Master Thesis for Degree in *Computational Engineering* of
Uni. Erlangen-Nuernberg, Germany

Pricing Bermudan and American Options Using the FFT Method

by Fang FANG

Martrikelnummer: 21119128

Advisor: Dr. Ir. Kees Oosterlee (TU Delft, Netherlands)
and Prof. Ulrich Ruede

Cooperation with: Ir. Drs. Roger Lord (Rabobank, Utrecht, Netherlands)

May 25, 2006

Special Thanks

This thesis work was carried out in the Delft Center for Computational Science and Engineering, TU Delft, the Netherlands. The research started in September 2005 and finished in May 2006.

During the whole period, from literature review, numerical implementation, to remedy seeking for discovered problems, and final thesis writing, the work was under supervision of Dr.Ir. Kees Oosterlee. Without his great help, the thesis work could never have been finished successfully.

The work is performed in a cooperation with Ir. Drs. Roger Lord from Rabobank.

The starting point of the methods involved in this thesis is based on his research. The idea of replacing FFT by FRFT is also proposed by him. Besides, I've benefited a lot from his comments to my work and from discussions and brain-storming together with him and Dr.Ir. Kees Oosterlee.

Mr. Frank Bervoets helped a lot in the aspect of high efficiency Matlab programming. Besides, discussions on parameter choosing and other problems in implementation with him gave me a better understanding of the methods and their behaviors.

Without the full support and help of prof. Ulrich Ruede, it was not possible for me to focus on this thesis work in the Netherlands.

My husband is the one who's always standing by my side, with encouragement and support.

Preface

The aim of this thesis work is to develop a fast and accurate solution method for (multi)asset Bermudan and American option pricing problems by means of Fourier transforms.

The outline of this thesis is as follows:

In Chap. 1 basic definitions and terminology of financial derivatives, especially of options, are given, as well as the notation used in later chapters.

In Chap. 2 some fundamental mathematical models and theorems in the field of computational finance are introduced, such as the models for asset prices and option prices, the famous Black-Scholes partial differential equation(PDE), and the Risk-neutral valuation theory.

In Chap. 3 literature on the recently developed "Transform methods", which employs the fast Fourier transform (FFT) algorithms in option pricing problems, is reviewed.

First, basic Transform methods are presented and their merits are examined and compared. Then, a highly efficient method, the Lord method, is described and chosen as the basic method for this thesis work. Based on the Lord method, we propose to use the fractional Fourier transform (FRFT) algorithm instead of FFT for a better accuracy. The FRFT algorithm is derived and explained.

In Chap. 4 implementation details of the Lord method and the Lord-FRFT method are presented. Besides, the Richardson extrapolation is introduced, and is proposed to be included as a posteriori technique applied to the results generated by the Lord/Lord-FRFT methods.

In Chap. 5 the four methods, the Lord, the Lord-FRFT, the Lord-extrapolation (the Lord method combined with 2-point Richardson extrapolation), and the Lord-FRFT-extrapolation method, are employed to value one-asset European, Bermudan and American options.

Two stochastic models, the Brownian motion and the Variance Gamma process, for the asset price dynamics are chosen.

The accuracy and the computational time of each method is summarized and compared.

In Chap. 6 the Lord and Lord-FRFT method are generalized to value two-asset options.

Definitions for the high-dimensional Fourier transform are given; then derivations of the high-dimensional formulas of the Lord method follow; Finally, a two-asset minimum call option is priced using the generalized methods, and the results are summarized and discussed.

Chap. 7 summarizes. The thesis is finished with a discussion of problems and an outlook to future work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Assets of various kinds are bought and sold in all types of markets. As markets have become more sophisticated, more complex contracts than simple buy/sell trades have been introduced.

These contracts include financial derivatives, derivative securities, derivative products, contingent claims and derivatives, which are traded on futures and options markets. They can give investors a great range of opportunities to tailor their dealings to their investment needs.

A derivative instrument (or simply derivative) is a financial instrument which derives its value from another "underlying" financial security.

Options, futures, swaps, swaptions, structured notes are all examples of derivative securities. And the subject of this thesis is options.

First of all, let's get ourselves familiar with necessary financial jargon, and see how options work.

1.1 What is Option?

An option is a contract to buy or sell an underlying at a prescribed price at a future time point.

The idea of trading options began to be widespread in 1973, when derivatives started to be occasionally traded on the Chicago Board Options Exchange.

The simplest option contracts are calls and puts. Referring to technicality in the option contracts, they are further classified as European, American or Bermudan option contracts. The options that can only be exercised at expiry are called European. An American option is the one that may be exercised at any time prior to expiry. If there are several early exercisable dates during the option's life time, then it is named Bermudan option.

A European call option, for example, is a contract, which gives the holder the right to purchase a prescribed asset, known as the "underlying asset" or "underlying", for a prescribed amount, known as the "exercise price" or the "strike price", at a prescribed time in the future, known as the "expiry date" or "expiration date".

For the holder of the option, this contract is a right and not an obligation. The other party to the contract, the writer, has a potential obligation: he must sell the asset if the holder chooses to buy it.

Since the option confers on its holder a right with no obligation it has some value. Conversely, the writer of the option must be compensated for the obligation he has. To value the options is also the aim of this thesis.

Options have two primary uses: speculation and hedging.

For example, an investor speculates that a particular stock is going to rise, hence he wants to purchase some shares in that company. Alternatively, he could buy a call option contract with an exercise price lower than his speculation, so that he could make money by executing the call contract buying in some shares for a lower price while selling them in the market for a higher price, when his guessing is correct. On the other hand, if he thinks that the shares are going to fall, he can, conversely, sell shares or buy puts. If he speculates by selling shares that he does not own(which in certain circumstances is legal in many markets) he is selling short[1].

The holder of a call option has the possibility of an arbitrarily large payoff with limited loss of the initial premium, while the writer has the possibility of an arbitrary large loss with the limited profit to the initial premium. Similarly, writing a put option exposes the writer to large potential losses for a profit limited to the initial premium. Then why would anyone write an option? The first likely answer is that the writer of an option expects to make a profit by taking a view on the market. Writers of calls are, in fact, taking a short position in the underlying: they expect its value to fall. It is usually argued that such people must be present in the market, for if everyone expected the value of a particular asset to rise its market price would be higher than, in fact, it is. Similarly, there must also be people who believe that the value of the underlying will rise.

However, one usually writes an option for the purpose of hedging. Since the value of a put option rises when an asset price falls, the value of a portfolio containing both assets and puts remain financially stable with a proper choice of the ratio of assets and options in the portfolio. A portfolio that contains only assets falls when the asset price falls, while one that contains only put options rises. Somewhere between these two extremes a ratio exists at which a small unpredictable movement in the asset does not result in any unpredictable movement in the value of the portfolio. This ratio is called instantaneously risk-free. The reduction of risk by taking advantage of such correlations between the asset and option price movements is called hedging. If a market maker can sell an option for more than it is worth and then hedge away all the risk for the rest of the option's life, he has locked in a profit.

Call and put options are only a small section of the available derivative products. Nowadays most traded options are American options.

Other types of option include the so-called exotic or path-dependent options. These options have values which depend on the history of an asset price, not only on its value on exercise. An example is an option to purchase an asset for the arithmetic average value of that asset over the month before expiry.

1.2 Notations in This Thesis

The value of an underlying is denoted by S throughout the thesis; when time specification is important, $S(t)$ is used to denote the underlying price at time t.

The value of an option is denoted by V ; when the distinction is important, C and P are used to denote a call and a put, respectively. When it is important to show that they're functions of S and time t, notations $V(S,t)$, $C(S,t)$ or $P(S,t)$ are used, instead.

The value of an option also depends on the volatility of the underlying asset σ, the exercise price E, the expiry time T, the interest rate r, and the dividend rate D_0.

Let $F(X)$ denote the cumulative distribution function of a stochastic variable X, $f(x)$ the probability density function, and $\phi(\omega)$ the corresponding characteristic function with

$$\phi(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} \cdot f(x)\, dt, \tag{1.1}$$

where i is the imaginary unit, $i = \sqrt{-1}$.

$\mathbb{E}[\cdot]$ denotes the operator of expectation of a stochastic variable; when a specific time t and a specific measure $\mathbb{Q}$ are important for understanding, the notation $\mathbb{E}_t^{\mathbb{Q}}$ is used.

[1]The opposite of a short position is a long position.

Bold big letters are used for matrices, like $\mathbf{F}$ and $\mathbf{M}$. Vectors are denoted by small letters with an arrow, like $\vec{\alpha}$ and $\vec{x}$.

Chapter 2

Mathematical Modeling in Computational Finance

The basic assumption in this thesis, common in most of option pricing theory, is that we do not know and cannot predict tomorrow's values of asset prices. Instead of precise future information, we can say something meaningful about expected future values. Interest rates, however, are discussed deterministic. We can get useful information like the mean and variance of the series, and the likely distribution of future asset prices.

An important concept concerning interest rates is that of present value or discounting. How much would one pay now to receive a guaranteed amount E at the future time T?

If we assume that interest rates are constant, the answer to this question is found by discounting the future value, E, using continuously compounded interest. With a constant interest rate, r, money in the bank $M(t)$ grows exponentially according to

$$\frac{dM}{M} = rdt. \tag{2.1}$$

The solution of this is simply

$$M = Ce^{rt}, \tag{2.2}$$

where C is the constant of integration. Since $M = E$ at $t = T$, the value at time T of the certain payoff is

$$M = Ee^{-r(T-t)}. \tag{2.3}$$

If the interest rate is a known function of time $r(t)$, then M can be modified trivially and results in

$$M = Ee^{-\int_t^T r(s)ds}. \tag{2.4}$$

2.1 Random Walk of Asset Prices

It is often stated that asset prices must move randomly because of the efficient market hypothesis[Paul Wilmott,1995]. There are several different forms of this hypothesis with restrictive assumptions, but basically they all say two things:

- The past history is fully reflected in the present asset price, which does not hold any further information;
- Markets respond immediately to any new information about assets.

With the two assumptions above, unanticipated changes in the asset price can be represented by a Markov process [1].

Since a relative measure of the change in price is a more useful indicator than an absolute measure, asset price models are typically defined for the return, dS/S.

Samuelson's model decomposes dS/S into two parts. One part is a deterministic and anticipated return as the return on money invested in a risk-free bank, μdt, where μ is a measure of the average rate of growth of the asset price, also known as the drift. In basic models μ is taken constant, whereas in more involved models, for exchange rates, for example, μ can be a function of S and t. The other part is the random change in the asset price in response to external effects (such as unexpected news); σdX, where σ is a parameter called the volatility, which measures the standard deviation of the returns. The quantity dX is a random variable drawn from a normal distribution, whose mean is zero and variance is dt. It is known as a so-called Wiener process [2].

Putting these two contributions together, the model is

$$\frac{dS}{S} = \mu dt + \sigma dX, \tag{2.5}$$

which is also known as the Brownian motion [Samuelson, 1965].

One way of writing the dX is $dX = \Phi\sqrt{dt}$, with Φ denoting a random variable drawn from a standardized normal distribution with zero mean and unit variance. The reason to scale dX with $\sqrt{dt}$ is that any other choice for the magnitude of dX would lead to a problem that is either meaningless or trivial in the limit $dt \to 0$. If dX were not scaled in this way, the variance of the random walk for S would have a limiting value of 0 or ∞.

Equation (2.5) is a particular example of a random walk, it fits real time series data very well, at least for equities and indices. There are some discrepancies; for instance, real data appears to have a greater probability of large rises or falls than the model predicts. The model does not allow large rises or falls. But, on the whole, it has stood the test of time remarkably well and can be the starting point for more sophisticated models. As an example of such generalization, the coefficients of dX and dt in (2.5) can be any functions of S and/or t.

If S follows the log normal random walk given by (2.5), then the probability density function is represented by a skewed bell-shaped curve.

We can interpret (2.5) as a recipe for generating a time series - each time the series is restarted a different path results. Each path is called a realization of the random walk.

Firstly, we have the mean of dS

$$\mathbb{E}[dS] = \mathbb{E}[\sigma S dX + \mu S dt] = \mu S dt,$$

since $\mathbb{E}[dX] = 0$.

Secondly, the variance of dS is

$$\mathrm{Var}[dS] = \mathbb{E}[dS^2] - \mathbb{E}[dS]^2 = \mathbb{E}[\sigma^2 S^2 dX^2] = \sigma^2 S^2 dt.$$

The square root of the variance is the standard deviation, which is thus proportional to σ and $\sqrt{t}$.

If we compare two random walks with different values for the parameters μ and σ, we see that the one with the larger value of μ usually rises more steeply and the one with

[1] Markov process: A random process whose future probabilities are determined by its most recent values. A stochastic process $x(t)$ is called Markov if for every n and $t_1 < t_2 \ldots < t_n$, we have

$$\mathbb{P}\left(x(t_n) \leq x_n | x(t_{n-1}), \ldots, x(t_1)\right) = \mathbb{P}\left(x(t_n) \leq x_n | x(t_{n-1})\right).$$

[2] Wiener process: A continuous-time stochastic process $W(t)$ for $t \geq 0$ with $W(0) = 0$ and such that the increment $W(t) - W(s)$ is Gaussian with mean 0 and variance $t - s$ for any $0 \leq s < t$, and increments for non-overlapping time intervals are independent.

the larger value of σ appears more jagged. Typically, for stocks and indices the value of σ is in the range 0.05 to 0.4 (the units of σ^2 are per annum). Government bonds are examples of assets with low volatility, while shares in high-tech companies generally have high volatility. The volatility is often quoted as a percentage, so that $\sigma = 0.2$ would be a 20% volatility.

2.2 Random Walk of Option Prices

2.2.1 Itô's Lemma

In real life asset prices are quoted at discrete intervals of time. Mathematical models however are defined in the continuous time limit $dt \to 0$; it is more efficient to solve the resulting differential equations than to value options by direct simulation of the random walk on a practical time scale. We need some technical machinery that enables us to handle the random term as $dt \to 0$.

Itô's Lemma is an important result for the manipulation of random variables. It is like Taylor's expansion for functions of deterministic variables, as it relates the small change in a function of a random variable to the small change in the random variable itself.

Before stating Itô's Lemma, we need the result that

$$dX^2 \to dt \quad \text{as} \quad dt \to 0. \tag{2.6}$$

Suppose that $f(S)$ is a smooth function of S (and forget for the moment that S is stochastic). If we vary S by a small amount dS then f also varies by a small amount, provided we are not close to singularities of f. From the Taylor series expansion we can write

$$df = \frac{df}{dS}dS + \frac{1}{2}\frac{d^2 f}{dS^2}dS^2 + \cdots, \tag{2.7}$$

where the dots denote a remainder which is smaller than any of the terms that retain. Now recall that dS is given by the random walk equation

$$dS = \mu S dt + \sigma S dX.$$

Squaring it we find that

$$dS^2 = (\mu S dt + \sigma S dX)^2 = \mu^2 S^2 dt^2 + 2\mu S^2 \sigma dt dX + \sigma^2 S^2 dX^2. \tag{2.8}$$

Examine the order of the magnitude of each of the terms in (2.8). Since $dX = O(\sqrt{dt})$, the last term is the largest for small dt and dominates the other two terms. Thus, to leading order,

$$dS^2 \approx \sigma^2 S^2 dX^2. \tag{2.9}$$

We substitute this into (2.7) and keep only those terms which are at least of $O(dt)$. Using also the definition of dS, we find that

$$\begin{aligned} df &= \frac{df}{dS}(\sigma S dX + \mu S dt) + \frac{1}{2}\sigma^2 S^2 \frac{d^2 f}{dS^2} dt \\ &= \sigma S \frac{df}{dS} dX + \left(\mu S \frac{df}{dS} + \frac{1}{2}\sigma^2 S^2 \frac{d^2 f}{dS^2}\right) dt. \end{aligned} \tag{2.10}$$

This results is related to Itô's lemma, relating the small change in a function of a random variable to the small change in the random variable itself.

Because the order of magnitude of dX is $O(\sqrt{dt})$, the second derivative of f with respect to S appears in the expression for df at order dt. It can be shown that any other order of magnitude for dX leads to unrealistic properties for the random walk in the limit

$dt \to 0$; if $dX \gg \sqrt{dt}$ the random variable goes immediately to zero or infinity, and if $dX \ll \sqrt{dt}$ the random component of the walk vanishes in the limit 0.

Observe that (2.10) is made up of a random component proportional to dX and a deterministic component proportional to dt. In this respect it bears a resemblance to equation (2.5). Equation (2.10) is also a recipe, this time for the behavior of f , where f follows the random walk.

Result (2.10) can be further generalized by assuming a function of the random variable S and of time, $f(S,t)$. With the use of partial derivatives (since there are two independent variables, S and t), we can expand $f(S+dS, t+dt)$ in a Taylor series about (S,t) to get

$$df = \frac{\partial f}{\partial S} dS + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} dS^2 + \cdots . \tag{2.11}$$

Using the expressions (2.5) for dS and (2.6) for dX^2 we find the expression for df is

$$df = \sigma S \frac{\partial f}{\partial S} dX + \left(\mu S \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + \frac{\partial f}{\partial t} \right) dt \tag{2.12}$$

As a simple example of the theory above, consider the function

$$f(S) = \log(S). \tag{2.13}$$

Differentiation of this function gives $\frac{df}{dS} = \frac{1}{S}$ and $\frac{d^2 f}{dS^2} = -\frac{1}{S^2}$. Thus using (2.10), we arrive at

$$df = \sigma dX + \left(\mu - \frac{1}{2} \sigma^2 \right) dt. \tag{2.14}$$

This is a constant coefficient stochastic differential equation, which says that the jump df is normally distributed. Now consider f itself: it is the sum of the jumps df (in the limit, the sum becomes an integral). Since a sum of normal variables is also normal, f has a normal distribution with mean $\left(\mu - \frac{1}{2}\sigma^2\right) t$ and variance $\sigma^2 t$ (Here, of course, $f_0 = \log S_0$ is the initial value of f.).

2.2.2 Arbitrage & the Value of an Option

Arbitrage is one of the fundamental concepts underlying the theory of financial derivative pricing and hedging.

It says that there are no opportunities to make an instantaneous risk-free profit higher than guaranteed payments from a bank account. More correctly, such opportunities cannot exist for a significant length of time before prices move to eliminate them.

The so-called risk-free investments are the investments that give guaranteed returns with no chance of default. A good approximation for such an investment is a government bond or a deposit in a bank.

The key words in the definition of arbitrage are 'instantaneous' and 'risk-free'. It tells us that, if one would like to have a greater return then one must accept a greater risk.

Two types of risk are usually defined in this context: specific and non-specific risk. Specific risk is the component of risk associated with a single asset(or a sector of the market, for example chemicals), whereas non-specific risk is associated with factors affecting the whole market. It is usually possible to diversify away specific risk by having a portfolio with a large number of assets from different sectors of the market; however, it is not possible to diversify away non-specific risk.

Now consider a European call option. If $S > E$ at expiry, one should exercise a call option to get the profit $(S - E)$. In converse situations, when $S < E$, one should not exercise the call option to avoid the loss of $(E - S)$. Hence, the value of the call option at expiry can be written as

$$C(S,T) = \max(S - E, 0); \tag{2.15}$$

As we get closer to the expiry date, T, the value of the call option approaches (2.15).

Similarly, the value/payoff of a put option at expiry is

$$P(S,T) = \max(E - S, 0). \tag{2.16}$$

The payoff for a so-called digital option is $BH(S,E)$, where B is a positive constant and $H(\cdot)$ is the Heaviside function, which has value 0 when its argument is negative but 1 otherwise.

By combining calls and puts with various exercise prices one can construct portfolios with a great variety of payoffs. For example:

For a portfolio of long one asset, long one put and short one call, its value is

$$\begin{aligned} \Pi &= S + P - C \qquad (2.17) \\ &= S + \max(E-S,0) - \max(S-E,0) \\ &= \begin{cases} S + E - S - 0 = E & \text{if} S \leq E \\ S + 0 - (S - E) = E & \text{if} S \geq E \end{cases}, \qquad (2.18) \end{aligned}$$

i.e., whether S is greater or less than E at expiry the payoff at expiry is always E. Thus this portfolio now worth $Ee^{-r(T-t)}$ at any time previous to the expiry. Hence,

$$S + P - C = Ee^{-r(T-t)}, \tag{2.19}$$

this relationship is known as put-call parity; The appeal of such strategies is in their ability to redirect risk.

For a "bullish vertical spread", buying a call and writing a call option with the same expiry date but a larger exercise price, its payoff can be written as

$$\max(S - E_1, 0) - \max(S - E_2, 0),$$

with $E_1 > E_2$;

2.3 Black-Scholes PDE

The Black-Scholes partial differential equation(PDE) can be derived under the following assumptions:

a. The asset price follows the log-normal random walk;

b. The risk-free interest rate r and the asset volatility σ are known functions of time over the life of the option;

c. There are no transaction costs associated with hedging a portfolio;

d. The underlying asset pays no dividends during the life of the option;

e. There are no arbitrage possibilities, which means all risk-free portfolios must earn the same return;

f. Trading of the underlying asset can take place continuously;

g. Short selling is permitted and the assets are divisible.

Itô's lemma then states that the random walk followed by the value of a portfolio reads

$$dV = \sigma S \frac{\partial V}{\partial S} dX + \left(\mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) + \frac{\partial V}{\partial t} dt; \tag{2.20}$$

Note that we require V to have at least one t derivative and two S derivatives.

Now one constructs a portfolio consisting of long one option and short a number, Δ, of the underlying asset. The value of this portfolio is thus

$$\Pi = V - \Delta S. \tag{2.21}$$

The jump in the value of this portfolio in one time-step therefore is

$$d\Pi = dV - \Delta dS. \tag{2.22}$$

Keeping Δ fixed during the time-step (otherwise $d\Delta$ would contain terms in $d\Pi$) and combining (2.5),(2.20) and (2.22) , we have

$$\begin{aligned} d\Pi &= \sigma S \left(\frac{\partial V}{\partial S} - \Delta \right) dX \\ &+ \left(\mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right) dt. \end{aligned} \tag{2.23}$$

Now, by taking

$$\Delta = \frac{\partial V}{\partial S}, \tag{2.24}$$

the random component in (2.3) can be eliminated to high order, and the resulting jump in the value of this portfolio is

$$d\Pi = \left(\frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right) dt. \tag{2.25}$$

Here we refer to the concepts of arbitrage. With the assumption of no transaction costs, the return on an amount Π invested in risk-less assets would give a growth of $r\Pi dt$ in time dt. Thus we have

$$r\Pi dt = \left(\frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} \right) dt. \tag{2.26}$$

Replacing Π by (2.21), Δ by (2.24), and dividing by dt we arrive at the famous Black-Scholes PDE:

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0. \tag{2.27}$$

Notice that the equation above does not contain the drift parameter μ. In other words, the value of an option is independent of how rapidly or slowly an asset grows in average. The only parameter from the stochastic differential equation (2.5) for the asset price that affects the option price is the volatility, σ. A consequence of this is that two people may differ in their estimates for μ yet still agree on the value of an option.

The Black-Scholes equation for the value of a European call (2.27) can be interpreted as an 'extended' diffusion equation.

The derivation of Black-Scholes Formula and its analytical solution are given in the Appendix.

2.4 Risk-neutral Valuation

The solution for the Black-Scholes equation for a European call reads

$$C(S,t) = SN(d_1) - Ee^{-r(T-t)}N(d_2), \tag{2.28}$$

with the cumulative distribution function $N(\cdot)$ for a standardized normal random variable, given by

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2} dy, \tag{2.29}$$

and

$$d_1 = \frac{\log(S-E) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \tag{2.30}$$

$$d_2 = \frac{\log(S-E) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \tag{2.31}$$

The form of the analytical solution reveals another important representation of Black-Scholes equation, the "risk-neutral valuation" formula:

$$V(t) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[V(T)], \tag{2.32}$$

where $\mathbb{Q}$ denotes risk-free measure. It says that the option value at time point t is simply the discounted expectation of the option value at the time point $T(T > t)$.

