Universiteit van Amsterdam
Amsterdam School of Economics

Numerical Methods for pricing Options under Stochastic Interest Rates

Terry Ringlever

Master's Thesis to obtain the degree in
Actuarial Science and Mathematical Finance
University of Amsterdam
Faculty of Economics and Business
Amsterdam School of Economics

Author:	Terry Ringlever
Student nr:	10626484
Email:	tdk_ringlever@hotmail.com
Date:	August 14, 2017
Supervisor:	Prof. Dr. Ir. M.H. Vellekoop
Second reader:	Prof. Dr. R.J.A. Laeven

Statement of Originality

This document is written by Student Terry Ringlever who declares to take full responsibility for the contents of this document. I declare that the text and the work presented in this document are original and that no sources other than those mentioned in the text and its references have been used in creating it. The Faculty of Economics and Business is responsible solely for the supervision of completion of the work, not for the contents.

Abstract

Low (or even negative) bond yields necessitate the extension of standard methods for option pricing to deal with stochastic interest rates. The main focus of this thesis is to make this possible by means of finite difference schemes. To this end, we employ the traditional Crank Nicolson method and an Alternating Direction Implicit Method. We present a detailed description of the derivation of these schemes and we apply these methods towards the valuation of European and American options where the underlying follows a geometric Brownian motion and the dynamics of the interest rate is determined by a one-factor Hull White model. The time-dependent parameter of the Hull-White model is calibrated by using an initial term structure. We validate the results by using the closed form solution for zero-coupon bonds and we draw conclusions about the sensitivity to the uncertainty in short rates.

Keywords Brownian motion, Hull-White model, European Options, Calibration, ADI method, Crank Nicolson method, Mathematics, Finance, Stochastic processes, Zero-Coupon Bonds, American Options, PSOR algorithm.

Contents

Acknowledgments

I wish to thank my supervisor Michel Vellekoop for his guidance throughout this project and advice on several occasions. Without his valuable input and patience this thesis could not have been succesfully conducted. I also would like to thank my friends and family for their support.

Chapter 1

Introduction

In 1973 a groundbreaking paper " the pricing of options and corporate liabilities " was published [1]. The authors, Fischer Black and Myron Scholes developed a framework that laid the foundation for current option pricing theory. Before their paper was published there wasn't any standard model that could fairly value an option. Options were back then regarded as a risky undertaking and therefore not as popular as today. One would think that a model in which options could be valued was therefore more than welcome. However Black and Scholes still had a hard time getting their paper published. After numerous rejections, they eventually succeeded. In their groundbreaking paper, they suggest that an option can have one price only. This price is determined by their Black-Scholes formula or Black-Scholes equation as we call it nowadays. Later on, Robert Merton expanded their pricing model in the paper "Theory of rational option pricing" [2]. From then on, a quantitative approach towards option pricing popularized the trading in these contracts. These revolutionary insights of the previous two papers are of such an importance that Robert Merton and Myron Scholes received the Nobel Prize in Economics. To this day, investors and traders still use their ideas. This mathematical approach was so successful that it opened up a whole new branch of research called 'financial engineering'.

The Black-Scholes model (BSM) has some underlying assumptions under which a replicating strategy, which makes the value of the derivative fair to both buyer and seller, can be determined. At the same time, many of these assumptions form a weakness of the model in comparison to real-world situations. For instance, the BSM assumes a constant deterministic interest rate. This is not a realistic assumption as interest rates change over time. In the past it was common to assume interest rates to be positive quantities. However in 2012, it has become apparent that interest rates could even drop below zero. These phenomena necessitate a modification of the classical BSM. Since constant deterministic interest rates turn out not to be suitable we regard a stochastic interest rate following the famous Hull White model. This Hull White model captures the mean reversion in short rates and allows interest rates to become negative. It is a popular model containing three parameters of which one is time-dependent. This time-dependent parameter needs to be calibrated by using the initial term structure.
Besides the constant interest rate assumption, the BSM uses a constant volatility for the underlying asset. There is a famous modification that incorporates stochastic volatility in the model, namely the Heston model. It has the attractive features of explaining the volatility smile along with mean reversion of the volatility. It is also possible to combine the two ingredients into a hybrid model in which both the interest rate as well as the volatility are stochastic [3][4]. These models are however outside the scope of this thesis.

1.1 Research Question

There are many methods developed throughout the years to calculate the value of an option. One approach is by means of numerical methods. These methods have the advantage that they provide for a whole range of asset values a price. This makes such methods attractive in the valuation of options. However, one is often interested in cases where there is not only a variety in asset prices present, but also a variation in volatilities, interest rates or even both. Naturally, this leads for traditional numerical methods to computational difficulties and time consuming problems. To this end, there are more advanced methods that enable us to simplify the problem that we face in traditional numerical methods. In this thesis we restrict ourselves to two methods, namely the Crank Nicolson (CN) method and the Alternating Direction Implicit (ADI) method. Further, we consider for both methods two-dimensional problems. In particular, we take a stochastic interest rate into account. To obtain a thorough understanding of the differences between the two approaches we state the following research question: "For which options does the ADI method outperform the CN method in a model with stochastic interest rates?". In particular, we are interested in the differences between the two methods and how these differences are expressed in the particular case of stochastic interest rates.

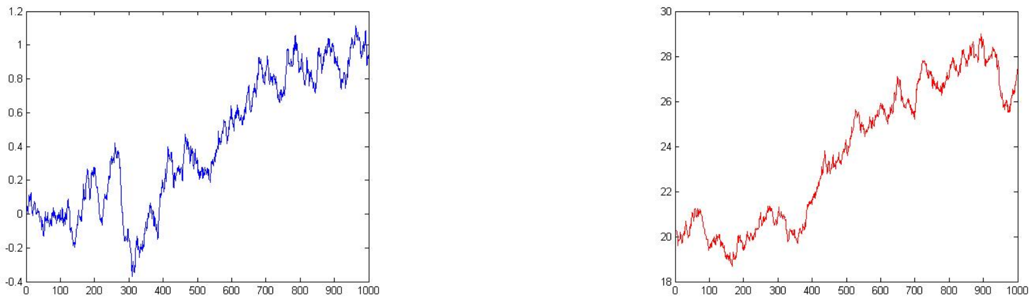
1.2 Outline

This thesis is organized as follows. Chapter 2 starts with a detailed explanation of the setting in which we will employ several numerical methods. The Black Scholes equation and the Hull White model will be introduced along with the basics of Linear Complementarity problems. In chapter 3, we introduce the Crank Nicolson method by applying it in the valuation of European and American options under a fixed interest rate. From there we move on to a setting with stochastic interest rates where we introduce more advanced methods for option pricing. These methods will then be applied to value zero-coupon bonds, European options, Binary options and American options. In chapter 4 we present our results and compare them with an explicit solution whenever possible. The thesis ends with chapter 5 in which we draw a conclusion and provide the reader with a recommendation.

Chapter 2

Model and Assumptions

Stochastic processes play an important role in many areas of research. In financial engineering there is one type of stochastic process that is of particular interest: the Wiener process. It all started with Robert Brown, who observerd that tiny particles in a liquid move and change directions due to collisions. Back then the movements of these molecules couldn't be explained. This behaviour is nowadays referred to as Brownian motion. It took until 1923, when Norbert Wiener developed successfully a model for Brownian motion. The process describing this Brownian motion is called a Wiener process. It can be defined 'rigorously' by defining it as a stochastic process $W(t)$ which for $t \geq 0$ satisfies (1) $W(0) = 0$, (2) $W(t) \sim N(0, t)$, (3) has independent increments for disjunct time intervals and (4) it has continuous paths $t \mapsto W(t)$. A sample path of such a Wiener process is presented in the figure below.



(a) Wiener process path

(b) Geometric Brownian Motion path

Figure 2.1

The blue graph in the left figure shows a irregular pattern which is also observed in the stock market. This process is often applied in finance by means of related stochastic processes such as the Geometric Brownian Motion (GBM). A Geometric Brownian Motion is given by the following stochastic differential equation

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$$

with $W(t)$ a Wiener process. The parameters μ and σ denote the drift and volatility. The drift is related to the change of the average value while the volatility is related to the variation. One of the characteristics of GBM is that $S(t)$ cannot become negative. This makes it in particular suitable for modeling stock prices. This stochastic differential equation can be solved, resulting in the analytical formula:

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

This makes modeling many asset paths easy. The figure above also presents a red graph. This graph is obtained by using the analytical formula with $\sigma = 0.2, \mu = 0.1$ and

$S(0) = 20$. The derivation of the analytic solution of GBM presented above relies on the use of Ito Calculus. It can be found for example in [5]. One of the technicalities used in this derivation is called Ito's Lemma. It yields a chain rule for random variables. Its form depends on the number of random variables considered. Below we state the multivariate case [6]:

Ito's Lemma: Let f be a twice continuously differentiable function $\mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ where $X = (X_1, \ldots, X_n)^T$ is n-dimensional that follows the stochastic process:

$$dX_t = M_t dt + N_t dW_t$$

with $M = (M_1, \ldots, M_n)^T$ a vector and $N = (a)_{i,j}$ a $n \times m$ matrix and W a vector of m independent Brownian motions. then $Y_t = f(X_t, t)$ follows a stochastic process of the same form and

$$dY_t = \left(f_t(X_t, t) + M_t f_x(X_t, t) + \frac{1}{2} tr \left[N_t N_t^t f_{xx}(t, X_t) \right] \right) dt + f_x(X_t, t) N_t dW_t$$

where we used the notation:

$$f_x(x,t) := \left(\frac{\partial f}{\partial x_i}(x,t), \ldots, \frac{\partial f}{\partial x_n}(x,t) \right) \quad \text{and} \quad f_{xx}(x,t) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x,t) \right)_{1 \le i,j \le n}$$

The proof of this result can be found in many textbooks on stochastic calculus. It is also mentioned in [5]. It is a magnificent result and is frequently used when solving stochastic differential equations. In addition it laid a mark in the derivation of the famous Black-Scholes equation which can be derived by invoking Ito's Lemma with one underlying stochastic process. We will encounter this result later on in a more general setting when deriving a Black-Scholes equation under stochastic interest rates.

2.1 Black-Scholes equation

The Black-Scholes equation changed the way we look at option valuation. Despite its fame, it still stands in a world far from reality. In the introduction we already mentioned a few assumptions underlying the BSM. Besides the deterministic volatility σ and constant interest rate r it is assumed that the no-dividend paying stock $S(t)$ has a log-normal distribution. In addition, the model is based on a liquid market without transaction costs. Every derivative, depending on the underlying and time, fulfilling the above mentioned assumption satisfies the Black-Scholes equation:

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

It is a backward parabolic partial differential equation describing the dynamics of the option price. A full derivation depends on Ito's Lemma and can be found in [5] [7]. At first glance, it isn't clear what the solution of this partial differential equation will look like. However, it is possible to transform it into a heat equation. This was already realized by Black, Scholes and Merton. And the solution of the heat equation was already known.
Many of the aforementioned assumptions aren't realistic. Therefore we will adjust it to a more appropriate setting where the fixed interest rate will turn into a stochastic process. Despite its deficiencies, it is still an important formula and we will also use it to explain several concepts.

2.2 European Options

Under the Black Scholes model it is possible to value European options. These options give the owner the right, but not the obligation, to buy or sell the underlying at maturity at a prescribed price called the strike price. If there is the possibility to buy the underlying asset at maturity one speaks of a European Call option. On the other hand, if there is solely the ability to sell at the strike price one speaks of a European Put option. Denoting by $C(S(t), t)$ the price of a call option at time t with underlying S and similarly $P(S(t), t)$ for a put option we can make these final conditions mathematically precise:

$$C(S(T), T) = max(S(T) - K, 0) \quad \text{and} \quad P(S(T), T) = max(K - S(T), 0).$$

By using the Black-Scholes equation along with the conditions at expiration date it is possible to derive the Black-Scholes formula:

$$C(S(t), t) = N(d_1)S(t) - N(d_2)Ke^{-r(T-t)} \quad \text{and} \quad P(S(t), t) = N(-d_2)Ke^{-r(T-t)} - N(-d_1)S(t)$$

where we denote with K the strike price, $T - t$ the time until maturity of the option and $N(\cdot)$ the standard normal cdf.

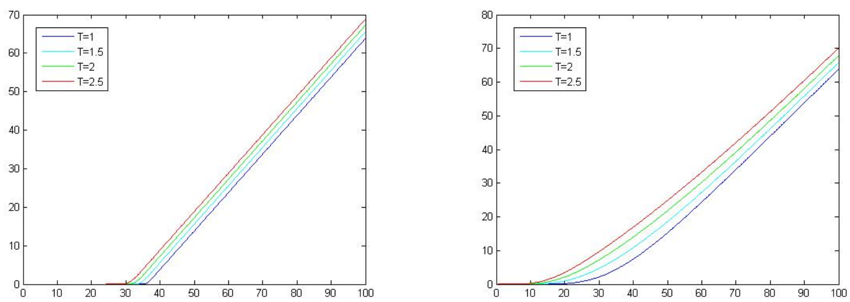
$$d_1 = \frac{ln\left(\frac{S(t)}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{T-t}$$

These expressions satisfy the so-called Put-Call parity:

$$C(S(t), t) + Ke^{-r(T-t)} = P(S(t), t) + S(t)$$

which describes the relation between the call option value and put option value with identical strike prices at every time instant. This makes it easy to obtain the European Call value when the European Put value is known and vice versa.

In order to get a feeling on how a European option depends on its parameters we show diagrams regarding a European Call option. We depicted a few graphs corresponding to parameters sets where the maturity level ranges from 1 to 2.5 while the other parameters are given by: $S(0) = 0, K = 40, r = 0.1$.



(a) European Call option: $\sigma = 0.02$ (b) European Call option: $\sigma = 0.35$

Figure 2.2

The results show how the value of a European Call option depends on both maturity and volatility. The payoff resembles the shape of a hockey stick. This shape is also present in the figures above. Notice that the longer the owner has until expiration date, the higher the value of the option. Also an increase in volatility flattens the payoff function.

2.3 American Options

The Black Scholes equation describes the dynamics of the value of European Options, i.e. options that can be exercised only at expiration date. There are however options with the possibility of an early exercise, such as American options. American options give the owner the right to exercise the option before maturity. This is a useful feature in practice, but at the same time it can make things more complicated when it comes to valuation. However, for American Call options on a non-dividend stock, its value equals that of a European call option. For these options we know that there is a closed-formula in a deterministic setting. The reason behind the identical value boils down to the fact that it is never optimal to exercise the option before maturity. Since due to the time-value of money the costs of exercising the option at an early time is higher than in the future. In addition, holding an in-the-money call option is as favourable as holding the asset. However when things turn for the worse and the asset moves below the strike price, holding a Call option option is advantageous since it protects the owner against these downward movements.

