THE MATHEMATICAL MODELLING AND NUMERICAL SOLUTION OF OPTIONS PRICING PROBLEMS

By

Sweta Rout

A thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
at
the University of Greenwich,
London, U.K.

June 2005

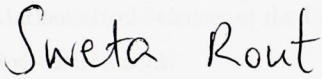
To my father, my mother

and

the loving memory of my grandparents.

Declaration

I certify that this work has not been accepted in substance for any degree, and is not concurrently submitted for any degree other than that of Doctor of Philosophy (PhD) at the University of Greenwich. I also declare that this work is entirely the result of my own investigations unless otherwise stated.



Sweta Rout

Acknowledgements

My deepest gratitude goes wholeheartedly to my supervisor, Professor Kevin Parrott. He has shown a genuine interest in my work, and has always been ready and eager to guide me through the problems I faced while working on this research. His friendly support has been a strength to me.

I am very grateful to Dr Choi-Hong Lai, my second supervisor, for his encouraging advice and guidance at various stages of the work.

Many thanks go to the staff and to the postgraduates at the School of Computing and Mathematical Sciences of the University of Greenwich for providing such a good working environment.

Let me take this opportunity to acknowledge the University of Greenwich research bursary which funded me.

Most importantly, I wish to thank my family and my friends - they loved me, protected me, supported me, taught me, and laughed with me. Especially my parents Prasanta Rauth & Manjushree Rout, my late grandparents Dr Sochi Routroy & Bhudevi Routroy, and my brother Dr Shantanu Rout. To them all, I dedicate this thesis.

Greenwich, London, U.K. Sweta Rout

June 2005

Abstract

Accurate and efficient numerical solutions have been described for a selection of financial options pricing problems. The methods are based on finite difference discretisation coupled with optimal solvers of the resulting discrete systems. Regular Cartesian meshes have been combined with orthogonal co-ordinate transformations chosen for numerical accuracy rather than reduction of the differential operator to constant coefficient form. They allow detailed resolution in the regions of interest where accuracy is most desired, and grid coarsening where there is least interest. These transformations are shown to be effective in producing accurate solutions on modest computational grids. The spatial discretisation strategy is chosen to meet accuracy requirements as well as to produce coefficient matrices with favourable sparsity and stability properties.

In the case of single factor European options, a modified Crank-Nicolson, second order accurate finite difference scheme is presented, which uses adaptive upwind differences when the mesh Peclet conditions are violated. The resulting tridiagonal system of equations is solved using a direct solver. A careful study of grid refinement displays convergence towards the true solution and demonstrates a high level of accuracy can be obtained with this approach. Laplace inversion methods are also implemented as an alternative solution approach for the one-factor European option. Results are compared to those produced by the direct solver algorithm and are shown to be favourable.

It is shown how Semi-Lagrange time-integration can solve the path-dependent Asian pricing problem, by integrating out the average price term and simplifying the finite difference equations into a parameterised Black-Scholes form. The implicit equations that result are unconditionally stable, second order accurate and can be solved using standard tridiagonal solvers. The Semi-Lagrange method is shown to be easily used in conjunction with co-ordinate transformations applied in both spatial directions. A variable time-stepping scheme is implemented in the algorithm. Early exercise is also easily incorporated, the resulting linear complementarity problem can be solved using a projection or penalty method (the penalty method is shown to be slightly more efficient). Second order accuracy has been confirmed for Asian options that must be held to maturity. A comparison with published results for continuous-average-rate put and call options, with and without early exercise, shows that the method achieves basis point accuracy and that Richardson extrapolation can also be applied.

Contents

Chapter 1

A Brief Introduction to Financial Derivatives

The success of financial markets depends largely on educated, informed investors knowing the risk they face, who understand not only the potential for profit but also the potential for loss. All organisations and individuals face financial risk induced by, say, changes in stock market prices, interest rates or exchange rates. Financial instruments for the management of such risk have been developed. These instruments are called **derivative securities** (also known as **financial derivatives, derivative products, contingent claims**, or just **derivatives** or **securities**). Their values are derived from the price of underlying assets, which could include stocks, bonds, interest rates, stock indices, foreign currencies and futures contracts. The modelling of financial derivative products is a fast growing area of applied mathematics with 'real-world' applications to problems originating in modern industry. Financial derivatives were developed to control the management of risk caused by adverse changes in the market. They also provide the opportunity to make a profit for those prepared to accept risk.

1.1 Forwards, Futures, Swaps and Options

Financial derivatives include forwards, futures, options and swaps. Forwards, futures and swaps involve commitments to exchanges of cash flows in the future at prices or rates determined in the present.

1.1.1 Forward contract

Forward contracts are the simplest contingent claim, they are generally traded over the counter. It is an agreement to buy or sell an asset at a certain future date for a specified price. It can be contrasted with a **spot contract** which is an agreement to buy or sell an asset today. On the delivery date of a forward contract, the asset has to be delivered and paid for. Forwards can relate to commitments in the short and medium term (maybe up to 5 years). An essential characteristic of this type of contract is that no money changes hands until delivery. It therefore costs nothing to enter into a forward contract.

1.1.2 Futures contract

A futures contract is essentially a forward contract with technical modifications. These are usually traded on an exchange, unlike forwards which can be set up between any two parties. As such, certain standard features of the contract are specified such as delivery date and contract size. Whereas the profit or loss from a forward contract is only realised at the delivery date, futures contracts are *marked to market*. In other words, an important feature of futures is their daily settlement. The value is calculated every day and the change in value is paid to one party by the other, so that the net profit or loss is paid across gradually over the lifetime of the contract. They tend to be short-term (less than a year).

1.1.3 Swaps

A swap is an agreement between two or more parties to exchange sets of cash flows over a certain period of time. They are the latest of the basic derivatives and were traded first in 1981. Swaps can be seen as a series of forward contracts. Indeed it turns out that one method for pricing swaps is to decompose them into forwards and options. They tend to be longer term than either forwards or futures (their maturities may be more than 10 years).

1.1.4 Option contract

An option provides the right, but not the obligation, to a future exchange at a price determined in the present. A **call option** gives its holder the right to buy the underlying asset by a certain date for a certain price. A **put option** gives its holder the right to sell the underlying asset by a certain date for a certain price. These have the advantage that the holder can choose to ignore them. If the price of

the underlying instrument moves so that the price stipulated by the option becomes unattractive, then the holder of the option may simply choose to disregard it. Options provide protection against adverse movements while preserving the ability to gain from beneficial price movements. However, these benefits of option contracts, relative to the other derivatives, have to be paid for at the outset in the form of an option premium. A corresponding payment is not required when using forwards, futures or swaps.

For example, vanilla options allow one to buy or sell at a preset **exercise price** $K \in \mathbb{R}^+$. Some options are contingent upon the trajectory of the underlying asset price and are thus described as **path-dependent**. One instance of this are **Asian options** (see Chapter 4), which have a payoff depending on the average security price over some time interval. Options may also confer different exercise rights upon the holder. Some options allow early exercise (**American options**), while for others the option right may only be exercised at the expiry date T of the contract (**European options**). Refer to [Hul03], [WDH93] for details on options.

Options have two primary uses: *speculation* and *hedging*. An investor who believes that a particular stock is going to rise can purchase some shares in that company. If he is correct, he will make money; if he is wrong, he will lose money. This investor is speculating. Alternatively, if he thinks that the share price is going to rise within a few months, he can speculate by buying a call option with a specific exercise price and expiry date. On the other hand, if the investor thinks a particular stock is going to fall, he can speculate by selling shares or buying puts. If an investor speculates in an option market, his profit or loss can be magnified to a very high percentage of his investment [WHD97]. Options allow speculators to obtain very high returns, but also exposes them to high risk. Reducing risk is the subject of hedging.

A portfolio is a collection of securities held by an investor. To form a portfolio, an investor needs to know the positions taken in each security under consideration. The value of a portfolio which only contains assets falls when the asset price falls, while one which is all put options rises. Somewhere in between these two extremes is a portfolio in which a small movement in the asset price does not result in any movement in the value of the portfolio. This portfolio is instantaneously risk-free. The reduction of risk by taking advantage of such correlations between the asset and option price movements is called **hedging**. Options provide an effective way to hedge a portfolio. By combining call options and put options with different exercise

prices and expiry dates, one can implement various investment strategies.

1.1.5 Details of the option contract

A standard financial option, the **European Call** option, is a contract with the following conditions:

At a prescribed time in the future, known as the ***expiry date****, the holder of the option may purchase a prescribed asset, known as the* ***underlying asset****, for a prescribed amount, known as the* ***strike price*** *or* ***exercise price****.*

Similarly, the **European Put** option gives the holder the right to sell a particular asset for a certain price at a prescribed time.

Options may be American or European, a distinction which has nothing to do with geography. American options can be exercised at any time up to the expiry date, whereas European options can only be exercised on the expiry date itself. Options of this nature are termed "plain vanilla" or "standard" derivatives.

Since the early 1980's, banks and other financial institutions have been very imaginative in designing non-standard derivatives to meet the needs of clients. Non-standard derivatives are sometimes called **exotic options** or just **exotics**. A description of different types of exotic options and their uses can be found in [Red97], [Hul03], [Wil98]. An example is the non-standard American option in which early exercise is restricted to certain dates. This instrument is called a **Bermudan option**. Another example is the **compound option** which are options on options. There are four main types of compound option: a call on a call, a put on a call, a call on a put, and a put on a put. Compound options have two strike prices and two exercise dates. There are **barrier options** where the payoff depends on whether the underlying asset's price reaches a certain level during a certain period of time. **Asian options** are options where the payoff depends on the average price of the underlying asset during at least some part of the option's life.

Since the option confers on its holder a right with no obligation it has some value. This must be paid for at the time of opening the contract. Conversely, the writer of the option has to be compensated for the obligation he has undertaken. The question is:

- How much would someone pay for this right? i.e., What is the price of the

premium on an option?

- How can the writer take on the minimum of risk associated with his obligation?

1.2 Option Positions

There are two sides to every option contract. On one side is the holder who has taken a *long position* (i.e., has bought the option). On the other side is the writer who has taken a *short position* (i.e., has sold the option). The four types of option positions are:

1. A long position in a call option.
2. A short position in a call option.
3. A long position in a put option.
4. A short position in a put option.

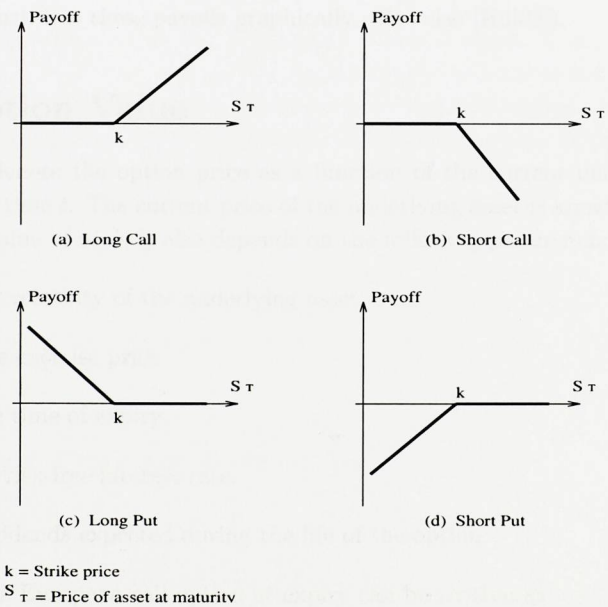


Figure 1.1: Payoffs from different positions in European options.

It is often useful to characterise European option positions in terms of the *terminal value* or **payoff** to the investor at maturity. Then the initial cost of the option is not included in the calculation. Let K be the strike price and S_T be the final price of the underlying asset. The payoff from a long position in a European call option is

$$\max(S_T - K, 0)$$

This reflects the fact that the option will be exercised if $S_T > K$ and will not be exercised if $S_T \leq K$. The payoff to the holder of a short position in a European call option is

$$-\max(S_T - K, 0) = \min(K - S_T, 0)$$

The payoff to the holder of a long position in a European put option is

$$\max(K - S_T, 0)$$

and the payoff from a short position in a European put option is

$$-\max(K - S_T, 0) = \min(S_T - K, 0)$$

Figure 1.1 illustrates these payoffs graphically. (See also [Hul03]).

1.3 Option Value

Let $u(S_t, t)$ denote the option price as a function of the current underlying asset value S_t, and time t. The current price of the underlying asset is known as the **spot price**. The value of $u(S_t, t)$ also depends on the following parameters:

- σ - The volatility of the underlying asset.
- K - The exercise price.
- T - The time of expiry.
- r - The risk-free interest rate.
- The dividends expected during the life of the option.

The value of a European call option at expiry can be written as

$$u(S_T, T) = \max(S_T - K, 0) \tag{1.1}$$

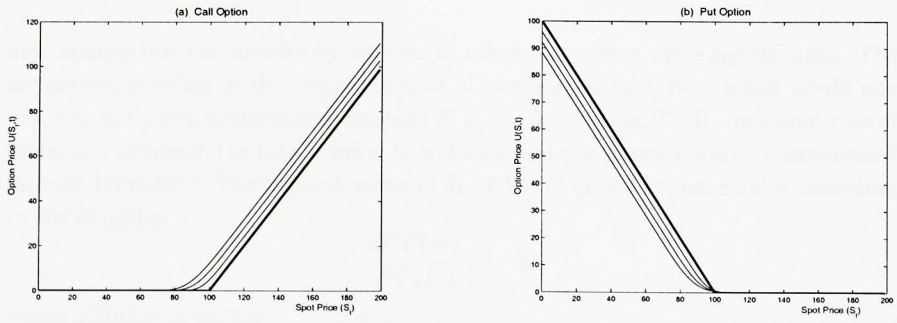


Figure 1.2: Payoff diagrams for vanilla options (strike price, $K = 100$). The bold line shows payoff at expiry, T. The fine lines show payoff at various times prior to expiry.

As the expiry date draws nearer, the value of the call option can be expected to approach the value as given by (1.1). Similarly, the payoff at expiry for a European put option is

$$u(S_T, T) = \max(K - S_T, 0) \tag{1.2}$$

Figure 1.2 displays the value of a call and put option at various times before expiry. Note that for $t < T$ the value of the function can be represented by a smooth continuous curve. Only at expiry (the bold line) does the option value become a piecewise linear function with a singularity at the strike price. Although the two most basic structures for the payoff are the call and the put, an option contract may well be written with a more general payoff. By combining calls and puts with various exercise prices one can construct portfolios with a variety of payoffs. The appeal of such strategies is in their ability to redirect risk. In exchange for the premium, which is the maximum possible loss and whose value is known from the start, one can construct portfolios to benefit from virtually any movement in the underlying asset.

1.4 Basic Pricing Principles

1.4.1 Present Value

Assume that the short-term bank deposit interest rate is a known function of time, not necessarily constant. This is a valid assumption since a typical equity option has a relatively short lifespan of about 9 months, say, within which time the interest rates

may change but not usually by enough to affect the option price significantly. The important question in this regard is that of **present value**. How much would one pay *now* to receive a guaranteed amount K at the future time T? If a constant rate of interest is assumed, the future value K is discounted using continuously compounded interest [WHD97]. The present value of K, $PV(K)$ grows exponentially according to the equation

$$\frac{dPV(K)}{PV(K)} = r\ dt$$

whose solution is simply

$$PV(K) = \mathcal{C}\ e^{rt}$$

where $\mathcal{C}$ is the constant of integration. Since $PV(K) = K$ at time $t = T$, the value at time t of the certain payoff is

$$PV(K) = K\ e^{-r(T-t)} \tag{1.3}$$

If the interest rate is a known function of time $r(t)$ then equation (1.3) can be trivially modified with the result

$$PV(K) = K\ e^{-\int_t^T r(s).ds}$$

1.4.2 Bounds on Option Prices

An option gives its holder a right with no obligation. At the same time the writer of the option has a potential obligation to comply with the contract. According to arbitrage theory the option holder has to pay the option writer for this right: the amount he has to pay is the option price. Arbitrage theory assumes that there is no opportunity to make a sure profit without any risk - or the opportunity would disappear the instant it arose.

Assume c is the price of a European call option and C the price of an American call option. The first bound is that a call option cannot be worth less than zero. Having the right but not the obligation to do something is worth either zero or some positive amount. That is,

$$c \geq 0 \quad \text{and} \quad C \geq 0$$

The next bound is that the current call price cannot exceed the current asset price,

$$c \leq S_t \quad \text{and} \quad C \leq S_t$$

This is because the right to buy an asset cannot be worth more than the asset itself at any time. After all, if the right is exercised, it just gives the asset and no more.

If this relationship were not true, an arbitrageur could make a risk-less profit by buying the underlying asset and selling the call option.

There is a lower bound for the price of a European call option which is derived by considering two alternative investment strategies [Hul03]:

1. Buy a share today at a price S_t, or
2. Buy a European call option on the share at a price of c today and, at the same time, deposit enough money at the risk-free interest rate to yield the exercise cost K at time T. The sum to deposit now is the present value of K, $PV(K)$ which is known from equation (1.3).

At time T the first portfolio will be worth S_T and the second will be worth

$$\max(S_T, K)$$

since, if $S_T > K$, the call option is exercised at maturity and the portfolio is worth S_T; else if $S_T < K$, the call option expires worthless and the portfolio is worth K. Therefore, at the options maturity, the second portfolio must have at least as much value as the first. It follows that in the absence of arbitrage opportunities the same must be true at time prior to T. That is,

$$c + PV(K) \geq S_t$$

or, rearranging,

$$c \geq S_t - PV(K) \tag{1.4}$$

The worst that can happen to an option is that it expires worthless, hence its value cannot be negative. Therefore,

$$c \geq \max(S_t - PV(K), 0) \tag{1.5}$$

Similarly, put options have the following price bounds

$$p \geq 0 \text{ and } P \geq 0 \tag{1.6}$$

$$p \leq K \text{ and } P \leq K \tag{1.7}$$

where p is the value of a European put option and P is the value of an American put option, respectively. The bounds imply that a put cannot have a negative value [(1.6)], and the price of a put cannot exceed the exercise price [(1.7)]. Similar to equation (1.4), a European put option has a lower bound given as

$$p \geq PV(K) - S_t$$

Also, similar to (1.5), the following holds:

$$p \geq \max(PV(K) - S_t, 0) \tag{1.8}$$

In the case of early exercise,

$$\begin{aligned} C &\geq c \\ P &\geq p \end{aligned}$$

since American options provide more exercise opportunities than the corresponding European options. It follows from (1.5) and (1.8) that

$$\begin{aligned} C &\geq S_t - PV(K) \\ P &\geq PV(K) - S_t \end{aligned}$$

Given $r > 0$, it follows that $C > S_t - K$. It is thus never optimal to exercise an American call option on a non-dividend-paying stock before the expiry date. (If it were optimal to exercise early, C would equal $S_t - K$). However, it may be optimal to exercise an American put option on a non-dividend-paying stock early. At any given time during the contract, a put option should always be exercised early if it is sufficiently deeply in the money. For an American put option, the stronger condition $P \geq K - S_t$ must always hold since immediate exercise is always possible. In general, the early exercise of a put option becomes more attractive as S_t decreases, as r increases and as the volatility decreases [Hul03].

1.4.3 Put-Call Parity

Although call and put options are superficially different, they can be combined in such a way as to be perfectly correlated. Here the **put-call parity** is demonstrated by following the argument in [WDH93]. Suppose that at time $t = 0$ one asset is bought and one European put option taken out on the asset, and at the same time one European call option is sold on the asset with the same expiry date, T and exercise price, K. Let Π_t denote the value of the portfolio at any time $t \in (0, T)$. It can be written that

$$\Pi_t = S_t + p - c \tag{1.9}$$

where p and c are the values of the put and call options respectively at time t. The payoff of this portfolio at the expiry date T is

$$S_T + \max(K - S_T, 0) - \max(S_T - K, 0) = \begin{cases} K & if \quad S_T \leq K \\ K & if \quad S_T \geq K \end{cases} \tag{1.10}$$

that is, this portfolio gives a guaranteed value K at time $t = T$ no matter what values of S_t, p and c at expiry. By following the same reasoning as in equation (1.3), the present value of the portfolio is $K\ e^{-r(T-t)}$. Thus

$$S_t + p - c = K\ e^{-r(T-t)} \tag{1.11}$$

which is known as the *put-call parity* for any time $t \in [0, T]$. Thus, the values of European put and call options on the same asset can be determined from each other if they have the same exercise price and expiry date.

1.5 Pricing Financial Derivatives

Assuming some known stochastic description of the evolution of the underlying variables which the derivative depends upon, a convection-diffusion partial differential equation (PDE) can be derived, which the financial derivative must satisfy under the further assumptions of frictionless, arbitrage-free markets. In the case where the derivative security permits early exercise, a free boundary problem (which can be formulated as a linear complementarity problem) is obtained. These convection-diffusion equations can also be derived by martingale-based arguments, see [RB96]. The formulation of the pricing problem in probabilistic terms can also be employed, where the solution is given as the discounted value of the expected payoff under the appropriate probability measure.

In general, analytic solutions are not readily available for either formulation of the problem and numerical methods have to be employed to determine the solution and hence the price of the financial derivative. Given the practical nature of the problem, the numerical solution needs to satisfy the criteria that it is accurate and quickly determined. Financial modelling typically requires large number of simulations and hence computing resources and efficiency of algorithms are very important to make evaluations and decisions before the agreement of a contract.

Pricing models can be solved mathematically by many methods. Numerical approximation by either finite difference or finite element methods is described in [WDH93] for single factor options, and either method can be readily extended to price two- or multi-factor problems. The paper [ZFV99] in particular deals with the application of finite element methods to stochastic volatility European lookback options. Another approach is detailed in [CDLL97], where finite volume methods are found to be of interest. This thesis seeks a numerical solution of the pricing

PDE using finite difference discretisation. The method and particular techniques used produce competitive results, which satisfy the performance criteria.

1.6 An outline of the thesis

The remainder of the thesis is organised as follows. In chapter 2, mathematical modelling of financial derivatives and the corresponding numerical methods are introduced. The chapter begins with a review of the tools used in continuous time financial modelling, namely stochastic differential equations and Ito's lemma. The geometric Brownian option model for security prices is discussed and the no-arbitrage approach to pricing a single factor option is outlined. The approach is extended to the case where the derivative security depends on multiple stochastic variables some of which are not tradable assets. The next section reviews some of the numerical methods used in the literature for the solution of the derivatives problem and the chapter concludes with a very brief outline of the general numerical approach pursued in this thesis.

In chapter 3, a method based on finite difference discretisation is used to solve the option pricing problem for European options. The mathematical problem is described followed by numerical techniques for computing the value of the option. The finite difference discretisation scheme is presented and a description of the resulting discrete problem follows. The algorithm with a direct solver is applied to the discrete system with emphasis placed on exploiting the special structure of the coefficient matrix. Numerical results are then presented which suggest that solutions to desired levels of accuracy can be obtained in seconds. The hedging parameters are also discussed. The next study in this chapter looks at ways of improving the accuracy of the solutions and efficiency of the solution method. Co-ordinate transformations are applied which ensure grid refinement in regions of interest. Once again, the discrete system is solved by exploiting the special structure of the coefficient matrix. A variety of options with different payoffs are priced using the co-ordinate transformed algorithm. It is shown by presenting numerical results that the required degree of accuracy can be achieved with fewer mesh points than when co-ordinate transformations are not applied. Furthermore, non-uniform time-stepping routines were used to enhance efficiency. To conclude this chapter a Laplace transform application to the pricing of a European put option is described. Results are compared to those produced by the direct solver algorithm and are shown to be comparable. It is suggested that the Laplace transform technique could be made use of for non-linear pricing problems.

In chapter 4, a Semi-Lagrangian time integration scheme is used to solve fixed strike average options. This problem is two-dimensional and is hyperbolic in the average direction. This hyperbolicity makes accurate solutions difficult to obtain but is adequately dealt with by using a Semi-Lagrangian integration scheme. The chapter begins with a review of the literature associated with average options. The mathematical model of the Asian option is then presented. Co-ordinate transformations are then used to ensure that regular discretisation of the transformed equation will result in detailed resolution in the areas of interest in the orthogonal co-ordinate system. Numerical results are then presented which suggest that accurate solutions can be obtained rapidly using this method.

Chapter 5 discusses the application of the Semi-Lagrangian scheme to the American style Asian problem. Spatial discretisation is outlined and the resulting discrete linear complementarity problems are presented. A non-uniform time stepping algorithm is introduced and implemented. A comparison of Projected Relaxation and Penalty method solution for the linear complementarity problem is shown.

Finally, in chapter 6, some conclusions are drawn which summarise the main points of the thesis and encapsulates its overall direction. Various suggestions are then given for the direction of future work.

Chapter 2

The Mathematical Modelling of Options Pricing

In this chapter, the mathematical modelling of financial derivatives and the corresponding numerical methods are introduced. The chapter begins with a review of the tools used in continuous time financial modelling, namely stochastic differential equations and Ito's lemma. The geometric Brownian option model for security prices is discussed and the no-arbitrage approach to pricing a single factor option is outlined. The approach is extended to the case where the derivative security depends on multiple stochastic variables some of which are not tradable assets. The next section reviews some of the numerical methods used in the literature for the solution of the derivatives problem and the chapter concludes with a very brief outline of the general numerical approach pursued in this thesis.

2.1 Introduction

Certain aspects of pricing derivatives set them apart from the general theory of asset valuation. By making simplifying assumptions, the arbitrage-free price of a derivative can be expressed as a function of some basic securities. A set of formulae can then be obtained which are used to price the option without having to consider any linkages to other financial markets or to the real side of the economy [Nef96].

There are specific ways to obtain such formulae. One method utilizes the notion of arbitrage to determine a probability measure under which financial assets behave as *martingales* once properly discounted. The tools of martingale arithmetic become available and one can calculate arbitrage-free prices by evaluating the implied expectations.

The second pricing method that uses arbitrage takes a somewhat more direct approach. A risk-free portfolio is constructed from which is obtained a PDE implied by the lack of arbitrage opportunities. This PDE is either solved analytically or numerically.

In either case, the problem of pricing derivatives is to find a function $u(S_t, t)$ that relates the price of the derivative product to S_t, the price of the underlying asset and possibly to some other market risk factors. When a closed-form formula cannot be determined, numerical methods are found to describe the dynamics of $u(S_t, t)$. The rest of the thesis is concerned with how to determine such pricing functions $u(S_t, t)$ for linear and non-linear derivatives. The material presented in this chapter is a short synthesis of established results needed to derive the pricing equations.

2.2 Definitions

Stochastic processes are systems which evolve probabilistically in time and they can have discrete or continuous sample paths [Gar94]. Brownian motion is a particular type of real-valued stochastic process. Financial modelling often involves calculations in continuous time where uncertainty is represented by Brownian motion. Some concepts in probability and stochastic calculus are introduced before Brownian motion is defined.

Results from stochastic calculus are used as tools for model manipulation enabling the practitioner to work with and make sense of the models. The concepts of the stochastic integral (interpreted in the Ito sense) and the stochastic differential are of central importance, and Ito's lemma is a major tool for manipulation of functions of stochastic variables. These concepts are briefly reviewed in order to set the framework for the pricing of financial derivatives. One can refer to [Gar94] for an introduction to stochastic methods. Advanced treatment of Brownian motion and stochastic calculus are dealt with in detail in [RY99] and [RW00]. A rigorous treatment of stochastic differential equations is to be found in [Fri75]. The relevance of these concepts to financial modelling is demonstrated in [Duf01], [Ing87] and [Mer92]. The application of stochastic calculus to the pricing of financial derivatives is covered in [Nef96] and [RB96]. Also [Hul03] and [WDH93] demonstrate the use of the results in derivative pricing.

2.2.1 Notation

Observe a family of random variables. Let time be continuous and deal with continuous-time stochastic processes. Let the observed process be denoted by $\{X_t,\ t \geq 0\}$. Let $\{I_t,\ t \geq 0\}$ represent a family of information sets that become continuously available to the decision maker as time passes. For $t_{i-1} < t_i < t_{i+1}$, this family of information sets satisfies

$$I_{t_{i-1}} \subseteq I_{t_i} \subseteq I_{t_{i+1}}$$

The set $\{I_t,\ t \geq 0\}$ is called a *filtration.*

Now consider the random price process X_t during the finite interval $[0, T]$. At some particular time t_i, the value of the price process will be X_{t_i}. If the value of X_t is included in the information set I_t at each $t \in [0, T]$, then it is said that $\{X_t,\ t \in [0, T]\}$ is *adapted* to $\{I_t,\ t \in [0, T]\}$. That is, X_t is 'known' at time t (see [Gar94]).

2.2.2 Continuous-Time Martingales

Using different information sets one can conceivably generate different "forecasts" of a process $\{X_t\}$. These forecasts are expressed using conditional expectations. In particular,

$$E_t[X_T] = E[X_T|I_t],\ t < T$$

is the formal way of denoting the forecast of a future value, X_T of X_t, using the information available as of time t [Nef96]. The defining property of a martingale relates to these conditional expectations.

DEFINITION 2.1 *A process $\{X_t,\ t \geq 0\}$ is said to be a* **martingale** *with respect to the family of information sets I_t and with respect to the probability P, if, for all $t > 0$,*

1. *X_t is known, given I_t. (X_t is I_t-adapted.)*
2. *Unconditional "forecasts" are finite:*
$$E[X_t] < \infty$$
3. *$E_t[X_T] = X_t$, for all $t < T$, with probability 1. That is, the best forecast of unobserved future values is the last observation on X_t.*

Here, all expectations $E[\cdot], E_t[\cdot]$ are assumed to be taken with respect to the probability P.

According to this definition, martingales are random variables whose future variations are completely unpredictable given the current information set. A martingale

is always defined with respect to some information set, and with respect to some probability measure. If the information content and/or the associated probabilities are changed, the process under consideration may cease to be a martingale. The opposite is also true. Given a process X_t which does not behave like a martingale, it may be possible to modify the relevant probability measure P and convert X_t into a martingale.

2.2.3 Wiener Process

In continuous time, "normal" events can be modelled using the Wiener process, or Brownian motion. This is a continuous stochastic process and can be used if markets are dominated by "ordinary" events while "extremes" occur only infrequently, according to the probabilities in the tail areas of a normal distribution. A Wiener process is appropriate if the underlying random variable, say W_t, can only change continuously. With a Wiener process, during a small time interval h, one in general observes "small" changes in W_t, and this is consistent with the events being "ordinary". The formal definition of Wiener processes approached as martingales is given (refer to [Nef96]) as follows:

DEFINITION 2.2 *A* **Wiener process** W_t, *relative to a family of information sets* I_t, *is a stochastic process such that*

1. *The pair* I_t, W_t *is a square integrable martingale with* $W_0 = 0$ *and*
$$E[(W_t - W_s)^2] = t - s, \quad s \leq t$$
2. *The trajectories of* W_t *are continuous over* t.

This definition indicates the following properties of a Wiener process:

- W_t has independent increments because it is a martingale, and because every martingale has unpredictable increments.
- W_t has zero mean because it starts at zero, and the mean of every increment equals zero.
- W_t has variance t.
- The process is continuous in the sense that in infinitesimal intervals, the movements of W_t are infinitesimal.

Note that in this definition nothing is said about increments being normally distributed. When the martingale approach is used, the normality follows from the

assumptions stated in the definition [1]. A Wiener process is the natural model for an asset price that has unpredictable increments but nevertheless moves over time continuously.

2.2.4 Brownian Motion

The definition of a Wiener process given above used the fact that W_t is a square integrable martingale. The definition of Brownian motion is now given. (See [Nef96]).

DEFINITION 2.3 *A random process* $\{B_t,\ t \in [0,T]\}$, *relative to a family of information sets* I_t, *is a (standard)* **Brownian motion** *if:*

1. *The process begins at zero,* $B_0 = 0$.
2. B_t *has stationary, independent increments.*
3. B_t *is continuous in time.*
4. *For any* $0 \leq s \leq t$, *the increments* $B_t - B_s$ *have a normal distribution with mean zero and variance* $|t-s|$*:*

$$(B_t - B_s) \sim N(0, |t-s|)$$

This definition is, in many ways, similar to that of the Wiener process. There is, however a crucial difference. W_t is assumed to be martingale, while no such statement is made about B_t. Instead, it is posited that B_t has a normal distribution.

The Lévy theorem states that there are no differences between the two processes.

THEOREM 2.4 (LÉVY) *Any Wiener process* W_t *relative to a family* I_t *is a Brownian motion process.*

This theorem is very explicit. The terms Wiener process and Brownian motion can be used interchangeably [Nef96]. Hence, no distinction is made between these two concepts in the remainder of the thesis.

Brownian Motion in Multiple Dimensions

Let $\{W_{t_1},\ t_1 \in [0,T]\}, \ldots, \{W_{t_N},\ t_N \in [0,T]\}$ be Brownian motions which are relative to a family of information sets $\{I_{t_i}\}$, $i = 1, 2, \ldots, N$, and are with respect to the probability P. The vector process

$$W_t = \{W_{t_1}, \ldots, W_{t_N}\}, \quad t \in [0,T]$$

[1] Suppose W_t is a process that is continuous, has finite variance (i.e., it is square integrable), and has increments that are unpredictable given the family of information sets $\{I_t\}$. Then, by a famous theorem of Lévy, these properties are sufficient to guarantee that the increments in W_t are normally distributed with mean zero and variance $\sigma^2 dt$.

is an N-dimensional Brownian motion with respect to the family of information sets I_t, and the probability P such that the vector increments $W_t - W_s$, $t \neq s$ are independent and have a multivariate normal distribution in $\mathbb{R}^N$ with mean zero and variance-covariance $(|t-s|)\mathtt{I}$, where $\mathtt{I}$ is the identity matrix.

2.3 Asset Price Evolution

A stock option pricing model requires certain assumptions about the evolution of stock prices over time. Almost all models of option pricing are founded upon one simple model for asset price movements, involving parameters derived from historical data, for example.

In a competitive market, a primary hypothesis is that asset prices accurately reflect all available information. Under this hypothesis, unanticipated changes in the asset value are a **Markov Process** [2]. A typical example of this kind of behaviour is the **random walk**. Percentage changes in the stock price in a short period of time are normally distributed.

Let the asset price be S_t at time t. In the time dt, let the asset price change to $S_t + dS_t$. Here, since the absolute change in the asset price is not a useful quantity by itself, let the *return* be the change in the price divided by the original value, denoted by dS_t/S_t (see [WHD97]). This relative measure of the change is a better indicator of its size than any absolute measure. Define the following:

$$\begin{aligned} \mu &: \text{average growth rate of the asset price} \\ \sigma &: \text{volatility of the stock price} \end{aligned}$$

μ is also known as the drift. In simple models it is taken to be constant. In more complex models, e.g for exchange rates, μ can be a function of S_t and t. It gives a contribution $\mu\ dt$ to the return dS_t/S_t. This contribution is the predictable, deterministic, anticipated return equivalent to the money invested in a risk-free bank account. The second contribution to dS_t/S_t models the random change in the asset price in response to external effects, such as unexpected news. A random sample drawn from a normal distribution with zero mean, dW_t, gives the term $\sigma\ dW_t$. Here, W_t is a standard Brownian motion.

[2] A **Markov Process** is a stochastic process where the expected value of a random variable, say X_i conditional upon all of the past events *depends only on the last value* X_{i-1} and not on the values prior to that.

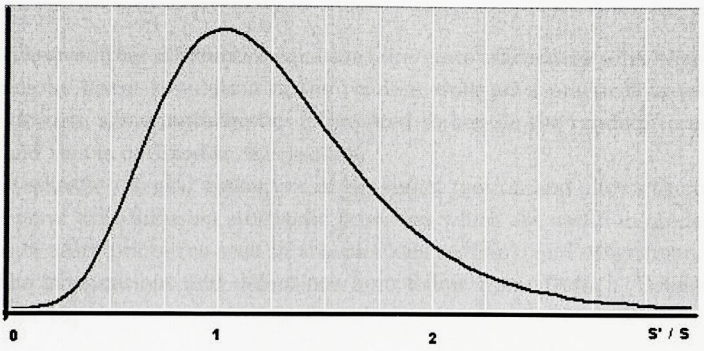


Figure 2.1: Probability Density Function for ${S_t}^+/S_t$

Thus the stochastic differential equation (SDE) is obtained,

$$\frac{dS_t}{S_t} = \sigma dW_t + \mu dt \tag{2.1}$$

which is the mathematical representation of the evolution of asset prices. The SDE is known as **geometric Brownian motion**.

Equation (2.1) is a particular example of a random walk. It cannot be solved to give a deterministic path for the share price, but it can give interesting and important information concerning the behaviour of S_t in a probabilistic sense. Suppose that today's asset price is S_t at time t. If the price at a later date t^+ is ${S_t}^+$, then ${S_t}^+$ will be distributed about S_t with a probability density function (pdf) of the form shown in Figure 2.1. The future asset price ${S_t}^+$ is thus most likely to be close to S_t and less likely to be far away. This distribution is more spread out the larger that t^+ is to t. If S_t follows the random walk given by equation (2.1) then the pdf represented by this skewed bell-shaped curve is the lognormal distribution and the random walk (2.1) is therefore known as a lognormal random walk.

2.4 Tools for Integration

In real life asset prices are quoted at discrete time intervals. Therefore, there is a practical lower bound for the basic time-step dt of the random walk (2.1). If options were valued using this time-step in practice, the data to deal with would be too large. Instead, the mathematical models are set up using the continuous time limit

$dt \to 0$. The resulting differential equations are more efficient to solve than valuing the options by direct simulation of the random walk on a practical timescale. In order to do this, some mathematics is required to handle the random term dW_t as $dt \to 0$, and this is outlined in this section.

The stochastic integral makes use of Brownian motion and allows the construction of general drift-diffusion stochastic processes which are useful in modelling the evolution of asset prices (as seen in the previous section), and other financial variables. The explanations and definitions here follow from [Duf01]. Details can be found in [Fri75].

2.4.1 The Ito Integral and Stochastic Integration

The procedure to obtain ordinary differential equations is not applicable in stochastic calculus. If unpredictable "news" arrives continuously (in time), and if equations representing the dynamics of the phenomena under consideration are a function of such noise, a meaningful notion of derivative cannot be defined. Yet, under some conditions, an integral can be obtained successfully. This permits replacing *ordinary* differential equations by *stochastic* differential equations

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t, \quad t \geq 0$$

where future movements are expressed in terms of differentials dX_t, dt and dW_t instead of derivatives such as dX_t/dt. These differentials are defined using another concept of integral. As h gets smaller, the increments

$$X_{t+h} - X_t = \int_t^{t+h} dX_u$$

can be used to give meaning to dX_t.

Consider the general SDE which represents dynamic behaviour of some asset price S_t

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t, \quad t \geq 0 \tag{2.2}$$

where $\mu(S_t, t)$ and $\sigma(S_t, t)$ are the drift and diffusion coefficients respectively. Equation (2.1) is a specific instance of this general SDE. Taking integrals on both sides gives,

$$\int_t^{t+h} dS_u = \int_t^{t+h} \mu(S_u, u)du + \int_t^{t+h} \sigma(S_u, u)dW_u$$

where the last term on the right hand side is an integral with respect to increments in the Wiener process W_t. The interpretation of this is not immediate as W_t is a

highly irregular process and its derivative does not exist in the sense of deterministic calculus. Obtaining a formal definition of the Ito integral makes the notion of a SDE more precise. Once the integral

$$\int_t^{t+h} \sigma(S_u, u)\, dW_u$$

is defined in some precise way (see [Fri75]), it is possible to integrate both sides of the SDE (2.2) to get

$$S_{t+h} - S_t = \int_t^{t+h} \mu(S_u, u)\, du + \int_t^{t+h} \sigma(S_u, u)\, dW_u$$

where h is some "small" finite time interval. Indeed, h is small enough that $\mu(S_t, t)$ and $\sigma(S_t, t)$ do not change very much during the time interval, especially if they are smooth functions of S_t and t. The equation could be rewritten as

$$S_{t+h} - S_t = \mu(S_t, t) \int_t^{t+h} du + \sigma(S_t, t) \int_t^{t+h} dW_u$$

Taking the integrals in a straightforward way, the forward difference approximation may be obtained,

$$S_{t+h} - S_t \cong \mu(S_t, t)h + \sigma(S_t, t)[W_{t+h} - W_t]$$

which is

$$\Delta S_t \cong \mu(S_t, t)h + \sigma(S_t, t)\Delta W_t$$

The approximation here is in the sense of dropping higher-order terms involving h in a Taylor series expansion of $E_t[S_{t+h} - S_t]$, setting it equal to a first-order Taylor series approximation

$$E_t[S_{t+h} - S_t] = \mu(S_t, t)h$$

Also, $\mu(S_u, u)$, $\sigma(S_u, u)$, $u \in [t, t+h]$ are approximated by their value at $u = t$. Both of these approximations require some smoothness conditions on $\mu(S_u, u)$ and $\sigma(S_u, u)$ (see [Nef96]). All of this implies that by writing

$$dS_t \;=\; \mu(S_t, t)dt \;+\; \sigma(S_t, t)dW_t$$

it means that in the integral equation,

$$\int_t^{t+h} dS_u = \int_t^{t+h} \mu(S_u, u)du + \int_t^{t+h} \sigma(S_u, u)dW_u$$

the second integral on the right-hand-side is defined in the Ito sense and that as $h \to 0$,

$$\int_t^{t+h} \sigma(S_u, u) dW_u \simeq \sigma(S_t, t) dW_t$$

That is, the diffusion terms of the SDE's are in fact Ito integrals approximated during infinitesimal time intervals. For these approximations to make sense, an integral with respect to W_t should first be defined formally. The ability to go from the SDE during the finite interval

$$\Delta S_k = \mu(S_k, k)h + \sigma(S_k, k)\Delta W_k; \quad k = 1, 2, \ldots, n$$

to the SDE

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t, \quad t \geq 0$$

and vice versa, is the ability to interpret dW_t by defining $\int_t^{t+h} \sigma(S_u, u) dW_u$, in a meaningful manner, which can only be done by constructing a stochastic integral. Conditions must also be imposed on the way $\mu(S_t, t)$ and $\sigma(S_t, t)$ move over time. Otherwise these parameters may be too erratic. See [Fri75] for regularity conditions.

