m)\right)!\left(\frac{1}{2}(n-m)\right)!} \\
& \approx\left(\frac{1}{2}\right)^{n} \frac{\sqrt{2 \pi} n^{n+\frac{1}{2}} e^{-n}}{\sqrt{2 \pi}\left(\frac{1}{2}(n+m)\right)^{\frac{1}{2}(n+m)+\frac{1}{2}} \sqrt{2 \pi}\left(\frac{1}{2}(n-m)\right)^{\frac{1}{2}(n-m)+\frac{1}{2}}} \\
& =\left(\frac{2}{\pi n}\right)^{\frac{1}{2}}\left(1+\frac{m}{n}\right)^{-\frac{1}{2}(n+m)-\frac{1}{2}}\left(1-\frac{m}{n}\right)^{-\frac{1}{2}(n-m)-\frac{1}{2}} \\
& \approx\left(\frac{2}{\pi n}\right)^{\frac{1}{2}}\left(1-\left(\frac{m}{n}\right)^{2}\right)^{-\frac{1}{2} n} \\
& =\left(\frac{2}{\pi n}\right)^{\frac{1}{2}}\left[\left(1-\left(\frac{m}{n}\right)^{2}\right)^{\left(\frac{n}{m}\right)^{2}}\right]^{-\frac{m^{2}}{2 n}} \\
& \approx\left(\frac{2}{\pi n}\right)^{\frac{1}{2}} \exp \left(-\frac{m^{2}}{2 n}\right)
\end{aligned}
$$

As $m \Delta \rightarrow x, n \Delta t \rightarrow t,(\Delta x)^{2} / \Delta t=\sigma$ fixed, we obtain

$$
\begin{aligned}
P\left(Z_{n}=m \Delta x \mid Z_{0}=0\right) / 2 \Delta x & \approx\left(\frac{2}{\pi n}\right)^{\frac{1}{2}} \frac{1}{2 \Delta x} e^{-\frac{m^{2}}{2 n}} \\
& =\left(\frac{1}{2 \pi n \Delta t}\right)^{\frac{1}{2}} e^{-\frac{(m \Delta x)^{2}}{n \Delta t \sigma^{2}}} \cdot \frac{\sqrt{\Delta t}}{2(\Delta x)} \\
& \rightarrow \frac{1}{\sqrt{2 \pi \sigma^{2} t}} e^{-\frac{x^{2}}{2 \sigma^{2} t}}
\end{aligned}
$$

This means that $Z_{n} \rightarrow z(t)$.

### 2.5.3 Properties of Brownian motion

By definition

$$
\mathcal{P}(z(t)=x \mid z(0)=0)=\frac{1}{\sqrt{2 \pi t}} e^{-\frac{x^{2}}{2 t}} .
$$

We can check

1. $\langle z(t)\rangle=0$
2. $\left\langle z(t)^{2}>=t\right.$
3. Independence of disjoint increments

$$
\begin{equation*}
\mathcal{P}(z(t)=x \mid z(0)=0)=\int_{-\infty}^{\infty} \mathcal{P}(z(t)=x \mid z(s)=y) \mathcal{P}(z(s)=y \mid z(0)=0) d y \tag{2.5}
\end{equation*}
$$

In particular, let us define an infinitesimal increment

$$
d z=z(t+d t)-z(t)
$$

We have

1. $\langle d z>=0$
2. $\left\langle(d z)^{2}\right\rangle=d t$

In fact we have more, we may think

$$
\begin{equation*}
d z=\epsilon \sqrt{d t} \tag{2.6}
\end{equation*}
$$

where $\epsilon$ is a random variable with standard Gaussian distribution $\mathcal{N}(0,1)$ (i.e. mean is 0 and variance is 1 ). And we have

$$
\begin{equation*}
(d z)^{2}=d t \text { with probability } 1 \tag{2.7}
\end{equation*}
$$

## Exercise

1. Check (2.5).

### 2.6 Itô's formula

In this section, we shall study differential equations which consist of deterministic part: $\dot{x}=b(x)$, and stochastic part $\sigma \dot{z}(t)$. Here, $z(t)$ is the Brownian motion. We call such an equation a stochastic differential equation and expressed as

$$
\begin{equation*}
d x(t)=b(x(t)) d t+\sigma(x(t)) d z(t) \tag{2.8}
\end{equation*}
$$

An important lemma for finding their solution is the following Itô's lemma.

Lemma 2.1 Suppose $x(t)$ satisfies the stochastic differential equation (2.8), and $f(x, t)$ is a smooth function. Then $f(x(t), t)$ satisfies the following stochastic differential equation:

$$
\begin{equation*}
d f=\left(f_{t}+b f_{x}+\frac{1}{2} \sigma^{2} f_{x x}\right) d t+\sigma f_{x} d z \tag{2.9}
\end{equation*}
$$

Proof. This is not a proof, rather an intuition why (2.9) is true. According to the Taylor expansion,

$$
d f=f_{t} d t+f_{x} d x+\frac{1}{2} f_{t t}(d t)^{2}+f_{x t} d x d t+\frac{1}{2} f_{x x}(d x)^{2}+\cdots
$$

Plug (2.8) into this equation. We recall that $d z=\epsilon \sqrt{d t}$, where $\epsilon$ is a random variable with standard Gaussian distribution $\mathcal{N}(0,1)$. In the Taylor expansion of $d f(x(t), t)$, the terms $(d t)^{2}, d t \cdot d z$ are relative unimportant as comparing with the $d t$ term and $d z$ term. Using (2.8) and noting $(d z)^{2}=d t$ with probability 1 , we obtain (2.9).

A simple application of Itô's lemma is to find the transition probability density function for the s.d.e.

$$
d x=a d t+\sigma d z
$$

where $a$ and $\sigma$ are constants. By letting $y=x-a t$, from Itô's lemma, $y$ satisfies $d y=\sigma d z$. Thus, the transition probability density function for $y$ is

$$
\mathcal{P}\left(y(t)=y \mid y(0)=y_{0}\right)=\frac{1}{\sqrt{2 \pi \sigma^{2} t}} e^{-\left(y-y_{0}\right)^{2} / 2 \sigma^{2} t}
$$

Or equivalently, the transition probability density function for $x$ is

$$
\mathcal{P}\left(x(t)=x \mid x(0)=x_{0}\right)=\frac{1}{\sqrt{2 \pi \sigma^{2} t}} e^{-\left(x-a t-x_{0}\right)^{2} / 2 \sigma^{2} t} .
$$

### 2.7 The solution of the continuous asset price model

In this section, we want to find the transition probability density function for the continuous asset price model:

$$
\begin{equation*}
d S=\mu S d t+\sigma S d z \tag{2.10}
\end{equation*}
$$

with initial data $S(0)=S_{0}$. We apply Itô's lemma with $x=f(S)=\log S$. Then $x$ satisfies the s.d.e.

$$
d x=\left(\mu-\frac{\sigma^{2}}{2}\right) d t+\sigma d z
$$

and $x(0)=x_{0}:=\log S_{0}$. From the discussion of the previous section, we obtain

$$
\mathcal{P}\left(x(t)=x \mid x(0)=x_{0}\right)=\frac{1}{\sqrt{2 \pi \sigma^{2} t}} e^{-\left(x-x_{0}-\left(\mu-\frac{\sigma^{2}}{2}\right) t\right)^{2} / 2 \sigma t}
$$

From

$$
\begin{aligned}
\mathcal{P}\left(x(t)=x \mid x(0)=x_{0}\right) d x & =\mathcal{P}\left(x(t)=x \mid x(0)=x_{0}\right) d S / S \\
& =\mathcal{P}\left(S(t)=S \mid S(0)=S_{0}\right) d S
\end{aligned}
$$

we obtain that the transition probability density function for $S(t)$ is

$$
\begin{equation*}
\mathcal{P}\left(S(t)=S \mid S(0)=S_{0}\right)=\frac{1}{\sqrt{2 \pi \sigma^{2} t} S} e^{-\left(\log \frac{S}{S_{0}}-\left(\mu-\frac{\sigma^{2}}{2}\right) t\right)^{2} / 2 \sigma^{2} t} \tag{2.11}
\end{equation*}
$$

This is called the lognormal distribution.

## Exercise

1. Find the mean and variance of the lognormal distribution.


### 2.8 Continuous model as a limit of the discrete model

We want to show that the continuous model (2.10) is the limit of the discrete model (2.1). The parameters in (2.10) are $\mu$ and $\sigma$. The parameters in (2.1) are $u, d$ and $p$. We may assume $p=1 / 2$. First, we relate $(\mu, \sigma)$ and $(u, d)$. Both models should have the same mean and variance. For the continuous model, we compute its mean under the condition $S((n-1) \Delta t)=S_{n-1}$. Then

$$
\begin{aligned}
E\left(S(n \Delta t) \mid S_{n-1}\right) & =\int S \mathcal{P}\left(S, n \Delta t \mid S((n-1) \Delta t)=S_{n-1}\right) d S \\
& =\int S\left(\frac{1}{\sqrt{2 \pi \sigma^{2} \Delta t} S} e^{-\left(\log \frac{S}{S_{n-1}}-\left(\mu-\frac{1}{2} \sigma^{2}\right) \Delta t\right)^{2} / 2 \sigma^{2} \Delta t}\right) d S
\end{aligned}
$$

$$
\begin{aligned}
& =S_{n-1} \int \frac{1}{\sqrt{2 \pi \sigma^{2} \Delta t}} e^{-\left(x-\left(\mu-\frac{1}{2} \sigma^{2}\right) \Delta t\right)^{2} / 2 \sigma^{2} \Delta t} e^{x} d x \\
& =S_{n-1} e^{\mu \Delta t} \int \frac{1}{\sqrt{2 \pi \sigma^{2} \Delta t}} e^{-\left(\frac{x}{\sqrt{2 \sigma^{2} \Delta t}}-\frac{\sqrt{2 \sigma^{2} \Delta t}}{2}\right)^{2}} d x \\
& =e^{\mu \Delta t} S_{n-1} .
\end{aligned}
$$

Here, we have used the change-of-variable: $x=\log \frac{S}{S_{n-1}}$. For the second moment for the continuous model, we have

$$
\begin{aligned}
E\left(S(n \Delta t)^{2} \mid S_{n-1}\right) & =\int S^{2} \mathcal{P}\left(S, \Delta t \mid S_{n-1}\right) d S \\
& =e^{\left(2 \mu+\sigma^{2}\right) \Delta t} S_{n-1}^{2}
\end{aligned}
$$

On the other hand, the mean and the second moment for the discrete model in one time step $\Delta t$ are

$$
\begin{gathered}
\left(\frac{1}{2} u+\frac{1}{2} d\right) S_{n-1} \\
\left(\frac{1}{2} u^{2}+\frac{1}{2} d^{2}\right) S_{n-1}^{2} .
\end{gathered}
$$

In order to have the same means and variances in one time step in both models, we should require

$$
\begin{aligned}
\frac{1}{2} u^{2}+\frac{1}{2} d^{2} & =e^{\left(2 \mu+\sigma^{2}\right) \Delta t} \\
\frac{1}{2} u+\frac{1}{2} d & =e^{\mu \Delta t}
\end{aligned}
$$

Or

$$
\begin{align*}
& u=e^{\mu \Delta t}\left(1+\sqrt{e^{\sigma^{2} \Delta t}-1}\right)  \tag{2.12}\\
& d=e^{\mu \Delta t}\left(1-\sqrt{e^{\sigma^{2} \Delta t}-1}\right) . \tag{2.13}
\end{align*}
$$

These relate $(u, d)$ and $(\mu, \sigma)$.
Theorem 2.1 Let us fix $(\mu, \sigma)$. Let us choose a $\Delta t$ and $a \Delta x$ with $(\Delta x)^{2} / \Delta t=\sigma^{2}$. Define $(u, d)$ by (2.12) and (2.13). Then

$$
P\left(S_{0} u^{p} d^{n-p} \mid S_{0}\right) / 2 \Delta x \longrightarrow \mathcal{P}\left(S(t)=S \mid S(0)=S_{0}\right)
$$

as $n \Delta t \rightarrow t, n \rightarrow \infty$ and $S_{0} u^{p} d^{n-p} \rightarrow S$.
Proof. Let us define $x=\log S, x_{0}=\log S_{0}$. Then

$$
\log S_{0} u^{p} d^{n-p}=x_{0}+p \log u+(n-p) \log d .
$$

Thus, what we want to show is equivalent to

$$
P\left(x=x_{0}+p \log u+(n-p) \mid x_{0}\right) / 2 \Delta x \rightarrow \mathcal{P}\left(x(t)=x \mid x(0)=x_{0}\right)
$$

where

$$
\mathcal{P}\left(x(t)=x \mid x(0)=x_{0}\right)=\frac{1}{\sqrt{2 \pi \sigma^{2} t}} e^{-\left(x-x_{0}-\left(\mu-\frac{\sigma^{2}}{2}\right) t\right)^{2} / 2 \sigma^{2} t}
$$

To show this, we define $m=2 p-n$. Then $p=\frac{1}{2}(n+m), n-p=\frac{1}{2}(n-m)$. Hence

$$
\begin{aligned}
p \log u+(n-p) \log d & =\frac{1}{2}(n+m) \log u+\frac{1}{2}(n-m) \log d \\
& =\frac{1}{2} n \log (u d)+\frac{1}{2} m \log \left(\frac{u}{d}\right)
\end{aligned}
$$

From (2.12) and (2.13),

$$
\begin{gathered}
u \cdot d=e^{2 \mu \Delta t}\left(2-e^{\sigma^{2} \Delta t}\right) \approx e^{2 \mu \Delta t} \cdot e^{-\sigma^{2} \Delta t} \\
\frac{u}{d}=1+2 \sigma \sqrt{\Delta t}+\sigma^{2} \Delta t \approx e^{2 \sigma \sqrt{\Delta t}}
\end{gathered}
$$

Hence

$$
\begin{aligned}
\frac{1}{2} n \log u d+\frac{1}{2} m \log \left(\frac{u}{d}\right) & \approx n\left(\mu-\frac{1}{2} \sigma^{2}\right) \Delta t+m \sigma \sqrt{\Delta t} \\
& =n\left(\mu-\frac{1}{2} \sigma^{2}\right) \Delta t+m \Delta x
\end{aligned}
$$

Define $\Delta x$ such that $\frac{(\Delta x)^{2}}{\Delta t}=\sigma^{2}$. Then

$$
p \log u+(n-p) \log d=n \Delta t\left(\mu-\frac{1}{2} \sigma^{2}\right)+m \Delta x
$$

Recall that the probability that the price moves up $p$ times is $\binom{n}{p}\left(\frac{1}{2}\right)^{n}$. Then the density is

$$
\binom{n}{p}\left(\frac{1}{2}\right)^{n} / 2 \Delta x \approx\left(\frac{2}{n \pi}\right)^{\frac{1}{2}} e^{-\frac{m^{2}}{2 n}} \longrightarrow \frac{1}{\sqrt{2 \pi \sigma^{2} t}} e^{-\left(x-x_{0}-\left(\mu-\frac{\sigma^{2}}{2}\right) t\right)^{2} / 2 \sigma^{2} t}
$$

### 2.9 Simulation of asset price model

Typically, $\mu=0.16, \sigma$ is $0.20 \sim 0.40$ for a stock. To simulate the model

$$
\begin{aligned}
\frac{d S}{S} & =\mu d t+\sigma d z \\
S(0) & =S_{0}
\end{aligned}
$$

We perform $N$ sample paths $\omega_{1}, \cdots, \omega_{N}$. In each path, we choose time step $\Delta t=0.01$, for instance. We obtain $S_{k+1}$ from $S_{k}$ by discretizing the s.d.e. and sample a number $\xi$ from the normalized Gaussian distribution $\mathcal{N}(0,1)$ :

$$
\begin{aligned}
\frac{S_{k+1}-S_{k}}{S_{k}} & =\mu \Delta t+\sigma \xi \sqrt{\Delta t} \\
& =0.16 \times 0.01+0.2 \times 0.5 \times 0.1
\end{aligned}
$$

Here, $\xi=0.5$ is the sampled number. Then the transition probability density function

$$
\int_{a}^{b} \mathcal{P}\left(S(t)=S \mid S(0)=S_{0}\right) d S \approx \#\left\{\omega \mid a \leq S_{n}(\omega) \leq b\right\} / N
$$

## Chapter 3

## Black-Scholes Analysis

### 3.1 The hypothesis of no-arbitrage-opportunities

The option pricing theory was introduced by Black and Scholes. The fundamental hypothesis of their analysis is that "there is no arbitrage opportunities in financial markets".

For simplicity, we shall also assume

1. There exists a risk-free investment that gives a guaranteed return with interest rate $r$. ( e.g. government bond, bank.)
2. Borrowing or lending at such riskless interest rate is always possible.
3. There is no transaction costs.
4. All trading profits are subject to the same tax rate.

We will use the following notations:

$$
\begin{array}{ll}
S & \text { current asset price } \\
E & \text { exercise price } \\
T & \text { expiry time } \\
t & \text { current time } \\
\mu & \text { growth rate of an asset } \\
\sigma & \text { volatility of an asset } \\
S_{T} & \text { asset price at } T \\
r & \text { risk-free interest rate } \\
c & \text { value of European call option } \\
C & \text { value of American call option } \\
p & \text { value of European put option } \\
P & \text { value of American put option } \\
\Lambda & \text { the payoff function }
\end{array}
$$

### 3.2 Basic properties of option prices

### 3.2.1 The relation between payoff and options

1. Recall that

$$
\begin{array}{ll}
\Lambda(t)=\max \left(S_{t}-E, 0\right) & \text { for call option } \\
\Lambda(t)=\max \left(E-S_{t}, 0\right) & \text { for put option }
\end{array}
$$

2. $c\left(S_{T}, T\right)=\Lambda(T)$.

Otherwise, there is a chance of arbitrage. For instance, if $c\left(S_{T}, T\right)<\Lambda$, then we can buy a call on price $c$, exercise it immediately. If $S_{T}>E$, then $\Lambda=S_{T}-E>0$ and $c<\Lambda$ by our assumption. Hence we have an immediate net profit $S_{T}-E-c$. This contradicts to our hypothesis. If $c\left(S_{T}, T\right)>\Lambda$, we can short a call and earn $c$. If the person who buy the call does not claim, then we have net profit $c$. If he does exercise his call, then we can buy an asset from the market on price $S_{T}$ and sell to that person with price $E$. The cost to us is $S_{T}-E$. By doing so, the net profit we get is $c-\left(S_{T}-E\right)>0$. Again, this is a contradiction.
3. Similarly, we have

$$
\begin{aligned}
p\left(S_{T}, T\right) & =\Lambda(T) \\
C\left(S_{t}, t\right) & =\Lambda(t) \\
P\left(S_{t}, t\right) & =\Lambda(t)
\end{aligned}
$$

### 3.2.2 European options

Lemma 3.2 We have the following for European options

$$
\begin{align*}
& \max \left\{S-E e^{-r(T-t)}, 0\right\} \leq c \leq S  \tag{3.1}\\
& \max \left\{E e^{-r(T-t)}-S, 0\right\} \leq p \leq E e^{-r(T-t)} \tag{3.2}
\end{align*}
$$

and the put-call parity

$$
\begin{equation*}
p+S=c+E e^{-r(T-t)} \tag{3.3}
\end{equation*}
$$

To show these, we need the following definition and lemmae.
Definition 2.1 A portfolio is a collection of investments.
For instance, a portfolio $I=c-\Delta S$ means that we long a call and short $\Delta$ amount of an asset $S$.

Lemma 3.3 Suppose $I(t)$ and $J(t)$ are two portfolios containing no American options such that $I(T) \leq J(T)$. Then under the hypothesis of no-arbitrage-opportunities, we can conclude that $I(t) \leq J(t), \forall t \leq T$.

Proof. Suppose the conclusion is false, i.e., there exists a time $t \leq T$ such that $I(t)>J(t)$. An arbitrageur can buy (long) $J(t)$ and short $I(t)$ and immediately gain a profit $I(t)-J(t)$. Since $I$ and $J$ containing no American options, nothing can be exercised before $T$. At time $T$, since $I(T) \leq J(T)$, he can use $J(T)$ (what he has) to cover $I(T)$ (what he shorts) and gains a profit $J(T)-I(T)$. This contradicts to the hypothesis of no-arbitrage-opportunities.

As a corollary, we have
Corollary 2.1 If $I(T)=J(T)$, then $I(t)=J(t), \forall t \leq T$.
Now, we can prove the basic properties of European options 1-5.

## Proof of Lemma 3.2.

1. Let $I=c$ and $J=S$. At $T$, we have

$$
I(T)=c_{T}=\max \left\{S_{T}-E, 0\right\} \leq \max \left\{S_{T}, 0\right\}=S_{T}=J(T)
$$

Hence, $I(t) \leq J(t)$ holds for all $t \leq T$.
Remark. The equality holds when $E=0$. In this case $c=S$
2. Consider $I=c+E e^{-r(T-t)}$ and $J=S$. At time $T$,

$$
I(T)=\max \left\{S_{T}-E, 0\right\}+E=\max \left\{S_{T}, E\right\} \geq S_{T}=J(T)
$$

This implies $I(t) \geq J(t)$.
3. Let $I=p$ and $J=E e^{-r(T-t)}$. At time $T$,

$$
I(T)=\max \left\{E-S_{T}, 0\right\} \leq E=J(T) .
$$

Hence, $I(t) \leq J(t)$.
4. Consider $I=p+S$ and $J=E e^{-r(T-t)}$. At time $T$,

$$
I(T)=\max \left\{S_{T}, E\right\} \geq E=J(T)
$$

. Hence, $I(t) \geq J(t)$.
5. Consider $I=c+E e^{-r(T-t)}$ and $J=p+S$. At time $T$,

$$
\begin{aligned}
& I(T)=c+E=\max \left\{S_{T}-E, 0\right\}+E=\max \left\{S_{T}, E\right\}, \\
& J(T)=p+S=\max \left\{E-S_{T}, 0\right\}+S_{T}=\max \left\{E, S_{T}\right\}
\end{aligned}
$$

Hence, $I(t)=J(t)$.

### 3.2.3 Basic properties of American options

Lemma 3.4 For American options, we have
(i) The optimal exercise time for American call option is $T$ and we have $C=c$.
(ii) The optimal exercise time for American put option is as earlier as possible, i.e. $t$, and we have $P \geq p$.
(iii) The put-call parity for American option:

$$
\begin{equation*}
S-E<C-P<S-E e^{-r(T-t)} \tag{3.4}
\end{equation*}
$$

As a consequence, $P \leq E$.
To prove these properties, we need the following lemma.
Lemma 3.5 Let I or $J$ be two portfolios that contain American options. Suppose $I(\tau) \leq$ $J(\tau)$ at some $\tau \leq T$. Then $I(t) \leq J(t)$, for all $t \leq \tau$.

Proof. Suppose $I(t)>J(t)$ at some $t \leq \tau$. An arbitrageur can long $J(t)$ and short $I(t)$ at time $t$ to make profit $I(t)-J(t)$ immediately. At later time $\tau$, he can use $J(\tau)$ to cover $I(\tau)$ with additional profit $J(\tau)-I(\tau)$, in case the person who owns $I$ exercises his American option.

Remark. The equality also holds if $I(\tau)=J(\tau)$.

## Proof of Lemma 3.4.

1. Firstly, we show $C \geq c$. If not, then $c(\tau)>C(\tau)$ for some time $\tau \leq T$, we can buy $C$ and sell $c$ at time $\tau$ to make a profit $c(\tau)-C(\tau)$. The right of $C$ is even more than that of $c$. This is an arbitrage opportunity which is a contradiction.
Secondly, we show $c \geq C$. Consider two portfolios $I=C+E e^{-r(T-t)}$ and $J=S$. Suppose we exercise $C$ at some time $\tau \leq T$, then $I(\tau)=\max \left\{S_{\tau}-E, 0\right\}+$ $E e^{-r(T-\tau)}$ and $J(\tau)=S_{\tau}$. This implies $I(\tau) \leq J(\tau)$. By our lemma, $I(t) \leq J(t)$ for all $t \leq \tau$. Since $\tau \leq T$ arbitrary, we conclude $I(t) \leq J(t)$ for all $t \leq T$. Combine this inequality with the inequality of 2 ) of section 3.2 , we conclude $c=C$. Further, early exercise results $C(\tau)+E e^{-r(T-\tau)}<S(\tau)$. Hence, the optimal exercise time for American option is $T$.
2. Example. Suppose $S=50, E=40$. If $C$ is exercised before expiration, then the investor needs to pay 40 to buy the share. However, he can instead invest $\$ 40$ into the bank to earn interest and there is a chance that the stock price may go up.
3. Suppose $p(t)>P(t)$. Then we can make an immediate profit by selling $p$ and buying $P$. We earn $p-P$ and gain more right. This is a contradiction.

Next, we show that if we have a $P$, we should exercise it immediately. We consider two portfolios $I=P+S$ and $J=E e^{-r(T-t)}$. If we exercise $P$ at some time $\tau$, $t \leq \tau \leq T$, then

$$
I(\tau)=\max \left\{E-S_{\tau}, 0\right\}+S_{\tau}=\max \left\{E, S_{\tau}\right\}=E
$$

Putting this money into bank we will receive $E e^{r(T-\tau)}$ at time $T$. On the other hand, $J(\tau)=E^{-r(T-\tau)}$. Hence, $I(\tau) \geq J(\tau)$. Therefore, $I(t) \geq J(t)$. Further, we see that if we exercise $P$ at $t$, then $I(T)=E e^{r(T-t)}$ is the maximum. Hence we should exercise $P$ as early as possible.
4. The second inequality follows from the put-call parity (3.3) and the facts that $c=C$ and $P \geq p$. To show the first inequality, we consider two portfolios: $I=C+E$ and $J=P+S$. Suppose $P$ is exercised at some time $\tau, t \leq \tau \leq T$. Then we must have $E \geq S_{\tau}$ (otherwise, we should not exercise our put option). Therefore,

$$
\begin{aligned}
J(\tau) & =\max \left\{E-S_{\tau}, 0\right\}+S_{\tau}=E \\
I(\tau) & =C(\tau)+E e^{r(\tau-t)} \\
& =\max \left\{S_{\tau}-E, 0\right\}+E e^{r(\tau-t)} \\
& =E e^{r(\tau-t)} .
\end{aligned}
$$

From lemma, we have $I(t)>J(t)$. Hence $C+E>P+S$.