For American call options on a divided-paying stock, the setting becomes different. To this end, suppose an American call option is exercised at an early stage. Then one has a position in the underlying stock, but the stock pays out dividends which before exercising the option one wouldn't have the right to receive. This suggests that some sort of premium must be involved in order to make up for the discrepancy between holding on the option or early exercise.
Regarding the American put option, the early exercise opportunity turns it into a more valuable derivative than the European put option. This can easily be seen when one considers the situation where a stock price is close to zero. In that situation the default risk needs to be taken into account and early exercise becomes attractive.

Recall that it is important to know when to exercise the American Call option on a dividend-paying stock at an early stage. This suggests that there is some unknown boundary that separates the situation in which it is attractive to exercise before expiration date or not. Therefore we can view the problem at hand as a free-boundary problem, meaning that there is a time-dependent boundary that indicates when the owner should exercise or not. This boundary divides our domain into two regions. One region in which the holder should exercise the option and another in which the holder needs to hold the option. In the literature one often takes the parallel to the obstacle problem which addresses the free-boundary problem [5] [7]. Our approach will be the same, but with the main purpose of introducing the technicalities that come into play.

2.3.1 The Obstacle Problem

Consider a rope with fixed endpoints at (-1,0) and (1,0) in the Euclidean plane. Suppose a smooth object depicted by $f(\cdot)$ is in between those two points and the rope falls on top of this object while it is stretched and under high tension, see figure 2.3. The left and right contact points of the rope with the obstacle, denoted by respectively M and N, are unknown in advance. The problem is to find the equilibrium position of this rope. Denoting the position of the rope by $u(\cdot)$, we make it mathematically precise by,

$$\begin{aligned} u(-1) = u(1) &= 0 \\ u(x) &\geq f(x) \\ u''(x) &\leq 0 \\ (u(x) - f(x))u''(x) &= 0 \end{aligned}$$

To tackle this problem we first discretize on a mesh $-1 = x_0 < x_1 < \ldots < x_{n-1} < x_n = 1$ by using for convenience a central difference scheme. With $u_i \approx u(x_i)$ this suggests that for all $i \in \{0, \ldots, n\}$

$$u_0 = u_n = 0, \quad u_i \geq f(x_i), \quad \frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta x^2} \leq 0 \text{ and } (u_i - f(x_i))\frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta x^2} = 0$$

holds. This is known as Linear Complementarity Problem (LCP) which in compact form is written as [8]:

$$A\mathbf{x} \geq \mathbf{b}$$
$$\mathbf{x} \geq \mathbf{c}$$
$$(\mathbf{x} - \mathbf{c})'(A\mathbf{x} - \mathbf{b}) = 0$$

The matrix A is a tridiagonal matrix with on its main diagonal 2 and first upper and lower diagonal -1. The vector **c** depends on the obstacle. Taking as an illustration $f(x) = 0.25 - x^2$, the indices of **c** are given by $f(x_i)$. The vector **b** is a zero vector in this particular problem.
This LCP problem can be solved using the PSOR algorithm [9] [10]. The PSOR algorithm is an extension of the SOR algorithm. The SOR algorithm is similar to the damped Gauss-Seidel algorithm. It has a parameter ω that serves as weight such that after a Gauss-Seidel step the solution is updated by taking an average of the solution before and after the Gauss-Seidel step. The PSOR algorithm deviates from the SOR algorithm where the constraint $\mathbf{x} \geq \mathbf{c}$ needs to be satisfied. This constraint is satisfied by taking at the end of an iteration the maximum of the SOR solution and the vector **c**. Below we present a pseudo code for the PSOR algorithm.

Projected SOR Algorithm

Data: Declare $A, \mathbf{f}, \omega$ and initial guess $\mathbf{x}_0$
Data: setup boolean **bool** $= 1$
while bool do
  for $i \leftarrow 1$ **to** n **do**
    xtemp $\leftarrow$ x(i)
    x(i) $\leftarrow [b(i) - A(i, 1 : i-1) \cdot u(1 : i-1) - A(i, i+1 : n) \cdot u(i+1 : n)]/A(i, i)$
    x(i) $\leftarrow \max(c(i), xtemp(i) + \omega \cdot [x(i) - xtemp(i)])$
  end
  if *norm(xtemp* $-$ *x)* $<$ *tol* **then**
    set **bool** $= 0$;
  end
end

The variable *xtemp* is an auxiliary variable to ensure that the algorithm stops whenever the distance between the iterates fall within a prespecified tolerance **tol**. The second line within the for-loop is the same as a single Gauss-Seidel step. The solution to the specific problem is given below, where the red graph is the obstacle and the blue graph the rope. We took $n = 52$ and tolerance $1e^{-4}$ along with $\omega = 1$.

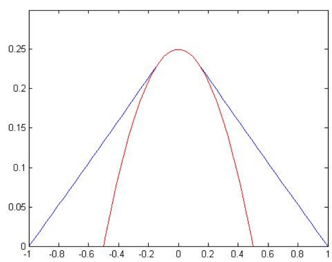


Figure 2.3: The Obstacle Problem

2.4 The Hull-White model

In reality interest rates are not fixed, they evolve over time. Since the phenomenon of negative interest rates has actually occurred, it has become more important to find an appropriate model which includes this feature. Luckily, there was already a model out there, called the Hull White model [11], developed by John Hull and Alan White. This model, which was introduced in 1990, is not the first model of future interest rates. Predecessors such as, the Vasicek model, the Cox, Ingersoll and Ross (CIR) model and the Ho-Lee model already succeeded in the modeling of future interest rates. However, these models all have their disadvantages. First, the CIR model cannot cope with negative interest rates. Back in the day, this was a desired feature as economists didn't think that interest rates could cross the zero lower bound. However, in 2012 it did happen in Denmark. Later on, it happened in other countries such as Japan and Switzerland as well. Secondly, the Vasicek model is not able to provide an exact fit of the initial term structure. These problems were solved after the introduction of the Ho-Lee model. But even this model has a drawback, namely that it doesn't incorporate the mean reversion of long term interest rates, i.e. it doesn't converge to a value in expectation over a long time period. A short rate model that provides us with all three of these properties is the Hull White model. The Hull White model was introduced in [11] where it was referred to as an extended Vasicek model. In this thesis we take the one-factor Hull White (HW) model which can be formulated as a stochastic process by:

$$dr(t) = [\theta(t) - \alpha r(t)]\, dt + \sigma dW(t), \tag{2.1}$$

where $r(t)$ is the short-term interest rate, σ its volatility, $W(t)$ a Wiener process, $\theta(t)$ a time-dependent drift and α denotes a mean-reversion speed parameter, i.e. it determines the curvature of the volatility structure.

A possible extension to this model it the two-factor Hull white model. This model contains a richer structure since there is given room for a correlation between forward rates. For more details, the reader is advised to take a look in [12]. For convenience, we only consider the one-factor HW model in this thesis.

In the one-factor HW model we need to calibrate the time-dependent parameter $\theta(t)$ along with the constant parameters α and σ such that it fits the initial term structure. It is common to let the parameters α and σ be chosen by the user of the HW model. To this end, one often chooses $\alpha \in [0, 0.25]$ and $\sigma \in [0, 0.05]$. In this thesis we follow common practice and only calibrate the time-dependent parameter. This is done by using an initial term structure along with the formula [12]:

$$\theta(t) = \frac{\partial f(0,t)}{\partial t} + \alpha f(0,t) + \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}), \tag{2.2}$$

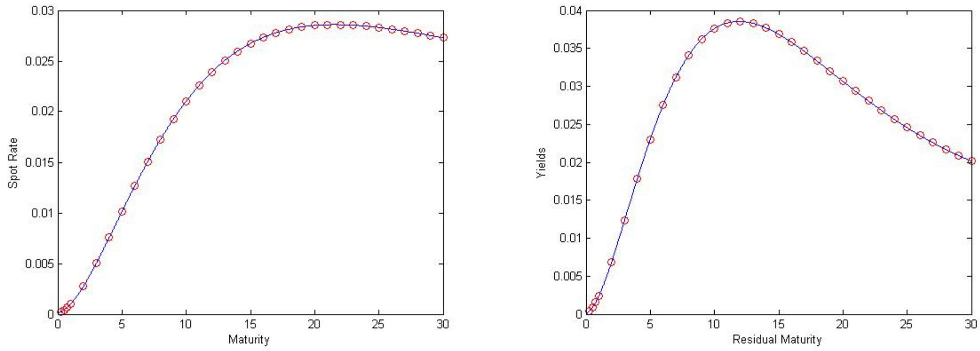
where, $f(t,T)$ denotes the instantaneous forward interest rate at time t for a zero-coupon bond paying 1 at maturity T. The instantaneous forward rate can be expressed in terms of the value of a zero-coupon bond with maturity T as follows,

$$f(t,T) := \lim_{S \to T^+} F(t;T,S) = -\frac{\partial ln P(t,T)}{\partial T} \quad \text{or} \quad P(t,T) = exp\left(-\int_t^T f(t,u)du\right). \tag{2.3}$$

In order to do the calibration we need some data. The data used here is provided by the European Central Bank and we intend to use the spot rates of triple A bonds on 1 july 2013. The spot rates along with their corresponding maturities are presented in the table below. The data in this table is multiplied by a factor 100.

Mat.	Spot Rate	Mat.	Spot Rate	Mat.	Spot Rate	Mat.	Spot Rate
0.25	0.022244	7	1.507070	16	2.731644	25	2.831694
0.5	0.041950	8	1.727866	17	2.778209	26	2.815317
0.75	0.068343	9	1.927226	18	2.812712	27	2.796483
1	0.100701	10	2.104018	19	2.836714	28	2.775723
2	0.277193	11	2.258247	20	2.851652	29	2.753490
3	0.504080	12	2.390689	21	2.858835	30	2.730171
4	0.755845	13	2.502617	22	2.859441		
5	1.014472	14	2.595599	23	2.854519		
6	1.267596	15	2.671353	24	2.844999		

The left figure below contains the data points as presented in the table above as red circles. The blue graph is the cubic spline approximation and provides a yield curve. the right figure shows the instantaneous forward curve. This was derived by using formula (2.3) along with a cubic spline approximation. Since we only have data points at discrete times, we need to interpolate to obtain the values in between the given data points. To this end, we used a cubic spline approximation. This method is very attractive since it provides us with smooth curves over the entire interval. It is also a very fast and stable method. On the other hand, it doesn't avoid overshooting at jumps. Since we don't have any jumps in our data, this method works. Again, the red circles denote the instantaneous forward rates at the maturities as given in the table above.



(a) Yield Curve

(b) Instantaneous Forward Curve

Figure 2.4

Note that in order to apply formula (2.2) we also need a second derivative of the yield curve, i.e. a derivative of the instantaneous forward curve. In order to obtain this derivative, we first determine the derivative at the red circle points of figure 2.4 (b). Afterwards, we apply a cubic spline approximation by using these points. The curve is presented below.

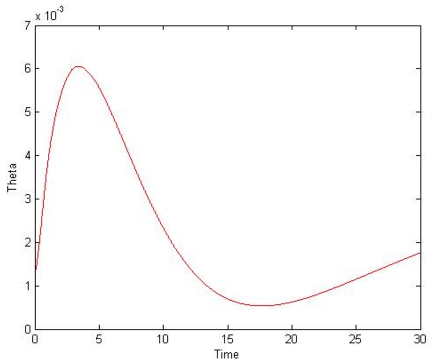


Figure 2.5: The time-dependent parameter $\theta(t)$

Note that in order to obtain such a smooth curve we only used the data points from the table. In addition, the parameters σ and α have to be calibrated. However, we choose to take $\alpha \approx 0.01$ and $\sigma \approx 0.01$. Since usually $\alpha \in [0, 0.25]$ and $\sigma \in [0, 0.05]$, it is a proper choice. The reason we fixed these two parameters is due to practical matters.

Later on, we also want to use a constant θ to simplify the model and to validate results of the models we intend to use. In this particular case, we can price a zero-coupon bond with maturity T at time t by [13]:

$$P(t,T) = exp\left(A(t,T) + B(t,T)r(t)\right) \tag{2.4}$$

where A and B are given by

$$A(t,T) = \int_t^T B(u,T)\theta(u)du + \frac{1}{2}\int_t^T B^2(u,T)\sigma^2 du \quad \text{and} \quad B(t,T) = \frac{1}{\alpha}\left(e^{-\alpha(T-t)} - 1\right)$$

These can be simplified further in the case of a constant θ by noticing that,

$$\begin{aligned} A(t,T) &= \frac{\theta}{\alpha}\int_t^T \left(e^{-\alpha(T-u)} - 1\right) du + \frac{\sigma^2}{2\alpha^2}\int_t^T \left(e^{-2\alpha(T-u)} - 2e^{-\alpha(T-u)} + 1\right) du \\ &= \frac{\theta}{\alpha}\left[\frac{1}{\alpha}e^{-\alpha(T-u)} - u\right]_t^T + \frac{\sigma^2}{2\alpha^2}\left[\frac{1}{2\alpha}e^{-2\alpha(T-u)} - \frac{2}{\alpha}e^{-\alpha(T-u)} + u\right]_t^T \\ &= \frac{\theta}{\alpha}\left[\frac{1}{\alpha} - T - \frac{1}{\alpha}e^{-\alpha(T-t)} + t\right] + \frac{\sigma^2}{2\alpha^2}\left[\frac{1}{2\alpha} - \frac{2}{\alpha} + T - \frac{e^{-2\alpha(T-t)}}{2\alpha} + \frac{2e^{-\alpha(T-t)}}{\alpha} - t\right] \end{aligned}$$

2.5 Black-Scholes equation under stochastic interest rates

To incorporate the stochastic interest rates there is a need for an adjusted Black-Scholes equation along with a model prescribing the dynamics of the stochastic interest rate. The model we use is the Hull-White model we discussed above.
The derivation of the coming partial differential equation follows a standard approach. We discuss it in detail. We start off by considering a derivative depending on the underlying asset price $S(t)$, the interest rate $r(t)$ and time t, i.e.

$$V(S(t), r(t), t).$$

Let the dynamics of the asset price $S(t)$ be given by a geometric Brownian motion:

$$dS(t) = \sigma_1 S(t)dW_1(t) + \mu S(t)dt \tag{2.5}$$

and the dynamics of the interest rate $r(t)$ by the Hull-White model

$$dr(t) = [\theta(t) - \alpha r(t)]\, dt + \sigma_2 dW_2(t) \tag{2.6}$$

where σ_1, σ_2 and α are constants. The terms $W_1(t)$ and $W_2(t)$ indicate Wiener processes that may be correlated, say

$$dW_1(t)dW_2(t) = Cov(dW_1(t), dW_2(t)) = \rho dt$$

Recall that equation (2.5) has an analytic solution:

$$S(t) = S(0)exp\left(\left(\mu - \frac{\sigma_1^2}{2}\right)t + \sigma W(t)\right). \tag{2.7}$$