Increments in a Wiener process, dW_t, represent random variables that are unpredictable, even in the immediate future. The value of the Wiener process at time t, written as W_t, is then a sum of an uncountable number of independent increments:

$$W_t = \int_0^t dW_u \tag{2.3}$$

At time zero, the Wiener process has a value of zero. This is the simplest stochastic integral one can write down. A more relevant stochastic integral is obtained by integrating the innovation term in the SDE

$$\int_0^t \sigma(S_u, u) dW_u \tag{2.4}$$

The integrals in equations (2.3) and (2.4) are summations of very erratic random variables. Two such variables that are $\epsilon > 0$ apart from each other, dW_t and $dW_{t+\epsilon}$, are still uncorrelated.

Consider the SDE (2.2) written over finite intervals of equal length h:

$$S_k - S_{k-1} = \mu(S_{k-1}, k)h + \sigma(S_{k-1}, k)\Delta W_k, \quad k = 1, 2, \ldots, n$$

Summing the increments,

$$\sum_{k=1}^{n}[S_k - S_{k-1}] = \sum_{k=1}^{n}[\mu(S_{k-1}, k)h] + \sum_{k=1}^{n}\sigma(S_{k-1}, k)[\Delta W_k]$$

Can a methodology similar to the Riemann-Stieltjes integration approach be used to define an integral with respect to the random variable S_t as (some type of) a limit

$$\int_0^T dS_u = \lim_{n\to\infty}\left\{\sum_{k=1}^{n}[\mu(S_{k-1}, k)h] + \sum_{k=1}^{n}\sigma(S_{k-1}, k)[\Delta W_k]\right\}$$

where it is assumed that $T = nh$?

The first term on the right-hand side does not contain any random terms once the information at time k becomes available. More importantly, the integral is taken with respect to increments in time h. Time is a smooth function and has finite variation. This means that the same procedure used for the Riemann-Stieltjes case can be applied to define an integral such as [3]

$$\int_0^T \mu(S_u, u)du = \lim_{n\to\infty}\sum_{k=1}^{n}[\mu(S_{k-1}, k)h]$$

However, the second term on the right-hand side contains random variables even after I_{k-1} is revealed. In fact, as of time $k-1$, the term

$$[W_k - W_{k-1}]$$

is a random variable, and the sum

$$\sum_{k=1}^{n}\sigma(S_{k-1}, k)[W_k - W_{k-1}] \tag{2.5}$$

is an integral with respect to a random variable. Under some conditions (see [Fri75]), it is possible to define a stochastic integral as the limit *in mean square* of the random sum (2.5). This integral would be a random variable.

[3] The sum on the right-hand side can be written in more detailed form as

$$\lim_{n\to\infty}\sum_{k=1}^{n}[\mu(S_{(k-1)h}, kh)][(k)h - (k-1)h]$$

with $kh = t_k$.

The use of *mean square convergence* implies that the difference between the sum (2.5) and the random variable called the **Ito integral**,

$$\int_0^T \sigma(S_u, u) dW_u$$

has a variance that goes to zero as n increases toward infinity. Formally, that is,

$$\lim_{n\to\infty} E\left[\sum_{k=1}^{n} \sigma(S_{k-1}, k)[W_k - W_{k-1}] - \int_0^T \sigma(S_u, u) dW_u\right]^2 = 0$$

A definition can now be provided for the Ito integral within the context of SDE's.

DEFINITION 2.5 *Consider the finite interval approximation*

$$S_k - S_{k-1} = \mu(S_{k-1}, k)h + \sigma(S_{k-1}, k)[W_k - W_{k-1}], \quad k = 1, 2, \ldots, n$$

where $[W_k - W_{k-1}]$ *is a standard Wiener process with zero mean and variance* h. *Let*

1. *The* $\sigma(S_t, t)$ *be non – anticipative, in the sense that they are independent of the future.*
2. *The random variables* $\sigma(S_t, t)$ *be non – explosive:*

$$E\left[\int_0^T \sigma(S_t, t)^2 dt\right] < \infty$$

Then, the **Ito integral** *is defined as,*

$$\int_0^T \sigma(S_t, t) dW_t$$

is the mean square limit

$$\sum_{k=1}^{n} \sigma(S_{k-1}, k)[W_k - W_{k-1}] \to \int_0^T \sigma(S_t, t) dW_t$$

as $n \to \infty$

According to this definition, as the number of intervals tends to infinity and the length of each interval becomes infinitesimal, the finite sum tends to the random variable represented by the Ito integral. Clearly, the definition makes sense only if such a limiting random variable exists. The assumption that $\sigma(S_{k-1}, k)$ is non-anticipating turns out to be a fundamental condition for the existence of such a limit

[Fri75]. One technical point is whether the limiting random variable, that is the Ito integral depends on the choice of how the partitioning of $[0, T]$ is carried out. It can be shown that the choice of partition does not influence the value of the Ito integral (see [Nef96]).

To summarise, there are some major differences between deterministic and stochastic integrations. The notion of limit used in stochastic integration is different. The Ito integral is defined for non-anticipative functions only. Finally, while integrals in standard calculus are defined using the actual paths followed by functions, stochastic integrals are defined within stochastic equivalence.

Properties of the Ito Integral

1. **Martingale property.** The Ito integral is a martingale. The condition that ensures this martingale property is the one that requires σ_t be non-anticipative given the information set I_t. This property is useful in modelling the innovation terms of asset prices in financial theory. The property is also important for practical calculations of asset prices.

2. **Existence.** The Ito integral of a general random function $F(X_t, t)$, given by

$$\int_0^t F(X_u, u) dX_u$$

exists if the function F is continuous, and if it is non-anticipating. In other words, the finite sums

$$\sum_{i=1}^{n} F(X_{t_i}, t_i)[X_{t_i} - X_{t_{i-1}}]$$

converge in mean square to "some" random variable that is called the Ito integral. Although it may exist, determining such a limit explicitly is not guaranteed.

3. **Correlation property.** Being a random variable (or more precisely, being a random process), the Ito integral has various moments. The martingale property gives the first moment of the integral of a non-anticipating F with respect to a Wiener process W_t (see [Nef96]),

$$E\left[\int_0^T F(W_t, t) dW_t\right] = 0$$

The second moments are given by the variance and covariances [4],

$$E\left[\int_0^t F_1(W_u,u)dW_u \int_0^t F_2(W_u,u)dW_u\right] = \int_0^t E[F_1(W_u,u)F_2(W_u,u)]du$$

and

$$E\left[\int_0^t F(W_u,u)dW_u\right]^2 = E\left[\int_0^t F(W_u,u)^2du\right]$$

4. **Addition.** The Ito integral also has some properties that are similar to those of the Riemann-Stieltjes integral. In particular, the integral of the sum of two random functions of X_t is equal to the sum of their integrals:

$$\int_0^T [F_1(X_t,t)+F_2(X_t,t)]dX_t = \int_0^T F_1(X_t,t)dX_t + \int_0^T F_2(X_t,t)dX_t$$

The Stochastic Differential

The stochastic integral can be used to build a class of processes which are useful in many areas including continuous time financial modelling. If $X_t, t \in [0,T]$ is a stochastic process such that for any $0 \le t_1 \le t_2 \le T$,

$$X_{t_2} - X_{t_1} = \int_{t_1}^{t_2} \mu(X_t,t)dt + \int_{t_1}^{t_2} \sigma(X_t,t)dW_t$$

where $\mu(X_t,t)$ and $\sigma(X_t,t)$ satisfy the regularity conditions which guarantee the existence of the integrals on the r.h.s (see [Fri75]), then X_t is said to satisfy the SDE,

$$dX_t = \mu(X_t,t)dt + \sigma(X_t,t)dW_t$$

For the purposes of this thesis, SDE's are denoted in this format rather than the integral format.

[4]Here, a result is used which may seem a bit unusual to one who is used to working with standard deterministic calculus, namely

$$\int_0^T (dx_t)^2 = \int_0^T dt$$

where the equality holds in the mean square sense. It is in this sense that, if W_t represents a Wiener process, for infinitesimal dt one can write

$$(dW_t)^2 = dt$$

In fact, in all practical calculations dealing with stochastic calculus, it is a common practice to replace the terms involving dW_t^2 by dt. The equality should be interpreted in the sense of mean square convergence.

Existence and Uniqueness of Solutions

If the vector of functions μ and the matrix of functions σ satisfy certain conditions of continuity and bounded growth (see [Fri75], [RW00]), the SDE's as defined above can be shown to have unique solutions. Continuity conditions are those of Lipschitz continuity. For square root processes, with a non-Lipschitz continuous diffusion coefficient, a special result is used to show uniqueness and existence (see [RW00]).

2.4.2 Ito's Lemma

Let X_t be a continuous time process which depends on the Wiener process W_t. Given a function of X_t, denoted by $F(X_t, t)$, suppose the change in F when dt amount of time passes needs to be calculated. Passing time would influence $F(X_t, t)$ in two different ways. There is a direct influence through the t variable in $F(X_t, t)$. Also, as time passes, one obtains new information about W_t and observes a new increment, dX_t. This will also make F change. The sum of these two effects is represented by the stochastic differential $dF(X_t, t)$ and is given by the stochastic equivalent of the chain rule in standard calculus.

To calculate Ito's formula, the Taylor series is used. In standard calculus, the Taylor series expansion of a smooth (i.e, infinitely differentiable) function $f(x)$ around some arbitrary point x_0 is given by,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + R \tag{2.6}$$

where R denotes the remainder.

Apply the Taylor series formula to $F(X_k, k)$, $k = 1, 2, \ldots$ where the X_k is assumed to obey

$$\Delta X_k = \mu(X_k, k)h + \sigma(X_k, k)\Delta W_k$$

Now fix k. Given the information set I_{k-1}, X_{k-1} is a known number. Apply the Taylors formula to expand $F(X_k, k)$ about X_{k-1} and $k - 1$,

$$\begin{aligned} F(X_k, k) = F(X_{k-1}, k-1) \quad + \quad & F_x[X_k - X_{k-1}] + F_t[h] + F_{xt}[h(X_k - X_{k-1})] + \\ & \frac{1}{2}F_{xx}[X_k - X_{k-1}]^2 + \frac{1}{2}F_{tt}[h]^2 + R \end{aligned} \tag{2.7}$$

where the partial derivatives $F_x, F_{xx}, F_t, F_{tt}, F_{xt}$ are all evaluated about $(X_{k-1}, k-1)$. R represents the remaining terms of the Taylor series expansion. Here the F_t, F_{xt}, F_{tt} notation is kept for convenience, although these partials are in k. The equality would have to be interpreted in the sense of mean square convergence. A shorthand is also

introduced at this point for the quantities $\mu(X_t, t)$ and $\sigma(X_t, t)$. They may be referred to, respectively, as μ_t and σ_t. The time subscript denotes that they change over time. Substitute for ΔX_k in the Taylor series expansion,

$$\begin{aligned}\Delta F(k) &= F_x[\mu_k h + \sigma_k \Delta W_k] + F_t[h] + F_{xt}(h)[\mu_k h + \sigma_k \Delta W_k] + \\ &\quad \frac{1}{2} F_{xx}[\mu_k h + \sigma_k \Delta W_k]^2 + \frac{1}{2} F_{tt}[h]^2 + R \end{aligned} \tag{2.8}$$

In order to obtain a chain rule in stochastic environments, the terms on the r.h.s are classified as negligible and non-negligible. In "small" time intervals, negligible terms can be dropped from the r.h.s and a chain rule formula can be obtained. Additionally, as $h \to 0$, a limiting argument can be used and a precise formula can be obtained in mean square sense. This formula is known as Ito's lemma.

LEMMA 2.6 (ITO'S LEMMA) *Let $F(X_t, t)$ be a twice-differentiable function of t and of the random process X_t, which obeys*

$$dX_t = \mu_t dt + \sigma_t dW_t, \quad t \geq 0 \tag{2.9}$$

with well-behaved drift and diffusion parameters, μ_t, σ_t [5]*. Then*

$$dF_t = \frac{\partial F}{\partial x} dX_t + \frac{\partial F}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} \sigma_t^2 dt$$

or, after substituting for dX_t using the relevant SDE (2.9),

$$dF_t = \left[\frac{\partial F}{\partial x} \mu_t + \frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} \sigma_t^2 \right] dt + \frac{\partial F}{\partial x} \sigma_t dW_t$$

where the equality holds in the mean square sense.

Ito's lemma is a fundamental result of Ito stochastic calculus and is of central importance to the pricing of financial derivatives. It can be seen as a vehicle that takes the SDE for X_t and determines the SDE that corresponds to $F(X_t, t)$. Financial derivatives are contracts written on underlying assets. Using Ito's lemma, the SDE for financial derivatives can be determined once the SDE for the underlying asset is given. For a market participant who wants to price a derivative asset, but who is willing to take the behaviour of the underlying as exogenous, Ito's formula is a necessary tool. In situations where the Ito formula has to be applied, in general, an SDE that drives the process X_t will also be given.

Furthermore, this can be extended to multiple dimensions. Refer to [Fri75].

[5] With this it is meant that the drift and diffusion parameters are not too irregular. Square integrability would satisfy this condition.

LEMMA 2.7 (ITO'S LEMMA IN MULTIPLE DIMENSIONS) *Let $u(z,t)$ be a continuous function in $(z,t) \in \mathbb{R}^N \times [0,\infty)$ together with its derivatives $u_t, u_{z_i}, u_{z_i z_j}$. Let $W_t, t \in \mathbb{R}^+$ be an N-dimensional Brownian motion relative to a family of information sets I_t. Let $X_t, t \in \mathbb{R}^+$ be an N-dimensional process with stochastic differential*

$$dX_t = \mu(X_t,t)dt + \sigma(X_t,t)dW_t$$

where $\mu = (\mu_1, \cdots, \mu_N)$, $\sigma = \sigma_{ij}, (1 \leq i,j \leq N)$ are respectively the N-dimensional drift and diffusion vectors. Also, suppose that the Brownian motion is correlated with correlation matrix $\rho = \rho_{ij}$. Then $u(X,t)$ (suppressing the dependence of X on t) has stochastic differential given by:

$$du_t = \left[\frac{\partial u}{\partial t} + \sum_{i=1}^{N} \mu_i \frac{\partial u}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \kappa_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j}\right] dt + \sum_{i=1}^{N} \sigma_{ii} \frac{\partial u}{\partial x_i} dW_{t_i} \tag{2.10}$$

where the matrix κ is given by $\kappa = \sigma\rho\sigma^T$ and represents the variance-covariance matrix of the process X_t.

2.5 Theoretical Models

Financial theory makes regular use of complex and sophisticated mathematical models. The basic principle underlying the pricing of financial instruments is the arbitrage pricing principle, introduced in the early 1970's by Black and Scholes [BS73] and R.C. Merton [Mer73]. In terms of impact on practice, the Black-Scholes model for option pricing is a milestone development.

There exist alternative approaches to obtaining the Black-Scholes formula, and also to pricing financial derivatives. Particularly, theories in continuous trading have been developed using probabilistic tools which do yield the Black-Scholes formula (see [Duf01], [HP81]). In the thesis, this way of thinking is referred to as the 'martingale approach'. It is also applicable in multiple dimensions.

2.5.1 Assumptions of the Model Market

Models approximate the real world and involve certain simplified assumptions about the operation of the market. The following properties are assumed for the model market in which financial derivatives will be priced. See [Duf01] and [Hul03].

1. The underlying asset is traded continuously.

2. There are no market frictions. That is, there are no margin requirements, taxes do not apply, and transaction costs and commissions do not exist.

3. All securities are perfectly divisible and any number $k \in \mathbb{R}^+$ of securities can be traded at any time.

4. The short selling of securities with full use of proceeds is permitted.

5. An investor cannot earn in excess of the risk-less rate of interest without undertaking some degree of risk. That is, there are no arbitrage opportunities.

Based upon these and other assumptions (given later as required), equations for financial derivatives can be derived. Therefore these assumptions will be referred to throughout this thesis.

2.5.2 The Geometric Brownian Motion Model

Under the assumption that log price changes within a trading day are characterised by (a) finite variance, (b) identical distributions, (c) same number of changes each day, (d) large number of changes, the central limit theorem can be used to imply that log price changes assume normal distributions in the limit. See [Tay86]. Furthermore, the variances of the distributions will scale with time. For continuous time modelling and the geometric Brownian motion model, this implies normal distributions for log price changes. Hence lognormal distribution for the security price is commonly used. An economic justification is presented in [Mer92]. The model has already been expressed in the form of the SDE (2.1). It can be written as

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2.11}$$

where $S_t, t \geq 0$ is the stochastic process representing the asset price; $W_t, t \geq 0$ is a Brownian motion relative to a family of information sets I_t; and μ, σ are positive constants known as the *drift* and *volatility* respectively. A realisation of this is shown in Figure 2.2. If S_t starts out positive it can never go negative. The simulation is started at 1.0, a non-zero value of S_t. The number of steps used is 10,000, up to a maximum time of $T = 1.0$. For this particular Matlab simulation of a one-dimensional Brownian motion using normally distributed steps, the values of $\mu = 0.0$ and $\sigma = 1.0$ were used.

The property of this random walk is seen clearly by defining $F(S_t, t) = \ln S_t$. Ito's lemma implies

$$dF(S_t, t) = \frac{\partial F}{\partial S} dS_t + \frac{\partial F}{\partial t} dt + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial S^2} dt \tag{2.12}$$

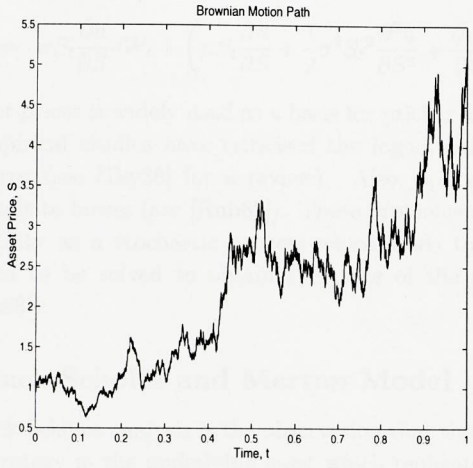


Figure 2.2: A realisation of $dS_t = \mu S_t dt + \sigma S_t dW_t$

$$= \frac{1}{S_t} dS_t + \frac{1}{2}\sigma^2 S_t^2 \left(\frac{-1}{S_t^2}\right) dt \tag{2.13}$$

$$= \frac{1}{S_t}(\mu S_t dt + \sigma S_t dW_t) - \frac{1}{2}\sigma^2 dt$$

$$\Rightarrow d \ln S_t = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t \tag{2.14}$$

The integral form of SDE (2.14) follows simply

$$\int_0^t d \ln S_s = \int_0^t (\mu - \frac{1}{2}\sigma^2)ds + \int_0^t \sigma dW_s$$

that is,

$$\ln S_t - \ln S_0 = (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t$$

So,

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \tag{2.15}$$

The SDE(2.11) will be particularly important in the modelling of many asset classes. Consider an option whose value $u(S_t, t)$ depends only on S_t and t, then from Ito's lemma it follows that

$$du = \frac{\partial u}{\partial t} dt + \frac{\partial u}{\partial S} dS_t + \frac{1}{2}\sigma^2 {S_t}^2 \frac{\partial^2 u}{\partial S^2} dt$$

$$= \sigma S_t \frac{\partial u}{\partial S} dW_t + \left(\mu S_t \frac{\partial u}{\partial S} + \frac{1}{2}\sigma^2 {S_t}^2 \frac{\partial^2 u}{\partial S^2} + \frac{\partial u}{\partial t} \right) dt \tag{2.16}$$

This model for asset prices is widely used as a basis for pricing financial derivatives, though various empirical studies have criticised the lognormality assumption for security price returns (see [Tay86] for a review). Also, option pricing based on this model often leads to biases (see [Rub85]). These problems are often addressed by modelling volatility as a stochastic process which leads to multi-dimensional equations that need to be solved to obtain the price of the particular financial derivative (see [Cla98]).

2.5.3 The Black-Scholes and Merton Model

The key to the Black-Scholes analysis is the observation that there exists a dynamic portfolio trading strategy in the underlying asset which replicates the returns from an option on that asset. Thus, to avoid the possibility of arbitrage, the value of the option must always be equal to the value of the replicating portfolio. This leads to a linear parabolic PDE for which an analytic solution exists. See [WDH93].

The replicating-portfolio approach of Black & Scholes can be generalised and applied to the pricing of general financial derivatives, contingent upon several underlying cash securities and other variables, with arbitrary payoffs. The multi-dimensional equations which these multi-factor models imply do not, in general, have readily available analytic solutions and so numerical methods have to be applied. In addition financial derivatives which permit early exercise do not have analytic solutions and numerical algorithms have to be employed to approximate the solutions to these problems.

The one-dimensional PDE for the pricing of one-dimensional options is derived using the Black-Scholes approach. The reader can refer to [BS73], [Hul03], [Nef96], [WDH93]. The 'martingale' approach is then summarised for completeness. For details on this approach, the reader can see [Duf01], [Nef96], [RB96].

The Black-Scholes Model Assumptions

The market conditions of section 2.5.1 are assumed to hold. The Black-Scholes analysis [BS73], which leads to the value of an option makes some additional assumptions.

1. The underlying asset price follows the lognormal random walk (2.11).

2. The price of the option u is a continuous, differentiable function of the asset price S_t and time t. i.e., $u \equiv u(S_t, t)$.

3. The risk-less rate of interest is a known function of time during the life of the option, investors can borrow and lend at this rate.

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

Assumption [1] can be modified. See [Jar87] and [CR85] for jump-diffusion models, and constant elasticity of variance models. In the former case, the underlying asset price random walk need not be continuous but can have random discontinuous jumps. In the latter, the volatility can be a function of S_t. [WDH93] and [Hul03] show how dividends can be easily incorporated into the model.

A portfolio is constructed with initial value Π_0 and value at time t, Π_t, which consists of one option and a number $-\Delta_t$, **delta** (as yet unspecified) of the security S_t. The value of the portfolio at time t is given by

$$\Pi_t = u(S_t, t) - \Delta_t S_t \tag{2.17}$$

The change in the value of the portfolio in time $[0, t]$ is therefore

$$\Pi_t - \Pi_0 = \int_0^t du(S_\tau, \tau) - \int_0^t \Delta_\tau dS_\tau$$

Note that Δ_t has not changed during the time-step. Ito's lemma implies that the portfolio follows the random walk

$$d\Pi_t = \frac{\partial u}{\partial t}dt + \frac{\partial u}{\partial S}dS_t + \frac{1}{2}\sigma^2 {S_t}^2 \frac{\partial^2 u}{\partial S^2}dt - \Delta_t dS_t \tag{2.18}$$

Delta Hedging

The r.h.s of equation (2.18) contains two types of terms, the deterministic and the random. The deterministic terms are those with the dt, and the random terms are those with the dS_t. These random terms are the risk in the portfolio. This risk can be reduced or even eliminated in theory (and *almost* in practice) by carefully choosing Δ_t. The random terms in (2.18) are

$$\left(\frac{\partial u}{\partial S} - \Delta_t\right) dS_t$$

By choosing

$$\Delta_t = \frac{\partial u}{\partial S} \tag{2.19}$$

the randomness is reduced to zero. This perfect elimination of risk, by exploiting correlation between two instruments (here, it is an option and its underlying) is generally called **delta hedging**. Delta hedging is an example of a *dynamic hedging strategy*. From one time-step to the next the quantity $\partial u/\partial S$ changes since it is, like u, a function of the ever-changing variables S_t and t. This means that the perfect hedge must be continuously re-balanced.

No Arbitrage

After choosing the quantity Δ_t as suggested above, the portfolio held is one whose value changes by the amount

$$d\Pi_t = \left(\frac{\partial u}{\partial t} + \frac{1}{2}\sigma^2 {S_t}^2 \frac{\partial^2 u}{\partial S^2}\right) dt \tag{2.20}$$

This change is completely risk-less. The *no-arbitrage* condition demands that this change in the portfolio value must be the same as the growth obtained if the equivalent amount of cash were put in a risk-free interest bearing account:

$$d\Pi_t = r\Pi_t dt \tag{2.21}$$

Hence, on substituting equations (2.17), (2.19) and (2.20) into (2.21), it can be shown that

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

which implies that the price of the option (suppressing the dependence of S on t) satisfies the PDE:

$$-\frac{\partial u}{\partial t} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2} + rS\frac{\partial u}{\partial S} - ru = \mathcal{L}_{BS}\, u \tag{2.22}$$

in $\Omega = \{S : S \geq 0\} \times [0, T]$ with final condition $u(S, T) = g(S)$.

The above PDE has a financial interpretation as a measure of the difference between the return on a hedged option portfolio (the first two terms) and the return on a bank deposit (the last two terms). Although this difference is identically zero for a European option, it is shown later that this is not necessarily the case for an American option (see section 2.5.5).

Finally, it is noted here that the Black-Scholes equation (2.22) does not contain the growth parameter μ. The value of an option is independent of how rapidly or

slowly an asset grows. A consequence of this is that two people may differ in their estimates for μ yet still agree on the value of an option.

2.5.4 European Call Option

Having derived the Black-Scholes equation for the value of an option, final and boundary conditions are required in order to obtain a unique solution. The European call option is considered here.

The European call option gives the owner the right to purchase the underlying asset at a fixed price $K \in \mathbb{R}^+$ on the expiry date $T > 0$. An analytic formula exists for the price of this contingent claim under the Black-Scholes model assumptions. The price of the corresponding put option, with the same strike price and expiry time, can be derived from the put-call parity result (1.11).

The final condition for any option applied at $t = T$ is $u(S,T) = g(S)$. For a call option,

$$g(S) = \max(S - K, 0) \; \forall \; S \in \mathbb{R}^+; \quad \text{and for some positive constant } K.$$

The asset-price boundary conditions are applied at zero asset price, $S = 0$ and at the final asset price, as $S \to \infty$. The SDE (2.11) satisfied by the asset price implies that if the asset price reaches zero at time t^*, then it remains zero for all $t > t^*$. Hence, the boundary condition at $S = 0$ for the call option is

$$u(0,t) = 0 \; \forall \; t \in [0,T]$$

As the asset price increases, the likelihood of exercising the call option increases. The price of the call option must tend towards the present value of the payoff in order to preclude arbitrage opportunities.

$$u(S,t) \to S - Ke^{-r(T-t)} \text{ as } S \to \infty \; \forall \; t \in [0,T]$$

When the interest rate and volatility are constant, equation (2.22) can be solved analytically (see [WDH93]). The exact solution for the European call is

$$u(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2) \tag{2.23}$$

where $N(.)$ is the cumulative distribution function for a standardised normal random variable given by

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

Here also,

$$\begin{aligned} d_1 &= \frac{\log(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ \text{and } d_2 &= \frac{\log(S/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ &= d_1 - \sigma\sqrt{T-t} \end{aligned}$$

The constraint that r and σ are constant can be dropped and more general formulae can be found.

2.5.5 American Put Option

The American put option gives the owner the right to sell the underlying asset at a fixed strike price $K \in \mathbb{R}^+$ at any time before the expiry date of the option, $T > 0$. The no-arbitrage condition demands that the price of the option is never less than the payoff value, $g(S) = \max(K - S, 0)$. This condition leads to a partial differential inequality equation satisfied by the value of the American. To derive this, a risk-less portfolio is constructed [as in (2.17) - (2.21)]. The possibility of early exercise demands that the appreciation in the value of the portfolio is bounded above by the risk-free rate of interest. The pricing function u satisfies the inequality

$$-\frac{\partial u}{\partial t} \geq \mathcal{L}_{BS}\, u \tag{2.24}$$

in $\Omega = \{S : S \geq 0\} \times [0, T]$ where $\mathcal{L}_{BS}$ is the Black-Scholes differential operator of equation (2.22).

The valuation of American options is a free boundary problem. Typically at each time t there is a value of the spot which marks the boundary between two regions; to one side one should hold the option and to the other side one should exercise it. That is, the domain Ω is demarcated into two regions $\mathcal{C}$ and $\mathcal{E}$ where

$$\mathcal{C} = \{(S,t) \in \Omega : u(S,t) > g(S)\}$$

is the *continuation* or *holding* region. Its complement is given by

$$\mathcal{E} = \{(S,t) \in \Omega : u(S,t) = g(S)\}$$

which is the *exercise* region. If the option is exercised then its value is determined by the payoff. If it is held then its value must be greater than the immediate payoff.

Since the option is either exercised or held, the pricing function satisfies the linear complementarity problem (see [Myn92], [Wil98]) given by,

$$\begin{aligned} -\frac{\partial u(S,t)}{\partial t} &\geq \mathcal{L}_{BS}\, u(S,t) \;\forall\; (S,t) \in \mathbb{R}^+ \times [0,T] \\ u(S,t) &\geq g(S) \;\forall\; (S,t) \in \mathbb{R}^+ \times [0,T] \\ \left(-\frac{\partial u}{\partial t} - \mathcal{L}_{BS}\, u\right).(u-g) &= 0 \;\forall\; (S,t) \in \mathbb{R}^+ \times [0,T] \end{aligned} \tag{2.25}$$

and final condition $u(S,T) = g(S)$. The complementarity formulation can be generalised to multi-factor financial derivatives.

The American option problem can also be formulated as a variational inequality problem and as an optimal stopping problem. A review of the American put problem is given in [Myn92] where the different formulations are discussed. Extensive analysis of the free boundary is contained in [Wil98].

2.6 Pricing Multi-Factor Derivatives

The Black-Scholes analysis can be extended to multiple dimensions (see [Duf01]). The martingale approach can also be applied in multiple dimensions and used to price a variety of financial derivatives (see [RB96]). The vanilla call and put options described thus far were assumed to depend on one underlying stochastic factor, namely the price of the underlying asset. In many instances, the pricing of financial derivatives requires the consideration of multiple stochastic factors. For instance, when pricing interest rate derivatives, two factor models are often used to describe the dynamics of the terms structure. In the pricing of options on stocks and currencies, volatility is often modelled as a stochastic variable. Two dimensional stochastic volatility problems are considered in [CP99], [JS87], [HW87], [MT90] and [HR98]. A review of two-factor models in the pricing of financial derivatives is given in [CHW85].

Multi-factor equations also arise in the pricing of certain types of path-dependent options where the state space is augmented to represent the path dependency. (Refer to [BP96], [ZFV99], [KV90], [Lev92], [RS95], [WDH93]). Average options which are included in this category of multi-factor derivatives are considered in chapter four of the thesis.

In this section, a general framework is presented for the pricing of multi-factor derivatives. The derivative is assumed to depend on a vector of factors whose

stochastic processes are assumed to satisfy known stochastic differentials [Hul03]. The market assumptions are as before in subsections 2.5.1 and 2.5.3.

2.6.1 Derivation of the General N-factor Equation

Consider a derivative security which depends on a N-vector of factors satisfying the SDE

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t \tag{2.26}$$

Here, $W_t, t \in \mathbb{R}^+$ is an N-dimensional Brownian motion relative to a family of information sets I_t. The Brownian motion is correlated with correlation matrix ρ. $\mu : \mathbb{R}^N \times [0, \infty) \to \mathbb{R}^N$ is an N-vector of functions and $\sigma : \mathbb{R}^N \times [0, \infty) \to \mathbb{R}^{N \times N}$ is a diagonal $N \times N$ matrix of functions $\{\sigma_{jj} : j = 1, \cdots, N\}$ such that μ and σ satisfy the regularity conditions which guarantee the existence of a unique solution to equation (2.26).

The derivative security under consideration has pricing function $u^{(1)}(X_t, t)$, with payoff $g(X_T)$ at time T. Further to the existing assumptions, it is assumed that there are at least N other securities (in addition to the one under consideration) whose prices depend on the N factors. A portfolio with value Π_t at time t is constructed consisting of a unit amount of the security $u^{(1)}(X_t, t)$ and ϕ_t^i amounts of each security $u^{(i)}(X_t, t)$, for $i = 2, \cdots, N + 1$ at time t. The market price of the portfolio at time t is therefore given by

$$\Pi_t = \sum_{i=1}^{N+1} \phi_t^i u^{(i)}(X_t, t)$$

where $\phi_t^1 = 1$. Ito's lemma (2.7) implies that the market price of the portfolio, Π_t has the stochastic differential

$$d\Pi_t = \left(\sum_{i=1}^{N+1} \phi_t^i \mu_i^* \right) dt + \sum_{i=1}^{N+1} \phi_t^i \sum_{j=1}^{N} \sigma_{ij}^* dW_{t_j} \tag{2.27}$$

where

$$\mu_i^* = \frac{\partial u^{(i)}}{\partial t} + \sum_{j=1}^{N} \mu_j \frac{\partial u^{(i)}}{\partial x_j} + \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} \sigma_{jj} \rho_{jk} \sigma_{kk} \frac{\partial^2 u^{(i)}}{\partial x_j \partial x_k} \text{ for } i = 1, \cdots, N + 1 \tag{2.28}$$

and

$$\sigma_{ij}^* = \sigma_{jj} \frac{\partial u^{(i)}}{\partial x_j} \text{ for } i = 1, \cdots, N + 1; \; j = 1, \cdots, N \tag{2.29}$$

At time t, the $\phi_t^i, i = 2, \cdots, N+1$ are chosen to eliminate the non-determinacy in the portfolio which implies

$$\sum_{i=1}^{N+1} \phi_t^i \sigma_{ij}^\star = 0 \text{ for } j = 1, \cdots, N \tag{2.30}$$

Since $\phi_t^1 = 1$, equation (2.30) represents N equations in N unknowns. It is assumed that the portfolio is constructed such that the coefficient matrix (of functions) is invertible, and (2.30) can be solved for the $\phi_t^i, i = 1, \cdots, N+1$.

The evolution of the portfolio is then deterministic and the no-arbitrage condition demands that the growth in the portfolio is determined by the risk-free rate of interest, r. This requires that

$$\sum_{i=1}^{N+1} \phi_t^i \mu_i^\star = r \sum_{i=1}^{N+1} \phi_t^i u^{(i)} \tag{2.31}$$

Therefore,

$$\sum_{i=1}^{N+1} \phi_t^i (\mu_i^\star - r u^{(i)}) = 0 \tag{2.32}$$

Given that the coefficient matrix is invertible and hence that $\phi_t^i \neq 0$ for at least some $i = 1, \cdots, N+1$, equations (2.30) and (2.32) are only consistent if there exist $\{\lambda_j, j = 1, \cdots, N\}$ such that

$$\mu_i^\star - r u^{(i)} = \sum_{j=1}^{N} \lambda_j \sigma_{ij}^\star \text{ for } i = 1, \cdots, N+1 \tag{2.33}$$

Substituting for $\sigma_{ij}^\star$ and $\mu_i^\star$ gives

$$\frac{\partial u^{(i)}}{\partial t} + \sum_{j=1}^{N} [\mu_j - \lambda_j \sigma_{jj}] \frac{\partial u^{(i)}}{\partial x_j} + \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} \sigma_{jj} \rho_{jk} \sigma_{kk} \frac{\partial^2 u^{(i)}}{\partial x_j \partial x_k} - r u^{(i)} = 0 \tag{2.34}$$

which is the PDE satisfied by the i^{th} security $u^{(i)}$. See [Cla98].

2.7 Alternative Approaches

The geometric Brownian motion is not always the model of choice for the behaviour of the stock price. The probability distribution may have to be modified. Alternatives to the basic concept of lognormal distributions exist. Examples include *displaced diffusion models* or *models based on binomial trees* and *lattice models.* Alternative models also include those based on martingales. These concepts are very briefly outlined below.

2.7.1 Martingale Approach

The Brownian motion in the SDE (2.11) is defined with respect to the family of information sets I_t, and the probability P. The security price process (2.15) thus has a positive drift under the measure P and hence is not a martingale with respect to that measure. Neither is the process for the discounted security price, $Z_t = e^{-rt}S_t$.

The martingale approach begins by using Girsanov's theorem (see [RY99], [RW00]) to construct a measure Q equivalent to P. This means that P and Q operate on the same sample space and have the same null sets. As such, the discounted price process Z_t has zero drift and is a martingale under Q.

Another process Y_t is constructed with the properties that it is a martingale under Q and has value equal to the payoff function at time $t = T$; $Y_t = E_Q[e^{-rT}g(S_T)|I_t]$. For examples, see [RB96]. Given that Y_t and Z_t are martingales with respect to Q, the martingale representation theorem (refer to [RW00], [Duf01]) is used to represent Y_t in terms of Z_t and a unique process ϕ_t^1. That is, $dY_t = \phi_t^1 dZ_t$.

A trading strategy is constructed

$$Y_t - Y_0 = \int_0^t \phi_\tau^1 dZ_\tau + \int_0^t r\phi_\tau^2 d\tau$$

where $\phi_t^1 = dY_t/dZ_t$ and $\phi_2^t = Y_t - \phi_t^1 Z_t$. A portfolio Π_t consisting of ϕ_t^1 units of the stock and ϕ_2^t units of an interest bearing security with process $B_t = e^{rt}$ (where r is the constant instantaneous risk-free rate of interest), is continuously re-balanced at all times $t \in [0, T]$. Therefore,

$$\Pi_t = \phi_1^t S_T + \phi_2^t B_T = B_T Y_T = g(S_T)$$

and additionally, the portfolio can be shown to be self-financing (see [RB96]). That is, all changes in the portfolio are due to capital gains and no new cash is deposited nor withdrawn in $[0, T]$, with

$$d\Pi_t = \phi_1^t dS_t + \phi_2^t dB_t$$

Since the portfolio is self-financing and its value at time T is equal to the value of the option at time T, then the value of the option u_t must equal the value of the portfolio at all other times to preclude arbitrage opportunities. That is,

$$u_t = \Pi_t = B_t Y_t = e^{-r(T-t)} E_Q[g(S_T)|I_t]$$

The value of the option at time t is given as the discounted value of the expected payoff, conditional on the information available at time t, with the expectation being evaluated under the probability measure Q. The evaluation of this expectation yields the formula (2.23) and further analysis shows that ϕ_t^1 is given by $\frac{\partial u}{\partial S}(S_t, t)$.

The expression of the price in terms of an expectation suggests that Monte Carlo techniques can be used to determine the option price numerically where expectations cannot be evaluated analytically. This numerical technique is widely used in financial derivatives and it is briefly reviewed in section 2.7.3. For detailed development of the martingale approach, refer to [Duf01], [HP81], [RB96].

In many situations, numerical methods have to be employed to obtain approximate solutions to the equations for the price of financial derivatives. The numerical methods which are generally used fall into three main categories: Monte Carlo simulation, lattice methods and methods based on numerical solution of the pricing PDE. In this thesis, methods based on the discretisation of the PDE are used. However, the alternate approaches are briefly reviewed and compared in this section.

2.7.2 Numerical Methods: Lattice Models

Consider a call option C_t written on the underlying asset S_t, expiring at time T with $t < T$. The time interval $(T - t)$ is divided into n smaller intervals, each of size Δ. Assume that during Δ the only possible changes in S_t are an *up* movement of size $\sigma\sqrt{\Delta}$ or a *down* movement of size $-\sigma\sqrt{\Delta}$:

$$S_{t+\Delta} = \begin{cases} S_t + \sigma\sqrt{\Delta} \\ S_t - \sigma\sqrt{\Delta} \end{cases} \tag{2.35}$$

Clearly, the size of the parameter σ determines how far $S_{t+\Delta}$ can wander during a time interval of length Δ, hence it is called the *volatility*. The dynamics described by equation (2.35) represents a lattice or a binomial tree. Figure 2.3 displays these dynamics in the case of multiplicative up and down movements. The binomial method and the trinomial method is outlined in detail in [Hul03] and [WDH93].

Now, given the risk-free interest rate r for the period Δ, it is known that the risk-adjusted probabilities $\tilde{P}_{up}$ and $\tilde{P}_{down}$ must satisfy

$$S_t = \frac{1}{1+r}\left[\tilde{P}_{up}(S_t + \sigma\sqrt{\Delta}) + \tilde{P}_{down}(S_t - \sigma\sqrt{\Delta})\right]$$

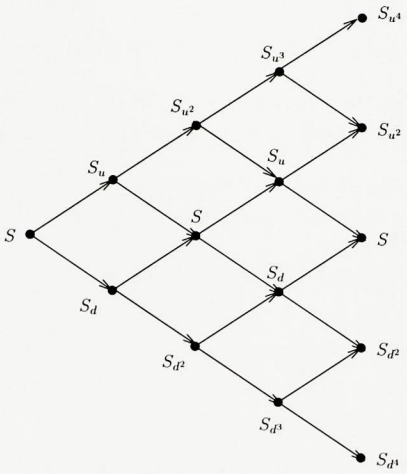


Figure 2.3: A multiplicative binomial lattice.

In this equation, r, S_t, σ, Δ are known. The only unknown is $\tilde{P}_{up}$ [6], which can be determined easily. Once this is done, the $\tilde{P}_{up}$ can be used to calculate the current arbitrage-free value of the call option. In fact, the equation

$$C_t = \frac{1}{1+r}\left[\tilde{P}_{up}C^{up}_{t+\Delta} + \tilde{P}_{down}C^{down}_{t+\Delta}\right] \tag{2.36}$$

"ties" two arbitrage-free values of the call option at any time $t+\Delta$ to the arbitrage-free value of the option as of time t. At this point, the $\tilde{P}_{up}$ is known. In order to make the equation usable, the two values $C^{up}_{t+\Delta}$ and $C^{down}_{t+\Delta}$ are needed. Given these, the value of the call option C_t at time t can indeed be calculated. Figure 2.4 shows the multiplicative lattice for the option price C_t. The arbitrage-free values of C_t are at this point indeterminate except for the expiry nodes. Given the lattice for S_t, the values of C_t can be determined from the boundary condition, which is the payoff. Once this is done, using equation (2.36) backwards, the initial node is eventually reached which gives the current value of the option.