It is the starting point of a few important option pricing methods, like binomial method [Cox, Ross and Rubinstein, 1979], and the transform methods that form the basis for the solution methods in later chapters.

2.4.1 Risk-neutral Measure

The risk neutral probability, which is involved in the risk-neutral valuation formula, is the probability of an future event or state that both trading parties in the market agree upon.

What kind of information is offered from Risk-neutral probability?

Suppose that two parties A and B enter into a contract for a future event. The contract obliges A to pay B 1€ if the event happens and 0€ if it doesn't. For such an agreement, B has to pay for the right granted by the contract. If they agree that B pays 0.4€ to A, this means the two parties think that the probability of the event that happens is 40%. Otherwise, they won't reach that agreement and sign a contract. This price reflects the common beliefs towards the probability that the event happens. 40% is the risk neutral probability of the event that happens.

It is not any historical statistic or prediction of any kind. It is not the realistic probability, either.

How can we find this risk-neutral measure in the real world?

For the simple example mentioned above, once the price is established, the risk-neutral measure is also determined. That is because, whenever one has a pricing problem in which the event is measurable under this measure, one has to use this measure to avoid arbitrage. If one doesn't, arbitrage opportunities will occur.

A writer of an option wants to replicate the option value by trading in the underlying and a savings account. He wants the replicating portfolio to match the option value for any possible outcome. He doesn't care if the probability of an event is low or high.

Hence, in option pricing, one is able to change to another measure which assigns different measures than the real world measure to the same events.

The only restriction is that events which were impossible in the real world measure remain impossible in the changed measure(and vice versa). These measures are then called equivalent.

The price of the future claim is the cost of its replication; This is the essential idea of Cox-Rubinstein's binomial model in pricing an option.

2.4.2 Risk-neutral Valuation Theory

The change in probability measure from the true one to the risk-neutral one is due to the Cameron-Martin-Girsanov theorem [Baxter and Rennie, 1996].

In the risk-neutral probability measure, the expected stock return is r, risk-free interest rate from a bank.

Then the stock price process(2.5) changed to

$$dS_t = rS_t dt + \sigma S_t dW. \tag{2.33}$$

In this risk-neutral valuation, the price of a European call follows[3]

$$\begin{aligned} C(S_t, t) &= e^{-r\Delta t}\mathbb{E}_t^{\mathbb{Q}}[C(S_T, T)] \\ &= e^{-r\Delta t}\int_{-\infty}^{\infty} \max(S_T - E, 0) f(S_T|S_t) dS_T, \end{aligned} \tag{2.34}$$

where the transition probability density function is given by

$$f(S_T, T|S_t, t) = \frac{1}{S_T\sqrt{2\pi\sigma^2\Delta t}} e^{-\frac{[\log\frac{S_T}{S_t} - (r-\frac{1}{2}\sigma^2)\Delta t]^2}{2\sigma^2\Delta t}} \tag{2.35}$$

for underlying that follows Brownian motion.

With the stochastic calculus approach, the option pricing problem turns out to be identifying the transition probability then solving an integration; And this is the starting point of the so-called Transform methods, which will be under main concern in the following chapters.

2.5 American and Bermudan Option

An option that can be exercised anytime during its life is called American option, which is the majority of exchange-traded options. Since investors have the freedom to exercise their American options at any point during the life of the contract, they are more valuable than European options which can only be exercised at maturity.

A Bermudan option is a call or put option which can be exercised on several specified days during the life of the option. It is reasonable to say that Bermudan options are a hybrid of European options, which can only be exercised on the option expiry date, and American options, which can be exercised at any time during the option life time.

As a consequence, under same conditions, the value of a Bermudan option is greater than (or equal to) a European option but less than (or equal to) an American option.

There are basically two ways to solve an option pricing problem. One is by solving Black-Scholes PDE with proper boundary conditions; Another is starting from the risk-neutral valuation formula.

2.5.1 Free Boundary Problem

American options have the additional feature that exercise is permitted at any time during the life of the option. Due to this additional right, American options potentially have higher values than European options.

Take a put option as an example. The arbitrage concept imposes the constraint

$$V(S, t) \geq \max(S - E, 0), \tag{2.36}$$

for otherwise there would be an arbitrage risk-free profit $S - E - V$.

[3] [Harrison-Pliska, 1981]

Further, there may be some optional values of $S(t)$ for which the holders would exercise the American option early. If it were not so, then the option would have the European value, since the Black-Scholes equation would then also hold for all S.

Thus, the valuation of American options is not only about determining the option value, but also about at which value of S one should exercise the option. This kind of problem belong to the so-called free-boundary problems:

- The option value must be greater than or equal to the payoff function;

  This says that the arbitrage profit obtainable from early exercise must be equal to zero. For each time t, we must divide the S axis into two distinct regions: $0 \leq S \leq S_f(t)$, and $S_f(t) \leq S \leq \infty$.

  To a American put, its value, P, satisfies $P = E - S$ when S within the former interval between zero and the optimal exercise price, but $P > E - S$ when S rises up to the latter interval between the optimal exercise price and infinity.

- The Black-Scholes equation is replaced by an inequality, instead of the equality;

  We could set up the delta-hedged portfolio as before, with exactly the same choice of delta. The arbitrage argument for the European option no longer leads to a unique value for the return on the portfolio, only to an inequality. And the return from the portfolio cannot be greater than the return from a bank deposit.

  For an American put, this gives

  $$\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP \leq 0. \tag{2.37}$$

  More precisely,

  $$\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP = 0 \quad \text{when} \quad 0 \leq S \leq S_f(t), \tag{2.38}$$

  and

  $$\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP < 0 \quad \text{when} \quad S_f(t) \leq S \leq \infty. \tag{2.39}$$

- The option value must be a continuous function of S;

  This follows from simple arbitrage. If there were a discontinuity in the option value, and if this discontinuity persisted for more than infinitesimal time, a portfolio of options only would make a risk-free profit with probability 1 should the asset price ever reach the value at which the discontinuity occurred. The curve of the American option price versus asset price should be tangent to the payoff line of the corresponding European option at the point of optimal exercise price.

  For a American put, this tells us that

  $$\frac{\partial P}{\partial S}\big(S_f(t), t\big) = -1, \tag{2.40}$$

  for otherwise the option values for prices next to the optimal exercise price would contradict to the first constraint of $P > \max(S - E, 0)$ when the optimal exercise price $P = \max(S - E, 0)$ at $S_f(t)$.

- The option delta(the slope) must be continuous.

The American option pricing problem can be reformulated into a linear complementarity problem.(See Appendix for detailed derivation).

2.5.2 Lattice Mode

Another way to value American options is to approximate by the value of N-time exercisable Bermudan options, with N big. Because the more exercising days a Bermudan option has, the closer its value approaches an American option.

Notice that between two adjacent early exercise dates, a Bermudan option can be treated as a European option, whose 'initial time point' and the 'expiry' are respectively the two dates under discussion. Thus the risk-neutral valuation theory is applicable.

Here is the brief idea of pricing Bermudan or American options by risk-neutral valuation formula and the lattice mode:

1. Set up a lattice of the underlying price by quantizing the time dimension;

2. Risk-neutral valuation formula is applied in between two adjacent lattices. Value obtained by risk-neutral valuation formula are called continuation value and need to be compared with the present payoff for the elimination of possible arbitrage opportunities. The final value of the option on that lattice is thus the maximum of the continuation value and the present payoff.

3. The valuation procedure starts between the lattice on expiry and the lattice before expiry, and continues backwards in time, until the initial time is reached.

There are two general types of lattice modes. One quantizes the time dimension with very small time steps, like the binomial method; The other quantizes the time dimension only on the early exercise dates, thus the time step is not necessarily small, like Transform methods.

The shape of lattices used by different Transform methods also differ from one to another, the way to construct a lattice is closely related with the specific pricing method being used.

Chapter 3

Using Fourier Transform to Value an Option

The Black-Scholes partial differential equation(PDE)(2.27) describes the dynamics of an option price driven by its underlying. Hence, a straight forward idea is to solve the PDE (2.27) directly, with proper boundary conditions.

Another branch of methods departs from risk-neutral valuation formulae. For example, the widely applied binomial method falls into this category.

Recent papers[Heston 1993, Carr Madan 1999, Andricopoulos' 2002, O'Sullivan 2004, Bervoets 2005, and Lord 2006] presented new ways in solving risk-neutral valuation formula by means of fast Fourier transform techniques, hence named as "Transform Methods", a short term for "Fourier and Fast Fourier Transform Methods".

In practice, parameter calibration[1] needs to be performed fast, and the transform methods turn out to be typically faster than PDE methods, and have no stability concerns as in the binomial method.

We chose the most recently developed Transform method, the Lord method, as the starting point of the thesis research.

There is, however, a constraint between grid sizes in the price domain and in the frequency domain(presented in detail in 4.1) in the Lord method. Hence, the Fractional Fourier transform(FRFT) method is incorporated into the Lord method as to solve the problem.

Early pioneering articles on the application of transform techniques are concerned with vanilla option pricing problems. They include Heston (1993), Carr and Madan(1999).

Then Andricopoulos' (2002) introduced a 'QUAD' method to value a wide variety of options with exotic features, assuming the underlying follows geometric Brownian motion. The method is, however, slow as the discretized integration in the risk-neutral valuation formula is solved straight forwardly.

O'Sullivan (2004) combined the QUAD method with existing transform methods and came up with a 'QUAD-FFT' method that can be used to price Bermudan and American options with a wide variety of underlying processes. The density function can be recovered very fast due to the employment of FFT algorithm, but the overall solution on the risk-neutral valuation formula is still slow, $O(N^2)$ complexity (N the number of grid points).

Bervoets (2005) found a way to employ the FFT algorithm twice, once for recovery of the density function and once for solution of the risk-neutral valuation formula, so as to

[1]The crucial parameter volatility(σ) that is fed into PDE or other valuation formulas is supposed to be estimated from historical data. But it is hard to measure directly. However, it is plainly true that option prices are quoted in the market. Therefore one idea is to take the option prices in history back into the option pricing formulas to recover the market's opinion on the value of σ. This procedure is one example of parameter calibration.

have achieved an overall $O(NlogN)$ complexity.

Lord (2006) provided an even faster method that can be used to value early exercisable options for a wide variety of underlying processes. Its overall complexity is $O(NlogN)$, too, but is actually two to four times faster than Bervoets' method.

3.1 Fourier Transforms

There exist various conventions for the definition of Fourier transforms. In this thesis we use the following definitions:

- Continuous Fourier transform(CFT):

$$\mathcal{F}\{F(x)\} = \int_{-\infty}^{\infty} e^{-i\omega x} F(x)\, dx \triangleq \hat{F}(\omega), \tag{3.1}$$

  where the '$-$' sign in front of the imaginary unit i is a convention.

- Inverse continuous Fourier transform:

$$\mathcal{F}^{-1}\{\hat{F}(\omega)\} = \frac{1}{(2\pi)} \int_{-\infty}^{\infty} e^{ix\omega} F(x)\, dx; \tag{3.2}$$

- Discrete Fourier transform(DFT):

$$\mathrm{D}_k\{F_n\} = \sum_{n=0}^{N-1} e^{-i\frac{2\pi}{N}nk} F_n, \tag{3.3}$$

- Inverse DFT:

$$\mathrm{D}_k^{-1}\{F_n\} = \frac{1}{N} \sum_{n=0}^{N-1} e^{i\frac{2\pi}{N}nk} F_n. \tag{3.4}$$

- Fractional Fourier transform(FRFT):

$$G_k(\vec{x}, \gamma) = \sum_{n=0}^{N-1} x_n e^{-2\pi i n k \gamma}. \tag{3.5}$$

Before applying the Fourier transform to value options, it is important to be aware of the existence conditions of the Fourier transforms. If these conditions are not satisfied, we need to pre-multiply some well-behaved damping functions to the original ones to assure the existence.

- Existence of CFT

  Conditions for the existence of the continuous Fourier transform are complicated to state in general [D. C. Champeney,1987], but it is sufficient for $x(t)$ to be absolutely integrable, i.e.,

$$\| x \|_1 \triangleq \int_{-\infty}^{\infty} |x(t)|\, dt < \infty. \tag{3.6}$$

  This requirement can be stated as $x \in L^1$, meaning that x belongs to the set of all functions having a finite L^1 norm ($\| x \|_1 < \infty$). It is similarly sufficient for $x(t)$ to be square integrable, i.e.,

$$\| x \|_2^2 \triangleq \int_{-\infty}^{\infty} |x(t)|^2\, dt < \infty, \tag{3.7}$$

  or, $x \in L^2$. More generally, it suffices to show $x \in L^p$ for $1 \leq p \leq 2$.

- Existence of DFT

  The DFT operates on a sampled, or discrete, function value in the time domain, and generates from this a sampled, or discrete, spectrum in the frequency domain. The resulting spectrum is an approximation of the Fourier Series.

  If the sampling rate is high enough to ensure a reasonable representation of the shape of the function value, the DFT produces a spectrum very close to a theoretically true spectrum. This spectrum is also discrete.

3.2 Important Functions and Relationships

There are three important functions involved in all Transform methods. They are the risk-neutral probability density function, the cumulative distribution function and the characteristic function of a random variable.

The relationship between the three plays an important role in deriving various option pricing formulas departing from the risk-neutral valuation equation.

It is thus necessary to give their definitions and relations before we go deep into the Transform methods:

- Consider an option at time t with strike E and expiry T; let S denote the stock price at time $t(t \in [0, T])$; let S_T denote the stock price at time T.
- Define transformed prices as

$$s = \log(S_T), \quad k = \log(E). \tag{3.8}$$

- Let $f(s)$ denote the risk-neutral density function of the transformed stock price s.
- Let $\phi(v)$ denote its Fourier transform, or, in other words, the characteristic function of the transformed stock price s:

$$\phi(v) = \int_{-\infty}^{\infty} e^{ivs} \cdot f(s)\, ds. \tag{3.9}$$

- Let $F(s)$ denote the cumulative distribution function of s. It is the integration of the density function:

$$F(s) = \int_{-\infty}^{s} f(t)\, dt, \tag{3.10}$$

- Relationship between the cumulative distribution function $F(s)$ and the characteristic function $\phi(v)$ is given by Gil-Palaez Inversion[Gil-Palaez,1951]:

$$F(s) = \frac{1}{2} - \frac{1}{\pi} \int_{0}^{\infty} \mathrm{Re}\bigg(\phi(v)\frac{e^{-ivs}}{iv}\bigg) dv. \tag{3.11}$$

- Relationship between the density function $f(s)$ and the characteristic function $\phi(v)$ can be obtained by substituting $F(s)$ by $f(s)$ via (3.10) in (3.11), then take the first derivative with respect to v[Carr Madan, 1999]:

$$f(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{Re}\Big(\phi(v)e^{-ivs}\Big) dv = \frac{1}{\pi} \int_{0}^{\infty} \mathrm{Re}\Big(\phi(v)e^{-ivs}\Big) dv. \tag{3.12}$$

The relationship between the three important functions, $F(s)$, $f(s)$ and $\phi(v)$ is summarized in Fig. (3.1):

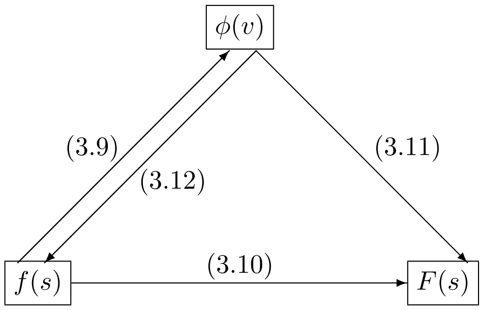


Figure 3.1: Relationships in between $F(s)$, $f(s)$ and $\phi(v)$

Among the three the characteristic function of a stochastic variable, say, z, is easiest to get, because:

- If $z = x + y$, then $\phi_z(v) = \phi_x(v)\phi_y(v)$.
- If $z = x/n$, then $\phi_z(v) = \phi_x(v/n)$.
- If $z = (x - \mu)/\sigma$, then $\phi_z(v) = e^{-iv\mu/\sigma}\phi_x(v/\sigma)$.

3.3 Basic Transform Methods

Let's take a European call as an example to illustrate how the methods based on Fourier transformation work.

Recall that the arbitrage-free price of a vanilla European option is given by the discounted expected payoff provided that the expectation is taken with respect to a risk-neutral measure:

$$\begin{aligned} C(t) &= e^{-r(T-t)} \cdot \mathbb{E}_t^{\mathbb{Q}}[\max(S_T - E, 0)], \\ &= e^{-r(T-t)} \cdot \int_{-\infty}^{\infty} \max(e^s - e^k, 0) \cdot f(s)\, ds, \end{aligned} \tag{3.13}$$

where $\mathbb{Q}$ denotes a risk-neutral measure.

$f(s)$ is only known for a limited number of stock price processes, including geometric Brownian motion, arithmetic Brownian motion and Ornstein-Uhlenbeck processes [O'Sullivan 2004].

In contrast, $\phi(v)$ is easier to obtain for many underlying processes[Hestion, 1993; Bakshi, 1997; Carr, Madan, 1999].

Therefore, the basic ideas of early transform methods are:

- Start from risk-neutral valuation, and substitute $f(s)$ in the integrand by expression of $\phi(v)$;
- Then apply a numerical integration rule to approximate the integral;
- Figure out DFT forms in the discretized formula, then solve it by employing FFT algorithm.

3.3.1 Gil-Palaez Inversion

The idea of the so-called Gil-Palaez inversion based Transform method is to replace the cumulative distribution function $F(s)$ in the analytical solution of Black-Scholes PDE for European call option by $\phi(v)$ via Gil-Palaez inversion.

Let's repeat the analytical solution of European call options in a more generalized form:

$$C(S,t) = \Delta \cdot S - Ee^{-r(T-t)} \cdot \mathrm{P}(S_T > E), \tag{3.14}$$

where $\mathrm{P}(S_T > E)$ is the risk-neutral probability of in-the-money, and delta Δ is $\partial C/\partial S$ (see Chap. 2.3).

Via Gil-Palaez inversion(3.11), $\mathrm{P}(S_T > E)$ can determined as [2] :

$$\mathrm{P}(S_T > E) = 1 - F(k) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\bigg(\frac{e^{-ivk}\phi(v)}{iv}\bigg)dv. \tag{3.15}$$

Similarly, the delta of the option can also be obtained as:

$$\Delta = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\bigg(\frac{e^{-ivk}\phi(v-i)}{iv\phi(-i)}\bigg)dv. \tag{3.16}$$

Thus in total, the value of a European call option is:

$$\begin{aligned} C(t) \;=\; & S\bigg[\frac{1}{2} + \frac{1}{\pi}\mathrm{Re}\bigg(\int_0^\infty e^{-ivk}\Pi_1(v)dv\bigg)\bigg] \\ & -Ke^{-r(T-t)}\bigg[\frac{1}{2} + \frac{1}{\pi}\mathrm{Re}\bigg(\int_0^\infty e^{-ivk}\Pi_2(v)dv\bigg)\bigg], \end{aligned} \tag{3.17}$$

with

$$\Pi_1(v) = \frac{1}{iv}\frac{\phi(v-i)}{\phi(-i)}, \quad \Pi_2(v) = \frac{1}{iv}\phi(v). \tag{3.18}$$

In practical applications, the characteristic functions of the logarithm of stock price are usually available, which means that $\Pi_1(v)$ and $\Pi_2(v)$ are available. Therefore, the problem of pricing a European call option turns out to lead to two integrations in (3.17).

In order to be able to compute the integral, quadrature rules are involved to discretize the integrals. We choose the trapezoidal rule.

However, the applicability of this rule is limited because that the existence of the Fourier transform is not guaranteed.

3.3.2 Carr-Madan Inversion

Carr and Madan(1999) included an exponential damping factor in their application of Fourier analysis to ensure the existence of the Fourier transform of European option prices.

Let's again take a European call option as an example to illustrate how the Fourier method based on Carr-Madan inversion works.

Define a so-called damped option value as

$$c(t) = e^{\alpha k}C(t), \quad \alpha > 0, \tag{3.19}$$

where the condition $\alpha > 0$ ensures the existence of the Fourier transform of modified price $c(t)$.

If let $\hat{c}(v)$ denote the Fourier transform of the damped price $c(t)$, then the Carr-Madan

[2]Detailed derivation see Bakshi and Madan(1999) and Scott(1997)

inversion based Transform method is defined by

$$\begin{aligned}
\hat{c}(v) &= \int_{-\infty}^{\infty} e^{ivk} c(t) dk = \int_{-\infty}^{\infty} e^{ivk} e^{\alpha k} C(t) dk \\
&= \int_{-\infty}^{\infty} e^{(\alpha+iv)k} \left[e^{-r(T-t)} \int_{-\infty}^{\infty} \max\left(e^s - e^k, 0\right) f(s) ds \right] dk \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} f(s) \int_{-\infty}^{s} \left(e^{s+\alpha k} - e^{(1+\alpha)k} \right) e^{ivk} dk ds \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} f(s) \left[\frac{e^{(\alpha+1+iv)s}}{\alpha+iv} - \frac{e^{(\alpha+1+iv)s}}{\alpha+1+iv} \right] ds \\
&= \frac{e^{-r(T-t)}}{(\alpha+iv)(1+\alpha+iv)} \phi\Big(v - (\alpha+1)i\Big), \qquad (3.20)
\end{aligned}$$

which shows a linear relationship between the characteristic function $\phi(v)$ and the Fourier transform of the damped option value $\hat{c}(v)$.

In other words, by Carr-Madan inversion, the Fourier transform of the damped option price is as easy to obtain as the characteristic function of the logarithm of the stock price.

Then, by the inverse Fourier transform one can recover the value of the option price:

$$C(t) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-ivk} \hat{c}(v) dv, \qquad (3.21)$$

This can be done fast by using an existing FFT algorithm, as long as the integral is discretized.

Early transform methods were able to handle general stochastic processes for asset prices and could also be applied for valuing exotic options such as barrier options, digital options etc. However, they cannot be applied to options with early exercise features.

3.4 QUAD-FFT Method

Alternatives have appeared in the literature: O'Sullivan's QUAD-FFT method[O'Sullivan, 2004], which was developed based on Andricopoulos' QUAD method[Andricopoulos, 2002] and basic transform methods.

3.4.1 QUAD Method

Andricopoulos (2002) presented a recursive model, which he named "QUAD method". He showed that options can be priced accurately using a discounted integration of the payoff, where the payoff is segmented so that the integration is only carried out over continuous segments. Integrations over each of the segments are approximated by applying a numerical quadrature rule.

By segmentation it is guaranteed that the quadrature rule is only applied to functions that are continuous and have continuous higher derivatives and thus accuracy is kept.

If there is more than one time step needed to price the option, an asset price lattice is set up and the price at each point in the lattice at time t is calculated recursively from the prices at time $t+1$ using a quadrature routine.

Take a Bermudan option as an example: Let the sequence $\{t_i\}, t_i \in [0, T],\ i = (1, 2, \cdots, N)$ denote a sequence of all possible exercising dates; We simply denote the stock price at time t_i by S_i.

Define the transformed prices(state variable) as

$$x = \log(S_i/E), \quad y = \log(S_{i+1}/E). \qquad (3.22)$$

By assumption, the stock price process follows geometric Brownian motion. Then by Itô's lemma, y is normally distributed(for a detailed derivation we refer to the previous chapter). Here, the risk-neutral conditional probability density function of y given x comes into play, which reads

$$f(y|x) = \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} e^{-\frac{\left(y-x-(r-D_0-\frac{1}{2}\sigma^2)\Delta t^2\right)}{2\sigma^2\Delta t}}, \tag{3.23}$$

where, as usual, r denotes risk-neutral interest rate, D_0 a continuous dividend, σ the volatility of the stock price process, and $\Delta t = t_{i+1} - t_i$.