First we derive Ito's formula for two random variables. To this end, consider $V(S + dS, r + dr, t + dt)$ and apply Taylor's theorem:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial r}dr + \frac{1}{2}\left(\frac{\partial^2 V}{\partial S^2}dS^2 + 2\frac{\partial^2 V}{\partial S \partial r}dSdr + \frac{\partial^2 V}{\partial r^2}dr^2\right) + \ldots$$

Using the rule $dW_1^2 = dt$ and that dt^2 and $dtdW_1$ go faster to zero than dW_1^2 we find

$$dS^2 = \sigma_1^2 S^2 dW_1^2 = \sigma^2 S^2 dt \quad \text{and} \quad dr^2 = \sigma_2^2 dW_2^2 = \sigma_2^2 dt$$

Similarly,

$$dSdr = \sigma_1\sigma_2 S dW_1 dW_2 = \rho\sigma_1\sigma_2 S dt$$

Thus, Ito's Lemma becomes:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial r}dr + \frac{1}{2}\left(\sigma_1^2 S^2\frac{\partial^2 V}{\partial S^2} + 2\rho\sigma_1\sigma_2 S\frac{\partial^2 V}{\partial S \partial r} + \sigma_2^2\frac{\partial^2 V}{\partial r^2}\right)dt$$

Consider a portfolio consisting of an option $V(S, r, t)$, and investment in $-\Delta_1$ zero-coupon bonds $B(r, t)$ and in $-\Delta_2$ underlying assets $S(t)$, then:

$$\Pi = V - \Delta_1 B - \Delta_2 S$$

Consequently, a change in this portfolio in a timestep dt is:

$$d\Pi = dV - \Delta_1 dB - \Delta_2 dS = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial r}dr + \frac{1}{2}\left(\sigma_1^2 S^2\frac{\partial^2 V}{\partial S^2} + 2\rho\sigma_1\sigma_2 S\frac{\partial^2 V}{\partial S \partial r} + \sigma_2^2\frac{\partial^2 V}{\partial r^2}\right)dt$$

$$-\Delta_1\left(\frac{\partial B}{\partial t}dt + \frac{\partial B}{\partial r}dr + \frac{1}{2}\sigma_2^2\frac{\partial^2 B}{\partial r^2}dt\right) - \Delta_2 dS \qquad (2.8)$$

Recall that the zero-coupon bond pricing equation is given by (formula 17.11 in [5])

$$\frac{\partial B}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 B}{\partial r^2} + (u(r,t) - \lambda(r,t)\sigma(r,t))\frac{\partial B}{\partial r} - rB = 0 \qquad (2.9)$$

Assume a constant market price of risk $\lambda(r, t) = \lambda$ and take according to the Hull White model $u(r, t) = \theta(t) - \alpha r(t)$ and $\sigma(r, t) = \sigma_2$. Now, equation (2.9) can be written as

$$\frac{\partial B}{\partial t} + \frac{1}{2}\sigma_2^2\frac{\partial^2 B}{\partial r^2} = rB - (\theta(t) - \alpha r(t) - \lambda\sigma_2)\frac{\partial B}{\partial r}$$

Substitution into (2.8) yields,

$$d\Pi = \left[\frac{\partial V}{\partial t} + \frac{1}{2}\left(\sigma_1^2 S^2\frac{\partial^2 V}{\partial S^2} + 2\rho\sigma_1\sigma_2 S\frac{\partial^2 V}{\partial S \partial r} + \sigma_2^2\frac{\partial^2 V}{\partial r^2}\right)\right]dt + \left[\frac{\partial V}{\partial S} - \Delta_2\right]dS + \left[\frac{\partial V}{\partial r} - \Delta_1\frac{\partial B}{\partial r}\right]dr$$

$$-\Delta_1\left[rB - [\theta(t) - \alpha r(t) - \lambda\sigma_2]\frac{\partial B}{\partial r}\right]dt$$

In order to eliminate the risk from our portfolio we choose,

$$\Delta_2 = \frac{\partial V}{\partial S} \quad \text{and} \quad \Delta_1 = \frac{\partial V}{\partial r}\bigg/\frac{\partial B}{\partial r}$$

Consequently,

$$d\Pi = \left[\frac{\partial V}{\partial t} + \frac{1}{2}\left(\sigma_1^2 S^2\frac{\partial^2 V}{\partial S^2} + 2\rho\sigma_1\sigma_2 S\frac{\partial^2 V}{\partial S \partial r} + \sigma_2^2\frac{\partial^2 V}{\partial r^2}\right)\right]dt - rB\frac{\partial V/\partial r}{\partial B/\partial r}dt + [\theta(t) - \alpha r(t) - \lambda\sigma_2]\frac{\partial V}{\partial r}dt$$

$$= r\Pi dt$$

where,

$$r\Pi dt = r\left[V - \Delta_1 B - \Delta_2 S\right] dt = r\left[V - \frac{\partial V/\partial r}{\partial B/\partial r}B - \frac{\partial V}{\partial S}S\right] dt$$

Thus, we find by rearranging terms

$$\frac{\partial V}{\partial t}+\frac{1}{2}\left(\sigma_1^2 S^2\frac{\partial^2 V}{\partial S^2} + 2\rho\sigma_1\sigma_2 S\frac{\partial^2 V}{\partial S\partial r} + \sigma_2^2\frac{\partial^2 V}{\partial r^2}\right)+\left[\theta(t) - \alpha r(t) - \lambda\sigma_2\right]\frac{\partial V}{\partial r}-r\left(V - \frac{\partial V}{\partial S}S\right) = 0 \tag{2.10}$$

Under risk-neutral pricing we can take $\lambda = 0$, such that

$$\frac{\partial V}{\partial t}+\frac{1}{2}\left(\sigma_1^2 S^2\frac{\partial^2 V}{\partial S^2} + 2\rho\sigma_1\sigma_2 S\frac{\partial^2 V}{\partial S\partial r} + \sigma_2^2\frac{\partial^2 V}{\partial r^2}\right)+\left[\theta(t) - \alpha r(t)\right]\frac{\partial V}{\partial r}-r\left(V - \frac{\partial V}{\partial S}S\right) = 0 \tag{2.11}$$

This last partial differential equation is suitable for pricing options. But it is actually easier to handle after applying the transformation $\tau = T - t$. This results in,

$$\frac{\partial V}{\partial \tau} = \frac{1}{2}\left(\sigma_1^2 S^2\frac{\partial^2 V}{\partial S^2} + 2\rho\sigma_1\sigma_2 S\frac{\partial^2 V}{\partial S\partial r} + \sigma_2^2\frac{\partial^2 V}{\partial r^2}\right) + \left[\theta(\tau) - \alpha r(\tau)\right]\frac{\partial V}{\partial r} - r\left(V - \frac{\partial V}{\partial S}S\right)$$

This transformation helps us as it basically turns our final conditions into initial conditions.

Chapter 3

Numerical Methods

In general it is hard to find an analytical solution of a partial differential equation. A way out is trying to find a numerical approximation of a solution. There are often many suitable numerical methods to solve a certain problem [14]. Nevertheless, it is still important to take for instance, convergence, accuracy and consistency into account. Here we will focus on a popular approach to tackle pde's, namely by means of finite difference methods. In other words, we will replace functions by their discrete counterparts. Thus, derivatives are replaced by finite differences. To this end, we need to divide a large enough domain into boxes in Euclidean space. Each box is then partitioned, often in an equidistant way. Doing so, we have a domain with discrete points. The finite difference approximations are then evaluated in these discrete points.

3.1 Finite Difference Schemes

A convenient way of obtaining finite difference approximations is by means of Taylor expansions. To this end, suppose we have a sufficiently smooth function in two variables, $f(x, y)$. Its Taylor expansion around (x_0, y_0) with $\Delta x := x - x_0$ and $\Delta y := y - y_0$ is given by,

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}\Delta x^2 + 2\frac{\partial^2 f}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}\Delta y^2\right] + \mathcal{O}(\Delta x^3, \Delta y^3)$$

Similarly one can find the Taylor expansion of a function in three variables, $f(x, y, t)$. In finite difference methods we have to truncate the approximation. Doing this, yields a truncation error of a particular order indicated by means of the Big-O notation. This truncation determines the accuracy of the method. As mentioned earlier, a grid is produced by considering a partition of the rectangles. Now, suppose we have a particular domain $D \subseteq \mathbb{R}^3$ where $D = [0, X] \times [0, Y] \times [0, T]$. We partition each of the intervals such that we obtain an equidistant grid in each direction. The interval lengths are denoted by Δx, Δy and Δt. These will usually differ from each other in magnitude. Introducing the notation $f^n_{i,j} = f(i\Delta x, j\Delta y, n\Delta t) = f(x_i, y_j, t_n)$ we will now derive finite difference approximations that are often used later on. We expand the functions $f^{n+1}_{i,j}$ and $f^n_{i,j}$ around $(x_i, y_j, t_{n+\frac{1}{2}})$

$$f^{n+1}_{i,j} = f(x_i,y_j,t_{n+\frac{1}{2}}) + \frac{\partial f}{\partial t}(x_i,y_j,t_{n+\frac{1}{2}})\frac{\Delta t}{2} + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(x_i,y_j,t_{n+\frac{1}{2}})\left(\frac{\Delta t}{2}\right)^2 + \frac{1}{6}\frac{\partial^3 f}{\partial t^3}(x_i,y_j,t_{n+\frac{1}{2}})\left(\frac{\Delta t}{2}\right)^3 + \mathcal{O}(\Delta t^4) \tag{3.1}$$

$$f_{i,j}^{n} = f(x_i, y_j, t_{n+\frac{1}{2}}) - \frac{\partial f}{\partial t}(x_i, y_j, t_{n+\frac{1}{2}})\frac{\Delta t}{2} + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(x_i, y_j, t_{n+\frac{1}{2}})\left(\frac{\Delta t}{2}\right)^2$$
$$-\frac{1}{6}\frac{\partial^3 f}{\partial t^3}(x_i, y_j, t_{n+\frac{1}{2}})\left(\frac{\Delta t}{2}\right)^3 + \mathcal{O}(\Delta t^4) \tag{3.2}$$

Subtracting these equations,

$$f_{i,j}^{n+1} - f_{i,j}^{n} = \Delta t\frac{\partial f}{\partial t}(x_i, y_j, t_{n+\frac{1}{2}}) + \frac{1}{24}\Delta t^3\frac{\partial^3 f}{\partial t^3}(x_i, y_j, t_{n+\frac{1}{2}}) + \mathcal{O}(\Delta t^5)$$

and rearranging yields the expression,

$$\frac{\partial f}{\partial t}(x_i, y_j, t_{n+\frac{1}{2}}) = \frac{f_{i,j}^{n+1} - f_{i,j}^{n}}{\Delta t} + \mathcal{O}(\Delta t^2) \tag{3.3}$$

Note that the time derivative is stated in $t + \frac{\Delta t}{2}$, while in an explicit or implicit method this would have been t or $t + \Delta t$ respectively. Therefore the spatial derivatives in x and y need to be expressed halfway as well. Using Taylor expansions,

$$f_{i+1,j}^{n+\frac{1}{2}} = f(x_i, y_j, t_{n+\frac{1}{2}}) + \frac{\partial f}{\partial x}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_i, y_j, t_{n+\frac{1}{2}})(\Delta x)^2$$
$$+\frac{1}{6}\frac{\partial^3 f}{\partial x^3}(x_i, y_j, t_{n+\frac{1}{2}})(\Delta x)^3 + \mathcal{O}(\Delta x^4) \tag{3.4}$$

$$f_{i,j}^{n+\frac{1}{2}} = f(x_i, y_j, t_{n+\frac{1}{2}}) \tag{3.5}$$

$$f_{i-1,j}^{n+\frac{1}{2}} = f(x_i, y_j, t_{n+\frac{1}{2}}) - \frac{\partial f}{\partial x}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_i, y_j, t_{n+\frac{1}{2}})(\Delta x)^2$$
$$-\frac{1}{6}\frac{\partial^3 f}{\partial x^3}(x_i, y_j, t_{n+\frac{1}{2}})(\Delta x)^3 + \mathcal{O}(\Delta x^4) \tag{3.6}$$

Adding equations (3.4) and (3.6) and subtracting two times equation (3.5) yields

$$f_{i+1,j}^{n+\frac{1}{2}} - 2f_{i,j}^{n+\frac{1}{2}} + f_{i-1,j}^{n+\frac{1}{2}} = \frac{\partial^2 f}{\partial x^2}(x_i, y_j, t_{n+\frac{1}{2}})(\Delta x)^2 + \mathcal{O}(\Delta x^4)$$

Thus,

$$\frac{\partial^2 f}{\partial x^2}(x_i, y_j, t_{n+\frac{1}{2}}) = \frac{f_{i+1,j}^{n+\frac{1}{2}} - 2f_{i,j}^{n+\frac{1}{2}} + f_{i-1,j}^{n+\frac{1}{2}}}{\Delta x^2} + \mathcal{O}(\Delta x^2) \tag{3.7}$$

Similarly, by subtracting (3.6) from (3.4) we find,

$$\frac{\partial f}{\partial x}(x_i, y_j, t_{n+\frac{1}{2}}) = \frac{f_{i+1,j}^{n+\frac{1}{2}} - f_{i-1,j}^{n+\frac{1}{2}}}{2\Delta x} + \mathcal{O}(\Delta x^2) \tag{3.8}$$

We emphasize that the gridpoints (x_i, y_j, t_j) are defined for $i, j, n \in S \subset \mathbf{N}$. Therefore, there is no node at $(i, j, n + \frac{1}{2})$ in our grid. This means we need approximations for the right hand side of equations (3.7) and (3.8) as well. These may be found by adding equations (3.1) and (3.2),

$$f_{i,j}^{n+\frac{1}{2}} = f(x_i, y_j, t_{n+\frac{1}{2}}) = \frac{f_{i,j}^{n+1} + f_{i,j}^{n}}{2} + \mathcal{O}(\Delta t^2) \tag{3.9}$$

Thus, after substituting equation (3.9) for $i-1$,i and $i+1$ into (3.7) and (3.8) we find,

$$\frac{\partial^2 f}{\partial x^2}(x_i, y_j, t_{n+\frac{1}{2}}) \approx \frac{f_{i+1,j}^{n+1} + f_{i+1,j}^{n} - 2f_{i,j}^{n+1} - 2f_{i,j}^{n} + f_{i-1,j}^{n+1} + f_{i-1,j}^{n}}{2\Delta x^2} \tag{3.10}$$

and

$$\frac{\partial f}{\partial x}(x_i, y_j, t_{n+\frac{1}{2}}) \approx \frac{f_{i+1,j}^{n+1} + f_{i+1,j}^{n} - f_{i-1,j}^{n+1} - f_{i-1,j}^{n}}{4\Delta x} \tag{3.11}$$