Hence, the procedure is to use the dynamics of S_t to go forwards and determine the expiry date values of the call option. Then, using the risk-adjusted probabilities, the current value C_t is determined by working backwards with the lattice for the call option. In this procedure, Figure 2.3 gives an approximation of all possible

[6]Remember that $\tilde{P}_{down} = 1 - \tilde{P}_{up}$.

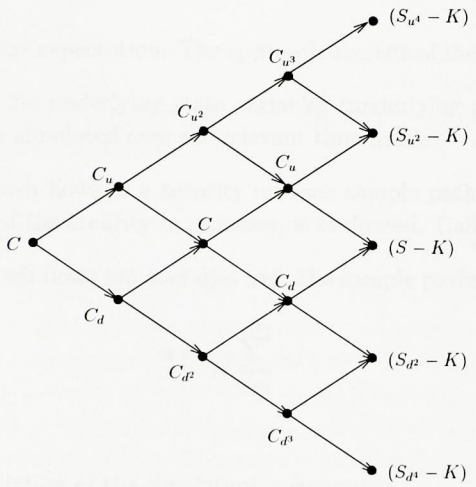


Figure 2.4: A multiplicative lattice for the option price of a call option C with payoff $\max(S - K, 0)$.

paths that S_t may take during the period $T - t$. The tree in Figure 2.4 gives an approximation of all possible paths that can be taken by the price of the call option written on S_t. If Δ is small, then the lattices will be close approximations to the true paths followed by S_t and C_t.

The binomial method is equivalent to an explicit finite difference scheme and therefore stability is only guaranteed for small values of Δ. This makes the method unattractive for longer dated contracts especially when full binary trees need to be used. In multiple dimensions, the constraints are more severe rendering the method relatively expensive.

2.7.3 Numerical Methods: Monte Carlo Simulation

The original work on Monte Carlo simulation in option pricing is [Boy77]. A detailed review of Monte Carlo methods applied to the pricing of financial derivatives is given in [BBG97]. This is one of many useful articles on Monte Carlo methodologies and application to pricing contained in [Dup98].

The Monte Carlo method is well suited once it is agreed that security pricing can be represented by expectations (see [Duf01]). Monte Carlo seeks to solve the pricing equation (2.34) by representing the solution as an expectation, then using numerical

simulation to evaluate the expectation. The approach consists of the following steps:

- Sample paths of the underlying state variables (underlying asset prices and interest rates) are simulated over the relevant time frame.
- The discounted cash flows of a security on each sample path, as determined by the structure of the security in question, is evaluated. Call this u_j.
- The discounted cash flows are averaged over the sample paths. That is,

$$\hat{u} = \frac{1}{N} \sum_{j=1}^{N} u_j$$

  is computed.
- The standard deviation of the simulation is computed,

$$\hat{\sigma}_{\hat{u}} = \sqrt{\frac{1}{N-1} \sum_{j=1}^{N} (u_j - \hat{u})^2}$$

As a simple example, refer back to Equation (2.15). Assume stock price evolves according to

$$\frac{dS_t}{S_t} = r\ dt + \sigma\ dW_t$$

By Ito's calculus,

$$S_t = S_0 e^{(r - \frac{\sigma^2}{2})t + \sigma W_t}$$

So the sample path could be generated by

$$S_{t+\Delta t} = S_t e^{(r - \frac{\sigma^2}{2})\Delta t + \sigma\sqrt{\Delta t}\ Z}$$

where Z is $N(0, 1)$.

Monte Carlo methods compute a multi-dimensional integral, the expected value of the discounted payoffs over the space of sample paths. The increase in the complexity of derivative securities in recent years has led to a need to evaluate high-dimensional integrals.

Monte Carlo methods are generally flexible and easy to implement. They can readily accommodate most forms of path-dependent options, and underlying factors can be allowed to evolve according to complex stochastic systems. Also, the computational complexity is linear in the number of state variables (unlike finite difference methods, or lattice models), and so the method is well suited to very high

dimensional problems. There are, however, difficulties in applying these methods to early-exercise options. Standard simulation programs are forward algorithms; however pricing American-style options generally requires a backward algorithm. Recent research is tackling this problem.

2.8 Conclusion

To conclude, the option pricing problem that is to be solved is usually stated in one of the three main forms: *pure diffusion*, *pure jump* or *jump-diffusion*. The Black-Scholes equation, arising from stochastic calculus is an example of the first type. A continuous motion is assumed by pure diffusion approaches such as lognormal distributions. Pure jump procedures assume a discontinuous motion, an example of this kind of procedure is the binomial tree approach. Jump-diffusion models, finally, are a mixture of the first two types.

This chapter has provided an overview of the theoretical background required (within the scope of this thesis). The remainder of the thesis focusses on the numerical approaches used to solve the pricing PDE's. A general summary of the numerical approaches pursued is as follows: The solution domain is truncated and approximate boundary conditions which admit the required solution are posed at these truncated boundaries. Orthogonal co-ordinate transformations are used to achieve detailed resolution in the areas of the computational domain which are of most interest (and consequently where the greatest accuracy is required), while coarsening the grid where there is least interest. Eulerian and Semi-Lagrange time integration schemes are used to approximate the time derivative which results in requiring discrete systems to be solved at each time-step. The results achieved are shown to be very favourable.

Chapter 3

Numerical Solutions of One-Factor Options

Depending on the nature of the boundary conditions which must be satisfied by the value of the contingent claim, the Black-Scholes PDE and its extensions may or may not have an analytic solution. In many realistic situations, analytic solutions do not exist, and the analyst must resort to other methods. Examples include numerical integration [Par77]; finite-difference approximation to the differential equation, employed extensively by Brennan and Schwartz [Bre79], [Sch77], [BS77b], [BS77a], [BS78]; and Monte Carlo methods advocated in [Boy77]. A number of analytic approximations are also available. [GJ84] proposed a compound-option analytic polynomial approximation. [BAW87] use a quadratic approximation. Others have used various problem-specific heuristic approximations. For a comprehensive review of numerical techniques, see [GS85] and [Hul03].

Throughout this study, the use of finite difference techniques is predominant. These are particularly well established when applied to one-dimensional problems. In the case of higher dimensional problems, they are well suited to those involving geometries with linear boundaries. When applied to problems in two and three dimensions involving irregular geometries, they lose their simplicity and ease of use.

The main idea in the calculus of finite differences is to replace derivatives with linear combinations of discrete function values, which are approximations based on Taylor series expansions of functions near the point or points of interest. Finite difference methods have the virtue of simplicity and they are used widely in applications because of this. Finite difference methods are very general and can be made quite robust. They are flexible, easily able to accommodate European options and American options as well as path-dependent options, and other exotics. They provide the price of the derivative at all the grid points which represent asset prices.

They also provide the parameters for hedging.

The first use of finite difference methods in finance was due to Brennan and Schwartz in 1977-78. [Sch77] proposed a model to solve for the value of a warrant or an option when a closed-form solution of the valuation equation cannot be obtained. This model is based on a difference approximation of the valuation equation and uses standard numerical methods. In [BS77a], the principles of the option pricing model are applied to the convertible bond. This paper extends the work of Black-Scholes and Merton to the pricing of convertible bonds. The differential equation and boundary conditions governing the value of the bond are derived, and the authors resort to finite differences as the numerical method to solve the differential equation. The finite difference approach of [Sch77] has been extended by [Cou82]. The same methods are used to derive a difference approximation of the solution of the valuation equation which has a greater level of accuracy than Schwartz's approximation. This is achieved by changing the point of interest about which the Taylor series expansions occur.

The implicit finite difference method relates the value of the derivative security at time t to three values at time $t + \Delta t$. The explicit finite difference method relates the value of the derivative security at time $t + \Delta t$ to three values at time t. [BS78] show that the explicit finite difference method is equivalent to a trinomial lattice approach. They also show that the implicit finite difference method corresponds to a multinomial lattice approach where, in the limit, the underlying variable can move from its value at time t to an infinity of possible values at time $t + \Delta t$.

None of the formulations of the American put problem yield analytic solutions (except for the perpetual American put [Mer73] where the option never expires). However various methods have been used to approximate the solution. In [BS77b], explicit finite differences are used to solve the linear complementarity problem (2.25). At each time-step, the scheme is applied in the continuation region $\mathcal{C}$ and the solution in the exercise region $\mathcal{E}$ is given by the payoff function. Since the regions are not known *a priori*, the boundary is updated at the end of each time-step, using projection. Explicit and implicit finite difference methods for valuing American put and call options with and without dividends are used in [GS85]. In the same paper, a comparison of different lattice and finite difference methods is also given and it is found that finite difference methods are more general and robust, and are more suitable when a large number of options needs to be priced. It is also shown that the explicit finite difference method's only disadvantage is that the numerical solution does not necessarily converge to the solution of the differential equations as the time step tends to zero (due to instabilities). A modification to the explicit finite difference method is suggested in [HW90] which ensures that, as smaller time intervals

are considered, the calculated values of the derivative security converge to the solution of the underlying differential equation. It can be used to value any derivative security dependent on a single state variable and can be extended to deal with many derivative security pricing problems where there are several state variables.

A lucid discussion of option pricing via numerical solution of PDE's is provided in [WD93]. The finite difference method is shown to be particularly effective for pricing a wide variety of exotic options. A detailed description of the finite difference method in option pricing is given in [WDH93], [WHD97], [Wil98], [TR00].

Methods based on linear programming are pursued in [DHR98]. Upon suitable domain truncation and the use of finite-difference approximations to the Black-Scholes operator, an ordinary linear program (LP) is obtained which can be solved numerically with standard techniques such as simplex or interior point algorithms. However, by exploiting the tridiagonal structure of the constraint matrices thus formed, a fast revised simplex method is derived for the solution of standard valuation problems in LP form. Specifically, this LP approach produces exotic option valuations to the accuracy of alternative methods in computing times which are at most several seconds and significantly less than alternatives such as multinomial trees, convolution methods and Monte Carlo.

The finite difference approximation technique remains the method of choice when pricing new derivative products. [AABR98] illustrates how a Crank-Nicolson finite-difference scheme is used to price a passport option. The passport option is a new contingent claim that, in effect, insures against cumulative losses resulting from repeated short-term trading activities. A passport option grants its holder the right to engage in a short/long trading strategy of his own choice, while obligating the option writer to cover any net losses on the strategy. The numerical scheme is shown to be applicable to a number of variations on the basic option contract, including American exercise and time-discrete trading strategies.

In this chapter, a method based on finite difference discretisation is used to solve the option pricing problem for a European put. Two different approaches were pursued. The study on co-ordinate transformations and non-uniform time-stepping looked at grid refinement as a means to improve numerical technique and solution accuracy. The study on Laplace transforms considered an alternative solution approach to the finite difference scheme.

The mathematical problem is described followed by numerical techniques for computation of the value of the option. The finite difference discretisation scheme is presented and a description of the resulting discrete problem follows. The algorithm with a direct solver is applied to the discrete system with emphasis placed on exploiting the special structure of the coefficient matrix. Numerical results are

then presented which suggest that solutions to desired levels of accuracy can be obtained in seconds. The hedging parameters are also discussed. The next study in this chapter looks at ways of improving the accuracy of the solutions and efficiency of the solution method. Co-ordinate transformations are applied which ensure grid refinement in regions of interest. It is shown by presenting numerical results that the required degree of accuracy can be achieved with fewer mesh points than when co-ordinate transformations are not applied. A non-uniform time-stepping algorithm is considered which uses small initial time steps and larger final time steps. This is shown to reduce the errors. Once again, an algorithm with a direct solver is applied to the discrete system with emphasis placed on exploiting the special structure of the coefficient matrix. A variety of options with different payoffs are priced using the grid refinement algorithm. To conclude this chapter a Laplace transform application to the pricing of a European put option is described. Results are compared to those produced by the direct solver algorithm and are shown to be comparable. It is suggested that the Laplace transform technique could be made use of for non-linear pricing problems.

3.1 European (Vanilla) Options

The European put option gives the owner the right to sell the underlying asset at a fixed price $K \in \mathbb{R}^+$ on the expiry date $T > 0$. An analytic formula exists for the price of this financial derivative under the Black-Scholes model assumptions [BS73]. The price of the corresponding call option (with same exercise price and expiry) can be derived by the put-call parity result, equation (1.11).

The Black-Scholes (BS) equation given by

$$\frac{\partial u(S,\tau)}{\partial \tau} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u(S,\tau)}{\partial S^2} + rS\frac{\partial u(S,\tau)}{\partial S} - ru(S,\tau) = 0;\ S \in [0,\infty];\ \tau \in [T,0] \quad (3.1)$$

is a *backward parabolic* partial differential equation, with unknown $u(S,\tau)$ representing an option price as a function of the current underlying asset value S, and time τ. This price also depends on the volatility of the underlying asset σ, the interest rate r, the time of expiry of the option T and the exercise or strike price K.

To obtain a unique solution, certain conditions must also be specified. The final conditions and boundary conditions associated with a European put option are given respectively by

$$u(S,T) \quad = \quad \max(K-S,0) \quad ; \quad S \in [0,\infty]$$

$$\begin{aligned} u(0,\tau) &= Ke^{-r(T-\tau)} \\ u(S,\tau) &\to 0 \text{ as } S \to \infty \quad ; \quad \tau \in [T,0] \end{aligned} \tag{3.2}$$

Equation (3.1) can be rewritten as a *forward parabolic* PDE by using the transformation $t = T - \tau$. This gives the equation

$$\frac{\partial u(S,t)}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u(S,t)}{\partial S^2} - rS\frac{\partial u(S,t)}{\partial S} + ru(S,t) = 0; \; S \in [0,\infty]; \; t \in [0,T] \tag{3.3}$$

with initial and boundary conditions

$$\begin{aligned} u(S,0) &= \max(K-S,0) \quad ; \quad S \in [0,\infty] \\ u(0,t) &= Ke^{-rt} \\ u(S,t) &\to 0 \text{ as } S \to \infty \quad ; \quad t \in [0,T] \end{aligned} \tag{3.4}$$

The domain is the positive real line. A closed space interval is defined, $S = [0, S_{\max}]$. In practice a finite time interval $[0,T]$ is used to work with, although T can be as large as required. The closed domain $[0, S_{\max}] \times [0,T]$ is divided by a set of lines parallel to the $S-$ and $t-axes$ to form a grid of state space (e.g, stock price or asset value) versus time. The time until maturity, T is divided into equally-spaced small intervals of length Δt, while the state space is divided into N equally-spaced subintervals of length $\Delta s = S_{\max}/N$. Let the mesh points be denoted by (S_j, t^m), where $S_j = j\ \Delta s$ and $t^m = m\ \Delta t$. Approximations of the solution are sought at these mesh points. These approximate values are denoted by

$$U_j^m \approx u(S_j, t^m) \tag{3.5}$$

The spatial derivatives in equation (3.3) are approximated by finite differences and the resulting system of equations are solved in an evolutionary manner at each time-step starting from $m = 1$. The diffusion term is approximated by second order central differences. Higher order approximations may be used as well, but they often fail to deliver improved accuracy when applied to these problems since the initial data often has a 'kink' (i.e., is C^0 continuous - see Figure 1.1). For parabolic problems, the drift term can be approximated by central differences provided the mesh Peclet number condition is satisfied. This condition ensures diagonal dominance in the coefficient matrix, therefore avoiding numerically unstable discrete systems. When the Peclet condition is violated, first-order upwind differences which take into consideration the local flow direction, are often used. This is discussed further in section 3.1.1.

The following approximations [Twi84] are used at node (j,m), where j is the integer index of the state variable corresponding to the number of space steps (each

of length Δs) and m is the integer number of time steps.

$$\begin{aligned} S^2\frac{\partial^2 u(S,t)}{\partial S^2} &\approx (S_j)^2\frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{\Delta s^2} \\ S\frac{\partial u(S,t)}{\partial S} &\approx (S_j)\frac{U_{j+1}^m - U_{j-1}^m}{2(\Delta s)} \\ u(S,t) &\approx U_j^m; \qquad\qquad j = 1,2,\cdots,N-1\ ; \quad m = 1,2,\cdots \end{aligned} \tag{3.6}$$

Discretisation of the spatial derivatives results in a semi-discrete problem of the form

$$\frac{\partial u(S,t)}{\partial t} = \mathcal{L}_j U_j \tag{3.7}$$

where U_j represents the discrete solution and $\mathcal{L}_j$ represents the discrete operator. This is integrated between two successive time levels at a particular space value, say $S = S_j$,

$$\int_{t^m}^{t^{m+1}} \frac{\partial u(S,t)}{\partial t} dt = U(S_j, t^{m+1}) - U(S_j, t^m) = \int_{t^m}^{t^{m+1}} \mathcal{L}_j U(S_j, t) dt \tag{3.8}$$

The right hand integral is approximated using simple quadrature,

$$\int_{t^m}^{t^{m+1}} \mathcal{L}_j U(S_j,t)dt \approx \Delta t[\theta\mathcal{L}_j U(S_j,t^{m+1}) + (1-\theta)\mathcal{L}_j U(S_j,t^m)]\ ;\ j = 0,1,\cdots,N\ ;\ m = 1,2,\cdots \tag{3.9}$$

where θ is a weighting factor. The quadrature error $E = \mathcal{O}\left((\theta - 0.5)\Delta t + \Delta t^2\right)$, for $0 < \theta < 1$ and $\Delta t = t^{m+1} - t^m$. In particular the cases $\theta = 0$ and $\theta = 1$ correspond respectively to the explicit and fully implicit case. In both of these cases, the accuracy is $\mathcal{O}(\Delta s^2, \Delta t)$. In the case $\theta = \frac{1}{2}$, the Crank-Nicolson method is obtained. The advantage of this method is that its accuracy is $\mathcal{O}(\Delta s^2, \Delta t^2)$. For our purposes a slight variation is used, namely $\theta = \frac{1}{2} + \epsilon$, where $\epsilon = 0.01 \times \Delta s$. This choice of θ follows the arguments in [MM94] which established a maximum principle and convergence for the case $\theta > \frac{1}{2}$ as long as Δt is of the same order as Δs. Although this choice of θ provides a theoretical gain even if not a practical gain, it does give the algorithm an extra degree of safety.

The stencil is the pattern of local mesh connections resulting from each finite difference equation. In this case it is a six-point stencil. A set of approximate discrete prices $\{U_j^m\}$ are made use of to calculate $\mathcal{L}_j U(S_j, t)$ at each time level.

$$\mathcal{L}_j U_j^m = \frac{1}{2}\sigma^2 (S_j)^2 \frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{\Delta s^2} + r(S_j)\frac{U_{j+1}^m - U_{j-1}^m}{2(\Delta s)} - rU_j^m$$

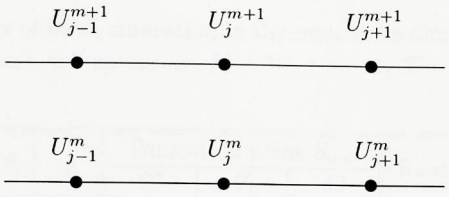


Figure 3.1: The six-point stencil.

and similarly for $\mathcal{L}_j U_j^{m+1}$, giving the $\theta-$method finite difference approximation to equation (3.9)

$$U_j^{m+1} - U_j^m = \Delta t[\theta \mathcal{L}_j U(S_j, t^{m+1}) + (1-\theta)\mathcal{L}_j U(S_j, t^m)] \tag{3.10}$$

which can be written out as

$$\begin{aligned} a_j U_{j-1}^{m+1} + b_j U_j^{m+1} + c_j U_{j+1}^{m+1} &= A_j U_{j-1}^m + B_j U_j^m + C_j U_{j+1}^m \ ; \\ & \quad j = 1, 2, \cdots, N-1; \ m = 1, 2, \cdots \end{aligned} \tag{3.11}$$

where

$$\begin{aligned} a_j &= -\frac{(\Delta t \cdot \theta)\ 0.5\ \sigma^2 (S_j)^2}{\Delta s^2} + \frac{(\Delta t \cdot \theta)\ r\ S_j}{2\Delta s} \\ b_j &= 1 + \frac{(\Delta t \cdot \theta)\ \sigma^2 (S_j)^2}{\Delta s^2} + (\Delta t \cdot \theta)\ r \\ c_j &= -\frac{(\Delta t \cdot \theta)\ 0.5\ \sigma^2 (S_j)^2}{\Delta s^2} - \frac{(\Delta t \cdot \theta)\ r\ S_j}{2\Delta s} \end{aligned}$$

and

$$\begin{aligned} A_j &= \frac{\Delta t \cdot (1-\theta)\ 0.5\ \sigma^2 (S_j)^2}{\Delta s^2} - \frac{\Delta t \cdot (1-\theta)\ r\ S_j}{2\Delta s} \\ B_j &= 1 - \frac{\Delta t \cdot (1-\theta)\ \sigma^2\ (S_j)^2}{\Delta s^2} - \Delta t \cdot (1-\theta)\ r \\ C_j &= \frac{\Delta t \cdot (1-\theta)\ 0.5\ \sigma^2 (S_j)^2}{\Delta s^2} + \frac{\Delta t \cdot (1-\theta)\ r\ S_j}{2\Delta s} \end{aligned} \tag{3.12}$$

If $a_j < 0$, $b_j \geq |a_j| + |c_j|$, $c_j < 0$, and for at least one j, $b_j > |a_j| + |c_j|$, then the system is positive definite and non-singular [MM94], so it can be solved either by factorisation or relaxation techniques.

Table 3.1: The effect of mesh truncation in the asset price direction. Model parameters for this European put option are $K = 10$, $r = 0.1$, $T = 4$ months. Grid size is 80.

S	σ	Truncation point $S_{\max}$			Exact
		20	40	80	
7	0.2	2.6731	2.6733	2.6735	2.6729
	0.45	2.7807	2.7808	2.7808	2.7809
8	0.2	1.6931	1.6932	1.6932	1.6934
	0.45	1.9779	1.9778	1.9776	1.9806
9	0.2	0.8414	0.8413	0.84	0.8468
	0.45	1.3336	1.3333	1.3339	1.3377
10	0.2	0.3049	0.3042	0.3042	0.3077
	0.45	0.8599	0.8593	0.8591	0.861
11	0.2	0.078	0.0783	0.0783	0.0794
	0.45	0.5308	0.5309	0.5308	0.5318
12	0.2	0.0142	0.0141	0.0143	0.0149
	0.45	0.3167	0.3164	0.3163	0.3174

Boundary Conditions

The option payoff provides the initial value. For a vanilla European put option,

$$U_j^1 = \max(K - S, 0) \quad ; \quad j = 0, 1, \cdots, N$$

is easily represented by a mesh function, although placing a mesh point exactly at the strike improves accuracy.

The option value at $S = 0$ is the discounted payoff Ke^{-rt}. Also, at $S = 0$, $\mathcal{L}_j U = -rU$ and the finite difference value of the option is given by

$$U_0^{m+1} = U_0^m - r\Delta t[\theta U_0^{m+1} + (1 - \theta)U_0^m)]$$

Truncation at some cutoff value of $S_{\max} \gg K$ is justified since prices far beyond the strike have less influence on prices around the strike. Tables 3.2 and 3.1 show the effect of mesh truncation for two different examples. It is evident that truncation has minimal effect on the solution in the regions of interest, provided that the truncations take place far enough away from the region of activity. The boundary condition at infinity is applied at this truncation point,

$$U_{S_{\max}}^m = 0$$

Table 3.2: The effect of mesh truncation in the asset price direction. Model parameters for this European call option are $K = 40$, $r = 0.1$, $T = 0.25$. Grid size is 100.

S	σ	Truncation point $S_{\max}$				Exact
		85	100	125	150	
35	0.2	0.2652	0.2648	0.2647	0.2641	0.2664
	0.4	1.3743	1.3747	1.3741	1.3739	1.3793
40	0.2	2.1209	2.1213	2.1220	2.1224	2.1321
	0.4	3.6802	3.6806	3.6811	3.6812	3.6867
45	0.2	6.1414	6.1412	6.1404	6.1403	6.1445
	0.4	7.1698	7.1701	7.1704	7.1708	7.1747

or a smoother solution is obtained if the following condition is used,

$$\frac{\partial^2 U^m_{S_{\max}}}{\partial S^2} = 0 \tag{3.13}$$

In the semi-implicit case this condition is incorporated into the matrix coefficients (3.11). The simplest way to do this is to modify them directly at the final spatial point,

$$\begin{aligned} a_N &\leftarrow a_N - c_N & A_N &\leftarrow A_N - C_N \\ b_N &\leftarrow b_N + 2c_N & B_N &\leftarrow B_N + 2C_N \\ c_N &\leftarrow 0 & C_N &\leftarrow 0 \end{aligned} \tag{3.14}$$

Errors, Stability and Convergence

The discrete finite difference error

$$e^m_j = u(S_j, t^m) - U^m_j$$

satisfies

$$e^{m+1}_j - e^m_j = \Delta t[\theta \mathcal{L}_j e(S_j, t^{m+1}) + (1-\theta)\mathcal{L}_j e(S_j, t^m)] + \mathrm{T}^m_j$$

which can be derived from equations (3.8) and (3.10). In deriving the local truncation errors T_j, the Taylor series is expanded about the point $t^{m+\frac{1}{2}}$. The truncation error term,

$$\mathrm{T}^m_j = \mathcal{O}\left((\theta - 0.5)\Delta t + \Delta t^2\right) + \mathcal{O}(\Delta s^2)$$

combines the Taylor series truncation terms and the quadrature errors.

The errors $\{e_j^m\}$ are initially zero. Under certain stability conditions the truncation terms T will only generate bounded error growth which will tend to zero as Δt and Δs tend to zero. These stability conditions are

$$\Delta t < \frac{\Delta s^2}{(1-2\theta)\sigma^2 S^2} \text{ if } \theta < \frac{1}{2}$$

$$\Delta s < \frac{\sigma^2 S}{r}$$

The first condition on the size of time-step does not apply when $\theta \geq \frac{1}{2}$. The second condition is too severe a constraint as $S \to 0$. It can be avoided by using a first-order difference approximation to the drift term. Violation of the second condition leads to spurious oscillations.

Under more restrictive conditions on the time-step it can be proved that the maximum $\{e_j^m\}$ is bounded by the maximum T_j^m.

Boundary Condition Stability

Stability proofs using Fourier methods do not deal with the boundary truncation condition (3.13). Although there are no reported instabilities, this truncation condition might have been expected to be unstable. The Black-Scholes equation reduces to the hyperbolic PDE

$$\frac{\partial u}{\partial t} = rS\frac{\partial u}{\partial S} - ru$$

and the central difference approximation to $\frac{\partial u}{\partial S}$ becomes a one-sided difference approximation.

If $r > 0$ then the information flow is in to the domain of computation and imposing an asymptotic value is technically required.

3.1.1 Adaptive Upwind Differencing

The drift term has a direction associated with it: as t moves away from expiry (t is decreasing), the drift moves towards smaller S. The algorithm changes from a central-difference to an upwind difference at mesh points where the stability criterion for the Peclet number is not met. The mesh Peclet number, Pe_S in the S-direction is defined by the ratio

$$\frac{|\text{coefficient of the first derivative } \frac{\partial U}{\partial S}|\ \Delta S}{|\text{coefficient of the second derivative } \frac{\partial^2 U}{\partial S^2}|}$$

It measures the ratio of the convective to diffusive fluxes in a particular direction. The Peclet condition,

$$Pe_S < 2$$

needs to be satisfied to ensure diagonal dominance of the coefficient matrix. If the condition is violated, numerical instabilities are introduced and the solution exhibits oscillations. When the condition is violated, the unconditionally stable first order upwind scheme, which takes into account the direction of flow of information, is applied.

$$\begin{aligned} \text{Use } \frac{\partial U(S,t)}{\partial S} &= \frac{U_{j+1}^m - U_{j-1}^m}{2\Delta s} \text{ if } \Delta s < \frac{\sigma^2 S}{r} \\ \text{Use } \frac{\partial U(S,t)}{\partial S} &= \frac{U_{j+1}^m - U_j^m}{\Delta s} \text{ if } \Delta s \geq \frac{\sigma^2 S}{r} \end{aligned} \tag{3.15}$$

Log normality implies that the risk tends to zero as the asset price tends to zero. The Black-Scholes PDE is drift dominated in this region so the difference approximation to the drift term may need to be modified to remove the second stability condition *where it applies*. Adaptive upwind differencing also restores the positive definiteness of the system, which is essential for efficient solution of the pricing equations.

3.1.2 θ-Method Time Stepping

Solving the equation (3.10) for the case $\theta = 0$ is special since the equations for the discrete prices are trivial to solve.

$$U_j^{m+1} - U_j^m = \Delta t \mathcal{L}_j U(S_j, t^m)$$

Simply by starting with the initial payoff and stepping gradually to the desired time renders the solution.

For $\theta > 0$, a system of equations needs to be solved. Since there are $N+1$ spatial mesh points $\{S_0, S_1, \ldots, S_N\}$, the discrete prices at any given time can be written as a vector,

$$\mathbf{U}^{m+1} = [U_0^{m+1}, U_1^{m+1}, \ldots, U_{N-1}^{m+1}, U_N^{m+1}]$$

The finite difference equations, plus the boundary conditions can then be written as follows

$$M\ \mathbf{U}^{m+1} = \mathbf{d} \tag{3.16}$$

where M is a $\{N+1\ \times\ N+1\}$ matrix with three non-zero diagonals. Referring to

equation (3.11), this can be assembled as follows:

$$\begin{bmatrix} b_0 & c_0 & & & & \\ a_1 & b_1 & c_1 & & & \\ & a_2 & b_2 & c_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{N-1} & b_{N-1} & c_{N-1} \\ & & & & a_N & b_N \end{bmatrix} \begin{bmatrix} U_0^{m+1} \\ U_1^{m+1} \\ U_2^{m+1} \\ \vdots \\ U_{N-1}^{m+1} \\ U_N^{m+1} \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_{N-1} \\ d_N \end{bmatrix}$$

The right hand side values, d_j are defined by

$$d_j^m = A_j U_{j-1}^m + B_j U_j^m + C_j U_{j+1}^m$$

once again referring to equation (3.11).

Thomas Algorithm Implementation

The system of equations given by (3.16) is a tridiagonal system. If the system of equations to be solved is very large, the matrix inversion method may use excessive computer time. Instead, the Thomas algorithm could be put to use. Rewrite the equations as

$$a_j U_{j-1} + b_j U_j + c_j U_{j+1} = d_j \qquad j = 1, 2, \ldots, N \tag{3.17}$$

where the superscript $m + 1$ has been eliminated, and where $c_N = 0$.
Let C be the set of solutions of the difference equation (3.17). The Thomas algorithm seeks two sets of quantities, e_j and f_j, such that for any member of the set C,

$$U_j = e_j U_{j+1} + f_j \tag{3.18}$$

In particular, if this is to be true for any member of the family of solutions, the boundary condition at $S = 0$ ensures that

$$e_0 = 0, \quad f_0 = Ke^{-rt}$$

Rewriting (3.18) as

$$U_{j-1} = e_{j-1} U_j + f_{j-1}$$

and substituting back into (3.17), what results is a relation between U_j and U_{j+1} which can be written as

$$U_j = \frac{-c_j}{b_j + a_j e_{j-1}} U_{j+1} + \frac{d_j - a_j f_{j-1}}{b_j + a_j e_{j-1}} ; \qquad j = 1, 2, \ldots, N.$$

Hence, given values for e_0 and f_0, it follows that

$$\begin{aligned} e_j &= \frac{-c_j}{b_j + a_j e_{j-1}}; \qquad j \geq 1 \\ f_j &= \frac{d_j - a_j f_{j-1}}{b_j + a_j e_{j-1}}; \qquad j \geq 1 \end{aligned}$$

which can be calculated inductively in order of increasing j. At the final index, this implies

$$U_N = e_N U_{N+1} + f_N$$

which is agreeable since

$$c_N = 0 \Rightarrow e_N = 0$$

This method is computationally efficient since it involves solving a general tridiagonal matrix equation by storing only the nonzero diagonal, sub- and super-diagonal elements. Also it involves explicit calculation only of the quantities e_j and f_j in order to find the solution quantities U_j.

3.1.3 Hedging Parameters

The sensitivity of the option price to the parameters involved in the governing pricing equation are collectively known as the Greeks (see [Wil98]). The simplest way of computing the Greeks is to use a finite difference discretisation to approximate the partial derivatives that constitute them.

The **delta**, given by

$$\Delta = \frac{\partial u}{\partial S} \approx \frac{U_{j+1}^m - U_{j-1}^m}{2\Delta s} \tag{3.19}$$

is the rate of change of the value of the option, or portfolio of options, with respect to the underlying asset S. It is of fundamental importance in both theory and practice. It is a measure of correlation between the movements of the option or options and those of the underlying.

The **gamma**, defined by

$$\Gamma = \frac{\partial^2 u}{\partial S^2} \approx \frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{(\Delta s)^2} \tag{3.20}$$

is the diffusion.

Since the finite difference calculations generate option values on the entire grid of S values, the calculations of Δ and Γ in this manner come at negligible additional computational cost.

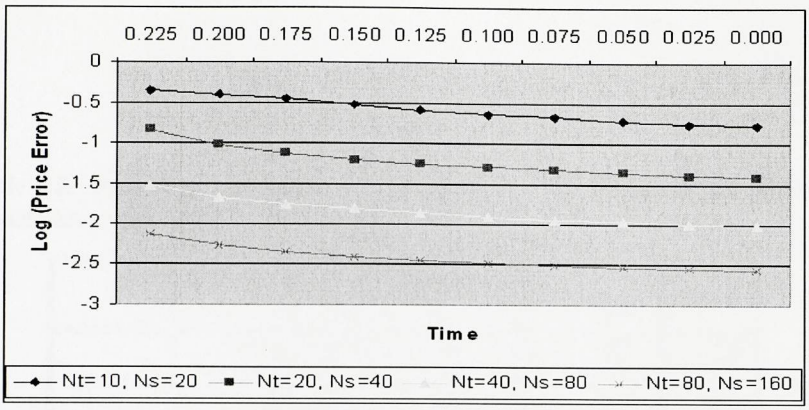


Figure 3.2: Log plot of l_∞ error per time-step.

3.1.4 Results and Discussions

The single-factor European call and put options have analytic solutions available, which are treated as the exact solution. The desired level of accuracy is taken to be 0.1%. The numerical algorithm is tested by comparing with the exact solution in each case. The algorithm for the exact solution has been incorporated into the option pricing programme. It uses an approximation to the Normal Cumulative function, accurate to eight digits.

The convergence behaviour of the method described in this chapter is displayed in Tables 3.3 and 3.4. Results corresponding to a rapidly varying region of the solution are displayed, as the grid is refined. The results suggest that the numerical scheme is convergent. Furthermore, the desired level of accuracy can be obtained on a solution grid of mesh size 160.

The maximum error at every time-step is assumed to occur at the strike, as that is the point in the payoff where the singularity lies. The l_∞ error is therefore evaluated at the strike. In Figure 3.2 a log plot of the maximum error in the price per time-step (evaluated at the strike) is shown. The parameters of this European put option are $\sigma = 0.4$, $r = 0.05$, $K = 40$, $S_{\max} = 100$, $T = 0.25$. The term *Price Error* in the figure is the error (|calculated solution - analytic solution|) for the price

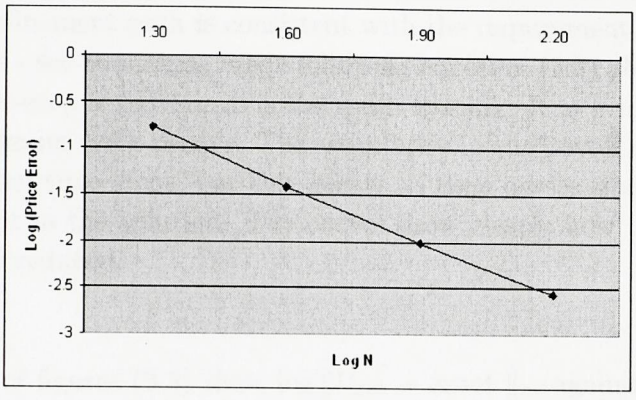
Table 3.3: Numerical convergence of the solution to the European put problem. Model parameters are $K = 10$, $r = 0.1$.

T	S	σ	Solution Grid				Exact
			20	40	80	160	
4 months	7	0.20	2.6752	2.6736	2.6731	2.673	2.6729
		0.45	2.7784	2.7802	2.7807	2.7809	2.7809
	8	0.20	1.6994	1.6951	1.6938	1.6935	1.6935
		0.45	1.9696	1.9778	1.9799	1.9804	1.9806
	9	0.20	0.83	0.8413	0.8454	0.8464	0.8468
		0.45	1.3198	1.3333	1.3365	1.3374	1.3376
	10	0.20	0.2542	0.2962	0.3049	0.307	0.3076
		0.45	0.8419	0.8563	0.8598	0.8607	0.861
	11	0.20	0.0601	0.0739	0.078	0.079	0.0794
		0.45	0.5158	0.5278	0.5308	0.5315	0.5318
	12	0.20	0.0123	0.0141	0.0147	0.0149	0.0149
		0.45	0.3063	0.3146	0.3167	0.3172	0.3174
8 months	7	0.20	2.3815	2.3768	2.3755	2.3751	2.375
		0.45	2.7088	2.7139	2.7152	2.7156	2.7157
	8	0.20	1.4809	1.4814	1.4819	1.4821	1.4822
		0.45	2.0429	2.0511	2.0531	2.0537	2.0538
	9	0.20	0.7662	0.785	0.7898	0.791	0.7914
		0.45	1.5096	1.5194	1.5218	1.5224	1.5226
	10	0.20	0.3243	0.3509	0.3569	0.3584	0.3589
		0.45	1.0981	1.1078	1.1102	1.1108	1.111
	11	0.20	0.1189	0.1346	0.1386	0.1395	0.1399
		0.45	0.7892	0.7979	0.8	0.8005	0.8009
	12	0.20	0.0397	0.0456	0.0472	0.0477	0.0478
		0.45	0.5622	0.5693	0.5711	0.5715	0.5721

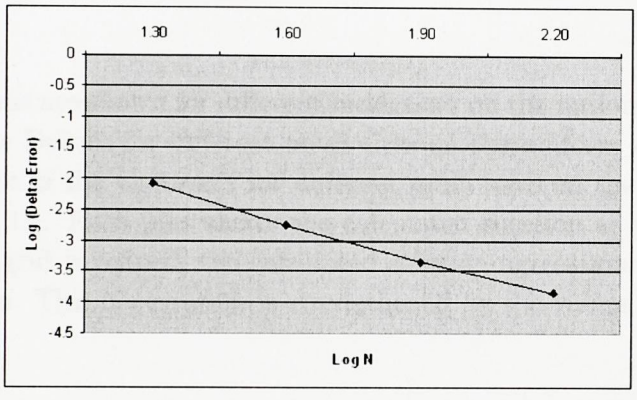
Table 3.4: Numerical convergence of the solution to the European call problem. Model parameters are $K = 40,\ r = 0.1$.

T	S	σ	Solution Grid				Exact
			20	40	80	160	
0.25	35	0.20	0.2588	0.2546	0.2583	0.2601	0.2608
		0.45	1.5744	1.6572	1.6779	1.6831	1.6848
	40	0.20	1.7325	2.0307	2.098	2.1132	2.1181
		0.45	3.9088	4.0197	4.0456	4.0521	4.0542
	45	0.20	6.0828	6.1137	6.1268	6.1306	6.1319
		0.45	7.3826	7.4597	7.4786	7.4834	7.485
0.5	35	0.20	0.8034	0.8551	0.8754	0.8808	0.8826
		0.45	3.1585	3.2268	3.2436	3.2478	3.2492
	40	0.20	3.0344	3.2521	3.2968	3.3076	3.3111
		0.45	5.8721	5.9449	5.9627	5.9672	5.9687
	45	0.20	7.1927	7.2608	7.281	7.2861	7.2878
		0.45	9.3502	9.4099	9.4247	9.4284	9.4297
1.0	35	0.20	2.1456	2.2441	2.2698	2.2762	2.2784
		0.45	5.6918	5.741	5.7533	5.7564	5.7575
	40	0.20	5.1283	5.2664	5.2976	5.3053	5.3079
		0.45	8.7844	8.8327	8.8448	8.8479	8.8491
	45	0.20	9.267	9.3448	9.3643	9.3691	9.3708
		0.45	12.3855	12.4279	12.4386	12.4413	12.4426

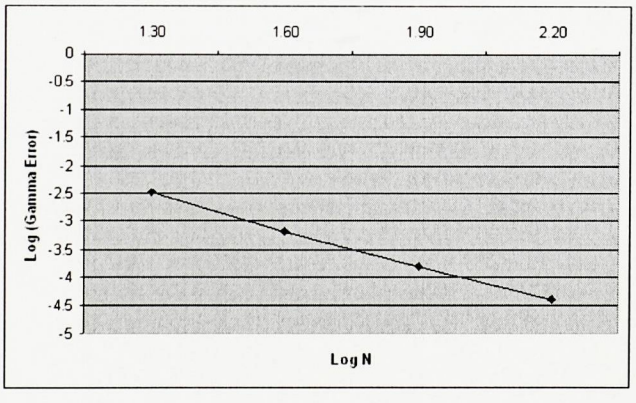
Figure 3.3: Convergence Plots for Uniform Meshing.



(a) In the Price.



(b) In the Delta.



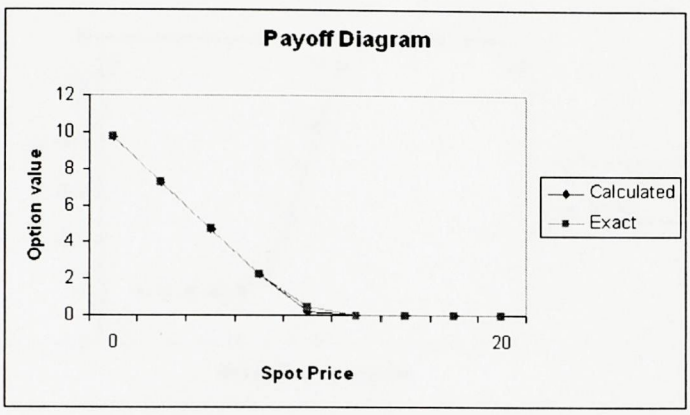
(c) In the Gamma.

values evaluated at the strike. This is plotted against time. The ratio of $\frac{\Delta t}{\Delta s}$ is kept constant. This refinement path is consistent with the requirement that Δt is of the same order as Δs - see comments made following equation (3.9) about the choice of θ. When the time-step is halved, so is the mesh spacing. It is evident that as time evolves, the maximum error decays. The singularity in the payoff causes the higher values at the initial time steps, but this decays as time passes since the singularity gets smoothed out in the solution. The curves show clearly how the error behaves as the grid size is reduced.