## Examples.

1. Suppose $S(t)=31, E=30, r=10 \%, T-t=0.25$ year, $c=3, p=2.25$. Consider two portfolios:

$$
\begin{aligned}
& I=c+E e^{-r(T-t)}=3+30 \times e^{-0.1 \times 0.25}=32.26, \\
& J=p+S=2.25+31=33.25 .
\end{aligned}
$$

We find $J(t)>I(t)$.
Strategy : long the security in portfolio $I$ and short the security in portfolio $J$. This results a cashflow: $-3+2.25+31=30.25$. Put this cash into a bank. We will get $30.25 \times e^{0.1 \times 0.25}=31.02$ at time $T$. Suppose at time $T, S_{T}>E$, we can exercise $c$, also we should buy a share for $E$ to close our short position of the stock. Suppose $S_{T}<E$, the put option will be exercised. This means that we need to buy the share for $E$ to close our short position. In both cases, we need to buy a share for $E$ to close the short position. Thus, the net profit is

$$
31.02-30=1.02
$$

2. Consider the same situation but $c=3$ and $p=1$. In this case

$$
\begin{aligned}
& I=c+E e^{-r(T-t)}=32.25 \\
& J=p+S=1+31=32
\end{aligned}
$$

and we see that $J$ is cheaper.
Strategy: We long $J$ and short $I$. To long $J$, we need an initial investment $31+1$, to short $c$, we gain 3 . Thus, the net investment is $31+1-3=29$ initially. We can finance it from the bank, and we need to pay $29 \times e^{0.1 \times 0.25}=29.73$ to the bank at time $T$. Now, at $T$, we must have that either $c$ or $p$ will be exercised. If $S_{T}>E$, then $c$ is exercised. We need to sell the share for $E$ to close our short position for $c$. If $S_{T}<E$, we exercise $p$. That is, we sell the share for $E$. In both cases, we sell the share for $E$. Thus, the net profit is $30-29.73=0.27$.

Remark. $P-p$ is called the time value of a put. The maximal time value is $E-E e^{-r(T-t)}$.

### 3.2.4 Dividend Case

Many stocks pay out dividends. These are payments to shareholders out of the profits made by the company. Since the company's wealth does not change after paying the dividends, the stock price, the strike prices fall as the dividends being paid. If a company declared a cash dividend, the strike price for options was reduced on the ex-dividend day by the amount of the dividend.

Lemma 3.6 Suppose a dividend $D$ will be paid during the life of an option. Then we have for European option

$$
\begin{gather*}
S-D-E e^{-r(T-t)}<c \leq S  \tag{3.5}\\
-S+D+E e^{-r(T-t)}<p \leq E e^{-r(T-t)} \tag{3.6}
\end{gather*}
$$

and the put-call parity:

$$
\begin{equation*}
c+E e^{-r(T-t)}=p+S-D \tag{3.7}
\end{equation*}
$$

For the American options, we have (i)

$$
\begin{equation*}
S-D-E<C-P<S-E e^{-r(T-t)} \tag{3.8}
\end{equation*}
$$

provided the dividend is paid before exercising the put option, or (ii)

$$
\begin{equation*}
S-E<C-P<S-E e^{-r(T-t)} \tag{3.9}
\end{equation*}
$$

if the put is exercised before the dividend being paid.
Proof. We consider two portfolios:

$$
\begin{aligned}
& I=c+D+E e^{-r(T-t)} \\
& J=S
\end{aligned}
$$

Then at time $T$,

$$
\begin{aligned}
& I(T)=\max \left\{S_{T}-E, 0\right\}+D+E=\max \left\{S_{T}, E\right\}+D \\
& J(T)=S_{T}+D
\end{aligned}
$$

Hence $I(T) \geq J(T)$. This yields $I(t) \geq J(t)$ for all $t \leq T$. This proves

$$
c \geq S-D-E e^{-r(T-t)}
$$

In other word, $c$ is reduced by an amount $D$. Similarly, we have

$$
p \geq D+E e^{-r(T-t)}-S
$$

That is, $p$ increases by an amount $D$.
For the put-call parity, we consider

$$
\begin{aligned}
I & =c+D+E e^{-r(T-t)} \\
J & =S+p
\end{aligned}
$$

At time $T$,

$$
I=J=\max \left\{S_{T}, E\right\}+D
$$

This yields the put-call parity for all time.
When there is no dividend, we have shown that

$$
C-P<S-E e^{-r(T-t)}
$$

When there is dividend payment, we know that

$$
C_{D}<C, \quad P_{D}>P
$$

Hence,

$$
C_{D}-P_{D}<C-P<S-E e^{-r(T-t)} .
$$

For the American call option, we should not exercise it early, because the dividend will cause the stock price to jump down, making the option less attractive. We should exercise it immediately prior to an ex-dividend date.

For the American put option, we consider

$$
I=C+D+E, \quad J=P+S
$$

If we exercise $P$ at $\tau \leq T$, then $S_{\tau}<E$ and

$$
\begin{aligned}
& I(\tau)=D+E e^{r(\tau-t)} \\
& J(\tau)=E+D
\end{aligned}
$$

We have $J(\tau) \leq I(\tau)$. Hence $J(t) \leq I(t)$ for all $t \leq \tau$.
If the put option is exercised before the dividend being paid, then we should consider $I=C+E$ and $J=P+S$. At $\tau$,

$$
\begin{aligned}
& I(\tau)=E e^{r(\tau-t)} \\
& J(\tau)=E
\end{aligned}
$$

Again, we have $J(\tau) \leq I(\tau)$. Hence $J(t) \leq I(t)$ for all $t \leq \tau$.

### 3.3 The Black-Scholes Equation

### 3.3.1 Black-Scholes Equation

The fundamental hypothesis of the Black-Scholes analysis is that there is no arbitrage opportunities. Besides, we make the following additional assumptions:
(1) The asset price follows the log-normal distribution.
(2) There exists a risk-free interest rate $r$.
(3) No transaction costs.
(4) No dividend paid.
(5) Shorting selling is permitted.

Our purpose is to value the price of an option (call or put). Let $V(S, t)$ denotes for the price of an option. The randomness of $V(S(t), t)$ would be fully correlated to that $S(t)$. Thus, we consider a portfolio which contains only $S$ and $V$, but in opposite position in order to cancel out the randomness. Then this portfolio becomes deterministic. To be more precise, let the portfolio be

$$
\Pi=V-\Delta S
$$

In one time step, the change of the portfolio is

$$
d \Pi=d V-\Delta d S
$$

Here $\Delta$ is held fixed during the time step. From Itôs lemma

$$
\begin{align*}
d \Pi= & \sigma S\left(\frac{\partial V}{\partial S}-\Delta\right) d z \\
& +\left(\mu S \frac{\partial V}{\partial S}-\mu \Delta S+\frac{\partial V}{\partial t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}\right) d t \tag{3.10}
\end{align*}
$$

Now, we can eliminate the randomness by choosing

$$
\Delta=\frac{\partial V}{\partial S}
$$

at the starting time of each time step. The resulting portfolio

$$
d \Pi=\left(\frac{\partial V}{\partial t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}\right) d t
$$

is wholly deterministic. From the hypothesis of no arbitrage opportunities, the return, $\frac{d \Pi}{\Pi}$, should be the same as $\Pi$ being invested in a riskless bank with interest rate $r$, i.e.

$$
\frac{d \Pi}{\Pi}=r d t
$$

Otherwise, there would be either a net loss or an arbitrage opportunity. Hence we must have

$$
\begin{aligned}
r \Pi d t & =\left(\mu S \frac{\partial V}{\partial S}-\mu \Delta S+\frac{\partial V}{\partial t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}\right) d t \\
& =\left(\frac{\partial V}{\partial t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}\right) d t
\end{aligned}
$$

or

$$
\begin{equation*}
\frac{\partial V}{\partial t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}=r\left(V-S \frac{\partial V}{\partial S}\right) \tag{3.11}
\end{equation*}
$$

This is the Black-Scholes partial differential equation (P.D.E.) for option pricing. Its lefthand side is the return from the hedged portfolio, while its right-hand side is the return from bank deposit. Note that the equation is independent of $\mu$.
Remark. Notice that the Black-Scholes equation is invariant under the change of variable $S \mapsto \lambda S$.

### 3.3.2 Boundary and Final condition for European options

- Final condition:

$$
\begin{aligned}
& c(S, T)=\max \{S-E, 0\} \\
& p(S, T)=\max \{E-S, 0\}
\end{aligned}
$$

In general, the final condition is

$$
V(S, T)=\Lambda(S)
$$

where $\Lambda$ is the payoff function.

- Boundary conditions:
(i) $\mathrm{On} S=0$ :

$$
c(0, \tau)=0, \forall t \leq \tau \leq T .
$$

This means that you wouldn't want to buy a right whose underlying asset costs nothing.
(ii) $\mathrm{On} S=0$ :

$$
p(0, \tau)=E e^{-r(T-\tau)}
$$

This follows from the put-call parity and $c(0, t)=0$.
(iii) For call option, at $S=\infty$ :

$$
c(S, t) \sim S-E e^{-r(T-t)}, \text { as } S \rightarrow \infty
$$

Since $S \rightarrow \infty$, the call option must be exercised, and the price of the option must be closed to $S-E e^{-r(T-t)}$.
(iv) For put option, at $S=\infty$ :

$$
p(S, t) \rightarrow 0, \text { as } S \rightarrow \infty
$$

As $S \rightarrow \infty$, the payoff function $\Lambda=\max \{E-S, 0\}$ is zero. Thus, the put option is unlikely to be exercised. Hence $p(S, T) \rightarrow 0$ as $S \rightarrow \infty$.

### 3.4 Exact solution for the B-S equation for European options

### 3.4.1 Reduction to parabolic equation with constant coefficients

Let us recall the Black-Scholes equation

$$
\begin{equation*}
\frac{\partial V}{\partial t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}+r S \frac{\partial V}{\partial S}-r V=0 \tag{3.12}
\end{equation*}
$$

This P.D.E. is a parabolic equation with variable coefficients. Notice that this equation is invariant under $S \rightarrow \lambda S$. That is, it is homogeneous in $S$ with degree 0 . We therefore make the following change-of-variable:

$$
d x=\frac{d S}{S}
$$

or equivalently,

$$
x=\log \frac{S}{E}
$$

The fraction $S / E$ makes $x$ dimensionless. The domain $S \in(0, \infty)$ becomes $x \in(-\infty, \infty)$ and

$$
\begin{aligned}
\frac{\partial V}{\partial x} & =\frac{\partial S}{\partial x} \frac{\partial V}{\partial S}=S \frac{\partial V}{\partial S} \\
\frac{\partial^{2} V}{\partial x^{2}} & =\frac{\partial}{\partial x}\left(S \frac{\partial V}{\partial S}\right) \\
& =\frac{\partial S}{\partial x} \frac{\partial V}{\partial S}+S \frac{\partial S}{\partial x} \frac{\partial^{2} V}{\partial S^{2}} \\
& =S \frac{\partial V}{\partial S}+S^{2} \frac{\partial^{2} V}{\partial S^{2}} \\
& =\frac{\partial V}{\partial x}+S^{2} \frac{\partial^{2} V}{\partial S^{2}}
\end{aligned}
$$

Next, let us reverse the time by letting

$$
\tau=T-t
$$

Then the Black-Scholes equation becomes

$$
\frac{\partial V}{\partial \tau}=\frac{1}{2} \sigma^{2} \frac{\partial^{2} V}{\partial x^{2}}+\left(r-\frac{1}{2} \sigma^{2}\right) \frac{\partial V}{\partial x}-r V .
$$

We can also make $V$ dimensionless by setting $v=V / E$. Then $v$ satisfies

$$
\begin{equation*}
\frac{\partial v}{\partial \tau}=\frac{1}{2} \sigma^{2} \frac{\partial^{2} v}{\partial x^{2}}+\left(r-\frac{1}{2} \sigma^{2}\right) \frac{\partial v}{\partial x}-r v \tag{3.13}
\end{equation*}
$$

The initial and boundary conditions for $v$ become

$$
\begin{aligned}
c(x, 0) & =\max \left\{e^{x}-1,0\right\} \\
p(x, 0) & =\max \left\{1-e^{x}, 0\right\} \\
c(-\infty, \tau) & =0, \\
p(-\infty, \tau) & =e^{-r \tau}, \\
c(x, \tau) & \rightarrow e^{x}-e^{-r \tau} \text { as } x \rightarrow \infty \\
p(x, \tau) & \rightarrow 0 \text { as } x \rightarrow \infty .
\end{aligned}
$$

Our goal is to solve $v$ for $0 \leq \tau \leq T$.

### 3.4.2 Further reduction

In investigating the equation (5.3), it is of the following form:

$$
\begin{equation*}
v_{t}+a v_{x}+b v=v_{x x} \tag{3.14}
\end{equation*}
$$

The part, $v_{t}+a v_{x}$ is call the advection part of (3.14). The term $b v$ is called the source term, and the tern $v_{x x}$ is called the diffusion term. Here, we have absorbed the diffusion coefficient $\frac{1}{2} \sigma^{2}$ in to time by setting $t=\tau /\left(\frac{1}{2} \sigma^{2}\right)$. (We somewhat abuse the notation here. The new $t$ here is different from the $t$ we used before.)

The advection part:

$$
v_{t}+a v_{x}=\left(\partial_{t}+a \partial_{x}\right) v
$$

is a direction derivative along the curve (called characteristic curve)

$$
\frac{d x}{d t}=a
$$

This suggests the following change-of-variable:

$$
\begin{aligned}
y & =x-a t \\
s & =t .
\end{aligned}
$$

Then the direction derivative become

$$
\begin{aligned}
\partial_{s} & =\partial_{t}+a \partial_{x} \\
\partial_{y} & =\partial_{x}
\end{aligned}
$$

Hence the equation is reduced to

$$
v_{s}+b v=v_{y y}
$$

Next, the equation $v_{s}+b v$ suggests that $v$ behaves like $e^{b s}$ along the characteristic curves. Thus, it is natural to make the following change-of-variable

$$
v=e^{b s} u
$$

Then the equation is reduced to

$$
u_{s}=u_{y y} .
$$

This is the standard heat equation. Its solution can be expressed as

$$
u(y, s)=\int \frac{1}{\sqrt{4 \pi s}} e^{-\frac{(y-z)^{2}}{4 s}} f(z) d z
$$

where $f$ is the initial data. A simple derivation of this solution is given in the Appendix of this chapter.

### 3.4.3 Black-Scholes formula

Lert us return to the Black-Scholes equation (5.3). Let us denote the rescaled payoff function by $\bar{\Lambda}(x)$. That is,

$$
\bar{\Lambda}(x)=\Lambda\left(E e^{x}\right) / E
$$

The change-of-variables above gives

$$
\begin{aligned}
s & =\tau /\left(\frac{1}{2} \sigma^{2}\right) \\
y & =x-a s \\
a & =1-r /\left(\frac{1}{2} \sigma^{2}\right) \\
b & =r /\left(\frac{1}{2} \sigma^{2}\right) \\
u & =e^{r \tau} v
\end{aligned}
$$

Then

$$
\begin{equation*}
v(x, \tau)=e^{-r \tau} \int \frac{1}{\sqrt{2 \pi \sigma^{2} \tau}} e^{-\frac{\left(x-z\left(r-\frac{1}{2} \sigma^{2}\right) \tau\right)^{2}}{2 \sigma^{2} \tau}} \bar{\Lambda}(z) d z \tag{3.15}
\end{equation*}
$$

In terms of the original variables, we have the following Black-Scholes formula:

$$
\begin{equation*}
V(S, t)=e^{-r(T-t)} \int \frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)} S^{\prime}} e^{-\frac{\left(\log \left(\frac{S}{J^{\prime}}\right)-\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)\right)^{2}}{2 \sigma^{2}(T-t)}} \Lambda\left(S^{\prime}\right) d S^{\prime} \tag{3.16}
\end{equation*}
$$

We may express it as

$$
\begin{equation*}
V(S, t)=e^{-r(T-t)} \int \mathcal{P}\left(S^{\prime}, T, S, t\right) \Lambda\left(S^{\prime}\right) d S^{\prime} \tag{3.17}
\end{equation*}
$$

Here,

$$
\begin{equation*}
\mathcal{P}\left(S^{\prime}, T, S, t\right):=\frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)} S^{\prime}} e^{-\frac{\left(\log \left(\frac{S}{S^{\prime}}\right)-\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)\right)^{2}}{2 \sigma^{2}(T-t)}} \tag{3.18}
\end{equation*}
$$

This is the transition probability density of an asset price model with growth rate $r$ and volatility $\sigma$. In other words, $V$ is the present value of the expectation of the payoff under an asset price model whose volatility is $\sigma$ and whose growth rate is $r$. We shall come back to this point later.

### 3.4.4 Special cases

1. European call option. The rescaled payoff function for a European call option is

$$
\bar{\Lambda}(z)=\max \left\{e^{z}-1,0\right\}
$$

Then

$$
v(x, \tau)=e^{-r \tau} \int_{0}^{\infty} \frac{1}{\sqrt{2 \pi \sigma^{2} \tau}} e^{-\left(x-z-\left(\frac{1}{2} \sigma^{2}-r\right) \tau\right)^{2} /\left(2 \sigma^{2} \tau\right)}\left(e^{z}-1\right) d z
$$

This can be integrated. Finally, we get the exact solution for the European call option

$$
\begin{align*}
c(S, t) & =S \mathcal{N}\left(d_{1}\right)-E e^{-r(T-t)} \mathcal{N}\left(d_{2}\right)  \tag{3.19}\\
\mathcal{N}(y) & =\frac{1}{\sqrt{2 \pi}} \int_{-\infty}^{y} e^{-\frac{z^{2}}{2}} d z  \tag{3.20}\\
d_{1} & =\frac{\log \left(\frac{S}{E}\right)+\left(r+\frac{1}{2} \sigma^{2}\right)(T-t)}{\sigma \sqrt{T-t}}  \tag{3.21}\\
d_{2} & =\frac{\log \left(\frac{S}{E}\right)+\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)}{\sigma \sqrt{T-t}} \tag{3.22}
\end{align*}
$$

Exercise. Prove the formula (3.19).
2. European put option. Recall the put-call parity

$$
c+E e^{-r(T-t)}=p+S
$$

We can obtain the price for $p$ from $c$ :

$$
\begin{equation*}
p(S, t)=E e^{-r(T-t)} \mathcal{N}\left(-d_{2}\right)-S \mathcal{N}\left(-d_{1}\right) \tag{3.23}
\end{equation*}
$$

Exercise. Show that $\mathcal{N}\left(d_{1}\right)-1=\mathcal{N}\left(-d_{1}\right)$. Use this to prove (3.23).
3. Forward contract Recall that a forward contract is an agreement between two parties to buy or sell an asset at certain time in the future for certain price. The payoff function for such a forward contract is

$$
\Lambda(S)=S-E
$$

The value $V$ for this contract also satisfies the B-S equation. Thus, its solution is given by

$$
V=E e^{-r \tau} u
$$

where

$$
\begin{aligned}
u(x, \tau) & =\frac{1}{\sqrt{2 \pi \sigma^{2} \tau}} \int_{-\infty}^{\infty} e^{-\frac{\left(y-z-\left(r-\frac{1}{2} \sigma^{2}\right) \tau\right)^{2}}{2 \sigma^{2} \tau}}\left(e^{z}-1\right) d z \\
& =e^{x+r \tau}-1
\end{aligned}
$$

Hence,

$$
\begin{equation*}
V(S, t)=S-E e^{-r(T-t)} \tag{3.24}
\end{equation*}
$$

This means that the current value of a forward contract is nothing but the difference of $S$ and the discounted $E$. Notice that this value is independent of the volatility $\sigma$ of the underlying asset.

Exercise. Show that the payoff function of a portfolio $c-p$ is $S-E$. From this and the Black-Scholes formula (3.16), show the formula of the put-call parity.
4. Cash-or-nothing. A contact with cash-or-nothing is just like a bet. If $S_{T}>E$, then the reward is $B$. Otherwise, you get nothing. The payoff function is

$$
\Lambda(S)= \begin{cases}B & \text { if } S>E \\ 0 & \text { otherwise }\end{cases}
$$

Using the Black-Scholes formula (3.16), we obtain the value of a cash-or-nothing contract to be

$$
\begin{equation*}
V(S, t)=B e^{-r(T-t)} \mathcal{N}\left(d_{2}\right) \tag{3.25}
\end{equation*}
$$

5. Supershare. Supershare is a binary option whose payoff function is defined to be

$$
\Lambda(S)= \begin{cases}B & \text { if } E_{1}<S<E_{2} \\ 0 \text { otherwise. }\end{cases}
$$

One can show that the value for this binary option is

$$
V(S, t)=B e^{-r(T-t)}\left(\mathcal{N}\left(d_{2}\left(E_{1}\right)\right)-\mathcal{N}\left(d_{2}\left(E_{2}\right)\right)\right)
$$

where $d_{2}(E)$ is given by (3.22).
6. Deterministic case ( $\sigma=0$ ). In this case, the Black-Scholes equation is reduced to

$$
V_{t}+r S V_{s}-r V=0 .
$$

Or in $\tau, x$ and $u$ variables:

$$
u_{\tau}-r u_{x}=0
$$

with initial data

$$
u(x, 0)=\Lambda\left(E e^{x}\right)
$$

Thus,

$$
u(x, \tau)=\Lambda\left(S e^{x+r \tau)}\right.
$$

Or

$$
V(S, t)=e^{-r(T-t)} \Lambda\left(S e^{r(T-t)}\right)
$$

This means that when the process is deterministic, the value of the option is the payoff function evaluated at the future price of $S$ at $T$ (that is $S e^{r(T-t)}$ ), and then discounted by the factor $e^{-r(T-t)}$.

### 3.5 Risk Neutrality

Notice that the growth rate $\mu$ does not appear in the Black-Scholes equation. The option may be valued as if all random walks involved are risk neutral. This means that the drift term (growth rate) $\mu$ in the asset pricing model can be replaced by $r$. The option is then valued by calculating the present value of its expected return at expiry. Recall the lognormal probability density function with growth rate $r$, volatility $\sigma$ is

$$
\begin{equation*}
\mathcal{P}\left(S^{\prime}, T, S, t\right):=\frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)} S^{\prime}} e^{-\frac{\left(\log \left(\frac{S}{S^{\prime}}\right)-\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)\right)^{2}}{2 \sigma^{2}(T-t)}} \tag{3.26}
\end{equation*}
$$

This is the transition probability density of an asset price model in a risk-neutral world:

$$
\begin{equation*}
\frac{d S}{S}=r d t+\sigma d z \tag{3.27}
\end{equation*}
$$

The expected return at time $T$ in this risk-neutral world is

$$
\int \mathcal{P}\left(S^{\prime}, T, s, t\right) \Lambda\left(S^{\prime}\right) d S^{\prime}
$$

At time $t$, this value should be discounted by $e^{-r(T-T)}$ :

$$
V(S, t)=e^{-r(T-t)} \int \mathcal{P}\left(S^{\prime}, T, S, t\right) \Lambda\left(S^{\prime}\right) d S^{\prime}
$$

We may reinvestigate the function $\mathcal{N}$ and the parameters $d_{i}$ in the Black-Scholes formula. After some calculation, we find

$$
\begin{equation*}
\mathcal{N}\left(d_{2}\right)=\int_{E}^{\infty} \mathcal{P}\left(S^{\prime}, T, S, t\right) d S^{\prime} \tag{3.28}
\end{equation*}
$$

This is the probability of the event $\{\tilde{S} \geq E\}$, where $\tilde{S}$ obeys the risk-neutral pricing model:

$$
\frac{d \tilde{S}}{\tilde{S}}=r d t+\sigma d z
$$

Similarly, one can show that

$$
\begin{equation*}
\mathcal{N}\left(d_{1}\right)=\frac{\int_{E}^{\infty} \mathcal{P}\left(S^{\prime}, T, S, t\right) S^{\prime} d S^{\prime}}{S e^{r(T-t)}} \tag{3.29}
\end{equation*}
$$

is the expectation of $\tilde{S}$ at $T$ when $S=1$ at $t$ and under the condition that $\tilde{S} \geq E$ at $T$.

### 3.6 The delta hedging

Hedging is the reduction of sensitivity of a portfolio to the movement of the underlying of asset by taking opposite position in different financial instruments. The Black-Scholes
analysis is a dynamical strategy. The delta hedge is instantaneously risk free. It requires a continuous rebalancing of the portfolio and the ratio of the holdings in the asset and the derivative product. The delta for a whole portfolio is $\Delta=\frac{\partial \Pi}{\partial S}$. This is the sensitivity of $\Pi$ against the change of $S$. By taking $d \Pi-\Delta \cdot d S$, the sensitivity of the portfolio to the asset price change is instantaneously zero.

Besides the delta helge, there are more sophisticated trading strategies such as:

$$
\begin{aligned}
\text { Gamma: } \Gamma & =\frac{\partial^{2} \Pi}{\partial^{2} S^{2}} \\
\text { Theta: } \theta & =-\frac{\partial \Pi}{\partial t}, \\
\text { Vega: } & =\frac{\partial \Pi}{\partial \sigma} \\
\text { rho: } \rho & =\frac{\partial \Pi}{\partial r}
\end{aligned}
$$

Hedging against any of these dependencies requires the use of another option as well as the asset itself. With a suitable balance of the underlying asset and other derivatives, hedgers can eliminate the short-term dependence of the portfolio on the movement in $t, S, \sigma, r$.

For the Delta-hedge for the European call and put options, we have the following propositions.