We now consider the approximation of a cross derivative

$$\frac{\partial^2 f}{\partial x \partial y}$$

Notice that the Taylor expansion in two variables x and y is given by,

$$f_{i+1,j+1}^{n+\frac{1}{2}} = f(x_i, y_j, t_{n+\frac{1}{2}}) + \frac{\partial f}{\partial x}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x + \frac{\partial f}{\partial y}(x_i, y_j, t_{n+\frac{1}{2}})\Delta y + \dots$$

$$\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(x_i, y_j, t_{n+\frac{1}{2}})\Delta y^2\right] + \dots \tag{3.12}$$

And,

$$f_{i-1,j-1}^{n+\frac{1}{2}} = f(x_i, y_j, t_{n+\frac{1}{2}}) - \frac{\partial f}{\partial x}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x - \frac{\partial f}{\partial y}(x_i, y_j, t_{n+\frac{1}{2}})\Delta y + \dots$$

$$\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(x_i, y_j, t_{n+\frac{1}{2}})\Delta y^2\right] + \dots \tag{3.13}$$

And,

$$f_{i-1,j+1}^{n+\frac{1}{2}} = f(x_i, y_j, t_{n+\frac{1}{2}}) - \frac{\partial f}{\partial x}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x + \frac{\partial f}{\partial y}(x_i, y_j, t_{n+\frac{1}{2}})\Delta y + \dots$$

$$\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x^2 - 2\frac{\partial^2 f}{\partial x \partial y}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(x_i, y_j, t_{n+\frac{1}{2}})\Delta y^2\right] + \dots \tag{3.14}$$

And,

$$f_{i+1,j-1}^{n+\frac{1}{2}} = f(x_i, y_j, t_{n+\frac{1}{2}}) + \frac{\partial f}{\partial x}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x - \frac{\partial f}{\partial y}(x_i, y_j, t_{n+\frac{1}{2}})\Delta y + \dots$$

$$\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x^2 - 2\frac{\partial^2 f}{\partial x \partial y}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(x_i, y_j, t_{n+\frac{1}{2}})\Delta y^2\right] + \dots \tag{3.15}$$

Adding equations (3.12) and (3.13) and subtracting equations (3.14) and (3.15) yields,

$$f_{i+1,j+1}^{n+\frac{1}{2}} + f_{i-1,j-1}^{n+\frac{1}{2}} - f_{i-1,j+1}^{n+\frac{1}{2}} - f_{i+1,j-1}^{n+\frac{1}{2}} = 4\frac{\partial^2 f}{\partial x \partial y}(x_i, y_j, t_{n+\frac{1}{2}})\Delta x \Delta y + \mathcal{O}(\Delta x^3, \Delta y^3)$$

Consequently,

$$\frac{\partial^2 f}{\partial x \partial y}(x_i, y_j, t_{n+\frac{1}{2}}) = \frac{f_{i+1,j+1}^{n+\frac{1}{2}} + f_{i-1,j-1}^{n+\frac{1}{2}} - f_{i-1,j+1}^{n+\frac{1}{2}} - f_{i+1,j-1}^{n+\frac{1}{2}}}{4\Delta x \Delta y} + \mathcal{O}(\Delta x^2, \Delta y^2) \tag{3.16}$$

Thus, after substituting equation (3.9) for $(i+1, j+1)$, $(i-1, j-1)$, $(i+1, j-1)$ and $(i-1, j+1)$ we find,

$$\frac{\partial^2 f}{\partial x \partial y}(x_i, y_j, t_{n+\frac{1}{2}}) \approx \frac{f_{i+1,j+1}^{n+1} + f_{i+1,j+1}^{n} + f_{i-1,j-1}^{n+1} + f_{i-1,j-1}^{n} - f_{i-1,j+1}^{n+1} - f_{i-1,j+1}^{n} - f_{i+1,j-1}^{n+1} - f_{i+1,j-1}^{n}}{8\Delta x \Delta y}$$

Combining equation (3.3) with (3.10), (3.11) and the last equation gives us enough machinery to obtain a so-called Crank-Nicolson scheme for coming problems. This numerical method has a local order of accuracy of two in time and two in space. This turns Crank-Nicolson into an attractive numerical scheme. Note that we used a central difference scheme regarding the space directions i and j, while using a middle point that is not a valid gridpoint here for the time direction. Another way to obtain the Crank Nicolson method is by using the average of a forward-time central-space method (FTCS) and backward-time central-space method (BTCS). This means that we rule out the usage of a middle point regarding the time direction, but rather use $n+1$ and n or n and $n-1$. The central difference schemes are still used regarding the space directions. The FCTS and BCTS are both first order in time and second in space. This last approach comes in handy when using a θ method as we will do in the next section.

3.2 Numerical Implementation

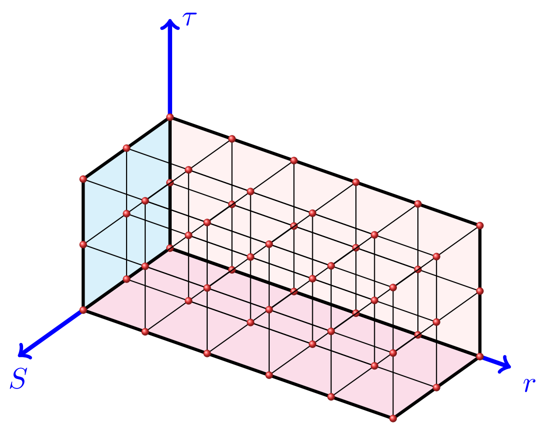
We intend to make use of the finite difference schemes as discussed above. To this end we need a domain on which we interchange the continuous operators with their discrete counterpart. The domain of interest is prescribed by $[S_0, S_{max}] \times [r_0, r_{max}] \times [0, T]$. This defines a 3-dimensional box shaped domain. We want to provide a numerical approximation within this domain. In order to do so we construct a mesh by partitioning the intervals in each direction. The interval $[S_0, S_{max}]$ is subdivided into N_S sub intervals. Similarly, the intervals $[r_0, r_{max}]$ and $[0, T]$ are partitioned into N_r and N_τ smaller intervals respectively. Therefore we have space steps and a time step defined by,

$$\Delta S = \frac{S_{max} - S_0}{N_S}, \qquad \Delta r = \frac{r_{max} - r_0}{N_r} \quad \text{and} \quad \Delta\tau = \frac{T}{N_\tau}$$

We shall often used the subscripts i, j and n in combination with S, r and τ to indicate the different gridpoints (S_i, r_j, τ_n) within our domain. Consequently, our mesh will be completely defined by the triple (i, j, n), where $i = 0, \ldots, N_S$, $j = 0, \ldots, N_r$ and $n = 0, \ldots, N_\tau$. At each of these gridpoints in our mesh we approximate a solution of the problem at hand. To make a distinction between the exact solution V of our differential equation and the numerical approximation u we introduce the notation,

$$u_{ij}^n = V(S_i, r_j, \tau^n) = V(S_0 + i\Delta S, r_0 + j\Delta r, n\Delta\tau)$$

It is good to keep in mind that we only approximate the solution V at the gridpoints. Our box-shaped region has $(N_S + 1) \cdot (N_r + 1) \cdot (N_\tau + 1)$ gridpoints. The figure below shows a typical mesh of such a region,



3.2.1 Boundary Conditions

In the following chapters we will use a number of different methods for tackling a certain problem. The problems we will face are often in the framework where we have a Black-Scholes type of partial differential equation along with a mesh as described above. Note that this mesh is only of finite size. Therefore, we need to prescribe the behaviour of our solutions at the boundaries, with the exception of the boundary regarding the initial state $\tau = 0$.

3.2.2 1-Dimensional problems

In the next section we will start by considering 1-Dimensional problems. Here we truncate the domain $[S_0, S_{max}] \times [r_0, r_{max}] \times [0, T]$ to $[S_0, S_{max}] \times [0, T]$, i.e. we use a constant interest rate. The endtime T will indicate the expiration date of the option. In addition we shall provide two boundary conditions and one final time condition. The boundary conditions we use will depend on the type of option. The final condition will equal the payoff at maturity.

3.2.3 2-Dimensional problems

Regarding the 2-dimensional problems, where we have two space dimensions and one time-dimension, we use four boundary conditions and one final condition. Our mesh is of the form $[S_0, S_{max}] \times [r_0, r_{max}] \times [0, T]$. The boundary conditions we propose are given by setting the second order derivatives equal to zero, i.e.

$$\frac{\partial^2 V}{\partial S^2}(S_0, r.\tau) = \frac{\partial^2 V}{\partial S^2}(S_{max}, r, \tau) = \frac{\partial^2 V}{\partial r^2}(S, r_0, \tau) = \frac{\partial^2 V}{\partial r^2}(S, r_{max}, \tau) = 0 \tag{3.17}$$

These boundary conditions are often referred to as 'linear boundary conditions'. It is common practice to use these type of boundary conditions on boundaries where no other boundary condition is preferred.
In physics we often have an initial condition in order to obtain a well posed problem. However, in option pricing problems we only know the payoff at maturity given a price of the underlying at that moment. Consequently, we need a final condition. This final condition depends on the type of option and will be presented below when needed.

3.3 Crank-Nicolson Method

In this section we take a brief look at a Black-Scholes equation with a constant interest rate. To price options in the Black-Scholes framework it is convenient to use a grid as described in the previous section. Since we have a constant interest rate this is to be a 1-dimensional problem. Therefore we discretize along the asset price and time direction. One of the most popular methods in pricing options on a grid is the Crank-Nicolson method. To introduce this method before applying it to a more difficult type of equation, we will price a European Call option in a 1-dimensional world.
The Crank-Nicolson scheme, often abbreviated by CN, is widely used for a broad spectrum of problems. It is an attractive method for obtaining an approximate solution due to its second order accuracy and stability properties. On the other hand, it is still possible that it shows signs of oscillatory behavior which is a feature we don't want to see. But before we dive into more details regarding its advantages and deficiencies, we will first apply the CN method to two problems. The first problem (to warm up), is a Black Scholes equation with a fixed interest rate,

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{3.18}$$

Before applying the CN-method, one should note that we will solve this problem on a confined region of our rectangular-shaped mesh. With this in mind and boundary conditions of Dirichlet type we will price a European Call option. The Call option will be out-of-the money in $i = 0$ and at the maximal stock level $i = N_S$ it equals the discounted payoff, i.e.

$$u_0^n = 0 \quad \text{and} \quad u_{N_S}^n = S_{max} - Ke^{-r(T-n\Delta t)} \tag{3.19}$$

The first boundary condition, $u_0^n = 0$ follows from the assumption $S_0 < K$. The second boundary condition $u_{N_S}^n = S_{max} - Ke^{-r(T-n\Delta t)}$ follows from the assumption $S_{max} > K$ and the put-call parity. Recall that the put-call parity is given by

$$C(S(t), t) + Ke^{-r(T-t)} = P(S(t), t) + S(t)$$

Thus, if $S_{max} > K$ then $P(S_{max}, t) = 0$. This leads to the boundary conditions as proposed at $i = N_S$. At maturity $t = T$ we have $n = N_t$. Since we consider a European Call option, the final condition will equal the payoff of this option, i.e.

$$u_i^{N_t} = max(S_0 + i\Delta S - K, 0) \tag{3.20}$$

Finally, we are ready to discretize according to the CN-method. The method boils down to taking the average of an explicit finite difference scheme (FTCS=forward in time, central in space) at gridpoint (i, n) and an implicit finite difference scheme (BTCS =backward in time, central in space) at gridpoint $(j, n+1)$, [15]. Thus,

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \frac{1}{2}\sigma^2 (S_0 + i\Delta S)^2 \left(\frac{u_{i+1}^{n+1} + u_{i+1}^n - 2u_i^{n+1} - 2u_i^n + u_{i-1}^{n+1} + u_{i-1}^n}{2\Delta S^2} \right)$$

$$+r (S_0 + i\Delta S) \left(\frac{u_{i+1}^{n+1} + u_{i+1}^n - u_{i-1}^{n+1} - u_{i-1}^n}{4\Delta S} \right) - ru_i^n = 0$$

Collecting terms involving the next time instant $n + 1$ in the left-hand side and terms involving time instant n on the right-hand side yields

$$u_i^{n+1} \left(\frac{1}{\Delta t} - \frac{1}{2}\sigma^2 \frac{(S_0 + i\Delta S)^2}{\Delta S^2} \right) +$$

$$u_{i+1}^{n+1} \left(\frac{1}{4}\sigma^2 \frac{(S_0 + i\Delta S)^2}{\Delta S^2} + \frac{1}{4}r\frac{(S_0 + i\Delta S)^2}{\Delta S} \right) + u_{i-1}^{n+1} \left(\frac{1}{2}\sigma^2 \frac{(S_0 + i\Delta S)^2}{2\Delta S^2} - r\frac{(S_0 + i\Delta S)^2}{4\Delta S} \right)$$

$$= u_i^n \left(\frac{1}{\Delta t} + \frac{\sigma^2}{2} \frac{(S_0 + i\Delta S)^2}{\Delta S^2} + r \right) + u_{i+1}^n \left(\frac{-\sigma^2}{4} \frac{(S_0 + i\Delta S)^2}{\Delta S^2} - \frac{r}{4} \frac{(S_0 + i\Delta S)^2}{\Delta S} \right)$$

$$+u_{i-1}^n \left(-\frac{\sigma^2}{2} \frac{(S_0 + i\Delta S)^2}{2\Delta S^2} + r\frac{(S_0 + i\Delta S)^2}{4\Delta S} \right)$$

This can be written in a more compact form as

$$au_i^{n+1} + bu_{i+1}^{n+1} + cu_{i-1}^{n+1} = \hat{a}u_i^n - bu_{i+1}^n - cu_{i-1}^n$$

where,

$$a = \tfrac{1}{\Delta t} - \tfrac{\sigma^2(S_0+i\Delta S)^2}{2\Delta S^2} \qquad b = \tfrac{\sigma^2(S_0+i\Delta S)^2}{4\Delta S^2} + \tfrac{r(S_0+i\Delta S)}{4\Delta S} \qquad c = \tfrac{\sigma^2(S_0+i\Delta S)^2}{4\Delta S^2} - \tfrac{r(S_0+i\Delta S)}{4\Delta S}$$

$$\hat{a} = \frac{1}{\Delta t} + \frac{\sigma^2(S_0 + i\Delta S)^2}{2\Delta S^2} + r$$