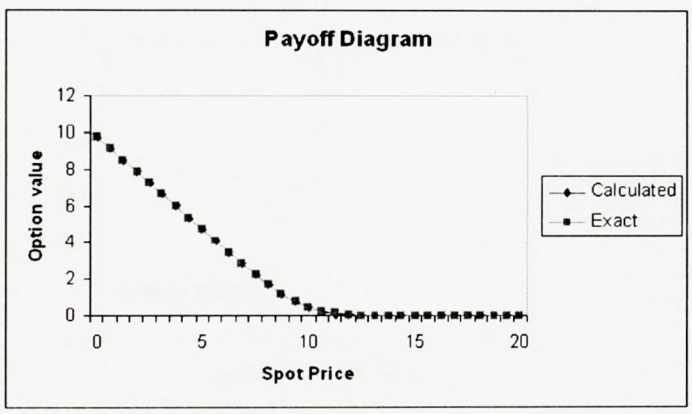
The next set of figures (3.3) show $\log \|U_{2kh} - \text{exact}\|_\infty$ against $\log(N)$ for $k = 1, 2, 3, 4$ representing the solutions calculated on mesh sizes of $N = 20, 40, 80, 160$ at time= 0. Similar log plots for the delta and gamma values are also shown. The plots suggest second order convergence.

Payoff diagrams are shown for different mesh sizes on the uniform mesh in Figure 3.4. A plot of the Delta's for different mesh sizes on the uniform mesh is shown in Figure 3.5. A plot of the Gamma's for different mesh sizes on the uniform mesh is shown in Figure 3.6. Each plot shows the calculated solution as well as the exact solution. As the grid is refined, the calculated solution agrees more and more with the exact solution. This observation is strengthened by the values in Table 3.3 and Table 3.4.

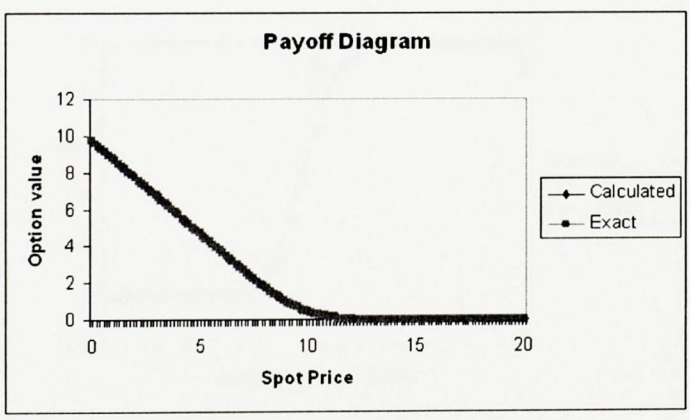
Figure 3.4: European put option plots as a function of S for several mesh sizes. Model parameters are $K = 10,\ r = 0.1,\ \sigma = 0.2,\ T = 6$ months.



(a) Mesh spacing $= \frac{1}{2^3}$.

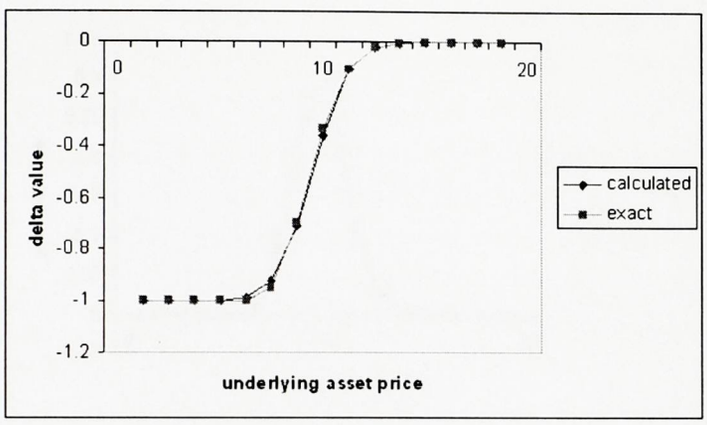


(b) Mesh spacing $= \frac{1}{2^5}$.

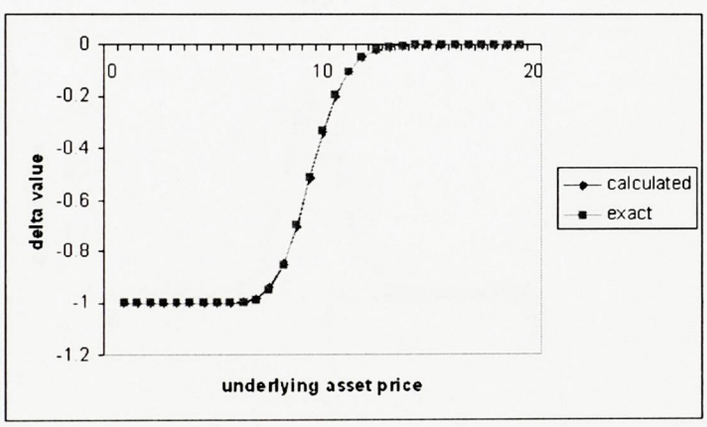


(c) Mesh spacing $= \frac{1}{2^7}$.

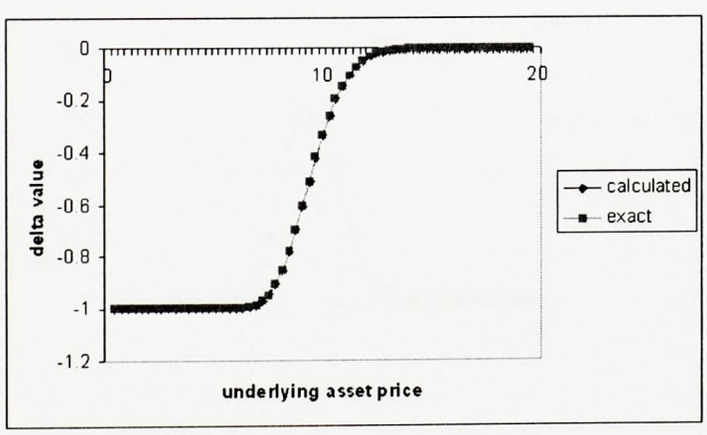
Figure 3.5: European put option Delta plots as a function of S for several mesh sizes. Model parameters are $K = 10$, $r = 0.1$, $\sigma = 0.2$, $T = 6$ months.



(a) Mesh spacing = $\frac{1}{2^4}$.

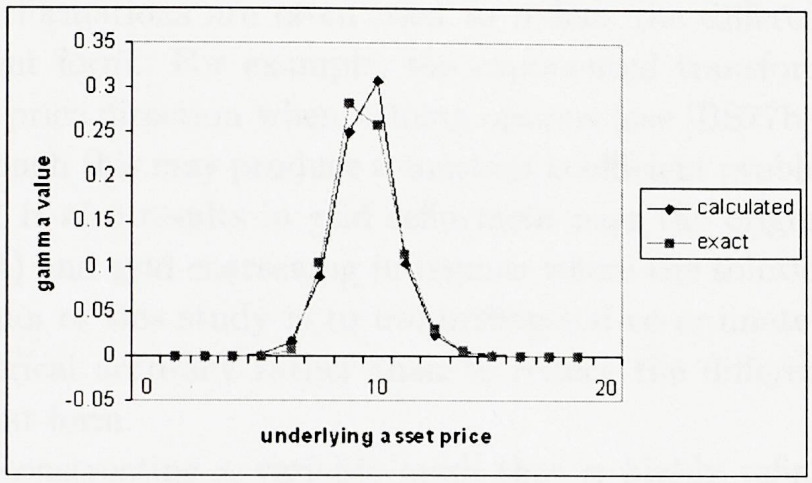


(b) Mesh spacing = $\frac{1}{2^5}$.

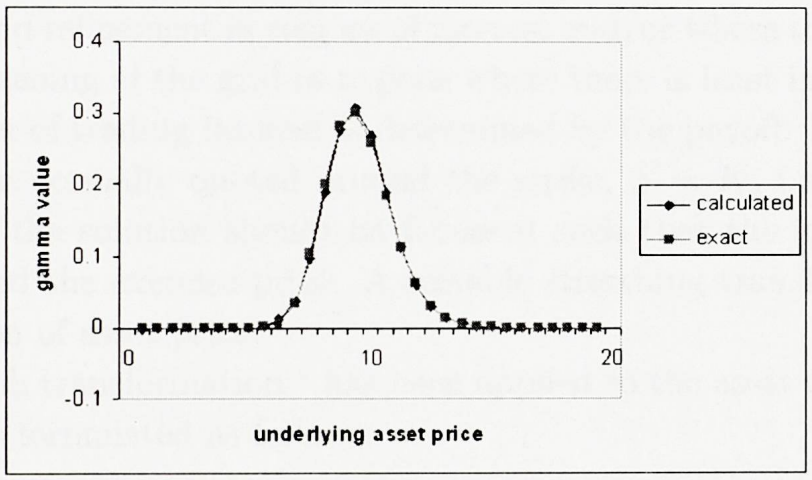


(c) Mesh spacing = $\frac{1}{2^6}$.

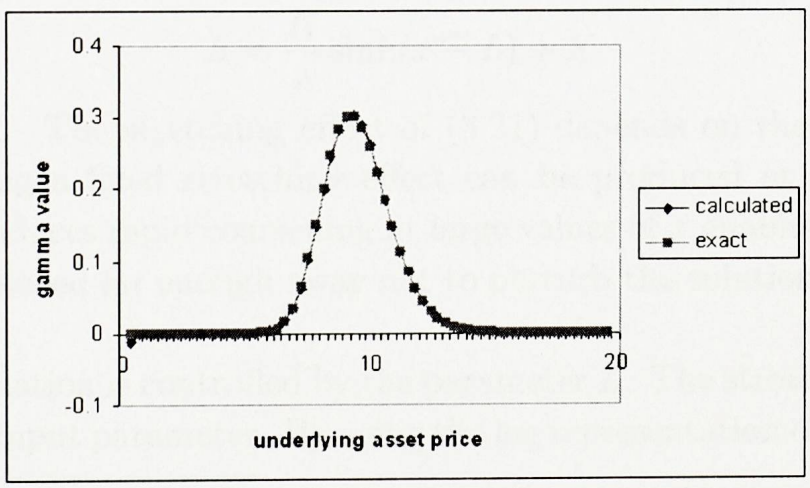
Figure 3.6: European put option Gamma plots as a function of S for several mesh sizes. Model parameters are $K = 10,\ r = 0.1,\ \sigma = 0.2,\ T = 6$ months.



(a) Mesh spacing $= \frac{1}{2^4}$.



(b) Mesh spacing $= \frac{1}{2^5}$.



(c) Mesh spacing $= \frac{1}{2^6}$.

3.2 Co-ordinate Transformation

Co-ordinate transformations are often used to reduce the differential operator to constant coefficient form. For example, the exponential transformation has been used in the asset price direction when valuing options (see [BS77b], [BS78], [GS85], [WDH93]). Although this may produce a constant coefficient problem and a simpler algebraic system, it also results in grid refinement near the origin (where there is not much interest) and grid coarsening in regions where the solution varies rapidly. In contrast the aim of this study is to use orthogonal co-ordinate transformations to improve numerical accuracy rather than to reduce the differential operator to constant coefficient form.

Rather than constructing a variable mesh that is highly refined in the region of interest, the use of co-ordinate transformations has been chosen to achieve the same effect. A co-ordinate transformation is chosen so that a uniformly spaced grid in the transformed space is mapped onto a grid in the original co-ordinate system which achieves grid refinement in regions of interest and/or where the solution varies rapidly, and coarsening of the grid in regions where there is least interest. For most options the region of trading interest is determined by the payoff. For a vanilla call or put, prices are normally quoted around the strike, $S = K$, i.e. at the money. In this instance, the solution should be focussed such that the greatest accuracy is provided around the exercise price. A suitable stretching transformation should enlarge this region of asset price.

An inverse-sinh transformation [1] has been applied to the asset price. The transformation used is formulated as follows:

$$x = \sinh^{-1}\left(\frac{\lambda}{K}(S - K)\right) + L \tag{3.21}$$

or, inversely,

$$S = \frac{K}{\lambda}\sinh(x - L) + K$$

where $\lambda = \sinh L$. The stretching effect of (3.21) depends on the value of K, but by suitable scaling a fixed stretching effect can be produced at any strike. The transformation induces rapid coarsening at large values of x enabling the truncation boundary to be placed far enough away not to perturb the solution for only a small increase in cost.

The transformation is controlled by the parameter L. The stretch $\tau = \cosh L$ is a more convenient input parameter. By using the log representation of the inverse cosh

[1] Andrew P. Weir carried out early work on the co-ordinate stretching transformation for two-factor European options while at Oxford University Computing Laboratory.

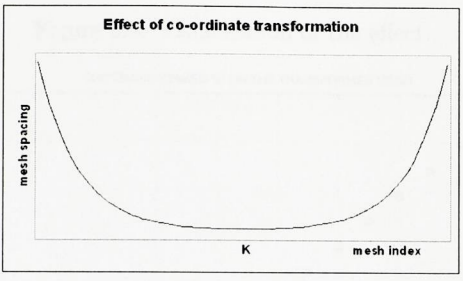


Figure 3.7: Grid refinement occurs in the region of interest (around the strike).

function, $L = \ln[\tau + \sqrt{(\tau^2 - 1)}]$. To simplify the notation, $\lambda = \sinh L = \sqrt{(\tau^2 - 1)}$ is introduced. A visualisation of the effects of the co-ordinate transformations are shown in Figure 3.7.
From equation (3.21),

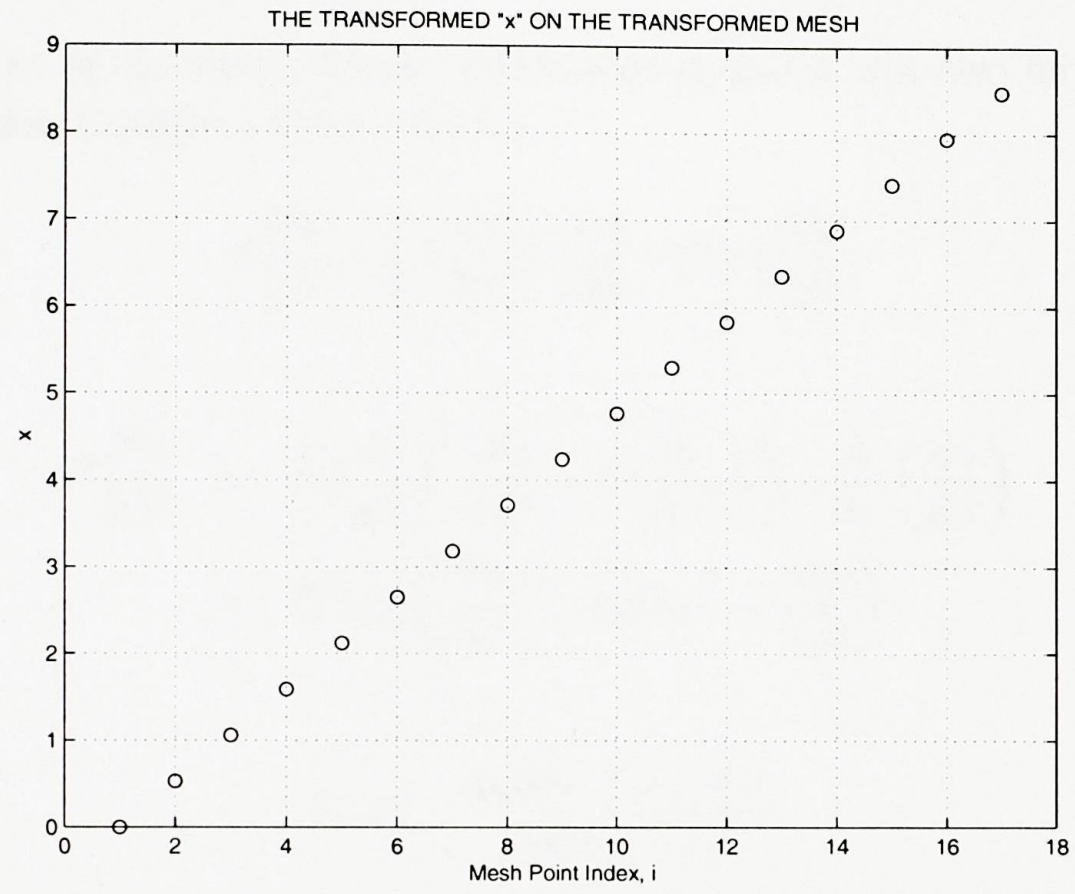
$$x = \sinh^{-1}(\lambda(S' - 1)) + \sinh^{-1}(\lambda); \qquad \text{where} \quad S' = \frac{S}{K} \tag{3.22}$$

Retrieving the values of S' from equation (3.22),

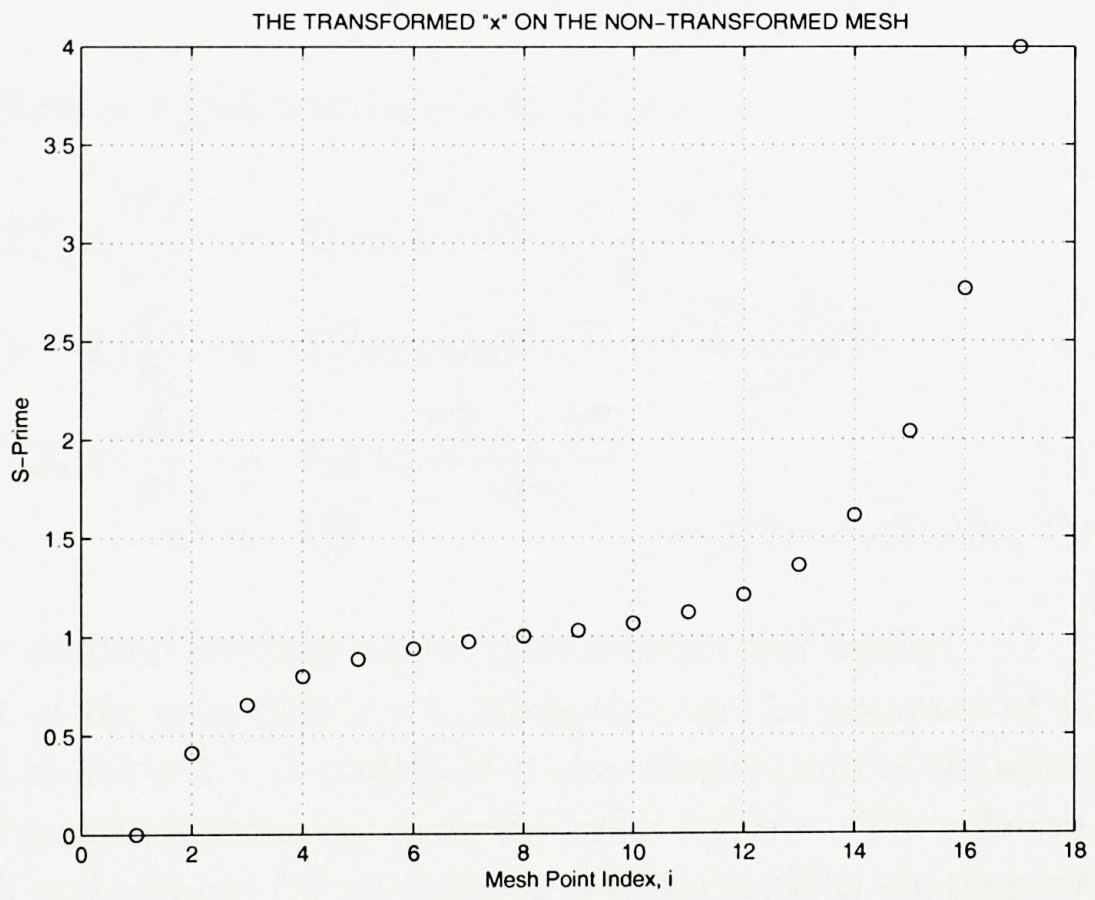
$$S' = \frac{\sinh(x - \sinh^{-1}(\lambda))}{\lambda} + 1 \tag{3.23}$$

The mesh would now be expected to take on small values around the point $S' = 1$, which is where the underlying asset price equals the strike; and large values toward the ends of the mesh. This effect can be seen when a plot is made of S' against the index of the mesh point in the non-transformed grid, Figure 3.8.

Figure 3.8: Visualisation of the effect.



(a) In the transformed grid, a plot of the transformed co-ordinates against the index of the mesh point produces equally spaced mesh intervals.



(b) In the original grid, the mesh is concentrated around the point $S' = 1$

3.2.1 European Options

The coefficients of the Black-Scholes differential operator will also be transformed. This application requires a little calculus.

$$S\frac{\partial u}{\partial S} = \left(S\frac{dx}{dS}\right)\cdot\frac{\partial u}{\partial x} = \mathcal{T}_\lambda(x)\frac{\partial u}{\partial x} \tag{3.24}$$

$$\begin{aligned} S^2\frac{\partial^2 u}{\partial S^2} &= \left(S\frac{dx}{dS}\right)^2\frac{\partial^2 u}{\partial x^2} + S^2\frac{\partial u}{\partial x}\cdot\frac{dx}{dS}\cdot\frac{\partial}{\partial x}\left(\frac{dx}{dS}\right) \\ &= \mathcal{T}_\lambda^2(x)\left(\frac{\partial^2 u}{\partial x^2} - \tanh(x-L)\frac{\partial u}{\partial x}\right) \end{aligned} \tag{3.25}$$

where

$$\mathcal{T}_\lambda(x) = \frac{\sinh(x-L)+\lambda}{\cosh(x-L)}$$

The transformed Black-Scholes equation is therefore written as

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x)\frac{\partial^2 u}{\partial x^2} &+ \frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x)\tanh(x-L)\frac{\partial u}{\partial x} \\ &- r\mathcal{T}_\lambda(x)\frac{\partial u}{\partial x} + ru = 0 \end{aligned} \tag{3.26}$$

The finite difference approximations used are now

$$\begin{aligned} \mathcal{T}_\lambda^2(x)\frac{\partial^2 u}{\partial x^2} &\approx \mathcal{T}_\lambda^2(x_j)\frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{\Delta x^2} \\ \mathcal{T}_\lambda^2(x)\tanh(x-L)\frac{\partial u}{\partial x} &\approx \mathcal{T}_\lambda^2(x_j)\tanh(x_j-L)\frac{U_j^m - U_{j-1}^m}{\Delta x} \\ \mathcal{T}_\lambda(x)\frac{\partial u}{\partial x} &\approx \mathcal{T}_\lambda(x_j)\frac{U_{j+1}^m - U_j^m}{\Delta x} \\ u &\approx U_j^m; \qquad j = 1,2,\cdots,N-1\,; \quad m = 1,2,\cdots \end{aligned} \tag{3.27}$$

and the same solution method can be used as described earlier.

The effect of the co-ordinate transformation on the accuracy of the solution is demonstrated in Table 3.5. An estimate of the relative error of the solutions on mesh sizes of 20, 40 and 80 is obtained using the exact solution. The relative errors of the solutions with and without the co-ordinate transformations are then compared. It is evident from the table that the co-ordinate transformations are effective in obtaining good accuracy on practical mesh sizes.

Table 3.5: Improvement in accuracy achieved by the co-ordinate transformations. The table shows the percentage error in the solutions on grids of sizes 20, 40 and 80, with and without the co-ordinate transformations. Model parameters are $K = 10$, $r = 0.1$, $\sigma = 0.2$, $T = 4$ months.

S	with/without transformation	Percentage Error		
		20	40	80
8	with transformation	0.28	0.07	0.02
	without transformation	0.35	0.10	0.02
9	with transformation	0.80	0.20	0.05
	without transformation	1.98	0.65	0.17
10	with transformation	0.88	0.20	0.03
	without transformation	17.36	3.71	0.87
11	with transformation	1.64	0.38	0.13
	without transformation	24.31	6.93	1.76

Quantisation Error

The structure of the error term directly shows that errors in the payoff will propagate back to the option prices. The initial data for a vanilla call or put is best represented when a mesh point is placed at the strike. This is easily arranged either by adjusting $S_{\max}$, or by constraining the transformation itself.

However the truncation error at the strike mesh point is very large initially due to the discontinuity of slope for vanilla puts or calls. Placing the strike in between mesh points also gives good results since it avoids this problem. This is necessary for digital options which have a discontinuity in the option value itself at the strike. Here the strike should definitely be midway between mesh points for similar reasons. This is discussed further in section 3.2.2.

Transformed Delta and Gamma

The hedging parameters are suitably transformed. Equation (3.24) implies that the delta becomes

$$\Delta = \frac{\partial u}{\partial S} = \frac{1}{S}\,\mathcal{T}_\lambda(x)\frac{\partial u}{\partial x}$$

and the gamma becomes

$$\Gamma = \frac{\partial^2 u}{\partial S^2} = \frac{1}{S^2}\,\mathcal{T}_\lambda^2(x)\left(\frac{\partial^2 u}{\partial x^2} - \tanh(x-L)\frac{\partial u}{\partial x}\right)$$

Numerical results are now presented for the co-ordinate transformed case. The effect of the transformation on the accuracy of the solution is demonstrated in Table 3.5. The relative errors of the solutions with and without the co-ordinate transformations are compared. It is evident from this table that the co-ordinate transformations are essential in obtaining good accuracy on practical mesh sizes.

Tables 3.6 and 3.7 show the numerical convergence of the solution when using co-ordinate transformations. The solutions at the discrete points were obtained by interpolating between grid points using a cubic spline interpolant. The results suggest the scheme is convergent. In this case, the desired level of accuracy can be obtained on a mesh size of 80.

In Tables 3.8 and 3.9 the effect of refining the time-step is seen for both put and call options.

In combination with spatial grid refinement, some time stepping refinement was also considered. Does a non-uniform time stepping strategy have any effect on the errors seen? In particular, are the errors more equi-distributed as a result of time stepping refinement?

The strategy was that only the first time-step was halved successively three times. This effectively produced two equal (very small) time-steps initially of length $\frac{dt}{8}$, then one of length $\frac{dt}{4}$, then one of length $\frac{dt}{2}$. For the purpose of displaying results, once these initial time-steps have added up to dt, the counting begins. All subsequent time-steps are of length dt.

As in the case of the uniform spatial grid, the l_∞ error is evaluated at the strike and Figure 3.9 shows a log plot of the maximum error in the price per time-step. The term *Price Error* in the figure is the error (|calculated solution - analytic solution|) for the price values at the strike. This is plotted against time. The two cases show a comparison between uniform and non-uniform time-stepping strategies. It is clear that a slight equi-distribution of the errors occurs as a result of using small initial time-steps, but it may be argued that the effect is not that significant.

The next set of figures (3.10) show $\log \|U_{2kh} - \text{exact}\|_\infty$ against $\log(N)$ for $k = 1, 2, 3, 4$, representing the solutions calculated on non-uniform mesh sizes of $N = 20, 40, 80, 160$ at time= 0. Similar log plots for the delta and gamma values are also shown.

Table 3.6: Numerical convergence of the solution to the European Put problem using the co-ordinate transformations. Model parameters are $K = 10,\ r = 0.1$.

T	S	σ	Solution Grid			Exact
			20	40	80	
4 months	8	0.20	1.6888	1.6923	1.6932	1.6935
		0.45	1.9719	1.9785	1.9801	1.9806
	9	0.20	0.84	0.8451	0.8464	0.8468
		0.45	1.334	1.3366	1.3374	1.3376
	10	0.20	0.3049	0.307	0.3075	0.3076
		0.45	0.8614	0.8611	0.861	0.861
	11	0.20	0.0781	0.0791	0.0793	0.0794
		0.45	0.533	0.5321	0.5318	0.5318
	12	0.20	0.0146	0.0149	0.0149	0.0149
		0.45	0.3176	0.3174	0.3174	0.3174
8 months	8	0.20	1.4701	1.4792	1.4815	1.4822
		0.45	2.0428	2.0511	2.0532	2.0538
	9	0.20	0.7825	0.7892	0.7908	0.7914
		0.45	1.517	1.5212	1.5222	1.5226
	10	0.20	0.3536	0.3583	0.3588	0.3589
		0.45	1.11	1.1107	1.1109	1.111
	11	0.20	0.1389	0.1396	0.1398	0.1399
		0.45	0.8015	0.8010	0.8008	0.8009
	12	0.20	0.047	0.0476	0.0477	0.0478
		0.45	0.572	0.5718	0.5716	0.5722

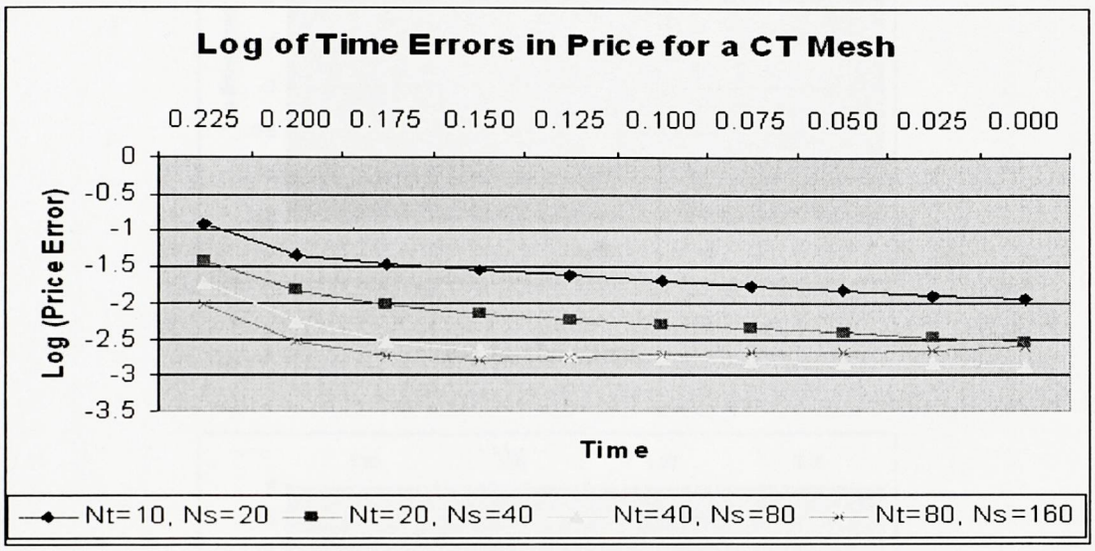
Table 3.7: Numerical convergence of the solution to the European Call problem using the co-ordinate transformations. Model parameters are $K = 40$, $r = 0.1$.

T	S	σ	Solution Grid			Exact
			20	40	80	
0.25	35	0.20	0.2571	0.2598	0.2606	0.2608
		0.45	1.6914	1.6865	1.6852	1.6848
	40	0.20	2.1194	2.1184	2.1182	2.1181
		0.45	4.068	4.0576	4.055	4.0542
	45	0.20	6.1383	6.1335	6.1323	6.1319
		0.45	7.4773	7.4831	7.4845	7.485
0.5	35	0.20	0.8792	0.8818	0.8825	0.8826
		0.45	3.2672	3.2537	3.2503	3.2492
	40	0.20	3.3313	3.3161	3.3123	3.3111
		0.45	5.9869	5.9733	5.9697	5.9687
	45	0.20	7.3073	7.2926	7.2889	7.2878
		0.45	9.4230	9.4281	9.4290	9.4297
1.0	35	0.20	2.2926	2.2819	2.2792	2.2784
		0.45	5.7847	5.7638	5.7581	5.7575
	40	0.20	5.3549	5.3196	5.3106	5.3079
		0.45	8.8631	8.8512	8.8471	8.8491
	45	0.20	9.4193	9.3829	9.3736	9.3708
		0.45	12.4204	12.4339	12.4352	12.4426

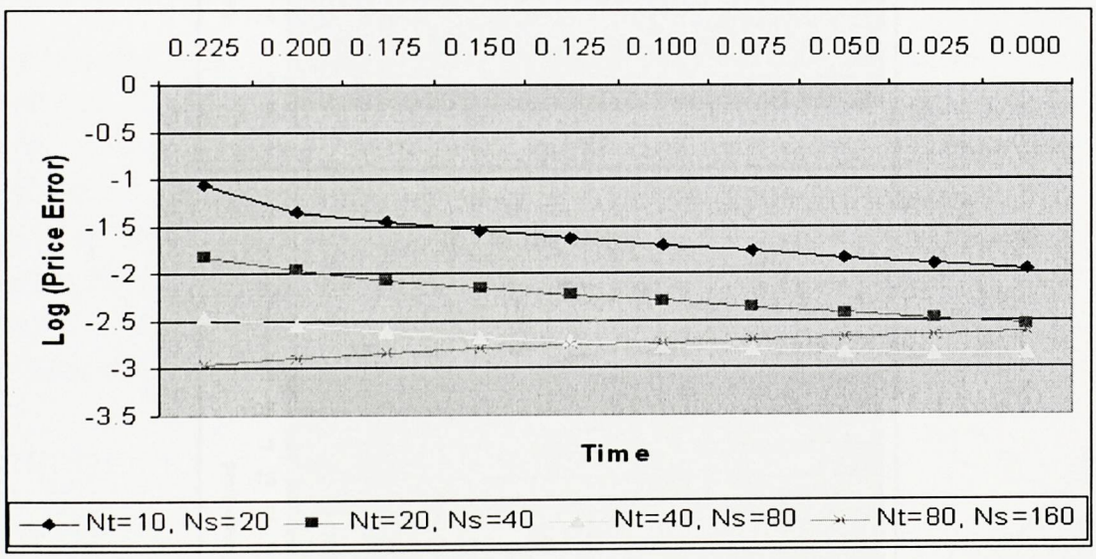
Table 3.8: Numerical convergence as the time-step is refined. Model parameters for this European Call option are $K = 40$, $r = 0.1$, $T = 0.5$. Grid size is 80.

S	σ	Time-step Δt					Exact
		16 days	8 days	4 days	2 days	1 day	
35	0.2	0.8752	0.8754	0.8754	0.8754	0.8754	0.8826
	0.45	3.2434	3.2433	3.2435	3.2436	3.2436	3.2492
40	0.2	3.2968	3.2967	3.2968	3.2968	3.2968	3.3111
	0.45	5.9599	5.9625	5.9626	5.9627	5.9628	5.9687
45	0.2	7.2807	7.2808	7.2809	7.281	7.281	7.2878
	0.45	9.4246	9.4244	9.4246	9.4247	9.4248	9.4297

Figure 3.9: Log plot of l_∞ error per time-step.

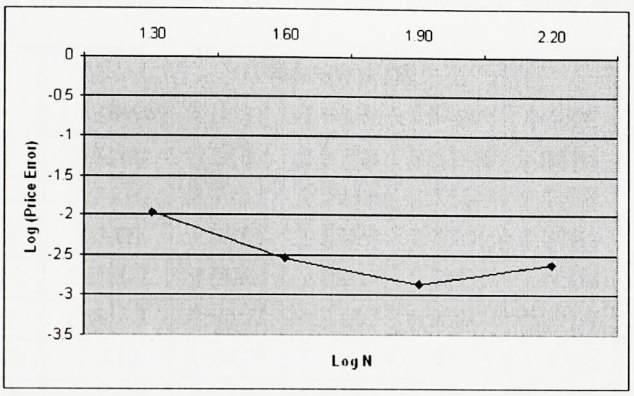


(a) Uniform Time-Steps.

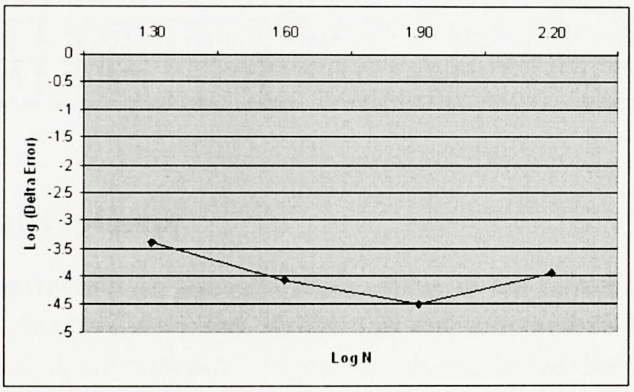


(b) Non-Uniform Time-Steps.

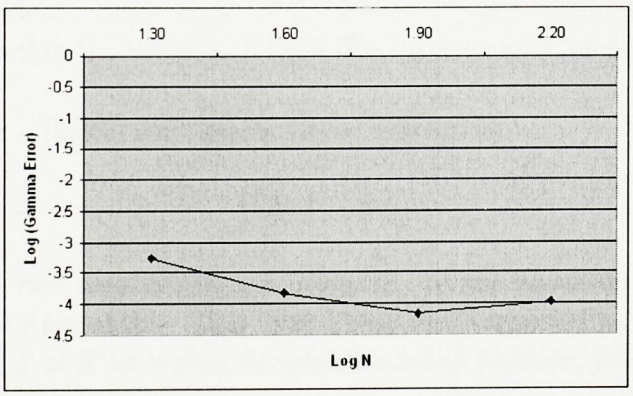
Figure 3.10: Convergence Plots for Non-Uniform Meshing.



(a) In the Price.



(b) In the Delta.



(c) In the Gamma.

Table 3.9: Numerical convergence as the time-step is refined. Model parameters for this European Put option are $K = 10$, $r = 0.1$, $T = 0.5$. Grid size is 80.

S	σ	Time-step Δt					Exact
		16 days	8 days	4 days	2 days	1 day	
7	0.2	2.5189	2.5188	2.5188	2.5188	2.5188	2.5184
	0.45	2.7416	2.7417	2.7418	2.7418	2.7418	2.7421
8	0.2	1.578	1.5781	1.5781	1.5781	1.5781	1.578
	0.45	2.0204	2.0204	2.0204	2.0204	2.0204	2.0211
9	0.2	0.8157	0.8157	0.8157	0.8157	0.8157	0.8173
	0.45	1.443	1.4428	1.4428	1.4428	1.4428	1.4437
10	0.2	0.3379	0.3378	0.3378	0.3378	0.3378	0.3401
	0.45	1.0004	1.0036	1.0035	1.0035	1.0035	1.0045
11	0.2	0.1128	0.1128	0.1128	0.1128	0.1128	0.1142
	0.45	0.6835	0.6833	0.6833	0.6833	0.6833	0.6841
12	0.2	0.0312	0.0312	0.0312	0.0312	0.0312	0.0317
	0.45	0.4575	0.4574	0.4574	0.4574	0.4574	0.4581

3.2.2 Digital Options

Many financial contracts have discontinuities, either in the payoff functions or their derivatives. The payoff of standard vanilla call and put options has a discontinuous first derivative. A characteristic of digital options is the discontinuity in their payoff itself. The digital option has a payoff at expiry that is discontinuous in the underlying asset price domain. The idea of smoothing the payoff has been discussed before in the finance literature (see [HZ00], [TR00], [PVF01]). Here, PDE methods are used to numerically price digital call options. Similar techniques to those that have been discussed in the thesis so far are also implemented for digital call options.

The payoff for a digital call option can be written as

$$\mathrm{B\ H}(S - K)$$

for some positive, real constant B, and where $\mathrm{H}(\cdot)$ is the heaviside function. Figure 3.11 is a representation of this. This option may be interpreted as a straight bet on the asset price, if $S > K$ at expiry the payoff is some positive, real constant B and otherwise it is 0.

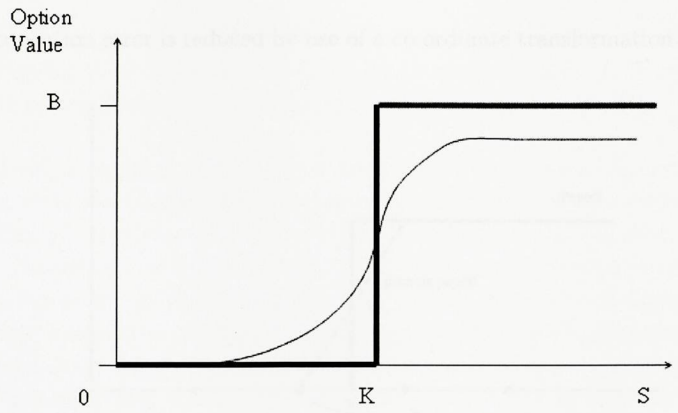


Figure 3.11: Payoff Diagram for a Digital Call option. The bold line represents payoff, the fine line is the contract value some time prior to expiry.

Assume the underlying asset follows geometric Brownian motion and that investors are risk-neutral. The value of the digital option satisfies the Black-Scholes pricing equation,

$$\frac{\partial u(S,t)}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u(S,t)}{\partial S^2} - rS\frac{\partial u(S,t)}{\partial S} + ru(S,t) = 0$$

When pricing forwards in time, the option payoff becomes the initial condition. For a digital call this is

$$u(S,T) = \begin{cases} \text{B} & \text{if } S \geq K \\ 0 & \text{otherwise} \end{cases}$$

As $S \to \infty$, the option is at most linear in the underlying and so

$$\frac{\partial^2 u}{\partial S^2} \to 0 \text{ as } S \to \infty$$

[TR00] remarks that having the strike price occur midway between mesh nodes generally increases the accuracy of the finite difference method. In this study too, the finite difference grid for this discontinuous initial payoff is generated such that the strike is placed in between two mesh nodes. The error $e = u - U$ satisfies a difference equation similar to the discrete option price itself, but with an additional source term. This term is the truncation error of the option price difference equation and measures the error made by replacing the derivatives in Black-Scholes by differences.