Proposition 1 For European call options, its $\Delta$ hedge is given by

$$
\Delta=\mathcal{N}\left(d_{1}\right) .
$$

Proof. By definition,

$$
\frac{\partial c}{\partial S}=\mathcal{N}\left(d_{1}\right)+S \cdot \mathcal{N}^{\prime}\left(d_{1}\right) \cdot d_{1 S}-E e^{-r(T-t)} \mathcal{N}^{\prime}\left(d_{2}\right) d_{2 S}
$$

Since

$$
d_{1}=\frac{\log \left(\frac{S}{E}\right)+\left(r+\frac{\sigma^{2}}{2}\right)(T-t)}{\sigma \sqrt{T-t}},
$$

we have

$$
\begin{aligned}
d_{1 S} & =\frac{1}{S \sigma \sqrt{(T-t)}}, \\
d_{2 S} & =\frac{1}{S \sigma \sqrt{(T-t)}}, \\
\mathcal{N}^{\prime}\left(d_{i}\right) & =\frac{1}{\sqrt{2 \pi}} e^{-d_{i}^{2}} .
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{\partial c}{\partial S} & =\mathcal{N}\left(d_{1}\right)+\left(S \mathcal{N}^{\prime}\left(d_{1}\right)-E e^{-r(T-t)} \mathcal{N}^{\prime}\left(d_{2}\right)\right) /(S \sigma \sqrt{T-t}) \\
& \equiv \mathcal{N}\left(d_{1}\right)+\mathcal{I} /(S \sigma \sqrt{T-t})
\end{aligned}
$$

We claim that $\mathcal{I}=0$. Or equivalently,

$$
\frac{S}{E} \frac{\mathcal{N}^{\prime}\left(d_{1}\right)}{\mathcal{N}^{\prime}\left(d_{2}\right)}=e^{-r \tau}
$$

This follows from the computation below.

$$
\frac{S}{E} \frac{\mathcal{N}^{\prime}\left(d_{1}\right)}{\mathcal{N}^{\prime}\left(d_{2}\right)}=e^{x} \cdot e^{-\left(d_{1}^{2}-d_{2}^{2}\right) / 2}
$$

From (3.21)(3.22),

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

Hence,

$$
\frac{S}{E} \frac{\mathcal{N}^{\prime}\left(d_{1}\right)}{\mathcal{N}^{\prime}\left(d_{2}\right)}=e^{x} \cdot e^{-x-r \tau}=e^{-r \tau}
$$

Proposition 2 For European put options, its $\Delta$ hedge is given by

$$
\Delta=\mathcal{N}\left(-d_{1}\right) .
$$

Proof. From the put-call parity,

$$
\Delta=\frac{\partial p}{\partial S}=\frac{\partial c}{\partial S}-1=\mathcal{N}\left(d_{1}\right)-1=-\mathcal{N}\left(-d_{1}\right)
$$

### 3.6.1 Time-Dependent $r, \sigma, \mu$

Suppose $r, \sigma, \mu$ are functions of $r$, but also deterministic. The Black-Scholes remains the same. We use the change-of-variables:

$$
S=E e^{x}, \quad V=E v, \quad \tau=T-t
$$

The Black-Scholes equation is converted to

$$
\begin{equation*}
v_{\tau}=\frac{\sigma^{2}(\tau)}{2} v_{x x}+\left(r(\tau)-\frac{\sigma^{2}(\tau)}{2}\right) v_{x}-r(\tau) v \tag{3.30}
\end{equation*}
$$

We look for a new time variable $\hat{\tau}$ such that

$$
d \hat{\tau}=\sigma^{2}(\tau) d \tau
$$

For instance, we can choose

$$
\hat{\tau}=\int_{0}^{\tau} \sigma^{2}(\tau) d \tau
$$

Then the equation becomes

$$
\begin{equation*}
v_{\hat{\tau}}=\frac{1}{2} v_{x x}+a(\hat{\tau}) v_{x}-b(\hat{\tau}) v \tag{3.31}
\end{equation*}
$$

To eliminate $a(\hat{\tau})$, we consider the characteristic equation:

$$
\frac{d x}{d \hat{\tau}}=-a(\hat{\tau})
$$

This can be integrated and yields

$$
x=-\int_{0}^{\hat{\tau}} a\left(\tau^{\prime}\right) d \tau^{\prime}+y
$$

Or equivalently,

$$
y=x+\int_{0}^{\hat{\tau}} a\left(\tau^{\prime}\right) d \tau^{\prime} \equiv x+A(\hat{\tau})
$$

Now, we consider the change-of-variable:

$$
\binom{x}{\hat{\tau}} \rightarrow\binom{y}{\hat{\tau}_{1}}
$$

Then,

$$
\left.\frac{\partial}{\partial x}\right|_{\hat{\tau}}=\left.\frac{\partial y}{\partial x}\right|_{\hat{\tau}} \frac{\partial}{\partial y}=\frac{\partial}{\partial y},
$$

and

$$
\left.\frac{\partial}{\partial \hat{\tau}_{1}}\right|_{y}=\frac{\partial \hat{\tau}}{\partial \hat{\tau}_{1}} \frac{\partial}{\partial \hat{\tau}}+\frac{\partial x}{\partial \hat{\tau}_{1}} \frac{\partial}{\partial x}=\frac{\partial}{\partial \hat{\tau}}-a(\hat{\tau}) \frac{\partial}{\partial x} .
$$

The equation (3.31)is transformed to

$$
v_{\hat{\tau}_{1}}=\frac{1}{2} v_{y y}-b\left(\hat{\tau}_{1}\right) v .
$$

Let $B\left(\hat{\tau}_{1}\right)=\int_{0}^{\hat{\tau}_{1}} b\left(\tau^{\prime}\right) d \tau^{\prime}$, and $u=e^{B\left(\hat{\tau}_{1}\right)} v$, then $u_{\hat{\tau}_{1}}=\frac{1}{2} u_{y y}$. And we can solve this heat equation explicitly.

### 3.7 Trading strategy involving options

The options whose payoff are $\max \left\{S_{T}-E, 0\right\}$ or $\max \left\{E-S_{T}, 0\right\}$ are called vanilla option. In this section, we shall discuss more general payoff functions. The goal is to design a portfolio involving vanilla option with a designed payoff function.

### 3.7.1 Strategies involving a single option and stock

There are four cases:
a. $\Pi=S-c$ (writing a covered call option). In this strategy, we short a call, long a share to cover $c$. The payoff of $\Pi$ is $\Lambda=S-\max (S-E, 0)=\min \{S, E\}$. In this case, we anticipate the stock price will increase.
b. $\Pi=c-S$ (reverse of a covered call). In this strategy, we anticipate the stock price will decrease. And $\Lambda=-\min \{S, E\}$.
c. $\Pi=p+S$ (protective put). In this portfolio, we long a $p$ and buy a share to cover $p$. We anticipate the stock price will increase. The payoff is $\Lambda=S+\max \{E-S, 0\}=$ $\max \{S, E\}$.
d. $\Pi=-p-S$ (reverse of a protective put). We do not anticipate the stock price will increase. The payoff is $-\max \{S, E\}$.

Below are the payoff functions for the above four cases.

(a)


## (c)

(d)

### 3.7.2 Bull spreads

In this strategy, an investor anticipates the stock price will increase. However, he would like to give up some of his right if the price goes beyond certain price, say $E_{2}$. Indeed, he does not anticipate the stock price will increase beyond $E_{2}$. Hence he does want to own a right beyond $E_{2}$. Such a portfolio can be designed as

$$
\Pi=C_{E_{1}}-C_{E_{2}}, \quad E_{1}<E_{2}
$$

where $C_{E_{i}}$ is a European call option with exercise price $E_{i}$ and $C_{E_{1}}, C_{E_{2}}$ have the same expiry. The payoff

$$
\begin{aligned}
\Lambda & =\max \left\{S_{T}-E_{1}, 0\right\}-\max \left\{S_{T}-E_{2}, 0\right\} \\
& = \begin{cases}0 & \text { if } S_{T}<E \\
S_{T}-E_{1} & \text { if } E_{1}<S_{T}<E_{2} \\
E_{2}-E_{1} & \text { if } S_{T}>E_{2}\end{cases}
\end{aligned}
$$



Since $E_{1}<E_{2}$, we have $C_{E_{1}}>C_{E_{2}}$. A bull spread, when created from $C_{E_{1}}-C_{E_{2}}$, requires an initial investment. We can describe the strategy by saying that the investor has a call option with a strike price $E_{1}$ and has chosen to give up some upside potential by selling a call option with strike price $E_{2}>E_{1}$. In return, the investor gets $E_{2}-E_{1}$ if the price goes up beyond $E_{2}$.
Example: $C_{E_{1}}=3, C_{E_{2}}=1$ and $E_{1}=30, E_{2}=35$. The cost of the strategy is 2. The payoff

$$
\begin{cases}0 & \text { if } S_{T} \leq 30 \\ S_{T}-30 & \text { if } 30<S_{T}<35 \\ 5 & \text { if } S_{T} \geq 35\end{cases}
$$

The bull spread can also be created by using put options

$$
\Pi=P_{E_{1}}-P_{E_{2}}, \quad E_{1}<E_{2} .
$$

### 3.7.3 Bear spreads

An investor entering into a bull spread is hoping that the stock price will increase. By contrast, an investor entering into a bear spread is expecting the stock price will go down. The bear spread is

$$
\Pi=C_{E_{2}}-C_{E_{1}}, \quad E_{1}<E_{2}
$$

There is cash flow entered $\left(C_{E_{2}}-C_{E_{1}}\right)$. The payoff is

### 3.7.4 Butterfly spread

If an investor anticipate the stock price will stay in certain region, say, $E_{1}<S_{T}<E_{3}$, he or she can have a butterfly spread such that the payoff function is positive in that region and he or she gives up the return outside that region.

1. Butterfly spread using calls: Define the portfolio:

$$
\Pi=C_{E_{1}}-2 C_{E_{2}}+C_{E_{3}}, \text { with } E_{1}<E_{2}<E_{3} .
$$

where $E_{3}=E_{2}+\left(E_{2}-E_{1}\right)$. Its payoff function is a piecewise linear function and is determined by $\Lambda\left(E_{1}\right)=\Lambda\left(E_{3}\right)=0, \Lambda\left(E_{2}\right)=E_{2}-E_{1}$. Below is the graph of its payoff function.


Example: Suppose a certain stock is currently worth 61. A investor who feels that it is unlikely that there will be significant price move in the next 6 month. Suppose the market of 6 month calls are

| $E$ | $C$ |
| :--- | :--- |
| 55 | 10 |
| 60 | 7 |
| 65 | 5 |

The investor creates a butterfly spread by

$$
\Pi=C_{E_{1}}-2 C_{E_{2}}+C_{E_{3}} .
$$

The cost is $10+5-2 \times 7=1$. The payoff is

2. Butterfly spread using puts.

$$
\begin{gathered}
P_{E_{1}}+P_{E_{3}}-2 P_{E_{2}}, \quad E_{1}<E_{3}, \quad E_{2}=\frac{E_{1}+E_{3}}{2} . \\
\varphi_{E_{2}}= \begin{cases}\text { linear } \\
\Delta E & \text { if } S=E_{2} \\
0 & \text { if } S<E_{2}-\Delta E, \text { or } S>E_{2}-\Delta E\end{cases}
\end{gathered}
$$

Remark 1. Suppose European options were available for every possible strike price $E$, then any payoff function could be created theoretically:

$$
\Lambda(S)=\sum \frac{\Lambda_{i}}{\Delta E} \varphi_{E_{i}}
$$

where $E_{i}=i \Delta E, \Lambda_{i}$ is constant. Then $\Lambda\left(E_{i}\right)=\Lambda_{i}$ and $\Lambda$ is linear on every interval $\left(E_{i}, E_{i+1}\right)$ and $\Lambda$ is continuous. As $\Delta E \rightarrow 0$, we can approximate any payoff function by using butterfly spreads.
Remark 2. One can also use cash-or-nothing to create any payoff function:

$$
\Lambda(S)=\sum \frac{\Lambda_{i}}{\Delta E} \psi S-E_{i}
$$

where

$$
\psi(S):=H(S)-H(S-\Delta E)
$$

The value for such a portfolio is

$$
\begin{aligned}
V & =e^{-r(T-t)} \int \mathcal{P}\left(S^{\prime}, T, S, t\right) \Lambda\left(S^{\prime}\right) d S^{\prime} \\
& =e^{-r(T-t)} \Sigma \Lambda_{i} \mathcal{P}\left(E_{i} \leq S \leq E_{i+1}\right)
\end{aligned}
$$

## Chapter 4

## Variations on Black-Scholes models

### 4.1 Options on dividend-paying assets

Dividends are payments to the shareholders out of the profits made by the company. We will consider two "deterministic" models for dividend. One has constant dividend yield. The other has discrete dividend payments.

### 4.1.1 Constant dividend yield

Suppose that in a short time $d t$, the underlying asset pays out a dividend $D_{0} S d t$, where $D_{0}$ is a constant, called the dividend yield. This continuous dividend structure is a good model for index options and for short-dated currency options. In the latter case, $D_{0}=r_{f}$, the foreign interest rate.

As the dividend is paid, the return $\frac{d S}{S}$ must fall by the amount of the dividend payment $D_{0} d t$. It follows the s.d.e. for the asset price is

$$
\frac{d S}{S}=\left(\mu-D_{0}\right) d t+\sigma d z
$$

For a portfolio : $\Pi=V-\Delta S$, we choose $\Delta=\frac{\partial V}{\partial S}$ in order to eliminate the randomness of $d \Pi$. In one time step, the change of portfolio is

$$
d \Pi=d V-\Delta d S-\Delta D_{0} S d t
$$

the last term $-\Delta D_{0} S d t$ is the dividend our assets received. Thus

$$
\begin{aligned}
d \Pi= & d V-\Delta\left(d S+D_{0} S d t\right) \\
= & \sigma S\left(\frac{\partial V}{\partial S}-\Delta\right) d z \\
& +\left(\left(\mu-D_{0}\right) S \frac{\partial V}{\partial S}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}+V_{t}-\left(\mu-D_{0}\right) \Delta S-\Delta D_{0} S\right) d t \\
= & \left(V_{t}-D_{0} S \frac{\partial V}{\partial S}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}\right) d t,
\end{aligned}
$$

Here, we have chosen $\Delta=\frac{\partial V}{\partial S}$ to eliminate the random term. From the absence of arbitrage opportunities, we must have

$$
d \Pi=r \Pi d t
$$

Thus,

$$
V_{t}-D_{0} S \frac{\partial V}{\partial S}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}=r\left(V-S \frac{\partial V}{\partial S}\right)
$$

i.e.,

$$
V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}+\left(r-D_{0}\right) S \frac{\partial V}{\partial S}-r V=0
$$

This is the Black-Scholes equation when there is a continuous dividend payment.
The boundary conditions are:

$$
\begin{aligned}
c(0, t) & =0 \\
c(S, t) & \sim S e^{-D_{0}(T-t)}
\end{aligned}
$$

The latter is the asset price $S$ discounted by $e^{-D_{0}(T-t)}$ from the payment of the dividend. The payoff function $c(S, T)=\Lambda(S)=\max \{S-E, 0\}$.

To find the solution, let us consider

$$
c(S, t)=e^{-D_{0}(T-t)} c_{1}(S, t)
$$

Then $c_{1}$ satisfies the original Black-Scholes equation with $r$ replaced by $r-D_{0}$ and the same final condition. The boundary conditions for $c_{1}$ are

$$
\begin{aligned}
c_{1}(0, t) & =0 \\
c_{1}(S, t) & \sim S \text { as } S \rightarrow \infty
\end{aligned}
$$

Hence,

$$
c_{1}(S, t)=S \mathcal{N}\left(d_{1,0}\right)-E e^{-\left(r-D_{0}\right)(T-t)} \mathcal{N}\left(d_{2,0}\right)
$$

or

$$
c(S, t)=S e^{-D_{0}(T-t)} \mathcal{N}\left(d_{1,0}\right)-E e^{-r(T-t)} \mathcal{N}\left(d_{2,0}\right)
$$

where

$$
\begin{aligned}
& d_{1,0}=\frac{\ln \frac{S}{E}+\left(r-D_{0}+\frac{1}{2} \sigma^{2}\right)(T-t)}{\sigma \sqrt{T-t}}, \\
& d_{2,0}=d_{1,0}-\sigma \sqrt{T-t}
\end{aligned}
$$

Remark. $c \searrow$ as $D_{0} \nearrow$.
Exercise. Derive the put-call parity for the European options on dividend-paying assets.

### 4.1.2 Discrete dividend payments

Suppose our asset pays just one dividend during the life time of the option, say at time $t_{d}$. The dividend yield is a constant. At $t_{d}+$, the asset holder receiver a payment $d_{y} S\left(t_{d}-\right)$. Hence,

$$
S\left(t_{d}+\right)=S\left(t_{d}-\right)-d_{y} S\left(t_{d}-\right)=\left(1-d_{y}\right) S\left(t_{d}-\right) .
$$

We claim that across the jumps, $V$ should be continuous, i.e.,

$$
V\left(S\left(t_{d}-\right), t_{d}-\right)=V\left(S\left(t_{d}+\right), t_{d}+\right)
$$

Reason : Otherwise, there is a net loss or gain from buying $V$ before $t_{d}$ then sell it right after $t_{d}$. To find $V(S, t)$, here is a procedure.

1. Solve the Black-Scholes from $T$ to $T_{d}+$ to obtain $V\left(S, t_{d}+\right)$ (using the payoff function $\Lambda$ )
2. Adjusting $V$ by

$$
V\left(S, t_{d}-\right)=V\left(\left(1-d_{y}\right) S, t_{d}+\right)
$$

3. Solve Black-Scholes equation from $t_{d}$ to $t$ with the final condition $V\left(\left(1-d_{y}\right) S, t_{d}+\right)$.

Let $c_{d}$ be the European option for this dividend-paying asset. Then

$$
\begin{gathered}
c_{d}(S, t)=c(S, t, E) \text { for } t_{d}+\leq t \leq T \\
c_{d}\left(S, t_{d}-\right)=c_{d}\left(S\left(1-d_{y}\right), t_{d}+\right)=c\left(S\left(1-d_{y}\right), t, E\right)
\end{gathered}
$$

Note that

$$
c\left(S\left(1-d_{y}\right), T, E\right)=\max \left\{S\left(1-d_{y}\right)-E, 0\right\}=\left(1-d_{y}\right) \max \left\{S-\left(1-d_{y}\right)^{-1} E, 0\right\}
$$

and the linearity of the Black-Scholes equation, we obtain

$$
c\left(S\left(1-d_{y}\right), t, E\right)=\left(1-d_{y}\right) c\left(S, t,\left(1-d_{y}\right)^{-1} E\right) .
$$

### 4.2 Warrants

An European warrant is a right to purchase an underlying stock at price $X$ at expiry. We want to determine the price of a warrant. Suppose a company has $N$ outstanding shares and $M$ outstanding European warrants. Suppose each warrant entitles the holder to purchase $\gamma$ share from the company at time $T$ at price $X$ per share. Let $V_{T}$ be the value of the company's equity at $T$. If the warrant holders exercise, then the company received a cash inflow $M \gamma X$ and the company's equity increases to $V_{T}+M \gamma X$. This value is distributed to $N+M \gamma$ shares. Hence the share price becomes

$$
\frac{V_{T}+M \gamma X}{N+\gamma M}
$$

The payoff to the warrant holder is

$$
\max \left\{\gamma\left[\frac{V_{T}+M \gamma X}{N+M \gamma}-X\right], 0\right\}=\frac{N \gamma}{N+M \gamma} \max \left\{\frac{V_{T}}{N}-X, 0\right\}
$$

This is exactly the payoff function for a European call. Thus, The value of the warrant at time $t$ should be

$$
w=\frac{N \gamma}{N+M \gamma} c\left(\frac{V}{N}, t, X\right)
$$

where $V$ is the value of the company's equity at time $t, c(S, t, X)$ is the value of a European call with strike price $X$. Since $V=N S+M w$, i.e., $\frac{V}{N}=S+\frac{M}{N} w$. We obtain a nonlinear algebraic equation for $w$ :

$$
\begin{aligned}
w & =\frac{N \gamma}{N+M \gamma} c\left(S+\frac{M}{N} w, t, X\right) \\
& =\frac{N \gamma}{N+M \gamma}\left(\left(S+\frac{M}{N} w\right) \mathcal{N}\left(d_{1}\right)-X e^{-r(T-t)} \mathcal{N}\left(d_{2}\right)\right)
\end{aligned}
$$

where

$$
\begin{aligned}
& d_{1}=\frac{\log \left(\left(S+\frac{M}{N} w\right) / X\right)+\left(r+\frac{1}{2} \sigma^{2}\right)(T-t)}{\sigma \sqrt{T-t}} \\
& d_{2}=d_{1}-\sigma \sqrt{T-t}
\end{aligned}
$$

This algebraic equation can be solved numerically.

### 4.3 Futures and futures options

### 4.3.1 Forward contracts

Recall that a forward contract is an agreement between two parties to buy or sell an underlying asset on a certain price $E$ at a certain future time $T$. Here, $E$ is called the delivery price. The payoff function for this forward contract is $\Lambda=S_{T}-E$. Based on the no arbitrage opportunity, the price for this forward contract is

$$
f=S-E e^{-r(T-t)}
$$

Definition 3.2 The forward price F for a forward contract is defined to be the delivery price which would make that contract have zero value, i.e.,

$$
F_{t}=S_{t} e^{r(T-t)}
$$

One can take another point of view. Consider a party who is short the contract. He can borrow an amount of money $S_{t}$ at time $t$ to buy an asset and use it to close his short position at $T$. The money he received at expiry, $F$, is used to pay the loan. If no arbitrage opportunities, then

$$
F=S_{t} e^{r(T-t)}
$$

### 4.3.2 Futures

Futures are very similar to the forward contracts, except they are traded in an exchange, thus, they are required to be standardized. This includes size, quality, price, expiry, . . etc. Let us explain the characters of a future by the following example.

1. Trading future contracts

- Suppose you call your broker to buy one July corn futures contract (5,000 bushels) on the Chicago Board of Trade (CBOT) at current market price.
- The broker send this signal to traders on the floor of the exchange.
- The trader signal this to ask other traders to sell, if no one want to sell, the trader who represents you will raise the price and eventually find someone to sell
- Confirmation: Price obtained are sent back to you.

2. Specification of the futures: In the above example, the specification of this future is

- Asset : quality
- Contract size: 5,000 bushel
- Delivery arrangement: delivery month is on December
- price quotes
- Daily price movement limits: these are specified by the exchange.
- Position limits: the maximum number of contracts that a speculator may hold.

3. Operation of margins

- Marking to market: Suppose an investor who contacts his or her broker on June 1, 1992, to buy two December 1992 gold futures contracts on New York Commodity Exchange. We suppose that the current future price is $\$ 400$ per ounce. The contract size is $\$ 100$ ounces, the investor want to buy $\$ 200$ ounces at this price. The broker will require the investor to deposit funds in a "margin account". The initial margin, say is $\$ 2,000$ per contract. As the futures prices move everyday, the amount of money in the margin account also changes. Suppose, for example, by the end of June 1, the futures price has dropped from $\$ 400$ to $\$ 397$. The investor has a loss of $\$ 200 \times 3=600$. This balance in the margin account would therefore be reduced by $\$ 600$. Maintaining margin needs to deposit. Certain account of money to keep that futures contract.
- Maintenance margin: To insure the balance in the margin account never becomes negative, a maintenance margin, which is usually lower than the initial margin, is set.

Theorem 4.2 Forward price and futures price are equal when the interest rates are constant.

Proof. Suppose a futures contract lasts for $n$ days. Let the future prices are

$$
F_{0}, \cdots, F_{n}
$$

at the end of each business day. Let $\delta$ be the risk-free interest rate per day. Consider the following two strategies:

1. Invest $G_{0}$ in a risk-free bond and take a long position of amount $e^{n \delta}$ forward contract. At day $n, G_{0} e^{n \delta}$ is used to buy the underlying asset at price $S_{T} e^{n \delta}$.
2.     - Invest $F_{0}$ amount of money in a risk-free bond.

- Take a long position of future $e^{\delta}$ amount of at the end of day 0 .
- Take a long position of future $e^{2 \delta}$ amount of at the end of day 1 .

| Day | 0 | 1 | 2 | $\cdots$ | $n-1$ | $n$ |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| futures price | $F_{0}$ | $F_{1}$ | $F_{2}$ | $\cdots$ | $F_{n-1}$ | $F_{n}$ |
| position | $e^{\delta}$ | $e^{2 \delta}$ | $e^{3 \delta}$ | $\cdots$ | $e^{n \delta}$ | 0 |
| gain/loss | 0 | $e^{\delta}\left(F_{1}-F_{0}\right)$ | $e^{2 \delta}\left(F_{2}-F_{1}\right)$ | $\cdots$ | $\cdots$ | $e^{n \delta}\left(F_{n}-F_{n-1}\right)$ |
| compound | $e^{\delta}\left(F_{1}-F_{0}\right) e^{(n-1) \delta}$ | $e^{2 \delta}\left(F_{2}-F_{0}\right) e^{(n-2) \delta}$ | $\cdots$ | $\cdots$ | $\cdots$ | $\left(F_{n}-F_{n-1}\right) e^{n \delta}$ |

The total gain/loss from the long position of the futures is

$$
\begin{aligned}
& \sum_{i=1}^{n}\left(F_{i}-F_{i-1}\right) e^{i \delta} \cdot e^{(n-i) \delta} \\
= & \left(F_{n}-F_{0}\right) e^{n \delta} \\
= & \left(S_{T}-F_{0}\right) e^{n \delta} .
\end{aligned}
$$

If we invest $F_{0}$ initially, at $T$, we received $F_{0} e^{n \delta}, N_{0}$ investment is required for all the long future positions. The payoff of strategy 2 is

$$
F_{0} e^{n \delta}+\left(S_{T}-F_{0}\right) e^{n \delta}=S_{T} e^{n \delta}
$$

Since both strategies have the same payoff, we conclude their initial investments must be the same, i.e., $F_{0}=G_{0}=S_{T} e^{r(T-t)}$.

### 4.3.3 Futures options

Options on futures are traded in many different exchanges. They require the delivery of an underlying futures contract when exercised. When a call futures option is exercised, the holder acquires a long position in the underlying futures contract plus a cash amount equal to the current futures price minus the exercise price.
Example. An investor who has a September futures call option on 25,000 pounds of copper with exercise price $E=70$ cents/pound. Suppose the current future price of copper
for delivery in September is 80 cents/pound. If the option is exercised, the investor received 10 cents $\times 25,000+$ long position in futures contract to buy 25,000 pound of copper in September at price 80 cents/pound.

The maturity date of futures option is generally on, or a few days before, the earlist delivery date of the underlying futures contract.

Futures options are more attractive to investors than options on the underlying assets when it is cheaper or more convenient to deliver futures contracts rather than the asset itself. Futures options are usually more liquid and involved lower transaction costs.

### 4.3.4 Black-Scholes analysis on futures options

As we have seen that the futures price is identical to the forward price when the interest rate is a constant, i.e., $F=S e^{r(T-t)}$. From Itôs lemma, we obtain a pricing model for $F$ :

$$
\begin{aligned}
d F & =\left(F_{t}+\frac{1}{2} F_{S S} \sigma^{2} S^{2}\right) d t+F_{S} d S \\
& =\left(-r e^{r(T-t)} S\right) d t+S e^{r(T-t)} \frac{d S}{S} \\
& =(-r F) d t+F(\mu d t+\sigma d z) \\
& =((\mu-r) F) d t+F \sigma d z .
\end{aligned}
$$

Hence,

$$
\begin{equation*}
\frac{d F}{F}=(\mu-r) d t+\sigma d z \tag{4.1}
\end{equation*}
$$

This means that the futures price is the same as a stock paying a dividend yield at rate $r$.
Next, we study the value $V$ of a futures option. It is a function of $F, t$. Consider a portfolio

$$
\Pi=V-\Delta F
$$

We choose $\Delta=\frac{\partial V}{\partial F}$ to eliminate randomness of $d \Pi$. Then

$$
\begin{aligned}
d \Pi & =d V-\Delta d F \\
& =\left(\frac{\partial V}{\partial F} \mu_{F} F+\frac{\partial V}{\partial t}+\frac{1}{2} \frac{\partial^{2} V}{\partial F^{2}} \sigma^{2} F^{2}\right) d t+\frac{\partial V}{\partial F} \sigma F d z-\frac{\partial V}{\partial F}\left(\mu_{F} F d t+\sigma F d z\right) \\
& =\left(\frac{\partial V}{\partial t}+\frac{1}{2} \frac{\partial^{2} V}{\partial F^{2}} \sigma^{2} F^{2}\right) d t .
\end{aligned}
$$

Since it costs nothing to enter into a future contract, the cost of setting up the above portfolio is just $V$. Thus based on the no arbitrage opportunity,

$$
d \Pi=r V d t,
$$

Thus, we obtain

$$
V_{t}+\frac{1}{2} \sigma^{2} F^{2} \frac{\partial^{2} V}{\partial F^{2}}=r V
$$

The payoff function for a call option is $\Lambda=\max \{F-E, 0\}$. This is because at time $T$, $S_{T}=F_{T}$.