The risk-neutral valuation of options gives the relation between option price at t_i and t_{i+1} in the following recursive form:

$$V(x, t_i) = e^{-r\Delta t}\mathbb{E}^{\mathbb{Q}}_{t_i}\left[V(y, t_{i+1})\right], \tag{3.24}$$

$$= e^{-r\Delta t}\int_{-\infty}^{\infty} V(y, t_{i+1}) f(y|x) dy. \tag{3.25}$$

Substitution of (3.23) into (3.25) gives the recursive formula for option prices:

$$V(x,t) = \frac{e^{-r\Delta t}}{\sqrt{2\pi\sigma^2\Delta t}}\int_{-\infty}^{\infty} V(y, t+\Delta t) e^{-\frac{\left(y-x-(r-q-\frac{1}{2}\sigma^2)\Delta t^2\right)}{2\sigma^2\Delta t}} dy, \tag{3.26}$$

which can be solved using any of the numerical integration rules.

For European options, only one step backward in time is needed to obtain the solution.

For multiply observed American or Bermudan options, the same formula can be applied for the time period from the last possibility of early exercise until maturity of the option. At each observation time t_i the option is priced for all x-values. These x-values then become the y-values used in the integration to find the x-values at a previous time step, t_{i+1}, until the option is valued completely.

The lattice mode employed in QUAD method is shown in Fig.(3.2). The lattices only lie on the exercise dates. And on each lattice, the stock price space is also quantized into grids.

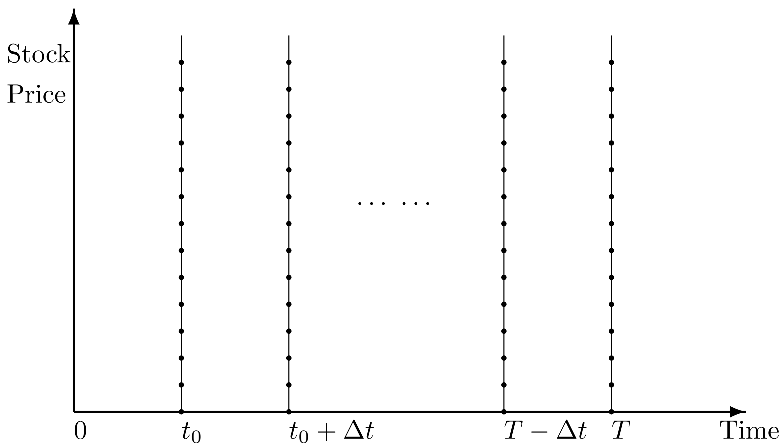


Figure 3.2: Lattice Mode for Bermudan and American Option Pricing

The QUAD method can be applied to value a wide variety of options.

There is, however, an important assumption in the QUAD method. The density function of the state variable is assumed to be known, like it is, for example, for geometric Brownian motion. Its applicability is thus limited due to the fact that the probability density $f(y|x)$ is only known for a restrictive class of underlying processes.

Moreover, the overall complexity of the algorithm is $\mathrm{O}(MN^2)$, with M the total number of time steps and N the grid size in price dimension.

3.4.2 QUAD-FFT Method

Inspired by the Fourier transform methods described previously, O'Sullivan(2004) obtained a simple representation of the probability density function, $f(x)$, of a state variable, x, via its characteristic function $\phi(v)$, with the help of the first derivative of the Gil-Palaez inversion:

$$\begin{aligned} f(y|x) &= \frac{dF(y|x)}{dy} = \Big(\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\Big(\phi(v|x)\frac{e^{-ivy}}{iv}\Big)dv\Big)' \\ &= \frac{1}{\pi}\int_0^\infty \mathrm{Re}\Big(e^{-ivy}\phi(v|x)dv\Big). \end{aligned} \tag{3.27}$$

A damping function is not used, because the density function obtained in this way is in L^1 thus needs not to be damped.

The y and v are discretized as

$$v_m = (m - \frac{N}{2})\Delta v, \qquad y_n = (n - \frac{N}{2})\Delta y,$$

with $n, m = 0, 1, \cdots, N$.

Appealing to the midpoint rule, the density can be approximated by

$$f\,(y_n|x) \approx \frac{1}{\pi}\mathrm{Re}\sum_{m=0}^{N}\Big(e^{-iv_m y_n}\phi(v_m|x)\Delta v\Big). \tag{3.28}$$

Set $\Delta v \Delta y = \frac{2\pi}{N}$, the density can then be fast resolved by the FFT algorithm:

$$f\,(y_n|x) \approx \frac{1}{\pi}(-1)^n\mathrm{Re}\sum_{m=0}^{N}\Big(e^{-i2\pi Nmn}\underbrace{(-1)^m\phi(v_m|x)\,\Delta v}_{z(m)}\Big) \tag{3.29}$$

$$= \frac{1}{\pi}(-1)^n\mathrm{Re}\left[\mathrm{DFT}(\vec{z})\right]. \tag{3.30}$$

Once the density is recovered, it is inserted back into the risk-neutral valuation formula, and the option value at a previous time point can be obtained by solving the integration numerically.

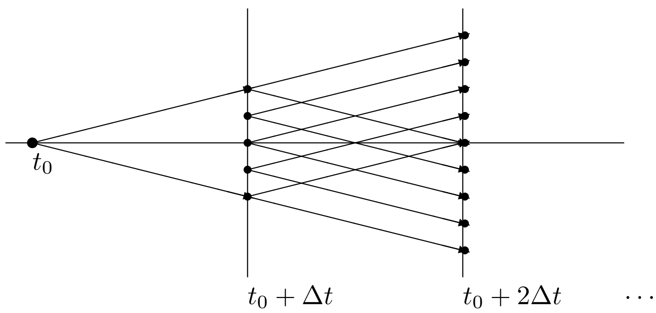


Figure 3.3: Lattice Mode of QUAD-FFT method

In pricing early exercisable options, lattices are set up on each exercisable dates, the above procedure is repeated backwards in time over each lattice, until the option value at initial time point is obtained.

The lattice mode in O'Sullivan's method is in a fashion as shown in Fig.(3.3). The number of grid points on each lattice is increased from 1 at initial time point to N at first early exercise date, to $2N-1$ at the second early exercise date, etc.. The grids are set centered at the initial stock price on each lattice.

Suppose that there are two time steps needed to price an option. Let x_0 be the initial transformed price at time t, $\vec{y}_1$ be the transformed price vector at time $t+\Delta t$ and $\vec{y}_2$ be the transformed price vector at time $t+2\Delta t$. Then $\vec{y}_1$ and $\vec{y}_2$ are defined as follows

$$\begin{aligned}\vec{y}_1 &= \{x_0 - q^*, x_0 - q^* + \Delta y, \cdots x_0 + q^* - \Delta y, x_0 + q^*\},\\ \vec{y}_2 &= \{x_0 - 2q^*, x_0 - 2q^* + \Delta y, \cdots x_0 + 2q^* - \Delta y, x_0 + 2q^*\}.\end{aligned}$$

The advantage of this type of lattice mode is that, rather than calculating $f(y_{2j}|y_{1i})$ for each $y_{2j} \in \vec{y}_2$ given each $y_{1i} \in \vec{y}_1$, one can simply use the density function estimated for $\vec{y}_1$ given x_0 and change the location of this density function so that it is centered at each $y_{1i} \in \vec{y}_1$ with a range of L standard deviations either side of y_{1i}. Of course, this can only be done if the time step between successive lattices is equal.

Due to the $\mathrm{O}(N \log N)$ complexity of the FFT algorithm, where N denotes the number of grid points in price dimension, the computational speed in solving the density function is also of $\mathrm{O}(N \log N)$.

However, the complexity of the computation at the second step of inserting back the recovered density into risk-neutral valuation formula is $\mathrm{O}(N^2)$.

Thus the overall complexity of QUAD-FFT method is $O(N^2)$ for European options, and $\mathrm{O}(MN^2)$ for M-time exercisable Bermudan options.

3.5 Bervoets Algorithm

Bervoets (2005) found out a way to employ the FFT in both the density recovery step and integration (in risk-neutral valuation formula) approximation step. Hence an overall $O(NlogN)$ complexity is obtained for European options, and $O(MN\log(N))$ complexity for M-time exercisable options.

The insight is, when discretising the integral in (3.31),

$$V(x_n, t_i) = e^{-r\Delta t}\int_{-\infty}^{\infty} V(y, t_{i+1}) f(y|x_n) dy, \quad n = 1, 2, \cdots, N \tag{3.31}$$

by a quadrature rule, e.g., Trapezoidal rule as

$$\begin{aligned}V(x_n, t_i) \approx\ & e^{-r\Delta t}\Delta y\Big(\sum_{i=0}^{N} V(y_k, t_{i+1}) f(y_k|x_n)\\ & -\frac{1}{2}(V(y_0, t_{i+1}) f(y_0|x_n) + V(y_N, t_{i+1}) f(y_N|x_n))\Big),\end{aligned} \tag{3.32}$$

it is possible to rewrite the equation in a matrix-vector product form:

$$\vec{v}_i \approx e^{-r\Delta t}\Delta y\left(\mathbf{F}\vec{v}_{i+1} - \frac{1}{2}\Big(V(y_0, t_{i+1})\vec{f}_1 + V(y_N, t_{i+1})\vec{f}_N\Big)\right), \tag{3.33}$$

with

$$\vec{v}_i = \begin{pmatrix} V(y_0, t_i) \\ V(y_1, t_i) \\ \vdots \\ V(y_{N-1}, t_i) \\ V(y_N, t_i) \end{pmatrix} \tag{3.34}$$

$$\vec{f_1} = \begin{pmatrix} f(y_0|y_0) \\ f(y_0|y_1) \\ \vdots \\ f(y_0|y_{N-1}) \\ f(y_0|y_N) \end{pmatrix}, \qquad \vec{f_N} = \begin{pmatrix} f(y_N|y_0) \\ f(y_N|y_1) \\ \vdots \\ f(y_N|y_{N-1}) \\ f(y_N|y_N) \end{pmatrix} \tag{3.35}$$

and

$$\mathbf{F} = \begin{pmatrix} f(y_0|y_0) & f(y_1|y_0) & \cdots & \cdots & f(y_N|y_0) \\ f(y_0|y_1) & f(y_1|y_1) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & f(y_{N-1}|y_{N-1}) & f(y_N|y_{N-1}) \\ f(y_0|y_N) & \cdots & \cdots & f(y_{N-1}|y_N) & f(y_N|y_N) \end{pmatrix} \tag{3.36}$$

Matrix $\mathbf{F}$ is a Toeplitz matrix if the process $y(k)$ is a so-called Lévy process[3], or a Brownian motion process.

That is,

$$\mathbf{F}_{j,k} = \mathbf{F}_{j+1,k+1}. \tag{3.37}$$

Since a Toeplitz matrix can easily be represented as a circulant matrix, the product of the Toeplitz matrix and a vector can efficiently be performed by applying a FFT algorithm.

With respective to errors, we have the same situation as for the previous method; In the density calculation step, the error sources are

- Truncation of the infinite integral in density equation (3.27);
- Discretization of the (truncated)finite integral using quadrature rules;
- Interpolation when the stock price does not lie on the grid.

Hence there are three main steps for this algorithm (Fig. (3.4)).

STEP 1: Use FFT algorithms to solve the density function; the complexity for this step is O($N\log(N)$).
STEP 2: Use FFT algorithms again to solve the matrix-vector product; the complexity for this step is O($N\log(N)$).
STEP 3: If there are M time steps, repeat the former two steps to solve the problem recursively backwards in time, until the initial time point is reached.

Figure 3.4: Basic Steps of Bervoets Method

Be aware that the obtained continuation values in STEP 2 need to be compared with the present payoff. The greater one is the real value for the option at that time point.

3.6 The Lord Method

In the Lord method one does not solve for the density function explicitly, instead, one keeps it in the risk-neutral valuation formula.

[3]see Bervoets(2005) for detailed proof

And it is based on the following assumption:

$$f(y|x) = f(y-x), \tag{3.38}$$

for state variable y at time point T given x at time point $t, t < T$.

Here are the derivation details of the method:

Take an European option as an example, its risk-neutral valuation formula reads

$$V(t,x) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,S_t}\left[V(T,y)\right], \tag{3.39}$$

and x and y are state variables at time t and $T(T > t)$, respectively.

By the definition of expectation, we have

$$V(t,x) = e^{-r(T-t)}\int_{-\infty}^{\infty} V(T,y)\cdot f(y|x)\ dy, \tag{3.40}$$

Together with the assumption (3.38), and with the rule of changing of variable, we have

$$e^{r(T-t)}V(t,x) = \int_{-\infty}^{\infty} V(T,x+z)\cdot f(z)\ dz, \tag{3.41}$$

where z is a new state variable that equals $(y - x)$.

Equation (3.41) is in the form of "cross-correlation"[4]. So we can employ the fast Fourier transform to get a fast resolution of $V(t)$.

Apply the Fourier transform on $V(t,x)$, after damping by $e^{\beta x}$ (to ensure the existence of the Fourier transform by properly choosing the value of β, which will be discussed in Chapter 5.3), we have

$$\begin{aligned}
& e^{r(T-t)}\mathcal{F}\{e^{\beta x}V(t,x)\} && (3.44)\\
&= e^{r(T-t)}\int_{-\infty}^{\infty} e^{i\omega x}e^{\beta x}V(t,x)dx && (3.45)\\
&= \int_{-\infty}^{\infty} e^{i\omega x}\left[\int_{-\infty}^{\infty} e^{\beta x}V(T,x+z)f(z)dz\right]dx, &&\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\omega x+\beta x}V(T,x+z)f(z)dzdx, &&\\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} e^{i(\omega-i\beta)x}V(T,x+z)dx\right]f(z)dz, &&\\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} e^{i(\omega-i\beta)(y-z)}V(T,y)dy\right]f(z)dz, &&\\
&= \int_{-\infty}^{\infty} e^{i(\omega-i\beta)y}V(T,y)dy\ \int_{-\infty}^{\infty} e^{i(-\omega+i\beta)z}f(z)dz && (3.46)\\
&= \hat{V}_T(\omega-i\beta)\ \phi(-\omega+i\beta), && (3.47)
\end{aligned}$$

where, as usual, $\phi(v)$ is the characteristic function of the density function $f(x)$:

$$\phi(v) = \int_{-\infty}^{\infty} e^{ivx}f(x)dx. \tag{3.48}$$

[4]The cross-correlation of two complex functions $f(t)$ and $g(t)$ of a real variable t, denoted $f \star g$ is defined by

$$\begin{aligned}
f \star g &\triangleq \bar{f}(-t) * g(t) && (3.42)\\
&= \int_{-\infty}^{\infty}\bar{f}(\tau)g(t+\tau)d\tau, && (3.43)
\end{aligned}$$

where '$*$' denotes the operator of convolution.

Next, the option price can be recovered by inverse Fourier transform as:

$$e^{r(T-t)}V(t,x) = e^{-\beta x}\,\mathcal{F}^{-1}\{\hat{V}_T(\omega - i\beta)\,\phi(-\omega + i\beta)\}. \tag{3.49}$$

From a "domain-switching point of view", the algorithm is sketched in Fig.(3.5).

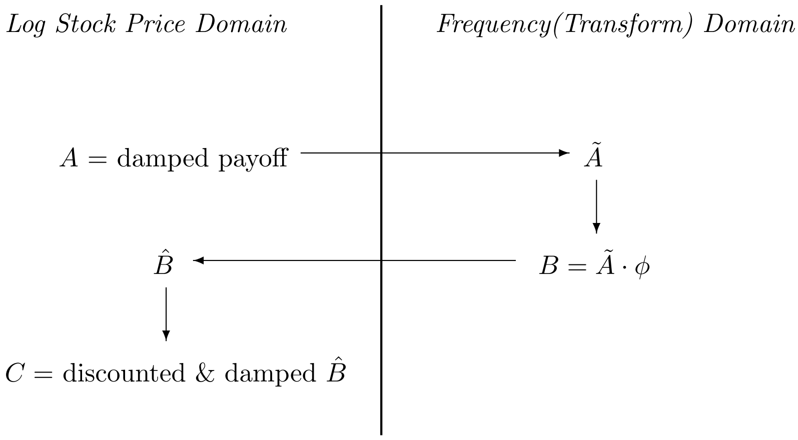


Figure 3.5: Domain Switching in the Lord Method

In order to realize the procedure in a digital computer, we need first to truncate the infinite integration in the Fourier transform in (3.49), then to discretize the finite integral by quadrature rules, and then to employ the FFT algorithm to solve the big sum part in the discretized formula (Details in the discretization and implementation are given in the next Chapter).

In summary, the algorithm of the Lord method is given by Figure 3.6:

> STEP 1: Derive the characteristic function $\phi(z)$;
>
> STEP 2: Apply Fourier transform on the damped option payoff/value;
>
> STEP 3: Take inverse Fourier transform on the product of the results from steps 1 and 2;
>
> STEP 4: If there are M time steps, repeat step 1 to 3 recursively and backwards in time. (Again, the true value of option is the maximum of the continuation value returned by STEP 3 and the present payoff.)

Figure 3.6: Algorithm of the Lord method

This method is most rapid among the choices tested[5]. Its complexity is $O(MN \log N)$, if there are M exercise dates during the option's life time.

With respect to error estimation, we have the same situation as for the previous methods; the error sources are

[5]Note that Bervoets' method is twice more expensive than the Lord method, because double sized vectors are dealt with in Bervoets' method.

- Truncation of the infinite integral;
- Discretization of the (truncated)finite integral using quadrature rules;
- Interpolation error if the stock price at initial time does not lie on the grid.

3.7 The FRFT Method

All previous methods suffer from the constraint on the mesh sizes in the option price dimension and in the frequency dimension, when the integral is discretized by quadrature rules, i.e.,

$$\Delta v = \frac{2\pi}{N\Delta x}, \tag{3.50}$$

Due to the dependence of Δv on Δy, the mesh size in the frequency domain Δv is fixed once the truncation range in the log-stock domain is fixed, and it has nothing to do with the number of grid points.

So a finer grid in the frequency domain is not possible with the Lord method.

One possible remedy is to use the fractional Fourier transform (FRFT) algorithm instead of the FFT[6].

3.7.1 FRFT Algorithm

The FRFT is a generalization of the DFT. It is defined on the N-long complex sequence $\vec{x} = (x_j, 0 \leq j < N)$ as

$$G_k(\vec{x}, \gamma) = \sum_{n=0}^{N-1} x_n e^{-2\pi i n k \gamma}. \tag{3.51}$$

The parameter γ is not restricted to rational numbers and in fact may be any complex number. Although this transform is defined for all integers k, we compute the first N nonnegative values, i.e. for $0 \leq k < N$. Straightforward evaluation of these N values using (3.51) required $8N^2$ floating point operations, provided the exponential factors have been precomputed.

The ordinary DFT and its inverse are special cases of the FRFT, where $\gamma = 1/N$ and $\gamma = -1/N$, respectively:

$$\begin{aligned} D_k(\vec{x}) &= \sum_{n=0}^{N-1} x_n e^{-2\pi i n k/N}, && (3.52) \\ &= G_k(\vec{x}, 1/N) \qquad 0 \leq k < N, && (3.53) \end{aligned}$$

and,

$$\begin{aligned} D_k^{-1}(\vec{x}) &= \frac{1}{N} \sum_{n=0}^{N-1} x_n e^{2\pi i n k/N}, && (3.54) \\ &= \frac{1}{N} G_k(\vec{x}, -1/N) \qquad 0 \leq k < N. && (3.55) \end{aligned}$$

Bailey(1995) introduced a fast version of the FRFT algorithm. It can be derived by replacing $2nk$ by $n^2 + k^2 - (k-n)^2$. The expression for the FRFT then becomes

[6] Another possibility is to use a zero-padding technique in the log-stock domain.

$$
\begin{aligned}
G_k(\vec{x},\gamma) &= \sum_{n=0}^{N-1} x_n e^{-\pi i[n^2+k^2-(k-n)^2]\gamma}, \\
&= e^{-\pi i k^2\gamma} \sum_{n=0}^{N-1} x_n e^{-\pi i n^2\gamma} e^{\pi i(k-n)^2\gamma}, \\
&= e^{-\pi i k^2\gamma} \sum_{n=0}^{N-1} y_n z_{k-n}, \qquad (3.56)
\end{aligned}
$$

where the N-long sequences $\vec{y}$ and $\vec{z}$ are defined by

$$
y_n = x_n e^{-\pi i n^2\gamma}, \quad z_n = e^{\pi i n^2\gamma}. \qquad (3.57)
$$

The summation (3.56) is in the form of a discrete convolution, which can be evaluated by DFT-based procedures. The usual DFT method evaluates circular convolutions with $z_{k-n} = z_{k-n+N}$. This condition is not satisfied here, but instead $z_{k-n} = z_{n-k}$ when $k - n < 0$. Hence there is a way to convert the summation into a circular convolution form.

One first selects an integer $p > N - 1$, and extends the sequences $\vec{y}$ and $\vec{z}$ to length $2p$ with

$$
\begin{aligned}
y_n &= 0 & N \le n < 2p, \qquad (3.58)\\
z_n &= 0 & N \le n < 2p - N \qquad (3.59)\\
z_n &= e^{\pi i(n-2p)^2\gamma} & 2p - N \le n < 2p.. \qquad (3.60)
\end{aligned}
$$

Notice that the first N values of $G_k(\vec{x},\gamma)$ satisfy

$$
G_k(\vec{x},\gamma) = e^{-\pi i k^2\gamma} \sum_{n=0}^{2p-1} y_n z_{k-n} \quad 0 \le k < N, \qquad (3.61)
$$

and also that the sequence $\vec{z}$ now satisfies the required property for a $2p$-point circular convolution.

Thus it follows that $2p$-point DFT's may be used to evaluate (3.61):

$$
\begin{aligned}
G_k(\vec{x},\gamma) &= e^{-\pi i k^2\gamma} D_k^{-1}(\vec{w}) \quad 0 \le k < N, \qquad (3.62)\\
\vec{w} &= D_k(\vec{y}) D_k(\vec{z}) \qquad 0 \le k < 2p. \qquad (3.63)
\end{aligned}
$$

It should be emphasized that $\vec{w}$, $\vec{y}$, and $\vec{z}$ in (3.62) and (3.63) are $2p$-long sequences, but after the final inverse DFT only the first N results are kept, while the remaining $2p - N$ results are discarded.

To compute a different N -long segment $G_{k+s}(\vec{x},\gamma), 0 \le k < N$, a slight modification of the above convolution procedure is needed. In this case, $\vec{z}$ is defined as

$$
\begin{aligned}
z_n &= e^{\pi i(n+s)^2\gamma} & 0 \le n < N, \qquad (3.64)\\
z_n &= 0 & N \le n < 2p - N, \qquad (3.65)\\
z_n &= e^{\pi i(n+s-2p)^2\gamma} & 2p - N \le n < 2p; \qquad (3.66)\\
&. & \qquad (3.67)
\end{aligned}
$$

STEP 1: Expand the vector of the sampled input function into double-sized one with zeros;

STEP 2: Construct two vectors $\vec{y}$ and $\vec{z}$ by means of Eq.(3.58) $\backsim$ (3.60);

STEP 3: Use FFT to both of the vectors generated in step 2;

STEP 4: Use IFFT to the product of the two outputs of step 3, and multiply with pre-computed coefficients;

Figure 3.7: Algorithm of FRFT

Now (3.61), (3.62) and (3.63) become

$$G_{k+s}(\vec{x},\gamma) = e^{-\pi i(k+s)^2\gamma}\sum_{n=0}^{2p-1} y_n z_{k-n} \quad 0 \le k < N, \tag{3.68}$$

$$G_{k+s}(\vec{x},\gamma) = e^{-\pi i(k+s)^2\gamma} D_k^{-1}(\vec{w}) \qquad 0 \le k < N, \tag{3.69}$$

$$\vec{w} = D_k(\vec{y})D_k(\vec{z}) \qquad 0 \le k < 2p;. \tag{3.70}$$

The remainder of the algorithm is unchanged.