This truncation error is reduced by use of a co-ordinate transformation as described before.

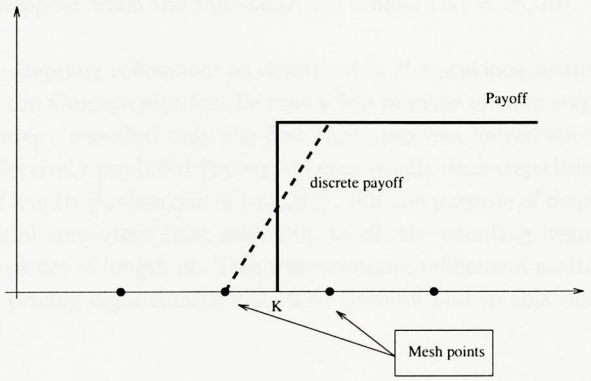


Figure 3.12: Error behaviour.

However the difference equation for the pricing error propagates the error in time in the same way that the Black-Scholes equation propagates the final payoff. Consequently the error in the mesh representation of the payoff also needs to be made as small as possible. The payoff for a digital put is a discontinuous function so the mesh representation error is minimised (in an area weighted norm) by placing the discontinuity midway between mesh points, as can be seen from Figure 3.12. For a vanilla call/put option payoff a mesh point would be placed on the strike itself.

Numerical results are now presented for the problem whose parameters are as set out in Table 3.10. Figure 3.13 shows a plot of the solution curve as well as

Table 3.10: Parameters of the Problem

Option Type	Digital Call
Time to expiry	0.5 Years
Strike Price	40
Volatility	0.3
Interest Rate	0.05
Theta	0.505
Stretch	20

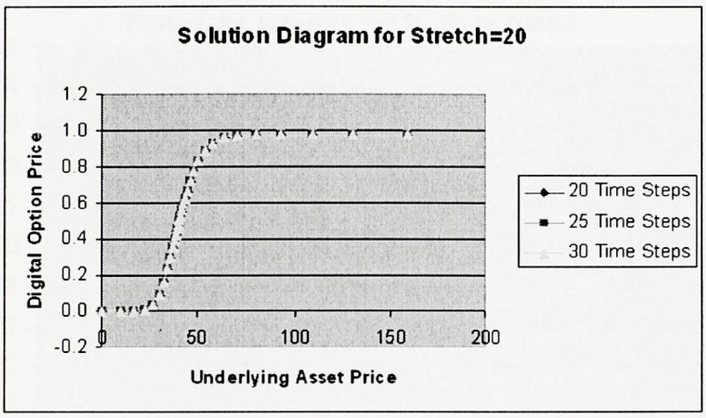
the Delta and Gamma curves for this problem. The Gamma plot exhibits some

spurious oscillations for the case $Nt = 20$ which is a consequence of Crank-Nicolson time-stepping when twice the maximum stable explicit time-step is exceeded. The oscillations disappear when the time-steps are refined ($Nt = 25, 30$).

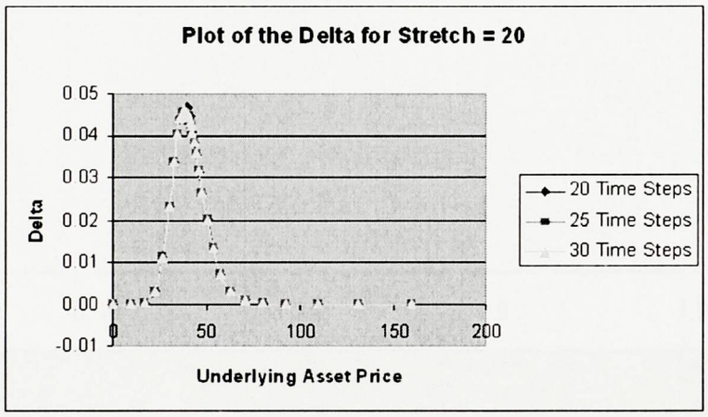
If the time stepping refinement as described in the previous section is used, the oscillations in the Gamma plot for the case when number of time steps = 20 disappear. The strategy was that only the first time-step was halved successively three times. This effectively produced two equal (very small) time-steps initially of length $\frac{dt}{8}$, then one of length $\frac{dt}{4}$, then one of length $\frac{dt}{2}$. For the purpose of displaying results, once these initial time-steps have added up to dt, the counting begins. All subsequent time-steps are of length dt. This time stepping refinement strategy makes the digital option pricing algorithm robust. The Gamma plot in this case can be seen in Figure 3.14.

Convergence results are presented for a sequence of grids starting with 8 nodes, refined successively to 64. Constant time stepping is used. Also, the same coordinate stretching as described previously is made use of. Table 3.11 shows the benefits of placing the strike in between mesh points. The modified Crank-Nicolson method alone produces linear convergence whereas in combination with placing the strike in between mesh points, the method produces better than quadratic convergence for fewer time steps. This can be seen clearly in Figure 3.15. Table 3.12 is used as a comparison of results with those published elsewhere. These numbers are produced from [PVF01].

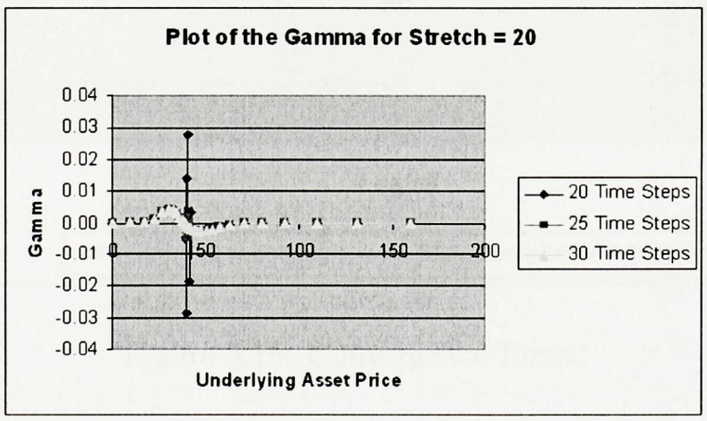
Figure 3.13: Digital Call option plots as a function of S for several values of time-step. Model parameters are as set out in Table 3.10.



(a) Solution Plot.



(b) Delta Plot.



(c) Gamma Plot.

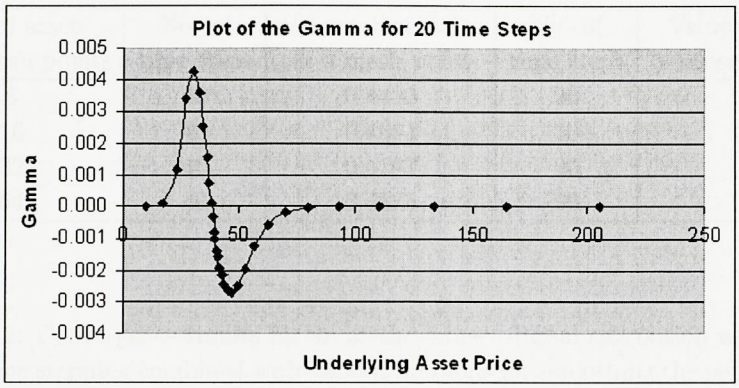


Figure 3.14: Gamma plot using non-uniform initial time steps.

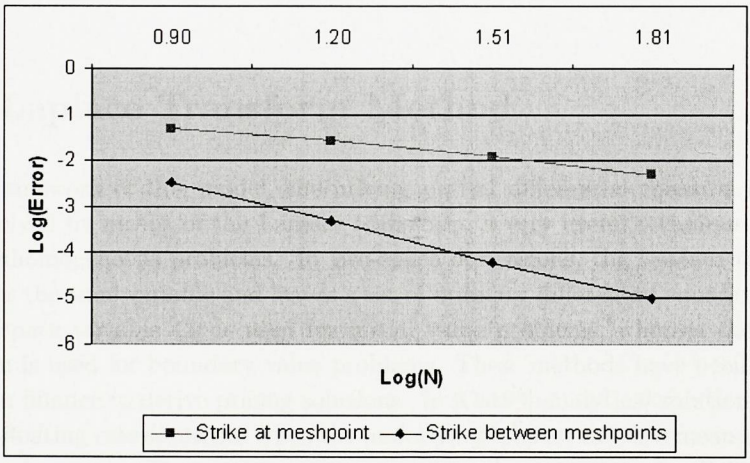


Figure 3.15: Convergence Plots.

Table 3.11: Solution values of the digital call for different mesh point strategies. The exact value is 0.4922403

No. of asset price mesh points	No. of time steps	Value when strike at a mesh point	No. of time steps	Value when strike between mesh points
8	40	0.4423	20	0.4890
16	40	0.4651	20	0.4918
32	40	0.4797	20	0.4921
64	40	0.4875	20	0.4922

Table 3.12: Convergence results for an at-the-money digital call option using Rannacher time stepping combined with various methods for smoothing the payoff. This table is reproduced from [PVF01]. The exact value is 0.4922

Nodes	Value when averaging	Value when shifting mesh	Value when using projection
41	0.4929	0.4927	0.4929
81	0.4924	0.4924	0.4923
161	0.4923	0.4923	0.4923
321	0.4922	0.4923	0.4922

3.3 Laplace Transform Method

In the framework of this model, the pricing partial differential equation obtained can be solved by means of the Laplace transform, a very useful technique for handling nonhomogeneous problems. In two-space dimensions, the Laplace transform eliminates the time variable and leaves a set of ordinary differential equations in the Laplace space variable. It is used for initial value problems, whereas the Fourier transform is used for boundary value problems. These methods have been recently applied in finance to derive pricing solutions. In [Cat98] analytical solutions are derived for floating rate securities when the underlying interest rate is a mean-reverting square-root process. The model is also applied to the valuation of callable bonds. [GY93] computed the Laplace transform of the Asian option price, by means of Bessel processes. The approximate closed form solution fails to adapt to more complex volatility models, as well as to American type features. The paper [FMW99] shows the authors investigate and contrast two numerical methods for pricing Asian

options: Laplace transform inversion and Monte Carlo simulation. Further, the double Laplace transform of the call option price for the continuous Asian is derived, in both its strike and its maturity. [Fus04] prices Asian options by computing a Laplace transform with respect to time-to-maturity and a Fourier transform with respect to the logarithm of the strike.

Different approaches to the problem of numerically inverting Laplace transforms in finance are considered in [CHP00]. [GY93] derived an analytic expression for the Laplace transform in maturity for the continuous call option case when the asset price follows geometric Brownian motion, and numerical inversion of this transform was considered only briefly. Numerically inverting the [GY93] Laplace transform is a nontrivial problem, as it involves the transform parameter in the index of a Bessel function, or as an argument of a gamma function. As shown in [FMW99], naive implementation may lead to reasonable looking results that are grossly inaccurate. For the discrete-time case, no similar analytic results exist, although various approximations have been proposed.

Several methods are available for the numerical inversion of Laplace transforms,

$$F(q) = \int_0^\infty e^{-qt} f(t) dt$$

[Cat98] says the best known approach for deriving values of $f(t)$ from values of $F(q)$ is the numerical evaluation of the Bromwich integral

$$f(t) = \frac{1}{2\Pi i} \int_L e^{qt} F(q) dq$$

where L is defined as the line Re $q = c$ in the complex plane and c is chosen so that L lies to the right of all singularities of $F(q)$ but is otherwise arbitrary. To implement this technique, for fixed t, a change of variable was made $qt = u$ and $G(u) = u^q F(u/t)$ was set so that the integral is approximated by

$$\frac{1}{2\Pi i} \int_{L^*} e^u u^{-q} G(u) du \approx \sum_{k=1}^{N} A_k G(u_k)$$

where L^* is now Re $u = ct$. The weights A_k and the abscissas u_k are complex numbers. [Cat98] used this Gaussian quadrature in combination with a new integration formula derived by the optimal addition of abscissas. The only variable of choice is N and sets of corresponding values of coefficients and abscissas have been tabulated in [Cat98]. Satisfactory accuracy was obtained with $N = 6$.

[FMW99] find that inversion of the Laplace transform via numerical integration can lead to numerical problems for low volatilities and short maturities, with the

quantity $\sigma^2 t$ providing a good relative yardstick for difficulty. Difficulties seem to begin when $\sigma^2 t < 0.01$. These problems appear to be independent of the technique used for the inversion, being a result of the slowly decaying oscillatory nature of the integrand for such parameter values, although the point at which each integration routine begins to degenerate may differ. Among the numerical inversion techniques considered, the Euler method provides a high degree of accuracy for a reasonable amount of computation time. The Euler method [AW95] applies the trapezoidal rule over intervals of an appropriate length that reduce the cosine function to alternating signs. Euler summation is employed to speed the convergence of the alternating series generated by this method.

In order to numerically invert the double transform of [Fus04] and obtain the option price, a multivariate version of the Fourier-Euler algorithm is used. Given the double transform, first the Fourier inverse is computed numerically. Then the Laplace transform is numerically inverted. It results in the numerical inversion being highly accurate even for low volatility levels.

A description now follows of a study on a distributed algorithm for the numerical solution of a European option [CDR+00]. The application of Laplace transform to the mathematical model for pricing options [Shi95] leads to a set of mutually independent linear ordinary differential equations, and these equations may then be solved concurrently in a distributed computing environment. The scalability of the algorithm is studied theoretically in [Cra96] and [CDLL97].

The numerical inverse Laplace transform is done here by means of an inversion algorithm developed by Stehfast [Ste70]. This approximate inverse Laplace transform is not the most accurate one by any means. The Stehfast method is selected due to previous experience with the method for linear problems [CDLL97], [Ste70]. In the pricing context - for time varying $\sigma(t)$ and $r(t)$, it is possible to make suitable co-ordinate transformation to the Black-Scholes equation and obtain a time independent heat like equation [WHD97]. Hence the Laplace transform method described here may still be applied.

3.3.1 Application to Option Pricing

It is possible to apply the Laplace transform [Wid46], [Wat81] to the Black-Scholes equation. For a European put option,

$$\frac{\partial U(S,t)}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U(S,t)}{\partial S^2} - rS\frac{\partial U(S,t)}{\partial S} + rU(S,t) = 0;\ S \in [0,\infty];\ t \in [0,T] \tag{3.28}$$

with initial and boundary conditions

$$\begin{aligned} U(S,0) &= \max(K-S,0) \quad ; \quad S \in [0,\infty] \\ U(0,t) &= Ke^{-rt} \\ U(S_{\max},t) &= 0\,; \quad t \in [0,T] \end{aligned} \tag{3.29}$$

Let

$$\mathrm{L}\{U(S,t)\} \equiv \int_0^\infty e^{-\lambda t} U(S,t)\, dt = \overline{U}(S;\lambda)$$

be the Laplace transform of the function $U(S,t)$, where λ is a constant parameter assumed to be positive and large enough to make the product $e^{-\lambda t}U(S,t)$ converge to zero as $t \to \infty$.

Then the Laplace transform of the Black-Scholes equation leads to

$$\frac{1}{2}\sigma^2 S^2 \frac{d^2\overline{U}}{dS^2} + rS\frac{d\overline{U}}{dS} - (r+\lambda)\overline{U} = -U(S,0);\ S \in [0,\infty] \tag{3.30}$$

subject to the boundary condition

$$\begin{aligned} \overline{U}(0;\lambda) &= K \cdot \frac{1}{\lambda + r} \\ \overline{U}(S_{\max};\lambda) &= 0 \end{aligned}$$

Here $\lambda \in \{\lambda_j\}$ is a finite set of transformation parameters defined by

$$\lambda_j = j\frac{\ln 2}{t} \quad ; \ j = 1, 2, \ldots, m \tag{3.31}$$

for *even values of* m [Ste70]. In this way, the original problem (3.28) is converted to m independent parametric boundary value problems as described by (3.30). These problems may be distributed and solved independently in a computational environment which consists of a number of processors linked by a network. From experience, the value of m is usually a small even number not larger than ten [CDLL97]. Numerical experiments which follow also confirm such experience.

An approximate inverse Laplace transform as described by [Ste70] may be used to retrieve the value of $U(S,t)$ according to

$$U(S,t) \approx \frac{\ln 2}{t} \sum_{j=1}^{m} w_j \overline{U}(S;\lambda_j)$$

where

$$w_j = (-1)^{\frac{m}{2}+j} \sum_{k=\frac{(1+j)}{2}}^{\min(j,\frac{m}{2})} \frac{k^{\frac{m}{2}}(2k)!}{(\frac{m}{2}-k)!\ k!\ (k-1)!\ (j-k)!\ (2k-j)!}$$

is known as the weighting factor.

3.3.2 Numerical Solution for Constant Volatility

An example of the method applied to a European put option with constant volatility is given here. The spatial domain is truncated at $S_{\max} = 320$, so that $\{S : 0 \leq S \leq 320\}$. The expiry date is set as $T = 0.25$ (three months). The remaining model parameters are strike price $K = 100$, volatility $\sigma = 0.4$, interest rate $r = 0.5$ throughout the simulation. The mesh size is chosen to be $\Delta S = h = \frac{320}{2^9}$.

A second order finite volume method is applied to each parametric equation (3.30). The Laplace transformed set of equations is solved sequentially in the same computational environment, for different values of m (see equation 3.31). An approximation to $U(S,t)$ corresponding to each value of m is found by using the inverse Laplace transform as described earlier, and is denoted by U_{IL}.

As a comparison, the forward Black-Scholes equation given by (3.28) is solved by means of an Euler marching scheme along the temporal axis with time-step length set as $\Delta t = \frac{1}{365}$, i.e. 1 day, in conjunction with the above finite volume scheme applied along the spatial axis S. The discretisation leads to a tridiagonal system of equations at every time-step, which is then solved by a direct method. The numerical solution for $U(S,t)$ obtained in this case is denoted by U_{TD}.

3.3.3 Results

Discrepancies between solutions, $\|U_{TD} - U_{IL}\|_2$ are recorded and shown in Table 3.13 for comparison. Solutions and timings were obtained on a Sun Ultra-5 workstation using an F90 program which implements the two methods described. As a comparison, the time taken to compute U_{TD} is 0.133.

The timing of each run observed in this example consists of two parts. The second order finite volume solver is first timed. Also timed are the overheads due to the computation of the inverse Laplace transforms. A crude estimate of the distributed computing processing time can be obtained by dividing the timings for the inverse Laplace algorithm. If there are as many processors available as the value of m, the scalability of the algorithm can be easily observed from the sequential timings recorded in Table 3.13 using the crude estimates of the corresponding distributed timings.

3.3.4 Discussion

Observing from Table 3.13, the discrepancy $\|U_{TD} - U_{IL}\|_2$ approaches an asymptotic value of 0.0032 when $m \geq 8$. Therefore it is not necessary to take a value of m very

Table 3.13: A table showing the timing of computation, and discrepancy comparisons for the two different solution methods.

m	$\|U_{TD} - U_{IL}\|_2$	Time (U_{IL})
2	1.0767	0.006
4	0.0812	0.009
6	0.0111	0.014
8	0.0037	0.017
10	0.0032	0.018
12	0.0032	0.021
14	0.0032	0.028
16	0.0032	0.028
18	0.0032	0.035

much larger than 10. This result confirms published tests on a linear heat conduction problem [CDLL97].

Since the total timings shown in Table 3.13 are very small, one may argue that the distributed algorithm is not necessary. However, as discussed in [CDR+00], the situation for non-linear problems is very different due to the linearisation steps, and hence the total computation increases. Preliminary tests in non-linear volatility can be found in [LPRH05].

The study (of which the work described in this section forms a part [Cra96], [CDLL97], [CDR+00]) suggests that inverse Laplace techniques have advantages in solving non-linear option pricing problems. Further investigation into alternative methods of the inverse Laplace transform is currently being undertaken by other researchers at the University of Greenwich.

An algorithm based on the Laplace transform of the time domain into a set of parametric equations is developed for a European put option with one spatial variable, S. Distributed computing may be applied to solve the parametric equations concurrently. An inverse Laplace transform based on the Stehfast method [Ste70] is applied to retrieve the solution.

3.4 Overview

Accurate solutions have been obtained for European call and put options, as well as for Digital call options.

A modified Crank-Nicolson finite difference scheme is presented for single factor European put options. The Black-Scholes equation is defined on a finite computational domain which is then discretised. Second order central differences are used as much as possible to approximate the equation at the discrete nodes. First order upwind differences are used when the mesh Peclet conditions are violated. Appropriate asymptotic boundary conditions are imposed at the truncation point in the spatial direction. The resulting tridiagonal system of equations is solved using a direct solver. Grid refinement displays convergence towards the true solution.

Results using the same algorithm are also presented for European call options. However, corresponding results for European call options can easily be obtained by put-call parity.

An essential component of the solutions strategy is the co-ordinate transformations used to produce non-uniform grids in the original co-ordinate system with detailed resolution in the regions where there is greatest interest and where accuracy is required. The co-ordinate transformations ensure that accurate solutions can be obtained on practical mesh sizes allowing rapid computations of option prices, using a solution strategy which exploits the structure of the coefficient matrix.

An application of Laplace inversion methods has also been implemented for the one-factor European option with constant volatility. The results show the scalability of the method when it is implemented in parallel computation. Preliminary tests in non-linear volatility can be found in [LPRH05].

Chapter 4

Numerical Solution of Asian Options

Conventional call and put options have payoffs that depend on the instantaneous price of the underlying security, either at maturity (European case) or throughout the contract (American case); this creates an incentive for traders to attempt price manipulation on the underlying security. Asian options avoid this by using payoffs that depend on the average price of the underlying security, rather than the instantaneous price. Asian options are path-dependent and exotic. They are particularly interesting from a mathematical standpoint. A variety of methods can be applied to the pricing of Asian options [KV90, BP96, RS95, ZFV98a, FMW99, Vec01].

Payoffs for Asians depend on the history of the random walk of the asset price via some sort of average. Several factors affect the definition of the average. Different choices of average lead to different values for the option. Typically, the exercise price is some form of average of the price of the underlying over some period of time prior to exercise. This average could be **geometric or arithmetic**, which may be measured either **continuously or discretely**.

The path-dependent pricing problem can be cast into the classical Black-Scholes valuation framework through inclusion of the path dependent variables into the state space. In a few simple cases, the resulting augmented PDE admits a closed form solution (see [BP96]). However the use of numerical techniques is required in most practical solutions. Standard techniques tend to be impractical, slow or inaccurate. For example, traditional binomial lattice methods require such enormous amounts of computer memory (owing to the necessity of keeping track of every possible path throughout the tree) that they are effectively unusable. PDE methods, as traditionally implemented in the finance literature, are inaccurate; see [BP96] for a discussion.

[GY93] computed the Laplace transform of the Asian option price, but numerical inversion remains a problem for low-volatility and/or cases of short-maturity (see [FMW99]). Monte Carlo simulation works well, but it can be computationally expensive without the enhancement of variance-reduction techniques, and one must account for the inherent discretisation bias resulting from the approximation of continuous-time processes through discrete sampling. See [KV90] for use of Monte Carlo simulation with a specific variance reduction method to compute the price of fixed-strike average-rate options.

In general, the price of an Asian option can be found by solving a PDE in two space dimensions, which is prone to oscillatory solutions. See [Ing87] where it is also observed that the two-dimensional PDE for a floating-strike Asian option can be reduced to a one-dimensional PDE. [RS95] have formulated a one-dimensional PDE which can model both floating- and fixed-strike Asians. They reduced the dimension of the problem by dividing $K - \bar{S}_t$ (where K is the strike and $\bar{S}_t$ is the average stock price over $[0, t]$) by the stock price S_t. However, this one-dimensional PDE is difficult to solve numerically since the diffusion term is very small for values of interest on the finite-difference grid. They propose a numerical technique to compute a tight lower bound on European average-rate option prices. The method is accurate and considerably faster than Monte Carlo simulation. The Dirac delta function also appears as a coefficient of the PDE in the case of the floating-strike option.

Finite difference methods are very flexible with regard to the asset price model, but encounter difficulty when applied to PDE models of Asian options because of the parabolic degeneracy in the average-price direction. This has led a number of authors to make use of techniques deriving from the numerical solution of hyperbolic PDE's e.g. [ZFV98a], such as TVD approximations. These techniques all have the disadvantage of introducing nonlinearity into the two-factor time-implicit finite difference equations. In [PC99] it is shown how Semi-Lagrange (S-L) time-integration, developed for numerical weather forecasting, is an elegant choice of technique which integrates out the average price term and simplifies the finite difference equations into a parameterised Black-Scholes form. Boundary conditions are unnecessary in the average-price direction. The implicit equations that result are unconditionally stable, second order accurate and can be solved using standard tridiagonal solvers. Uniform meshes are not efficient however the S-L method is shown to be easily applied in conjunction with co-ordinate transformations. The S-L method is more effective with small number of time steps since it is increasingly diffusive as the number of time steps becomes large. Early exercise is also easily incorporated, the resulting linear complementarity problem can be solved using a projection or penalty method. The S-L time integration method has been shown to dramatically simplify

the finite difference approximation of Asian options. Second order accuracy has been confirmed for Asian options that must be held to maturity. A comparison with published results for continuous-average-rate put and call options, with and without early exercise, shows that the method achieves basis point accuracy and that Richardson extrapolation can also be applied. This S-L time-integration method can be applied to multi-factor options in the same way leading to parameterised forms of multi-factor Black-Scholes equations.

4.1 Mathematical Model

The conventional geometric Brownian motion model is adopted for the evolution of the asset price, S_t:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{4.1}$$

where W_t is a standard Brownian motion. More general models (e.g. volatility surface) are also easily incorporated into this approach.

Let $[0, T]$ be the time interval within which a 2-factor Asian option, $u(S, A, t)$ is structured. It is assumed that:

1. All options are issued at time $t = 0$.

2. All options expire at time $T > 0$.

Asian options depend on some form of an average of the underlying stock price over time. In order to price such options a state variable A_t, the continuously sampled arithmetic average of S_t over $[0, t]$, is introduced. It is defined by

$$A_t = \frac{1}{t} \int_0^t S_\tau \, d\tau \tag{4.2}$$

and evolves according to the equation

$$dA_t = \frac{1}{t}\Big(S_t - A_t\Big)dt \tag{4.3}$$

According to Ito's lemma (2.10) and the analysis in subsection 2.6.1, the pricing function $u(S, A, t)$ changes by

$$du = \frac{\partial u}{\partial t}dt + \frac{\partial u}{\partial S}dS + \frac{\partial u}{\partial A}dA + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2}dt \tag{4.4}$$

Now consider a portfolio with initial value Π_0 and value at time t, Π_t, consisting of one long option position and a short position in a number Δ of the underlying asset S_t. The value of the portfolio at time t is given by

$$\Pi_t = u(S_t, A_t, t) - \Delta S_t \tag{4.5}$$

The change in the value of the portfolio in time $[0, t]$ is therefore

$$d\Pi_t = du(S_t, A_t, t) - \Delta dS_t \tag{4.6}$$

Note that Δ has not changed during the time-step. Substituting (4.4) into (4.6) gives

$$d\Pi = \frac{\partial u}{\partial t}dt + \frac{\partial u}{\partial S}dS + \frac{\partial u}{\partial A}dA + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2}dt - \Delta dS \tag{4.7}$$

This portfolio can be delta-hedged by choosing

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

thereby eliminating the random terms. Equation (4.7) becomes

$$d\Pi = \frac{\partial u}{\partial t}dt + \frac{\partial u}{\partial A}dA + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2}dt$$

The portfolio held now is risk-less. The no-arbitrage assumption demands that the value of the portfolio changes at the same growth rate as money invested in a bank.

$$\begin{aligned}
&\Rightarrow \quad d\Pi \;=\; r\ \Pi\ dt \\
&\Rightarrow \quad \left[\frac{\partial u}{\partial t} + \frac{\partial u}{\partial A}\frac{dA}{dt} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2}\right] dt \;=\; r\ \Pi\ dt \\
&\qquad\qquad\qquad\qquad\qquad\qquad = r\left(u - \frac{\partial u}{\partial S}S\right) dt \\
&\Rightarrow \quad \frac{\partial u}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2} + rS\frac{\partial u}{\partial S} - ru + \frac{\partial u}{\partial A}\frac{dA}{dt} = 0
\end{aligned}$$

Substitute equation (4.3) into the above to achieve the final result.

The PDE for pricing an Asian option is an "ultra-parabolic equation" given in this case (see [BP96]) by

$$\begin{aligned}
-\frac{\partial u}{\partial t} \;&=\; \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2} + rS\frac{\partial u}{\partial S} + \frac{1}{t}(S - A)\frac{\partial u}{\partial A} - ru \\
&=\; \mathcal{L}_{BS}\ u + \frac{1}{t}(S - A)\frac{\partial u}{\partial A} \quad ; \quad S \geq 0\ , A \geq 0\ , t \in (0, T]
\end{aligned} \tag{4.8}$$

for a given payoff with final conditions at time $t = T$,

$$u(S, A, T) = g(S, A, T) \tag{4.9}$$

where g is a known function, and $\mathcal{L}_{BS}$ is the Black-Scholes differential operator of equation (2.22).

If the option is exercised early at some time $t \in (0, T]$, the same payoff function applies i.e.

$$u(S, A, t) = g(S, A, t)$$

At $t = 0$, the average price is equal to the current asset price. That is, $A_0 = S_0$ and the PDE reduces to the parabolic Black-Scholes PDE.

4.1.1 Boundary Conditions

As expected, the data specified at the boundary determines the type of Asian option being priced.

Fixed Strike Call

An average-rate fixed-strike (rate) call has payoff

$$g(S, A, T) = \max(A - K,\ 0) \tag{4.10}$$

If at some time t^* the asset price $S_{t^*} = 0$ then it remains zero; hence the final average price is known

$$A_T = \frac{1}{T}\int_0^T S_t\, dt = \frac{1}{T}\int_0^{t^*} S_t\, dt = \frac{t^*}{T} A_{t^*}$$

and hence the payoff, $\forall A \geq 0$ is also known,

$$u(0, A, t) = e^{-r(T-t)} \max(\frac{t}{T} A - K, 0) \tag{4.11}$$

Alternatively use the exact solution of the PDE for $S = 0$ (it becomes linear hyperbolic). Payoffs a long way from the strike have diminishing influence.
As $S \to \infty$,

$$\frac{\partial^2 u(S, A, t)}{\partial S^2} \to 0 \qquad \forall\ A \geq 0; \quad t \in (0, T] \tag{4.12}$$

As $A \to \infty$,

$$\frac{\partial^2 u(S, A, t)}{\partial A^2} \to 0 \qquad \forall\ S \geq 0; \quad t \in (0, T] \tag{4.13}$$

At zero average value, the asset price as well must be zero. Thus, the line $A = 0$ on the computational domain has no physical meaning except at the point $S = 0$, $A = 0$, at which point the Black-Scholes equation is solved instead:

$$u(0, 0, t) = Ke^{-r(T-t)} \qquad t \in [0, T]$$

However, at $A = 0$, information flow is outward (this is influenced by the equation being integrated backwards in time, giving a negative advection term in the average

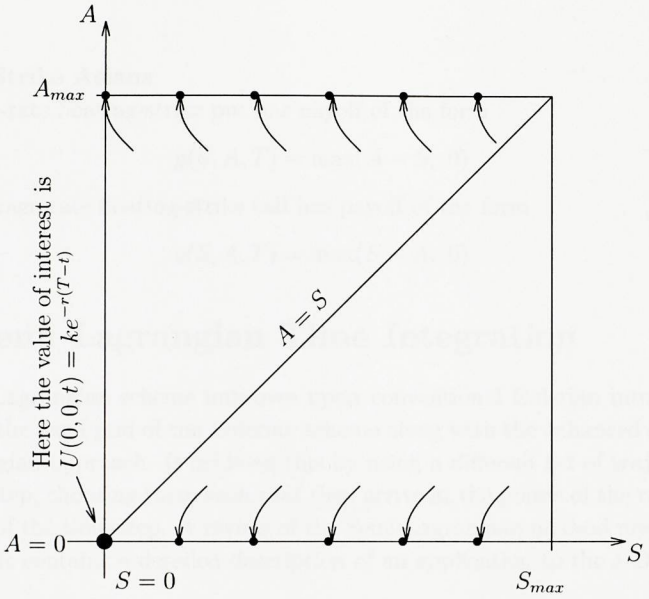


Figure 4.1: Boundary Conditions at $A = 0$.

price direction). At this boundary outflow conditions are prescribed.
At $A = 0$,

$$\frac{\partial u(S,0,t)}{\partial A} = 0 \qquad \forall\, S \geq 0; \quad t \in (0,T] \tag{4.14}$$

Fixed Strike Put

An average-rate fixed-strike (rate) put has terminal payoff

$$g(S,A,T) = \max(K - A,\ 0) \tag{4.15}$$

Similar arguments as in the case of a fixed-strike call produce boundary conditions:

$$u(0,A,t) = e^{-r(T-t)} \max(K - \frac{t}{T}A, 0);\ \forall A \geq 0;\ \forall t \in (0,T] \tag{4.16}$$

$$\frac{\partial u(S,0,t)}{\partial A} = 0 \qquad \forall\, S \geq 0; \quad t \in [0,T] \tag{4.17}$$

$$\frac{\partial^2 u(S,A,t)}{\partial S^2} \to 0 \quad \text{as } S \to \infty;\ \forall\, A \geq 0; \quad t \in [0,T] \tag{4.18}$$

$$\frac{\partial^2 u(S,A,t)}{\partial A^2} \to 0 \quad \text{as } A \to \infty;\ \forall\, S \geq 0; \quad t \in [0,T] \tag{4.19}$$

Floating Strike Asians
An average-rate floating-strike put has payoff of the form

$$g(S, A, T) = \max(A - S,\ 0)$$

and an average-rate floating-strike call has payoff of the form

$$g(S, A, T) = \max(S - A,\ 0)$$

4.2 Semi-Lagrangian Time Integration

The Semi-Lagrangian scheme improves upon conventional Eulerian integration by combining the fixed grid of the Eulerian scheme along with the enhanced stability of the Lagrangian approach. It achieves this by using a different set of trajectories at each time-step, choosing them such that they arrive at the points of the regular grid at the end of the time-step. A review of the Semi-Lagrangian method may be found in [SC91]. It contains a detailed description of an application to the 1-D advection equation

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial x}\,\frac{dx}{dt} = 0$$

where

$$\frac{dx}{dt} = g(x, t)$$

and $g(x, t)$ is a given function.

The Lagrangian Derivative:
By definition,

$$\frac{du}{dt} = \lim_{\Delta t \to 0} \frac{u\Big(t + \Delta t, A(t + \Delta t)\Big) - u\Big(t, A(t)\Big)}{\Delta t} \tag{4.20}$$

as illustrated in Figure (4.2). This implies that

$$\frac{du}{dt} = \lim_{\Delta t \to 0} \frac{u\Big(t + \Delta t, A(t + \Delta t)\Big) - u\Big(t + \Delta t, A(t)\Big) + u\Big(t + \Delta t, A(t)\Big) - u\Big(t, A(t)\Big)}{\Delta t} \tag{4.21}$$

which can be re-written as

$$\begin{aligned} \frac{du}{dt} &= \lim_{\Delta t \to 0} \frac{u\Big(t + \Delta t, A(t)\Big) - u\Big(t, A(t)\Big)}{\Delta t} \\ &+ \frac{u\Big(t + \Delta t, A(t + \Delta t)\Big) - u\Big(t + \Delta t, A(t)\Big)}{\Delta A} \cdot \frac{\Delta A}{\Delta t} \end{aligned} \tag{4.22}$$

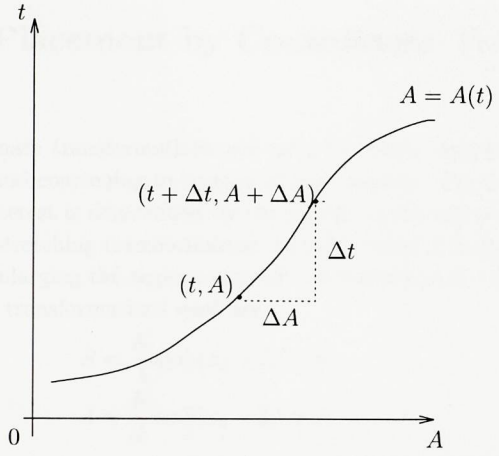


Figure 4.2: Lagrangian Derivative

where $\Delta A = A(t + \Delta t) - A(t)$. Therefore,

$$\text{In the limit as } \Delta t \to 0, \quad \frac{du}{dt} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial A}\frac{dA}{dt} \tag{4.23}$$

The Lagrangian derivative of u along any trajectory $A(t)$ in the $A - t$ plane is

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial A}\frac{dA}{dt}\bigg|_{trajectory} \tag{4.24}$$

It can be seen by comparing the coefficients of equation (4.8) that an appropriate choice of trajectory will eliminate the $\frac{\partial}{\partial A}$ terms. Choosing the paths $\mathcal{P}(A, t; S)$ such that

$$\frac{dA}{dt} = \frac{1}{t}(S - A) \tag{4.25}$$

gives an A-parameterised pricing PDE with identical spatial derivatives to Black-Scholes, namely

$$-\frac{du}{dt} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2} + rS\frac{\partial u}{\partial S} - ru = \mathcal{L}_{BS}\, u \tag{4.26}$$

The right-hand-side of this equation can be approximated in the usual way with second order finite differences, although the time integration needs special care.

4.3 Mesh Placement by Co-ordinate Transformation

Orthogonal co-ordinate transformations are used to obtain detailed resolution in regions of interest and coarsening in regions of least interest. For most options the region of trading interest is determined by the payoff. In the environment of Asian options, a suitable stretching transformation should be used in both the asset-price and average axes, enlarging the regions around S = strike and A = strike.

The co-ordinate transformations used are

$$S = \frac{K}{\lambda}\sinh(x_1 - L) + K \tag{4.27}$$

$$A = \frac{K}{\lambda}\sinh(x_2 - L) + K \tag{4.28}$$

where K is the strike price, x_1, x_2 are the transformed co-ordinates, L is a parameter controlling the degree of stretch, and $\lambda = \sinh L$.

In order to apply the transformation successfully to equation (4.8), some calculus and algebra are required. It has been shown earlier (see chapter 3) that

$$\begin{aligned} S\frac{\partial u}{\partial S} &= \mathcal{T}_\lambda(x_1)\frac{\partial u}{\partial x_1} \\ S^2\frac{\partial^2 u}{\partial S^2} &= \mathcal{T}_\lambda^2(x_1)\left(\frac{\partial^2 u}{\partial {x_1}^2} - \tanh(x_1 - L)\frac{\partial u}{\partial x_1}\right) \end{aligned}$$

where

$$\mathcal{T}_\lambda(x) = \frac{\sinh(x - L) + \lambda}{\cosh(x - L)}$$

Under co-ordinate transformations, equation (4.26) becomes

$$-\frac{du}{dt} = \frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)\left(\frac{\partial^2 u}{\partial {x_1}^2} - \tanh(x_1 - L)\frac{\partial u}{\partial x_1}\right) + r\mathcal{T}_\lambda(x_1)\frac{\partial u}{\partial x_1} - ru = \widehat{\mathcal{L}}_{BS}\, u \tag{4.29}$$

4.3.1 Trajectory Integration

The solution is carried out backwards in time, so beginning with values at T, values at t^{m-1} are obtained, and so on. The S-L integration from time t^m to t^{m-1} is along the path $\mathcal{P}_k^m(A, t; S_j) \equiv (\tilde{A}_k, t^m) \curvearrowright (A_k, t^{m-1})$. The trajectory is taken backwards *from a mesh point* A_k at time t^{m-1} to some departure point $\tilde{A}_k$ which is not necessarily a mesh point at time t^m. This is illustrated in Figure (4.3). Integrating along this trajectory gives

$$u(S_j, A_k, t^{m-1}) - u(S_j, \tilde{A}_k, t^m) = \Delta t \int_{\mathcal{P}_k^m(A,t;S_j)} \mathcal{L}_{BS}\, u(S, A, t)\, dt \tag{4.30}$$

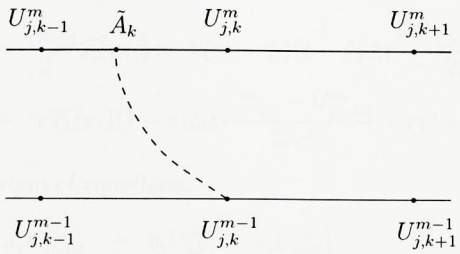


Figure 4.3: Typical integration trajectories in the $A-t$ plane

If $\mathcal{L}_j$ is an $O(h^2)$ approximation to $\mathcal{L}_{BS}$, where $h = \min(\Delta S)$, and $\{U^m_{j,k}\}$ are finite difference mesh prices on $\Omega = \{S_0, ...S_N\} \otimes \{A_0, ...A_N\}$ such that

$$U^{m-1}_{j,k} - \tilde{U}^m_{j,k} = \Delta t\Big(\theta\mathcal{L}_j(U^{m-1}_{j,k}) + (1-\theta)\mathcal{L}_j(\tilde{U}^m_{j,k})\Big) \tag{4.31}$$

then $U^m_{j,k} \approx u(S_j, A_k, t^m)$ to $O(\Delta t^2) + O(h^2)$ for $\theta = 0.5$ and is unconditionally stable for $\theta \geq 0.5$. $\tilde{U}^m_{jk}$ is the mesh price interpolated to $\tilde{A}_k$.

Under co-ordinate transformations, $U^m_{j,k} \approx u(x_{1_j}, x_{2_k}, t^m)$ denotes the finite difference approximation at the non-uniform mesh points (x_{1_j}, x_{2_k}, t^m). Equation (4.31) becomes

$$U^{m-1}_{j,k} - \tilde{U}^m_{j,k} = \Delta t\Big(\theta\widehat{\mathcal{L}}_j(U^{m-1}_{j,k}) + (1-\theta)\widehat{\mathcal{L}}_j(\tilde{U}^m_{j,k})\Big) \tag{4.32}$$

where $\widehat{\mathcal{L}}_j$ is the transformed finite difference approximation to the Black-Scholes pricing operator, as given in equation (4.29).