To solve this equation, we recall the option price equation for stock paying dividend is

$$
V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}+\left(r-D_{0}\right) S \frac{\partial V}{\partial S}-r V=0
$$

In our case, $D_{0}=r$, so the futures call option

$$
c(F, t)=e^{-r(T-t)} c_{1}(F, t)
$$

where $c_{1}$ satisfies Black-Scholes equation with $r$ replaced by $r-r=0$. This gives

$$
\begin{aligned}
c_{1}(F, t) & =F \mathcal{N}\left(d_{1}\right)-E \mathcal{N}\left(d_{2}\right), \\
d_{2} & =\frac{\ln \frac{F}{E}-\frac{1}{2} \sigma^{2}(T-t)}{\sigma \sqrt{T-t}}, \\
d_{1} & =d_{2}+\sigma \sqrt{T-t}
\end{aligned}
$$

Notice that this $V$ is the same as $\tilde{V}(S(F, t), t)$, where $\tilde{V}$ is the solution of the option corresponding to the underlying asset $S$. That is

$$
\begin{aligned}
\tilde{V}(S, t) & =\mathcal{N}\left(d_{1}\right)-E e^{-r(T-t)} \mathcal{N}\left(d_{2}\right) \\
d_{2} & =\frac{\ln \frac{S}{E}+\left(r-\frac{1}{2} \sigma^{2}\right)(T-t)}{\sigma \sqrt{T-t}}, \\
d_{1} & =d_{2}+\sigma \sqrt{T-t}
\end{aligned}
$$

We can write this $\tilde{V}$ in terms of $F$ by $S=F e^{-r(T-t)}$. Plug this into the above equation to obtain

$$
\begin{aligned}
d_{2} & =\frac{\ln \frac{F e^{-r(T-t)}}{E}-\frac{1}{2} \sigma^{2}(T-t)}{\sigma \sqrt{T-t}} \\
& =\frac{\ln \frac{F}{E}-\frac{1}{2} \sigma^{2}(T-t)}{\sigma \sqrt{T-t}} \\
\tilde{V}(S(F, t), t) & =S \mathcal{N}\left(d_{1}\right)-E e^{-r(T-t)} \mathcal{N}\left(d_{2}\right) \\
& =F e^{-r(T-t)} \mathcal{N}\left(d_{1}\right)-E e^{-r(T-t)} \mathcal{N}\left(d_{2}\right) \\
& =V(F, t)
\end{aligned}
$$

## Conclusion:

1. Futures price $F=S_{t} e^{r(T-t)}$.
2. Futures price is the same as a stock paying dividend at yield rate $r$.
3. The price for future options is the same as the price for options on the underlying assets.

Finally, let us find the put-call parity for futures options.

## Proposition 3

$$
\begin{equation*}
c+E e^{-r(T-t)}=p+F e^{-r(T-t)} \tag{4.2}
\end{equation*}
$$

Proof. Consider two portfolios:

$$
\begin{aligned}
& A=c+E e^{-r(T-t)} \\
& B=p+F e^{-r(T-t)}+\text { a futures contract }
\end{aligned}
$$

At time $T$,

$$
\begin{aligned}
& \Lambda_{A}=\max \left\{F_{T}-E, 0\right\}+E=\max \left\{F_{T}, E\right\}, \\
& \Lambda_{B}=\max \left\{E-F_{T}, 0\right\}+F+\left(F_{T}-F\right)=\max \left\{E, F_{T}\right\} .
\end{aligned}
$$

Hence we obtain $A=B$.

## Chapter 5

## Numerical Methods

### 5.1 Monte Carlo method

We recall that the value of a European option is given by

$$
V(S, t)=e^{-r(T-t)} \int \mathcal{P}(\tilde{S}, T, S, t) \Lambda(\tilde{S}) d \tilde{S}
$$

where $\Lambda$ is the payoff function, $\mathcal{P}$ is the transition probability density function of $\tilde{S}$ which satisfies

$$
\begin{equation*}
\frac{d \tilde{S}}{\tilde{S}}=r d t+\sigma d z, \quad(\text { initial state } S(t)=S) \tag{5.1}
\end{equation*}
$$

i.e., it is the asset price model in the risk-neutral world. The Monte Carlo simulation is a numerical procedure to estimate $V$ based on this formula.

To find $V$, we sample, say, 10,000 paths from (5.1). We obtain $S_{i}(T), \quad i=1, \ldots, 10000$. We then approximate $V$ by

$$
V \approx e^{-r(T-t)} \frac{1}{N} \sum_{i=1}^{N} \Lambda\left(S_{i}(T)\right)
$$

To sample a path from (5.1), we divide the interval $[t, T]$ into $M$ subinterval with equal length $\Delta t=\frac{T-t}{M}$. We sample $M$ random numbers $\epsilon_{k}, k=1, \ldots, M$ with distribution $N(0,1)$ (i.e., the normal distribution with mean 0 , variance 1 ). We then define $S_{i}(t+k \Delta t)$ by

$$
\frac{S(t+k \Delta t)-S(t+(k-1) \Delta t)}{S(t+(k-1) \Delta t)}=r \Delta t+\sigma \epsilon_{k} \sqrt{\Delta t}
$$

Remark. The error of a Monte-Carlo method is $O\left(\frac{1}{\sqrt{N}}\right)$. If there is only one underlying asset, the Monte Carlo does not have any advantage. However, if there are many underlying assets, say more than three, the corresponding Black Scholes equation is a diffusion equation in high dimensions. In this case, finite difference method is very difficult and the Monte Carlo method wins.

### 5.2 Binomial Methods

In binomial method, we first simulate a risk-neutral asset price model forward in time by a binomial model, then we determine the option price from the expectation of the payoff function according to the price distribution of the asset in the risk-neutral world.

### 5.2.1 Binomial method for asset price model

We consider the underlying asset is risk-neutral, i,e.,

$$
\begin{equation*}
\frac{d S}{S}=r d t+\sigma d z \tag{5.2}
\end{equation*}
$$

We shall approximate this continuous model by the following discrete model.
First, we assume our discrete asset prices only take discrete values $S_{j}=S_{0} e^{j \Delta x}$, where $S_{0}$ is the asset price at current time $t$, and $\Delta x$ is a parameter to be determined later. We want to find the probability distribution of the asset price in a risk-neutral world at time $T$ whose current price is $S_{0}$.

Next, we discrete the continuous model in time, namely, we partition $[t, T]$ into $N$ subintervals with equal length $\Delta t=(T-t) / N$. The discrete asset price model is:
if the asset price is at $S_{j}$ at time step $n$, then the asset price will move up to $S_{j+1}=$ $S_{j} u$ with probability $p$ and move down to $S_{j-1}=S_{j} d$ with probability $1-p$. Here, $u=e^{\Delta x}$ and $d=e^{-\Delta x}$.

Let us denote the probability that the price is at $S_{j}$ at time step $n$ by $P_{j}^{n}$. Then $P_{0}^{0}=1$ and $P_{j}^{n}$ is exactly the binomial distribution:

$$
P_{j}^{n}= \begin{cases}\binom{n}{r} p^{r}(1-p)^{n-r} & n+j=2 r \\ 0 & \text { otherwise } .\end{cases}
$$

This discrete model depends on two parameters: $u$ and $p$. (The down ratio $d=1 / u$.) They are determined by the conditions so that the discrete model and the continuous model have the same mean and variance in one time step $\Delta t$. We recall that these conditions are

$$
\begin{aligned}
p u+(1-p) d & =e^{r \Delta t} \\
p u^{2}+(1-p) d^{2} & =e^{\left(2 r+\sigma^{2}\right) \Delta t}
\end{aligned}
$$

Thus, $p$ and $u$ can be expressed in terms of $r, \sigma$ and $\Delta t$. A simple calculation gives

$$
\begin{aligned}
u & =1+\sigma \sqrt{\Delta t}+O(\Delta t) \\
p & =\frac{1}{2}+O(\sqrt{\Delta t})
\end{aligned}
$$

We should require $\Delta t$ is chosen so that $0 \leq p \leq 1$.
Remark. If we denote $\log S / E$ by $x$, then the movement of $S$ on the discrete values $S_{j}$ corresponds to a movement of $x$ on $x_{j}=j \Delta x$. This movement is exactly the random walk we introduced in Chapter 2.

### 5.2.2 Binomial method for option

Since the asset price only takes discrete values $S_{j}$, we shall approximate $V\left(S_{j}, t+n \Delta t\right)$ by $V_{j}^{n}$. We recall that $V^{n}$ is the expected value of the option at $(n+1) \Delta t$ discounted by $e^{-r \Delta t}$. If $S$ takes value at $S_{j}$ at time step $n$, then $S$ takes values $S_{j+1}$ with probability $p$ and $S_{j-1}$ with probability $1-p$. Therefore, the expected value of $V$ at time step $n$ should satisfies

$$
e^{r \Delta t} V_{j}^{n}=p V_{j+1}^{n+1}+(1-p) V_{j-1}^{n+1}
$$

Example. For put option,

$$
\begin{aligned}
T & =5 \text { months }=0.4167 \text { year } \\
\Delta t & =1 \text { months }=0.0833 \text { year } \\
r & =0.1, \quad \sigma=0.4 \\
S & =\$ 50, \quad E=\$ 50 \\
u & =e^{\sigma \sqrt{\Delta t}}=1.1224 \\
d & =0.8909 \\
p & =0.5076 \\
e^{r \Delta t} & =1.0084
\end{aligned}
$$

We begin to generate a binomial tree from $S=50$ consisting of $S_{j}^{n}=S u^{r} d^{n-r}$, where $n+j=2 r,-n \leq j \leq n$. Then we compute $V_{j}^{n}$ inductively from $n=N$ to $n=0$ by

$$
e^{r \Delta t} V_{j}^{n}=p V_{j+1}^{n+1}+(1-p) V_{j-1}^{n+1}
$$

with $V_{j}^{N}$ being the payoff function. The value $V_{0}^{0}$ is our answer.

### 5.3 Finite difference methods (for the modified B-S eq.)

In this section, we shall solve the Black-Scholes equation by finite difference methods. Recall that the Black-Scholes equation is

$$
V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}=r\left(V-S \frac{\partial V}{\partial S}\right)
$$

Using dimensionless variables $S=E e^{x}, V=E v, \tau=T-t$, we have

$$
v_{\tau}=\frac{1}{2} \sigma^{2} \frac{\partial^{2} v}{\partial x^{2}}+\left(r-\frac{1}{2} \sigma^{2}\right) \frac{\partial v}{\partial x}-r v .
$$

Let $v=e^{-r \tau} u$, then $u$ satisfies

$$
\begin{equation*}
u_{\tau}=\frac{1}{2} \sigma^{2} \frac{\partial^{2} u}{\partial x^{2}}+\left(r-\frac{1}{2} \sigma^{2}\right) \frac{\partial u}{\partial x} \tag{5.3}
\end{equation*}
$$

The initial condition for $u$ is

$$
u(x, 0)= \begin{cases}\max \left\{e^{x}-1,0\right\} & \text { for call option } \\ \max \left\{1-e^{x}, 0\right\} & \text { for put option }\end{cases}
$$

The far field boundary condition for $u$ is

$$
u(-\infty, t)=0, \quad u(x, t)=e^{x} e^{r \tau} \text { as } x \rightarrow \infty
$$

for a call option, and

$$
u(-\infty, t)=1, \quad u(x, t)=0 \text { as } x \rightarrow \infty
$$

for a put option.

### 5.3.1 Discretization methods

To solve (5.3) numerically, we follow the following procedure:

1. Discretize space and time. We choose a proper finite domain $\left(x_{L}, x_{R}\right)$, discretize it into

$$
x_{j}=j \Delta x, \quad j=-N, \ldots, N, \text { where } \Delta x=\frac{x_{R}-x_{L}}{N} .
$$

Similarly, we discretize $[t, T]$ into $N$ steps, $\Delta t=\frac{T-t}{M}$.
We shall approximate $u\left(x_{j}, n \Delta \tau\right)$ by $U_{j}^{n}, V\left(x_{j}, n \Delta \tau\right)$ by $V_{j}^{n}$. From $v=e^{-r \tau} u$, we have

$$
V_{j}^{n}=e^{-r n \Delta \tau} U_{j}^{n} .
$$

2. Spatial discretization. We replace the spatial derivatives by finite differences:
(a) $u_{x}$ is replaced by one of the following three:

$$
u_{x} \leftarrow \begin{cases}\frac{\frac{u_{j+1}-u_{j-1}}{2_{0}}}{} \leftarrow \frac{u_{j}}{u_{j}-u_{j-1}} & \text { if } \frac{1}{2} \sigma^{2}-r>0 \\ \frac{u_{j+1}-u_{j}}{\Delta x} & \text { if } \frac{1}{2} \sigma^{2}-r \leq 0 .\end{cases}
$$

(b) $u_{x x} \leftarrow \frac{u_{j+1}-2 u_{j}+u_{j-1}}{(\Delta x)^{2}}$

Then the right-hand-side of (5.3) is discretized into

$$
(Q U)_{j} \equiv\left(\frac{\sigma^{2}}{2}\right) \frac{U_{j+1}-2 U_{j}+U_{j-1}}{(\Delta x)^{2}}+\left(r-\frac{\sigma^{2}}{2}\right) \frac{U_{j+1}-U_{j-1}}{2 \Delta x}
$$

3. Temporal discretization. For the temporal discretization, we introduce the following three methods:
(a) Forward Euler method:

$$
\frac{U_{j}^{n+1}-U_{j}^{n}}{\Delta t}=\left(Q U^{n}\right)_{j}
$$

(b) Backward Euler method:

$$
\frac{U_{j}^{n+1}-U_{j}^{n}}{\Delta t}=\left(Q U^{n+1}\right)_{j}
$$

(c) Crank-Nicolson method:

$$
\frac{U_{j}^{n+1}-U_{j}^{n}}{\Delta t}=\frac{1}{2}\left[\left(Q U^{n+1}\right)_{j}+\left(Q U^{n}\right)_{j}\right]
$$

### 5.3.2 Binomial method is a forward Euler finite difference method

We choose $x=S / S_{0}$, where $S_{0}$ is the current asset price value. Let $x_{j}=j \Delta x, j=$ $-N, \ldots, N$, where $\Delta x$ is a small parameter satisfying some stability constraint to be shown below. Let us partition the time interval $[t, T]$ into $N$ subintervals uniformally, and let $\Delta t=(T-t) / N$.

For the forward Euler method, we rewrite it as

$$
\begin{aligned}
U_{j}^{n+1} & =U_{j}^{n}+\Delta t\left(Q U^{n}\right)_{j} \\
& =a U_{j-1}^{n}+b U_{j}^{n}+c U_{j+1}^{n}
\end{aligned}
$$

In terms of $V_{j}^{n}$, we have

$$
\begin{aligned}
e^{r \Delta t} V_{j}^{n+1}= & V_{j}^{n}+\frac{1}{2} \frac{\Delta t}{(\Delta x)^{2}} \sigma^{2}\left(V_{j+1}^{n}-2 V_{j}^{n}+V_{j-1}^{n}\right) \\
& +\left(r-\frac{\sigma^{2}}{2}\right) \frac{\Delta t}{2 \Delta x}\left(V_{j+1}^{n}-V_{j-1}^{n}\right) \\
= & {\left[\frac{1}{2} \frac{\Delta t}{(\Delta x)^{2}} \sigma^{2}+\left(r-\frac{\sigma^{2}}{2}\right) \frac{\Delta t}{2 \Delta x}\right] V_{j+1}^{n}+\left(1-\frac{\Delta t}{(\Delta x)^{2}} \sigma^{2}\right) V_{j}^{n} } \\
& +\left[\frac{1}{2} \frac{\Delta t}{(\Delta x)^{2}} \sigma^{2}-\left(r-\frac{\sigma^{2}}{2}\right) \frac{\Delta t}{2 \Delta x}\right] V_{j-1}^{n} . \\
\equiv & a V_{j+1}^{n}+b V_{j}^{n}+c V_{j-1}^{n}
\end{aligned}
$$

We should require $a, b, c \geq 0$ for stability reason. This will be discussed later. Notice that $a+b+c=1$. Thus $V_{j}^{n+1}$ is the "average" of $V_{j-1}^{n}, V_{j}^{n}, V_{j+1}^{n}$ with weight $a, b, c$, then discounted by $e^{-r \Delta t}$.

The stability condition $a, b, c \geq 0$ reads

$$
\left|r-\frac{\sigma^{2}}{2}\right| \frac{\Delta t}{\Delta x} \leq \frac{\Delta t}{(\Delta x)^{2}} \sigma^{2} \leq 1
$$

Next, let us consider a special case: $b=0$. We can choose $\Delta \tau$ and $\Delta x$ properly so that $b=0$. i.e., $1=\frac{\Delta t}{(\Delta x)^{2}} \sigma^{2}$. In this case,

$$
e^{r \Delta t} V_{j}^{n+1}=p V_{j+1}^{n}+(1-p) V_{j-1}^{n}
$$

where $p=a=\frac{1}{2} \frac{\Delta t}{(\Delta x)^{2}} \sigma^{2}+\left(r-\frac{\sigma^{2}}{2}\right) \frac{\Delta t}{2 \Delta x}$. The stability condition is satisfied if and only if

$$
\begin{equation*}
0 \leq p \leq 1 \tag{5.4}
\end{equation*}
$$

We see that this finite difference is identical to the binomial method in the previous section.

### 5.3.3 Stability

Definition 3.3 A finite difference method is called consistent to the corresponding P.D.E. if for any solution of the corresponding P.D.E., it satisfies
F.D.E (finite difference equation) $+\epsilon(\Delta x, \Delta t)$
and $\epsilon \rightarrow 0$ as $\Delta t, \Delta x \rightarrow 0$

Definition 3.4 The truncation error of a finite difference method is defined to be the function $\epsilon(\Delta x, \Delta t)$ in the previous definition.

For instance, the truncation error for central difference is

$$
Q u=\frac{\sigma^{2}}{2} u_{x x}+\left(r-\frac{\sigma^{2}}{2}\right) u_{x}+O\left((\Delta x)^{2}\right)
$$

And the truncation for various temporal discretizations are

1. Forward Euler:

$$
\frac{u(j \Delta x,(n+1) \Delta t)-u(j \Delta x, n \Delta t)}{\Delta t}-(Q u)(j \Delta x, n \Delta t)=O\left((\Delta x)^{2}\right)+O(\Delta t)
$$

2. Backward Euler:

$$
\frac{u(j \Delta x,(n+1) \Delta t)-u(j \Delta x, n \Delta t)}{\Delta t}-(Q u)(j \Delta x,(n+1) \Delta t)=O\left((\Delta x)^{2}\right)+O(\Delta t)
$$

3. Crank-Nicolson method

$$
\begin{aligned}
& \frac{u(j \Delta x,(n+1) \Delta t)-u(j \Delta x, n \Delta t)}{\Delta t}-\frac{1}{2}[(Q u)(j \Delta x,(n+1) \Delta t)+(Q u)(j \Delta x, n \Delta t)] \\
= & O\left((\Delta x)^{2}\right)+O\left((\Delta t)^{2}\right) .
\end{aligned}
$$

The true error $U_{j}^{n}-u(j \Delta x, n \Delta t)$ is usually estimated in terms of the truncation error.
Definition 3.5 A finite difference equation is said to be ( $\left.L^{2}-\right)$ stable if the norm

$$
\left\|U^{n}\right\|^{2}:=\Sigma_{j}\left|U_{j}^{n}\right|^{2} \Delta x
$$

is bounded for all $n \geq 0$.
Definition 3.6 A finite difference method for a P.D.E. is convergent if its solution $U_{j}^{n}$ converges to the solution $u(j \Delta x, n \Delta t)$ of the corresponding P.D.E..

Theorem 5.3 (Lax) : For linear partial differential equations, a finite difference method is convergent if and only if it is consistent and stable.

This theorem is standard and its proof can be found in most numerical analysis text book. We therefore omit it here.

Since the consistency is easily to achieve, we shall focus on the stability issue. A standard method to analyze stability issue is the von Neumann stability analysis. It works for P.D.E. with constant coefficients. It also works "locally" and serves as a necessary condition for linear P.D.E. with variable coefficients and nonlinear P.D.E.. We describe his method below.

We take Fourier transform of $\left\{U_{j}\right\}_{j=-\infty}^{\infty}$ by defining

$$
\hat{U}(\xi)=\sum_{j=-\infty}^{\infty} U_{j} e^{-i j \xi}
$$

It is a well-known fact that

$$
\begin{aligned}
\sum_{j}\left|U_{j}\right|^{2} & =\frac{1}{\sqrt{2 \pi}} \int_{-\pi}^{\pi}|\hat{U}(\xi)|^{2} d \xi \\
& \equiv\|\hat{U}\|^{2}
\end{aligned}
$$

Thus, the boundedness of $\sum_{j}\left|U_{j}\right|^{2}$ can be estimated by using $\|\hat{U}\|^{2}$. The advantage of using $\hat{U}$ is that the finite difference operation becomes a multiplier in terms of $\hat{U}$. Namely,

$$
\begin{aligned}
\widehat{D U}(\xi) & =\sum_{j}\left(\frac{U_{j+1}-U_{j-1}}{2}\right) e^{-i j \xi} \\
& =\sum_{j}\left(\frac{U_{j} e^{i(j+1) \xi}-U_{j} e^{i(j-1) \xi}}{2}\right) \\
& =\left(\frac{e^{i \xi}-e^{-i \xi}}{2}\right) \sum_{j} U_{j} e^{-i j \xi} \\
& =(2 i \sin \xi) \hat{U}(\xi)
\end{aligned}
$$

For the finite difference operator $Q U, \mathrm{~m}$ we have

$$
\begin{aligned}
\widehat{(Q U)} & =\sum_{j}(Q u)_{j} e^{-i j \xi} \\
& =\left[\frac{\sigma^{2}}{2} \frac{1}{(\Delta x)^{2}}(2 \cos \xi-2)+\left(r-\frac{\sigma^{2}}{2}\right) \frac{1}{\Delta x}(2 i \sin \xi)\right] \hat{U} \\
& \equiv \hat{Q}(\xi) \hat{U}(\xi) .
\end{aligned}
$$

For forward Euler method,

$$
\begin{aligned}
\widehat{U^{n+1}}(\xi) & =(1+\Delta t \hat{Q}(\xi)) \widehat{U^{n}} \\
& =G(\xi) \widehat{U^{n}} \\
& =G(\xi)^{n+1} \widehat{U^{n}}
\end{aligned}
$$

We observe that

$$
\begin{aligned}
\int_{-p i}^{\pi}\left|\widehat{U^{n}}(\xi)\right|^{2} d \xi & =\left.\int|G(\xi)|^{2 n} \widehat{U^{0}}(\xi)\right|^{2} d \xi \\
& \leq \max _{\xi \in(-\pi, \pi)}|G(\xi)|^{2 n} \int\left|\widehat{U^{0}}(\xi)\right|^{2} d \xi
\end{aligned}
$$

If $|G(\xi)| \leq 1, \forall \xi \in(-\pi, \pi)$, then stability condition holds. On the other hand, if $|G(\xi)|>$ 1 at some point $\xi_{0}$, then by the continuity of $G$, we have that

$$
|G(\xi)| \geq 1+\epsilon
$$

for some small $\epsilon>0$ and for all $\xi$ with $\left|\xi-\xi_{0}\right| \leq \delta$ for some $\delta>0$. Let consider an initial condition such that

$$
\widehat{U^{0}}(\xi)= \begin{cases}1 & \left|\xi-\xi_{0}\right| \leq \delta \\ 0 & \text { otherwise }\end{cases}
$$

Then the corresponding $\widehat{U^{n}}$ will have

$$
\begin{aligned}
\int_{-p i}^{\pi}\left|\widehat{U^{n}}(\xi)\right|^{2} d \xi & =\left.\int|G(\xi)|^{2 n} \widehat{U^{0}}(\xi)\right|^{2} d \xi \\
& \rightarrow \infty
\end{aligned}
$$

as $n \rightarrow \infty$. We conclude the above discussion by the following theorem.
Theorem 5.4 For a finite difference equation with constant coefficients, suppose its fourier transform satisfies'

$$
\widehat{U^{n+1}}(\xi)=G(\xi) \widehat{U^{n}}(\xi)
$$

Then the finite difference equation is stable if and only if

$$
|G(\xi)| \leq 1 \forall \xi \in(-\pi, \pi] .
$$

Example: Let apply the forward Euler method for the heat equation: $u_{t}=u_{x x}$. Then

$$
\frac{U_{j}^{n+1}-U_{j}^{n}}{\Delta t}=\frac{1}{(\Delta x)^{2}}\left(U_{j+1}^{n}-2 U_{j}^{n}+U_{j-1}^{n}\right), \quad \Longrightarrow U^{n+1}=U^{n}+\frac{\Delta t}{(\Delta x)^{2}} D^{2} U^{n}
$$

From von Neumann analysis:

$$
\begin{aligned}
\widehat{U^{n+1}} & =\left[1+\frac{\Delta t}{(\Delta x)^{2}}(2 \cos \xi-2)\right] \widehat{U^{n}} \\
& =\left(1-4 \frac{\Delta t}{(\Delta x)^{2}} \sin ^{2} \frac{\xi}{2}\right) \widehat{U^{n}} \\
& \equiv G(\xi) \widehat{U^{n}} .
\end{aligned}
$$

Hence

$$
|G(\xi)| \leq 1 \Longrightarrow \frac{\Delta t}{(\Delta x)^{2}} \sin ^{2} \frac{\xi}{2} \leq \frac{1}{2} \Longrightarrow \frac{\Delta t}{(\Delta x)^{2}} \leq \frac{1}{2} \text { (stability condition) }
$$

If we rewrite the finite difference scheme by

$$
\begin{aligned}
U_{j}^{n+1} & =\frac{\Delta t}{(\Delta x)^{2}} U_{j+1}^{n}+\left(1-2 \frac{\Delta t}{(\Delta x)^{2}}\right) U_{j}^{n}+\frac{\Delta t}{(\Delta x)^{2}} U_{j-1}^{n} \\
& \equiv a U_{j+1}^{n}+b U_{j}^{n}+c U_{j-1}^{n}
\end{aligned}
$$

Then the stability condition is equivalent to

$$
a, b, c \geq 0 .
$$

Since we have $a+b+c=1$ from the definition, thus we see that the finite difference scheme is nothing but saying $U_{j}^{n+1}$ is the average of $U_{j+1}^{n}, U_{j}^{n}$ and $U_{j-1}^{n}$ with weights $a, b, c$. In particular, if we choose $\frac{\Delta t}{(\Delta x)^{2}}=\frac{1}{2}$, then $b=0$. If we rename $a=p, c=1-p$, then $U_{j}^{n+1}=p U_{j+1}^{n}+(1-p) U_{j-1}^{n}$. This can be related to the random walk as the follows.