This compact form gives rise to a matrix form,

$$Au^{n+1} = Bu^n + f, \qquad \text{for } n = 0, \ldots, N_\tau$$

with,

$$A = \begin{bmatrix} a & b & 0 & \ldots & \ldots & \ldots & \ldots \\ c & a & b & 0 & \ldots & \ldots & \ldots \\ 0 & c & a & b & 0 & \ldots & \ldots \\ \ldots & \ldots & \ddots & \ddots & \ddots & \ldots & \ldots \\ \ldots & \ldots & 0 & c & a & b & 0 \\ \ldots & \ldots & \ldots & 0 & c & a & b \\ \ldots & \ldots & \ldots & \ldots & 0 & c & a \end{bmatrix}, \quad B = \begin{bmatrix} \hat{a} & -b & 0 & \ldots & \ldots & \ldots & \ldots \\ -c & \hat{a} & -b & 0 & \ldots & \ldots & \ldots \\ 0 & -c & \hat{a} & -b & 0 & \ldots & \ldots \\ \ldots & \ldots & \ddots & \ddots & \ddots & \ldots & \ldots \\ \ldots & \ldots & 0 & -c & \hat{a} & -b & 0 \\ \ldots & \ldots & \ldots & 0 & -c & \hat{a} & -b \\ \ldots & \ldots & \ldots & \ldots & 0 & -c & \hat{a} \end{bmatrix}$$

and f a vector containing only zeros except for the last entry in which the non-zero Dirichlet boundary condition is processed. The contribution of this last entry is given by,

$$f_{last} = -b\left[S_{max} - Ke^{-r(T-(n+1)\Delta t)} + S_{max} - Ke^{-r(T-n\Delta t)}\right]$$

This can be seen by the fact that the last equation in the system of equations $Au^{n+1} = Bu^n + f$ is given by

$$cu^{n+1}_{N_S-2} + au^{n+1}_{N_S-1} + bu^{n+1}_{N_S} = -cu^n_{N_S-2} + \hat{a}u^n_{N_S-1} - bu^n_{N_S}$$

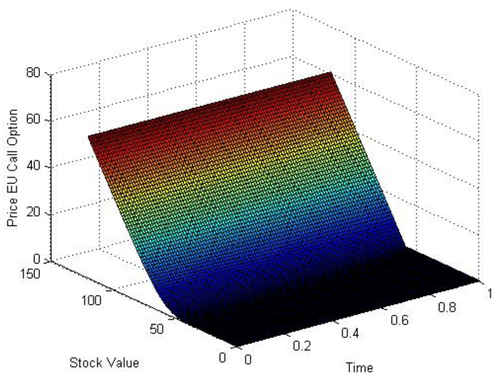
Hence, by applying the boundary condition at $i = N_S$ we find

$$cu^{n+1}_{N_S-2}+au^{n+1}_{N_S-1}+b\left[S_{max} - Ke^{-r(T-(n+1)\Delta t)}\right] = -cu^n_{N_S-2}+\hat{a}u^n_{N_S-1}-b\left[S_{max} - Ke^{-r(T-n\Delta t)}\right]$$

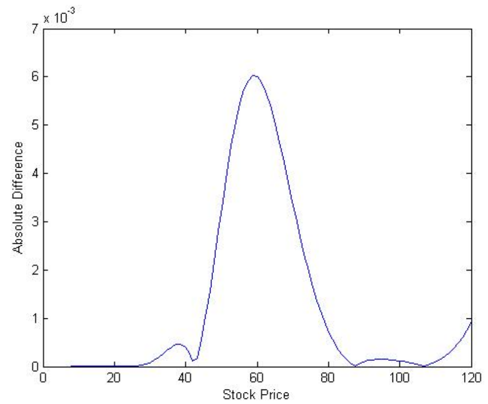
Rearranging terms leads to

$$cu^{n+1}_{N_S-2}+au^{n+1}_{N_S-1} = -cu^n_{N_S-2}+\hat{a}u^n_{N_S-1}-b\left[S_{max} - Ke^{-r(T-(n+1)\Delta t)} + S_{max} - Ke^{-r(T-n\Delta t)}\right]$$

The contribution in the right hand side due to the boundary condition is included in the vector f.
Notice that we have Dirichlet boundary conditions in this problem. This means that we already 'know' the solution at the boundaries. Our numerical scheme provides an approximation at the unknown gridpoints, i.e. without the boundary contributions.
In Matlab we can easily implement this scheme and solve $Au^{n+1} = Bu^n + f$. The parameters we use are, $T = 1$, $S_0 = 0$, $S_{max} = 120$, $K = 60$, $r = 0.01$, $\sigma = 0.2$, $N_S = 100$ and $N_t = 100$



(a) European Call option



(b) difference BS formula and CN method

Figure 3.1

In Figure 3.1 (a) we present a rather typical surface of a European Call option. Notice that the price of the option maintains its hockey stick shape as presented earlier in figure 2.2. Moreover, the surface plot looks similar to that of a heat equation. It is well known that the Black Scholes equation can be transformed into a heat equation, see [7]. Remember that we already presented a closed-form solution in section 2.2 for a European Call option. We used this formula to derive an error of our numerical approximation. Figure 3.1 (b) shows that around the discontinuity in the payoff function, that is around $S = 60$, there is a small error with a maximum magnitude of $6 \cdot 10^{-3}$. This oscillatory behavior of the Crank Nicolson method around a discontinuity makes it less attractive. However, the oscillations do not grow in time and can therefore be controlled. The absolute difference between the analytical solution and the numerical approximation has order of magnitude -3. This result confirms the second order accuracy of the Crank-Nicolson method.

The CN method can also be applied to the valuation of American options. Since American options give the owner an extra right in comparison to European-style options, we can't approach the problem as we did for European options. To this end, we rewrite the problem into a linear complementarity problem (LCP) as we did before in the obstacle problem, see section 2.3.1. This makes it possible to use finite difference methods in order to determine a value. We start our analysis by considering an American Put option. First we take the interest rate to be constant, which will simplify our analysis a bit. As discussed in section 2.3, an American Put option is potentially more worth than its European counterpart. By the no-arbitrage principle the following constraint holds

$$P(S(t), t) \geq max(K - S(t), 0) := (K - S)^+.$$

The stock value that separates our domain into two regions will be denoted by S_f where the subscript f refers to the free-boundary. It is usually called the optimal exercise price. Note that when a strict inequality occurs, i.e. $P(S(t), t) > (K - S)^+$, one should hold on to the option and the Black-Scholes equation holds. On the other hand, whenever $P(S(t), t) = (K - S)^+$ it is optimal to exercise. In summary, we face a problem which takes place on three parts:

1. If $S_f(t) < S(t) < \infty$ the owner holds the option and we have $P(S(t), t) > (K - S)^+$ along with the Black-Scholes equation:

$$\frac{\sigma^2 S^2}{2} \frac{\partial^2 P(S(t), t)}{\partial S^2} + rS \frac{\partial P(S(t), t)}{\partial S} - rP(S(t), t) + \frac{\partial P(S(t), t)}{\partial t} = 0$$

2. if $0 \leq S(t) \leq S_f$ the option is exercised and we have $P(S(t), t) = (K - S)^+$ along with the strict inequality,

$$\frac{\sigma^2 S^2}{2} \frac{\partial^2 P(S(t), t)}{\partial S^2} + rS \frac{\partial P(S(t), t)}{\partial S} - rP(S(t), t) + \frac{\partial P(S(t), t)}{\partial t} < 0$$

3. If $S_f(t) = S(t)$, i.e. the stock value is on the free-boundary, then

$$P(S_f(t), t) = K - S_f(t) \quad \text{and} \quad \frac{\partial P}{\partial S}(S_f(t), t) = -1.$$

along with continuity of P and its slope. Under 3. we stated that the first derivative with respect to S equals minus one. This requirement is also known as the smooth-pasting condition.

The above problem can be formulated as Linear Complementarity Problem (LCP):

$$-\left(\frac{\partial P}{\partial t}+\frac{\sigma^2 S^2}{2}\frac{\partial^2 P}{\partial S^2}+rS\frac{\partial P}{\partial S}-rP\right)\geq 0$$

$$P-max(K-S,0)\geq 0$$

$$\left(\frac{\partial P}{\partial t}+\frac{\sigma^2 S^2}{2}\frac{\partial^2 P}{\partial S^2}+rS\frac{\partial P}{\partial S}-rP\right)(P-max(K-S,0))=0$$

with final condition,

$$P(S(T),T)=max(K-S(T),0)$$

Our approach is similar to that of the obstacle problem. However it has some subtleties worth mentioning. Here, we use the θ-method. If $\theta=0.5$ we obtain the CN method.

$$-\frac{P_i^{n+1}-P_i^n}{\Delta t}-\theta\left(\frac{\sigma^2(S_0+i\Delta S)^2}{2}\frac{P_{i+1}^n-2P_i^n+P_{i-1}^n}{\Delta S^2}+r(S_0+i\Delta S)\frac{P_{i+1}^n-P_{i-1}^n}{2\Delta S}-rP_i^n\right)$$

$$-(1-\theta)\left(\frac{\sigma^2(S_0+i\Delta S)^2}{2}\frac{P_{i+1}^{n+1}-2P_i^{n+1}+P_{i-1}^{n+1}}{\Delta S^2}+r(S_0+i\Delta S)\frac{P_{i+1}^{n+1}-P_{i-1}^{n+1}}{2\Delta S}-rP_i^{n+1}\right)=0$$

Multiplying both sides by $-\Delta t$ and rearranging terms yields,

$$\begin{aligned}
P_i^{n+1}-P_i^n+\theta&\left(\frac{\sigma^2(S_0+i\Delta S)^2\Delta t}{2\Delta S^2}\left(P_{i+1}^n-2P_i^n+P_{i-1}^n\right)\right.\\
&\left.+\Delta t\frac{r(S_0+i\Delta S)}{2\Delta S}\left(P_{i+1}^n-P_{i-1}^n\right)-r\Delta tP_i^n\right)\\
+(1-\theta)&\left(\frac{\sigma^2(S_0+i\Delta S)^2\Delta t}{2\Delta S^2}\left(P_{i+1}^{n+1}-2P_i^{n+1}+P_{i-1}^{n+1}\right)\right.\\
&\left.+\Delta t\frac{r(S_0+i\Delta S)}{2\Delta S}\left(P_{i+1}^{n+1}-P_{i-1}^{n+1}\right)-r\Delta tP_i^{n+1}\right)=0
\end{aligned}$$

We abbreviate this equation by using the notation:

$$v_i=\frac{\sigma^2(S_0+i\Delta S)^2\Delta t}{2\Delta S^2}\qquad\text{and}\qquad w_i=\frac{\Delta tr(S_0+i\Delta S)}{2\Delta S}$$

Collecting similar terms and taking terms of the same time-instant on the same side:

$$\begin{aligned}
&P_i^n-\theta\left((v_i-w_i)P_{i-1}^n+(-2v_i-r\Delta t)P_i^n+(v_i+w_i)P_{i+1}^n\right)\\
&=P_i^{n+1}+(1-\theta)\left((v_i-w_i)P_{i-1}^{n+1}+(-2v_i-r\Delta t)P_i^{n+1}+(v_i+w_i)P_{i+1}^{n+1}\right)
\end{aligned}$$

We define the following terms,

$$\hat{a}_i=v_i-w_i,\qquad \hat{b}_i=-2v_i-r\Delta t\qquad\text{and}\qquad \hat{c}_i=v_i+w_i.$$

We can write our problem in a more compact form,

$$(I-\theta A)\mathbf{P}^n=(I+(1-\theta)A)\mathbf{P}^{n+1}+\mathbf{f}$$

where I is an identity matrix and A a tri-diagonal matrix:

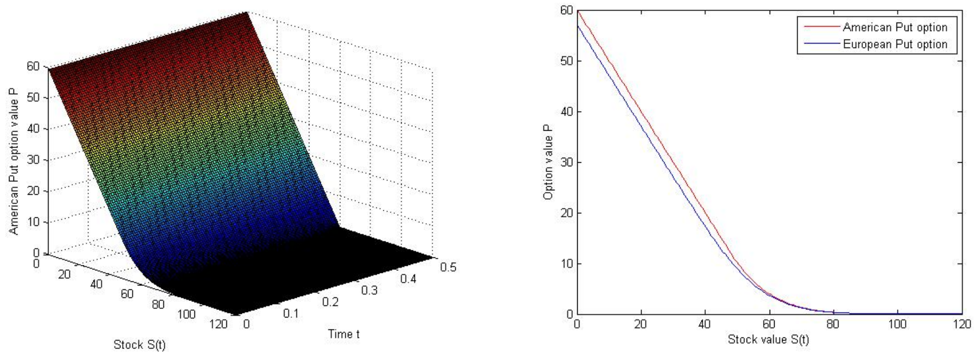
$$A=\begin{bmatrix}
\hat{b}_1 & \hat{c}_1 & 0 & \dots & \dots & \dots & \dots\\
\hat{a}_2 & \hat{b}_2 & \hat{c}_2 & 0 & \dots & \dots & \dots\\
0 & \hat{a}_i & \hat{b}_i & \hat{c}_i & 0 & \dots & \dots\\
\dots & \dots & \ddots & \ddots & \ddots & \dots & \dots\\
\dots & \dots & 0 & \hat{a}_{N_x-3} & \hat{b}_{N_x-3,j} & \hat{c}_{N_x-3,j} & 0\\
\dots & \dots & \dots & 0 & \hat{a}_{N_x-2} & \hat{b}_{N_x-2,j} & \hat{c}_{N_x-2,j}\\
\dots & \dots & \dots & \dots & 0 & \hat{a}_{N_x-1} & \hat{b}_{N_x-1,j}
\end{bmatrix}$$

and the $(N_x - 1) \times 1$ vector $\mathbf{f}$ contains the boundary contributions. For the boundary conditions we assume $S_0 = 0$ and $S_{max} > K$. Hence, the boundary conditions satisfy $P(S_0, t) = K$ and $P(S_{max}, t) = 0$. This leads to the following contributions,

$$f_1 = \theta \cdot \hat{a}_1 \cdot K + (1 - \theta) \cdot \hat{a}_1 \cdot K = \hat{a}_1 \cdot K$$

$$f_{N_x} = \theta \cdot \hat{c}_{N_x} \cdot 0 + (1 - \theta) \cdot \hat{c}_{N_x} \cdot 0 = 0$$

The other entries of the vector $\mathbf{f}$ are equal to zero. Using the PSOR algorithm we find the following results.