The technique of converting the summation (3.61) into a circular convolution can also be understood as the embedding of a Toeplitz matrix into a larger circulant matrix, which admits evaluation using an FFT.

In summary:

- Note that the ordinary DFT and its inverse are special cases of the FRFT.
- It in fact employs the FFT in a crucial step so as to be able to keep a fast computation speed.
- Note that the exponential factors in (3.60) can be precomputed; the DFT of the z sequence can also be precomputed. Thus the cost of an N-point FRFT is about four times the cost of an N-point FFT.

3.7.2 Applicability in Option Pricing

Bailey(1993) showed that the FRFT-based technique can be profitably applied to the numerical evaluation of any integral transform where

1. Both the input function values and the output transform values are equally spaced;
2. A large fraction of input function is either zero or within machine tolerance;
3. Only a limited range of output is required.

In option pricing, the functions involved satisfy the three points above, because

- The payoff function usually has zero values within a continuous range of stock value.
- The value of the characteristic function of stock price process is often zero when stock price is far away from its expectation.

- The value of the output(option price) that is of interest is in a limited range of stock prices.

Hence, it is possible to gain good performance by employing FRFT methods in option pricing.

Chapter 4

Implementation Details

Implementation details of the following methods are given in this chapter:

- The Lord method;
- The Lord-FRFT method (replacement of FFT by FRFT);
- The Lord-extrapolation (a posteriori on the Lord method with 2-point Richardson extrapolation technique);
- The Lord-FRFT-extrapolation methods (a posteriori on the Lord-FRFT method with 2-point Richardson extrapolation technique).

4.1 Implementation of the Lord Method

Recall that the Lord method can be represented by the following compact form:

$$C(t,x) = e^{-r(T-t)} \cdot e^{-\beta x} \, \underbrace{\mathcal{F}^{-1}\{ \underbrace{\mathcal{F}\{e^{\beta y} V(T,y)\}}_{\text{step 1}} \cdot \phi(-\omega + i\beta) \}}_{\text{step 2}}, \tag{4.1}$$

where $t < T$, and both "step 1" and "step 2" are continuous-time Fourier transforms. Hence the main task of implementing the Lord method turns out to be solving the two Fourier transforms numerically.

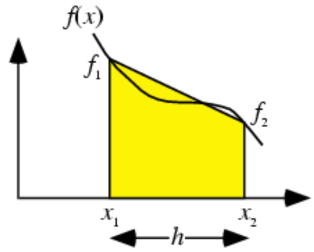
First of all, one needs to truncate the infinite integrals in the Fourier transforms. O'Sullivan (2004) has proposed a proper way to determine the truncation range, thus we follow his suggestions here: The total truncation range in the log-stock price domain is chosen to be $2L$ times the volatility (standard deviation[1] of logarithm) of the stock price with $L = 10$. The explanation is that, in the distance of 10 times standard deviation from the origin, the value of the probability density function would be smaller than the machine precision.

Then, one needs to discretize the integration by quadrature rules. We used the trapezoidal rule (Figure 4.1) throughout this thesis .

We can see that the "step 1" and "step 2" in (4.1) together actually return the results of a linear convolution on $V(T,y)$ and the density function. Hence, in the next step of

[1]The standard deviation is the square root of the variance, $\sqrt{var}$. And variance var can be computed for stochastic processes whose characteristic function $\phi(v)$ is known with

$$var = \left(-\frac{\partial^2 \phi(v)}{\partial v^2} \right)_{v=0} \tag{4.2}$$



$$\int_{x_1}^{x_2} f(x)dx = \frac{1}{2}h(f_1+f_2) - \frac{1}{12}h^3 f''(\xi),$$

where $f_i = f(x_i)$, h is the mesh width, and ξ is a point satisfying $x_1 \leq x_2$

Figure 4.1: Trapezoidal Rule

using DFT to approximate the continuous-time Fourier transform, we are about to use a circular convolution to approximate a linear one.

4.1.1 Linear Convolution

Consider a linear convolution of two sequences A and B,

$$A * B = \mathcal{F}^{-1}\{\mathcal{F}\{A\} \cdot \mathcal{F}\{B\}\}. \tag{4.3}$$

If we compute $A * B$ via a circular convolution by means of the DFT,

$$A \circledast B = DFT^{-1}\left(\ DFT(A) \cdot DFT(B)\ \right), \tag{4.4}$$

an artificial error would occur. A simple example illustrates why:
Suppose

$$A = (1 \quad 2 \quad 3 \quad 4)$$

and

$$B = (5 \quad 6 \quad 7 \quad 8),$$

then a direct calculation gives the outputs of their linear convolution as

$$A * B = (5 \quad 16 \quad 34 \quad 60 \quad 61 \quad 52 \quad 32).$$

However, their circular convolution through DFT gives

$$DFT^{-1}\left(DFT(A) \cdot DFT(B)\right) = (66 \quad 68 \quad 66 \quad 60).$$

The discrepancy between the two outputs is due to the fact that, there are only 4 digits available in the output sequence of the circular convolution while a 7-point sequence is given by the linear convolution, thus the circular convolution ‘overlaps’ the results of the linear one. In this example, point No. 5 overlaps point No. 1, No. 6 overlaps No. 2, and No. 7 overlaps No. 3:

$$(66 \quad 68 \quad 66 \quad 60) = (5 \quad 16 \quad 34 \quad 60)\ +\ (61 \quad 52 \quad 32 \quad 0).$$

One remedy to this problem is to pad zeros to one or both of the input sequences of the circular convolution so as to make room for the longer output from the linear convolution. For example, one could pad three zeros to both A and B to make room for the 7-length outcome[2] by setting

$$A' = (1 \quad 2 \quad 3 \quad 4 \quad 0 \quad 0 \quad 0)$$

[2]Zeros can also be padded in the beginning of a vector, or on both sides, but no in the middle

and

$$B' = (5 \quad 6 \quad 7 \quad 8 \quad 0 \quad 0 \quad 0),$$

then use the DFT to compute the circular convolution of A' and B' to get

$$A' \circledast B' = DFT^{-1}\left(DFT(A') \cdot DFT(B')\right) = (5 \quad 16 \quad 34 \quad 60 \quad 61 \quad 52 \quad 32) = A * B.$$

The correct values of the linear convolution are recovered in this way.

It is also shown, by the above example, that the length of the output vector of a linear convolution on two N-length vectors is $2N-1$, therefore $N-1$ zeros need to be padded to the two input vectors.

For pricing an option using the Lord method, this is not problematic. Padding fewer or even no zeros to the input vectors, we can still expect sufficiently accurate results. That is because of some pleasant properties of the two input functions, the payoff function and the probability density function:

- The payoff function is half zero and half non-zero for European options, so that the discretized payoff can be treated as having padded zeros to the non-zero half.

  We know that the payoff function is

  $$C = (S - E)^+, \tag{4.5}$$

  for a European call, and

  $$P = (E - S)^+. \tag{4.6}$$

  for a European put.

  Either call or put, the state variable $y = \log(S/E)$ is everywhere zero on one side of the origin. Thus no extra zeros need to be padded, and this makes life easier.

- The probability density function contains a large fraction of zeros on both sides of its center, hence, when the truncation range is big enough, the discretized sequence can be seen as zero-padding to both ends of a non-zero sequence.

- Even for early-exercisable options, no extra zeros need to be padded as long as a big truncation range of integration is taken.

  Early-exercisable (Bermudan and American) options can be treated as European options in between two adjacent exercise dates. Hence, option prices on the earlier date can be obtained by feeding the prices on the latter date, which also contain a large portion of zeros in the out-of-the-money direction, into the Lord formula (4.1). So if the accuracy is not required to be extremely high, like 10^{-7}, then padding a few or even no extra zeros would return accurate results with no extra memory allocation or computer time.

We'll confirm the above arguments by various experiments in the next chapter.

4.1.2 Grids and Lattices

The probability distribution of the logarithm of the stock price (named "log-stock" hereafter) is usually easier to get than the distribution density of the stock price itself. So we work in the log-stock domain by defining a state variable as

$$y = \log(S/E), \tag{4.7}$$

where S denotes the stock price and E the exercise price.

When setting up a grid in the log-stock domain, one needs to take care of the non-continuous and the non-smooth points. When applying the trapezoidal rule to approximate the integrals, extra(non-linearity) errors will enter if the non-continuous and non-smooth points do not lie on the grid. The non-smoothness in the payoff functions of vanilla options occurs wherever $S = E$, therefore we arrange the grid of y centered at zero to make sure that those points always lie on the grid:

$$y_p = (p - N/2)\Delta y, \quad p = 0, 1, \cdots, N-1, \tag{4.8}$$

where Δy is the mesh width, and

$$\Delta y = (2L \cdot vol)/N, \quad L \geq 10. \tag{4.9}$$

In the frequency domain, we center the grid around zero as well. Denoting the frequency variable by v, we have

$$v_n = (n - N/2)\Delta v, \quad n = 0, 1, \cdots, N-1. \tag{4.10}$$

where Δv needs to satisfy

$$\Delta v \Delta y = 2\pi/N, \tag{4.11}$$

as explained in the previous chapter.

There are two reasons to build the grids in the above manner.

1. The characteristic function of Geometric Brownian Motion has the shape of an up-side-down bell (sometimes with side tails) with the center at the origin. Hence, by centering the grids around zero all the non-zero values are placed in the middle of the sequence and the zeros at the two ends. If the truncation range is sufficiently large, the discretized data sequence will contain enough zeros to have the same effect as zero-padding in a later convolution computation.

2. The method is now easy to implement. With (4.8) and (4.10), we have symmetric structures in the log-stock and the frequency grids. Thus, there is no need to specify the direction of Fourier transforms, like 'log-stock to frequency' or 'frequency to log-stock'. Once the coding of the Fourier transform in one direction is done, it can be used to compute the other.

For Bermudan and American options, lattices are needed, and we use the same lattice as in the QUAD method (Figure 3.2): On each exercisable date, a lattice for the log-stock is set up, and the number of the grid points on all the lattices remains the same.

4.1.3 Algorithm

The main part of the implementation of the Lord method is to construct a sub-program that takes in a sample of an input function and returns a discretized data sequence of the continuous-time Fourier transform of the input function. Then the main program is constructed as shown in Figure 4.2, where the sub-program is called twice, once for returning the Fourier transform of the damped payoff and once for the inverse Fourier transform of the product of the characteristic function and the transformed payoff.

Now let's start constructing the sub-program. The continuous-time Fourier transform reads

$$\hat{f}(y) = \int_{-\infty}^{\infty} \underbrace{e^{-ivy} f(v)}_{\triangleq F(v)} dv, \tag{4.12}$$

then truncating the infinite integral we have,

$$\hat{f}(y) \approx \int_{U_1}^{U_2} F(v) dv, \tag{4.13}$$

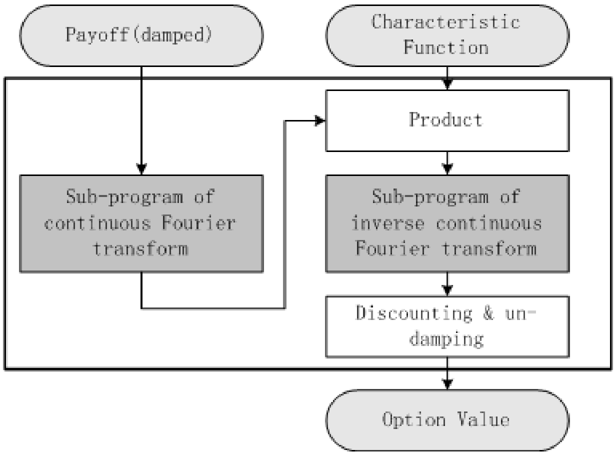


Figure 4.2: Implementation of the Lord Method in Pricing European Options

and applying the trapezoidal rule we have,

$$\hat{f}(y_p) = \sum_{n=0}^{N-1} F(v_n)\Delta v - \underbrace{\frac{1}{2}\Delta v[F(v_0) + F(v_{N-1})]}_{\triangleq \text{Tail}} + \mathrm{O}(\Delta v)^2. \tag{4.14}$$

Re-organizing the terms gives

$$\hat{f}(y_p) \approx \sum_{n=0}^{N-1} \left[e^{-i\cdot(n-\frac{N}{2})\Delta v\cdot(p-\frac{N}{2})\Delta y)} f(v_n) \right] \Delta v - \text{Tail} \tag{4.15}$$

$$= \sum_{n=0}^{N-1} \left[e^{-i\cdot np\cdot\Delta v\Delta y} \cdot e^{-i\Delta v\Delta y[-(n+p)\frac{N}{2}]} e^{-i\Delta v\Delta y\frac{N^2}{4}} f(v_n) \right] \Delta v$$
$$-\text{Tail}. \tag{4.16}$$

By including

$$\Delta v\Delta y = \frac{2\pi}{N}, \tag{4.17}$$

we finally have

$$\hat{f}(y_p) = (-1)^p \cdot \sum_{n=0}^{N-1} \left[e^{-i\cdot np\cdot\frac{2\pi}{N}} \cdot (-1)^{n+\frac{N}{2}} f(v_n) \right] \Delta v - \text{Tail}, \tag{4.18}$$

where the sum part can be rapidly computed by the FFT algorithm, such as the one provided by Matlab.

The subprogram algorithm is summarized in Figure 4.3.

STEP 1:	For n=0 : N-1 f[n] = (-1)^(n+N/2) * f[n]; End
STEP 2:	result = FFT(f) * dv;
STEP 3:	For p=0 : N-1 result[p] = (-1)^p * result[p] - Tail; End

Figure 4.3: Algorithm of sub-program that computes (4.12)

The integration

$$\tilde{f}(y) = \int_{-\infty}^{\infty} e^{ivy} f(v) dv, \tag{4.19}$$

that has a '+' sign in the exponent can be computed by the inverse DFT routine, analogously. In fact, we have the following relationship:

$$e^{ivy} = \overline{e^{-ivy}}. \tag{4.20}$$

Simple derivation leads to

$$\tilde{f}(y) = \int_{-\infty}^{\infty} \overline{e^{-ivy}} f(v) dv = \overline{\int_{-\infty}^{\infty} e^{-ivy} f(v) dv} = \overline{\hat{f}(y)}, \tag{4.21}$$

for a real valued function $f(v)$.

With the subprogram ready and the main program constructed as Figure (4.2) shows, we finished the implementation of the Lord method.

Due to the dependency of Δv on Δy, the mesh size in the frequency domain Δv is fixed once the truncation range in the log-stock domain is fixed, and it has nothing to do with the number of grid points:

$$\Delta v \cdot N = \frac{2\pi}{\Delta y}. \tag{4.22}$$

So a finer grid in the frequency domain is not possible with the Lord method. One remedy to this problem is to use the FRFT instead of the FFT.

One more advantage of the FRFT method is that zero-padding is automatically included in the FRFT algorithm. Higher accuracy can threfore be expected.

4.2 Implementation of the Lord-FRFT Method

The Lord-FRFR method is a short for "replacing FFT by FRFT in the Lord method".

The only thing that needs to be modified is the subprogram that computes the continuous-time Fourier transform.

In constructing the subprogram, we need to modify the grid in the frequency domain, as the mesh size is no longer dependent on the mesh size in log-stock domain. The grid setting in the log-stock domain remains.

4.2.1 Grids and Lattices

Set the grid in the frequency domain centered at zero:

$$v_n = (n - N/2)\Delta v, \quad n = 0, 1, \cdots, N-1,$$

with the mesh size Δv given by

$$\Delta v = \frac{v_{N-1} - v_0}{N}. \tag{4.23}$$

Fine grids in both log-stock and frequency domain can be obtained by increasing the number of grid points.

Recall the 'domain-switching' procedure in the Lord method: in the frequency domain the transformed payoff function is first multiplied by the characteristic function, then the product is transformed back to the log-stock domain. The truncation range in the frequency domain is therefore depending on the combination of the nonzero portions of the characteristic function and the nonzero portions of the transformed payoff.

Let's get an image impression via an example:

Example 1: an European option with

- $E = 100$(USD)
- $\Delta t = 1.0$(year)
- $\sigma = 0.2$
- $r = 0.03$
- $D_0 = 0.07$
- $\beta = 0$
- $L = 10$
- Stock price follows Brownian Motion process, whose $vol = \sigma\sqrt{\Delta t}$
- No. of grid points $N = 1024$

The characteristic function of the example option has a very narrow nonzero part, and as shown in Figure 3.5, out of the range of [-40, 40] the value of the characteristic function is zero.

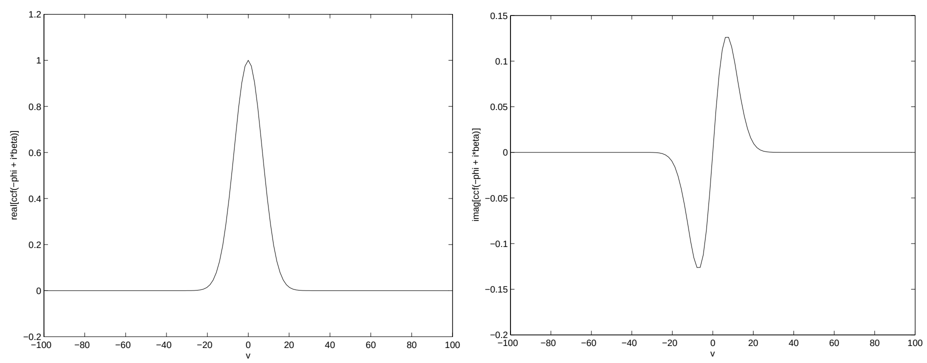


Figure 4.4: Real and Image of ϕ in Example 1

Thus we set the truncation range in the frequency domain as

$$v \in [-40,\ 40], \quad \text{when employing the FRFT algorithm.}$$

As a comparison, by $\Delta v = 2\pi/(N\Delta y)$ in the Lord method, the total range involved in the frequency domain would be

$$v \in [-803.4816,\ 801.9123], \quad \text{when employing the FFT algorithm.}$$

The mesh size in the given example is $\Delta y = 2L \cdot vol/N = 0.0039$, and

- $\Delta v = 1.5693$ by the Lord method, but
- $\Delta v = 0.07813$ by the Lord-FRFT method.

For very small Δt, like 0.1 year, the nonzero part of the characteristic function would grow, in that case the truncation range in the frequency domain should be set to the minimum of the width of the nonzero portion in the characteristic function and the width obtained by the Lord method.

4.2.2 Algorithm

The FRFT-based algorithm is derived basically in the same way as the FFT-based. Starting from the definition of the continuous-time Fourier transform in (4.12), truncating the integration range and applying the trapezoidal rule we get to (4.16). Then define

$$\gamma = \frac{\Delta v \Delta y}{2\pi}. \tag{4.24}$$

Replacing $(\Delta v\Delta y)$ by $(2\pi\gamma)$ in (4.16), we have

$$\begin{aligned} \hat{f}(y_p) &= \sum_{n=0}^{N-1}\left[e^{-i\cdot np\cdot 2\pi\gamma}\cdot \underbrace{e^{-i2\pi\gamma[-(n+p)\frac{N}{2}+\frac{N^2}{4}]}}_{\triangleq \text{Coeff}} f(v_n)\right]\Delta v - \text{Tail} && (4.25)\\ &= G_k(\text{Coeff}\cdot f,\ \gamma)\cdot \Delta v - \text{Tail}, && (4.26)\end{aligned}$$

where

$$G_k(\vec{x},\gamma) = \sum_{n=0}^{N-1} x_n e^{-2\pi i n k\gamma}$$

is the fractional Fourier transform defined in (3.51). Together with the algorithm of FRFT in Figure 3.7, the algorithm can be constructed as in Figure 4.5:

STEP 1: f = Coeff * f;
STEP 2: γ = dv * dy /2π;
STEP 3: result = FRFT(f, γ) * dv;
STEP 4: result = result - Tail;

Figure 4.5: Algorithm of the Sub-program by FRFT

Then the algorithm of the Lord-FRFT method can be constructed as in Figure 4.6:

4.3 Richardson Extrapolation

Both the accuracy of the numerical results from the Lord method and the Lord-FRFT method can be further improved by including a 2-point Richardson extrapolation technique in valuing European options, or by a repeated Richardson extrapolation technique in valuing American options.

STEP 1: A = damped payoff;
STEP 2: $\tilde{A}$ = conjugate of subprogram(A);
STEP 3: $B = \tilde{A} \cdot \phi$
STEP 4: $\hat{B}$ = subprogram(B)
STEP 5: result = un-damping and discounting on $\hat{B}$.

Figure 4.6: Algorithm of FRFT-Lord Method

4.3.1 Richardson Extrapolation Formulas

Richardson Extrapolation is an efficient technique to generate highly accurate results by using low-order formulas, for which the order of the truncation error is known. It is applicable to such situations that the error depends on a parameter, like the mesh size h.

To illustrate the procedure assume we have an approximation $N(h)$ to some quantity Q. And we know what the explicit expression is for the first few terms in the truncation error.

$$Q = N(h) + K_1 h + K_2 h^2 + K_3 h^3 \cdots \tag{4.27}$$

where K_i's are known constants. Here $N(h)$ is an approximation to Q, to $O(h)$. Assume $h > 0$ arbitrarily chosen such that for $h \to 0$ a better approximation is obtained.

Consider the case when the step size is halved to $h/2$, then

$$Q = N\left(\frac{h}{2}\right) + K_1 \frac{h}{2} + K_2 \frac{h^2}{4} + K_3 \frac{h^3}{8} + \cdots \tag{4.28}$$

Multiply (4.28) by 2 and subtracting (4.27) we obtain

$$Q = \left[N\left(\frac{h}{2}\right) + \left(N\left(\frac{h}{2}\right) - N(h)\right)\right] + K_2\left(\frac{h^2}{2} - h^2\right) + K_3\left(\frac{h^3}{4} - h^3\right) + \cdots$$

In this way the $O(h)$ term is eliminated.

Let

$$N_1(h) \triangleq N(h) \tag{4.29}$$

and

$$N_2(h) \triangleq N_1(\frac{h}{2}) + \left[N_1(\frac{h}{2}) - N_1(h)\right]. \tag{4.30}$$

Now, to the order of $O(h^2)$ we have

$$Q = N_2\left(h\right) - \frac{K_2}{2}h^2 - \frac{3K_3}{4}h^3 \cdots \tag{4.31}$$

Replacing h by $h/2$, we have

$$Q = N_2\left(\frac{h}{2}\right) - \frac{K_2}{8}h^2 - \frac{3K_3}{32}h^3 \cdots \tag{4.32}$$

Subtracting (4.31) from 4 times (4.32) the $O(h^2)$ is thus eliminated:

$$3Q = 4N_2\left(\frac{h}{2}\right) - N_2(h) + \frac{3K_3}{4}\left(-\frac{h^3}{2} + h^3\right) + \cdots \tag{4.33}$$

Now we have $O(h^3)$ approximation:

$$N_3(h) \triangleq N_2\left(\frac{h}{2}\right) + \frac{N_2(h/2) - N2(h)}{3}. \tag{4.34}$$

The process can be continued by constructing the $O(h^4)$ approximation:

$$N_4(h) = N_3\left(\frac{h}{2}\right) + \frac{N_3(h/2) - N_3(h)}{7} \tag{4.35}$$

Then $O(h^5)$ approximation is

$$N_5(h) = N_4\left(\frac{h}{2}\right) + \frac{N_4(h/2) - N_4(h)}{15} \cdots \tag{4.36}$$

And so on. In summary, the repeat Richardson extrapolation reads:

$$N_j(h) = N_{j-1}\left(\frac{h}{2}\right) + \frac{N_{j-1}(h/2) - N_{j-1}(h)}{2^{j-1} - 1}. \tag{4.37}$$

4.3.2 European Option Pricing

Pricing European options using the Lord or the Lord-FRFT method, the following errors are brought in during the discretization process[3]:

1. Truncation error when truncating the infinite integration. It can be well controlled by taking a reasonably large integration range.
2. The quantization error introduced by the trapezoidal rule, which is known to be of order $O(h^2)$ and is typically larger than the errors in item 1.