Consider a particle move randomly on the grid points $j \Delta x$. In one time step, the particle moves toward right with probability $p$ and left with probability $1-p$. Let $U_{j}^{n}$ be the probability of the particle at $j \Delta x$ at time step $n$ for a random walk.

$$
U_{j}^{n+1}=p U_{j-1}^{n}+(1-p) U_{j+1}^{n}
$$

We can also apply the above stability analysis to backward Euler method and the CrankNicolson method. Let us only demonstrate the analysis for the heat equation. We left the analysis for the Black-Scholes equations as exercises.

1. For backward Euler method,

$$
\frac{U_{j}^{n+1}-U_{j}^{n}}{\Delta t}=\frac{1}{(\Delta x)^{2}}\left(U_{j-1}^{n+1}-2 U_{j}^{n+1}+U_{j+1}^{n+1}\right)
$$

Then

$$
\left(1+\left(4 \sin ^{2} \frac{\xi}{2}\right) \frac{\Delta t}{(\Delta x)^{2}}\right) \widehat{U^{n+1}}=\widehat{U^{n}} \Longrightarrow \widehat{U^{n+1}}=G(\xi) \widehat{U^{n}}
$$

where

$$
G(\xi)=\frac{1}{1+4 \frac{\Delta t}{(\Delta x)^{2}}\left(\sin ^{2} \frac{\xi}{2}\right)} .
$$

We find that $|G(\xi)| \leq 1$, for all $\xi$. Hence, the backward Euler method is always stable.
2. For Crank-Nicolson method,

$$
\frac{U_{j}^{n+1}-U_{j}^{n}}{\Delta t}=\frac{1}{2(\Delta x)^{2}}\left[\left(U_{j+1}^{n+1}-2 U_{j}^{n+1}+U_{j-1}^{n+1}\right)+\left(U_{j+1}^{n}-2 U_{j}^{n}+U_{j-1}^{n}\right)\right]
$$

Its Fourier transform satisfies

$$
\frac{\widehat{U^{n+1}}-\widehat{U^{n}}}{\Delta t}=\frac{1}{2(\Delta x)^{2}}\left[-\left(4 \sin ^{2} \frac{\xi}{2}\right) \widehat{U^{n+1}}-\left(4 \sin ^{2} \frac{\xi}{2}\right) \widehat{U^{n}}\right]
$$

We have

$$
\left(1+2 \frac{\Delta t}{(\Delta x)^{2}}\left(\sin ^{2} \frac{\xi}{2}\right)\right) \widehat{U^{n+1}}=\left(1-2 \frac{\Delta t}{(\Delta x)^{2}}\left(\sin ^{2} \frac{\xi}{2}\right)\right) \widehat{U^{n}}
$$

and hence

$$
\widehat{U^{n+1}}=\frac{1-2 \frac{\Delta t}{(\Delta x)^{2}} \sin ^{2} \frac{\xi}{2}}{1+2 \frac{\Delta t}{(\Delta x)^{2}} \sin ^{2} \frac{\xi}{2}} \widehat{U^{n}} .
$$

Let $\alpha=2 \frac{\Delta t}{(\Delta x)^{2}} \sin ^{2} \frac{\xi}{2}$, then $G(\xi)=\frac{1-\alpha}{1+\alpha}$. We find that for all $\alpha \geq 0,|G(\xi)| \leq 1$, hence Crank-Nicolson method is always stable for all $\Delta t, \Delta x>0$.

Exercise study the stability criterion for the modified Black-Scholes equation

$$
u_{\tau}=\frac{\sigma^{2}}{2} u_{x x}+\left(r-\frac{\sigma^{2}}{2}\right) u_{x},
$$

for the forward Euler method, back Euler method and Crank-Nicolson method.

### 5.3.4 Convergence

Let us study the convergence for finite difference schemes for the modified Black-Scholes equation. Let us take the forward Euler scheme as our example. The method below can also be applied to other scheme.

The forward Euler scheme is given by:

$$
\frac{U_{j}^{n+1}-U_{j}^{n}}{\Delta t}=\left(Q U^{n}\right)_{j}
$$

We have known that it has first-order truncation error, namely, suppose $u_{j}^{n}:=u(j \Delta x, n \Delta t)$, where $u$ is the solution of the modified Black-Scholes equation, then

$$
\frac{u_{j}^{n+1}-u_{j}^{n}}{\Delta t}=\left(Q u^{n}\right)_{j}+O(\Delta t)+O\left((\Delta x)^{2}\right)
$$

We subtract the above two equations, and let $e_{j}^{n}$ denotes for $u_{j}^{n}-U_{j}^{n}$ and $\epsilon_{j}^{n}$ denotes for the truncation error. Then we obtain

$$
\frac{e_{j}^{n+1}-e_{j}^{n}}{\Delta t}=\left(Q e^{n}\right)_{j}+\epsilon_{j}^{n}
$$

Or equivalently,

$$
\begin{equation*}
e_{j}^{n+1}=a e_{j+1}^{n}+b e_{j}^{n}+c e_{j-1}^{n}+\Delta t \epsilon_{j}^{n} \tag{5.5}
\end{equation*}
$$

Here, $a, b, c \geq 0$ and $a+b+c=1$. We can take Fourier transformation $\widehat{e^{n}}$ of $e^{n}$. It satisfies

$$
\widehat{e^{n+1}}(\xi)=G(\xi) \widehat{e^{n}}(\xi)+\Delta t \widehat{\epsilon^{n}}(\xi)
$$

where

$$
G(\xi)=a e^{i \xi}+b+c e^{-i \xi}
$$

Recall that the stability $|G(\xi)| \leq 1$ is equivalent to $a, b, c \geq 0$. Thus, by applying the above recursive formula, we obtain

$$
\begin{aligned}
\left\|\widehat{e^{n}}\right\| & \leq\left\|\widehat{e^{n-1}}\right\|+\Delta t\left\|\widehat{\epsilon^{n-1}}\right\| \\
& \leq\left\|\widehat{e^{n-2}}\right\|+\Delta t\left(\left\|G \widehat{\epsilon^{n-2}}\right\|+\left\|\widehat{\epsilon^{n-1}}\right\|\right) \\
& \leq\left\|\widehat{e^{n-2}}\right\|+\Delta t\left(\left\|\widehat{\epsilon^{n-2}}\right\|+\left\|\widehat{\epsilon^{n-1}}\right\|\right) \\
& \leq\left\|\widehat{e^{0}}\right\|+\Delta t \sum_{k=0}^{n-1}\left\|\widehat{\epsilon^{k}}\right\| \\
& \leq O(\Delta t)+O\left((\Delta x)^{2}\right) .
\end{aligned}
$$

Here, we have used the estimate for the truncation error

$$
\left\|\epsilon^{n}\right\|=O(\Delta t)+O\left((\Delta x)^{2}\right.
$$

and that $n \Delta t=O(1)$. We conclude the error analysis as the following theorem.
Theorem 5.5 The error $e_{j}^{n}:=u(j \Delta x, n \Delta t)-U_{j}^{n}$ for the Euler method has the following convergence rate estimate:

$$
\left(\sum_{j}\left|e_{j}^{n}\right|^{2} \Delta x\right)^{1 / 2} \leq O(\Delta t)+O\left((\Delta x)^{2}\right), \text { for all } n
$$

It is simpler to ontain the maximum norm estimate. Let $E(n):=\max _{j}\left|e_{j}^{n}\right|$ be the maximum error. From (5.5), we have

$$
\begin{aligned}
\left|e_{j}^{n+1}\right| & \leq a\left|e_{j+1}^{n}\right|+b\left|e_{j}^{n}\right|+c\left|e_{j-1}^{n}\right|+\Delta t\left|\epsilon_{j}^{n}\right| \\
& \leq a E(n)+b E(n)+c E(n)+\Delta t \epsilon \\
& =E(n)+\Delta t \epsilon
\end{aligned}
$$

where

$$
\epsilon:=\max _{j, n}\left|\epsilon_{j}^{n}\right|=O(\Delta t)+O\left((\Delta x)^{2}\right) .
$$

Hence,

$$
E(n+1) \leq E(n)+\Delta t \epsilon .
$$

Since we take $U_{j}^{0}=u_{j}^{0}$, there is no error initially. Hence, we have

$$
\begin{aligned}
E(n) & \leq \sum_{k=0}^{n-1} \Delta t \epsilon \\
& \leq n \Delta t \epsilon
\end{aligned}
$$

Since $n \Delta t$ is a fixed number, as we take the limit $n \rightarrow \infty$, we obtain the error is bounded by the truncation error. We summarize the above discussion as the following theorem.

Theorem 5.6 The error $e_{j}^{n}:=u(j \Delta x, n \Delta t)-U_{j}^{n}$ for the Euler method has the following convergence rate estimate:

$$
\max _{j}\left|e_{j}^{n}\right| \leq O(\Delta t)+O\left((\Delta x)^{2}\right)
$$

Exercise. Prove that the true error of the Crank-Nicolson scheme is $O\left((\Delta t)^{2}\right)+O\left((\Delta x)^{2}\right)$.

### 5.3.5 Boundary condition

For the modified Black-Scholes equation, we have

$$
u(-\infty, t)=0, \quad u(x, t)=e^{x} e^{r \tau} \text { as } x \rightarrow \infty
$$

for a call option, and

$$
u(-\infty, t)=1, \quad u(x, t)=0 \text { as } x \rightarrow \infty
$$

In computation, we can choose a finite domain $\left(x_{L}, x_{R}\right)$ with $x_{L} \ll-1$ and $x_{R} \gg 1$. The boundary condition at the boundary points are an approximation to the above far field boundary condition.

In practice, we don't even use this boundary condition. Indeed, if we want to know $u\left(x_{k}, n \Delta t\right)$, we can find the numerical domain of this quantity, which is the triangle

$$
\{(j \Delta x, m \Delta t)||j-k| \leq n-m\}
$$

We only need to compute $u$ in this domain, which needs no boundary data.

### 5.4 Converting the B-S equation to finite domain

The transformation $x=\log (S / E)$ converts the B-S equation to a heat equation. However, the domain of $x$ is the whole real line. For numerical computation, it is desirable to have a finite computation domain. The transformation in this section converts $S$ to $\xi$ with $\xi \in$ $(0,1)$. The price is that the resulting equation has variable coefficients. But this is not a problem for numerical computation.

We define the transformation:

$$
\begin{align*}
\xi & =\frac{S}{S+E}  \tag{5.6}\\
\bar{V} & =\frac{V(S, t)}{S+E}  \tag{5.7}\\
\tau & =T-t \tag{5.8}
\end{align*}
$$

Notice that $\xi$ is dimensionless and important values of $\xi$ are near $1 / 2$. With this, the inverse transformation is

$$
S=\frac{E \xi}{1-\xi}, \frac{d \xi}{d S}=\frac{(1-\xi)^{2}}{E}
$$

We plug this transformation to the B-S equation. We allow $\sigma$ depend on $S$. Define $\bar{\sigma}(\xi)=$ $\sigma(E \xi / 1-\xi)$. Then the resulting equation is

$$
\begin{equation*}
\frac{\partial \bar{V}}{\partial \tau}=\frac{1}{2} \bar{\sigma}^{2}(\xi) \xi^{2}(1-\xi)^{2} \frac{\partial^{2} \bar{V}}{\partial \xi^{2}}+r \xi(1-\xi) \frac{\partial \bar{V}}{\partial \xi}-r(1-\xi) \bar{V} \tag{5.9}
\end{equation*}
$$

for $0 \leq \xi \leq 1$ and $\tau>0$. The initial data reads

$$
\begin{equation*}
\bar{V}(\xi, 0)=\frac{1-\xi}{E} \Lambda\left(\frac{E \xi}{1-\xi}\right) \tag{5.10}
\end{equation*}
$$

For a call option, the payoff is $\Lambda(S)=\max (S-E, 0)$. The corresponding

$$
\begin{aligned}
\bar{V}(\xi, 0) & =\max (S-E, 0)(1-\xi) / E \\
& =\max (2 \xi-1,0)
\end{aligned}
$$

Similarly, $\bar{V}(\xi, 0)=\max (1-2 \xi, 0)$ for a put option.
On the boundaries $\xi=0$ and $\xi=1$, the diffusion coefficients are degenerate. If the solution is smooth up to the boundaries, then on the boundary, the equation is degenerate to the following ordinary differential equations:

$$
\begin{gathered}
\frac{\partial \bar{V}(0, \tau)}{\partial \tau}=-r \bar{V}(0, \tau) \\
\frac{\partial \bar{V}(1, \tau)}{\partial \tau}=0
\end{gathered}
$$

The corresponding solutions are

$$
\begin{align*}
\bar{V}(0, \tau) & =\bar{V}(0,0) e^{r \tau}  \tag{5.11}\\
\bar{V}(1, \tau) & =\bar{V}(1,0) . \tag{5.12}
\end{align*}
$$

We can discretize equation (5.9) by finite difference method. Let $\Delta \xi$ and $\Delta \tau$ are the spatial and temporal mesh sizes, respectively. Let $\xi_{j}=j \Delta \xi, \tau^{n}=n \Delta \tau$. The boundaries points are $\xi_{0}$ and $\xi_{M}$. We use central difference for $\partial^{2} \bar{V} / \partial \xi^{2}$ and $\partial \bar{V} / \partial \xi$. The resulting finite difference equation reads

$$
\begin{aligned}
\frac{d v_{j}}{d \tau}= & \frac{1}{2} \sigma_{j}^{2} \xi_{j}^{2}\left(1-\xi_{j}\right)^{2} \frac{v_{j+1}-2 v_{j}+v_{j-1}}{\Delta \xi^{2}} \\
& +r \xi_{j}\left(1-\xi_{j}\right) \frac{v_{j+1}-v_{j-1}}{2 \Delta \xi}-r\left(1-\xi_{j}\right) v_{j}
\end{aligned}
$$

We can discretize this equation in the time direction by forward Euler method. The stability constraint is

$$
\begin{equation*}
\left|r \xi_{j}\left(1-\xi_{j}\right)-\frac{1}{2} \sigma_{j}^{2} \xi_{j}^{2}\left(1-\xi_{j}\right)^{2}\right| \frac{\Delta \tau}{\Delta \xi} \leq \sigma_{j}^{2} \xi_{j}^{2}\left(1-\xi_{j}\right)^{2} \frac{\Delta \tau}{2 \Delta \xi^{2}} \leq 1 \tag{5.13}
\end{equation*}
$$

Remark. Many options have non-smooth payoff functions. This causes low order accuracy for finite difference scheme. Fortunately, many simple payoff function has exact solution. For instance, the European call option. For general payoff function, we may subtract its non-smooth part for which an exact solution is available. The remainder is smooth, and a finite difference scheme can yield high-order accuracy.

### 5.5 Fast algorithms for solving linear systems

In the backward Euler method and the Crank-Nicolson method, we need to solve linear systems of the form

$$
A U=F .
$$

For the backward Euler scheme,

$$
A=\operatorname{diag}(-a, 1+a+c,-c)
$$

$$
:=\left(\begin{array}{ccccccc}
1+a+c & -c & 0 & \cdots & & & \\
-a & 1+a+c & -c & 0 & \cdots & & \\
0 & -a & 1+a+c & -c & 0 & \cdots & \\
& \ddots & \ddots & \ddots & \ddots & & \\
& \cdots & 0 & -a & 1+a+c & -c & 0 \\
& & \cdots & 0 & -a & 1+a+c & -c \\
& & & \cdots & 0 & -a & 1+a+c
\end{array}\right)
$$

and

$$
U=\left(\begin{array}{c}
U_{j_{L}+1}^{n} \\
\vdots \\
U_{j_{R}-1}^{n}
\end{array}\right), F=\left(\begin{array}{c}
a U_{j_{L}}^{n+1} \\
\vdots \\
c U_{j_{R}}^{n+1}
\end{array}\right)
$$

For the Crank-Nicolson scheme We have

$$
A U^{n+1}=B U^{n}+b^{n+1 / 2}
$$

where $A=\operatorname{diag}\left(-\frac{a}{2}, 1+\frac{a}{2}+\frac{c}{2},-\frac{c}{2}\right) B=\operatorname{diag}\left(\frac{a}{2}, 1-\frac{a}{2}-\frac{c}{2}, \frac{c}{2}\right)$, and

$$
b^{n+1 / 2}=\left(\begin{array}{c}
a \frac{U_{j_{L}}^{n+1}+U_{j_{L}}^{n}}{2} \\
0 \\
\vdots \\
0 \\
c \frac{U_{j_{R}}^{n+1}+U_{j_{R}}^{n}}{2}
\end{array}\right) .
$$

Now, we concentrate on solving the linear system

$$
A x=f
$$

The matrix $A$ is tridiagonal and diagonally dominant. Let us rewrite $A=\operatorname{diag}(a, b, c)$. Here, the constants $a, b, c$ are different from the average weights we had before. We may assume $b>0$. We say that $A$ is diagonally dominant if $b>|a|+|c|$. More generally, $A$ may takes the form $A=\operatorname{diag}\left(a_{j}, b_{j}, c_{j}\right)$. and $\left|b_{j}\right|>|a|+\left|c_{j}\right|$. Without loss of generality, we may normalize the $j-t h$ so that $b_{j}=1$.

There are two classes of methods to solve the above linear systems. One is called direct methods, the other is called iterative methods. For one-dimensional case as we have here, direct method is usually better. However, for high-dimensional cases, iterative methods are better.

### 5.5.1 Direct methods

## Gaussian elimination

Let us illustrate this method by the simple example: $A=\operatorname{diag}(a, 1, c)$. We multiple the first equation by $-a$ and add it into the second equation to eliminate the term $x_{j_{L}+1}$ in the
second equation. Then the resulting equation becomes

$$
\left(\begin{array}{cccccc}
1 & c & 0 & 0 & \cdots & 0 \\
0 & 1-a c & c & 0 & \cdots & 0 \\
0 & a & 1 & c & \cdots & 0 \\
& & & & \vdots & 0 \\
& & & & \cdots & 1
\end{array}\right)\left(\begin{array}{c}
x_{j_{L}+1} \\
x_{j_{L}+2} \\
x_{j_{L}+3} \\
\vdots \\
x_{j_{R}-1}
\end{array}\right)=\left(\begin{array}{c}
b_{j_{L}+1} \\
-a b_{j_{L}+1}+b_{j_{L}+2} \\
b_{j_{L}+3} \\
\vdots \\
b_{j_{R}-1}
\end{array}\right)
$$

We continue to eliminate the term $a$ in the third equation, and so on. Finally, we arrive

$$
\left(\begin{array}{cccccc}
1 & c & 0 & 0 & \cdots & 0 \\
0 & 1-a c & c & 0 & \cdots & 0 \\
0 & 0 & 1-c /(1-a c) & c & \cdots & 0 \\
0 & 0 & 0 & & \vdots & 0 \\
0 & 0 & 0 & 0 & \cdots & 1
\end{array}\right)\left(\begin{array}{c}
x_{j_{L}+1} \\
x_{j_{L}+2} \\
x_{j_{L}+3} \\
\vdots \\
x_{j_{R}-1}
\end{array}\right)=\left(\begin{array}{c}
b_{j^{\prime}+1}^{\prime} \\
b_{j_{j}+2}^{\prime} \\
b_{j_{L}+3}^{L} \\
\vdots \\
b_{j_{R}-1}^{\prime}
\end{array}\right)
$$

Then $x_{j}$ can be solved easily. The diagonal dominance condition guarantee that the reduced matrix is also diagonally dominant. Thus, this scheme is numerical stable.

## LU decomposition

We decompose $A=L U$, where

$$
L=\left(\begin{array}{ccccc}
1 & 0 & 0 & \cdots & 0 \\
\ell_{j_{L}+2} & 1 & \ddots & & \vdots \\
0 & \ddots & \ddots & \ddots & 0 \\
\vdots & & \ddots & \ddots 0 \\
0 & \cdots 0 & \ell_{j_{R}-1} & 1 &
\end{array}\right), U=\left(\begin{array}{ccccc}
u_{j_{L}+1} & v_{j_{L}+1} & 0 & \cdots & 0 \\
0 & u_{j_{L}+2} & \ddots & & \vdots \\
0 & \ddots & \ddots & \ddots & 0 \\
\vdots & & \ddots & \ddots & v_{j_{R}-2} \\
0 & \cdots 0 & 0 & u_{j_{R}-1} &
\end{array}\right)
$$

It is easy to find a recursion formula to find the coefficients $\ell, u$ and $v$ 's. Once these are found, we can find $x$ by solving

$$
L y=b, U x=y
$$

These two equations are easy to solve. One can show that both $L$ and $U$ are diagonally dominant if $A$ is.

If we watch carefully, LU-decomposition is equivalent to the Gaussian elimination.

## Cyclic reduction method

Let us take the case $A=\operatorname{diag}(a, 1, c)$ to illustrate this method. Consider three consecutive equations

$$
\begin{aligned}
a x_{2 j-2}+x_{2 j-1}+c x_{2 j} & =b_{2 j-1} \\
a x_{2 j-1}+x_{2 j}+c x_{2 j+1} & =b_{2 j} \\
a x_{2 j}+x_{2 j+1}+c x_{2 j+2} & =b_{2 j+1}
\end{aligned}
$$

We can eliminate the odd-index terms $x_{2 j-1}$ and $x_{2 j+1}$. Namely, $-a \times(2 j-1)-$ eq $+(2 j)-$ eq $-c \times(2 j+1)$-eq: After normalization, we obtain

$$
a^{\prime} x_{2 j-2}+x_{2 j}+c^{\prime} x_{2 j+2}=b_{j}^{\prime}
$$

Here,

$$
\begin{gathered}
a^{\prime}=-\frac{a^{2}}{1-2 a c}, c^{\prime}=-\frac{c^{2}}{1-2 a c}, \\
b_{j}^{\prime}=\left(b_{2 j}-a b_{2 j-1}-c b_{2 j+1}\right) /(1-2 a c) .
\end{gathered}
$$

If we rename $x_{j}^{\prime}=x_{2 j}$. Then we have $A^{\prime} x^{\prime}=b^{\prime}$, where $A^{\prime}=\operatorname{diag}\left(a^{\prime}, 1, c^{\prime}\right)$. Notice that the system is reduced to half and with the same form. One can show that the iterative mapping

$$
\binom{a}{c} \mapsto\binom{a^{\prime}}{c^{\prime}}
$$

converges to $(0,0)^{t}$ quadratically fast, provided $|a|+|c|<1$ initially. Thus, for few iteration, the matrix $A$ is almost an identity matrix. We can invert it trivially. Once $x_{2 j}$ are found, the odd-index $x+2 j+1$ can be found from the equation:

$$
a x_{2 j}+x_{2 j+1}+c x_{2 j+2}=b_{2 j+1} .
$$

A careful reader should find that the cyclic reduction is also a version of the Gaussian elimination method.

### 5.5.2 Iterative methods

Most iterative methods can be viewed as a proper decomposition of $A$, then solve an important and treat the rest as a perturbation term.

## Jacob method

In Jacobi method, wer decompose

$$
A=D+B
$$

where $D$ is the diagonal part and $B$ is the off diagonal part. Since $A$ is diagonally dominant, we may approximate $x$ by the sequence $x^{n}$, where $x^{n}$ is defined by the following iteration scheme:

$$
D x^{n+1}+B x^{n}=b .
$$

Let the error $e^{n}:=x^{n+1}-x^{n}$. Then

$$
D e^{n}=-B e^{n-1}
$$

Or

$$
e^{n}=-D^{-1} B e^{n-1}
$$

Let us define the maximum norm

$$
\left\|e^{n}\right\|:=\max _{j}\left|e_{j}^{n}\right|
$$

Then

$$
\begin{aligned}
\left|e_{j}^{n}\right| & =\left|-\frac{a}{b} e_{j-1}^{n-1}-\frac{c}{b} e_{j+1}^{n-1}\right| \\
& \leq \left\lvert\, \frac{|a|}{|b|}\left\|e^{n-1}\right\|+\frac{|c|}{|b|}\left\|e^{n-1}\right\|\right. \\
& =\frac{|a|+|c|}{|b|}\left\|e^{n-1}\right\|
\end{aligned}
$$

Hence,

$$
\left\|e^{n}\right\| \leq \rho\left\|e^{n-1}\right\|
$$

where $\rho=\frac{|a|+|c|}{|b|}<1$ from the fact that $A$ is diagonally dominant. This yields the convergence of the sequence $x_{n}$. The limit $x$ satisfies the equation $A x=b$.

## Gauss-Seidel method

In Gauss-Seidel method, $A$ is decomposed into $A=(D+L)+U$, where $D$ is the diagonal part, $L$, the lower triangular part, and $U$, the upper triangular part of $A$. The approximate solution sequence is given by

$$
(D+L) x^{n+1}+U x^{n}=b .
$$

As before, the error $e^{n}:=x^{n+1}-x^{n}$ satisfies

$$
e^{n}=-(D+L)^{-1} U e^{n-1}
$$

To analyze the decay of $e^{n}$, we use Fourier method. Let

$$
\widehat{e^{n}}(\xi):=\sum_{j} e_{j}^{n} e^{-i j \xi} .
$$

Then we have

$$
\begin{aligned}
\widehat{e^{n}}(\xi) & =G(\xi) \widehat{e^{n-1}} \\
G(\xi) & =-\frac{c e^{i \xi}}{b+a e^{-i \xi}}
\end{aligned}
$$

It is easy to see that the amplification matrix $G$ satisfies

$$
\begin{equation*}
\max _{\xi}|G(\xi)|:=\rho<1, \text { provided }|b|>|a|+|c| . \tag{5.14}
\end{equation*}
$$

This shows that the Gauss-Seidel method also converges for diagonally dominant matrix. Exercise. Show the above statement (5.14).

## Successive over-relaxation method (SOR)

In the methods of Jacobi and Gauss-Seidel, the approximate sequence $x^{n}$ is usually convergent monotonely. We therefore have a chance to speed them up by an extrapolation procedure described below.

$$
\begin{aligned}
& y^{n+1}=-D^{-1}(L+U) x^{n}+b \\
& x^{n+1}=x^{n}+\omega\left(y^{n+1}-x^{n}\right)
\end{aligned}
$$

Here, $\omega$ is a parameter. In order to speed up, we require $\omega>1$. We also need to require $\omega<2$ for stability. The optimal $\omega$ is chosen to minimize the amplification matrix $G_{\omega}(\xi)$.
Exercise. Find the amplification matrix $G_{\omega}$ and the optimal $\omega$ for the matrix $A=\operatorname{diag}(a, b, c)$.
Also, determine the rate

$$
\rho:=\min _{\omega} \max _{\xi}\left|G_{\omega}(\xi)\right| .
$$

## Multigrid method

Probably the most powerful method in higher dimension is the multigrid method.