(a) American Put option (b) Put option values at $t = 0$

Figure 3.2

The parameters used in order to obtain the figures are: $S_0 = 0$, $S_{max} = 120$, $K = 60$, $r = 0.1$, $\sigma = 0.3$, $N_S = 200$, $N_\tau = 100$ and $T = 0.5$. From figure 3.2 (b) we see that the American Put option has a value that is at least as high as that of its European counterpart.

3.4 Crank-Nicolson Method for Multiple Dimensions

We successfully applied the CN-method to a Black-Scholes equation in only one space dimension. Here we will generalize this procedure to incorporate a stochastic interest rate. The generalization towards two space dimensions is mathematically easier to make if we consider the θ-method. As the name suggests we have a parameter θ which gives us the freedom to switch between fully explicit finite difference schemes and fully implicit finite difference schemes. Taking $\theta = 0.5$ we obtain the CN-method.
We already introduced a time-dependent parameter called theta in the Hull-White model. Therefore we choose to write, $\hat{\theta}$, for the time-independent parameter of the method. Now, we start off in much the same manner as before. We take a weighted average of an FTCS and BTCS scheme.

$$\frac{u_{i,j}^{n+1} - u_{i,j}^{n}}{\Delta\tau} = \hat{\theta}\left[\frac{\sigma_1^2}{2}(S_0 + i\Delta S)^2 \frac{u_{i+1,j}^{n+1} - 2u_{i,j}^{n+1} + u_{i-1,j}^{n+1}}{\Delta S^2}\right.$$

$$+\rho\sigma_1\sigma_2(S_0 + i\Delta S)\frac{u_{i+1,j+1}^{n+1} - u_{i+1,j-1}^{n+1} + u_{i-1,j-1}^{n+1} - u_{i-1,j+1}^{n+1}}{4\Delta S\Delta r}$$

$$+\frac{\sigma_2^2}{2}\frac{u_{i,j-1}^{n+1} - 2u_{i,j}^{n+1} + u_{i,j+1}^{n+1}}{\Delta r^2} + (\theta^n - \alpha(r_0 + j\Delta r))\frac{u_{i,j+1}^{n+1} - u_{i,j-1}^{n+1}}{2\Delta r}$$

$$\left.-(r_0 + j\Delta r)\left(u_{i,j}^{n+1} - (S_0 + i\Delta S)\frac{u_{i+1,j}^{n+1} - u_{i-1,j}^{n+1}}{2\Delta S}\right)\right]$$

$$+(1-\hat{\theta})\left[\frac{\sigma_1^2}{2}(S_0+i\Delta S)^2\frac{u^n_{i+1,j}-2u^n_{i,j}+u^n_{i-1,j}}{\Delta S^2}\right.$$
$$+\rho\sigma_1\sigma_2(S_0+i\Delta S)\frac{u^n_{i+1,j+1}-u^n_{i+1,j-1}+u^n_{i-1,j-1}-u^n_{i-1,j+1}}{4\Delta S\Delta r}$$
$$+\frac{\sigma_2^2}{2}\frac{u^n_{i,j-1}-2u^n_{i,j}+u^n_{i,j+1}}{\Delta r^2}+[\theta^n-\alpha(r_0+j\Delta r)]\frac{u^n_{i,j+1}-u^n_{i,j-1}}{2\Delta r}$$
$$\left.-(r_0+j\Delta r)\left(u^n_{i,j}-(S_0+i\Delta S)\frac{u^n_{i+1,j}-u^n_{i-1,j}}{2\Delta S}\right)\right]$$

Collecting terms regarding the new time instant on the left hand side and those of the old time instant on the right hand side yields,

$$u^{n+1}_{i,j}-\hat{\theta}\Delta\tau\left[u^{n+1}_{i-1,j}\left(\frac{\sigma_1^2(S_0+i\Delta S)^2}{2\Delta S^2}-\frac{(r_0+j\Delta r)(S_0+i\Delta S)}{2\Delta S}\right)\right.$$
$$+u^{n+1}_{i,j}\left(-\frac{\sigma_1^2(S_0+i\Delta S)^2}{\Delta S^2}-\frac{\sigma_2^2}{\Delta r^2}-(r_0+j\Delta r)\right)+u^{n+1}_{i+1,j}\left(\frac{\sigma_1^2(S_0+i\Delta S)^2}{2\Delta S^2}+\frac{(r_0+j\Delta r)(S_0+i\Delta S)}{2\Delta S}\right)$$
$$u^{n+1}_{i-1,j-1}\left(\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}\right)+u^{n+1}_{i,j-1}\left(\frac{\sigma_2^2}{2\Delta r^2}-\frac{(\theta-\alpha(r_0+j\Delta r))}{2\Delta r}\right)+u^{n+1}_{i+1,j-1}\left(-\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}\right)$$
$$\left.+u^{n+1}_{i-1,j+1}\left(-\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}\right)+u^{n+1}_{i,j+1}\left(\frac{\sigma_2^2}{2\Delta r^2}+\frac{(\theta-\alpha(r_0+j\Delta r))}{2\Delta r}\right)+u^{n+1}_{i+1,j+1}\left(\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}\right)\right]$$
$$=u^n_{i,j}+(1-\hat{\theta})\Delta\tau\left[u^n_{i-1,j}\left(\frac{\sigma_1^2(S_0+i\Delta S)^2}{2\Delta S^2}-\frac{(r_0+j\Delta r)(S_0+i\Delta S)}{2\Delta S}\right)\right.$$
$$+u^n_{i,j}\left(-\frac{\sigma_1^2(S_0+i\Delta S)^2}{\Delta S^2}-\frac{\sigma_2^2}{\Delta r^2}-(r_0+j\Delta r)\right)+u^n_{i+1,j}\left(\frac{\sigma_1^2(S_0+i\Delta S)^2}{2\Delta S^2}+\frac{(r_0+j\Delta r)(S_0+i\Delta S)}{2\Delta S}\right)$$
$$u^n_{i-1,j-1}\left(\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}\right)+u^n_{i,j-1}\left(\frac{\sigma_2^2}{2\Delta r^2}-\frac{(\theta-\alpha(r_0+j\Delta r))}{2\Delta r}\right)+u^n_{i+1,j-1}\left(-\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}\right)$$
$$\left.+u^n_{i-1,j+1}\left(-\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}\right)+u^n_{i,j+1}\left(\frac{\sigma_2^2}{2\Delta r^2}+\frac{(\theta-\alpha(r_0+j\Delta r))}{2\Delta r}\right)+u^n_{i+1,j+1}\left(\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}\right)\right]$$

This equation can be written as,

$$u^{n+1}_{i,j}-\hat{\theta}\Delta\tau\left[a_{i,j}u^{n+1}_{i-1,j}+b_{i,j}u^{n+1}_{i,j}+c_{i,j}u^{n+1}_{i+1,j}+\hat{a}_iu^{n+1}_{i-1,j-1}+\hat{b}_ju^{n+1}_{i,j-1}+\hat{c}_iu^{n+1}_{i+1,j-1}+\hat{\hat{a}}_iu^{n+1}_{i-1,j+1}+\hat{\hat{b}}_ju^{n+1}_{i,j+1}\right.$$
$$\left.+\hat{\hat{c}}_iu^{n+1}_{i+1,j+1}\right]=u^n_{i,j}+(1-\hat{\theta})\Delta\tau\left[a_{i,j}u^n_{i-1,j}+b_{i,j}u^n_{i,j}+c_{i,j}u^n_{i+1,j}+\hat{a}_iu^n_{i-1,j-1}+\hat{b}_ju^n_{i,j-1}+\hat{c}_iu^n_{i+1,j-1}\right.$$
$$\left.+\hat{\hat{a}}_iu^n_{i-1,j+1}+\hat{\hat{b}}_ju^n_{i,j+1}+\hat{\hat{c}}_iu^{n+1}_{i+1,j+1}\right]$$

where,

$$a_{i,j}=\frac{\sigma_1^2(S_0+i\Delta S)^2}{2\Delta S^2}-\frac{(r_0+j\Delta r)(S_0+i\Delta S)}{2\Delta S}\qquad b_{i,j}=-\frac{\sigma_1^2(S_0+i\Delta S)^2}{\Delta S^2}-\frac{\sigma_2^2}{\Delta r^2}-(r_0+j\Delta r)$$
$$c_{i,j}=\frac{\sigma_1^2(S_0+i\Delta S)^2}{2\Delta S^2}+\frac{(r_0+j\Delta r)(S_0+i\Delta S)}{2\Delta S}\qquad \hat{a}_i=\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}$$
$$\hat{b}_j=\frac{\sigma_2^2}{2\Delta r^2}-\frac{(\theta-\alpha(r_0+j\Delta r))}{2\Delta r}\qquad \hat{c}_i=-\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}$$
$$\hat{\hat{a}}_i=-\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}\qquad \hat{\hat{b}}_j=\frac{\sigma_2^2}{2\Delta r^2}+\frac{(\theta-\alpha(r_0+j\Delta r))}{2\Delta r}$$
$$\hat{\hat{c}}_i=\frac{\rho\sigma_1\sigma_2(S_0+i\Delta S)}{4\Delta S\Delta r}$$

Using the last equation, we construct matrices A_j,B_j and C_j as follows

$$A_j = \begin{bmatrix} b_{0,j}+2a_{0,j} & c_{0,j}-a_{0,j} & 0 & \dots & \dots & \dots & \dots \\ a_{1,j} & b_{1,j} & c_{1,j} & 0 & \dots & \dots & \dots \\ 0 & a_{2,j} & b_{2,j} & c_{2,j} & 0 & \dots & \dots \\ \dots & \dots & \ddots & \ddots & \ddots & \dots & \dots \\ \dots & \dots & 0 & a_{N_x-2,j} & b_{N_x-2,j} & c_{N_x-2,j} & 0 \\ \dots & \dots & \dots & 0 & a_{N_x-1,j} & b_{N_x-1,j} & c_{N_x-1,j} \\ \dots & \dots & \dots & \dots & 0 & a_{N_x,j}-c_{N_x,j} & b_{N_x,j}+2c_{N_x,j} \end{bmatrix}$$

The matrix B_j is obtained in a similar way by replacing all $a_{:,j}, b_{:,j}$ and $c_{:,j}$ by $\hat{a}_{:,j}, \hat{b}_{:,j}$ and $\hat{c}_{:,j}$ respectively. Similarly for C_j we replace the hats by double hats. Note that we already incorporated the boundary conditions in the matrix A_j. This is done by using ghost points at $(-1, j)$ and $(S_{max}+1, j)$. For completeness we will discuss this in more detail. By using a central difference scheme for the linear boundary condition at $(0, j)$, we find

$$\frac{u_{-1,j} - 2u_{0,j} + u_{1,j}}{\Delta S^2} = 0 \implies u_{-1,j} = 2u_{0,j} - u_{1,j}$$

Consequently, we derive for the first row of A_j:

$$a_{0,j}u_{-1,j} + b_{0,j}u_{0,j} + c_{0,j}u_{1,j} \implies a_{0,j}\left(2u_{0,j} - u_{1,j}\right) + b_{0,j}u_{0,j} + c_{0,j}u_{1,j}$$

We can write this last expression in terms of only $u_{0,j}$ and $u_{1,j}$ to obtain the contribution as presented in the first row of A_j

$$(b_{0,j} + 2a_{0,j})u_{0,j} + (c_{0,j} - a_{0,j})u_{1,j}$$

The same procedure can be done by using a ghost point at $(S_{max}+1, j)$ to derive the last row of A_j.

The next step consists of constructing a big tridiagonal block matrix such that the matrices B_j, A_j and C_j form the subdiagonal, diagonal and superdiagonal. To this end, let

$$\mathbf{X} = \begin{bmatrix} A_0+2B_0 & C_0-B_0 & 0 & \dots & \dots & \dots & \dots \\ B_1 & A_1 & C_1 & 0 & \dots & \dots & \dots \\ 0 & B_2 & A_2 & C_2 & 0 & \dots & \dots \\ \dots & \dots & \ddots & \ddots & \ddots & \dots & \dots \\ \dots & \dots & 0 & B_{N_y-2} & A_{N_y-2} & C_{N_y-2} & 0 \\ \dots & \dots & \dots & 0 & B_{N_y-1} & A_{N_y-1} & C_{N_y-1} \\ \dots & \dots & \dots & \dots & 0 & B_{N_y}-C_{N_y} & A_{N_y}+2C_{N_y} \end{bmatrix}$$

and let $\mathbf{I}$ denote a $(N_x + 1) \times (N_y + 1)$ identity matrix. Then we have the following functional relation

$$I\mathbf{u}^{n+1} - \hat{\theta}\Delta\tau\mathbf{X}\mathbf{u}^{n+1} = I\mathbf{u}^n + (1-\hat{\theta})\Delta\tau\mathbf{X}\mathbf{u}^n \tag{3.21}$$

Hence,

$$\mathbf{u}^{n+1} = \left(I - \hat{\theta}\Delta\tau\mathbf{X}\right)^{-1}\left(I + (1-\hat{\theta})\Delta\tau\mathbf{X}\right)\mathbf{u}^n \tag{3.22}$$

Since we know the vector $\mathbf{u}^{N_\tau}$ at maturity we can go from time instant n to $n+1$ which is backward in time to price a certain option at every time instant. A big disadvantage of this method are the big matrices of dimension $(N_x+1)\cdot(N_y+1)\times(N_x+1)\cdot(N_y+1)$. This leads to a lot of memory that is used when implementing this method. Therefore we are limited in the magnitude of the stepsize. Also due to the large bandwith, it is inefficient to compute the inverse of the matrix as presented above. Besides, it could give rise to an ill-conditioned matrix and consequently to bad results.