Thus the total error is dominated by the trapezoidal rule. Therefore the $O(h^2)$ error can be eliminated by applying the 2-point Richardson extrapolation formula (4.37) with $j = 2$.

We hence constructed algorithms for 'Lord-extrapolation' and 'Lord-FRFT-extrapolation' methods, by executing the Lord or the Lord-FRFT method twice, once with mesh size h and once with $h/2$, then we employ (4.37) with j set as 2 to get more accurate results.

4.3.3 American Option Pricing

Numerically, the value of an American option can be approximated either by a Bermudan option with many early exercise dates or by Richardson extrapolation on a series of Bermudan options with increasing early exercise dates.

The idea is an important contribution by Geske and Johnson (1984). In their original paper, Geske-Johnson showed that an American put option can be calculated to a high degree of accuracy using a Richardson approximation. If let $P(n)$ denote the price of a Bermudan option with n exercise dates, then, for example, using $P(1)$, $P(2)$ and $P(3)$, the price of the American put is approximately

$$V = P(3) + \frac{7}{2}(P(3) - P(2)) - \frac{1}{2}(P(2) - P(1)). \tag{4.38}$$

Omberg(1987) suggested a modified Geske-Johnson formula as

$$V = P(4) + \frac{5}{3}(P(4) - P(2)) - \frac{1}{3}(P(2) - P(1)), \tag{4.39}$$

[3]There would be an additional interpolation step needed as long as the stock price of interest does not lie on the grid. In that case the interpolation error would be brought in, too.

which is aimed to handle situations when $P(3) < P(2) > P(1)$ and the original Geske-Johnson formula (4.38) is not applicable.

Moreover, several researchers further modified the above two extrapolations for a smooth convergence in the binomial model, like Omberg(1987), Breen(1991) and Broadie and Detemple(1996).

Besides, Chang, Chung, and Stapleton (2001) proposed to employ a repeated Richardson approximation to improve the accuracy of an approximation to the unknown true option price. Their algorithm departs from the assumption that the value of an American option has the form

$$V(t) = a_0 + a_1 t^{\eta} + a_2 t^{\eta \cdot 2} + \cdots + a_k t^{\eta \cdot k} + \cdots , \tag{4.40}$$

with known η but unknown $a_0, a_1, \ldots, a_k, \ldots$. They employed the repeated Richardson extrapolation and ended up with the following algorithm:
For $j = 1, 2, 3, \ldots$, set $A_{j,0} = V(t_j)$, and compute for $m = 1, 2, 3, \ldots, k-1$ [4]:

$$A_{j,m} = A_{j+1,m-1} + \frac{A_{j+1,m-1} - A_{j,m-1}}{(t_i/t_{i+m})^{\eta} - 1} \tag{4.41}$$

Taking $\eta = 1$, $m = 3$ and $k = 4$, and $t_1 = T$, $t_2 = T/2$, $t_3 = T/4$, and $t_4 = T/8$, where T denotes the maturity of option. The algorithm is shown in Figure 4.7

$A_{1,0} = P(1)$	$A_{1,1} = 2A_{2,0} - A_{1,0}$	$A_{1,2} = \frac{4A_{2,1} - A_{1,1}}{3}$	$A_{1,3} = \frac{8A_{2,2} - A_{1,2}}{7}$
$A_{2,0} = P(2)$	$A_{2,1} = 2A_{3,0} - A_{2,0}$	$A_{2,2} = \frac{4A_{3,1} - A_{2,1}}{3}$	
$A_{3,0} = P(4)$	$A_{3,1} = 2A_{4,0} - A_{3,0}$		
$A_{4,0} = P(8)$			

Figure 4.7: Approximate American option values with 4 Bermudan options

[4]The constraint on $0 < m \leq k - 1$ is according to Schmidt(1968).

Chapter 5

One-Asset Option Pricing

Numerical results of pricing one-asset European, Bermudan and American options using the following methods are given in this chapter:

- The Lord method
- The Lord-FRFT method
- The Lord-extrapolation method[1]
- The Lord-FRFT-extrapolation method

Two stochastic models for the underlying stock prices are used, Brownian motion and the Variance Gamma.

Accuracy and computational time of the four methods are summarized and compared.

Then discussions on the choices of the damping factors are given.

The chapter ends up with a conclusion on the pros and cons of the methods.

5.1 Geometric Brownian Motion

In Chapter 2.1 Geometric Brownian Motion(GBM) has been introduced as one basic model for the movements of the underlying asset(stock) price. Equation (2.5) is the dynamic model of stock price's random walk that follow GBM, and under the risk-neutral measure it becomes (2.33):

$$dS = rSdt + \sigma SdW,$$

where W is a Wiener process.

5.1.1 Characteristic Function

Suppose f is a function of the stock price, by Itô's lemma we have

$$\begin{aligned} df &= \frac{df}{dS}\left(\sigma SdW + rSdt\right) + \frac{1}{2}\sigma^2 S^2 \frac{d^2 f}{dS^2} dt \\ &= \sigma S \frac{df}{dS} dW + \left(rS\frac{df}{dS} + \frac{1}{2}\sigma^2 S^2 \frac{d^2 f}{dS^2}\right) dt. \end{aligned} \tag{5.1}$$

If $f = \log(S)$, we have

$$df = \sigma dW + \left(r - \frac{1}{2}\sigma^2\right) dt. \tag{5.2}$$

[1]Here 'extrapolation' denotes the usage of the 2-point Richardson extrapolation technique in (4.37) with $j = 2$.

Hence, $\log(S)$ follows the normal distribution with mean value equal to $\left(r - \frac{1}{2}\sigma^2\right)\Delta t$ and the standard deviation equal to $\sigma\sqrt{\Delta t}$.

So the characteristic function of $\log(S)$ is

$$\phi(v) = e^{iv\bar{\mu} - \frac{1}{2}\bar{\sigma}^2 v^2}, \tag{5.3}$$

with

$$\bar{\mu} = \left(r - \frac{1}{2}\sigma^2\right)\Delta t, \quad \bar{\sigma} = \sigma\sqrt{\Delta t}. \tag{5.4}$$

5.1.2 Results of European Call

Now consider three European call options. Assume their underlying stock prices all follow GBM, their exercise prices E are all 100€, the life time of the options is 0.5 years, and their (σ, r, D_0) are equal to

- (0.2, 0.03, 0.07),
- (0.4, 0.03, 0.07), and
- (0.3, 0.0, 0.07),

respectively.

The initial stock prices are assumed to be 80€, 90€, 100€, 110€, and 120€, respectively.

Numerical results are compared with the analytical solution of the Black-Scholes PDE, and the root mean squre error(RMSE)[2] is calculated as an indicator of the accuracy of the methods.

In the first step of implementation, the number of grid points are set the same as O'Sullivan(2004) did in his paper, then we evaluate the results generated by the Lord, the Lord-FRFT, the Lord-extrapolation, and the Lord-FRFT-extrapolation methods, and compare them with O'Sullivan's.

Table 5.1: Prices of European call options under Geometric Brownian Motion with σ =0.2, r =0.03 and D_0 =0.07

S	BS	Lord	FRFT	Lord+	FRFT+	Q-FFT
80	0.2148	0.2148	0.2147	0.2148	0.2148	0.2149
90	1.3451	1.3449	1.3449	1.3451	1.3451	1.3444
100	4.5778	4.5774	4.5774	4.5778	4.5778	4.5778
110	10.4208	10.4204	10.4204	10.4208	10.4208	10.4205
120	18.3024	18.3022	18.3022	18.3024	18.3024	18.3017
RMSE		2.50e-04	2.51e-04	4.24e-06	2.69e-07	4.65e-04

(E = €100, $\Delta t = 0.5$ years, $L = 10$ and $N = 512$)

The results are listed in Tables 5.1, 5.2 and 5.3, where 'BS' represents the analytical solution of the Black-Scholes equation, 'Lord' represents results of the Lord method,

[2] The RMSE is calculated w.r.t. the analytical solution from Black-Scholes PDE. For example, RMSE for results computed by the Lord method is

$$RMSE_{Lord} = \sqrt{\frac{\sum_{n=0}^{N-1}(V_{BS}[n] - V_{Lord}[n])^2}{n}}.$$

Table 5.2: Prices of European call options under Geometric Brownian Motion with σ =0.4, r =0.03 and D_0 =0.07

S	BS	Lord	FRFT	Lord+	FRFT+	Q-FFT
80	2.6506	2.6502	2.6502	2.6506	2.6506	2.6511
90	5.6221	5.6215	5.6215	5.6221	5.6220	5.6228
100	10.0211	10.0204	10.0214	10.0211	10.0211	10.0211
110	15.7676	15.7669	15.7669	15.7676	15.7676	15.7678
120	22.6502	22.6496	22.6496	22.6502	22.6502	22.6509
RMSE		6.20e-04	6.20e-04	1.34e-07	7.86e-08	9.37e-04

(E = €100, $\Delta t = 0.5$ years, $L = 10$ and $N = 512$)

Table 5.3: Prices of European call options under Geometric Brownian Motion with σ =0.3, r =0.0 and D_0 =0.07

S	BS	Lord	FRFT	Lord+	FRFT+	Q-FFT
80	1.0064	1.0062	1.0062	1.0064	1.0064	1.0058
90	3.0041	3..0037	3.0037	3.0041	3.0041	3.0040
100	6.6943	6.6938	6.6938	6.6943	6.6943	6.6943
110	12.1661	12.1655	12.1655	12.1661	12.1661	12.1666
120	19.1555	19.1550	19.1550	19.1554	19.1554	19.1555
RMSE		4.29e-04	4.40e-04	5.58e-07	9.18e-08	3.52e-04

(E = €100, $\Delta t = 0.5$ years, $L = 10$ and $N = 512$)

'FRFT' the Lord-FRFT method, 'Lord +' the Lord-extrapolation method, 'FRFT +' the Lord-FRFT-extrapolation method, and 'Q-FFT' represents results from the QUAD-FFT method by O'Sullivan's.

Then the RMSE of all the results generated by each method is calculated and summarized in Table 5.4. It shows that all our four methods give very accurate results, which are more accurate than the QUAD-FFT method.

With $N = 512$ the errors by the Lord and Lord-FRFT methods are of degree 10^{-4}, the errors by the Lord-extrapolation method are of degree 10^{-6}, and the ones by the Lord-FRFT-extrapolation method are of degree 10^{-7}.

Hence, the Lord-FRFT-extrapolation method stands out with remarkably high accuracy; followed by the Lord-extrapolation method, the Lord method and the Lord-FRFT method.

Table 5.4: RMSE comparison

	Lord	FRFT	Lord+	FRFT+	Q-FFT
RMSE	4.58e-04	4.62e-04	2.47e-06	1.72e-07	6.37e-04

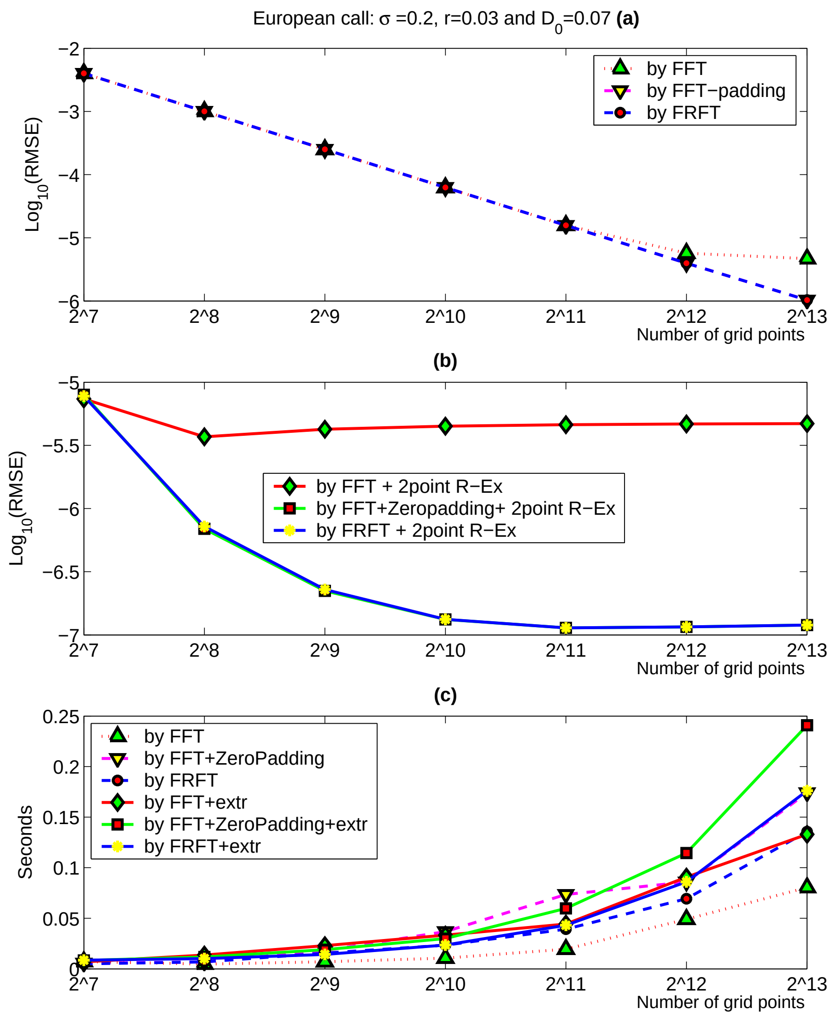


Figure 5.1: Errors and the computer time in Pricing European call under GBM with $\sigma = 0.2$, $r = 0.03$ and $D_0 = 0.07$

In the second step of implementation, we further examined their error convergence and computer time.

Besides, we also included a zero-padding technique to the Lord method to confirm our arguments in chapter 4.1.1, that no explict zero-padding needs to be incorporated to the Lord method to obtain sufficiently accurate results.

The computer we used for the experiments has a Pentium-4 CPU with 2.8 GHz frequency, and 512M DDR memory and 1M cache.

The results are plotted into Figure 5.1, 5.2 and 5.3. They show that:

1. No explict zero-padding needs to be incorporated in the Lord method and sufficiently accurate results (e.g., 10^{-4}) can be obtained by the Lord method.

2. The Lord-FRFT method gives almost the same results as the Lord method combined with explicit zero-padding. If very high accurate results (e.g., 10^{-7}) are expected, the Lord-FRFT method is preferable because it takes less CPU time than the zero-padding Lord method.

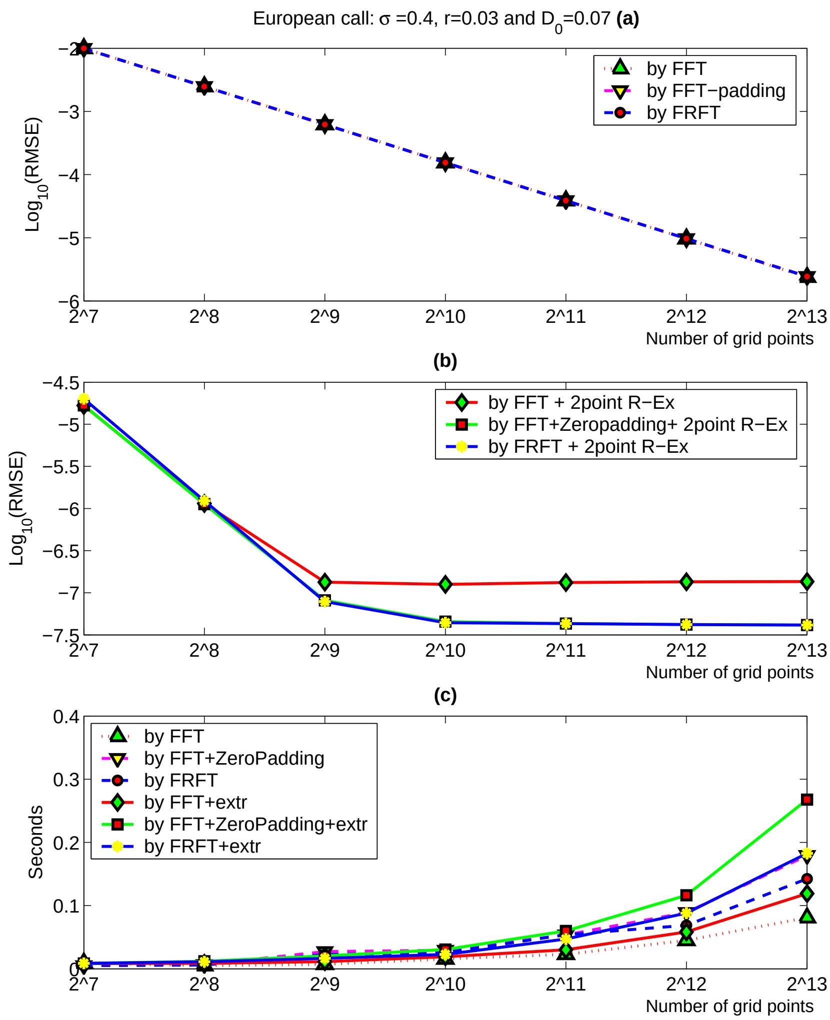


Figure 5.2: Errors and the computer time in pricing European call under GBM with $\sigma = 0.4$, $r = 0.03$ and $D_0 = 0.07$

3. The Lord-method stops to converge when the truncation error is small. Then 2-point Richardson extrapolation can't improve the accuracy further.

   For example, in Figure 5.1, the Lord-extrapolation method ('FFT+ 2point R-Ex' in (b) plots) gives a constant error of 10^{-5}, while the Lord-FRFT-extrapolation method ('FRFT+ 2point R-Ex' in (b) plots) continues to converge till the error introduced by truncating the infinite integral dominates.

4. The Lord-FRFT-extrapolation method gives remarkably accurate results with only a small number of grid points.

   For example, its error is of degree 10^{-6} in Figure 5.1, and 10^{-5} in Figures 5.2 and 5.3, with 128 grid points.

5. All four methods are remarkably fast! All of them finished the computation with 1024 grid points within 0.03 seconds(which can be seen from **(c)** plot in the three Figures). The differences in computer times among the four are small.

If ordered from the fastest to the slowest, the Lord method is followed by the Lord-FRFT, the Lord-extrapolation, and the Lord-FRFT-extrapolation method.

The experiment results confirm the complexity analysis that the overall complexity of the Lord method, $O(NlogN)$, is faster than O'Sullivan's method ($O(N^2)$).

6. If considering the overall performances (both accuracy and computational speed), the preferred choice is the Lord-FRFT-extrapolation method, followed by the Lord-extrapolation method, the Lord method, and the Lord-FRFT method.

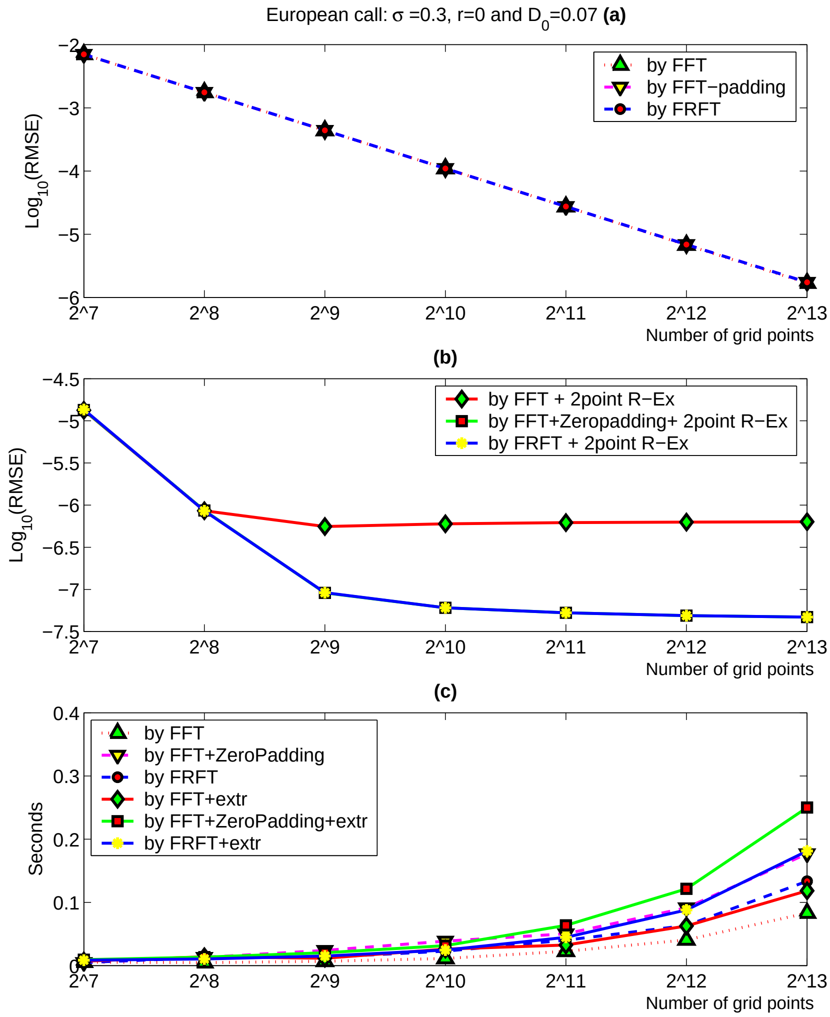


Figure 5.3: Errors and the computer time in pricing European call under GBM with $\sigma = 0.3$, $r = 0.0$ and $D_0 = 0.07$

5.1.3 Results of American Call

We apply the Lord method and the Lord-FRFT method to value American options.

If $P(n)$ denotes an n-time exercisable Bermudan option price, the price of an American option is approximated by inserting $P(1)$, $P(2)$, $P(4)$, and $P(8)$ into repeated Richardson extrapolation, as shown in Figure 4.7.

When using the Lord/Lord-FRFT method to price Bermudan options, the one-step Lord/Lord-FRFT method is used recursively backwards in time: First, the payoff at expiry is set in the Lord/Lord-FRFT method to get the continuation value at the last early exercise date, then the result is compared to the payoff at that time point and the minimum is placed in the Lord/Lord-FRFT method to generate the option prices on the last-but-one early exercise date. Again the solution is compared to the payoff at that time point and the minimum is used in the Lord/Lord-FRFT method, and so on $\cdots$; the procedure continues till the option prices at the initial time point are obtained (as shown in Figure 5.4).

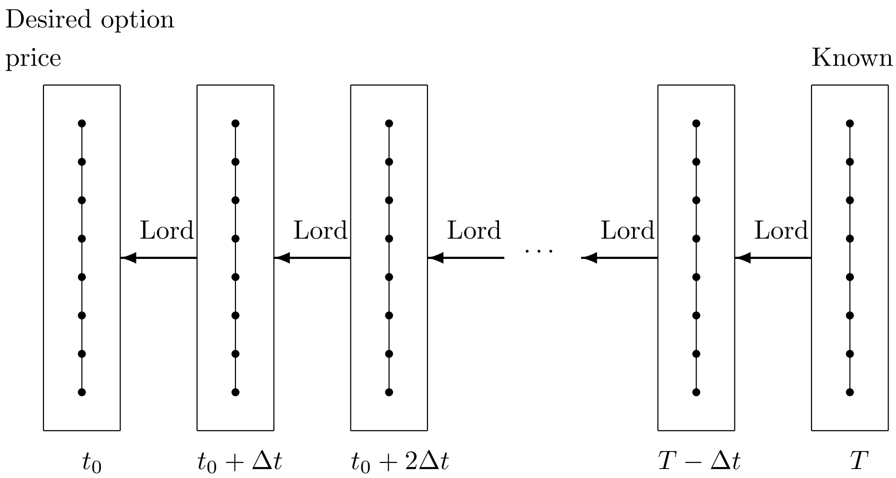


Figure 5.4: Multi-steps in pricing Bermudan options using the Lord/Lord-FRFT method

Numerical results are reported in Tables 5.5, 5.6 and 5.7. 'True' represents the values calculated with a binomial tree with 10,000 steps. 'Lord' represents the values by inserting $P(1)$, $P(2)$, $P(4)$ and $P(8)$ values calculated by the Lord method in the 4-times repeated Richardson extrapolation scheme. 'FRFT' represents the values with $P(1)$, $P(2)$, $P(4)$ and $P(8)$ calculated by the Lord-FRFT method inserted in the 4-time repeated Richardson extrapolation scheme. 'Q-FFT R4' represents the values calculated by the 4-point Richardson extrapolation scheme and O'Sullivan's QUAD-FFT method.