## Chapter 6

## American Option

### 6.1 Introduction

An American option has the right to exercise any time during the life of the option. The first important thing we should note is that the value of an American option is greater than or equal to the payoff function: $V(S, t) \geq \Lambda(S, t)$. Otherwise, there is an arbitrage opportunity because we can buy the American option then sell it immediately to gain a net profit $V-\Lambda$.

We recall that the value of an American call option is equal to that of a European call option. However, for other cases like the the American put option or the American call option on dividend-paying asset, the American options do cost more. We explain why it is so below. The figure below is the value of a European put.


Notice that $P(S, t)<\max \{E-S, 0\}$ in some region in the $S$ - $t$ plane. In this region, the corresponding American option must be higher than the European option, otherwise for $S$, we can buy a put $P(S, t)$, then exercise it immediately. We make a riskless profit: $E-P-S>0$. Another example is the American call option on a dividend-paying asset. Its value is shown in the Figure below.

$C(S, t) \sim S e^{-D_{0}(T-t)}$ for $S \gg 1$, hence $C(S, t)<\max (S-E, 0)$ for some $S_{f}(t)$.
Since $C(S, t) \sim S e^{-D_{0}(T-t)}$ for large $S$, there is a region in $(S, t)$-plane where $C(S, t)<$ $\max \{S-E, 0\}$. In this region, if we could exercise the call option, then based on the same argument above, there would be an arbitrage opportunity. Hence the corresponding American call option should also satisfy

$$
C(S, t) \geq \max \{S-E, 0\}
$$

### 6.2 American options as a free boundary value problem

### 6.2.1 American put option

We can view an American option as a free boundary value problem. Let us take the American put option as an example.

First, there must be some value of $S$ for which it is optimal from the holder's point of view to exercise the American option. Otherwise, we should hold the option for all possible $S$. Then this option is identical to a European option. But we have seen that this is not the case. In other words, there is a $S_{f}(t)$, if $S<S_{f}(t)$, one should exercise the put option, which maximize the payoff function $E-S$. And for $S>S_{f}(t)$, we should hold the
option. This $S_{f}(t)$ is referred as the optimal exercise price. In other word, we should have $P(S, t) \equiv \max (E-S, 0)$ for $S<S_{f}(t)$, and $P(S, t)$ satisfies the Black-Sholes equation for $S>S_{f}(t)$.

However, we do not know $S_{f}(t)$ a priori. We should treat $S_{f}(t)$ as a new unknown (called free boundary and we should impose boundary condition to determine it.) We claim that the proper boundary condition on $S_{f}(t)$ are

1. $P(S, t)$ is continuous across $\left(S_{f}(t), t\right)$,
2. $\partial P(S, t) / \partial S$ is also continuous across $\left(S_{f}(t), t\right)$.

Remark. For $S<S_{f}(t)$, we should exercise the American put option because the corresponding payoff $\Lambda=\max \{E-S, 0\}$ is higher. Thus, for $S<S_{f}(t)$, the value of the put option should be the payoff function $\Lambda=E-S$. Its derivative in $S$ is -1 . Thus, the second boundary condition is equivalent to saying that $\frac{\partial P}{\partial S}\left(S_{f}(t), t\right)$ is continuous across $\left(S_{f}(t), t\right)$.

## Reasons.

1. If $S(t)=S_{f}(t)$ then $S(t+\Delta t)>S_{f}(t)$ with probability 1 . This follows from $\frac{d S}{S}=$ $\mu d t+\sigma d z$ and $\mu>0$. If $P$ is discontinuous across $\left(S_{f}(t), t\right)$, then $P(S(t+\Delta t), t+$ $\Delta t) \neq P(S(t), t)$ with probability 1 . This would make an arbitrage opportunity by buying $P$ at $t$ then selling it at $t+\Delta t$.
2. We prove this by contradiction.
(a) If $\frac{\partial P}{\partial S}\left(S_{f}(t), t\right)<-1$ then as $S$ increases from $S_{f}(t), P(S, t)$ drops below the payoff $E-S$, this contradicts to $P(S, t) \geq \max \{E-S, 0\}$. At $S_{f}(t)$, $P\left(S_{f}(t), t\right)=E-S_{f}(t)$.
(b) Suppose $\frac{\partial P}{\partial S}\left(S_{f}(t), t\right)>-1$. First, for $S>S_{f}(t), P$ satisfies the Black-Scholes equation (for put option) and its solution curve should lie above the payoff function $E-S$ on the $(S, P)$-plane with $P\left(S_{f}(t), t\right)=E-S_{f}(t)$. This curve moves up as $S_{f}(t)$ moves down, and the corresponding $\frac{\partial P}{\partial S}\left(S_{f}(t), t\right)$ decreases. If $\frac{\partial P}{\partial S}\left(S_{f}(t), t\right)>-1$, then we can move down $S_{f}(t)$ to another $\tilde{S}_{f}(t)<S_{f}(t)$ where $\frac{\partial P}{\partial S}\left(\tilde{S}_{f}(t), t\right)=-1$. In this movement, the curve stays above the payoff function. Now, if we exercise the put option for $S \leq \tilde{S}_{f}(t)$, the payoff $E-\tilde{S}_{f}(t)$ is higher than $E-S_{f}(t)$. This means that $S_{f}(t)$ is not the optimal exercise price. This is a contradiction.

Thus, we treat the American put option as the following free boundary value problem. There exists an optimal exercise price $S_{f}(t)$ such that

1. for $S<S_{f}(t)$, early exercise is optimal, and $P(S, t)=E-S$;
2. for $S>S_{f}(t)$, one should hold the put option and $P$ satisfies the Black-Scholes equation:

$$
\frac{\partial P}{\partial t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} P}{\partial S^{2}}+r S \frac{\partial P}{\partial S}-r P=0
$$

3. across the free boundary $\left(S_{f}(t), t\right)$, both $P$ and $\frac{\partial P}{\partial S}$ are continuous.

### 6.2.2 American call option on a dividend-paying asset

As we have seen in the introduction of this chapter that an American call option $C(S, t)$ on a dividend-paying asset has asymptotic value $C(S, t) \sim S e^{-D_{0}(T-t)}$ for large $S$. This value is below the payoff function $\Lambda \equiv \max (S-E, 0)$. Therefore, there must an optimal $S_{f}(t)$ such that we should exercise this call option when $S>S_{f}(t)$ and hold it when $S<S_{f}(t)$. On the free boundary $S=S_{f}(t)$, based on the no-arbitragy hypothesis, we should have both $C(S, t)$ and $\partial C(S, t) / \partial S$ are continuous the free boundary $(S(t), t)$ for $0<t<T$. We summarize this by the following equations

$$
\begin{gathered}
C_{t}+\frac{\sigma^{2}}{2} S^{2} \frac{\partial^{2} C}{\partial S^{2}}+\left(r-D_{0}\right) S \frac{\partial C}{\partial S}-r C=0,0<S<S_{f}(t) \\
C(S, t) \equiv \Lambda(S) \equiv \max \{S-E, 0\}, S>S_{f}(t)
\end{gathered}
$$

On the free boundary $S=S_{f}(t)$, the boundary condition is required

$$
\begin{gathered}
C\left(S_{f}(t), t\right)=S_{f}(t)-E \\
\frac{\partial C}{\partial S}\left(S_{f}(t), t\right)=1,0 \leq t \leq T
\end{gathered}
$$

### 6.3 American option as a linear complementary problem

The American option can also be formulated as a linear complementary problem, where the free boundary is treated implicitly. To illustrate this linear complementary problem, first we notice that an American option should satisfy the following conditions:
(i) $V \geq \Lambda$,
(ii) $V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}} \leq r\left(V-S \frac{\partial V}{\partial S}\right)$,
(iii) either $V=\Lambda$, or $V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}=r\left(V-S \frac{\partial V}{\partial S}\right)$ should hold,
(iv) both $V$ and $\frac{\partial V}{\partial S}$ are continuous.

Here, $V$ is the value of the American option, $\Lambda$ is the corresponding payoff function.
We have seen the reasons for (i), and (iv). We explain the reasons of (ii) below. Let us consider the portfolio, $\Pi=V-\Delta S$. As we have seen that the Delta hedge eliminate the randomness of $\Pi$ and yields

$$
d \Pi=V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}
$$

When it is optimal to hold the option, then

$$
d \Pi=r\left(V-\frac{\partial V}{\partial S} S\right)
$$

Otherwise, we should have

$$
d \Pi \leq r\left(V-\frac{\partial V}{\partial S} S\right)
$$

based on no arbitrage opportunities. Thus, the Black-Scholes is replaced by the BlackScholes inequality.

To show (iii), we know that if we exercise the option, then $V=\Lambda$, otherwise, we hold the option and its value should satisfy the Black-Scholes equation.

Properties (i)-(iv) can be formulated as the following linear complementary problem:
(i) $V-\Lambda \geq 0$,
(ii) $V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}-r\left(V-S \frac{\partial V}{\partial S}\right) \leq 0$,
(iii) $(V-\Lambda)\left(V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}-r\left(V-S \frac{\partial V}{\partial S}\right)\right)=0$,
(iv) both $V$ and $\frac{\partial V}{\partial S}$ are continuous.

Such a problem is called a linear complementary problem. The advantage of this formulation is that the free boundary is treated implicitly.

We can reformulate this problem in terms of $x$ variable. As before, we use the following change of variables: $V=E v, S=E e^{x}, \tau=T-t$. The free boundary now in $x$-variable is $x_{f}(t)$. The free boundary value problem is formulated as
(i) for $-\infty<x<x_{f}(t), v=1-e^{x}$, and

$$
v_{\tau}-\frac{\sigma^{2}}{2} v_{x x}-\left(r-\frac{\sigma^{2}}{2}\right) v_{x}+r v \geq 0
$$

(ii) for $x_{f}(t)<x<\infty, v>1-e^{x}$, and

$$
v_{\tau}-\frac{\sigma^{2}}{2} v_{x x}-\left(r-\frac{\sigma^{2}}{2}\right) v_{x}+r v=0
$$

(iii) both $v$ and $\frac{\partial v}{\partial x}$ are continuous.

The linear complementary problem is formulated as:
(i) $v-\left(1-e^{x}\right) \geq 0$,
(ii) $v_{\tau}-\frac{\sigma^{2}}{2} v_{x x}-\left(r-\frac{\sigma^{2}}{2}\right) v_{x}+r v \geq 0$,
(iii) $\left(v_{\tau}-\frac{\sigma^{2}}{2} v_{x x}-\left(r-\frac{\sigma^{2}}{2}\right) v_{x}+r v\right)\left(v-\left(1-e^{x}\right)\right)=0$,
(iv) $v$ and $v_{x}$ are continuous.
with initial condition $v(x, 0)=\Lambda(x)=1-e^{x}$.
We may replace $v$ by $u e^{-r \tau}$ to eliminate the term $r v$. Then we have

$$
\begin{aligned}
\left(u_{\tau}-\frac{\sigma^{2}}{2} u_{x x}-\left(r-\frac{\sigma^{2}}{2}\right) u_{x}\right)(u-g) & =0 \\
u_{\tau}-\frac{\sigma^{2}}{2} u_{x x}-\left(r-\frac{\sigma^{2}}{2}\right) u_{x} & \geq 0
\end{aligned}
$$

$$
\begin{aligned}
u-g & \geq 0 \\
u(x, 0) & =g(x, 0)
\end{aligned}
$$

with $u, u_{x}$ being continuous. Here $g(x, \tau)=\max \left\{e^{r \tau}\left(1-e^{x}\right), 0\right\}$. The far field boundary conditions are

$$
u(x, \tau) \rightarrow 0, \text { as } x \rightarrow \infty, \quad u(x, \tau) \rightarrow e^{r \tau}, \text { as } x \rightarrow-\infty
$$

A mathematical theory called parabolic variational inequality gives construction, existence, uniqueness of the solution. (see reference: A. Friedman, Variational Inequality). Let us demonstrate this theory briefly. The method we shall use is called the penalty method. Let us consider the following penalty function:

$$
\phi_{N}(v)=-e^{-N v} .
$$

It has the properties: (i) $\phi_{N}>0$, (ii) $\phi_{N}(v) \rightarrow 0$ whenever $v>0$.. We consider the following penalized P.D.E.:

$$
\begin{aligned}
u_{\tau}-\frac{1}{2} \sigma^{2} u_{x x}-\left(r-\frac{\sigma^{2}}{2}\right) u_{x}+\phi_{N}(u-g) & =0 \\
u(x, 0) & =g(x, 0),
\end{aligned}
$$

with $u \rightarrow 0$ as $x \rightarrow+\infty, \quad u \rightarrow e^{r \tau}$ as $x \rightarrow-\infty$.
From a standard theory of nonlinear P.D.E. (by monotone method, for instance), one can show that the solution $u_{N}$ exists for all $N>0$. We then need an estimate for $u_{N}$ and $\frac{\partial u_{N}}{\partial x}$. The boundedness of these two gives that $u_{N}$ has a convergent subsequence, say $u_{N_{i}}$ such that

$$
u_{N_{i}} \rightarrow u,
$$

with $u_{N_{i}}, u, \frac{\partial}{\partial x} u_{N_{i}}, u_{x}$ being continuous. Moreover,

$$
\left|\phi_{N_{i}}\left(u_{N_{i}}-g\right)\right| \leq \text { constant }
$$

As $N_{i} \rightarrow \infty$, we conclude $u-g \geq 0$. Further, on the set $\{u-g>0\}, \phi_{N_{i}}\left(u_{N_{i}}-g\right) \rightarrow 0$. Hence we have

$$
u_{\tau}-\frac{1}{2} u_{x x}-\left(r-\frac{\sigma^{2}}{2}\right) u_{x}=0 \text { on }\{u-g>0\} .
$$

### 6.4 Numerical Methods

### 6.4.1 Projective method for American put

We recall that the solution of the Black-Scholes equation can be discretized by the following binomial method (or the forward Euler method):

$$
e^{r \Delta} V_{j}^{n}=p V_{j+1}^{n+1}+q V_{j-1}^{n+1}, \quad p, q \geq 0, \quad p+q=1
$$

Similarly, the linear complementary problem can be discretized as

$$
\left(e^{r \Delta t} V_{j}^{n}-p V_{j+1}^{n+1}-q V_{j-1}^{n+1}\right)\left(V_{j}^{n}-h_{j}^{n}\right)=0
$$

$$
\begin{aligned}
e^{r \Delta t} V_{j}^{n}-p V_{j+1}^{n+1}-q V_{j-1}^{n+1} & \geq 0, \quad V_{j}^{n}-\Lambda_{j}^{n} \geq 0 \\
V_{j}^{N} & =\Lambda_{j}^{N}
\end{aligned}
$$

Here, $\Lambda_{j}^{n}$ is the discretized payoff function after changing variable.
This discretized linear complementary problem can be solved by the following projected forward Euler method. Define

$$
V_{j}^{n}=\max \left\{e^{-r \Delta t}\left(p V_{j+1}^{n+1}+q V_{j-1}^{n+1}\right), \Lambda_{j}^{n}\right\} .
$$

One can show that this method converges. (The main tool to prove this is a theory for monotone operator. One can show that the scheme is monotone, $\|V\|_{\infty},\left\|V_{x}\right\|_{\infty}$ are bounded. Reference. Majda \& Crandell, Math. Comp..) Furthermore, from the construction, we have $V_{j}^{n} \geq h_{j}^{n}$. Thus the limiting function satisfies $V \geq \Lambda$. At those points $V(S, t)>\Lambda(S, t)$, we have $V_{j}^{n}>\Lambda_{j}^{n}$, for large $n$, where $n \Delta t \sim t$ and $E e^{j \Delta x} \sim S$. In this case, we always have

$$
V_{j}^{n}=e^{-r \Delta t}\left(p V_{j+1}^{n+1}+q V_{j-1}^{n+1}\right)
$$

Hence, the limiting function satisfies the Black-Scholes equation whenever $V(S, t)>$ $\Lambda(S, t)$. The regularity result (i.e. continuity of $V$ and $V_{S}$ ) follows from the theory of monotone operator.

### 6.4.2 Projective method for American call

The linear complementary problem for this American call option is
(i) $V(S, t) \geq \Lambda(S) \equiv \max \{S-E, 0\}$
(ii) $V_{t}+\frac{\sigma^{2}}{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}+\left(r-D_{0}\right) S \frac{\partial V}{\partial S}-r V \leq 0$,
(iii) $\left(V_{t}+\frac{\sigma^{2}}{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}+\left(r-D_{0}\right) S \frac{\partial V}{\partial S}-r V\right)(V-\Lambda)=0$.
(iv) $V$ and $V_{S}$ are continuous.

The binomial approximation for the B-S equation is

$$
e^{r \Delta t} V_{j}^{n}=p V_{j+1}^{n+1}+q V_{j-1}^{n+1}
$$

where

$$
p=\frac{\sigma^{2}}{2} \frac{\Delta t}{(\Delta x)^{2}}+\left(r-D_{0}-\frac{\sigma^{2}}{2}\right) \frac{\Delta t}{2 \Delta x}, \quad q=\frac{\sigma^{2}}{2} \frac{\Delta t}{(\Delta x)^{2}}-\left(r-D_{0}-\frac{\sigma^{2}}{2}\right) \frac{\Delta t}{2 \Delta x},
$$

and $p+q=1$. We choose $\Delta t$ and $\Delta x$ so that $p>0$ and $q>0$. For American option, $V$ has to be greater than $\Lambda(S, t)$, the payoff function at time $t$. Hence, we should require

$$
V_{j}^{n}=\max \left\{e^{-r \Delta t}\left(p V_{j+1}^{n+1}+q V_{j-1}^{n+1}\right), \Lambda_{j}^{n}\right\} .
$$

The above is the projective forward Euler method.

For the corresponding binomial model, first we determine the up/down ratios $u$ and $d$ for the riskless asset price by

$$
p u+(1-p) d=e^{\left(r-D_{0}\right) \Delta t}, \quad p u^{2}+(1-p) d^{2}=e^{\left(2\left(r-D_{0}\right)+\sigma^{2}\right) \Delta t}
$$

or equivalently,

$$
\begin{aligned}
u & =A+\sqrt{A^{2}-1}, d=1 / u \\
A & =\frac{1}{2}\left(e^{-\left(r-D_{0}\right) \Delta t}+e^{\left(r-D_{0}+\sigma^{2}\right) \Delta t}\right) \\
p & =\frac{e^{\left(r-D_{0}\right) \Delta t}-d}{u-d}, q=1-p
\end{aligned}
$$

Then the binomial model is given by

$$
\begin{aligned}
S_{j}^{n} & = \begin{cases}u S_{j-1}^{n-1}, & \text { with probability } p \\
d S_{j+1}^{n-1}, & \text { with probability } q\end{cases} \\
S_{0}^{0} & =S .
\end{aligned}
$$

and

$$
V_{j}^{n}=\max \left\{e^{-r \Delta t}\left(p V_{j+1}^{n+1}+q V_{j-1}^{n+1}\right), S_{j}^{n}-E\right\} .
$$

### 6.4.3 Implicit method

For implicit method like backward Euler or Crank-Nicolson method, we need to add the constraints $u^{n+1} \geq g^{n+1}$ for American option. It is important to know that if an iterative method is used, then we should require this condition hold in each iteration steps. For instance, in the SOR iteration method,

$$
\begin{aligned}
V^{n,(k+1)} & =\max \left\{V^{n,(k)}+\omega\left(y^{n,(k+1)}-V^{n,(k)}\right), \Lambda^{n}\right\}, \\
y^{n,(k+1)} & =(D+L)^{-1}\left(-U V^{n,(k)}+f^{n+1}\right), k=0, \cdots K \\
V^{n} & \equiv V^{n,(K)}
\end{aligned}
$$

This guarantees that $V^{n+1} \geq \Lambda^{n+1}$.

### 6.5 Converting American option to a fixed domain problem

### 6.5.1 American call option with dividend paying asset

We consider the American call option on a dividend paying asset:

$$
\begin{gathered}
V_{t}+\frac{\sigma^{2}}{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}+\left(r-D_{0}\right) S \frac{\partial V}{\partial S}-r V=0 \\
V(S, t) \geq \Lambda(S) \equiv \max \{S-E, 0\}
\end{gathered}
$$

$$
0 \leq S \leq S_{f}(t), 0 \leq t \leq T
$$

where $S_{f}(t)$ is the free boundary. On this free boundary, the boundary condition is required

$$
\begin{aligned}
V\left(S_{f}(t), t\right) & =S_{f}(t)-E \\
\frac{\partial V}{\partial S}\left(S_{f}(t), t\right) & =1,0 \leq t \leq T
\end{aligned}
$$

We also need a condition for $S_{f}$ at final time

$$
S_{f}(T)=\max \left(E, r E / D_{0}\right)
$$

Before converting the problem, we first remove the singularity of the final data (i.e. nonsmoothness of the payoff function) as the follows. We may substract $V$ by an European call option $c$ with the same payoff data. Notice that $c(S, t)$ has exact solution. The new variable $V-c$ satisfies the same equation, yet it has smooth final data.

To convert the free boundary problem to a fixed domain problem, we introduce the following change-of-variables:

$$
\begin{cases}\xi & =S / S_{f}(t) \\ \tau & =T-t \\ u(\xi, \tau) & =(V(S, t)-c(S, t)) / E \\ s_{f}(\tau) & =S_{f}(t) / E\end{cases}
$$

The new equations for these new variables are

$$
\begin{cases}\frac{\partial u}{\partial \tau}=\frac{\sigma^{2}}{2} \xi^{2} \frac{\partial^{2} u}{\partial \xi^{2}}+\left(\left(r-D_{0}\right)+\frac{1}{s_{f}} \frac{d s_{f}}{d \tau}\right) \xi \frac{\partial u}{\partial \xi}-r u, & 0 \leq \xi \leq 1,  \tag{6.1}\\ u(\xi, 0)=0, & 0 \leq \tau \leq T \\ u(1, \tau)=g\left(s_{f}(\tau), \tau\right), & 0 \leq \xi \leq 1, \\ \frac{\partial u}{\partial \xi}(1, \tau)=h\left(s_{f}(\tau), \tau\right), & 0 \leq \tau \leq T \\ s_{f}(0)=\max \left(1, r / D_{0}\right), & 0 \leq \tau \leq T\end{cases}
$$

where

$$
\begin{aligned}
g\left(s_{f}(\tau), \tau\right) & =s_{f}(\tau)-1-c\left(E s_{f}(\tau), T-\tau\right) \\
h\left(s_{f}(\tau), \tau\right) & =s_{f}(\tau)\left[1-\frac{\partial c}{\partial S}\left(E s_{f}(\tau), T-\tau\right)\right]
\end{aligned}
$$

At the boundary $\xi=0$, the Black-Sholes equation is degenerate to

$$
\frac{\partial u}{\partial \tau}=-r u
$$

With the trivial initial condition yields

$$
u(0, \tau)=0,0 \leq \tau \leq T
$$

In practice, we can solve the modified Black-Sholes equation (6.1) with the boundary conditions

$$
\begin{cases}u(0, \tau)=0 & 0 \leq \tau \leq T  \tag{6.2}\\ \frac{\partial u}{\partial \xi}(1, \tau)=h\left(s_{f}(\tau), \tau\right), & 0 \leq \tau \leq T\end{cases}
$$

We can differentiate the other boundary condition in $\tau$ and yield an ODE for the free boundary:

$$
\frac{\partial u}{\partial \tau}(1, \tau)=\frac{\partial g}{\partial \xi} \frac{d s_{f}}{d \tau} .
$$

with $s_{f}(0)=\max \left(1, r / D_{0}\right)$.

### 6.5.2 American put option

For American put option $P$, the B-S equation is on the infinite domain $S>S_{f}(t), 0 \leq t \leq$ $T$. Through the change-of-variable

$$
\left\{\begin{array}{l}
\eta=\frac{E^{2}}{S} \\
u(\eta, t)=\frac{E P(S, t)}{S}
\end{array}\right.
$$

the infinity domain problem is converted to a finite domain problem:

$$
\begin{cases}\frac{\partial u}{\partial t}+\frac{1}{2} \sigma^{2} \eta^{2} \frac{\partial^{2} u}{\partial \eta^{2}}-r \eta \frac{\partial u}{\partial \eta}=0, & 0 \leq \eta \leq \eta_{f}(t) \\ u(\eta, T)=\max (\eta-E, 0), & 0 \leq t \leq T \\ u\left(\eta_{f}(t), t\right)=\eta_{f}(t)-E, & 0 \leq t \leq T, \\ \frac{\partial u}{\partial \eta}\left(\eta_{f}(t), t\right)=1, & 0 \leq t \leq T \\ \eta_{f}(T)=\max (E, 0) & \end{cases}
$$

We can further convert it to a fixed domain problem as that in the last section.

## Chapter 7

## Exotic Options

Option with more complicated payoff then the standard European or American calls and puts are called exotic options. They are usually traded over the counter. Their prices are usually not quoted on an exchange. We list some common exotic options below.

1. Binary options
2. compound options
3. chooser options
4. barrier options
5. Asian options
6. Lookback options

In the last two, the payoff depends on the history of the asset prices, for instance, the averages, the maximum, etc., we shall call these kinds of options, the path-dependent options, and will be discussed in the next Chapter.

### 7.1 Binaries

The payoff function $\Lambda(S)$ is an arbitrary function. One particular binary option is the cash-or-nothing call, whose payoff is

$$
\Lambda(S)=B H(S-E)
$$

This option can be interpreted as a simple bet on an asset price: if $S>E$ at expiry the payoff is $B$, otherwise zero. We have seen its value is

$$
V=e^{-r(T-t)} \int_{0}^{\infty} \mathcal{P}\left(S^{\prime}, T, S, t\right) \Lambda\left(S^{\prime}\right) d S^{\prime}=e^{-r(T-t)} B \mathcal{N}\left(d_{2}\right),
$$

where

$$
\mathcal{P}\left(S^{\prime}, T, S, t\right)=\frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)}} e^{-\frac{\log \left(\frac{S^{\prime}}{S}\right)-\left(r-\frac{\sigma^{2}}{2}\right)(T-t)}{2 \sigma^{2}(T-t)}}
$$

is the transition probability density for asset price in risk-neutral world.

### 7.2 Compounds

A compound option may be described as an option on an option. We consider the case where the underlying option is a vanilla put or call and the compound option is vanilla put or call on the underlying option. The extension to more complicated option on more complicated option is relatively straightforward. There are four different classes of basic compound options:

1. call-on-call,
2. call-on-put,
3. put-on-call,
4. put-on-put.