3.5 Alternating Direction Implicit Method

Previously we derived a CN scheme for the Black-Scholes equation under stochastic interest rates. The approach we took wasn't very practical due to the many dimensions that were needed to take care of. Here we will discuss an alternative, in which we handle every dimension separately. This method is called the Alternating Direction Implicit method (ADI). The method is especially attractive for solving higher dimensional problems without mixed derivatives. The method is based on reducing the problem to problems that are easier to handle, namely one-dimensional problems. The main idea behind the ADI method is to alternate between the various dimensions present in the problem such that simpler one-dimensional problems arise which can then be solved separately. There are many different ADI schemes available. In this section we will treat only one of those. The derivation is inspired by [16].
In order to introduce the method we consider a general problem in the form of the equation:

$$\frac{\partial V}{\partial \tau} = \mathcal{L}V$$

where $\mathcal{L}$ is an operator that can be written as a sum of three other operators, i.e. $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3$. Each of these operators should treat one space dimension only. Since the Black-Scholes equation under stochastic interest contains a mixed derivative term we reserve the operator $\mathcal{L}_1$ for this term. The term $-rV$ is spread evenly over the operators $\mathcal{L}_2$ and $\mathcal{L}_3$. By using central difference schemes we end up with

$$\mathcal{L}_1 u_{i,j}^n = \rho\sigma_1\sigma_2(S_0 + i\Delta S)\frac{u_{i+1,j+1}^n - u_{i+1,j-1}^n - u_{i-1,j+1}^n + u_{i-1,j-1}^n}{4\Delta S\Delta r}$$

$$\mathcal{L}_2 u_{i,j}^n = \frac{\sigma_1^2(S_0 + i\Delta S)^2}{2}\frac{u_{i+1,j}^n - 2u_{i,j}^n + u_{i-1,j}^n}{\Delta S^2} + (r_0 + j\Delta r)(S_0 + i\Delta S)\frac{u_{i+1,j}^n - u_{i-1,j}^n}{2\Delta S} - \frac{(r_0 + j\Delta r)}{2}u_{i,j}^n$$

$$\mathcal{L}_3 u_{i,j}^n = \frac{\sigma_2^2}{2}\frac{u_{i,j+1}^n - 2u_{i,j}^n + u_{i,j-1}^n}{\Delta r^2} + [\theta - \alpha(r_0 + j\Delta r)]\frac{u_{i,j+1}^n - u_{i,j-1}^n}{2\Delta r} - \frac{(r_0 + j\Delta r)}{2}u_{i,j}^n$$

Applying the Crank-Nicolson method yields,

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} = \frac{1}{2}\left(\mathcal{L}_2 u_{i,j}^{n+1} + \mathcal{L}_2 u_{i,j}^n\right) + \frac{1}{2}\left(\mathcal{L}_3 u_{i,j}^{n+1} + \mathcal{L}_3 u_{i,j}^n\right) + \frac{1}{2}\left(\mathcal{L}_1 u_{i,j}^{n+1} + \mathcal{L}_1 u_{i,j}^n\right)$$
$$+\mathcal{O}(\Delta r^2) + \mathcal{O}(\Delta S^2) + \mathcal{O}(\Delta t^2)$$

Rearranging terms leads to the expression,

$$\left(I - \frac{\Delta t}{2}\mathcal{L}_2 - \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} = \left(I + \frac{\Delta t}{2}\mathcal{L}_2 + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^n + \frac{\Delta t}{2}\left(\mathcal{L}_1 u_{i,j}^{n+1} + \mathcal{L}_1 u_{i,j}^n\right)$$
$$+\mathcal{O}(\Delta t\Delta r^2) + \mathcal{O}(\Delta t\Delta S^2) + \mathcal{O}(\Delta t^3)$$

The next step is a crucial part of the method. We add and subtract the terms $\frac{\Delta t^2}{4}\mathcal{L}_2\mathcal{L}_3 u_{i,j}^{n+1}$ and $\frac{\Delta t^2}{4}\mathcal{L}_2\mathcal{L}_3 u_{i,j}^n$ on both sides. Doing this enables us to factorize a part of the equation.

$$\left(I - \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I - \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} = \left(I + \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^n$$

$$+\frac{\Delta t^2}{4}\mathcal{L}_2\mathcal{L}_3\left(u_{i,j}^{n+1} - u_{i,j}^n\right) + \frac{\Delta t}{2}\left(\mathcal{L}_1 u_{i,j}^{n+1} + \mathcal{L}_1 u_{i,j}^n\right) + \mathcal{O}\left(\Delta t\Delta r^2\right) + \mathcal{O}(\Delta t\Delta S^2) + \mathcal{O}(\Delta t^3)$$

We simplify the expression by noting that,

$$u_{i,j}^{n+1} = u_{i,j}^n + \mathcal{O}(\Delta t) \implies \frac{\Delta t^2}{4}\mathcal{L}_2\mathcal{L}_3\left(u_{i,j}^{n+1} - u_{i,j}^n\right) = \mathcal{O}(\Delta t^3)$$

Hence,

$$\left(I - \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I - \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} = \left(I + \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n}$$

$$+\frac{\Delta t}{2}\left(\mathcal{L}_1 u_{i,j}^{n+1} + \mathcal{L}_1 u_{i,j}^{n}\right) + \mathcal{O}\left(\Delta t \Delta r^2\right) + \mathcal{O}(\Delta t \Delta S^2) + \mathcal{O}(\Delta t^3)$$

Following [16], we would end up with the so-called Douglas, Peaceman Rachford scheme which aims to solve the last equation for u by the two step procedure:

$$\left(I - \frac{\Delta t}{2}\mathcal{L}_2\right)u_{i,j}^{n+\frac{1}{2}} = \left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n} + \frac{\Delta t}{4}\left(\mathcal{L}_1 u_{i,j}^{n+1} + \mathcal{L}_1 u_{i,j}^{n}\right)$$

$$\left(I - \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} = \left(I + \frac{\Delta t}{2}\mathcal{L}_2\right)u_{i,j}^{n+\frac{1}{2}} + \frac{\Delta t}{4}\left(\mathcal{L}_1 u_{i,j}^{n+1} + \mathcal{L}_1 u_{i,j}^{n}\right)$$

Here, the unknown $u_{i,j}^{n+\frac{1}{2}}$ in the first step must be read as an auxiliary variable. In contrast to [16] where the second term in the right hand side of both steps denotes a source term, we have a mixed derivative term that depends on the unknown solution. This makes it impossible to apply the scheme in this manner. A possible adjustment is to treat the mixed derivative term either explicitly or implicitly instead of applying the Crank-Nicolson scheme to all three terms in our splitting. We choose to treat the mixed derivative term explicitly, i.e. the first step after introducing the operators at the beginning of this section reads

$$\frac{u_{i,j}^{n+1} - u_{i,j}^{n}}{\Delta t} = \frac{1}{2}\left(\mathcal{L}_2 u_{i,j}^{n+1} + \mathcal{L}_2 u_{i,j}^{n}\right) + \frac{1}{2}\left(\mathcal{L}_3 u_{i,j}^{n+1} + \mathcal{L}_3 u_{i,j}^{n}\right) + \mathcal{L}_1 u_{i,j}^{n}$$

A similar derivation leads to

$$\left(I - \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I - \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} = \left(I + \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n}$$

$$+\Delta t \mathcal{L}_1 u_{i,j}^{n} + \mathcal{O}\left(\Delta t \Delta r^2\right) + \mathcal{O}(\Delta t \Delta S^2) + \mathcal{O}(\Delta t^2)$$

Consequently, we have global error of order two in both space dimensions as well as the time dimension. Now, the Douglas Peaceman Rachford scheme is given by the following two steps.

Step 1: $$\left(I - \tfrac{\Delta t}{2}\mathcal{L}_2\right)u_{i,j}^{n+\frac{1}{2}} = \left(I + \tfrac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n} + \tfrac{\Delta t}{2}\mathcal{L}_1 u_{i,j}^{n}$$

Step 2: $$\left(I - \tfrac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} = \left(I + \tfrac{\Delta t}{2}\mathcal{L}_2\right)u_{i,j}^{n+\frac{1}{2}} + \tfrac{\Delta t}{2}\mathcal{L}_1 u_{i,j}^{n}$$

In order to see that this scheme indeed solves our equation we apply the operator $(I - \frac{\Delta t}{2}\mathcal{L}_2)$ to the left hand side of step 2. By using step 1 it follows that

$$\left(I - \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I - \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} = \left(I - \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I + \frac{\Delta t}{2}\mathcal{L}_2\right)u_{i,j}^{n+\frac{1}{2}} + \frac{\Delta t}{2}\left(I - \frac{\Delta t}{2}\mathcal{L}_2\right)\mathcal{L}_1 u_{i,j}^{n}$$

$$= \left(I + \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n} + \frac{\Delta t}{2}\left(I + \frac{\Delta t}{2}\mathcal{L}_2\right)\mathcal{L}_1 u_{i,j}^{n} + \frac{\Delta t}{2}\left(I - \frac{\Delta t}{2}\mathcal{L}_2\right)\mathcal{L}_1 u_{i,j}^{n}$$

$$= \left(I + \frac{\Delta t}{2}\mathcal{L}_2\right)\left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n} + \Delta t \mathcal{L}_1 u_{i,j}^{n}$$

This shows that the two step scheme indeed solves the problem. From the two step scheme it can be seen that the first step treats the r-dimension explicitly and the S-dimension implicitly (apart from the mixed derivative). In the second step it is the

other way around, the r-dimension is treated implicitly while the S-dimension is treated explicitly. This structure justifies the name Alternating Direction Implicit method or abbreviated ADI.
Since we need an extra step in order to go from time instant n to time instant $n+1$, it looks harder to implement at first sight. However, each step handles only one dimension, which makes it much easier to implement. To show this we write out step 1 and rearrange terms to obtain.

$$u_{i-1,j}^{n+\frac{1}{2}}\left(-\frac{\Delta t\sigma_1^2(S_0+i\Delta S)^2}{4\Delta S^2}+\frac{\Delta t(r_0+j\Delta r)(S_0+i\Delta S)}{4\Delta S}\right)+$$

$$u_{i,j}^{n+\frac{1}{2}}\left(1+\frac{\Delta t\sigma_1^2(S_0+i\Delta S)^2}{2\Delta S^2}+\frac{\Delta t(r_0+j\Delta r)}{4}\right)+$$

$$u_{i+1,j}^{n+\frac{1}{2}}\left(-\frac{\Delta t\sigma_1^2(S_0+i\Delta S)^2}{4\Delta S^2}-\frac{\Delta t(r_0+j\Delta r)(S_0+i\Delta S)}{4\Delta S}\right)=u_{i-1,j-1}^{n}\left(\frac{\Delta t\rho\sigma_1\sigma_2(S_0+i\Delta S)}{8\Delta S\Delta r}\right)+$$

$$u_{i,j-1}^{n}\left(\frac{\Delta t\sigma_2^2}{4\Delta r^2}-\frac{\Delta t}{4\Delta r}\left[\theta-\alpha(r_0+j\Delta r)\right]\right)+u_{i+1,j-1}^{n}\left(-\frac{\Delta t\rho\sigma_1\sigma_2(S_0+i\Delta S)}{8\Delta S\Delta r}\right)+$$

$$u_{i,j}^{n}\left(1-\frac{\Delta t\sigma_2^2}{2\Delta r^2}-\frac{\Delta t(r_0+j\Delta r)}{4}\right)+u_{i-1,j+1}^{n}\left(-\frac{\Delta t\rho\sigma_1\sigma_2(S_0+i\Delta S)}{8\Delta S\Delta r}\right)+$$

$$u_{i,j+1}^{n}\left(\frac{\Delta t\sigma_2^2}{4\Delta r^2}+\frac{\Delta t}{4\Delta r}\left[\theta-\alpha(r_0+j\Delta r)\right]\right)+u_{i+1,j+1}^{n}\left(\frac{\Delta t\rho\sigma_1\sigma_2(S_0+i\Delta S)}{8\Delta S\Delta r}\right)$$

Rewritten into a more compact form:

$$\alpha_{i,j}u_{i-1,j}^{n+\frac{1}{2}}+\beta_{i,j}u_{i,j}^{n+\frac{1}{2}}+\gamma_{i,j}u_{i+1,j}^{n+\frac{1}{2}}=$$

$$\hat{\alpha}_i u_{i-1,j-1}^{n}+\hat{\beta}_j u_{i,j-1}^{n}+\hat{\gamma}_i u_{i+1,j-1}^{n}+\eta_j u_{i,j}^{n}+\hat{\hat{\alpha}}_i u_{i-1,j+1}^{n}+\hat{\hat{\beta}}_j u_{i,j+1}^{n}+\hat{\hat{\gamma}}_i u_{i+1,j+1}^{n}$$

where,

$$\alpha_{i,j}=-\tfrac{\Delta t\sigma_1^2(S_0+i\Delta S)^2}{4\Delta S^2}+\tfrac{\Delta t(r_0+j\Delta r)(S_0+i\Delta S)}{4\Delta S} \qquad \beta_{i,j}=1+\tfrac{\Delta t\sigma_1^2(S_0+i\Delta S)^2}{2\Delta S^2}+\tfrac{\Delta t(r_0+j\Delta r)}{4}$$

$$\gamma_{i,j}=-\tfrac{\Delta t\sigma_1^2(S_0+i\Delta S)^2}{4\Delta S^2}-\tfrac{\Delta t(r_0+j\Delta r)(S_0+i\Delta S)}{4\Delta S} \qquad \hat{\alpha}_i=\tfrac{\Delta t\rho\sigma_1\sigma_2(S_0+i\Delta S)}{8\delta S\Delta r}$$

$$\hat{\beta}_j=\tfrac{\Delta t\sigma_2^2}{4\Delta r^2}-\tfrac{\Delta t}{4\Delta r}\left[\theta-\alpha(r_0+j\Delta r)\right] \qquad \hat{\gamma}_i=-\tfrac{\Delta t\rho\sigma_1\sigma_2(S_0+i\Delta S)}{8\Delta S\Delta r}$$

$$\eta_j=1-\tfrac{\Delta t\sigma_2^2}{2\Delta r^2}-\tfrac{\Delta t(r_0+j\Delta r)}{4} \qquad \hat{\hat{\alpha}}_i=-\tfrac{\Delta t\rho\sigma_1\sigma_2(S_0+i\Delta S)}{8\Delta S\Delta r}$$

$$\hat{\hat{\beta}}_j=\tfrac{\Delta t\sigma_2^2}{4\Delta r^2}+\tfrac{\Delta t}{4\Delta r}\left[\theta-\alpha(r_0+j\Delta r)\right] \qquad \hat{\hat{\gamma}}_i=\tfrac{\Delta t\rho\sigma_1\sigma_2(S_0+i\Delta S)}{8\Delta S\Delta r}$$

We can write the functional relation into the form

$$Au_{:,j}^{n+\frac{1}{2}}=Bu_{:,j-1}^{n}+Cu_{:,j}^{n}+Du_{:,j+1}^{n} \tag{3.23}$$