Table 5.5: Prices of American call options under GBM with $\sigma = 0.2$, $r = 0.03$, and $D_0 = 0.07$

S	'True'	Lord	FRFT	Q-FFT R4
80	0.2194	0.2191	0.2191	0.2194
90	1.3864	1.3854	1.3854	1.3849
100	4.7825	4.7841	4.7841	4.7829
110	11.0978	11.0897	11.0897	11.0832
120	20.0004	20.0055	20.0055	20.0088
RMSE		0.0044	0.0044	0.0077

($E =$ €100, $\Delta t = 0.5$ years, $L = 10$ and $N = 512$)

From the results we can see that both, the Lord and Lord-FRFT methods, give more accurate results than 'Q-FFT R4'. With $N = 512$, the errors are of degree 10^{-3}.

We have only used the Lord method and the Lord-FRFT method in each valuation step. We used neither the Lord-extrapolation(the Lord-extrapolation(the Lord method +

Table 5.6: Prices of American call options under GBM with $\sigma = 0.4$, $r = 0.03$, and $D_0 = 0.07$

S	'True'	Lord	FRFT	Q-FFT R4
80	2.6889	2.6875	2.6875	2.6896
90	5.7223	5.7206	5.7206	5.7181
100	10.2385	10.2397	10.2397	10.2431
110	16.1812	16.1840	16.1840	16.1801
120	23.3598	23.3538	23.3538	23.3247
RMSE		0.0032	0.0032	0.0357

($E =$ €100, $\Delta t = 0.5$ years, $L = 10$ and $N = 512$)

Table 5.7: Prices of American call options under GBM with $\sigma = 0.3$, $r = 0.0$, and $D_0 = 0.07$

S	'True'	Lord	FRFT	Q-FFT R4
80	1.0373	1.0362	1.0362	1.0337
90	3.1233	3.1222	3.1222	3.1278
100	7.0354	7.0374	7.0374	7.0375
110	12.9552	12.9448	12.9448	12.9287
120	20.7173	20.7255	20.7255	20.7424
RMSE		0.0108	0.0108	0.0166

($E =$ €100, $\Delta t = 0.5$ years, $L = 10$ and $N = 512$)

2 point Richardson extrapolation technique), nor the Lord-FRFT-extrapolation method, although the latter two brought remarkable high accuracy with a few number of grid points in the European call pricing case. The reason is that, in order to be able to employ the 2-point Richardson extrapolation and eliminate the $O(h^2)$ error term from the calculated values on one lattice, we also need values from a finer or coarser grid. This is difficult with early exercise cases, in which the continuation value is sought.

We also evaluated the error convergence and the computational time of both methods. The results are summarized in Figures 5.5, 5.6 and 5.7. The tables and the figures show that:

- Accurate results are obtained by 'Lord' and 'FRFT'. They are more accurate than 'Q-FFT R4'.
- The results stop to converge when the number of grid points on each lattice exceed some value, e.g. $N = 1024$. The reason is that, when the grid is fine enough, the error is dominated by the repeated Richardson extrapolation scheme error.
- The methods are very fast. With 1024 grid points, both the Lord and the Lord-FRFT methods finish the computations within 0.3 seconds.
- Another advantage of the Lord or the Lord-FRFT methods is that, once the exercise price E is given, all the values of the options with different initial underlying stock prices can be obtained with one-time execution only.
- As for the computing time, the Lord method is faster than the Lord-FRFT method because of fewer memory allocation. For example, when $N = 2^{13}$, the Lord-FRFT method is running out of cache thus needs a longer time than the Lord method.

In total, the Lord method is preferable over the Lord-FRFT method when pricing American options, as the former gives similar accurate results but uses less CPU time than the latter.

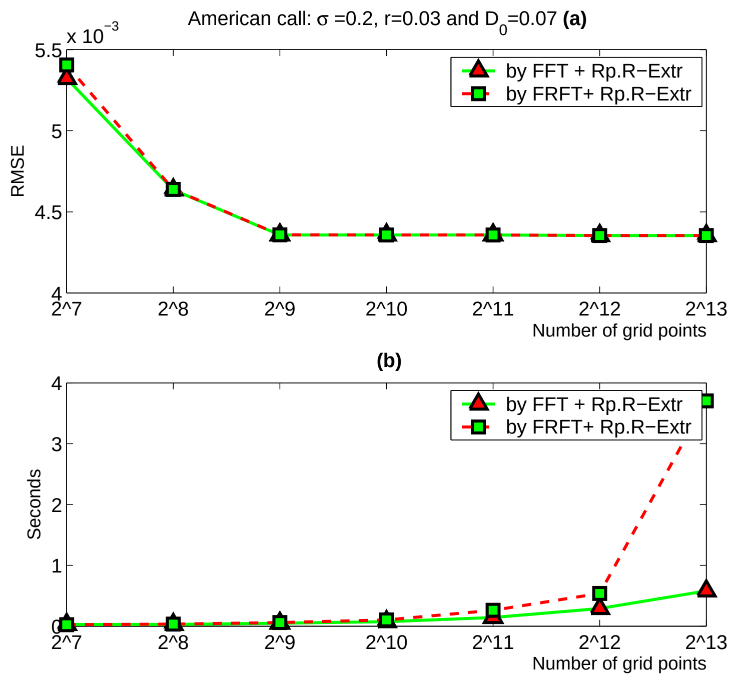


Figure 5.5: Errors and the computer time in pricing European call under GBM with $\sigma = 0.2$, $r = 0.03$ and $D_0 = 0.07$

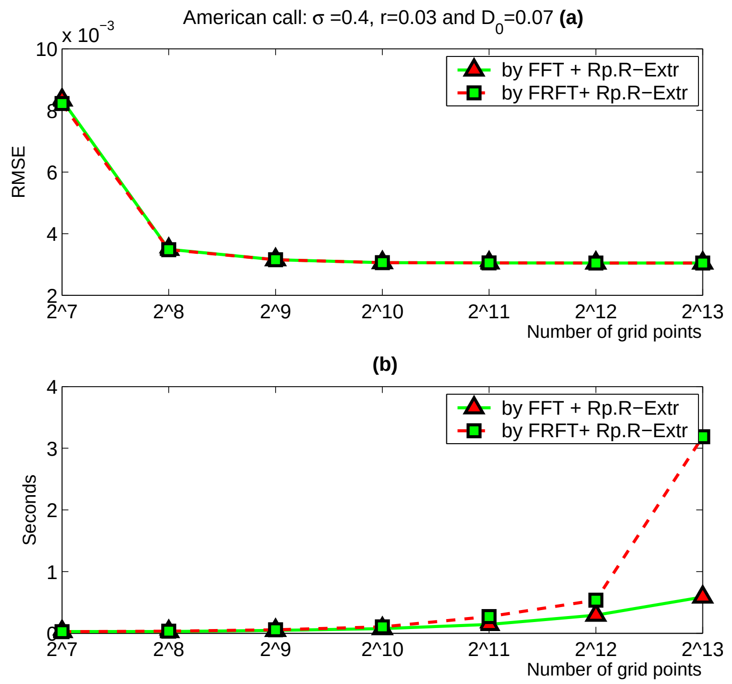


Figure 5.6: Errors and the computer time in pricing European call under GBM with $\sigma = 0.4$, $r = 0.03$ and $D_0 = 0.07$

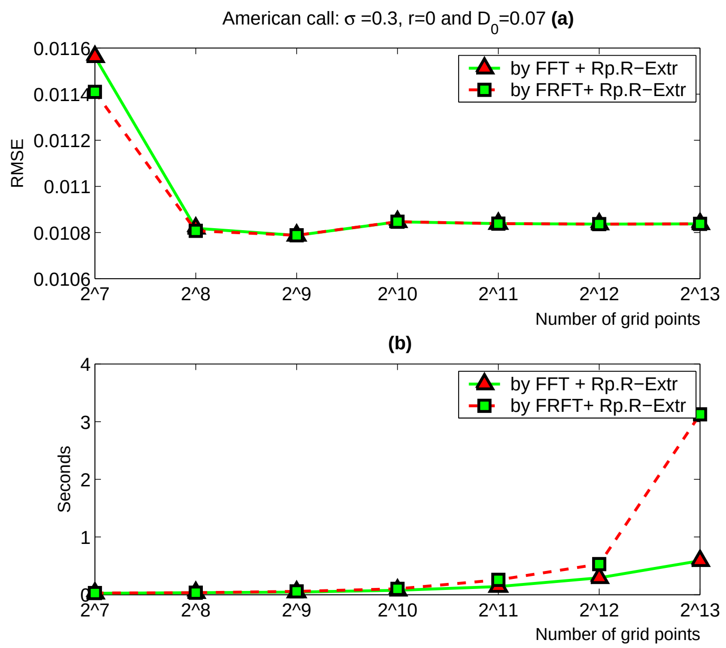


Figure 5.7: Errors and the computer time in pricing European call under GBM with $\sigma = 0.3$, $r = 0.0$ and $D_0 = 0.07$

5.2 The Variance Gamma Process

We repeat the experiments under a different random walk of the underlying stock prices: the Variance Gamma (VG) process.

Though GBM remains to be the most widely used model by practitioners, it has known biases. Therefore researchers have come up with alternative underlying process models including finite or infinite jumps. The VG model falls into the Lévy processes with an infinite jump arrival rate. The VG process for asset returns was first proposed by Madan and Seneta(1990), and extended by Madan and Milne (1991), Madan, Carr and Chang (1998) and Carr, Geman, Madan and Yor (2003).

As long as the characteristic function of the underlying stock price is known, and the conditional density equals to a transition density (Equation (3.38)), then we can use the Lord/Lord-FRFT method to price the options.

So, we need to get the characteristic function of the VG process.

5.2.1 Characteristic Function

The VG process is obtained by evaluating Brownian motion with drift at a random time given by a gamma process.

It has three parameters: σ the volatility of the Brownian motion, ν the variance rate of the gamma time change and θ the drift in the Brownian motion.

Let

$$b(t; \theta, \sigma) = \theta t + \sigma W(t), \tag{5.5}$$

where $W(t)$ is a standard Brownian motion. The process $b(t; \theta, \sigma)$ is a Brownian motion with drift θ and volatility σ.

Then a VG process $X(t; \sigma, \nu, \theta)$ is defined in terms of the Brownian motion with drift

$b(t;\theta,\sigma)$ and the gamma process with unit mean rate, $\gamma(t;1,\nu)$ as

$$X(t;\sigma,\nu,\theta) = b(\gamma(t;1,\nu);\theta,\sigma). \tag{5.6}$$

The density function for the VG process at time t can be expressed conditional on the realization of the gamma time change g as a normal density function[3]. The unconditional density may then be obtained on integrating out g employing the density of the gamma process for the time change g. This gives us the density for $X(t)$ as

$$f_{X(t)}(X) = \int_0^\infty \frac{1}{\sigma\sqrt{2\pi g}} \exp\left(-\frac{(X-\theta g)^2}{2\sigma^2 g} - \frac{g}{\nu}\right) \frac{g^{t/\nu-1}}{\nu^{t/\nu}\Gamma(t/\nu)} dg, \tag{5.7}$$

where $\Gamma(x)$ is the gamma function.

The characteristic function for the VG process reads

$$\phi_{X(t)}(v) = \mathbb{E}\left[e^{ivX(t)}\right] = \left(1 - i\theta\nu v + \frac{1}{2}\sigma^2\nu v^2\right)^{-t/\nu}. \tag{5.8}$$

With $x_t = \ln(S_t/E)$, the characteristic function of the log stock price conditioned on a given x_t has the following closed form:

$$\phi(v|x_t) = \exp\left[ivx_t + iv(r - D_0 + w)\Delta t - \Psi(v,\Delta t)\right], \tag{5.9}$$

where

$$\Psi(v,\Delta t) = \frac{1}{\nu}\ln\left[1 - i\theta\nu v + \frac{1}{2}\sigma^2\nu v^2\right], \tag{5.10}$$

and

$$w = \frac{1}{\nu}\left[1 - \theta\nu - \frac{1}{2}\sigma^2\nu\right]. \tag{5.11}$$

5.2.2 Results of European Call

We price a series of European call options, with exercise prices E as 90€, 95€, $\cdots$, 115€, and 120€, respectively. Their life time, Δt, is set to 1 year, the interest rate, r, is set to 0.10, the divident rate, D_0, is set to 0, and the number of grid points in log stock dimension, N, is set to 1024. The initial stock price is set to 100€.

The stock price is assumed to follow a VG process in (5.6) with

$$\sigma = 0.12; \quad \theta = -0.14; \quad \nu = 0.2. \tag{5.12}$$

The results are listed in Table 5.8.

'VG exact' represents analytical solution, 'Lord' represents results from the Lord method, 'FRFT' represents results from the Lord-FRFT method, 'Lord +' represents results from Lord-extrapolation method, 'FRFT +' represents results from Lord-FRFT-extrapolation method. Results from two existing methods are also involved for benchmark and comparison purposes: 'Q-FFT' represents results from O'Sullivan (2004), and 'KW' represents results from Kellezi and Webber (2003).

The RMSE of the results are calculated w.r.t. the analytical prices for the European call prices.

From the results we could see that, all four methods, the Lord, the Lord-FRFT, the Lord-extrapolation, and the Lord-FRFT-extrapolation method, give very accurate results. Among them, the Lord-FRFT-extrapolation method gives remarkably accurate results with errors of degree 10^{-6} for $N = 1024$, which is as accurate as the KW method.

We further examine the error convergence and the computer time.

Results are plotted in Figure 5.8. From the figure we can see that

[3] Madan, Carr, and Chang(1998)

Table 5.8: Prices of European call options under Variance Gamma process

E	VG exact	Lord	FRFT	Lord+	FRFT+	KW	Q-FFT
90	19.09935	19.09933	19.09933	19.09936	19.09934	19.09936	19.09937
95	15.07047	15.07044	15.07043	15.07048	15.07048	15.07048	15.07049
100	11.37002	11.36998	11.36996	11.37004	11.37003	11.37002	11.37003
105	8.11978	8.11972	8.11969	8.11981	8.11977	8.11978	8.11986
110	5.42960	5.42956	5.42948	5.42967	5.42960	5.42960	5.42963
115	3.36543	3.36547	3.36530	3.36559	3.36542	3.36544	3.36523
120	1.92110	1.92130	1.92097	1.92140	1.92110	1.92110	1.92078
RMSE		8.3e-05	9.4e-05	1.4e-04	9.3e-06	6.5e-06	1.5 e-04

(S_0 = 100€, $\Delta t = 1$ year, σ =0.12, r =0.10 D_0 =0, $L = 10$ and $N = 1024$)

- The Lord method generates accurate results (e.g. 10^{-4}) without explicit zero-padding.
- The Lord-FRFT method can be used to generate highly accurate (e.g. 10^{-7}) results.
- The combination of the Lord-FRFT method with the 2-point Richardson extrapolation technique gives very accurate results. For example, with $N = 128$, the errors are of order 10^{-5}.

  However, the combination of the Lord method with the 2-point Richardson extrapolation technique does not converge when the number of grid points exceeds some value (e.g. $N = 1024$). This is because Richardson extrapolation cannot remove the errors arising in the frequency domain (Recall that the mesh size in frequency domain is fixed in the Lord method) because of the use of the FFT.
- All the methods are fast! They compute the whole set of option prices, with 1024 grid points, within 0.3 seconds. Ordering the methods from the fastest to the slowest, we have the Lord, the Lord-FRFT, the Lord-extrapolation, and the Lord-FRFT-extrapolation method.
- All the methods are faster than the other existing methods we compared with, such as the QUAD-FFT method from O'Sullivan. This is due to the $O(Nlog(N))$ complexity of the Lord method.

5.2.3 Results of Bermudan Put

We value a series of 10-times exercisable Bermudan put options, whose exercise prices range from 90 to 120€. The life time, T, is 1 year, the interest rate, r, is 0.10, the dividend rate, D_0, is 0, and the number of grid points in log stock dimension, N, is set to 1024. The initial stock price S_0 is 100€. The RMSE is taken w.r.t. the results given by Kellezi and Webber (2003) as the benchmarck.

The results are given in Table 5.9.

We further examine the error convergence and the computational time of the Lord and the Lord-FRFT methods. The testing results are plotted in Figure 5.9.

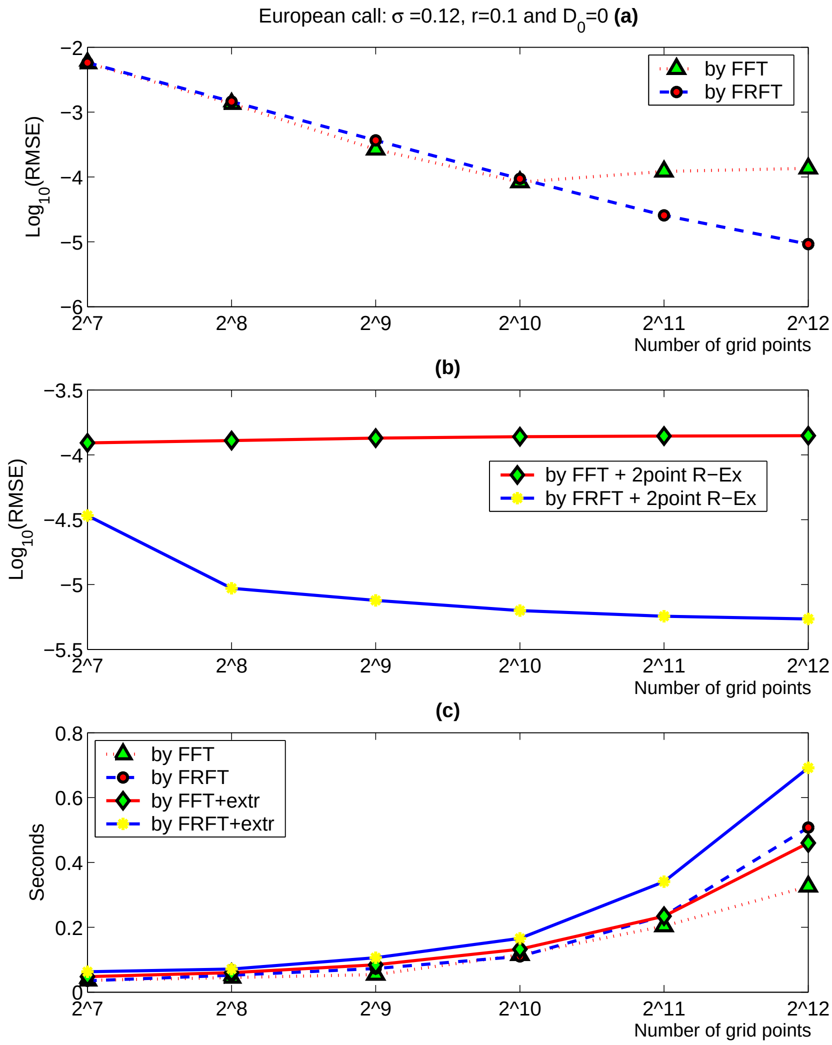


Figure 5.8: Errors and the computer time in Pricing European call under VG process

Table 5.9: Prices of Bermudan put options under Variance Gamma process

E	KW	Lord	FRFT	Q-FFT
90	0.76115	0.76115	0.76110	0.75787
95	1.52574	1.52574	1.52578	1.53866
100	2.88152	2.88153	2.88154	2.87782
105	5.17036	5.17040	5.17042	5.19537
110	9.04064	9.04067	9.04074	9.06040
115	13.87623	13.87624	13.87626	13.89456
120	18.80965	18.80965	18.80964	18.82561
RMSE		2.18e-05	5.20e-05	1.60e-02

(S_0 = 100€, $\Delta t = 1$ year, σ =0.2, r =0.03 D_0 =0.07, $L = 10$ and $N = 1024$)

From the table and the figure we can see that

- The Lord method and Lord-FRFT method give very accurate results (10^{-5}), and are

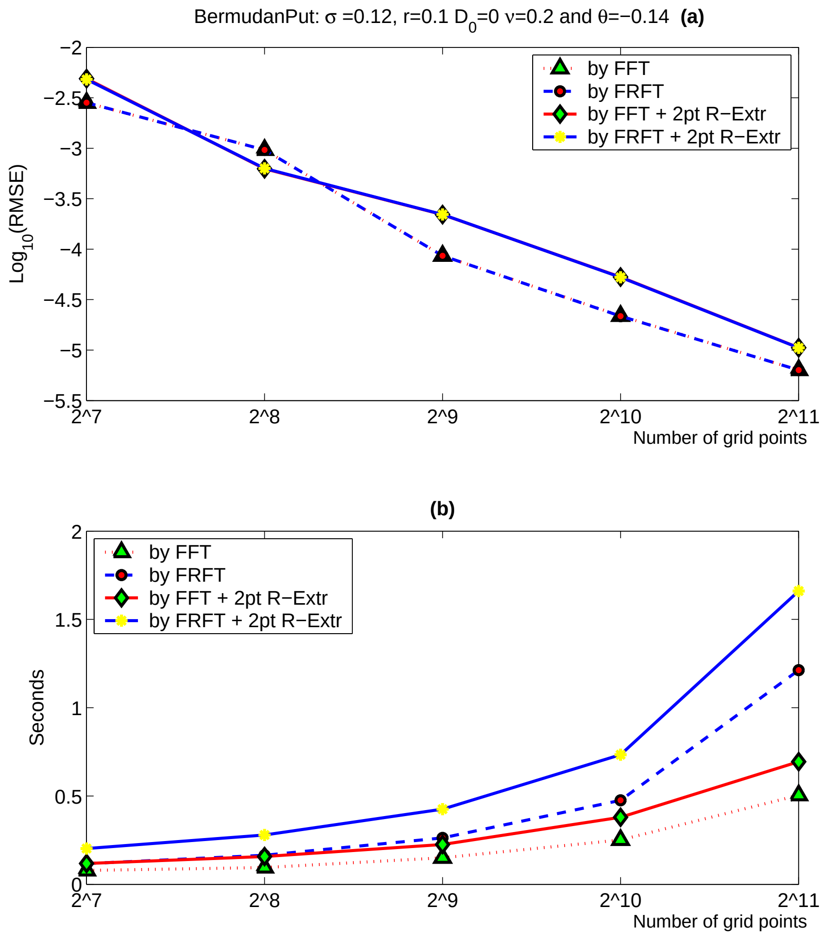


Figure 5.9: Errors and the computer time in Pricing Bermudan put under VG process

more accurate than the QUAD-FFT method (10^{-2}) in pricing Bermudan options.

- Their fast computational speed is again shown. With $N = 1024$, the Lord method completes the computations within 0.25 seconds, and the Lord-FRFT method in 0.5 seconds.
- If combining the 2 point Richardson extrapolation with the Lord and the Lord-FRFT method, we do not obtain higher accurate results.
- Considering the over all performance, when pricing Bermudan options, the Lord method beats the Lord-FRFT method with the same accuracy but a faster computational speed.