Let us investigate the case call-on-call. Other cases can be treated similarly. The underlying option is

$$
\text { Expiry : } T_{2}, \quad \text { Strike price : } E_{2} .
$$

The compound option on this option

$$
\text { Expiry: } T_{1}<T_{2}, \quad \text { Strike price : } E_{1} .
$$

The underlying option has value $C\left(S, t, T_{2}, E_{2}\right)$. At time $T_{1}$, its value $C\left(S, T_{1}, T_{2}, E_{2}\right)$. The payoff for the compound call option is $\max \left\{C\left(S, T_{1}, T_{2}, E_{2}\right)-E_{1}, 0\right\}$. Because the compound options value is governed only by the randomness of $S$, according to the BlackScholes analysis, it also must satisfy the same Black-Scholes equation. We then solve the Black-Scholes equation with payoff

$$
\max \left\{C\left(S, T_{1}, T_{2}, E_{2}\right)-E_{1}, 0\right\}
$$

### 7.3 Chooser options

A regular chooser option gives its owner the right to purchase, for an amount $E_{1}$ at time $T_{1}$, either a call or a put with exercise price $E_{2}$ at time $T_{2}$. Thus, it is a "call on a call or put". Certainly, we have $T_{1}<T_{2}$. The payoff at $T_{1}$ for this call-on-a-call-or-put" is

$$
\Lambda=\max \left\{C\left(S, T_{1}\right)-E_{1}, P\left(S, T_{1}\right)-E_{1}, 0\right\} .
$$

The compound option also satisfies the Black-Scholes equation for the same reason as above. From this and payoff function at $T_{1}$, we can value $V$ at $t$. The contract can be made more general by having the underlying call and put with different exercise prices and expiry dates, or by allowing the right to sell the vanilla put or call. By using the Black-Scholes formula for vanilla option, there is no difficulty to value these complex chooser options.

### 7.4 Barrier option

Barrier options differ from vanilla options in that part of the option contract is triggered if the asset price hits some barrier, $S=X$, say at some time prior to $T$. As well as being either calls or puts, barrier options are categorized as follows.

1. up-and-in: the option expires worthless unless $S$ reaches $X$ from below before expiry.
2. down-and-in: the option expires worthless unless $S$ reaches $X$ from above before expiry.
3. up-and-out: the option expires worthless if $S$ reaches $X$ from below before expiry.
4. down-and-out: the option expires worthless if $S$ reaches $X$ from above before expiry.

### 7.4.1 down-and-out call(knockout)

A European option whose value becomes zero if $S$ ever goes as low as $S=X$. Sometimes in the knockout options, one can have boundary of time, or one can have rebate if the barrier is crossed. In the latter case, the option holder receives a specific amount $Z$ for compensation.

Let us consider the case of a European style down-and-out option without relate. We assume $X<E$. The boundary conditions are

$$
V(X, t)=0, \text { (boundary condition) }, \quad V(S, t) \sim S \text { as } S \rightarrow \infty .
$$

The final condition, $V(S, T)=\max \{S-E, 0\}$. For $S>X$, the option becomes a vanilla call, it satisfies the Black-Scholes equation.

Let us find its explicit solution. Let $S=E e^{x}, t=T-\frac{\tau}{\left(\sigma^{2} / 2\right)}, V=E v$. The BlackScholes equation is transformed into

$$
v_{\tau}=v_{x x}+(k-1) v_{x}-k v,
$$

where $k=\frac{r}{\sigma^{2} / 2}$. We make another change of variable :

$$
v=e^{\alpha x+\beta \tau} u
$$

We choose $\alpha, \beta$ to eliminate the lower order terms in the derivatives of $x$ :

$$
\begin{aligned}
\beta e^{\alpha x+\beta \tau} u+e^{\alpha x+\beta \tau} u_{\tau}= & \alpha^{2} e^{\alpha x+\beta \tau} u+2 \alpha e^{\alpha x+\beta \tau} u_{x}+e^{\alpha x+\beta \tau} u_{x x} \\
& +(k-1)\left(\alpha e^{\alpha x+\beta \tau} u+e^{\alpha x+\beta \tau} u_{x}\right)-k e^{\alpha x+\beta \tau} u .
\end{aligned}
$$

This implies that $\alpha=-\frac{1}{2}(k-1)$ and $\beta=-\frac{1}{4}(k+1)^{2}$ and equation becomes $u_{\tau}=u_{x x}$.
Let $x_{0}=\log \left(\frac{x}{E}\right)$, or $X=E e^{x_{0}}$. The boundary condition becomes $u\left(x_{0}, \tau\right)=0$ and $u(x, \tau) \sim e^{(1-\alpha) x-\beta \tau}$ as $x \rightarrow \infty$. The initial condition becomes

$$
u(x, 0)=u_{0}(x)=\max \left\{e^{\frac{1}{2}(k+1) x}-e^{\frac{1}{2}(k-1) x}, 0\right\}
$$

This follows from the payoff function being

$$
\Lambda=\max \{S-E, 0\}=E \max \left\{\frac{S}{E}-1,0\right\}=E \max \left\{e^{x}-1,0\right\}
$$

and $V=e^{\alpha x+\beta \tau} u\left(x, T-\frac{\tau}{\left(\sigma^{2} / 2\right)}\right)$, with $u(x, 0)=u_{0}(x)=e^{-\alpha x} \max \left\{e^{x}-1,0\right\}=$ $\max \left\{e^{(-\alpha+1) x}-e^{-\alpha x}, 0\right\}$. Notice that because $X<E$, we have $x_{0}<0$, and

$$
u_{0}(x)= \begin{cases}0 & \text { for } x_{0}<x<0 \\ \max \left\{e^{(-\alpha+1) x}-e^{-\alpha x}, 0\right\} & \text { otherwise }\end{cases}
$$

We use method-of-reflection to solve above heat equation with zero boundary condition. We reflect the initial condition about $x_{0}$ as

$$
u(x, 0)= \begin{cases}u_{0}(x), & \text { for } x_{0}<x<\infty \\ -u_{0}\left(2 x_{0}-x\right), & \text { for }-\infty<x<x_{0}\end{cases}
$$

The equation and the initial condition are unchanged under the change-of-variable: $x \rightarrow$ $2 x_{0}-x, u \rightarrow-u$. From the uniqueness of the solution, the solution has the property:

$$
u\left(2 x_{0}-x, t\right)=-u(x, t) .
$$

From this, we can obtain that $u\left(x_{0}, t\right)=-u\left(x_{0}, t\right)=0$.
Since $C=E e^{\alpha x+\beta \tau} u_{1}$ is the vanilla call, where $u_{1}$ satisfies the heat equation with the initial condition:

$$
u_{1}(x, 0)= \begin{cases}e^{\frac{1}{2}(k+1) x}-e^{\frac{1}{2}(k-1) x} & \text { for } x>0 \\ 0, & \text { for } x \leq 0\end{cases}
$$

Using this and the method of reflection, we may express $V$ in terms of $C$ as the follows. First, we may write $V=E e^{\alpha x+\beta \tau}\left(u_{1}+u_{2}\right)$, where the initial condition for $u_{2}$ is the reflected condition from $u_{1}$ :

$$
u_{2}(x, 0)= \begin{cases}u_{0}\left(2 x_{0}-x, 0\right) & \text { for } x \leq 0 \\ 0 & \text { for } x>0\end{cases}
$$

The solution $u_{1}$ corresponds to $C(S, t)$. The solution $u_{2}$ is corresponds to $e^{2 \alpha\left(x-x_{0}\right)} C\left(x^{2} / S, t\right)$. We conclude

$$
V=C(S, t)-\left(\frac{S}{X}\right)^{-(k-1)} C\left(X^{2} / S, t\right)
$$

### 7.4.2 down-and-in(knock-in) option

An "in" option becomes worthless unless the asset price reaches the barrier before expiry. If $S$ crosses the line $S=X$ at some time prior to expiry, then the option becomes a vanilla option. It is common for in-type barrier option to give a rebate, usually a fixed amount, if the barrier is not hit. This compensates the holder for the loss of the option.

The boundary condition for an "in" option is the follows. The option is worthless as $S \rightarrow \infty$, i.e., $V(S, t) \rightarrow 0$ as $S \rightarrow \infty$. At $T$, if $S>X$, then $V(S, T)=0$. For $t<T$, $V(X, t)=C(X, t)$. Since the option immediately turns into a vanilla call and must have
the same value of this vanilla call. For $S \leq X, V(S, t)=C(X, t)$. We only need to solve $V$ for $S>X . V$ still satisfies the same Black-Scholes equation for all $S$, $t$, because its randomness is fully correlated to the randomness of $S$.

We may write $V=c-\bar{V}$, where $c$ is the value of a vanilla call. Then the boundary condition for $\bar{V}$ is $\bar{V}(S, t)=c-V \sim S-0=S$ as $S \rightarrow \infty$. And

$$
\bar{V}(X, t)=c(X, t)-V(X, t)=0, \quad \bar{V}(S, T)=c(S, T)-V(S, T)=c(S, T)=\Lambda(S) .
$$

We observe that $\bar{V}$ is indeed a "down-and-out" barrier option. In other words, 1(down-andin) plus 1 (down-and-out) equal to 1 vanilla call. This is because one and only one of the two barrier options can be active at expiry and whichever it is, its value is the value of a vanilla call.

### 7.5 Asian options and lookback options

In Asian options and lookback options, their payoff functions depend on the history of the underlying asset. For example,

1. a European-type average strike option has the following payoff function

$$
\max \left\{S_{T}-\frac{1}{T} \int_{0}^{T} S(\tau) d \tau, 0\right\}
$$

2. an American-type average strike option,

$$
\Lambda(S, t)=\max \left\{S-\frac{1}{t} \int_{0}^{t} S(\tau) d \tau, 0\right\}
$$

3. geometric mean,

$$
\Lambda(S, T)=\max \left\{S-e^{\int_{0}^{T} \log S(\tau) d \tau}, 0\right\}
$$

4. Lookback call,

$$
\Lambda(S, T)=\max \{S-J, 0\}, \quad J=\max _{0 \leq \tau \leq T} S(\tau)
$$

In general, the payoff depends on $I$, which is defined by

$$
I=\int_{0}^{t} f(S(\tau), \tau) d \tau
$$

where $f$ is a smooth function. The payoff function is $\Lambda(S, I)$. It is important to notice that $I(t)$ is independent of $S(t)$. The value of an asian option should depend on $S, t \mathrm{~s}$ well as $I$. Indeed, we shall see in the next chapter that $d I=f d t$. The only randomness is through $S$, therefore $V$ can be valued through a delta hedge.

For the lookback option, it will be treated as a limiting case of an asian option. We shall discuss this in the next chapter.

## Chapter 8

## Path-Dependent Options

### 8.1 Introduction

If the payoff depends on the history of the underlying asset price, such an option is called a path-dependent option. The Asian options and the Russian options (Lookback options) are the typical examples. The payoff functions for these options are, for example,

1. average strike call option: $\Lambda=\max \left\{S-\frac{1}{T} \int_{0}^{T} S(\tau) d \tau, 0\right\}$,
2. average rate call option: $\Lambda=\max \left\{\frac{1}{T} \int_{0}^{T} S(\tau) d \tau-E, 0\right\}$
3. geometric mean: the arithmetic mean $\frac{1}{T} \int_{0}^{T} S(\tau) d \tau$ above is replaced by $e^{\int_{0}^{T} \log (S(\tau)) d \tau}$.
4. lookback strike put: $\Lambda=\max \left\{\max _{0 \leq \tau \leq T} S(\tau)-S, 0\right\}$.
5. lookback rate put: $\Lambda=\max \left\{E-\max _{0 \leq \tau \leq T} S(\tau), 0\right\}$.

### 8.2 General Method

Let $f$ be a smooth function, define

$$
I(t)=\int_{0}^{t} f(S(\tau), \tau) d \tau
$$

In previous examples, $f(S(\tau), \tau)=S(\tau)$ for arithmetic mean and $f(S(\tau), \tau)=\log S(\tau)$ for geometric mean.

Notice that $I(t)$ is a random variable and is independent of $S(t)$. (This is because $I(t)$ is the sum of increment of functions of $S$ before time $t$, and each increment of $S(\tau), \tau<t$, is independent of $S(t)$.) Therefore, we should introduce another independent variable $I$ besides $S$ to value the derivative $V(S, I, t)$.

The stochastic differential equations governed by $S$ and $I$ are

$$
\begin{aligned}
\frac{d S}{S} & =\mu d t+\sigma d z \\
d I(t) & =f(S(t), t) d t
\end{aligned}
$$

Notice that there is no noize term in $d I$. The only randomness is from $d S$. Therefore, we can use delta hedge to eliminate this randomness. Namely, we consider the portfolio

$$
\Pi=V-\Delta S
$$

as before. We have

$$
d \Pi=d V-\Delta d S=\left(V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} S}{\partial V^{2}}\right) d t+V_{I} d I+V_{S} d S-\Delta d S
$$

We choose $\Delta=\frac{\partial V}{\partial S}$ to eliminate the randomness in $d \Pi$. From the arbitrage assumption, we arrive

$$
V_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}+f \frac{\partial V}{\partial I}=r\left(V-S \frac{\partial V}{\partial S}\right)
$$

### 8.3 Average strike options

### 8.3.1 European calls

Let us consider an average strike call option with European exercise feature. Its payoff function is defined by

$$
\max \left\{S-\frac{1}{T} \int_{0}^{T} S(\tau) d \tau, 0\right)
$$

Or, in terms of $I, \Lambda(S, I, T)=\max \left\{S-\frac{I}{T}, 0\right\}$.
We notice that the modified Black-Scholes equation

$$
V_{t}+S \frac{\partial V}{\partial I}+\frac{\sigma^{2}}{2} S^{2} \frac{\partial^{2} V}{\partial S^{2}}+r S \frac{\partial V}{\partial S}-r V=0
$$

and the initial data for the average strike options are invariant under the transformation: $(S, I) \rightarrow \lambda(S, I)$. Therefore, we expect that its solution is a function of the scale-invariant variable $R=I / S$. Notice that this is also reflected in that

$$
d R=\left(1+\left(\sigma^{2}-\mu\right) R\right) d t-\sigma R d z
$$

depends on $R$ only. Since the initial data can be expressed as $\Lambda=S \max \{1-R / T, 0\}$, we may also expect that $V=S H(R, t)$. This reduces one independent variable.

Plug $V=S H$ into the above modified Black-Scholes equation:

$$
S H_{t}+\frac{1}{2} \sigma^{2} S^{2}\left(2 \frac{\partial H}{\partial S}+S \frac{\partial^{2} H}{\partial S^{2}}\right)+S \cdot S \frac{\partial H}{\partial I}+r S \frac{\partial}{\partial S}(S H)-r(S H)=0
$$

From

$$
\begin{aligned}
\frac{\partial}{\partial S} & =\frac{\partial R}{\partial S}, \quad \frac{\partial}{\partial R}=-\frac{R}{S} \frac{\partial}{\partial R} \\
\frac{\partial^{2}}{\partial S^{2}} & =\frac{2 I}{S^{3}} \frac{\partial}{\partial R}+\frac{I^{2}}{S^{4}} \frac{\partial^{2}}{\partial R^{2}}=\frac{1}{S^{2}}\left(2 R \frac{\partial}{\partial R}+R^{2} \frac{\partial^{2}}{\partial R^{2}}\right)
\end{aligned}
$$

we obtain

$$
\begin{aligned}
& S H_{t}+\frac{1}{2} \sigma^{2} S^{2}\left(2\left(-\frac{R}{S} \frac{\partial H}{\partial R}\right)+S \cdot \frac{1}{S^{2}}\left(2 R \frac{\partial H}{\partial R}+R^{2} \frac{\partial^{2} H}{\partial R^{2}}\right)\right) \\
& +S^{2} \frac{1}{S} H_{R}+r S H+r S^{2}\left(-\frac{R}{S} H_{R}-r S H\right)=0
\end{aligned}
$$

Finally, we arrive

$$
H_{t}+\frac{\sigma^{2}}{2} R^{2} \frac{\partial^{2} H}{\partial R^{2}}+(1-r R) \frac{\partial H}{\partial R}=0
$$

The payoff function

$$
\Lambda(R, T)=\max \left\{1-\frac{R}{T}, 0\right\}=H(R, T),(\text { final condition })
$$

we should require the boundary conditions.

- $H(\infty, t)=0$. Since as $R \rightarrow \infty$ implies $S \rightarrow 0$, then $V \rightarrow 0$, then $H \rightarrow 0$.
- $H(0, t)$ is finite. When $R=0$, we have $S(\tau)=0$, for all $\tau$ with probability 1 . That implies that $\Lambda=0$, and consequently, $H$ is finte at $(0, t)$.

Next, we expect that the solution is smooth up to $R=$. This implies that $H_{R}, H_{R R}$ are finite at $(0, t)$. We have the following two cases: (i) If $R^{2} \frac{\partial^{2} H}{\partial R^{2}}=O(1) \neq 0$ as $R \rightarrow 0$ then $H=O(\log R)$ for $R \rightarrow 0$. Or (ii) if $R \frac{\partial H}{\partial R}=O(1) \neq 0$ as $R \rightarrow 0$, then $H=O(\log R)$. Both cases contradict to $H(0, t)$ being finite. Hence, we have $R H_{R}(0, t), R^{2} H_{R R}(0, t)$ are zeros as $R \rightarrow 0$. Hence the boundary condition $H(0, t)$ is finite is equivalent to $H_{t}(0, t)+$ $H_{R}(0, t)=0$.

This equation with boundary condition can be solved by using the hypergeometric functions. However, in practice, we solve it by numerical method.

### 8.3.2 American call options

We consider the average strike call option with American exercise feature. In this case,

$$
\begin{aligned}
H_{t}+\frac{1}{2} \sigma^{2} R^{2} \frac{\partial^{2} H}{\partial R^{2}}+(1-r R) \frac{\partial H}{\partial R} & \leq 0 \\
H-\Lambda & \geq 0 \\
\left(H_{t}+\frac{1}{2} \sigma^{2} R^{2} \frac{\partial^{2} H}{\partial R^{2}}+(1-r R) \frac{\partial H}{\partial R}\right)(H-\Lambda) & =0
\end{aligned}
$$

where $\Lambda(R, t)=\max \left\{1-\frac{R}{t}, 0\right\}, R(t)=I(t) / S(t), I(t)=\int_{0}^{t} S(\tau) d \tau$.

### 8.3.3 Put-call parity for average strike option

We study the put-call parity for average strike options with European exercise feature. Consider a portfolio is $C-P$. The corresponding payoff function is

$$
S \max \left\{1-\frac{R}{T}, 0\right\}-S \max \left\{\frac{R}{T}-1,0\right\}=S\left(1-\frac{R}{T}\right) \equiv S \cdot H(R, T)
$$

Since the Black-Scholes equation in linear in $R$, we only need to solve the equation with final condition (i) $H(R, T)=1$, (ii) $H(R, T)=-\frac{R}{T}$. For (i), $H(R, t) \equiv 1$. For (ii), since the final condition and the P.D.E. is linear in $R$, we expect that the solution is also linear in $R$. Thus, we consider $H$ is of the following form

$$
H(t, R)=a(t)+b(t) R
$$

Plug this into the equation, we obtain

$$
\frac{d}{d t} a+\frac{d}{d t} b R+(1-r R) b=0
$$

and

$$
\frac{d}{d t} a+b=0, \quad \frac{d}{d t} b-r b=0
$$

with $a(T)=0, b(T)=-\frac{1}{T}$. This differential equation can be solved easily:

$$
b(t)=-\frac{1}{T} e^{-r(T-t)}, \quad a(t)=-\frac{1}{r T}\left(1-e^{-r(T-t)}\right) .
$$

Consequently, we obtain the put-call parity:

$$
\begin{aligned}
C-P & =S\left(1-\frac{1}{r T}\left(1-e^{-r(T-t)}-\frac{1}{T} e^{-r(T-t)} \frac{1}{S} \int_{0}^{t} S(\tau) d \tau\right)\right. \\
& =S-\frac{S}{r T}\left(1-e^{-r(T-t)}\right)-e^{-r(T-t)} \frac{1}{T} \int_{0}^{t} S(\tau) d \tau
\end{aligned}
$$

### 8.4 Lookback Option

A lookback option is a derivate product whose payoff depends on the maximum or minimum of its underlying asset price. For instance, the payoff function for a lookback option with European exercise feature is

$$
\Lambda=\max _{0 \leq \tau \leq T} S(\tau)-S(T)
$$

Such an option is relatively expansive because it gives the holder an extremely advantageous payoff.

As before, let us introduce $J(t)=\max _{0 \leq \tau<t} S(\tau)$. Since $S(\tau), \tau<t$ are independent of $S(t)$, we see that $J(t)$ is independent of $S(t)$. This suggest that we should introduce another independent variable $J$ to value the lookback option in addition to $S$ and $t$. We can derive a stochastic differential equation for $J$ as before. Indeed, it is $d J=0$. However, we shall give a more careful approach. We shall use the fact that for a continuous function $S(\cdot)$,

$$
\max _{0 \leq \tau \leq t}|S(\tau)|=\lim _{n \rightarrow \infty}\left(\int_{0}^{t}|S(\tau)|^{n} d \tau\right)^{\frac{1}{n}}
$$

We leave its proof as an exercise.
Remark. For the minimum, we have

$$
\lim _{n \rightarrow-\infty}\left(\int_{0}^{t}(S(\tau))^{n} d \tau\right)^{\frac{1}{n}}=\min _{0 \leq \tau \leq t} S(\tau) .
$$

Let us introduce

$$
I_{n}=\int_{0}^{t}(S(\tau))^{n} d \tau, \quad J_{n}=I_{n}^{\frac{1}{n}}
$$

The s.d.e. for $J_{n}$,

$$
\begin{aligned}
d J_{n} & =\left(\int_{0}^{t+d t}(S(\tau))^{n} d \tau\right)^{\frac{1}{n}}-\left(\int_{0}^{t}(S(\tau))^{n} d \tau\right)^{\frac{1}{n}} \\
& =\frac{1}{n} \frac{S(t)^{n}}{J_{n}^{n-1}} d t
\end{aligned}
$$

Now, as before we consider the delta hedge:

$$
\Pi=P-\Delta S
$$

From the arbitrage assumption, we can derive the equation for $P\left(S, J_{n}, t\right)$ :

$$
\begin{aligned}
d \Pi & =P_{t} d t+\frac{1}{n} \frac{S^{n}}{J_{n}^{n-1}} \frac{\partial P}{\partial J_{n}} d t+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} P}{\partial S^{2}} d t \\
& =r\left(P-\frac{\partial P}{\partial S} S\right) d t
\end{aligned}
$$

Taking $n \rightarrow \infty$, using the facts that $J_{n} \rightarrow J$ and $\frac{S}{J} \leq 1$, we arrive

$$
P_{t}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} P}{\partial S^{2}}+r S \frac{\partial P}{\partial S}-r P=0
$$

This is the usual Black-Scholes equation. The role of $J$ here is only a parameter. This is consistent to the fact that

$$
d J=0 .
$$

### 8.4.1 A lookback put with European exercise feature

The range for $S$ is $0 \leq S \leq J$. This is because $S \leq J$, for $0 \leq t \leq T$. We claim that

$$
P(0, J, t)=J e^{-r(T-t)}
$$

Firstly, we have that $\Lambda(0, J, T)=\max \{J-S, 0\}=J$. Secondly, if $S(t)=0$, then $S(\tau)=0$ for $t \leq \tau \leq T$. The asset price process becomes deterministic. Therefore, the value of $P$ is the discounted payoff: $P(0, J, t)=J e^{-r(T-t)}$.

Next, we claim that

$$
\frac{\partial P}{\partial J}(J, J, t)=0
$$

From $\mu>0$, the current maximum cannot be the final maximum with probability 1 . The value of $P$ must be insensitive to a small change of $J$.

We can use method of image to solve this problem. Its solution is given by

$$
P=S\left(-1+N\left(d_{7}\right)\left(1+k^{-1}\right)\right)+J e^{-r(T-t)} N\left(d_{5}\right)-k^{-1}\left(\frac{S}{J}\right)^{1-k} N\left(d_{6}\right)
$$

where

$$
\begin{aligned}
d_{5} & =\left[\ln \left(\frac{J}{S}\right)-\left(r-\frac{\sigma^{2}}{2}\right)(T-t)\right] / \sigma \sqrt{T-t} \\
d_{6} & =\left[\ln \left(\frac{S}{J}\right)-\left(r-\frac{\sigma^{2}}{2}\right)(T-t)\right] / \sigma \sqrt{T-t} \\
d_{7} & =\left[\ln \left(\frac{J}{S}\right)-\left(r+\frac{\sigma^{2}}{2}\right)(T-t)\right] / \sigma \sqrt{T-t} \\
k & =\frac{r}{\sigma^{2} / 2}
\end{aligned}
$$

### 8.4.2 Lookback put option with American exercise feature

We have the following linear complementary equation,

$$
L_{B S} P \leq 0, \quad P-\Lambda \geq 0, \quad\left(L_{B S} P\right)(P-\Lambda)=0
$$

where

$$
L_{B S}=\frac{\partial}{\partial t}+\frac{\sigma^{2} S^{2}}{2} \frac{\partial^{2}}{\partial S^{2}}+r S \frac{\partial}{\partial S}-r
$$

The final condition

$$
P(S, J, T)=\Lambda(S, J, T)=J-S
$$

The boundary condition

$$
\frac{\partial P}{\partial J}(J, J, t)=0
$$

We require $P, \frac{\partial P}{\partial S} \frac{\partial P}{\partial J}$ are continuous.
For lookback call option, we simply replace $\max _{0 \leq \tau \leq t} S(\tau)$ by $\min _{0 \leq \tau \leq t} S(\tau)$. For instance, its payoff is $\Lambda=S(T)-\min _{0 \leq \tau \leq T} S(\tau)$.

## Chapter 9

## Bonds and Interest Rate Derivatives

### 9.1 Bond Models

A bond is a long-term contract under which the issuer promises to pay the bondholder coupon payment (usually periodically) and principal (at the maturity dates). If there is no coupon payment, the bond is called a zero-coupon bond. The principal of a bond is called its face value.

### 9.1.1 Deterministic bond model

The value of a bond certainly depends on the interest rate. Let us first assume that the interest rate is deterministic temporarily, say $r(\tau), t \leq \tau \leq T$, is known. Let $B(t, T)$ be the bond value at $t$ with maturity date $T, k(t)$ be its coupon rate. This means that in a small $d t$, the holder receives coupon payment $k(t) d t$. From the no-arbitrage argument,

$$
d B+k(t) d t=r(t) B d t
$$

together with the final condition:

$$
B(T, T)=F(\text { face value })
$$

the bond value can be solved and has the following expression:

$$
B(t, T)=e^{-\int_{t}^{T} r(\tau) d \tau}\left[F+\int_{t}^{T} k(\tau) e^{\int_{\tau}^{T} r(s) d s} d \tau\right]
$$

Thus, the bond value is the sum of the present face value and the coupon stream.
However, the life span of a bond is long (usually 10 years or longer), it is unrealistic to assume that the interest rate is deterministic. In the next subsection, we shall provide a stochastic model.