The solution is carried out backwards in time, so that values at t^m are known, from which the values at t^{m-1} have to be obtained. A set of tridiagonal equations are solved at each time-step (in the case of no early exercise). The governing equation follows from (4.32),

$$\begin{aligned}
U^{m-1}_{j,k} &- \frac{1}{2}\sigma^2\mathcal{T}^2_\lambda(x_1)(\theta\Delta t)\frac{U^{m-1}_{j+1,k} - 2U^{m-1}_{j,k} + U^{m-1}_{j-1,k}}{{\Delta x_1}^2} \\
&+ \frac{1}{2}\sigma^2\mathcal{T}^2_\lambda(x_1)\tanh(x_1 - L)(\theta\Delta t)\frac{U^{m-1}_{j+1,k} - U^{m-1}_{j-1,k}}{2\Delta x_1} \\
&- r\mathcal{T}_\lambda(x_1)(\theta\Delta t)\frac{U^{m-1}_{j+1,k} - U^{m-1}_{j-1,k}}{2\Delta x_1} + r(\theta\Delta t)U^{m-1}_{j,k} \\
= \tilde{U}^m_{j,k} &+ \frac{1}{2}\sigma^2\mathcal{T}^2_\lambda(x_1)(1-\theta)\Delta t\frac{\tilde{U}^m_{j+1,k} - 2\tilde{U}^m_{j,k} + \tilde{U}^m_{j-1,k}}{{\Delta x_1}^2}
\end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)\tanh(x_1-L)(1-\theta)\Delta t\frac{\tilde{U}^m_{j+1,k}-\tilde{U}^m_{j-1,k}}{2\Delta x_1}\\
&+ r\mathcal{T}_\lambda(x_1)(1-\theta)\Delta t\frac{\tilde{U}^m_{j+1,k}-\tilde{U}^m_{j-1,k}}{2\Delta x_1} - r(1-\theta)\Delta t\tilde{U}^m_{j,k} \qquad (4.33)
\end{aligned}
$$

which leads to a system of equations:

$$
\begin{aligned}
a_j U^{m-1}_{j-1,k} &+ b_j U^{m-1}_{j,k} + c_j U^{m-1}_{j+1,k}\\
&= aa_j\tilde{U}^m_{j-1,k} + bb_j\tilde{U}^m_{j,k} + cc_j\tilde{U}^m_{j+1,k} \qquad (4.34)
\end{aligned}
$$

where

$$
\begin{aligned}
a_j &= -(\theta\Delta t)\left(\frac{\frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)}{\Delta {x_1}^2} + \frac{\frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)\tanh(x_1-L)}{2\Delta x_1} - \frac{\frac{1}{2}r\mathcal{T}_\lambda(x_1)}{\Delta x_1}\right)\\
b_j &= 1-(\theta\Delta t)\left(\frac{-\sigma^2\mathcal{T}_\lambda^2(x_1)}{\Delta {x_1}^2} - r\right)\\
c_j &= -(\theta\Delta t)\left(\frac{\frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)}{\Delta {x_1}^2} - \frac{\frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)\tanh(x_1-L)}{2\Delta x_1} + \frac{\frac{1}{2}r\mathcal{T}_\lambda(x_1)}{\Delta x_1}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
aa_j &= (1-\theta)\Delta t\left(\frac{\frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)}{\Delta {x_1}^2} + \frac{\frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)\tanh(x_1-L)}{2\Delta x_1} - \frac{\frac{1}{2}r\mathcal{T}_\lambda(x_1)}{\Delta x_1}\right)\\
bb_j &= 1+(1-\theta)\Delta t\left(\frac{-\sigma^2\mathcal{T}_\lambda^2(x_1)}{\Delta {x_1}^2} - r\right)\\
cc_j &= (1-\theta)\Delta t\left(\frac{\frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)}{\Delta {x_1}^2} - \frac{\frac{1}{2}\sigma^2\mathcal{T}_\lambda^2(x_1)\tanh(x_1-L)}{2\Delta x_1} + \frac{\frac{1}{2}r\mathcal{T}_\lambda(x_1)}{\Delta x_1}\right)
\end{aligned}
$$

The right hand side values, $d^m_{j,k}$ are defined by

$$
d^m_{j,k} = aa^m_j\tilde{U}^m_{j-1,k} + bb^m_j\tilde{U}^m_{j,k} + cc^m_j\tilde{U}^m_{j+1,k} \qquad (4.35)
$$

4.3.2 Determining $\tilde{A}_k$

Trajectories $A = A(t)$ are chosen such that

$$
\frac{dA}{dt} = \frac{S-A}{t}
$$

Separating the variables, and integrating with respect to time along the path $\mathcal{P}^m_k(A,t;S)$, gives

$$
\int_{\tilde{A}_k}^{A_k}\frac{1}{S_j-A}dA = \int_{t^m}^{t^{m-1}}\frac{1}{t}dt \qquad (4.36)
$$

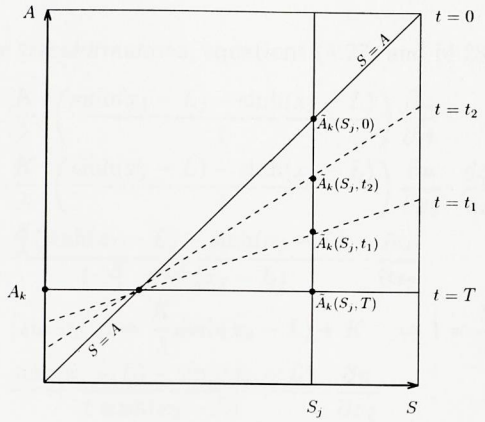


Figure 4.4: $\tilde{A}_k$ dependence on S at different times

(keep in mind that S_j is constant). That is,

$$\int_{\tilde{A}_k}^{A_k} \frac{1}{S_j - A} dA = \int_{t^m}^{t^{m-1}} \frac{1}{t} dt \tag{4.37}$$

which gives,

$$\begin{aligned} -\Big[\ln|S_j - A|\Big]_{\tilde{A}_k}^{A_k} &= \Big[\ln|t|\Big]_{t^m}^{t^{m-1}} \\ \Rightarrow \; -\Big[\ln|S_j - A_k| - \ln|S_j - \tilde{A}_k|\Big] &= \big[\ln|t^{m-1}| - \ln|t^m|\big] \\ \Rightarrow \; \ln\left|\frac{S_j - \tilde{A}_k}{S_j - A_k}\right| &= \ln\left|\frac{t^{m-1}}{t^m}\right| \end{aligned}$$

so that

$$\begin{aligned} \tilde{A}_k &= S_j - \left(\frac{t^{m-1}}{t^m}\right)(S_j - A_k) \\ &= (1-\alpha)S_j + \alpha A_k \end{aligned} \tag{4.38}$$

where

$$\alpha = \frac{t^{m-1}}{t^m}$$

Figure 4.4 shows the dependence of $\tilde{A}_k$ on S at different times.

Under co-ordinate transformations, equations (4.27) and (4.28) give

$$\begin{aligned}
\frac{1}{t}(S-A)\frac{\partial u}{\partial A} &= \frac{K}{\lambda}\left(\frac{\sinh(x_1-L)-\sinh(x_2-L)}{t}\right)\frac{\partial u}{\partial A} \\
&= \frac{K}{\lambda}\left(\frac{\sinh(x_1-L)-\sinh(x_2-L)}{t}\right)\frac{\partial u}{\partial x_2}\cdot\frac{dx_2}{dA} \\
&= \frac{\frac{K}{\lambda}(\sinh(x_1-L)-\sinh(x_2-L))}{t\cdot\frac{K}{\lambda}\cdot\cosh(x_2-L)}\cdot\frac{\partial u}{\partial x_2} \\
&\quad \lceil \text{since, } A=\frac{K}{\lambda}\sinh(x_2-L)+K \quad \Rightarrow 1=\frac{K}{\lambda}\cosh(x_2-L)\cdot\frac{dx_2}{dA} \rfloor \\
\therefore \frac{1}{t}(S-A)\frac{\partial u}{\partial A} &= \frac{\sinh(x_1-L)-\sinh(x_2-L)}{t\cosh(x_2-L)}\cdot\frac{\partial u}{\partial x_2}
\end{aligned}$$

In the transformed context, the Lagrangian derivative is

$$\frac{du}{dt}=\frac{\partial u}{\partial t}+\frac{\partial u}{\partial x_2}\frac{dx_2}{dt}$$

and the trajectory chosen is now

$$\frac{dx_2}{dt}=\frac{\sinh(x_1-L)-\sinh(x_2-L)}{t\cosh(x_2-L)}$$

Separate the variables and integrate over a time-step, as before

$$\int_{\tilde{x_{2k}}}^{x_{2k}}\frac{\cosh(x_2-L)}{\sinh(x_1-L)-\sinh(x_2-L)}\cdot dx_2=\int_{t^m}^{t^{m-1}}\frac{1}{t}\cdot dt$$

Using the substitution $S'=\sinh(x_1-L)$ and $A'=\sinh(x_2-L)$ gives

$$\int_{\tilde{A'_k}}^{A'_k}\frac{dA'}{S'-A'}=\int_{t^m}^{t^{m-1}}\frac{1}{t}\cdot dt$$

which implies that (see section 4.3.2)

$$\tilde{A'_k}=(1-\alpha)S'_j+\alpha A'_k$$

where

$$\alpha=\frac{t^{m-1}}{t^m}$$

Substituting back for S' and A' gives

$$\sinh(\tilde{x_{2k}}-L)=(1-\alpha)\sinh(x_{1j}-L)+\alpha\sinh(x_{2k}-L)$$

which implies

$$\begin{aligned}\tilde{x_{2k}} - L &= \sinh^{-1}\left((1-\alpha)\sinh(x_{1j} - L) + \alpha\sinh(x_{2k} - L)\right) \\ &= \sinh^{-1}\left((1-\alpha)S'_j + \alpha A'_k\right) \\ &= \sinh^{-1}(\tilde{A'_k})\end{aligned}$$

Substituting $\tilde{A'_k} = \sinh(\tilde{x_{2k}} - L)$ into equation (4.28) gives

$$\begin{aligned}\tilde{A_k} &= \frac{K}{\lambda}\tilde{A'_k} + K \\ \Rightarrow \tilde{A'_k} &= \frac{\lambda}{K}(\tilde{A_k} - K)\end{aligned}$$

which gives that

$$\tilde{x_{2k}} = \sinh^{-1}\left(\frac{\lambda}{K}(\tilde{A_k} - K)\right) + L \tag{4.39}$$

Values at trajectory endpoints are obtained by cubic-spline interpolation in the A-direction. Since these interpolated values are not readily differenced in the S-direction the combined t^m-term is regarded as a function and interpolated rather than the approximate prices. The trajectory endpoints can be evaluated analytically and are always contained within the computational region.

4.3.3 Truncation Conditions

The original domain is the infinite quarter plain and must be truncated before it can be meshed: e.g. at $S = S_{\max}$ and $A = A_{\max}$. The problem is now solved on the truncated region

$$[0, S_{\max}] \times [0, A_{\max}] \times (0, T]$$

The asymptotic boundary condition at $S \to \infty$ is replaced by,

$$\frac{\partial^2 U(S_{\max}, A, t)}{\partial S^2} = 0 \qquad \forall\, A \in [0, A_{\max}]; \quad t \in (0, T]$$

Recall that boundary conditions are not needed in the A-direction. At $S_{\max}$, equation (4.26) reduces to the following:

$$-\frac{dU}{dt} = rS\frac{\partial U}{\partial S} - rU$$

which can then be expressed in terms of the co-ordinate transformations as

$$-\frac{dU}{dt} = r\mathcal{T}_\lambda(x_{1_{\max}})\frac{\partial U}{\partial x} - rU$$

The one-sided finite difference operator is used for the drift term, resulting in the coefficients at the S-boundary given as:

$$\begin{aligned} a_{x_{\max}} &= \frac{r\mathcal{T}_\lambda(x_{\max})(\theta\Delta t)}{dx_1} \\ b_{x_{\max}} &= 1 - r(\theta\Delta t)\left(\frac{\mathcal{T}_\lambda(x_{\max})}{dx_1} - 1\right) \\ c_{x_{\max}} &= 0 \end{aligned}$$

and

$$\begin{aligned} aa_{x_{\max}} &= -\frac{r\mathcal{T}_\lambda(x_{\max})(1-\theta)\Delta t}{dx_1} \\ bb_{x_{\max}} &= 1 + r(1-\theta)\Delta t\left(\frac{\mathcal{T}_\lambda(x_{\max})}{dx_1} - 1\right) \\ cc_{x_{\max}} &= 0 \end{aligned}$$

It is noted here that the diffusion term in the Black-Scholes equation

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2}$$

gets transformed to

$$\frac{1}{2}\sigma^2 \mathcal{T}_\lambda^2(x_1)\left(\frac{\partial^2 U}{\partial {x_1}^2} - \tanh(x_1 - L)\frac{\partial U}{\partial x_1}\right)$$

which is approximated by finite differences as

$$a2 * U_{x_{\max}-1} + b2 * U_{x_{\max}} + c2 * U_{x_{\max}+1}$$

where

$$\begin{aligned} a2 &= \frac{\sigma^2\mathcal{T}_\lambda^2(x_1)}{2{dx_1}^2} + \frac{\sigma^2\mathcal{T}_\lambda^2(x_1)\tanh(x_1 - L)}{4dx_1} \\ b2 &= -\frac{\sigma^2\mathcal{T}_\lambda^2(x_1)}{{dx_1}^2} \\ c2 &= \frac{\sigma^2\mathcal{T}_\lambda^2(x_1)}{2{dx_1}^2} - \frac{\sigma^2\mathcal{T}_\lambda^2(x_1)\tanh(x_1 - L)}{4dx_1} \end{aligned}$$

For Asian options, where the solution is observed to rotate progressively along the diagonal, a more appropriate boundary condition at $S_{\max}$ is found to be

$$\frac{\partial^2 U}{\partial l^2} = 0$$

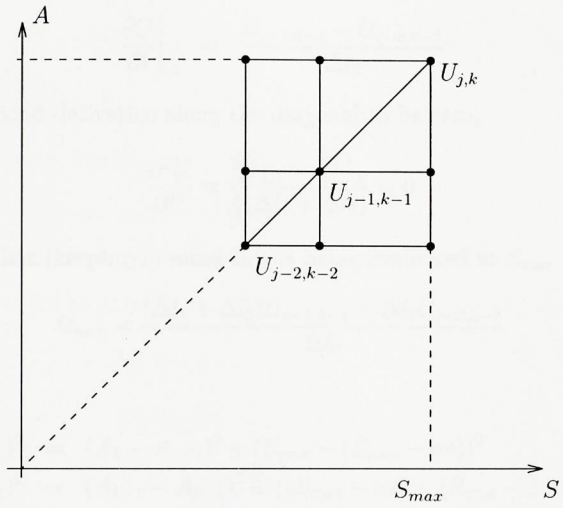


Figure 4.5: At $S_{\max}$, the second derivative along the diagonal is set to be equal to zero.

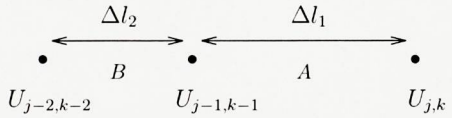


Figure 4.6: Figure showing which mesh points are used when differencing along the diagonal.

where l is the distance between mesh points along the diagonal. The second derivative is more appropriately considered along the diagonal. For a uniform mesh, $\Delta l_1 = \Delta l_2$. This is a simple linear extrapolation (See Figure 4.6).

$$\begin{aligned} \frac{\partial^2 U}{\partial l^2} &\approx \frac{U_{j,k} - 2U_{j-1,k-1} + U_{j-2,k-2}}{\Delta l^2} = 0 \\ \Rightarrow U_{j,k} &= 2u_{j-1,k-1} - U_{j-2,k-2} \end{aligned}$$

For a non-uniform mesh, this is weighted (Figure 4.5). Firstly,

$$\left.\frac{\partial U}{\partial l}\right|_A \approx \frac{U_{j,k} - U_{j-1,k-1}}{\Delta l_1}$$

$$\left.\frac{\partial U}{\partial l}\right|_B \approx \frac{U_{j-1,k-1} - U_{j-2,k-2}}{\Delta l_2}$$

Now set the second derivative along the diagonal to be zero,

$$\frac{\partial^2 U}{\partial l^2} \approx \frac{\left.\frac{\partial U}{\partial l}\right|_A - \left.\frac{\partial U}{\partial l}\right|_B}{\frac{1}{2}(\Delta l_1 + \Delta l_2)} = 0$$

which implies that (keeping in mind this is being evaluated at $S_{\max}$ for all $k > 2$),

$$U_{ns,k} = \frac{(\Delta l_1 + \Delta l_2)U_{j-1,k-1} - \Delta l_1 U_{j-2,k-2}}{\Delta l_2}$$

where

$$\begin{aligned}(\Delta l_1)^2 &= (A_k - A_{k-1})^2 + (S_{\max} - (S_{\max} - ns))^2 \\ (\Delta l_2)^2 &= (A_{k-1} - A_{k-2})^2 + ((S_{\max} - ns) - (S_{\max} - 2\,ns))^2\end{aligned}$$

4.4 S-L Algorithm for the Asian Option Solution

For every time-step Δt

For every asset-price step S_j

Define cubic spline interpolant:
$d_j^m = I_{CS}(d_{j,0}^m, d_{j,1}^m, d_{j,2}^m, \cdots, d_{j,N_k}^m)$

Interpolate in k, average price direction

For every $k-$step

Evaluate $\tilde{A}_k(S_j, A_k, t^m)$

$\tilde{U}_{j,k}^m \longleftarrow I_{CS}(\tilde{A}_{j,k}^m)$

End For

End For

For every $k-$step

Solve in $j-$direction:

Either a tridiagonal system,
$U_{j,k}^{m-1} \longleftarrow \tilde{d}_{j,k}^m$ [Ref: Equation (4.33)]

Or, if early exercise features,
apply PSOR or use the penalty method.

End For

For every $j-$step

For every $k-$step

$U^m \longleftarrow U^{m-1}$

End For

End For

End For

4.5 Results and Conclusions for European Asians

Numerical results are now presented. Just as in previous chapters, the desired level of accuracy is set to be 0.1% of S_0. For example, in a case where $S = 10$, this requirement is equivalent to 'penny accuracy'.

4.5.1 Numerical Convergence

Table 4.1 displays results corresponding to the fixed strike European call Asian when $r = 0.1$, and prices quoted at $S_0 = 100$. σ, K and expiry, T vary as shown in the table. The solution is calculated as the grid is refined. The solutions at the discrete asset prices were obtained by interpolating between grid points. Richardson extrapolations are also shown here. This extrapolation works well and gives basis point accuracy using meshes of 40 and 80 points. The results suggest that the scheme is convergent. Further, the desired level of accuracy can be obtained on a 80×80 mesh using a stretching parameter of ten, and 10 time steps. Note that the number of time steps is relatively small since the S-L method is increasingly diffusive as the number of time steps becomes large.

Corresponding results for the fixed strike European put Asian is shown is Table 4.2. Again, the solution is calculated as the grid is refined. The solutions at the discrete asset prices were obtained by interpolating between grid points. Richardson extrapolations are also shown here. This extrapolation works well and gives basis point accuracy using meshes of 40 and 80 points. The results suggest that the scheme is convergent. Further, the desired level of accuracy can again be obtained on a 80×80 mesh using a stretching parameter of ten, and 10 time steps.

A slightly different problem is displayed in Table 4.3. The fixed strike European call Asian is again considered. The parameters of this problem are $r = 0.15$, $T = 1$ year. Prices are quoted at $S_0 = 100$. σ and K vary as shown in the table. The solution is calculated as the grid is refined. The solutions at the discrete asset prices were obtained by interpolating between grid points. Richardson extrapolations are also shown here. This extrapolation works well and gives basis point accuracy using meshes of 40 and 80 points. The results suggest that the scheme is convergent. Further, the desired level of accuracy can be obtained on a 80×80 mesh using a stretching parameter of ten, and 10 time steps.

4.5.2 Surface Plots

For a fixed-strike Asian put option, Figure 4.7 shows the surface plots of the evolution of the solution over time. Figure 4.8(a) displays the initial data for the Asian

put problem when $K = 100$. As time progresses, the solution is seen to rotate completely along the $S = A$ diagonal: it starts as a hockey stick shaped plane (with a singularity) in the average direction only, and results in a hockey stick shaped plane in the spot direction only. A similar evolution can be seen in the case of fixed-strike Asian call options (Figure 4.8). This figure displays the initial data for the Asian call problem when $K = 100$, and displays the evolution of the solution.

4.5.3 Numerical Tests

The numerical algorithm is tested by comparing the results with those obtained by others. Results obtained with the S-L method are compared to published results [BP96], [RS95], [ZFV98a], [Vec01] where available. It is noted here that the published results themselves do not agree with each other, making it difficult to conduct a concrete validation against them. However, the S-L results agree with those published within the same ballpark as they disagree. The S-L method has been shown to be rigorous and robust, and convergence tests suggest it to work extremely well for the problems considered. The results quoted for the S-L case are those obtained by using Richardson extrapolation on 40×40 and 80×80 mesh discrete prices. In Table 4.4 for fixed strike European call Asians, S-L results are compared to those in [RS95], [ZFV98a], [BP96]. The comparisons are seen to be satisfactory.

In Table 4.5 for fixed strike European put Asians, S-L results are compared to those in [BP96]. The comparisons are seen to be satisfactory.

Table 4.6 for fixed strike European call Asians displays S-L results compared to those in [RS95], [ZFV98a], [Vec01]. The comparisons are seen to be satisfactory.

In Table 4.7 the mean execution time is displayed along with the solution grid size used. Note that the number of time steps is relatively small since the S-L method is increasingly diffusive as the number of time steps becomes large. Average CPU times are shown since the performance of the iterative method varies with the volatility chosen.

Table 4.1: Convergence results for Fixed Strike Call Asians where $r = 0.1$, at a spot price of $S_0 = 100$. The Semi-Lagrangian (S-L) results were calculated on meshes of size 20x20, 40x40, 80x80 and 160x160; the Richardson extrapolated results are also shown. nt, the number of time steps used, is 10 for all maturities. λ, the stretching parameter, is also 10.

σ	T	K	Calculated by S-L				Richardson Extrapolation		
			20x20	40x40	80x80	160x160	(20,40)	(40,80)	(80,160)
0.10	0.25	95	6.146	6.122	6.120	6.119	6.114	6.119	6.119
		100	1.679	1.816	1.844	1.850	1.861	1.853	1.852
		105	0.143	0.150	0.151	0.152	0.152	0.152	0.152
	0.50	95	7.246	7.225	7.221	7.220	7.218	7.220	7.220
		100	2.987	3.054	3.067	3.070	3.076	3.071	3.071
		105	0.685	0.706	0.714	0.717	0.713	0.717	0.717
	1.00	95	9.343	9.299	9.288	9.286	9.285	9.285	9.285
		100	5.257	5.256	5.255	5.255	5.256	5.255	5.255
		105	2.264	2.289	2.296	2.298	2.297	2.298	2.298

σ	T	K	Calculated by S-L				Richardson Extrapolation		
			20x20	40x40	80x80	160x160	(20,40)	(40,80)	(80,160)
0.20	0.25	95	6.462	6.474	6.477	6.478	6.478	6.478	6.478
		100	2.884	2.922	2.931	2.933	2.935	2.934	2.934
		105	0.920	0.942	0.948	0.950	0.949	0.951	0.951
	0.50	95	7.898	7.894	7.893	7.893	7.893	7.893	7.893
		100	4.502	4.505	4.506	4.506	4.506	4.506	4.506
		105	2.197	2.206	2.208	2.209	2.209	2.209	2.209
	1.00	95	10.338	10.305	10.297	10.295	10.295	10.294	10.294
		100	7.086	7.054	7.046	7.044	7.043	7.043	7.043
		105	4.534	4.518	4.513	4.512	4.512	4.512	4.512
0.40	0.25	95	8.096	8.102	8.104	8.104	8.104	8.104	8.104
		100	5.174	5.171	5.171	5.170	5.170	5.170	5.170
		105	3.067	3.066	3.066	3.066	3.066	3.066	3.066
	0.50	95	10.357	10.351	10.349	10.349	10.349	10.349	10.349
		100	7.600	7.581	7.576	7.575	7.575	7.575	7.575
		105	5.398	5.383	5.378	5.377	5.377	5.377	5.377
	1.00	95	13.751	13.733	13.729	13.728	13.727	13.727	13.727
		100	11.172	11.139	11.131	11.129	11.129	11.129	11.129
		105	8.961	8.930	8.921	8.919	8.919	8.918	8.918

Table 4.2: Convergence results for Fixed Strike Put Asians where $r = 0.1$, at a spot price of $S_0 = 100$. The Semi-Lagrangian (S-L) results were calculated on meshes of size 20x20, 40x40, 80x80 and 160x160; the Richardson extrapolated results are also shown. nt, the number of time steps used, is 10 for all maturities. λ, the stretching parameter, is also 10.

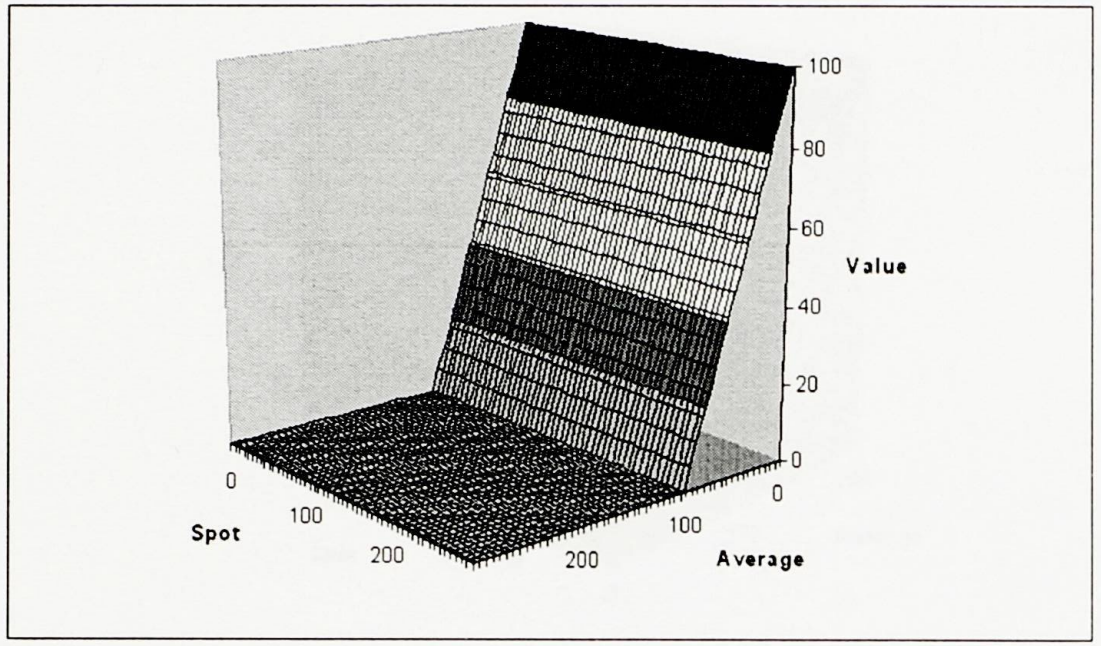
σ	T	K	Calculated by S-L				Richardson Extrapolation		
			20x20	40x40	80x80	160x160	(20,40)	(40,80)	(80,160)
0.10	0.25	95	0.020	0.012	0.013	0.013	0.009	0.013	0.013
		100	0.428	0.581	0.613	0.621	0.632	0.624	0.623
		105	3.771	3.792	3.797	3.799	3.799	3.799	3.799
	0.50	95	0.031	0.041	0.044	0.045	0.044	0.046	0.046
		100	0.525	0.625	0.646	0.652	0.658	0.654	0.653
		105	2.984	3.033	3.050	3.054	3.050	3.055	3.056
	1.00	95	0.062	0.078	0.082	0.083	0.083	0.083	0.083
		100	0.495	0.558	0.572	0.576	0.578	0.577	0.577
		105	2.037	2.115	2.137	2.143	2.141	2.145	2.145

σ	T	K	Calculated by S-L				Richardson Extrapolation		
			20x20	40x40	80x80	160x160	(20,40)	(40,80)	(80,160)
0.20	0.25	95	0.342	0.364	0.370	0.372	0.372	0.372	0.372
		100	1.634	1.687	1.700	1.703	1.705	1.704	1.704
		105	4.544	4.583	4.594	4.597	4.595	4.598	4.598
	0.50	95	0.694	0.713	0.718	0.719	0.719	0.719	0.719
		100	2.043	2.077	2.085	2.088	2.088	2.088	2.088
		105	4.493	4.533	4.544	4.546	4.546	4.547	4.547
	1.00	95	1.074	1.088	1.092	1.093	1.093	1.093	1.093
		100	2.331	2.357	2.364	2.365	2.366	2.366	2.366
		105	4.305	4.344	4.355	4.357	4.357	4.358	4.358
0.40	0.25	95	1.990	1.996	1.998	1.998	1.998	1.999	1.998
		100	3.928	3.938	3.940	3.941	3.941	3.941	3.941
		105	6.688	6.706	6.711	6.713	6.712	6.713	6.713
	0.50	95	3.177	3.175	3.175	3.175	3.175	3.175	3.175
		100	5.152	5.156	5.157	5.157	5.157	5.157	5.157
		105	7.694	7.709	7.714	7.715	7.715	7.715	7.715
	1.00	95	4.527	4.526	4.526	4.526	4.526	4.526	4.526
		100	6.441	6.449	6.451	6.451	6.451	6.451	6.451
		105	8.738	8.758	8.763	8.765	8.764	8.765	8.765

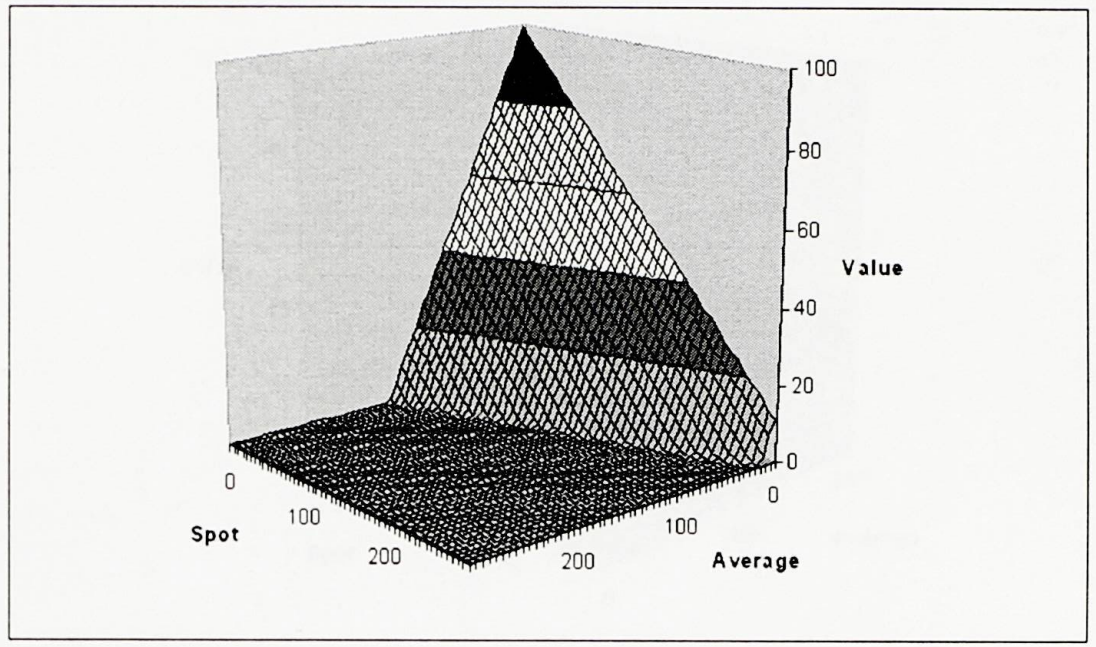
Table 4.3: Convergence results for Fixed Strike Call Asians where $r = 0.15$, $T = 1$ year, at a spot price of $S_0 = 100$. The Semi-Lagrangian (S-L) results were calculated on meshes of size 20x20, 40x40, 80x80 and 160x160; the Richardson extrapolated results are also shown. nt, the number of time steps used, is 10 for all maturities. λ, the stretching parameter, is 15. For comparison, three other published results are also indicated.

σ	K	Calculated by S-L				Richardson Extrapolation		
		20x20	40x40	80x80	160x160	(20,40)	(40,80)	(80,160)
0.05	95	11.251	11.132	11.101	11.093	11.092	11.091	11.091
	100	6.933	6.825	6.800	6.793	6.790	6.791	6.791
	105	2.611	2.708	2.731	2.737	2.740	2.739	2.738
0.10	90	15.548	15.433	15.405	15.398	15.395	15.395	15.395
	100	7.117	7.051	7.032	7.027	7.029	7.026	7.026
	110	1.403	1.418	1.423	1.424	1.423	1.425	1.425
0.20	90	15.742	15.665	15.646	15.641	15.640	15.639	15.639
	100	8.496	8.433	8.415	8.411	8.412	8.409	8.409
	110	3.590	3.569	3.563	3.561	3.562	3.561	3.561
0.30	90	16.570	16.527	16.516	16.513	16.513	16.512	16.512
	100	10.287	10.232	10.217	10.213	10.214	10.212	10.212
	110	5.788	5.750	5.739	5.736	5.737	5.735	5.735

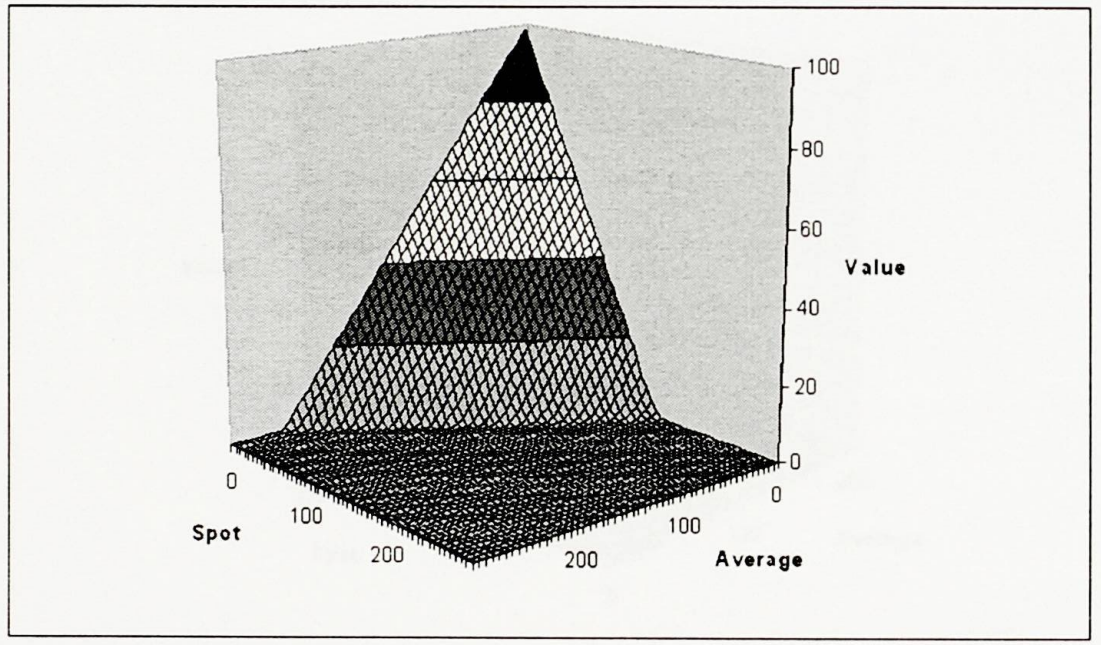
Figure 4.7: Evolution over time of the solution for a fixed-strike European put option. Model parameters are $K = 100$, $r = 0.1, \sigma = 0.2, T = 3$ months.



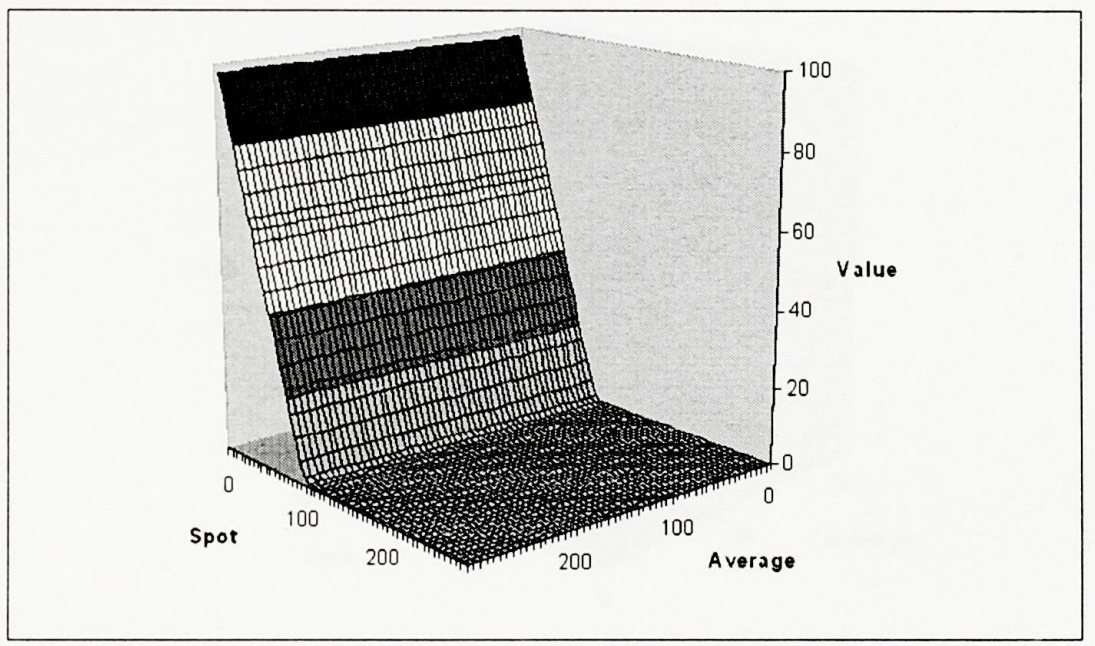
(a) At time = expiry (3 months).



(b) At time = 2 months.

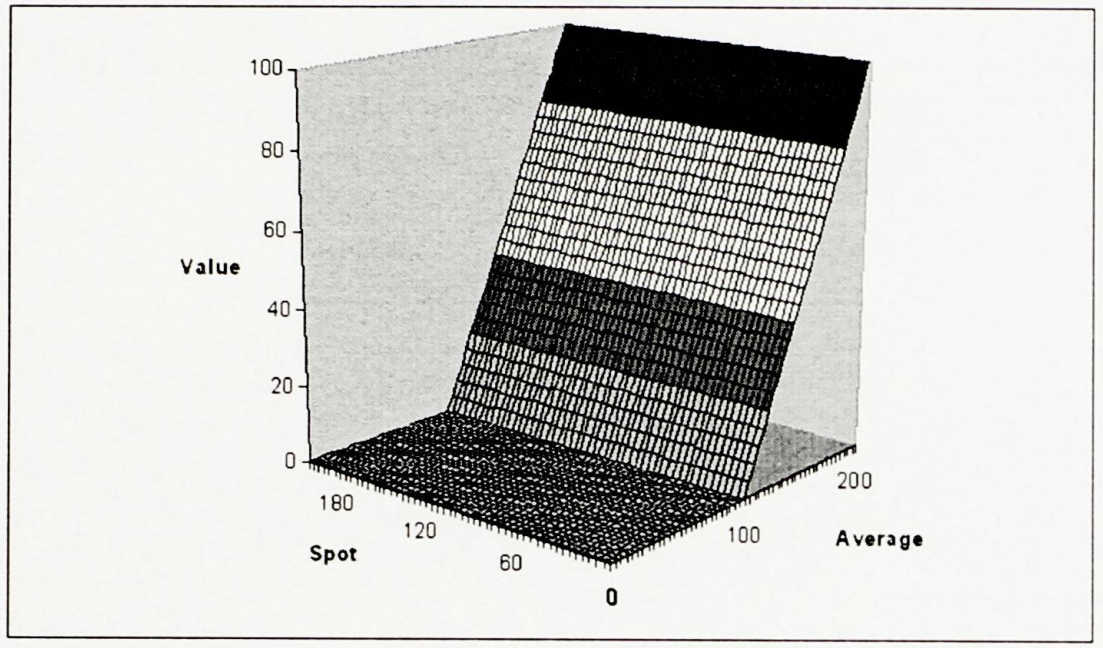


(c) At time = 1 month.