5.3 Value of The Damping Factor

Up to now, we did not clarify what value should be assigned to the damping factor β in the Lord algorithm (Equation 4.1):

$$C(t,x) = e^{-r(T-t)} \cdot e^{-\beta x}\mathcal{F}^{-1}\{\ \mathcal{F}\{e^{\beta y}V(T,y)\} \cdot \phi(-\omega + i\beta)\ \}. \tag{5.13}$$

In the implementations in this thesis, we have set $\beta = 0$. Experiments (Figure 5.10 with the Lord-FRFT method) show that, with $\Delta t = 0.5$ year, β should be smaller than 2 for call options or should be larger than -2 for put options. The number of floating point operations is least when $\beta = 0$. Hence, we took $\beta = 0$ aiming at both a fast resolution and a high accuracy.

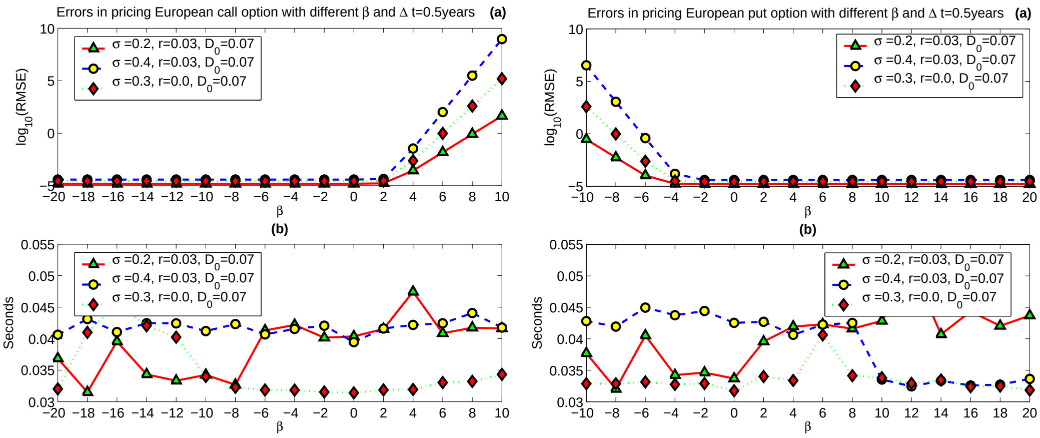


Figure 5.10: Errors in pricing European call (left) and put (right) options with different damping factor β for $\Delta t = 0.5$ years

There are many suggestions in the literature on how to choose a proper damping factor, such as, Carr and Maddan (1999), Raible (2000) and Lee (2004). However, one should be aware that their research is based on Carr-Madan-inversion based transform method, which damps the option value with $e^{\alpha k}$ with $k = \log(E)$, e.g. for a European call:

$$\begin{aligned}
\hat{c}(v) &= \mathcal{F}\{e^{\alpha k}C(t)\} = \int_{-\infty}^{\infty} e^{ivk}e^{\alpha k}C(t)dk \\
&= \int_{-\infty}^{\infty} e^{(\alpha+iv)k}\left[e^{-r(T-t)}\int_{-\infty}^{\infty}\max\left(e^s - e^k, 0\right)f(s)ds\right]dk \\
&= e^{-r(T-t)}\int_{-\infty}^{\infty} f(s)\int_{-\infty}^{s}\left(e^{s+\alpha k} - e^{(1+\alpha)k}\right)e^{ivk}dkds \\
&= e^{-r(T-t)}\int_{-\infty}^{\infty} f(s)\left[\frac{e^{(\alpha+1+iv)s}}{\alpha+iv} - \frac{e^{(\alpha+1+iv)s}}{\alpha+1+iv}\right]ds \\
&= \frac{e^{-r(T-t)}}{(\alpha+iv)(1+\alpha+iv)}\phi\Big(v-(\alpha+1)i\Big), \qquad (5.14)
\end{aligned}$$

where $C(t)$ denotes the option price at time t and $s = \log(S_t)$; In contrast, the Lord method damps the payoff function with $e^{\beta x}$ with $x = \log(S_t/E)$ and $y = \log(S_T/E)$:

$$
\begin{aligned}
& e^{r(T-t)}\mathcal{F}\{e^{\beta x}C(t,x)\} \\
&= e^{r(T-t)}\int_{-\infty}^{\infty} e^{i\omega x}e^{\beta x}C(t,x)dx \\
&= \int_{-\infty}^{\infty} e^{i\omega x}\left[\int_{-\infty}^{\infty} e^{\beta x}C(T,x+z)f(z)dz\right]dx, \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\omega x+\beta x}C(T,x+z)f(z)dzdx, \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} e^{i(\omega-i\beta)x}C(T,x+z)dx\right]f(z)dz, \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} e^{i(\omega-i\beta)(y-z)}C(T,y)dy\right]f(z)dz, \\
&= \int_{-\infty}^{\infty} e^{i(\omega-i\beta)y}C(T,y)dy \int_{-\infty}^{\infty} e^{i(-\omega+i\beta)z}f(z)dz \qquad (5.15) \\
&= \hat{C}_T(\omega-i\beta)\ \phi(-\omega+i\beta). \qquad (5.16)
\end{aligned}
$$

Due to this discrepancy, suggestions on damping values from the literature are usually not applicable to the Lord method.

If we substitute $C(T, y)$ in (5.15) by the payoff function of a European call option, $(Ee^y - E)^+$, we get

$$
\begin{aligned}
& e^{r(T-t)}\mathcal{F}\{e^{\beta x}C(t,x)\} \\
&= \underbrace{\int_{0}^{\infty} e^{i\omega y}e^{\beta y}\left[E(e^y-1)\right]dy}_{\mathcal{F}\{e^{\beta y}[E(e^y-1)]\}} \underbrace{\int_{-\infty}^{\infty} e^{i(-\omega+i\beta)z}f(z)dz}_{\phi(-\omega+i\beta)}. \qquad (5.17)
\end{aligned}
$$

To ensure the existence of the Fourier transform of $e^{\beta y}\left[E(e^y - 1)\right]$, it has to be in L^1, namely,

$$\int_0^{\infty} e^{\beta y}\left[E(e^y-1)\right]dy < \infty, \qquad (5.18)$$

or equivalently:

$$\int_0^{\infty} e^{(\beta+1)y}\left[E(1-e^{-y})\right]dy < \infty, \qquad (5.19)$$

from which we can see that when $\beta \le -1$ the above relationship holds.

Analogously, β should be larger or equal to 1 for put options.

On the other hand, when the life time of an option is extremely short, like 0.02 years or one week, the choice of β becomes more important (see Figure 5.11). When $\Delta t = 0.02$ year, β should be smaller than -1 for call options but cannot be too small as the errors would grow; β should be larger than 1 for put options but cannot be too big, either, due to the same consideration on accuracy.

We provide as a rule of thumb:

- $\beta \in [-4, -1]$ for call options; and
- $\beta \in [1, 4]$ for put options.

5.4 Conclusions

Based on the experiments we could draw the following conclusions:

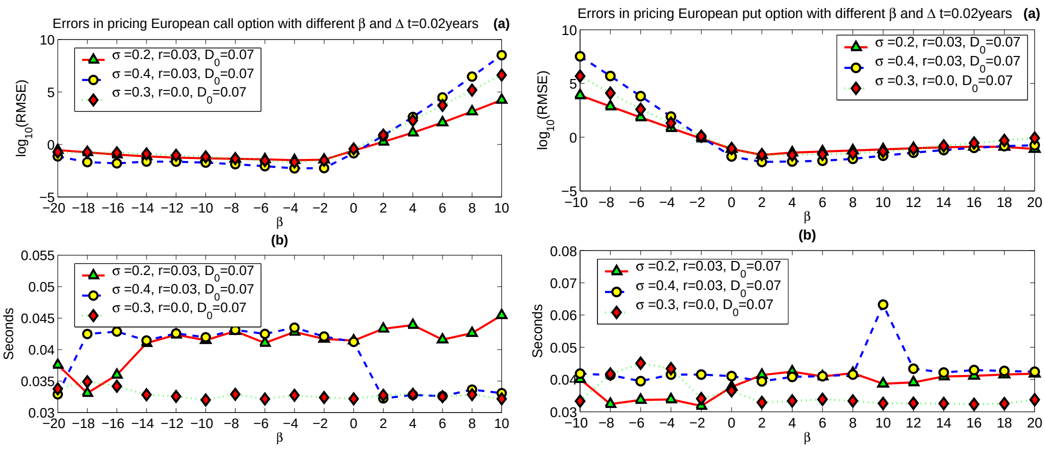


Figure 5.11: Errors in pricing European call (left) and put (right) options with different damping factor β for $\Delta t = 0.02$ years

- The Lord method gives very accurate and fast resolution in pricing European, Bermudan and American options.
- The Lord-FRFT method can be used to generate high accurate results, because by increasing the number of grid points higher accuracy is obtained in both the log-stock domain and the frequency domain. At the same time, aliasing artifacts of approximating linear convolution via circular are prevented.
- The best choice for pricing European options is the combination of the Lord-FRFT method with the 2-point Richardson extrapolation technique. There is no need to take large number of grid points, as only 128 points can give very accurate results, like 10^{-4} error.
- The best choice for pricing Bermudan options are the Lord method and the Lord-FRFT method. Both give similar accurate results, but the Lord method is a bit faster. It is not trivial to determine the truncation range in the frequency domain in the Lord-FRFT method, because the best truncation range differs from case to case. Usually 512 points on each lattice give sufficiently high accuracy.
- The Lord method and the Lord-FRFT method are faster than existing methods for pricing early exercisable options, and they also generate the most accurate results with the same number of grid points.

Chapter 6

Two-Assets Option Pricing

An interesting feature of the Lord/Lord-FRFT methods is that they can be easily generalized to high dimensions for solving multi-assets option pricing problems.

In this chapter, high dimensional Fourier transforms and fast algorithms are introduced; After that the Lord method and the Lord-FRFT method are generalized to two dimensions; Numerical experiments in pricing a two-asset minimum call option using the methods follow.

6.1 High-dimensional Fourier Transforms

To be consistent with 1-D case, we use the following definitions for high-dimensional Fourier transforms:

- Continuous-time Fourier transform in m-dimensions:

$$\mathcal{F}^m\{F(\vec{x})\} = \int_{\mathbb{R}^m} e^{i\vec{\omega}\cdot\vec{x}} F(\vec{x})\, d\vec{x} \triangleq \hat{F}(\vec{\omega}), \tag{6.1}$$

  where "$\cdot$" denotes inner product of two vectors,

$$\vec{a}\cdot\vec{b} = \sum_{j=1}^{N} a_j b_j, \quad \vec{a} = \{a_j\}_{j=1}^N,\ \vec{b} = \{b_j\}_{j=1}^N; \tag{6.2}$$

- Inverse continuous-time Fourier transform in m-dimensions:

$$\mathcal{F}^{-m}\{\hat{F}(\vec{\omega})\} = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} e^{-i\vec{x}\cdot\vec{\omega}} F(\vec{x})\, d\vec{x}; \tag{6.3}$$

- Discrete Fourier transform(DFT):

$$DFT^m_{k_1\ldots k_m}\{F_{n_1\ldots n_m}\} = \sum_{n_1=0}^{N_1-1} e^{i\frac{2\pi}{N_1}n_1k_1} \cdots \sum_{n_m=0}^{N_m-1} e^{i\frac{2\pi}{N_m}n_mk_m} F_{n_1\ldots n_m}, \tag{6.4}$$

  which can be more compactly expressed in vector notation by

$$DFT^m_{k_1\ldots k_m}\{F_{n_1\ldots n_m}\} = \sum_{\vec{n}=0}^{\vec{N}-1} e^{i2\pi\frac{\vec{n}}{\vec{N}}\cdot\vec{k}} F_{n_1\ldots n_m}, \tag{6.5}$$

  where $\vec{n} = (n_1, n_2, \cdots, n_m)$ and $\vec{k} = (k_1, k_2, \cdots, k_m)$ are m-dimensional vectors, $\vec{N} - 1 = (N_1 - 1, N_2 - 1, \cdots, N_m - 1)$, the division $\frac{\vec{n}}{\vec{N}}$ is performed element-wise, and the sum denotes a set of nested summations;

- Inverse DFT:

$$DFT^{-m}_{k_1 \dots k_m}\{F_{n_1 \dots n_m}\} = \frac{1}{\prod_{d=1}^{m} N_d} \sum_{\vec{n}=0}^{\vec{N}-1} e^{-i2\pi \frac{\vec{n}}{\vec{N}} \cdot \vec{k}} F_{n_1 \dots n_m}. \tag{6.6}$$

The main step in the Lord/Lord-FRFT method, from the scientific computing point-of-view, is to approximate the continuous-time Fourier transform by the DFT using the existing FFT algorithm. Two issues turn out to be relevant: how to discrete the continuous-time Fourier transform formula; and how to generalize the 1-D FFT algorithm to solve high dimensional DFT.

Consider a 2-D DFT,

$$DFT^2_{k_1 k_2}\{F_{n_1 n_2}\} = \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} e^{i2\pi \frac{n_1}{N_1} k_1 + i2\pi \frac{n_2}{N_2} k_2} F_{n_1 n_2}, \tag{6.7}$$

we then have

$$DFT^2_{k_1 k_2}\{F_{n_1 n_2}\} = \sum_{n_1=0}^{N_1} e^{i2\pi \frac{n_1}{N_1} k_1} \left(\sum_{n_2=0}^{N_2} e^{i2\pi \frac{n_2}{N_2} k_2} F_{n_1 n_2} \right) \tag{6.8}$$

$$= DFT^1_{k_1}\{DFT^1_{k_2}\{F_{n_1 n_2}\}\}, \tag{6.9}$$

Hence, the 2-D DFT is equivalent to two nested 1-D DFT's, one of which is executed along the rows and the other along the columns. The algorithm is shown in Figure 6.1:

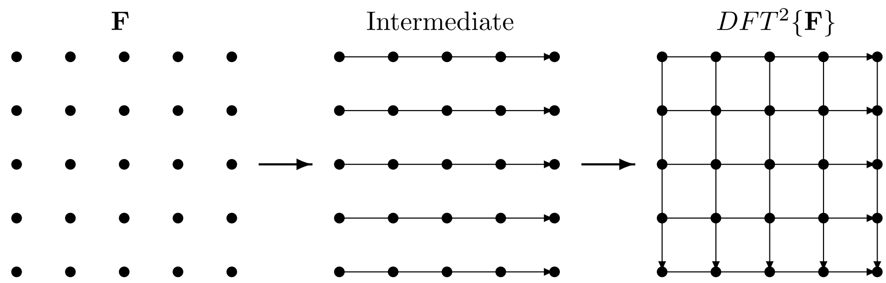


Figure 6.1: Algorithm of 2-D DFT

The complexity of a M-dimension DFT is thus $O(MNlogN)$, if N denotes the number of grid points in each dimension.

6.2 The Lord Method Extended to High Dimensions

In order to use the Lord method, we need an important assumption that, the multi-variate conditional probability density function, f, of the vector of state variables $\{y_1, y_2, \cdots, y_n\}$ conditioned on $\{x_1, x_2, \cdots, x_n\}$ is equal to the transition probability density $\{y_1 - x_1, y_2 - x_2, \cdots, y_n - x_n\}$:

$$f(y_1, y_2, \cdots, y_n | x_1, x_2, \cdots, x_n) = f(y_1 - x_1, y_2 - x_2, \cdots, y_n - x_n). \tag{6.10}$$

Bjørk(1998) summaries that the solution to multi-dimensional Black-Scholes PDEs has a risk neutral valuation form:

Theorem 1 (Bjørk) *Let $\vec{s} = \{s_1, s_2, \cdots, s_m\}^*$ denote the vector of stock prices at time t and $F(t, \vec{s})$ denote the pricing function of a multi-asset option, then the option pricing Black-Scholes PDE has the following equivalent representation:*

$$F(t, \vec{s}) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}_{t,\vec{s}} \left[V(\vec{S}(T)) \right]. \tag{6.11}$$

Here the expectation $\mathbb{E}^{\mathbb{Q}}_{t,\vec{s}}$ is to be taken with respect to the martingale (risk-neutral) measure $\mathbb{Q}$, and $V(\vec{S}(T))$ indicates the value or payoff/value of the option at time $T(T > t)$ with $\vec{S}(T) = \{S_1, S_2, \cdots, S_m\}^*$ indicating the vector of stock prices at time T; r denotes the risk-neutral interest rate.

With the definition of expectation, we get

$$F(t, \vec{x}) = e^{-r(T-t)} \int_{\mathbb{R}^m} V(\vec{y}) \cdot f(\vec{y}|\vec{x}) \quad d\vec{y}, \tag{6.12}$$

where $\mathbb{R}^m$ denotes the Euclidean space, $\vec{y}$ and $\vec{x}$ denote vectors of state variables at time T and $t(t < T)$, respectively.
Making use of the assumption in (6.10), and changing variables with $\vec{z} = \vec{y} - \vec{x}$, we have

$$e^{r(T-t)} F(t, \vec{x}) = \int_{\mathbb{R}^m} V(\vec{x} + \vec{z}) \cdot f(\vec{z}) \quad d\vec{z}. \tag{6.13}$$

Damping $F(t, \vec{x})$ with $e^{\vec{\beta}\cdot\vec{x}}$ then taking Fourier transform on both sides of (6.13), we have

$$\begin{aligned}
e^{r(T-t)} \mathcal{F}^m\{e^{\vec{\beta}\cdot\vec{x}} F(t, \vec{x})\} &= \int_{\mathbb{R}^m} e^{i\vec{\omega}\cdot\vec{x}} e^{\vec{\beta}\cdot\vec{x}} F(t, \vec{x}) d\vec{x} && (6.14) \\
&= \int_{\mathbb{R}^m} e^{i\vec{\omega}\cdot\vec{x}} \left[\int_{\mathbb{R}^m} e^{\vec{\beta}\cdot\vec{x}} V(\vec{x} + \vec{z}) f(\vec{z}) d\vec{z} \right] d\vec{x}, \\
&= \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} e^{i\vec{\omega}\cdot\vec{x} + \vec{\beta}\cdot\vec{x}} V(\vec{x} + \vec{z}) f(\vec{z}) d\vec{z} d\vec{x}, \\
&= \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^m} e^{i(\vec{\omega} - i\vec{\beta})\cdot\vec{x}} V(\vec{x} + \vec{z}) d\vec{x} \right] f(\vec{z}) d\vec{z}, \\
&= \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^m} e^{i(\vec{\omega} - i\vec{\beta})\cdot(\vec{y} - \vec{z})} V(\vec{y}) d\vec{y} \right] f(\vec{z}) d\vec{z}, \\
&= \int_{\mathbb{R}^m} e^{i(\vec{\omega} - i\vec{\beta})\cdot\vec{y}} V(\vec{y}) d\vec{y} \int_{\mathbb{R}^m} e^{i(-\vec{\omega} + i\vec{\beta})\cdot\vec{z}} f(\vec{z}) d\vec{z} \\
&= \hat{V}(\vec{\omega} - i\vec{\beta})\ \phi(-\vec{\omega} + i\vec{\beta}), && (6.15)
\end{aligned}$$

where

$$\phi\left(\vec{t}\right) = \int_{\mathbb{R}^m} e^{i\vec{t}\cdot\vec{x}} f(\vec{x}) d\vec{x} \tag{6.16}$$

is the characteristic function (Fourier transform) of multi-variate probability density function $f(\vec{x})$.

The option price can be recovered by inverse Fourier transform and un-damping:

$$e^{r(T-t)} F(t, \vec{x}) = e^{-\vec{\beta}\cdot\vec{x}}\ \mathcal{F}^{-m}\{\hat{V}(\vec{\omega} - i\vec{\beta})\ \phi(-\vec{\omega} + i\vec{\beta})\}. \tag{6.17}$$

(6.17) is the basis of the Lord method in high dimensions. Before we discuss implementation, two key questions need to be answered:

- What is the multi-variate characteristic function?

- How to build a subprogram that calculates the high-dimensional continuous-time Fourier transforms?

Let's assume that the underlying asset prices, $\vec{S}$, follow Geometric Brownian Motion. Suppose we have n given assets("stocks") with price processes $S_1(t), \cdots, S_n(t)$. The asset price vector is denoted by $\vec{S}(t)$,

$$\vec{S}(t) = \begin{pmatrix} S_1(t) \\ \vdots \\ S_n(t) \end{pmatrix}. \tag{6.18}$$

Under the objective probability measure $\mathbb{P}$, the $\vec{S}$-dynamics are given by Geometric Brownian Motion:

$$dS_i(t) = \mu_i S_i(t)dt + S_i(t)\sigma_i d\bar{W}_i(t), \quad i = 1, \cdots, n, \tag{6.19}$$

where $\bar{W}_1, \cdots, \bar{W}_n$ are correlated $\mathbb{P}$-Wiener processes.

6.2.1 The Characteristic Function

Under the risk-neutral measure $\mathbb{Q}$, the drift of underlying assets prices $\{\mu_i\}$ should be the risk-less return $(r - D_i)$, with r denoting the risk-neutral interest rate from a bank and D_i the dividend rate from the underlying stock. Then the dynamic model under the $\mathbb{Q}$ measure for $\vec{S}$ becomes

$$\begin{aligned} dS_i &= (r - D_i)S_i dt + S_i\sigma_i dW_i, \quad i = 1, \cdots, n, \\ dW_i dW_j &= \rho_{ij} dt, \end{aligned} \tag{6.20}$$

where W_i is a $\mathbb{Q}$-Wiener process and ρ_{ij} is the correlation between S_i and S_j.

We can derive mathematical models for functions of asset prices with the help of the multidimensional Itô's formula.

Theorem 2 (The multidimensional Itô's formula) *Take a vector Wiener process $\vec{W} = (W_1, \cdots, W_n)$ with correlation matrix ρ given, and assume that the vector process $\vec{X} = (X_1, \cdots, X_n)$ has a structure*

$$dX_i = \mu_i dt + \sigma_i dW_i, i = 1, \cdots, n. \tag{6.21}$$

Then the differential of the process $f(t, \vec{X}(t))$ is given by

$$df = \left\{ \frac{\partial f}{\partial t} + \sum_{i=1}^{n} \mu_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \sigma_i \sigma_j \rho_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} \right\} dt + \sum_{i=1}^{n} \sigma_i \frac{\partial f}{\partial x_i} dW_i. \tag{6.22}$$

Apply (6.22) to $f = log(\vec{S})$, we find that the marginal density function of $log(S_i)$ under risk-neutral measure follows the normal distribution:

$$\log(S_i) \sim N(\mu_i, \bar{\sigma}_i), \tag{6.23}$$

with

$$\mu_i = \left(log(S_i^0) + r - D_0 - \frac{1}{2}\sigma_i^2 \right)(T - t), \tag{6.24}$$

and

$$\bar{\sigma}_i = \sigma_i \sqrt{T - t}. \tag{6.25}$$

Hence $\vec{S}$ follows the multivariate Normal distribution, whose characteristic function is given in the literature.

For example, the characteristic function of a bivariate Normal distribution is given by:

$$\phi(v_1, v_2) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(v_1x_1+v_2x_2)} f(x_1, x_2)dx_1dx_2 \tag{6.26}$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(v_1x_1+v_2x_2)} \exp\left[-\frac{z}{2(1-\rho^2)}\right] dx_1dx_2, \tag{6.27}$$

with

$$z \equiv \left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]. \tag{6.28}$$

By rewriting the exponential e^{ivx} as $(\cos(vx) + i\sin(vx))$ and evaluating the Gaussian integral

$$\int_{-\infty}^{\infty} e^{ivx}e^{-ax^2}dx = \int_{-\infty}^{\infty} \cos(vx)e^{-ax^2}dx = \sqrt{\frac{\pi}{a}}e^{-v^2/4a}, \tag{6.29}$$

we finally obtain the explicit form of the characteristic function as

$$\phi(v_1, v_2) = \exp\left[i(v_1\mu_1 + v_2\mu_2) - \frac{1}{2}(\sigma_1^2v_1^2 + 2\rho\sigma_1\sigma_2v_2v_2 + \sigma_2^2v_2^2)\right]. \tag{6.30}$$

6.2.2 The Subprogram

In the 1-D implementation, we have built up a subprogram in Chap.4.1 that takes in a sampled input function and returns the value of the 1-D continuous-time Fourier transform on the input function at discrete points.