### 9.1.2 Stochastic bond model

Let us assume that the interest rate satisfies the following s.d.e.:

$$
d r=u(r, t) d t+w(r, t) d z,
$$

where $d z$ is the standard Wiener process. The drift $u$ and the variance $w^{2}$ are proposed by many researchers. We shall discuss these issues later. To find the equation for $B$ with stochastic property of $r$, we consider a portfolio containing bonds with different maturity dates:

$$
\Pi=V\left(t, r, T_{1}\right)-\Delta V\left(t, r, T_{2}\right) \equiv V_{1}-\Delta V_{2}
$$

The change $d \Pi$ in a small time step $d t$ is

$$
d \Pi=d V_{1}-\Delta d V_{2}
$$

where

$$
\begin{aligned}
\frac{d V_{i}}{V_{i}} & =\mu_{i} d t+\sigma_{i} d z \\
\mu_{i} & =\frac{1}{V_{i}}\left(V_{i, t}+u V_{i, r}+\frac{1}{2} w^{2} V_{i, r r}\right) \\
\sigma_{i} & =\frac{1}{V_{i}} w V_{i, r}
\end{aligned}
$$

We we choose $\Delta=V_{1, r} / V_{2, r}$, then the random term is canceled in $d \Pi$. From the noarbitrage argument, $d \Pi=r \Pi d t$. We obtain

$$
\mu_{1} V_{1} d t-\Delta \mu_{2} V_{2} d t=r\left(V_{1}-\Delta V_{2}\right) d t
$$

This yields

$$
\left(\mu_{1}-r\right) V_{1} / V_{1, r}=\left(\mu_{2}-r\right) V_{2} / V_{2, r}
$$

or equivalently

$$
\frac{\mu_{1}-r}{\sigma_{1}}=\frac{\mu_{2}-r}{\sigma_{2}}
$$

Since the left-hand side is a function of $T_{1}$, while the right-hand side is a function of $T_{2}$. Therefore, it is independent of $T$. Let us express it as a known function $\lambda(r, t)$ :

$$
\frac{\mu-r}{\sigma}=\lambda(r, t) .
$$

Plug $\mu_{i}$ and $\sigma_{i}$ back to this equation, and drop the index $i$, we obtain

$$
V_{t}+\frac{1}{2} w^{2} V_{r r}+(u-\lambda w) V_{r}-r V=0 .
$$

The function $\lambda(r, t)=\frac{\mu-r}{\sigma}$ is called the market price, since it gives the extra increase in expected instantaneous rate of return on a bond per an additional unit of risk.

This stochastic bond model depends on three parameter functions $u(r, t), w(r, t)$ and $\lambda(r, t)$. In the next section, we shall provide some model to determine them.

### 9.2 Interest models

There are many interest rate models. We list some of them below.

- Merton (1973): $d r=\alpha d t+\sigma d z$.
- Vasicek (1977): $d r=\beta(\alpha-r) d t+\sigma d z$.
- Dothan (1978): $d r=\sigma r d z$.
- Marsh-Rosenfeld (1983): $d r=\left(\alpha r^{\delta-1}+\beta r\right) d t+\sigma r^{\delta / 2} d z$.
- Cox-Ingersoll-Ross (1985) $d r=\beta(\alpha-r) d t+\sigma r^{\frac{1}{2}} d z$.
- Ho-Lee (1986): $d r=\alpha(t) d t+\sigma d z$.
- Black-Karasinski (1991): $d \ln r=(a(t)-b(t) \ln r) d t+\sigma d z$.

The main requirements for an interest rate model are

- positivity: $r(t) \geq 0$ almost surely,
- mean reversion: $r$ should tends to increase (or to decrease) and toward a mean.

The C-I-R and B-K models have these properties.
Below, we shall illustrate a unified approach proposed by Luo, Yen and Zhang.

### 9.2.1 A functional approach for interest rate model

The idea is to design $r$ to be a function of $x(t)$ and $t$, (i.e. $r=r(x(t), t)$ ) with $x(t)$ governed by a simple stochastic process. We notice that the Ornstein-Uhlenbeck process $d x=-\eta x d t+\sigma d z$ has the property to tend to its mean (which is 0 ) time asymptotically. While the Bessel process $d x=\epsilon / x d t+\sigma d z$ has positive property. We then design the underlying basic process is the sum of these two processes:

$$
d x=\left(-\eta x+\frac{\epsilon}{x}\right) d t+\sigma d z
$$

In general, we allow $\eta, \epsilon$, and $\sigma$ are given functions of $t$. With this simple process, we can choose $r=r(x(t), t)$. Then all interest models mentioned above correspond to different choices of $r(x, t), \eta, \epsilon$ and $\epsilon$.

- Merton's model: we choose $\epsilon=\eta=0, r=x+\alpha t$.
- Dothan's model: $\epsilon=\eta=0, r=e^{x-\sigma^{2} / 2 t}$.
- Ho-Lee: $\epsilon=\eta=0, r=x+\int_{0}^{t} \alpha(s) d s$.
- Vasicek: $\epsilon=0, \eta=\beta, r=x+\alpha$.
- C-I-R: $\beta=2 \eta, \alpha=\left(\sigma^{2}+2 \epsilon\right) /(8 \eta), r=\frac{1}{4} x^{2}$.
- Black-Karasinski: $\epsilon=0, r=\exp (g(t, x))$, where $g=x+\int_{0}^{t} a(s) d s, \eta=b(t)$.

With the interest rate model, the zero-coupon bond price $V$ is given by

$$
V(x, t)=E\left(e^{\int_{t}^{T} r(s, x(s)) d s} \mid x_{t}=x\right), t<T
$$

From the Feymann-Kac formula, $V$ satisfies

$$
\left[\frac{\partial}{\partial t}+\frac{1}{\sigma^{2}} \frac{\partial^{2}}{\partial x^{2}}+\left(-\eta x+\frac{\epsilon}{x}\right) \frac{\partial}{\partial x}-r\right] V=0 .
$$

This model depends three parameter functions $\epsilon(t), \eta(t), \sigma(t)$, and $r(x, t)$. There is no unified theory available yet with this approach and the approach of the previous subsection.

### 9.3 Convertible Bonds

A convertible bond is a bond plus a call option under which the bond holder has the right to convert the bond into a common shares. Thus, it is a function of $r, S, t$ and $T$. Let the stochastic processes governed by $S$ and $r$ are

$$
\begin{aligned}
\frac{d S}{S} & =\mu d t+\sigma d z_{S} \\
d r & =u d t+w d z_{r}
\end{aligned}
$$

Suppose the correlation between $d z_{S}$ and $d z_{R}$ is

$$
d z_{S} d z_{r}=\rho(S, r, t) d t
$$

The final value of the convertible bond $V(r, S, T)=F$, the face value of the bond. Suppose the bond can be converted to $n S$ at any time priori to $T$. Then we have

$$
V(r, S, t) \geq n S
$$

We also have the boundary conditions:

$$
\begin{aligned}
\lim _{S \rightarrow \infty} V(r, S, t) & =n S \\
\lim _{r \rightarrow \infty} V(r, S, t) & =0
\end{aligned}
$$

At $S=0$ or $r=0$, we should require $V(r, 0, t)$ or $V(0, S, t)$ to be finite.
The Black-Scholes analysis for a convertible bond is similar to the analysis for a bond. Let $V_{i}=V\left(r, S, t, T_{i}\right), i=1,2$. Consider a portfolio

$$
\Pi=\Delta_{1} V_{1}+\Delta_{2} V_{2}+\Delta_{S} S
$$

In a small time step $d t$, the change of $d \Pi$ is

$$
d \Pi=\Delta_{1} d V_{1}+\Delta_{2} d V_{2}+\Delta_{S} d S
$$

where

$$
\begin{aligned}
& \quad \frac{d V_{i}}{V_{i}}=\mu_{i} d t+\sigma_{r, i} d z_{r}+\sigma_{S, i} d z_{S} \\
& \mu_{i}=\frac{1}{V_{i}}\left(V_{i, t}+\frac{\sigma^{2}}{2} S^{2} V_{i, S S}+\rho S w V_{i, S r}+\frac{w^{2}}{2} V_{i, r r}+\mu S V_{i, S}+u V_{i, r}\right) \\
& \sigma_{S, i}= \frac{1}{V_{i}} \sigma S V_{i, S} \\
& \sigma_{r, i}= \frac{1}{V_{i}} w V_{i, r}
\end{aligned}
$$

This implies

$$
\begin{aligned}
d \Pi= & \left(\Delta_{1} \mu_{1} V_{1}+\Delta_{2} \mu_{2} V_{2}+\Delta \mu S\right) d t \\
& +\left(\Delta_{1} \sigma_{S, 1} V_{1}+\Delta_{2} \sigma_{S, 2} V_{2}+\Delta \sigma S\right) d z_{S} \\
& +\left(\Delta_{1} \sigma_{r, 1} V_{1}+\Delta_{2} \sigma_{r, 2} V_{2}\right) d z_{r}
\end{aligned}
$$

We choose $\Delta_{1}, \Delta_{2}$ and $\Delta_{S}$ to cancel the randomness terms $d z_{r}$ and $d z_{S}$. This means that

$$
\begin{aligned}
\left(\Delta_{1} \sigma_{S, 1} V_{1}+\Delta_{2} \sigma_{S, 2} V_{2}+\Delta \sigma S\right) d z_{S} & =0 \\
\left(\Delta_{1} \sigma_{r, 1} V_{1}+\Delta_{2} \sigma_{r, 2} V_{2}\right) d z_{r} & =0
\end{aligned}
$$

And it yields

$$
\begin{aligned}
d \Pi & =\left(\Delta_{1} \mu_{1} V_{1}+\Delta_{2} \mu_{2} V_{2}+\Delta \mu S\right) d t \\
& =r\left(\Delta_{1} V_{1}+\Delta_{2} V_{2}+\Delta S\right) d t .
\end{aligned}
$$

Or equivalently,

$$
\Delta_{1}\left(\mu_{1}-r\right) V_{1}+\Delta_{2}\left(\mu_{2}-r\right) V_{2}+\Delta(\mu-r) S=0
$$

This equality together with the previous two give that there exist $\lambda_{r}$ and $\lambda_{S}$ such that

$$
\begin{aligned}
\left(\mu_{1}-r\right) & =\lambda_{S} \sigma_{S, 1}+\lambda_{r} \sigma_{r, 1} \\
\left(\mu_{2}-r\right) & =\lambda_{S} \sigma_{S, 2}+\lambda_{r} \sigma_{r, 2} \\
(\mu-r) & =\lambda_{S} \sigma
\end{aligned}
$$

The functions $\lambda_{r}$ and $\lambda_{S}$ are called the market prices of risk with respect to $r$ and $S$, respectively. Plug the formulae for $\mu_{i}, \sigma_{r, i}$ and $\sigma_{S, i}$, we obtain the Black-Scholes equation for a convertible bond:

$$
\left[\frac{\partial}{\partial t}+\frac{\sigma^{2}}{2} S^{2} \frac{\partial^{2}}{\partial S^{2}}+\rho S w \frac{\partial^{2}}{\partial S \partial r}+\frac{w^{2}}{2} \frac{\partial^{2}}{\partial r^{2}}+r S \frac{\partial}{\partial S}+\left(u-\lambda_{r} w\right) \frac{\partial}{\partial r}-r\right] V=0
$$

## Appendix A

## Basic theory of stochastic calculus

## A. 1 Brownian motion

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A process is a function $X:[0, \infty) \times(\Omega, \mathcal{F}) \mapsto$ $(E, \mathcal{B})$, such that for each $t \geq 0, X(t)$ is a random variable. Here, $E=R^{d}, \mathcal{B}$ is the Borel sets. Let

$$
\begin{aligned}
\mathcal{F}_{t} & =\sigma\{X(s), s \leq t\} \\
\mathcal{F}^{t} & =\sigma\{X(s), s \geq t\}
\end{aligned}
$$

A process is called Markov if

$$
P\left(A \mid \mathcal{F}_{t}\right)=P(A \mid X(t)), \forall A \in \mathcal{F}^{t} .
$$

This is equivalent to

$$
P\left\{X(r) \in B \mid \mathcal{F}_{t}\right\}=P\{X(r) \in B \mid X(t)\} \forall r>t
$$

A markov process is characterized by its transition probability:

$$
P(t, x, s, B):=P\{X(s) \in B \mid X(t)=x\}
$$

with initial distribution

$$
P\{X(0) \in B\}=\nu(B)
$$

Theorem 1.7 If $X$ is a Markov process, then the corresponding transition probability $P$ satisfies Chapman-Kolmogorov equation:

$$
\int P\left(t_{0}, x_{0}, t_{1}, d x_{1}\right) P\left(t_{1}, x_{1}, t_{2}, B\right)=P\left(t_{0}, x_{0}, t_{2}, B\right) .
$$

Conversely, if $P$ is a function satisfies Chapman-Kolmogorov equation, then there is a Markov process whose transition probability is $P$.

Two standard Markov processes are the Wiener process and the Poison process. The Wiener process has the transition probability density function

$$
p(t, x, s, y)=\frac{1}{(2 \pi(t-s))^{d / 2}} \exp \left(-\frac{|x-y|^{2}}{s-t}\right)
$$

Such a distribution is called a normal distribution with mean $x$ and variance $s-t$.
Definition 1.7 Brownian motion: A process is called a Brownian motion(or a Wiener process) if

1. $B_{t}-B_{s}$ has normal distribution with mean 0 and variance $t-s$,
2. it has independent increments: that is $B_{t}-B_{s}$ is independent of $B_{u}$ for all $u \leq s<t$,
3. $B_{t}$ is continuous in $t$.

It is easy to see that $B_{t}$ is markovian and its transition probability is

$$
p(t, x, s, y)=\frac{1}{(2 \pi(t-s))^{d / 2}} \exp \left(-\frac{|x-y|^{2}}{s-t}\right)
$$

Definition 1.8 A process $\left\{X_{t} \mid t \geq 0\right\}$ is called martingale if
(i) $E X_{t}<\infty$,
(ii) $E\left(X_{u} \mid X_{s}, 0 \leq s \leq t\right\}=X_{t}$.

This means that if we know the value of the process up to time $t$ and $X_{t}=x$, then the future expectation of $X_{u}$ is $x$.

Theorem 1.8 1. $B_{t}$ is a martingale.
2. $V_{t}^{2}-t$ is a martingale.

## Proof.

1. $E\left(B_{t}\right)=0$. From the fact that $B_{t+s}-B_{t}$ is independent of $B_{t}$, for all $s>0$, we obtain $E\left(B_{t+s}-B_{t} \mid B_{t}\right)=0$, for all $s>0$. Hence,

$$
\begin{aligned}
E\left(B_{t+s} \mid B_{t}\right) & =E\left(\left(B_{t+s}-B_{t}\right)+B_{t} \mid B_{t}\right) \\
& =E\left(\left(B_{t+s}-B_{t}\right) \mid B_{t}\right)+E\left(B_{t} \mid B_{t}\right) \\
& =B_{t}
\end{aligned}
$$

2. $E\left(B_{t}^{2}\right)=t<\infty$.
3. Use

$$
\begin{aligned}
B_{t+s}^{2} & =\left(\left(B_{t+s}-B_{t}\right)+B_{t}\right)^{2} \\
& =\left(B_{t+s}-B_{t}\right)^{2}+2 B_{t}\left(B_{t+s}-B_{t}\right)+B_{t}^{2}
\end{aligned}
$$

and the fact that $B_{t+s}-B_{t}$ is independent of $B_{t}$, we obtain

$$
E\left(B_{t+s}^{2} \mid B_{t}\right)=s+B_{t}^{2}
$$

Hence,

$$
E\left(B_{t+s}^{2}-(t+s) \mid B_{t}\right)=B_{t}^{2}-t .
$$

We can show that the Brownian motion has infinite total variation in any interval. This means that

$$
\lim \sum\left|B\left(t_{i}\right)-B\left(t_{i-1}\right)\right|=\infty
$$

However, its quadratic variation, defined by

$$
[B, B](a, b):=\lim \sum\left|B\left(t_{i}\right)-B\left(t_{i-1}\right)\right|^{2}
$$

is finite. Here, $a=t_{0}<t_{1}<\cdots<t_{n}=b$ is a partition of $(a, b)$, and the limit is taken to be $\max \left(t_{i}-t_{i-1}\right) \rightarrow 0$.

Theorem 1.9 We have $[B, B](0, t)=t$ almost surely.
Proof. Let us partition $(0, t)$ evenly into $2^{n}$ subintervals. Let $T_{n}=\sum_{i=1}^{2^{n}}\left|B\left(t_{i}\right)-B\left(t_{i-1}\right)\right|^{2}$. We see that

$$
\begin{aligned}
& \left.E T_{n}=\sum E\left(\left|B\left(t_{i}\right)-B\left(t_{i-1}\right)\right|^{2}\right)\right)=\sum\left|t_{i}-t_{i-1}\right|=t \\
& \qquad \begin{aligned}
\operatorname{Var}\left(T_{n}\right) & =\operatorname{Var}\left(\sum\left|B\left(t_{i}\right)-B\left(t_{i-1}\right)\right|^{2}\right) \\
& =\sum \operatorname{var}\left(\left(B\left(t_{i}\right)-B\left(t_{i-1}\right)\right)^{2}\right) \\
& =2 \sum\left(t_{i}-t_{i-1}\right)^{2} \\
& =2 t^{2} 2^{-n}
\end{aligned}
\end{aligned}
$$

Hence, $\sum_{n=1}^{\infty} \operatorname{Var}\left(T_{n}\right)<\infty$. From Fubini theorem,

$$
E\left(\sum_{n=1}^{\infty}\left(T_{n}-E T_{n}\right)^{2}\right)<\infty
$$

This implies $E\left(\left(T_{n}-E T_{n}\right)^{2}\right) \rightarrow 0$ and hence, $T_{n} \rightarrow E T_{n}$ almost surely.

## A. 2 Stochastic integral

We shall define the integral

$$
\int_{0}^{t} f(s) d B(s)
$$

The method can be applied with $B$ replaced by a martingale, or a martingale plus a function with finite total variation. We shall require that $f \in H^{2}$, which means:
(i) $f(t)$ depends only on the history $\mathcal{F}_{t}$ of $B_{s}$, for $s \leq t$,
(ii) $\int_{0}^{t} E|f|^{2}<\infty$.

For this kind of functions, we can define its Itô's integral as the follows.

1. We define Itô's integral for $f \in H^{2}$ and being step functions. Its Itô's integral is define by

$$
\int_{0}^{t} f(s) d B(s)=\sum_{i=0}^{n} f\left(t_{i-1}\right)\left(B\left(t_{i}\right)-B\left(t_{i-1}\right)\right) .
$$

2. We use above step functions $f_{n}$ to approximate general function $f \in H^{2}$. Using the fact that, for step functions $g \in H^{2}$,

$$
E\left(\left(\int_{0}^{t} g(s) d B(s)\right)^{2}\right)=\int_{0}^{t} E|g(s)|^{2} d s
$$

one can show that

$$
\lim _{n \rightarrow \infty} \int_{0}^{t} f_{n}(s) d B(s)
$$

almost surely.
An typical example is

$$
\int_{0}^{t} B(s) d B(s)=\frac{1}{2} B(t)^{2}-\frac{t}{2}
$$

From the definition, the integral can be approximated by

$$
I_{n}=\sum_{i=1}^{n} B\left(t_{i-1}\right)\left(B\left(t_{i}\right)-B\left(t_{i-1}\right) .\right.
$$

We have

$$
\begin{aligned}
I_{n} & =\frac{1}{2} \sum_{i=1}^{n}\left[B^{2}\left(t_{i}\right)-B^{2}\left(t_{i-1}\right)-\left(B\left(t_{i}\right)-B\left(t_{i-1}\right)\right)^{2}\right] \\
& =\frac{1}{2} B^{2}(t)-\frac{1}{2} \sum_{i=1}^{n}\left(B\left(t_{i}\right)-B\left(t_{i-1}\right)\right)^{2}
\end{aligned}
$$

We have seen that the second on the right-hand side tends to $\frac{1}{2} t$ almost surely.

## A. 3 Stochastic differential equation

A stochastic differential equation has the form

$$
\begin{equation*}
d X_{t}=a\left(X_{t}, t\right) d t+b\left(X_{t}, t\right) d B(t) \tag{A.1}
\end{equation*}
$$

A Markov process $X$ is said to be a strong solution of this s.d.e. if it satisfies

$$
X_{t}-X_{0}=\int_{0}^{t} a\left(X_{s}, s\right) d s+\int_{0}^{t} b\left(X_{s}, s\right) d B(s) .
$$

Theorem 1.10 (Itô's formula) If $X$ satisfies the s.d.e. $d X=a d t+b d B$, then

$$
d f(X(t))=\left(a f^{\prime}(X(t))+\frac{1}{2} b^{2} f^{\prime \prime}(X(t))\right) d t+f^{\prime}(X(t)) b d B(t) .
$$

We shall give a brief idea of the proof. In a small time step $\left(t_{i-1}, t_{i}\right)$, let $\Delta B=B\left(t_{i}\right)-$ $B\left(t_{i-1}\right), \Delta t=t_{i}-t_{i-1}$. We have

$$
f\left(X\left(t_{i-1}\right)+\Delta X\right)-f\left(X\left(t_{i-1}\right)\right)=f^{\prime} \Delta X+\frac{1}{2} f^{\prime \prime}(\Delta X)^{2}+o\left((\Delta X)^{2}\right)
$$

We notice that

$$
\begin{aligned}
(\Delta X)^{2} & =(a \Delta t+b \Delta B)^{2} \\
& =a^{2}(\Delta t)^{2}+2 a b \Delta t+b^{2}(\Delta B)^{2} \\
& \approx b^{2} \Delta t+o(\Delta t) .
\end{aligned}
$$

Plug $\Delta X$ and $(\Delta X)^{2}$ into the Taylor expansion formula for $f$. This yields the Inô's formula.

## A. 4 Diffusion process

For a s.d.e.(A.1), we define the associated semigroup $T_{t}$ by

$$
T_{t} f=E_{x, t}(f(X(s)), s>t
$$

From the Markovian property, one can show that $T_{t}$ is a semi-group. Indeed, in terms of the transition probability density function $p(t, x, s, y)$,

$$
T_{t} f=\int p(t, x, s, y) f(y) d y
$$

For a semigroup $T_{t}$, we define its generator as

$$
L f:=\lim _{t \rightarrow 0} \frac{T_{t} f-f}{t} .
$$

From Itô's formula,

$$
L f:=a f^{\prime}+\frac{1}{2} b^{2} f^{\prime \prime}
$$

This follows from Itô's formula and $E\left(\int_{0}^{t} g(s) d B(s)=0\right)$.
With a fixed $s>t$ and $f$, we define

$$
u(x, t):=\left(T_{t} f\right)(x, t)=E_{x, t}(f(X(s)) .
$$

Theorem 1.11 If $X$ satisfies the s.d.e. (A.1), then the associate $u(x, t):=E_{x, t}(f(X(s))$, $s>t$, satisfies the backward diffusion equation:

$$
u_{t}+L u=0,
$$

and has final condition $u(s, x)=f(x)$.
Proof. First, we notice that

$$
\begin{aligned}
u(x, t-h) & =E_{x, t-h}(f(X(s)) \\
& =E_{x, t-h} E_{X(t), t}(f(X(s)) \\
& =E_{x, t-h}(u(X(t), t))
\end{aligned}
$$

This shifts final time from $s$ to $t$. Now, we consider

$$
\begin{aligned}
\frac{u(x, t-h)-u(x, t)}{h} & =\frac{1}{h} E_{x, t-h}(u(X(t), t)-u(X(t-h), t)) \\
& =\frac{1}{h} \int_{t-h}^{t} L u(X(s), t) d s \\
& \rightarrow L u(x, t)
\end{aligned}
$$

as $h \rightarrow 0+$. Next, $u(x, s-h)=E_{x, s-h} f(X(s))$, we have $X(s) \rightarrow x$ as $h \rightarrow 0+$ almost surely. Hence $u(x, s-h) \rightarrow f(x)$ as $h \rightarrow 0+$.

Since $u$ can be represent as

$$
u(x, t)=\int p(x, t, s, y) f(y) d y
$$

we obtain that $p$ satisfies

$$
p_{t}+L_{x} p=0,
$$

and

$$
p(s, x, s, y)=\delta(x-y)
$$

For diffusion equation with source term, its solution can be represented by the following Feymann-Kac formula.

Theorem 1.12 (Feymann-Kac) If $X$ satisfies the s.d.e. (A.1), then the associate

$$
\begin{equation*}
u(x, t):=E_{x, t}\left[f(X(s)) \exp \int_{t}^{s} g(X(\tau), \tau) d \tau\right], s>t \tag{A.2}
\end{equation*}
$$

solves the backward diffusion equation:

$$
u_{t}+L u+g u=0,
$$

with final condition $u(s, x)=f(x)$.

Proof. As before, we shift final time from $s$ to $t$ :

$$
\begin{aligned}
u(x, t-h) & =E_{x, t-h}\left[f(X(s)) \exp \int_{t-h}^{s} g(X(\tau), \tau) d \tau\right] \\
& =E_{x, t-h} E_{X(t), t}\left[f(X(s)) \exp \int_{t-h}^{s} g(X(\tau), \tau) d \tau\right] \\
& =E_{x, t-h}\left[E_{X(t), t}\left[f(X(s)) \exp \int_{t}^{s} g(X(\tau), \tau) d \tau\right] \cdot E_{X(t), t}\left[\exp \int_{t-h}^{t} g(X(\tau), \tau) d \tau\right]\right] \\
& =E_{x, t-h}\left[u(X(t), t) \exp \int_{t-h}^{t} g(X(\tau), \tau) d \tau\right]
\end{aligned}
$$

Here, we have used independence of $X$ in the regions $(t-h, t)$ and $(t, s)$. Now, from Itô's formula, we have

$$
\begin{aligned}
u(x, t-h)-u(x, t) & =E_{x, t-h}\left[u(X(t), t) \exp \int_{t-h}^{t} g(X(\tau), \tau) d \tau-u(X(t-h), t)\right] \\
& =E_{x, t-h} \int_{t-h}^{t} d\left(u(X(s), t) \exp \int_{t-h}^{s} g(X(\tau), \tau) d \tau\right) \\
& =E_{x, t-h} \int_{t-h}^{t}(L u(X(s), t)+u(X(s), t) g(X(s), s)) d s
\end{aligned}
$$

From this, it is easy to see that

$$
\lim _{h \rightarrow 0+} \frac{u(x, t-h)-u(x, t)}{h}=L u(x, t)+g(x, t) u(x, t)
$$