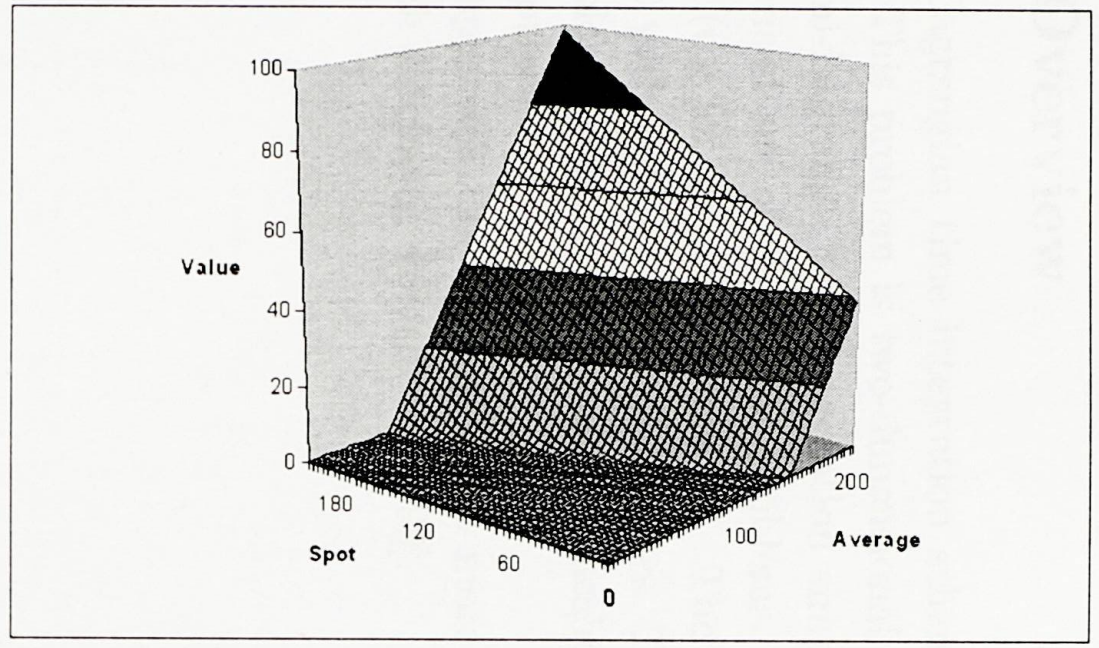


(d) At time = 0.

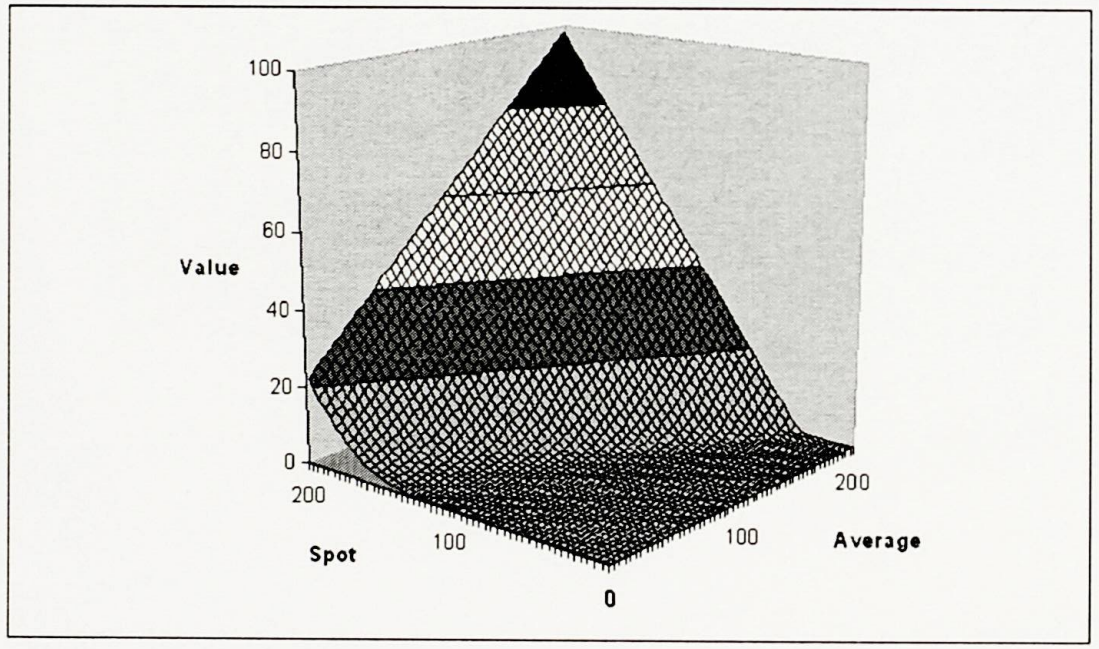
Figure 4.8: Evolution over time of the solution for a fixed-strike European call option. Model parameters are $K = 100$, $r = 0.15, \sigma = 0.4, T = 6$ months.



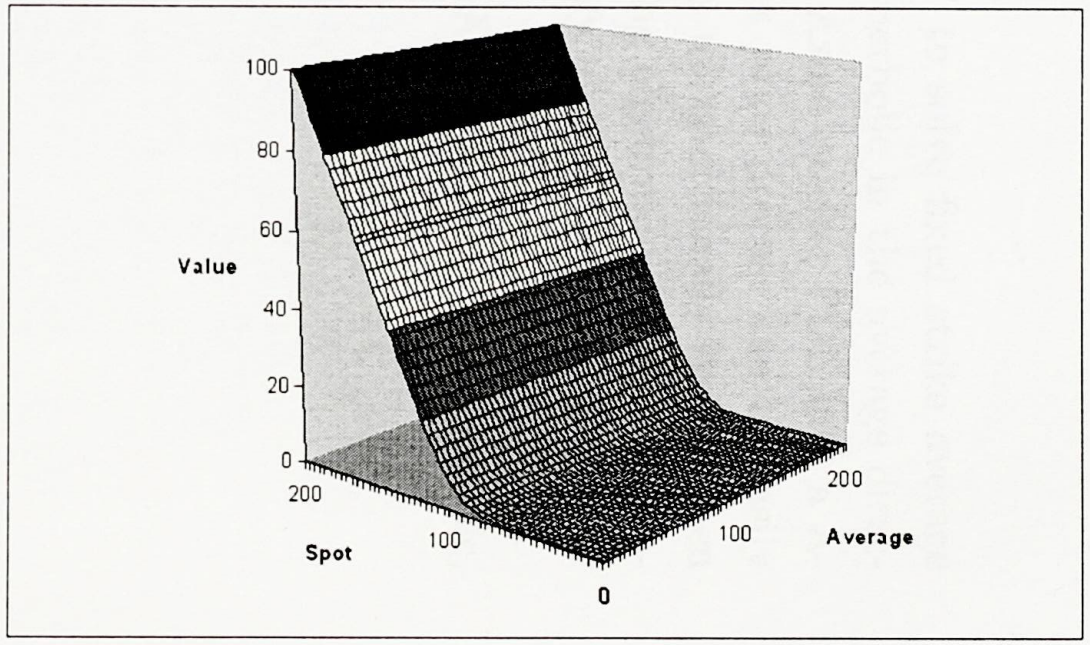
(a) At time = expiry (6 months).



(b) At time = 4 months.



(c) At time = 2 months.



(d) At time = 0.

4.6 Overview

A Semi-Lagrangian time integration scheme is used to solve fixed strike average options. This problem is two-dimensional and is hyperbolic in the average direction. Semi-Lagrange time integration simplifies the Asian pricing PDE into a A-parameterised set of one-factor problems. The asset price process can be easily extended (e.g for volatility surfaces). The method is unconditionally stable when combined with implicit finite differences. The resulting algebraic problems are linear and block tridiagonal. Stretched meshes are easily incorporated and give very accurate prices.

This approach can be applied to more complex Asian options. For example, assets with stochastic volatility [CP99].

Table 4.4: Calculated Semi-Lagrangian (S-L) results for Fixed Strike Call Asians where $r = 0.1$, at a spot price of $S_0 = 100$. For comparison, three other published results are also indicated. nt, the number of time steps used, is 10 for all maturities. λ, the stretching parameter, is also 10.

σ	T	K	S-L	[RS95]	[ZFV98a]	[BP96]
0.10	0.25	95	6.119	6.118	6.133	6.132
		100	1.853	1.851	1.793	1.869
		105	0.152	0.148	0.162	0.151
	0.50	95	7.220	7.22	7.244	7.248
		100	3.071	3.104	3.052	3.1
		105	0.717	0.714	0.726	0.727
	1.00	95	9.285	9.285	9.316	9.313
		100	5.255	5.255	5.261	5.279
		105	2.298	2.294	2.314	2.313
0.20	0.25	95	6.478	6.476	6.501	6.5
		100	2.934	2.932	2.928	2.96
		105	0.951	0.947	0.971	0.966
	0.50	95	7.893	7.891	7.921	7.793
		100	4.506	4.505	4.511	4.548
		105	2.209	2.211	2.229	2.241
	1.00	95	10.294	10.295	10.309	10.336
		100	7.043	7.042	7.042	7.079
		105	4.512	4.509	4.519	4.539
0.40	0.25	95	8.104	8.102	8.123	8.151
		100	5.170	5.168	5.175	5.218
		105	3.066	3.063	3.082	3.106
	0.50	95	10.349	10.346	10.357	10.425
		100	7.575	7.572	7.574	7.65
		105	5.377	5.371	5.384	5.444
	1.00	95	13.727	13.721	13.721	13.825
		100	11.129	11.121	11.115	11.213
		105	8.918	8.91	8.912	8.989

Table 4.5: Calculated Semi-Lagrangian (S-L) results for Fixed Strike Put Asians where $r = 0.1$, at a spot price of $S_0 = 100$. For comparison, one other published result is also indicated. nt, the number of time steps used, is 10 for all maturities. λ, the stretching parameter, is also 10.

σ	T	K	S-L	[BP96]
0.10	0.25	95	0.013	0.013
		100	0.624	0.626
		105	3.799	3.785
	0.50	95	0.046	0.046
		100	0.654	0.655
		105	3.055	3.039
	1.00	95	0.083	0.084
		100	0.577	0.577
		105	2.145	2.137
0.20	0.25	95	0.372	0.379
		100	1.704	1.716
		105	4.598	4.598
	0.50	95	0.719	0.731
		100	2.088	2.102
		105	4.547	4.552
	1.00	95	1.093	1.099
		100	2.366	2.369
		105	4.358	4.356
0.40	0.25	95	1.999	2.025
		100	3.941	3.97
		105	6.713	6.735
	0.50	95	3.175	3.215
		100	5.157	5.197
		105	7.715	7.748
	1.00	95	4.526	4.55
		100	6.451	6.465
		105	8.765	8.767

Table 4.6: Calculated Semi-Lagrangian (S-L) results for Fixed Strike Call Asians where $r = 0.15$, $T = 1$ year, at a spot price of $S_0 = 100$. nt, the number of time steps used, is 10 for all maturities. λ, the stretching parameter, is 15. For comparison, three other published results are also indicated.

σ	K	S-L	[Vec01]	[RS95]	[ZFV98a]
0.05	95	11.091	11.094	11.094	11.094
	100	6.791	6.795	6.794	6.793
	105	2.739	2.744	2.744	2.744
0.10	90	15.395	15.399	15.399	15.399
	100	7.026	7.029	7.028	7.030
	110	1.425	1.415	1.413	1.410
0.20	90	15.639	15.643	15.641	15.643
	100	8.409	8.410	8.408	8.409
	110	3.561	3.558	3.554	3.554
0.30	90	16.512	16.515	16.512	16.514
	100	10.212	10.213	10.208	10.210
	110	5.735	5.734	5.728	5.729

Table 4.7: Average times required to obtain a result for varying mesh size, when a relaxation solver is used. nt, the number of time steps used, is 10 for all maturities. λ, the stretching parameter, is also 10.

Size of Grid	Mean execution time (sec)
40×40	0.8
80×80	3.1
160×160	14.4

Chapter 5

Asian Options with Early Exercise

The previous chapter detailed the solution method for pricing Asian options using a Semi-Lagrange time integration method. Early exercise is also easily incorporated.

Early exercise options allow the holder to exercise (that is, buy for a call, sell for a put) before maturity. This is assuming that there is a well-defined payoff for early exercise. It has already been shown in section 1.4.2 that the American option has a potentially higher value than the equivalent European option. (This is since the American option gives its holder greater rights than the European option, due to its right of early exercise). Some discussion on the American put option has already been detailed in the thesis. The reader is asked to refer to section 2.5.5.

If the Asian option permits early exercise then the co-ordinate transformed, A-parameterised pricing PDE (4.29) is now replaced by a partial differential inequality

$$-\frac{du}{dt} \geq \widehat{\mathcal{L}}_{BS}\, u \tag{5.1}$$

in $\Omega = \{x_1, x_2 : x_1 \geq 0, x_2 \geq 0\} \times [0, T]$ where $\widehat{\mathcal{L}}_{BS}$ is given in (4.29) and final condition $u(S, A, T) = g(S, A, T)$ applies.

If the option is exercised, then its value is determined by the payoff. If the option is held, then its value must be greater than the payoff to preclude arbitrage opportunities. Therefore,

$$u(S, A, t) \geq g(S, A, t) \; \forall \; (S, A, t) \in \Omega \times [0, T] \tag{5.2}$$

Conditions (5.1) and (5.2) combine into a linear complementarity problem

$$(u(S, A, t) - g(S, A, t)) \times \; (\; -\frac{du}{dt} - \widehat{\mathcal{L}}_{BS}) = 0 \; \forall \; (S, A, t) \in \Omega \times [0, T] \tag{5.3}$$

where $\widehat{\mathcal{L}}_{BS}$ is as given in (4.29).

5.0.1 Solution by Relaxation

For this section, change notation and call **U** (which comes from approximating the values of u at finite difference mesh points, to get U) by "**x**" to conform to standard matrix notation. The finite difference equations combined with the boundary conditions can be written in matrix form as

$$M\,\mathbf{x}^{m-1} = \mathbf{d} \tag{5.4}$$

where M is a mostly zero $\{N \times N\}$ matrix with three non-zero diagonals.

Classical relaxation methods involve splitting the sparse matrix which arises from the finite difference scheme and then iterating until a solution is found. A standard (triangular) splitting of M as

$$M = L + D + U$$

where L is the lower triangular part of M with zeros on the diagonal,

$$L = \begin{bmatrix} 0 & 0 & & \cdots & & 0 \\ a_1 & 0 & 0 & & & \\ & a_2 & 0 & 0 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & a_{N-1} & 0 & 0 \\ 0 & & \cdots & & a_N & 0 \end{bmatrix}$$

D is the diagonal part of M,

$$D = \begin{bmatrix} b_0 & 0 & & \cdots & & 0 \\ 0 & b_1 & 0 & & & \\ & 0 & b_2 & 0 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & 0 & b_{N-1} & 0 \\ 0 & & \cdots & & 0 & b_N \end{bmatrix}$$

and U is the upper triangular part of M with zeros on the diagonal,

$$U = \begin{bmatrix} 0 & c_0 & & \cdots & & 0 \\ 0 & 0 & c_1 & & & \\ & 0 & 0 & c_2 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & 0 & 0 & c_{N-1} \\ 0 & & \cdots & & 0 & 0 \end{bmatrix}$$

Note that L and U are not the LU-factors of M. The splitting must be regular. That is, element by element,

$$(L+D)^{-1} \geq 0 \quad \text{and} \quad U \geq 0$$

In the Jacobi method, the k^{th} step of iteration is

$$D \cdot \mathbf{x}^{k+1} = -(L+U) \cdot \mathbf{x}^k + \mathbf{d}$$

The right hand side is known at each step, so the solution is

$$\mathbf{x}^{k+1} = -D^{-1}(L+U) \cdot \mathbf{x}^k + D^{-1} \cdot \mathbf{d}$$

For Gauss-Seidel iteration, the k^{th} step is

$$(L+D) \cdot \mathbf{x}^{k+1} = -U \cdot \mathbf{x}^k + \mathbf{d} \tag{5.5}$$

The right hand side is known at each step and a lower triangular system of equations is obtained which can be solved by forward substitution. The solution is

$$\mathbf{x}^{k+1} = -(D+L)^{-1}U \cdot \mathbf{x}^k + (D+L)^{-1} \cdot \mathbf{d} \tag{5.6}$$

A more practical algorithm is obtained if an *overcorrection* is made to the value of x^{k+1} at the $k+1^{th}$ stage of Gauss-Seidel iteration, thus anticipating future corrections. Adding and subtracting $\mathbf{x}^k$ on the right hand side of equation (5.6) gives

$$\mathbf{x}^{k+1} = \mathbf{x}^k - (L+D)^{-1}[(L+D+U) \cdot \mathbf{x}^k - \mathbf{d}]$$

The term in the square brackets is just the residual vector, $\mathbf{r}^k = \mathbf{d} - M\mathbf{x}^k$. Therefore,

$$\mathbf{x}^{k+1} = \mathbf{x}^k - (L+D)^{-1}\mathbf{r}^k$$

Now for the over corrections step. Define

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \omega(L+D)^{-1}\mathbf{r}^k$$

where ω is the over relaxation parameter. The optimal over relaxation factor lies between 1.0 and 2.0. The stopping criteria are clearly

$$\|\mathbf{x}^{k+1} - \mathbf{x}^k\| \leq \epsilon_1$$

and

$$\|\mathbf{r}^{k+1}\| \leq \epsilon_2$$

The relaxation error is $\mathbf{e}^k = \mathbf{x} - \mathbf{x}^k$. The error equations for the classical methods all take the form $\mathbf{e}^{k+1} = G\,\mathbf{x}^k$ where G is the iteration matrix of the given method. Convergence requires

$$\lim_{m \to \infty} \|G^m\| = 0$$

which is equivalent to $\rho(G) \to 0$ where $\rho(G)$ is the spectral radius of G.

5.1 Projected Relaxation

The finite difference equations also become inequalities.

$$U_{j,k}^{m-1} - \tilde{U}_{j,k}^{m} \geq \Delta t\Big(\theta\widehat{\mathcal{L}}_j(U_{j,k}^{m-1}) + (1-\theta)\widehat{\mathcal{L}}_j(\tilde{U}_{j,k}^{m})\Big)$$

$$U_{j,k}^{m-1} \geq g(S_j, A_k, t^{m-1})$$

leading to a *discrete* linear complementarity problem

$$\big(U_{j,k}^{m-1} - g(S_j, A_k, t^{m-1})\big) \times \Big((U_{j,k}^{m-1} - \tilde{U}_{j,k}^{m}) - \Delta t\Big(\theta\widehat{\mathcal{L}}_j(U_{j,k}^{m-1}) + (1-\theta)\widehat{\mathcal{L}}_j(\tilde{U}_{j,k}^{m})\Big)\Big) = 0$$

that can be solved using a projection method. The discrete linear complementarity problem can be solved using relaxation combined with a simple projection step. A projected Gauss-Seidel iteration method is exploited and solved.

Written out in component form, where i now represents the iteration step, equation (5.5) is

$$\begin{aligned}
U_{0,k}^{i+1} &= \frac{1}{b_0}(-c_j U_{1,k}^{i} + \mathbf{d}_0) \\
U_{j,k}^{i+1} &= \frac{1}{b_j}(-a_j U_{j-1,k}^{i} - c_j U_{j+1,k}^{i} + \mathbf{d}_j);\ j \in [1, N-1] \\
U_{N,k}^{i+1} &= \frac{1}{b_N}(-a_N U_{N-1,k}^{i} + \mathbf{d}_N)
\end{aligned}$$

At the end of every relaxation step, the option price $U_{j,k}^{i+1}$ is corrected by adjusting any values less than the payoff to be equal to the payoff. That is,

$$U_{j,k}^{i+1} = \max(U_{j,k}^{i+1}, g(S_j, A_k, t^m))$$

5.2 Penalty Method

In current practice, the most common method of handling the early exercise condition is simply to advance the discrete solution over a time-step, ignoring the constraint, and then to apply the constraint explicitly. One method for incorporating the algebraic constraint (due to the early exercise feature) is to view the problem as a linear complementarity problem and then use projected SOR to solve the discrete algebraic equations. Alternatively, it is well known that a linear complementarity problem can be solved by a penalty method [EO82, Fri88, Sch86]. The use of penalty functions can convert a constrained problem to an unconstrained function. The objective is expanded to include equality constraints multiplied by a penalty

parameter. If this is minimized for increasing values of the parameter, the solution converges to the desired result.

The penalty method for solving option problems was introduced by Zvan, Forsyth and Vetzal in [ZFV98b]. This approach was developed further in [FV02, NST02]. In the penalty approach, a small, continuous, non-linear penalty term is added to the pricing equation. This gives a fixed solution domain, removing the difficulties associated with the free and moving boundary imposed by the early exercise feature of the contract. As the solution approaches the payoff function at expiry, the penalty term forces the solution to stay above it. When the solution is far from the payoff, the term is small and thus the pricing equation is approximately satisfied in this region.

The advantage of the penalty method is that a single technique can be used for one dimensional or multidimensional problems, and standard sparse matrix solution techniques can be used to solve the Jacobian matrix. This technique can be used for any type of discretisation, in any dimension, and on unstructured meshes.

5.2.1 Formulation

The method used in this section essentially uses a non-smooth Newton iteration to solve the penalised problem. The advantages of this approach are (under certain conditions, see [FV02]) that:

- This method has finite termination. That is, for an iterate sufficiently close to the solution, the algorithm converges satisfactorily in one iteration. This is of special advantage when dealing with American option pricing, since an excellent initial guess exists from the previous time step. In fact, [FV02] has shown that for typical grids and time-steps, the algorithm takes, on average, less than two iterations per time-step to converge. Finite termination also implies that the number of iterations required for convergence is insensitive to the size of the penalty factor until the limits of machine precision is reached.

- The iteration is globally convergent using full Newton steps.

The real advantage of the penalty method is that this technique takes full advantage of the fact that a good initial iterate is available, and of sparsity, which is important in multi-factor problems.

This portion of the study aims to answer the question: How effective is the penalty method with stretched finite difference methods for Asian options? Subsequently, is the penalty method a reliable and efficient alternative to projection for these cases?

The penalty method solution method involves replacing problem (5.1) by the nonlinear PDE

$$-\frac{du}{dt} = \widehat{\mathcal{L}}_{BS}\, u + \rho \max(g - u, 0) \tag{5.7}$$

The penalty parameter ρ is selected so that

$$|u - g| < tol \text{ when } u < g$$

where *tol* is a user specified tolerance.

Discretisation

The solution is carried out backwards in time, so beginning with values at T, values at t^{m-1} are obtained, and so on. Take the integral of (5.7) along the path $\mathcal{P}_k^m(A, t; S_j) \equiv (\tilde{A}_k, t^m) \curvearrowright (A_k, t^{m-1})$,

$$\begin{aligned} u(S_j, A_k, t^{m-1}) \;&-\; u(S_j, \tilde{A}_k, t^m) \\ &= \int_{\mathcal{P}_k^m(A,t;S_j)} \widehat{\mathcal{L}}_{BS} u(S, A, t)\; dt + \rho \max(g - u(S_j, A_k, t^{m-1}), 0) \end{aligned}$$

If $\widehat{\mathcal{L}}_j$ is an $O(h^2)$ approximation to $\widehat{\mathcal{L}}_{BS}$, where $h = \min(\Delta x_1)$, and $\{U_{j,k}^m\}$ are finite difference mesh prices on $\Omega = \{S_0, ...S_N\} \otimes \{A_0, ...A_N\}$ such that

$$U_{j,k}^{m-1} = \tilde{U}_{j,k}^m + \Delta t \Big(\theta \widehat{\mathcal{L}}_j(U_{j,k}^{m-1}) + (1-\theta)\widehat{\mathcal{L}}_j(\tilde{U}_{j,k}^m) + \rho \max(g - U_{j,k}^{m-1}, 0) \Big)$$

then $U_{j,k}^m \approx U(S_j, A_k, t^m)$ to $O(\Delta t^2) + O(h^2)$ for $\theta = 0.5$ and is unconditionally stable for $\theta \geq 0.5$. $\tilde{U}_{jk}^m$ is the mesh price interpolated to $\tilde{A}_k$. Section 4.3.2 outlines how to determine $\tilde{A}_k$.

The penalty method proposed uses a Newton iteration to solve the penalised problem. The non-linear system of equations to be solved can be written as

$$U_{j,k}^{m-1} - \Delta t \theta \widehat{\mathcal{L}}_j(U_{j,k}^{m-1}) - \rho \max(g - U_{j,k}^{m-1}, 0) = c_{j,k}^m$$

where $c_{j,k}^m$ represents the whole right hand side. Set $U_{j,k}^{m-1} = V_j$ and $c_{j,k}^m = c_j$. This implies that,

$$V_j - \Delta t\ \theta\ \widehat{\mathcal{L}}_j V_j - \rho \max(g - V_j, 0) = c_j$$

The system of equations is

$$G(V) = AV + f(V) = c$$

where

$$f(V_i) = -\rho \max(g - V_i, 0)$$

The Newton iteration algorithm can then be written as

$$V^{r+1} = V^r - J^{-1}[G(V^r) - c]$$

where J is the Jacobian matrix, given by

$$J = \frac{\partial G}{\partial V} = A + \frac{\partial f_i}{\partial V_i}$$

and

$$\frac{\partial f_i}{\partial V_i} = \begin{cases} \rho & : \quad V_i < g \\ 0 & : \quad V_i > g \end{cases}$$

Thus, a suitable form for the discrete penalty term is

$$\frac{\partial}{\partial U}\left(\max(g - U_j^{m-1}, 0)\right) \equiv \xi_j^{(m-1)} = \begin{cases} 1 & : \quad U_j^{m-1} < g \\ 0 & : \quad U_j^{m-1} > g \end{cases} \tag{5.8}$$

5.2.2 Non-Uniform Time-Steps

The quantity $\tilde{A}_k$ depends on a ratio α (equation 4.38). One obvious criticism of a uniform time-stepping algorithm is that the ratio α above varies rather abruptly with time. At expiry, $\alpha = \frac{t^{m-1}}{T} \approx 1$, and thus $\tilde{A}_k \approx A_k$. At the final time-step, $t_0 = 0$, implying that $\alpha = \frac{t_0}{t_1} = 0$, and thus $\tilde{A}_k = S_j$. Figure 5.1 shows the variation in α as the solution progresses through time. Refer back to Figure 4.4 which shows the dependence of $\tilde{A}_k$ on S at different times. The implication is that in the S-L integration, information gets passed along the S_j line, terminating at the diagonal where $\tilde{A}_k = S_j$. However, with uniform dt, the intervals of α are far from uniform, and the final (rather large) "α-steps" are the concern, since they are taking information from rather large distances away and forcing that information onto the diagonal.

In an endeavour to rectify this shortcoming, non-uniform time-steps were chosen. dt is chosen to follow a cubic polynomial. Figure 5.2 shows the variation in α as the solution progresses through time for non-uniform time-steps. It can be seen that this is a more acceptable distribution of "α-steps".

The time-stepping algorithm is now:

```
Set time T = Expiry
Store the value of constant time-steps, dt
```

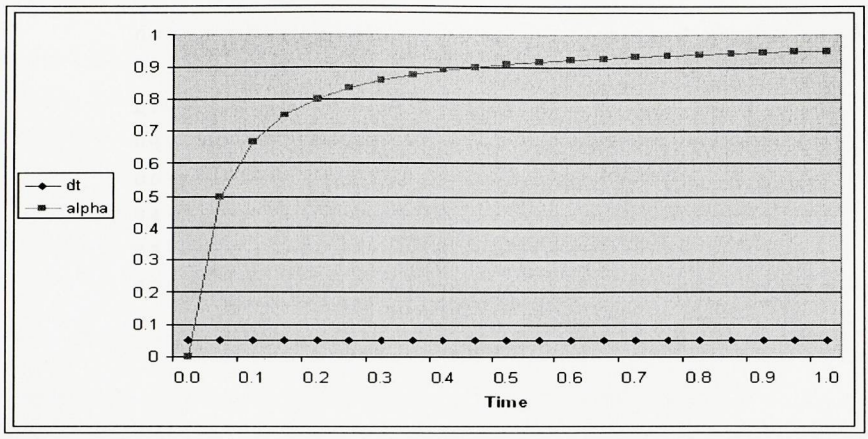


Figure 5.1: Effect of uniform time-steps on alpha.

Define non-uniform time-steps to follow a cubic polynomial:

$$dt_m = \beta dt(1 + cm^3) \ ; \ m = 1, \cdots, nt$$

(for a user-defined constant, c, and a function β), such that

$$\sum_{m=1}^{nt} dt_m = T$$

Thus, β is defined by

$$\beta = \frac{T}{dt \sum_{m=1}^{nt}(1 + cm^3)}$$

An interpolation error occurs at each time-step, so this method ensures the number of time-steps does not increase while making them non-uniform.

5.3 Results and Conclusions for American Asians

Numerical results are now presented. As previously, the desired level of accuracy is set to be 0.1% of S_0.

5.3.1 Numerical Convergence

Table 5.1 displays results corresponding to the fixed strike Asian put with early exercise for *uniform* time-steps using *projection* when $r = 0.1$, and prices are quoted

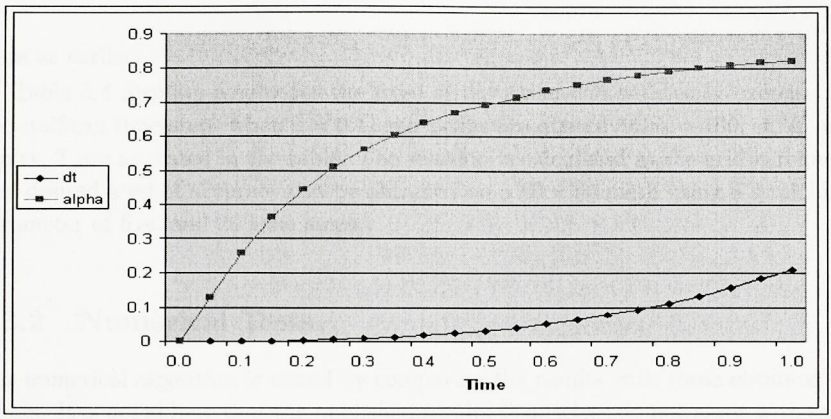


Figure 5.2: Effect of non-uniform time-steps on alpha.

at $S_0 = 100$. σ, K and expiry, T vary as shown in the table. The solution is calculated as the grid is refined. The solutions at the discrete asset prices were obtained by interpolating between grid points. The results suggest that the scheme is convergent. Further, the desired level of accuracy can be obtained on a 80×80 mesh using a stretching parameter of five, and 40 time steps.

Table 5.2 displays results corresponding to the fixed strike Asian put with early exercise for *uniform* time-steps using the *penalty method* when $r = 0.1$, and prices are quoted at $S_0 = 100$. σ, K and expiry, T are as shown in the table. The solution is calculated as the grid is refined. Although the results quoted are only for the case $\sigma = 0.2$, the penalty method was used to systematically check all cases in $\sigma = 0.1$, $\sigma = 0.4$ also. The penalty coefficient was specified as $\rho = 10^4$, $\rho = 10^5$, $\rho = 10^6$. It was found that when the penalty coefficient of $\rho = 10^5$ or $\rho = 10^6$ is used, the prices are consistently comparable to the prices obtained when projection is used. The results suggest that the penalty method is indeed a reliable alternative to projection. For the remainder of quoted results let it be understood that the penalty method was also chosen as the solution method for verification and the results were comparable.

Table 5.3 shows results for the fixed strike Asian put with early exercise for *non-uniform* time-steps when $r = 0.1$, and prices are quoted at $S_0 = 100$. σ, K and expiry, T vary as shown in the table. The table shows the solution calculated as the grid is refined. The desired level of accuracy can be obtained on a 80×80 mesh using a stretching parameter of five, and 20 time steps (half the number of time

steps as earlier).

Table 5.4 displays results for the fixed strike Asian call with early exercise for non-uniform time-steps when $r = 0.1$, and prices are quoted at $S_0 = 100$. σ, K and expiry, T are as stated in the table. The solution is calculated as the grid is refined. The desired level of accuracy can be obtained on a 80×80 mesh using a stretching parameter of five, and 20 time steps.

5.3.2 Numerical Tests

The numerical algorithm is tested by comparing the results with those obtained by others. It is noted here that the published results themselves do not agree with each other, making it difficult to conduct a concrete validation against them. However, the S-L results agree with those published within the same ballpark as they disagree. The S-L method has been shown to be rigorous and robust, and convergence tests suggest it to work extremely well for the problems considered.

Results obtained with the S-L method in the case of fixed strike Asian put with early exercise are compared to results published in [BP96] in Table 5.5 and 5.6. The results quoted for the S-L case are those of the 160×160 mesh using uniform time-steps and non-uniform time-steps respectively. The comparisons are seen to be satisfactory, more so in the case of non-uniform time-stepping.

Table 5.7 compares results obtained with the S-L method to results published in [BP96] and [ZFV98a] for the case of fixed strike Asian call with early exercise. The results quoted for the S-L case are those of the 160×160 mesh using non-uniform time-steps. The comparison is seen to be satisfactory.

In Table 5.8 the mean execution time for obtaining results while using a projection solver ($nt = 40$) is displayed along with the solution grid size used. Average CPU times are shown since the performance of the iterative method varies with the volatility chosen.

Table 5.9 shows the mean execution time elapsed while using a penalty method to solve the pricing problem. Again, $nt = 40$, in keeping with the parameters of the problem. The results suggest that the penalty method is actually a more efficient alternative than projection in solving the early exercise problem.

Table 5.10 quotes the mean execution time required to obtain a solution when non-uniform time steps are used. As noted earlier, the desired level of accuracy can be obtained with half the number of time steps, so here $nt = 20$. Once again, average CPU times are shown since the performance of the iterative method varies with the volatility chosen.

5.3.3 Surface Plots

For a fixed-strike Asian put with early exercise, Figure 5.3 shows the surface plots of the evolution of the solution over time. Typical to projection, the initial conditions are frozen in. As time progresses, the solution is seen to rotate along the $S = A$ diagonal. A similar evolution can be seen in the case of fixed-strike Asian call options with early exercise (Figure 5.4). This figure displays the initial data for the Asian call problem with early exercise when $K = 100$, and displays the evolution of the solution.

Table 5.1: Convergence results for Fixed Strike Asian put with early exercise, using uniform dt. $r = 0.1$, at a spot price of $S_0 = 100$. The Semi-Lagrangian (S-L) results are calculated using projection on meshes of size 40x40, 80x80 and 160x160. nt, the number of time steps used, is 40 for all maturities. λ, the stretching parameter, is 5.

σ	T	K	Calculated by S-L		
			40x40	80x80	160x160
0.10	0.25	95	0.0734	0.0273	0.0140
		100	0.6654	0.7787	0.8475
		105	5.6138	5.6527	5.5556
	0.50	95	0.1271	0.0605	0.0506
		100	0.8380	0.9622	1.0075
		105	5.7502	5.6927	5.6266
	1.00	95	0.1808	0.1117	0.1058
		100	0.9579	1.1036	1.1255
		105	5.8134	5.7040	5.6700
0.20	0.25	95	0.4523	0.4059	0.4018
		100	1.9592	2.0843	2.0900
		105	6.5956	6.4608	6.4268
	0.50	95	0.8417	0.8145	0.8135
		100	2.6210	2.6671	2.6673
		105	6.8111	6.7723	6.7539
	1.00	95	1.3374	1.3274	1.3277
		100	3.2432	3.2557	3.2547
		105	7.1317	7.0855	7.0696
0.40	0.25	95	2.2171	2.2095	2.2094
		100	4.6186	4.6145	4.6141
		105	8.4509	8.4285	8.4189
	0.50	95	3.5996	3.5961	3.5954
		100	6.1284	6.1237	6.1223
		105	9.6848	9.6844	9.6800
	1.00	95	5.2924	5.2885	5.2872
		100	7.8619	7.8565	7.8550
		105	11.1991	11.1897	11.1869

Table 5.2: Convergence results for Fixed Strike Asian put with early exercise, using uniform dt. $r = 0.1$, at a spot price of $S_0 = 100$. The Semi-Lagrangian (S-L) results are calculated using the penalty method on meshes of size 40x40, 80x80 and 160x160. nt, the number of time steps used, is 40 for all maturities. λ, the stretching parameter, is 5.

σ	T	K	ρ	Calculated by S-L		
				40x40	80x80	160x160
0.20	0.25	95	10^4	0.4518	0.4058	0.4017
			10^5	0.4522	0.4059	0.4018
			10^6	0.4523	0.4059	0.4018
		100	10^4	1.9597	2.0833	2.0895
			10^5	1.9592	2.0842	2.0900
			10^6	1.9592	2.0843	2.0900
		105	10^4	6.5909	6.4573	6.4238
			10^5	6.5951	6.4605	6.4265
			10^6	6.5955	6.4608	6.4267
	0.50	95	10^4	0.8415	0.8144	0.8135
			10^5	0.8417	0.8145	0.8135
			10^6	0.8417	0.8145	0.8135
		100	10^4	2.6205	2.6666	2.6669
			10^5	2.6210	2.6671	2.6673
			10^6	2.6210	2.6671	2.6673
		105	10^4	6.8082	6.7703	6.7521
			10^5	6.8108	6.7720	6.7538
			10^6	6.8111	6.7722	6.7539
	1.00	95	10^4	1.3373	1.3274	1.3277
			10^5	1.3374	1.3274	1.3278
			10^6	1.3374	1.3274	1.3278
		100	10^4	3.2430	3.2554	3.2545
			10^5	3.2432	3.2557	3.2548
			10^6	3.2432	3.2557	3.2548
		105	10^4	7.1297	7.0840	7.0684
			10^5	7.1315	7.0853	7.0695
			10^6	7.1317	7.0855	7.0697

Table 5.3: Convergence results for Fixed Strike Asian put with early exercise, using non-uniform dt. $r = 0.1$, at a spot price of $S_0 = 100$. The Semi-Lagrangian (S-L) results are calculated on meshes of size 40x40, 80x80 and 160x160. nt, the number of time steps used, is 20 for all maturities. λ, the stretching parameter, is 5.

σ	T	K	Calculated by S-L		
			40x40	80x80	160x160
0.10	0.25	95	0.0384	0.0206	0.0136
		100	0.6815	0.6822	0.8350
		105	5.6604	5.6654	5.5407
	0.50	95	0.0830	0.0524	0.0499
		100	0.8283	0.8393	1.0141
		105	5.7937	5.7032	5.5821
	1.00	95	0.1304	0.1046	0.1051
		100	0.8885	0.9951	1.1235
		105	5.8517	5.7183	5.6043
0.20	0.25	95	0.3791	0.4000	0.4015
		100	1.7797	2.1125	2.0843
		105	6.6332	6.3903	6.3760
	0.50	95	0.7709	0.8107	0.8133
		100	2.4871	2.6833	2.6592
		105	6.8310	6.6894	6.6609
	1.00	95	1.2947	1.3247	1.3262
		100	3.2049	3.2363	3.2363
		105	7.0447	6.9507	6.9251
0.40	0.25	95	2.1939	2.2088	2.2095
		100	4.6762	4.6095	4.6084
		105	8.3535	8.3784	8.3581
	0.50	95	3.5988	3.5948	3.5945
		100	6.1282	6.1140	6.1125
		105	9.6683	9.6238	9.6199
	1.00	95	5.2882	5.2833	5.2825
		100	7.8386	7.8374	7.8360
		105	11.1721	11.1223	11.1189

Table 5.4: Convergence results for Fixed Strike Asian call with early exercise, using non-uniform dt. $r = 0.1$, at a spot price of $S_0 = 100$. The Semi-Lagrangian (S-L) results are calculated on meshes of size 40x40, 80x80 and 160x160. nt, the number of time steps used, is 20 for all maturities. λ, the stretching parameter, is 5.

σ	T	K	Calculated by S-L		
			40x40	80x80	160x160
0.10	0.25	95	6.7274	6.5570	6.5906
		100	1.7729	1.9048	1.9557
		105	0.2146	0.1649	0.1522
	0.50	95	7.5840	7.6946	7.6573
		100	2.9632	3.2059	3.1944
		105	0.7490	0.7251	0.7237
	1.00	95	9.6617	9.6294	9.6279
		100	5.4150	5.3833	5.3853
		105	2.2774	2.3193	2.3228
0.20	0.25	95	7.4967	7.5296	7.4827
		100	3.0404	3.2285	3.2179
		105	0.9751	0.9873	0.9886
	0.50	95	8.9004	8.8988	8.8897
		100	4.9399	4.8884	4.8866
		105	2.2813	2.3125	2.3142
	1.00	95	11.3344	11.3012	11.2937
		100	7.5408	7.5475	7.5456
		105	4.7325	4.7312	4.7315
0.40	0.25	95	9.6527	9.5858	9.5756
		100	5.8900	5.8401	5.8372
		105	3.3115	3.3280	3.3287
	0.50	95	12.0488	12.0371	12.0328
		100	8.5289	8.5331	8.5312
		105	5.8976	5.8911	5.8905
	1.00	95	15.8127	15.7973	15.7936
		100	12.5341	12.5240	12.5203
		105	9.8486	9.8397	9.8370

Table 5.5: Calculated Semi-Lagrangian (S-L) results for Fixed Strike Asian puts with early exercise, using uniform dt. $r = 0.1$, at a spot price of $S_0 = 100$. For comparison, one other published result is also indicated.

σ	T	K	S-L	[BP96]
0.10	0.25	95	0.014	0.013
		100	0.848	0.832
		105	5.556	5.337
	0.50	95	0.051	0.051
		100	1.008	0.978
		105	5.627	5.287
	1.00	95	0.106	0.104
		100	1.126	1.079
		105	5.670	5.230
0.20	0.25	95	0.402	0.407
		100	2.090	2.066
		105	6.427	6.108
	0.50	95	0.814	0.820
		100	2.667	2.629
		105	6.754	6.338
	1.00	95	1.328	1.318
		100	3.255	3.181
		105	7.070	6.596
0.40	0.25	95	2.209	2.223
		100	4.614	4.581
		105	8.419	8.168
	0.50	95	3.595	3.610
		100	6.122	6.078
		105	9.680	9.438
	1.00	95	5.287	5.263
		100	7.855	7.761
		105	11.187	10.927

Table 5.6: Calculated Semi-Lagrangian (S-L) results for Fixed Strike Asian puts with early exercise, using non-uniform dt. $r = 0.1$, at a spot price of $S_0 = 100$. For comparison, another published result is also indicated.

σ	T	K	S-L	[BP96]
0.10	0.25	95	0.014	0.013
		100	0.835	0.832
		105	5.541	5.337
	0.50	95	0.050	0.051
		100	1.014	0.978
		105	5.582	5.287
	1.00	95	0.105	0.104
		100	1.124	1.079
		105	5.604	5.230
0.20	0.25	95	0.402	0.407
		100	2.084	2.066
		105	6.376	6.108
	0.50	95	0.813	0.820
		100	2.659	2.629
		105	6.661	6.338
	1.00	95	1.326	1.318
		100	3.236	3.181
		105	6.925	6.596
0.40	0.25	95	2.210	2.223
		100	4.608	4.581
		105	8.358	8.168
	0.50	95	3.595	3.610
		100	6.113	6.078
		105	9.620	9.438
	1.00	95	5.283	5.263
		100	7.836	7.761
		105	11.119	10.927

Table 5.7: Calculated Semi-Lagrangian (S-L) results for Fixed Strike Asian calls with early exercise, using non-uniform dt. $r = 0.1$, at a spot price of $S_0 = 100$. For comparison, two other published results are also indicated.