Due to the fact that a M-dimensional Fourier transform is equivalent to M nested 1-dimensional Fourier transforms, we can make use of the 1-D subprogram for the implementation in high dimensions.

For example, a 2-D subprogram, with input a sampled $N \times N$ matrix that returns the value of a 2-dimensional Fourier transform on the input function at the $N \times N$ discrete points in the frequency domain, can be constructed as described in 6.2.

STEP1: Take in the matrix that contains the sampled value of an input function, say, $\mathbf{F}_{N\times N}$.

STEP2: Call the 1-D subprogram, for all row vector of $\mathbf{F}$, and store the returned results in an intermediate matrix.

STEP3: Transpose the intermediate matrix.

STEP4: Call 1-D subprogram again, pass in all row vectors of the intermediate matrix, and store the final results.

Figure 6.2: Algorithm of the subprogram in 2 dimensional case

For cases higher than 2D, the algorithm can be set up in a similar manner, only that, each time N^{M-1} row vectors are passed into the 1-D subprogram, and the 1-D subprogram is called M times. Hence, the complexity of the M dimensional Lord method will be $O(N^M \cdot \log(N))$ in total.

The Lord-FRFT method in high dimensions is constructed in exactly the same way as the Lord method, only that the subprogram is based on the 1-D FRFT subprogram.

6.3 Pricing a 2-asset Min-Call

In this section, the accuracy and computational time of the methods are evaluated for a two-asset minimum call option.

The payoff function of a 2-asset minimum call option is given by

$$\text{Payoff} = \max\{\min(S_1, S_2) - E, 0\}, \tag{6.31}$$

with S_1 and S_2 denoting the two underlying assets prices, and E the exercise price (Figure 6.3).

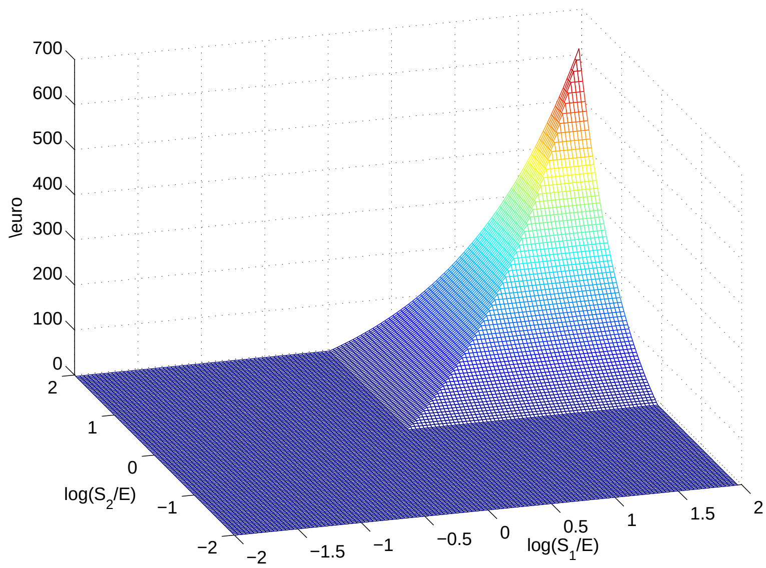


Figure 6.3: Payoff of a 2-asset min-call

The analytical solution for a European style 2-asset minimum call option is given by[1]:

$$\begin{aligned} M &= S_2 N_2(\gamma_2 + \sigma_2\sqrt{\Delta t}, \frac{(\log(S_1/S_2)) - \frac{1}{2}\sigma^2\sqrt{\Delta t}}{\sigma\sqrt{\Delta t}}, \frac{(\rho\sigma_1 - \sigma_2)}{\sigma}) \\ &\quad + S_1 N_2(\gamma_1 + \sigma_1\sqrt{\Delta t}, \frac{(\log(S_2/S_1)) - \frac{1}{2}\sigma^2\sqrt{\Delta t}}{\sigma\sqrt{\Delta t}}, \frac{(\rho\sigma_2 - \sigma_1)}{\sigma}) \\ &\quad - E e^{-R\Delta t} N_2(\gamma_2, \gamma_1, \rho), \end{aligned} \tag{6.32}$$

where $N_2(\alpha, \beta, \theta)$ is the bivariate cumulative standard normal distribution with upper limits of integration α and β, and coefficient of correlation θ, with

$$\gamma_1 = \left[\log(S_1/E) + (R - \frac{1}{2}\sigma_1^2)\Delta t\right] / \sigma_1\sqrt{\Delta t}, \tag{6.33}$$

$$\gamma_2 = \left[\log(S_2/E) + (R - \frac{1}{2}\sigma_2^2)\Delta t\right] / \sigma_2\sqrt{\Delta t}, \tag{6.34}$$

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2. \tag{6.35}$$

[1]Stulz, 1982

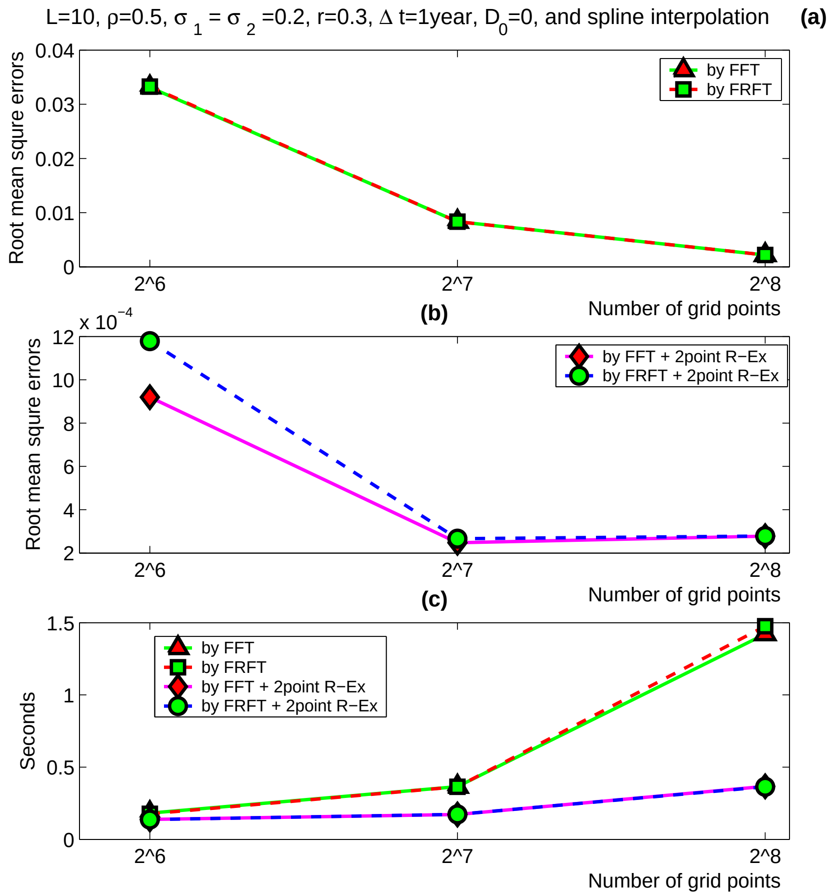


Figure 6.4: Numerical results in pricing a 2-asset min-call option

We work again in the log-stock domain and set up grids in two dimensions $y_1 = \log(S_1/E)$ and $y_2 = \log(S_2/E)$ the same way as in 1-D.

Both the Lord method and the Lord-FRFT method are employed in pricing a two-asset minimum call option, as well as their combinations with the 2-point Richardson extrapolation technique.

The RMSE is calculated as w.r.t. the analytical solution and taken as an indicator of the method accuracy.

Numerical results are summarized in Figure 6.4. They show that,

- Both the Lord and the Lord-FRFT method bring a smooth error convergence as the number of grid points increases, up to the point where error due to truncating the infinite integration (in the continuous-time Fourier transform formula) dominate.
- Both the Lord-extrapolation and the Lord-FRFT-extrapolation method give accurate results with only a few grid points.
- The methods are still very fast.
- The methods that incorporate Richardson extrapolation technique beat the other two without extrapolation in terms of CPU time and accuracy. From the Figure (6.4), we could see that, with 128 grid points the Lord-FRFT-extrapolation and the

Lord-extrapolation methods finish the computation within 0.5 seconds and return very accurate results of order 10^{-4}.

The advantages in both accuracy and computational speed of the Lord and the Lord-FRFT method remain in solving 2-D problems. The 2-point Richardson extrapolation works very well together with the Lord and the Lord-FRFT methods. That is due to the fact that the leading error is from the trapezoidal rule, whose order is known exactly ($O(h^2)$). Thus higher order of accuracy can be expected by using more accurate quadrature rules, like Simpson's rule.

However,the truncation error brought by the action of truncating the infinite integration domain is more significant in 2-D case than in 1-D case, given the same truncation range in the log-stock domain. That can be explained as, in the 2-D case with infinite integral is truncated twice because there is a double integration in 2-D continuous-time Fourier transform formula (6.1), whereas it is truncated only once in the 1-D case.

Hence, a larger truncation range L needs to be taken for high-dimensional cases.

Nevertheless, with $L = 10$ we obtain satisfactory accurate results (10^{-4}) in the 2-D case.

Chapter 7

Conclusions and Discussions

This thesis deals with a recently developed method in option pricing: the Lord method. It's performance is enhanced by the use of the fractional Fourier transform(FRFT) and Richardson extrapolation.

The basic idea of the Lord method is to depart from the risk-neutral valuation formula for European options, take Fourier transform, and rewrite the new formula in a form of linear convolution of the payoff and the density function of log-stock distribution. The Fourier transform of the density function, the characteristic function, is more often available than the density itself. The option value can be recovered by inverse Fourier transform on the product of the transformed payoff and the characteristic function.

Due to a large amount of zeros in both, the payoff function and the characteristic function, zero-padding needs not to be incorporated for satisfactory accuracy.

The idea of the Lord-FRFT method is to use the FRFT algorithm instead of the FFT. The usage of FRFT enables one to gain higher accuracy in the frequency domain by increasing the number of the grid points. This is not possible with the Lord method because the application of the FFT puts a restriction on the mesh sizes in log-stock and the frequency domain.

On the other hand, the FRFT involves explicit zero-padding, thus aliasing artifacts will not be present and a very high accuracy can be obtained by the Lord-FRFT method.

Due to the fact that the leading error term is from the trapezoidal rule and the order of the error is known as $O(h^2)$, the incorporation of 2-point Richardson extrapolation gives highly accurate results with only a small number of grid points. This is a conclusion w.r.t. European options, both under GBM and under the VG process.

All four methods (the Lord, the Lord-FRFT, the Lord-extrapolation, the Lord-FRFT-extrapolation) are employed in pricing European, Bermudan and American options.

Conclusions on the performance (accuracy, error convergence and the computational time) of the four methods in pricing European, Bermudan and American options are that: the Lord-FRFT method is the method of choice in pricing European options, as it can give highly accurate results, and its computation efforts can be greatly reduced by adding a 2-point Richardson extrapolation; the Lord method also gives accurate results in pricing European options, and it is superior in pricing Bermudan and American options due to its high computational speed in obtaining the required accuracy.

A discussion on the choice of the damping factor is also given. The damping factor β cannot be larger than -1 for call options, or cannot be smaller than 1 for put options. A rule of thumb is provided by, $\beta \in [-4, -1]$ for call options and $\beta \in [1, 4]$ for put options.

Finally the methods are generalized into two dimensions. A two-asset minimum call option is priced using the 2D variants of the methods.

Numerical results show that the advantage in both accuracy and computational speed of the Lord and the Lord-FRFT method in 1-D remains in 2-D case. The 2-point Richardson extrapolation works fairly well when incorporated with the two methods. But a large truncation range needs to be taken in 2-D case than in 1-D case.

Although the methods are very fast in one and two dimensional cases, they need tensor-product grids and therefore cannot break the ‘curse of dimensionality’.

The methods are not applicable to those cases where the conditional density does not equal to transition density.

7.1 Future Works

Error bounds of the Lord and Lord-FRFT method need to be analyzed.

In the next step of our research, high dimensional option pricing problems are to be solved using the methods proposed.

Also, in the multi-dimensional case, we need to study how correlation parameters and the damping factors are to be incorporated with the accuracy and error convergence.

Breaking of the ‘curse of dimensionality’ is also a topic of future research.

Bibliography

[1] Andricopoulos AD, Widdicks M.,Duck PW, Newton DP (2002), "Universal Option Valuation Using Quadrature Methods", Journal of Financial Economics 67(2003) 447-471.

[2] Andricopoulos AD, Widdicks M., Duck PW and Newton DP(2006), "Extending Quadrature Methods to Value Multi-asset and Complex Path Dependent Options", Journal of Financial Economics, 2006 (forthcoming).

[3] Bailey DH, Swarztrauber PN(1993), "A Fast Method for the Numerical Evaluation of Continuous Fourier and Laplace Transforms", SIAM J. on Scientific Computing, vol.15, no.5, pp. 1105-1110.

[4] Bailey DH, Swarztrauber PN(1995), "The Fractional Fourier Transform and Applications", SIAM Review, vol. 33 no.3(Sept. 1991), pg. 389-404.

[5] Bally V., Pages G., and Printems J.(2005), "A Quantization Tree Method for Pricing and Hedging Multidimensional American Options", Mathematical Finance, Vol. 15, No. 1, pp. 119-168.

[6] Baxter MW and Rennie AJO(1996), "An Introduction to Derivative Pricing", Financial Calculus, *Cambridge University Press (UK).*

[7] Breen R.(1991), "The Accelerated Binomial Option Pricing Model", Journal of Financial and Quantitative Analysis, Vol. 26, No. 2, pp. 153-164.

[8] Bervoets F.(2005), "Jump Diffusion Process, The pricing of European and American options", Internship thesis of Rabobank.

[9] Björk T.(2004), "Arbitrage Theory in Continuous Time", 2nd edition, *Oxford University Press.*

[10] Boyle PP, Evnine J. and Gibbs S.(1989), "Numerical Evaluation of Multivariate Contingent Claims", Review of Financial Studies, 1989, 2:241-250.

[11] Broadie M., Detemple J.(1996), "American option valuation: New bounds, approximations, and a comparison of existing methods", The Review of Financial Studies 9, 1211-1250.

[12] Broadie M., Detemple J.(1997), "The Valuation of American Options on Multiple Assets", Mathematical Finance, Vol. 7, nr. 3, pp. 241-286(46).

[13] Broadie M. and Yamamoto Y.(2002), "Application of the Fast Gauss Transform to Option Pricing", Management Sci. 49(8) 1071C1088.

[14] Carr P.(1995), "The Valuation of American Exchange Options with Application to Real Options Real Options in Capital Investments: Models, Strategies, and Aplications", Ed. by L. Trigeorgis, Praeger Publisher, Westport, Conn. pp.109-120.

[15] Carr P., Madan DB(1999), “Option Valuation Using the Fast Fourier Transform”, Journal of Computational Finance, 2, 4, 61-73.

[16] Champeney DC A. (1987), “Handbook of Fourier Theorems”, `http://ccrma.stanford.edu/~jos/mdft/Fourier_Theorems.html`, *Cambridge University Press.*

[17] Chang C., Chung S., and Stapleton RC (2001), “Richardson Extrapolation Techniques for Pricing American-style Options”, EFMA 2002 London Meetings.

[18] Chourdakis K.(2004), “Option Pricing Using the Fractional FFT”, Journal of Computational Finance (2004/05), 8(2).

[19] Detemple J., Feng S., and Tian W.(2003), “The Valuation of American Call Options on the Minimum of Two Dividend-paying Assets”, Ann. Appl. Probab., 13

[20] Dey AK and Ruymgaart FH(1999), “Direct Density Estimation as an Ill-posed Inverse Estimation Problem”, Statistica Neerlandica, Vol.53, nr.3, pp. 309-326.

[21] D’Halluin Y., Forsyth PA, and Labahn G.(2004), “A Penalty Method for American Options with Jump Diffusion Processes”, Numer. Math., 97.

[22] Duan JD, Ritchken P., Sun ZQ(2006), “Approximating Garch-jump Models, Jump-diffusion Processes, and Option Pricing”, Mathematical Finance, Vol. 16, No. 1(January 2006), 21-52.

[23] Duffie D., Pan J. and Singleton K.(2000), “Transform Analysis And Asset Pricing For Affine Jump-Diffusions”, Econometrica, Vol. 68, No. 6.

[24] Gerstner T. and Griebel M.(1998), “Numerical Integration using Sparse Grids”, Num. Alg., 18, pp. 209-232.

[25] Gerstner T. and Griebel M.(2003), “Dimension-adaptive tensor-product quadrature”, Computing 71, pp.65-87.

[26] Gil-Palaez J. (1951), “Note on the Inversion Theorem”, Biometrika, 38 (1951), pp. 481-482.

[27] Harrison JM, and Pliska SR(1981), “Martingales and Stochastic Integrals in the Theory of Continuous Trading”, Stochastic Processes and their Applications, 11:261–271.

[28] Higham DJ(2002), “Nine Ways to Implement the Binomial Method for Option Valuation in MATLAB”, SIAM REVIEW, vol.44. No. 4. pp.661-677.

[29] Hori M.(2001), “Inverse Analysis Method Using Spectral Decomposition of Green’s Function”, Geophysical Journal International, Vol. 147, nr. 1, pp. 77-87(11).

[30] Hull JC(2003), “Options, Futures, & Other Derivatives”, fifth Edition, *Prentice Hall International Editions.*

[31] Jerri AJ(1992), “Integral and Discrete Transforms With Applications and Error Analysis”, Published by *Marcel Dekker*, 848p, ISBN 0824782526 .

[32] Kay J., Davison M., Rasmussen H.(2003), “The Early Exercise Region for Bermudan Put Options on Multiple Underlyings”, submitted to Applied Mathematical Finance.

[33] Konikov M. and Madan D. (2002), “Option Pricing Using Variance Gamma Markov Chains”, Review of Derivatives Research, Vol. 5, nr. 1, pp. 81-115(35).

[34] Lee RW(2004), "Option Pricing by Transform Methods: Extensions, Unification, and Error Control", Journal of Computational finance, 7(3):51-86.

[35] Lord R.(2006), "Fast Pricing of Bermudans within 1D Models", working paper of Erasmus University Rotterdam and Rabobank International.

[36] Madan DB, Carr PP, Chang EC(1998), "The Variance Gamma Process and Option Pricing", European Finance Review 2: 79-105.

[37] Oetiker T. , Partl H.,Hyna I. and Schlegl E.(1999), "The Not So Short Introduction to Latex 2 ϵ".

[38] Omberg E.(1987), "A Note on the Convergence of the Binomial Pricing and Compound Option Models", Journal of Finance, 42, June, 463-469.

[39] O'Sullivan C.(2004), "Path Dependent Option Pricing Under Lévy Process, Applied to Bermudan Options", EFA 2005 Moscow Meetings Paper.

[40] Pinsky MA (1995), "Speed of Convergence of Two-dimensional Fourier Integrals", math.CA/9501218.

[41] Rosenberg J.(1998), "Pricing Multivariate Contingent Claims Using Estimated Risk-neutral Density Functions", Journal of International Money and Finance, 17, 229-247.

[42] Rosenberg J.(1999), "Semiparametric Pricing of Multivariate Contingent Claims", New York University, Leonard N. Stern School Finance Department Working Paper Seires.

[43] Rosenberg J.(2003), "Nonparametric Pricing of Multivariate Contingent Claims", Federal Reserve Bank of New York.

[44] Van Rooij A.(2002), "Fourier Inversion and the Hausdorff Distance", Statistica Neerlandica, Vol. 56 Page 206.

[45] Stulz RM(1982), "Options On the Minimum or the Maximum of Two Risky Assets", Journal of Financial Economics 10, 161-185.

[46] Wilmott P., Howison S., and Dewynne J.(2002), "The Mathematics of Financial Derivatives, A Student Introduction", *Cambridge University Press.*

[47] Van den Goorbergh R. WJ, Genest C., Bas JM Werker (2005), "Bivariate Option Pricing using Dynamic Copula Models", Mathematics and Economics 37, 101-114.

Appendix A

Derivation of Black-Scholes Formula

A.1 Boundary Conditions of European Options

In order to get a unique solution, final and boundary conditions for Equation (2.27) must be provided.

Equation (2.27) is a backward linear parabolic equation; its highest derivative with respect to S is a second derivative, and the highest derivative with respect to t is a first derivative. Thus, two conditions in S should be imposed, but only one in t.

It is important to remember that the parabolic equation cannot be solved in the wrong direction; that is, impose an initial condition on a forward equation, whereas final condition is needed for a backward equation. Otherwise, the problem is ill-posed.

For a European call option, the final condition is

$$C(S,T) = \max(S - E, 0), \tag{A.1}$$

and the 'spatial' or asset-price boundary conditions are applied at zero asset price, $S = 0$, and also for S positive infinity. From (2.5) we can see that if S is zero then dS is also zero, and remains zero for every t. This is the only deterministic case of the stochastic differential equation (2.5). Thus, the call option is worthless on $S = 0$:

$$C(0,t) = 0. \tag{A.2}$$

If the asset price increases it becomes ever more likely that the option will be exercised. So,

$$C(S,t) \sim S \quad \text{as} \quad S \to \infty. \tag{A.3}$$

Analogously, for a European put option, the final condition is

$$P(S,T) = \max(E - S, 0). \tag{A.4}$$

And boundary condition

$$P(0,t) = Ee^{-r(T-t)}; \tag{A.5}$$

when r is fixed; or

$$P(0,t) = Ee^{-\int_t^T r(\tau)d\tau} \tag{A.6}$$

and

$$P(S,t) \to 0 \quad \text{as} \quad S \to \infty. \tag{A.7}$$

A.2 Derivation of Black-Scholes Formula

Equation (2.27) can be transformed to the simple diffusion equation by changing of variables:
With

$$S = Ee^x, \tag{A.8}$$
$$t = T - \frac{\tau}{\frac{1}{2}\sigma^2}, \tag{A.9}$$
$$C = Ev(x,\tau), \tag{A.10}$$

we then get

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} + (k-1)\frac{\partial v}{\partial x} - kv, \tag{A.11}$$

with

$$k = \frac{r}{\frac{1}{2}\sigma^2}, \tag{A.12}$$
$$v(x,0) = \max(e^x - 1, 0). \tag{A.13}$$

With

$$v = e^{\alpha x + \beta\tau} u(x,\tau), \tag{A.14}$$

for some constants α and β differentiation gives

$$\beta u + \frac{\partial u}{\partial \tau} = \alpha^2 u + 2\alpha\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2} + (k-1)\left(\alpha u + \frac{\partial u}{\partial x}\right) - ku. \tag{A.15}$$

By choosing

$$0 = 2\alpha + (k-1),$$
$$\beta = \alpha^2 + (k-1)\alpha - k.$$

we get

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} \quad \text{for} - \infty < x < \infty, \tau > 0, \tag{A.16}$$

with

$$u(x,0) = u_0(x) = \max\left(e^{\frac{1}{2}(k+1)x} - e^{\frac{1}{2}(k-1)x}, 0\right). \tag{A.17}$$

We now find

$$u(x,\tau) = \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} u_o(s) e^{-\frac{(x-s)^2}{4\tau}} ds. \tag{A.18}$$

Replacing the variables of E and t, we get the exact solution of the Black-Scholes equation for a European call:

$$C(S,t) = SN(d_1) - Ee^{-r(T-t)}N(d_2), \tag{A.19}$$

with the cumulative distribution function $N(\cdot)$ for a standardized normal random variable, given by

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}} dy, \tag{A.20}$$

and

$$d_1 = \frac{\log(S-E) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \tag{A.21}$$

$$d_2 = \frac{\log(S-E) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \tag{A.22}$$