σ	T	K	S-L	[BP96]	[ZFV98a]
0.10	0.25	95	6.591	6.546	6.646
		100	1.956	1.967	1.903
		105	0.152	0.152	0.161
	0.50	95	7.657	7.632	7.687
		100	3.194	3.212	3.180
		105	0.724	0.735	0.733
	1.00	95	9.628	9.616	9.662
		100	5.385	5.394	5.398
		105	2.323	2.336	2.340
0.20	0.25	95	7.483	7.371	7.521
		100	3.218	3.219	3.224
		105	0.989	1.001	1.009
	0.50	95	8.890	8.805	8.908
		100	4.887	4.893	4.901
		105	2.314	2.337	2.337
	1.00	95	11.294	11.218	11.295
		100	7.546	7.521	7.548
		105	4.732	4.729	4.742
0.40	0.25	95	9.576	9.447	9.548
		100	5.837	5.826	5.846
		105	3.329	3.347	3.349
	0.50	95	12.033	10.927	11.997
		100	8.531	8.519	8.527
		105	5.891	5.913	5.899
	1.00	95	15.794	15.649	15.749
		100	12.520	12.439	12.497
		105	9.837	9.790	9.825

Table 5.8: Average times (seconds) required to obtain a solution, when a projection solver is used. nt, the number of time steps used, is 40. λ, the stretching parameter, is 5.

Size of Grid	Mean execution time (sec)
40×40	4.0
80×80	21.1
160×160	158.38

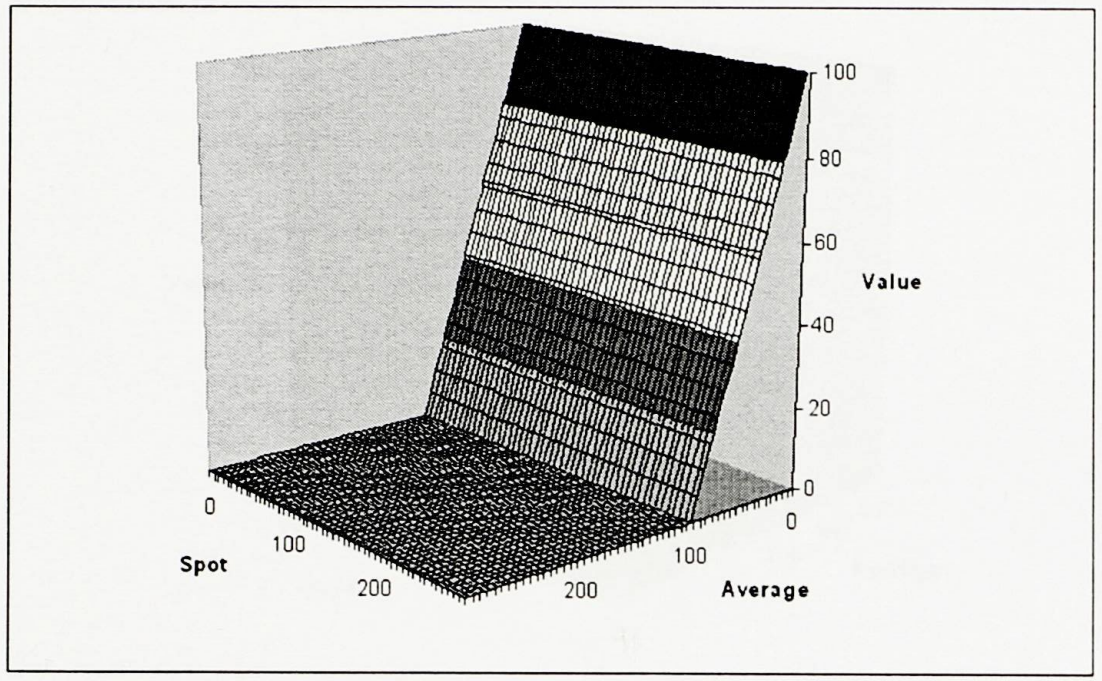
Table 5.9: Average times (seconds) required to obtain a solution, when a penalty method is used. nt, the number of time steps used, is 40. λ, the stretching parameter, is 5.

ρ	Size of Grid	Mean execution time (sec)
10^5	40×40	3.7
	80×80	14.8
	160×160	65.3
10^6	40×40	3.7
	80×80	14.8
	160×160	65.4

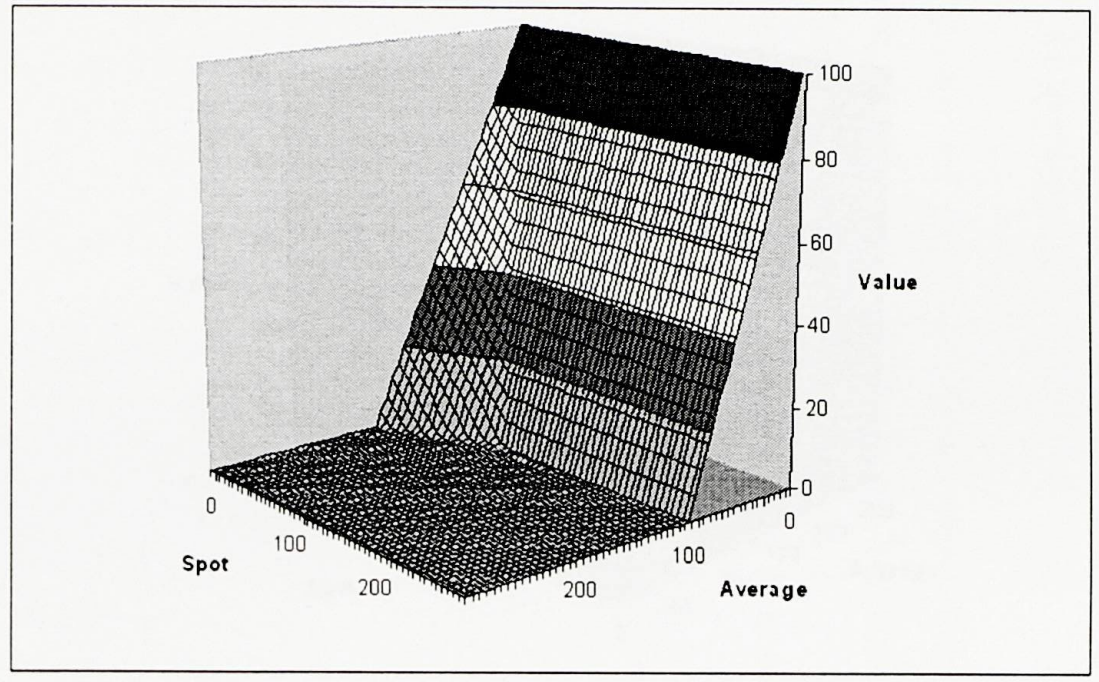
Table 5.10: Average times (seconds) required to obtain a solution, when non-uniform time-steps are used. nt, the number of time steps used, is 20. λ, the stretching parameter, is 5.

Size of Grid	Mean execution time (sec)
40×40	2.7
80×80	18.0
160×160	150.2

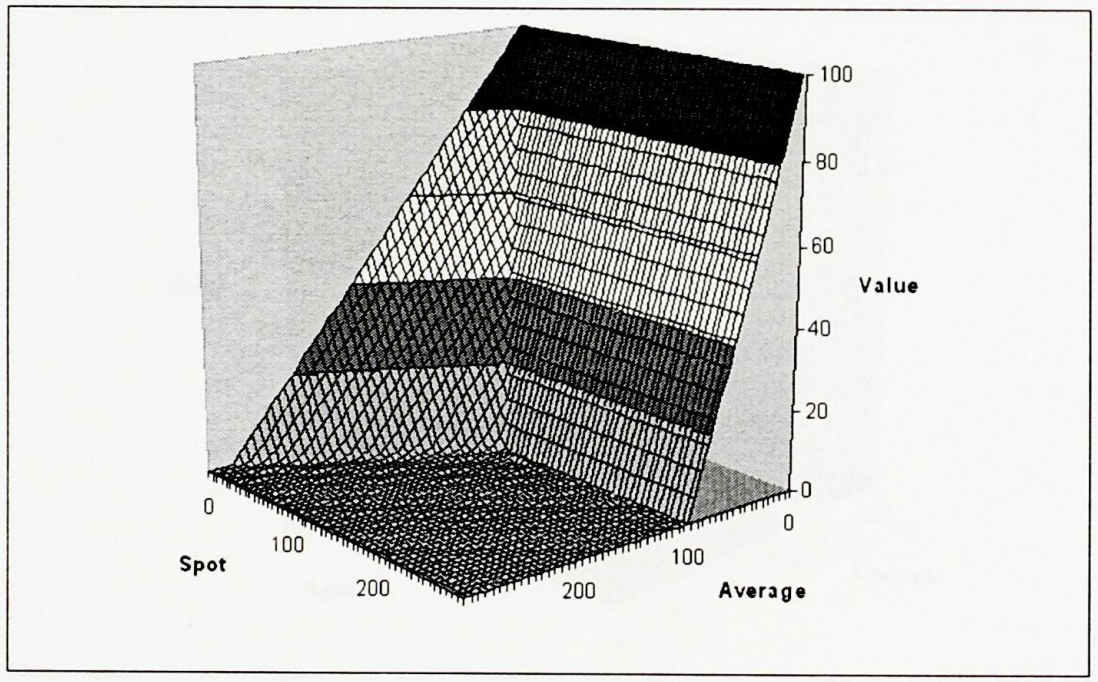
Figure 5.3: Evolution over time of the solution for a fixed-strike American put option. Model parameters are $K = 100$, $r = 0.1, \sigma = 0.2, T = 3$ months.



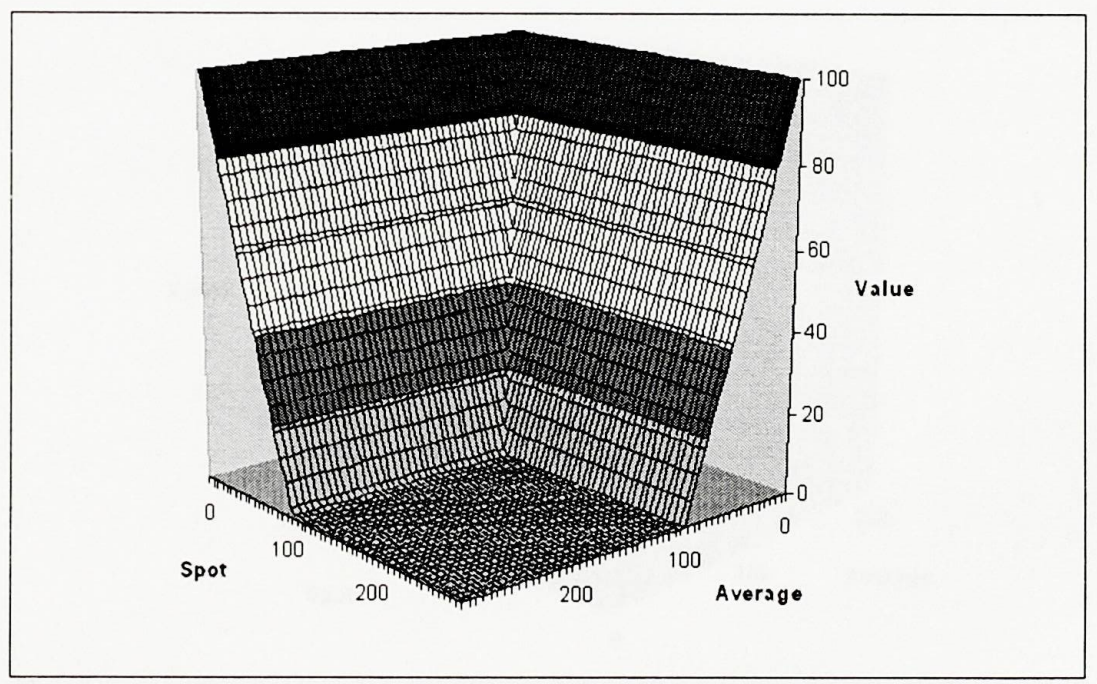
(a) At time = expiry (3 months).



(b) At time = 2 months.

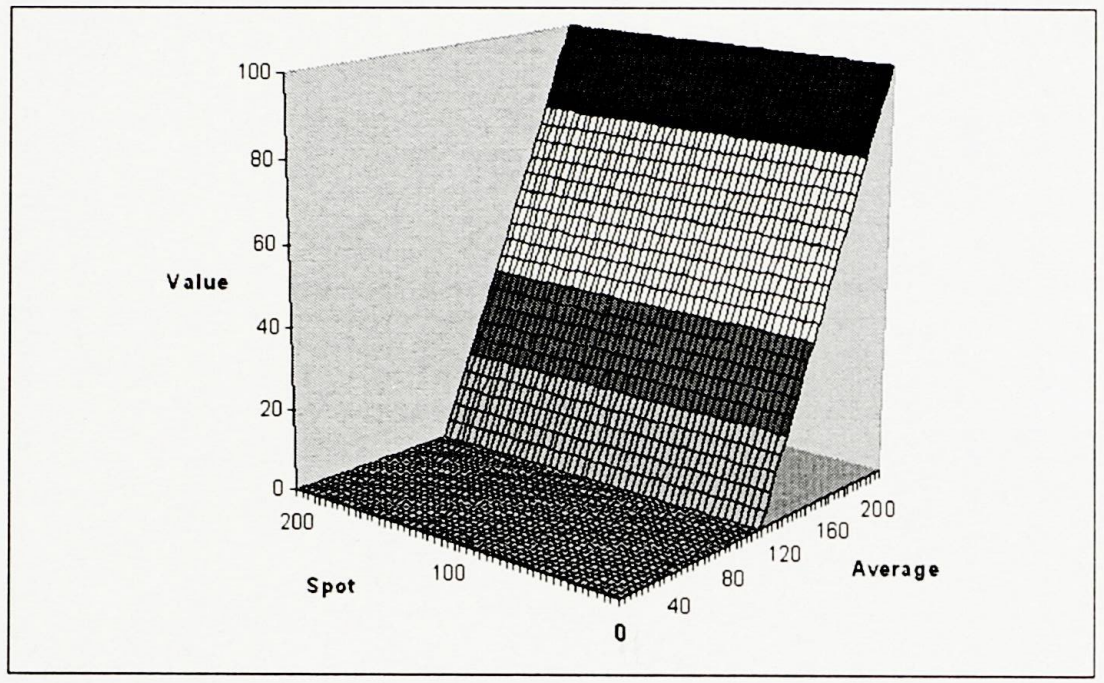


(c) At time = 1 month.

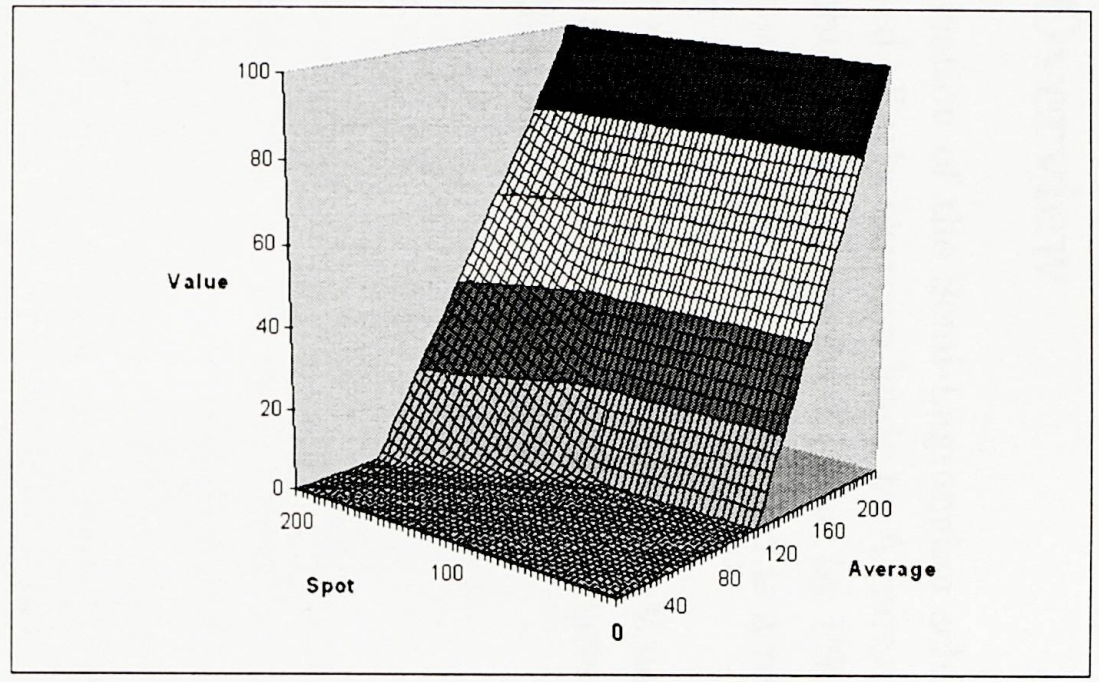


(d) At time = 0.

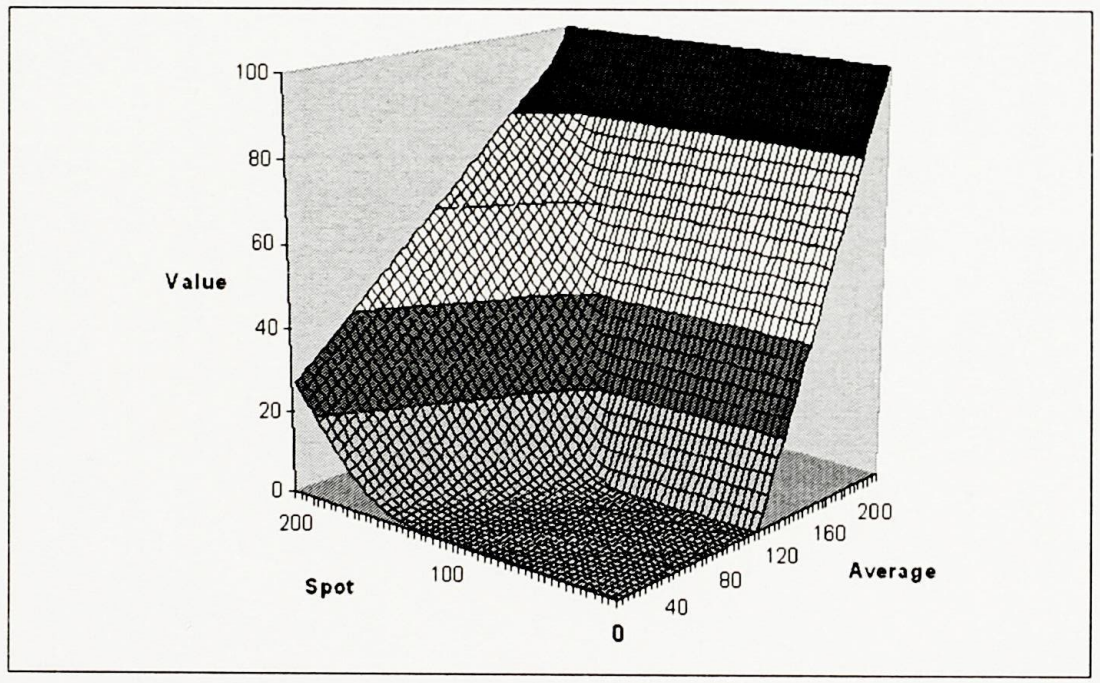
Figure 5.4: Evolution over time of the solution for a fixed-strike American call option. Model parameters are $K = 100$, $r = 0.15, \sigma = 0.4, T = 6$ months.



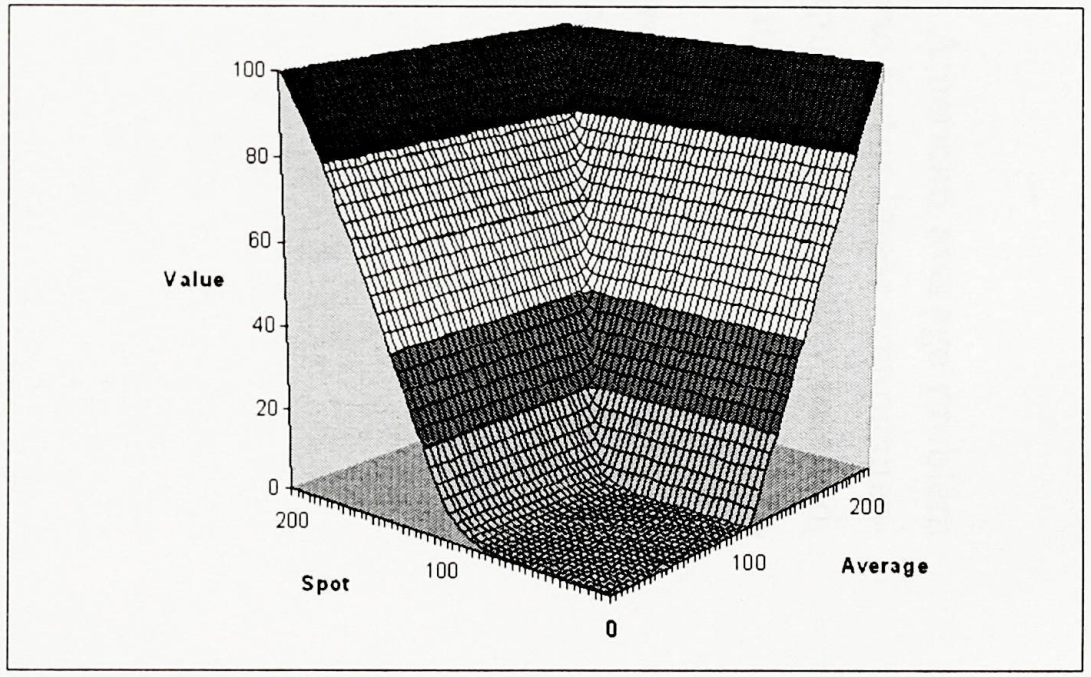
(a) At time = expiry (6 months).



(b) At time = 4 months.



(c) At time = 2 months.



(d) At time = 0.

5.4 Overview

The application of the Semi-Lagrangian scheme to the American average problem is described. Early exercise leads to A-parameterised one-factor linear complementarity problems that are solved either by PSOR techniques or the Penalty method. It was shown that the penalty method is a reliable and more efficient alternative to projection.

Better results were obtained when the time-stepping was made to be non-uniform.

These approaches can be applied to more complex Asian options. For example, assets with stochastic volatility [CP99].

Chapter 6

Conclusions and Further Work

In order to price a financial derivative, some stochastic description of the evolution of the underlying variables on which the financial derivative must be assumed. Portfolio replication arguments can be used to derive a convection-diffusion partial differential equation which the financial derivative must satisfy under the further assumptions of frictionless, arbitrage-free markets. If early exercise is permitted by the financial derivative, then a free boundary problem (which can be formulated as a linear complementarity problem) is obtained.

There are very few instances where the analytic solution is readily available. Generally, numerical methods have to be employed to determine the solution and hence the price of the financial derivative. Given the practical nature of this specific problem of options pricing, the numerical solution needs to satisfy the criteria that it is accurate and quickly determined. In this thesis, various numerical and computational strategies have been considered and developed, which produce competitive methods for a range of financial derivative problems.

The numerical methods are based on efficient finite difference discretisation in association with optimal solvers of the discrete system. These methods allow rapid computation of solutions.

The problems are defined on infinite domains, which are truncated to allow for numerical approximation on a finite grid. Outflow boundary conditions of a numerical origin which admit the required solution are placed on the truncated boundaries at large (but finite) distances from the region of interest. Regular Cartesian meshes are combined with orthogonal co-ordinate transformations thereby achieving detailed resolution in the regions of interest and grid coarsening in the regions of least interest. The transformations allow the domain truncation at large distances from the regions of interest while using practical grid sizes. These transformations have been essential in producing accurate solutions on modest computational grids.

The spatial discretisation strategy was chosen to meet accuracy requirements as well as to produce coefficient matrices with favourable sparsity and stability properties. Given the character of the equation, time integration schemes have been chosen to produce stable methods while satisfying accuracy criterion. Backward Eulerian time integration has been used where appropriate. Semi-Lagrangian time integration schemes have been developed for convection dominated problems. Furthermore, non-uniform time-stepping routines have been used to enhance efficiency. The methods produce accurate solutions which display approximately second-order convergence.

Additionally, a Laplace transform application to the pricing of a European put option was described. Results were compared to those produced by the direct solver algorithm and were shown to be comparable. It was suggested that the Laplace transform technique be made use of for higher order pricing problems.

6.1 Further Work

There are many possible directions which can be pursued as further work leading on from the research presented in this thesis. A discussion follows now in some detail of some of the more promising of these.

6.1.1 Volatility

The research in the thesis has followed the assumptions of the Black-Scholes world in that the volatility of the underlying is a single constant value. It is a parameter that cannot be directly observed. Stocks typically have a volatility between 20% and 50%. One way of estimating the volatility is from a history of the stock price. In practice, however, traders usually work with **implied volatilities** [Wil98]. These are volatility values implied by option prices observed in the market.

Markets suggest that actual prices for European options on the same underlying asset have associated implied volatilities that vary with both exercise price and time to expiry. A plot of the implied volatility of an option as a function of its strike price is known as a **volatility smile** [Hul03]. A three-dimensional plot of the implied volatility against both maturity and strike produces a **volatility surface**. It represents the constant value of volatility that gives each traded option a theoretical value equal to the market value.

Stochastic Volatility

The volatility of the underlying asset can also be modelled as a stochastic variable [Wil98], [HW87], [MT90], [Cla98], [ZFV98b]. The time series properties of the volatility of price returns has been investigated by several authors [Tay86] and the discrete-time (G)ARCH models [Wil98], [Hul03] have been widely used to model the volatility of financial time series. In continuous time, SDE's are used to model volatility and [MT90] argue that this approach provides a better description of empirical data than the constant volatility geometric Brownian motion models.

The value of an option with stochastic volatility is a function of three variables, $u(S, \sigma, t)$. In some cases [Cla98], [ZFV98b] an option which is a function of the asset price, S and the variance, $v = \sigma^2$ is considered instead, where S and v evolve according to:

$$\begin{aligned} dS &= \mu S dt + \sqrt{v} S dW_{1t} \\ dv &= \kappa(v,t)dt + \xi(v,t)dW_{2t} \end{aligned}$$

where $W_t = (W_{1t}, W_{2t})$ is a two-dimensional Brownian motion with correlation coefficient $|\rho| < 1$, and $\kappa : \mathbb{R} \times [0,T] \to \mathbb{R}$, $\xi : \mathbb{R} \times [0,T] \to \mathbb{R}$ are functions satisfying sufficient regularity conditions. Different functional forms of κ and ξ have been used in the finance literature. Following the usual steps, the stochastic volatility model leads to a two-dimensional option pricing equation which, in most cases, has to be solved numerically.

6.1.2 Jump Diffusion Models

The basic building block for the random walks considered so far is continuous Brownian motion based on the Normally-distributed increment. However, there is increasing evidence that the usual assumption of geometric Brownian motion should be augmented by discontinuous jump processes [dFL05], [Hul03], [Nef96], [Wil98]. A richer model which has attracted attention is based on the jump diffusion process, first suggested in [Mer76]. The extra building block needed for the jump-diffusion model for an asset price is the Poisson process.

The Semi-Lagrangian scheme described earlier for Asian options can be applied equally well to Asian options with jump diffusion [dFL05]. In this case, the price of the underlying asset (S) follows a jump diffusion process,

$$dS = (\mu - \lambda\kappa)Sdt + \sigma S dW + (\eta - 1)dq$$

where μ and σ are the drift and diffusion terms respectively, W is a standard Brownian motion, and

$$dq \quad \text{is the independent Poisson process,} \quad = \begin{cases} 0 & : \quad \text{with probability } 1 - \lambda dt \\ 1 & : \quad \text{with probability } \lambda dt \end{cases}$$

λ is the intensity of the Poisson process, and $(\eta - 1)$ is an impulse function producing a jump from S to $S\eta$. The processes dW and dq are assumed to be independent. The continuously sampled arithmetic average evolves according to

$$dA = \frac{1}{t}\Big(S - A\Big)dt$$

In general it is not possible to construct a hedging portfolio which eliminates jump risk. However, by adding options to the hedging portfolio, a hedging strategy can be constructed which minimizes jump risk.

The pricing equation for the Asian option under jump diffusion is then [dFL05],

$$\frac{\partial U}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + \frac{1}{t}(S - A)\frac{\partial U}{\partial A} + (r - \lambda\kappa)S\frac{\partial U}{\partial S} - rU + \Big(\lambda \int_0^\infty U(S\eta)g(\eta)d\eta - \lambda U\Big)$$

where $g(\eta)$ is the probability density function of the jump amplitude η. Thus, for all η: $g(\eta) \geq 0$ and $\int_0^\infty g(\eta)d\eta = 1$.

The methods outlined in chapters four and five would be well suited to this problem and the performance would be expected to be similar.

Another natural direction of further work is to apply the numerical methods developed in this thesis to other financial derivative problems. For example, a convertible bond is a coupon paying instrument which allows the investor to convert the bond into a fixed number of shares before maturity. Also, if conversion does not occur, the bond is redeemed for a fixed amount of cash (at par). There are certain conditions on the price of convertible securities which make them similar to the condition on American options.

Although the focus was on Asian options in this thesis, similar PDE's (no diffusion in one of the space-like directions) occur in certain interest rate models and hence the methods developed here will be applicable to these cases as well.

Bibliography

[AABR98] L Andersen, J Andreasen, and R Brotherton-Ratcliffe. The passport option. *Journal of Computational Finance*, 1(3):15–36, 1998.

[AW95] J Abate and W Whitt. Numerical inversion of Laplace transforms of probability distributions. *ORSA Journal of Computing*, 7:36–43, 1995.

[BAW87] G Barone-Adesi and R Whaley. Efficient analytic approximation of American option values. *The Journal of Finance*, 62, 1987.

[BBG97] P. P Boyle, M Broadie, and P Glasserman. *Monte Carlo Methods for Security Pricing*, pages 1267–1321. 1997.

[Boy77] P. P Boyle. Options: A Monte Carlo approach. *Journal of Financial Economics*, 4:323–338, 1977.

[BP96] J Barraquand and T Pudet. Pricing of American path-dependent contingent claims. *Mathematical Finance*, 6(1):17–51, 1996.

[Bre79] M. J Brennan. The pricing of contingent claims in discrete time models. *Journal of Finance*, 34:53–68, 1979.

[BS73] F Black and M Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[BS77a] M. J Brennan and E. S Schwartz. Convertible bonds: Valuation and optimal strategies for call and conversion. *The Journal of Finance*, 32, December 1977.

[BS77b] M. J Brennan and E. S Schwartz. The valuation of American put options. *The Journal of Finance*, 32, May 1977.

[BS78] M. J Brennan and E. S Schwartz. Finite difference methods and jump processes arising in the pricing of contingent claims: a synthesis. *Journal of Financial and Quantitative Analysis*, 13:462–474, 1978.

[Cat98] L Cathcart. The pricing of floating rate instruments. *Journal of Computational Finance*, 1(4):31–51, 1998.

[CDLL97] D Crann, A. J Davies, C-H Lai, and S. H Leong. Time domain decomposition for European options in financial modelling. In *Tenth International Conference on Domain Decomposition Methods*, August 1997.

[CDR+00] D Crann, A. J Davies, S Rout, C-H Lai, and A. K Parrott. A distributed algorithm for European options in financial modelling. In *Twelfth International Conference on Domain Decomposition Methods*, 2000.

[CHP00] M Craddock, D Heath, and E Platen. Numerical inversion of Laplace transforms: a survey of techniques with applications to derivative pricing. *Journal of Computational Finance*, 4(1), 2000.

[CHW85] L Clewlow, S Hodges, and N Webber. *Options Markets*. Prentice-Hall, 1985. ISBN: 0136382053.

[Cla98] N. A. L Clarke. *The Numerical Solution of Financial Derivatives*. PhD thesis, University of Oxford, 1998.

[Cou82] G Courtadon. A more accurate finite difference approximation for the valuation of options. *Journal of Financial and Quantitative Analysis*, 17:697–703, December 1982.

[CP99] N. A. L Clarke and A. K Parrott. Multigrid for American option pricing with stochastic volatility. *Applied Mathematical Finance*, 6:177–195, 1999.

[CR85] J. C Cox and M Rubinstein. *Options Markets*. Prentice-Hall, 1985. ISBN: 0136382053.

[Cra96] D Crann. The Laplace transform: Numerical inversion for computational methods. Technical Report 21, University of Hertfordshire, July 1996.

[dFL05] Y d'Halluin, P. A Forsyth, and G Labahn. A Semi-Lagrangian approach for American Asian options under jump diffusion. *to appear in the SIAM Journal of Scientific Computing*, 2005.

[DHR98] M. A. H Dempster, J. P Hutton, and D. G Richards. LP valuation of Exotic American options exploiting structure. *Journal of Computational Finance*, 2(1):61–84, 1998.

[Duf01] D Duffie. *Dynamic Asset Pricing Theory.* Princeton University Press, third edition, 2001. ISBN: 069109022X.

[Dup98] B Dupire, editor. *Monte Carlo: Methodologies and Applications for Pricing and Risk Management.* Risk Publications, London, 1998.

[EO82] C. M Elliott and J. R Ockendon. *Weak and Variational Methods for Free and Moving Boundary Problems.* Pitman, Boston, 1982.

[FMW99] M. C Fu, D. B Madan, and T Wang. Pricing continuous Asian options: a comparison of Monte Carlo and Laplace transform inversion methods. *Journal of Computational Finance*, 2(2):49–74, 1998/99.

[Fri75] A Friedman. *Stochastic Differential Equations with Applications.* Academic Press, 1975. ISBN: 0122682017.

[Fri88] A Friedman. *Variational Principles and Free Boundary Probelms.* Robert E. Krieger Publishing, Florida, 1988.

[Fus04] G Fusai. Pricing Asian options via Fourier and Laplace transforms. *Journal of Computational Finance*, 7(3), 2004.

[FV02] P. A Forsyth and K. R Vetzal. Quadratic convergence for valuing American options using a penalty method. *SIAM Journal on Scientific Computing*, 23(6):2095 – 2122, 2002.

[Gar94] C. W Gardiner. *Handbook of Stochastic Methods.* Springer-Verlag, reissued second edition, 1994. ISBN: 3540156070.

[GJ84] R Geske and H. E Johnson. The American option valued analytically. *The Journal of Finance*, 39, 1984.

[GS85] R Geske and K Shastri. Valuation by approximation: A comparison of alternative option valuation techniques. *Journal of Financial and Quanititative Analysis*, 20:45–71, 1985.

[GY93] H Geman and M Yor. Bessel processes, Asian options, and perpetuities. *Mathematical Finance*, 3(4):349–375, 1993.

[HP81] J. M Harrison and S. R Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications*, 11:215–260, 1981.

[HR98] D. G Hobson and L. C. G Rogers. Complete models with stochastic volatility. *Mathematical Finance*, 8:27–48, 1998.

[Hul03] J Hull. *Options, Futures and Other Derivatives.* Prentice Hall International, fifth edition, 2003. ISBN: 0130465925.

[HW87] J Hull and A White. The pricing of options with stochastic volatilities. *Journal of Finance*, 3:281–300, 1987.

[HW90] J Hull and A White. Valuing derivative securities using the explicit finite difference method. *Journal of Financial and Quanititative Analysis*, 25(1):87–100, March 1990.

[HZ00] S Heston and G Zhou. On the rate of convergence of discrete-time contingent claims. *Mathematical Finance*, 10:53–75, 2000.

[Ing87] J Ingersoll. *Theory of Financial Decision Making.* Roman & Littlefield, 1987.

[Jar87] R. A Jarrow. *Option Pricing.* Irwin, 1987. ISBN: 0870943782.

[JS87] S Johnson and D Shanno. Option pricing when the variance is changing. *Journal of Financial and Quantitative Analysis*, 22:143–151, 1987.

[KV90] A. G. Z Kemna and A. C. F Vorst. A pricing method for options based on average asset values. *Journal of Banking and Finance*, 14:113–129, 1990.

[Lev92] E Levy. Pricing European average rate currency options. *Journal of International Money and Finance*, 11:474–491, 1992.

[LPRH05] C-H Lai, A. K Parrott, S Rout, and M. E Honnor. A distributed algorithm for European options with nonlinear volatility. *Computers and Mathematics with Applications*, 49:885–894, 2005.

[Mer73] R. C Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4:141–183, 1973.

[Mer76] R. C Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

[Mer92] R. C Merton. *Continuous-Time Finance.* Blackwell Publishers, 1992. ISBN: 0631185089.

[MM94] K. W Morton and D. F Mayers. *Numerical Solution of Partial Differential Equations : An Introduction.* Cambridge University Press, 1994. ISBN: 0521429226.

[MT90] A Melino and S. M Turnbull. Pricing foreign exchange options with stochastic volatility. *Journal of Econometrics*, 45:239–265, 1990.

[Myn92] R Myneni. The pricing of the American option. *The Annals of Applied Probability*, 1992.

[Nef96] S. N Neftci. *An Introduction to the Mathematics of Financial Derivatives.* Academic Press, 1996. ISBN: 0-12-515390-2.

[NST02] B. F Nielsen, O Skavhaug, and A Tveito. Penalty and front-fixing methods for the numerical solution of American option problems. *Journal of Computational Finance*, 5(4):69–97, 2002.

[Par77] M Parkinson. Option pricing: The American put. *Journal of Business*, 50:21–36, 1977.

[PC99] A. K Parrott and N. A. L Clarke. The parallel solution of early-exercise Asian options with stochastic volatility. In *Eleventh International Conference on Domain Decomposition Methods*, pages 431–438, 1999.

[PVF01] D. M Pooley, K. R Vetzal, and P. A Forsyth. Digital projection. Technical report, University of Waterloo, Department of Computer Science, January 2001.

[RB96] A Rennie and M Baxter. *Financial Calculus: An Introduction to Derivative Pricing.* Cambridge University Press, 1996. ISBN: 0521552893.

[Red97] K Redhead. *Financial Derivatives: An Introduction to Futures, Forwards, Options and Swaps.* Prentice Hall, 1997. ISBN: 0-13-241399-X.

[RS95] L. C. G Rogers and Z Shi. The value of an Asian option. *Journal of Applied Probability*, 32:1077–1088, 1995.

[Rub85] M Rubinstein. Nonparametric tests of alternative option pricing models. *Journal of Finance*, 60:455 – 480, 1985.

[RW00] L. C. G Rogers and D Williams. *Diffusions, Markov Processes and Martingales: Ito Calculus.* Cambridge University Press, 2000. ISBN: 0521775930.

[RY99] D Revuz and M Yor. *Continuous Martingales and Brownian Motion.* Springer-Verlag, 1999. ISBN: 3540643257.

[SC91] A Staniforth and J Cote. Semi-Lagrangian integration schemes for atmospheric models - a review. *Monthly Weather Review*, 119:2206–2223, 1991.

[Sch77] E. S Schwartz. The valuation of warrants: Implementing a new approach. *Journal of Financial Economics*, 4:79–94, 1977.

[Sch86] R Scholz. Numerical solution of the obstacle problem by the penalty method. part ii. time-dependent problem. *Numerische Mathematik*, 49:255–268, 1986.

[Shi95] D Shimko. *Finance in Continuous Time: A Primer.* Blackwell Publishers, 1995. ISBN: 1878975072.

[Ste70] H Stehfast. Numerical inversion of Laplace transforms. *Comm ACM*, 13, 1970.

[Tay86] S Taylor. *Modelling Financial Time Series.* Wiley, 1986.

[TR00] D Tavella and C Randall. *Pricing Financial Instruments: The Finite Difference Method.* John Wiley & Sons, 2000. ISBN: 0-471-19760-2.

[Twi84] E. H Twizell. *Computational Methods for Partial Differential Equations.* Ellis Horwood Limited, 1984. ISBN: 0-85312-383-7.

[Vec01] J Vecer. A new pde approach for pricing arithmetic average Asian options. *Journal of Computational Finance*, 4(4):105–113, 2001.

[Wat81] E. J Watson. *Laplace Transforms and Applications.* Van Nostrand Reinhold Company Ltd., 1981. ISBN: 0-442-30176-6.

[WD93] P Wilmott and J Dewynne. Partial to the exotic. *Risk*, pages 38–46, March 1993.

[WDH93] P Wilmott, J Dewynne, and S Howison. *Option Pricing: Mathematical Models and Computation.* Oxford Financial Press, 1993. ISBN: 0952208202.

[WHD97] P Wilmott, S Howison, and J Dewynne. *The Mathematics of Financial Derivatives.* Cambridge University Press, 1997. ISBN: 0-521-49789-2.

[Wid46] D. V Widder. *The Laplace Transform.* Princeton University Press, 1946.

[Wil98] P Wilmott. *Derivatives: The Theory and Practice of Financial Engineering.* John Wiley & Sons Ltd., 1998. ISBN: 0-471-983 66-7.

[ZFV98a] R Zvan, P. A Forsyth, and K. R Vetzal. Robust numerical methods for pde models of Asian options. *Journal of Computational Finance*, 1(2):39–78, 1997/1998.

[ZFV98b] R Zvan, P. A Forsyth, and K. R Vetzal. Penalty methods for American options with stochastic volatility. *Journal of Computational and Applied Mathematics*, 91, 1998.

[ZFV99] R Zvan, P. A Forsyth, and K. R Vetzal. A finite volume approach for contingent claims valuation. Technical report, University of Waterloo, Department of Computer Science, October 1999.