Fix the index j, then we are able to form the matrices A, B, C and D. Here A is a tridiagonal matrix containing α on its lower diagonal, β on its diagonal and γ on its upper diagonal. Similarly, B is a tridiagonal matrix with $\hat{\alpha},\hat{\beta},\hat{\gamma}$ on its lower, middle and upper diagonal. the same holds for the matrix D where the hats are replaced by double hats. On the other hand, C is a diagonal matrix with only η on its diagonal. Here A is a tridiagonal matrix. Hence, the equation is easy to solve for $u^{n+\frac{1}{2}}$ for each j by using the so-called Thomas algorithm or just the backslash command in Matlab. This equation

also motivates our choice to take the mixed derivative in an explicit manner. If we chose to treat it implicitly, we would not have ended up with a tridiagonal matrix on the left hand side and this would make it more difficult to solve for $u^{n+\frac{1}{2}}$.
The second step can be solved in similar fashion. We can write it as

$$\kappa_j u_{i,j-1}^{n+1} + \lambda_j u_{i,j}^{n+1} + \mu_j u_{i,j+1}^{n+1} = \hat{\kappa}_i u_{i-1,j-1}^{n} + \hat{\mu}_i u_{i-1,j+1}^{n} + \hat{\lambda}_i u_{i-1,j}^{n+\frac{1}{2}} + \hat{\hat{\lambda}}_i u_{i,j}^{n+\frac{1}{2}}$$
$$+\hat{\mu}_i u_{i+1,j-1}^{n} + \hat{\kappa}_i u_{i+1,j+1}^{n} + \hat{\eta}_{i,j} u_{i+1,j}^{n+\frac{1}{2}}$$

where,

$$\kappa_j = -\tfrac{\Delta t\sigma_2^2}{4\Delta r^2} + \tfrac{\Delta t}{4\Delta r}\left[\theta - \alpha(r_0 + j\Delta r)\right] \qquad \lambda_j = 1 + \tfrac{\Delta t\sigma_2^2}{2\Delta r^2} + \tfrac{\Delta t(r_0+j\Delta r)}{4}$$

$$\mu_j = -\tfrac{\Delta t\sigma_2^2}{4\Delta r^2} - \tfrac{\Delta t}{4\Delta r}\left[\theta - \alpha(r_0 + j\Delta r)\right] \qquad \hat{\kappa}_i = \tfrac{\rho\sigma_1\sigma_2(S_0+i\Delta S)\Delta t}{8\Delta S\Delta r}$$

$$\hat{\mu}_i = -\tfrac{\rho\sigma_1\sigma_2(S_0+i\Delta S)\Delta t}{8\Delta S\Delta r} \qquad \hat{\lambda}_i = \tfrac{\Delta t\sigma_1^2(S_0+i\Delta S)^2}{4\Delta S^2} - \tfrac{\Delta t(r_0+j\Delta r)(S_0+i\Delta S)}{4\Delta S}$$

$$\hat{\hat{\lambda}}_i = 1 - \tfrac{\Delta t\sigma_1^2(S_0+i\Delta S)^2}{2\Delta S^2} - \tfrac{\Delta t(r_0+j\Delta r)}{4} \qquad \hat{\mu}_i = -\tfrac{\rho\sigma_1\sigma_2(S_0+i\Delta S)\Delta t}{8\Delta S\Delta r}$$

$$\hat{\kappa}_i = \tfrac{\rho\sigma_1\sigma_2(S_0+i\Delta S)\Delta t}{8\Delta S\Delta r} \qquad \hat{\eta}_{i,j} = \tfrac{\Delta t\sigma_1^2(S_0+i\Delta S)^2}{4\Delta S^2} + \tfrac{\Delta t(r_0+j\Delta r)(S_0+i\Delta S)}{4\Delta S}$$

As before we process the terms into matrices to find the equation.

$$\hat{A}u_{i,:}^{n+1} = \hat{B}u_{i-1,:}^{n} + \hat{C}u_{i-1,:}^{n+\frac{1}{2}} + \hat{D}u_{i,:}^{n+\frac{1}{2}} + \hat{E}u_{i+1,:}^{n} + \hat{F}u_{i+1,:}^{n+\frac{1}{2}} \tag{3.24}$$

Here $\hat{A}$ has κ, λ and μ on its sub,middle and super diagonal. The matrix $\hat{B}$ has $\hat{\kappa}$ and $\hat{\mu}$ on its sub and super diagonal, the matrices $\hat{C}$ and $\hat{D}$ are diagonal matrices with $\hat{\lambda}$ and $\hat{\hat{\lambda}}$ on each of their diagonals respectively. The first two terms of the second line in the above mentioned equation, $\hat{\mu}$ and $\hat{\kappa}$, fill up the subdiagonal and superdiagonal of $\hat{E}$. finally, the matrix $\hat{F}$ is a diagonal matrix with $\hat{\eta}$ on its diagonal. This equation can also be solved using the Thomas algorithm, since $\hat{A}$ is a tridiagonal matrix. Applying the two steps yields a solution for the next time instant. Repeating this procedure yields the solution at time zero.
In solving both steps one has to be cautious when it comes down to the treatment of boundary conditions. We use linear boundary conditions. This may not cause any difficulties since these are incorporated in the same manner as for the CN-method. However, if boundary conditions depend explicitly on the time, the first step needs some more attention, see [17] [18]. To this end, one should rewrite the equation in step one.

$$\frac{\Delta t}{2}\mathcal{L}_2 u_{i,j}^{n+\frac{1}{2}} = -\left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n} - \frac{\Delta t}{2}\mathcal{L}_1 u_{i,j}^{n} + u_{i,j}^{n+\frac{1}{2}}$$

Substitution into the equation of step 2 yields

$$\left(I - \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} = 2u_{i,j}^{n+\frac{1}{2}} - \left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n}.$$

Rearranging terms gives

$$u_{i,j}^{n+\frac{1}{2}} = \frac{1}{2}\left(I - \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} + \frac{1}{2}\left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n}.$$

Suppose the time-dependent boundary condition is given by a function $g(t)$ and denote its discrete counterpart by g^n. The previous equation provides us with the right boundary condition for the auxiliary variable,

$$u_{i,j}^{n+\frac{1}{2}} = \frac{1}{2}\left(I - \frac{\Delta t}{2}\mathcal{L}_3\right)g^{n+1} + \frac{1}{2}\left(I + \frac{\Delta t}{2}\mathcal{L}_3\right)g^{n}.$$

3.6 Higher-Dimensional Linear Complementarity Problems

In section 3.3 we addressed the valuation of American options in one space dimension. This led to a linear complementarity problem (LCP) which can be formally written in matrix-vector notation as

$$\begin{aligned} A\mathbf{x} &\geq \mathbf{b} \\ \mathbf{x} &\geq \mathbf{c} \\ (\mathbf{x}-\mathbf{c})'(A\mathbf{x}-\mathbf{b}) &= 0 \end{aligned}$$

where $A \in \mathbb{R}^{n\times n}$ and $\mathbf{x}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$. In the case of one space dimension we were able to solve this problem by using the PSOR algorithm. However, for problems in which we face multiple dimensions we will follow a different approach. We choose to apply the CN Ikonen-Toivanen (CN-IT) method as well as an ADI Ikonen-Toivanen (ADI-IT) method, see [19] and [20]. Both methods are adjustments of the previous discussed CN and ADI methods. The reason we choose for the CN-IT and ADI-IT methods instead of the PSOR algorithm, is because the PSOR algorithm tends to be very slow for higher-dimensional problems. Luckily, we only need to apply some minor adjustments to both the CN and ADI method. By using equation (3.21) we observe that applying the CN method leads to the following LCP,

$$\begin{aligned} I\mathbf{u}^{n+1} - \frac{\Delta\tau}{2}\mathbf{X}\mathbf{u}^{n+1} &\geq I\mathbf{u}^n + \frac{\Delta\tau}{2}\mathbf{X}\mathbf{u}^n \\ \mathbf{u}^{n+1} &\geq \mathbf{u}^0 \\ (\mathbf{u}^{n+1}-\mathbf{u}^0)'\left(I\mathbf{u}^{n+1} - \frac{\Delta\tau}{2}\mathbf{X}\mathbf{u}^{n+1} - \left(I\mathbf{u}^n + \frac{\Delta\tau}{2}\mathbf{X}\mathbf{u}^n\right)\right) &= 0 \end{aligned}$$

The vector $\mathbf{u}^0$ is the solution at maturity, i.e. the payoff function. For an American Put option this leads to a vector containing $(K-(S_0+i\Delta S))^+$ for each $i \in \{0,\ldots,N_S\}$. In order to solve this LCP we follow the CN-IT splitting procedure by introducing an auxiliary vector λ^n. This vector is initially taken to be equal to the zero vector and is updated after each time step. A step in the CN-IT splitting method starts by solving

$$\mathbf{u}^* = \left(I - \frac{\Delta\tau}{2}\mathbf{X}\right)^{-1}\left(I + \frac{\Delta\tau}{2}\mathbf{X} + \Delta t\lambda^n\right)\mathbf{u}^n \tag{3.25}$$

This is an ordinary CN step along with an adjustment in order to cope with the inequality in the LCP problem. The star in $\mathbf{u}^*$ indicates that it is an auxiliary variable. Before we update λ^n we update the solution vector $\mathbf{u}^n$ according to

$$\mathbf{u}^{n+1} = max(\mathbf{u}^* - \Delta t\lambda^n, \mathbf{u}^0)$$

This yields the solution one step further in time. At last, one needs to update λ^n according to

$$\lambda^{n+1} = max\left(0, \lambda^n + \frac{1}{\Delta\tau}\left(\mathbf{u}^0 - \mathbf{u}^*\right)\right)$$

The maxima in the steps for updating $\mathbf{u}^n$ and λ^n are taken componentwise. Repeating this procedure by using after each step the new λ^{n+1} and $\mathbf{u}^{n+1}$ yields the solution. The procedure for the ADI-IT method follows in a similar way. We adjust both steps in the two-step scheme as follows,

Step 1: $$\left(I - \tfrac{\Delta t}{2}\mathcal{L}_2\right)u_{i,j}^{n+\frac{1}{2}} = \left(I + \tfrac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^n + \tfrac{\Delta t}{2}\mathcal{L}_1 u_{i,j}^n + \Delta t\lambda_{i,j}^n$$

Step 2: $$\left(I - \tfrac{\Delta t}{2}\mathcal{L}_3\right)u_{i,j}^{n+1} = \left(I + \tfrac{\Delta t}{2}\mathcal{L}_2\right)u_{i,j}^{n+\frac{1}{2}} + \tfrac{\Delta t}{2}\mathcal{L}_1 u_{i,j}^n + \Delta t\lambda_{i,j}^n$$

Note that we adjust the ADI method in exactly the same way as the CN method of equation (3.25). This adjustment leads to a modification of the functional relations given in (3.23) and (3.24). Hence the adjustment of equation (3.23) reads,

$$Au_{:,j}^{n+\frac{1}{2}} = Bu_{:,j-1}^{n} + Cu_{:,j}^{n} + Du_{:,j+1}^{n} + \Delta t\lambda_{:,j}^{n} \tag{3.26}$$

Similarly, the adjustment of equation (3.24) is given by

$$\hat{A}u_{i,:}^{*} = \hat{B}u_{i-1,:}^{n} + \hat{C}u_{i-1,:}^{n+\frac{1}{2}} + \hat{D}u_{i,:}^{n+\frac{1}{2}} + \hat{E}u_{i+1,:}^{n} + \hat{F}u_{i+1,:}^{n+\frac{1}{2}} + \Delta t\lambda_{i,:}^{n} \tag{3.27}$$

In order to apply the ADI-IT method one starts by initializing the auxiliary variable λ^n by setting it equal to the zero vector. Then, equations (3.26) and (3.27) are solved for $u_{:,j}^{n+\frac{1}{2}}$ and $u_{i,:}^{*}$ respectively. Now, we obtain the solution at the next time instant by taking the following maximum componentwise,

$$\mathbf{u}^{n+1} = max(\mathbf{u}^{*} - \Delta t\lambda^{n}, \mathbf{u}^{0})$$

where $\mathbf{u}^0$ is given by the payoff function. At last, we update the auxiliary variable,

$$\lambda^{n+1} = max\left(0, \lambda^{n} + \frac{1}{\Delta\tau}\left(\mathbf{u}^{0} - \mathbf{u}^{*}\right)\right)$$

By repeating this procedure with the updated λ^{n+1} and $\mathbf{u}^{n+1}$ we find the solution to the LCP.

Chapter 4

Results and Discussion

The aim of this chapter is to provide the results obtained by applying the numerical methods treated in Chapter 3. We start off by the pricing of zero-coupon bonds. From there, we move on to equity and bonds. Next, we will examine European options, Binary options and American options. Besides the valuation of options it is also important to determine the sensitivity to a change in the parameters. To this end, we compute the Δ as well as the ρ for both a European Call option as well as for an American Put option. For each of these topics we discuss the results and emphasize the differences between the results obtained by both numerical methods.

4.1 Zero Coupon Bonds

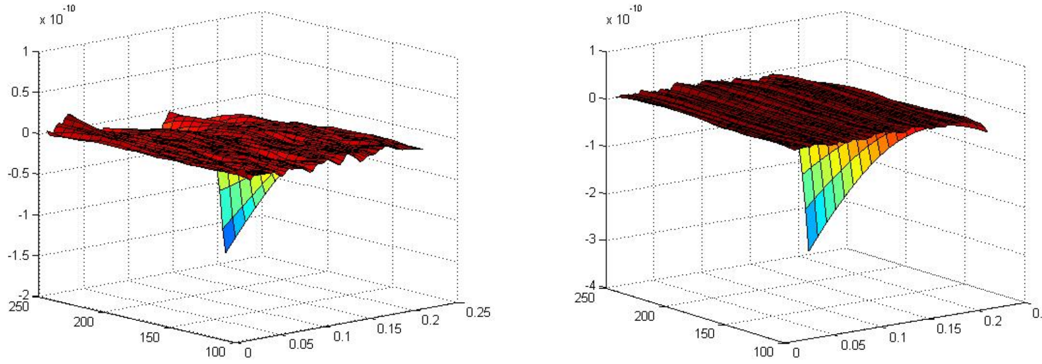
In section 2.4 we gave an analytic expression for the price of zero-coupon bonds under the HW model. Here we use this closed form solution along with the CN method and the ADI method in order to price zero-coupon bonds by specifying a payoff of a specific amount K at maturity. To this end we consider the following parameters, $\sigma_1 = 0.3$, $\sigma_2 = 0.01$, $S_0 = 100$, $S_{max} = 250$, $K = 100$, $r_{min} = 0.01$, $r_{max} = 0.21$, $\rho = 0.5$. We let the number of grid points be defined by $N_S = 20$, $N_r = 60$ and $N_\tau = 1000$. At last we take as Hull White model parameters, $\theta = 0.02$ and $\alpha = 0.01$. In the table below we present the analytic solution of a zero-coupon bond $P(0, T)$ regarding a payoff of $K = 100$ at maturity T for several values of r and T. In addition we show the results obtained by using the CN and ADI methods of the previous chapter with the same payoff at maturity. Note that we also value the bonds on the boundary of our grid, $r = 0.01$. This could lead to inaccurate results due to the linear boundary conditions.

	HW	CN	ADI
		T=1	
$r = 0.01$	98.0296	98.0294	98.0294
$r = 0.05$	94.2046	94.2046	94.2046
$r = 0.10$	89.6325	89.6325	89.6325
		T=5	
$r = 0.01$	74.6284	74.6222	74.6222
$r = 0.05$	61.4018	61.4015	61.4015
$r = 0.10$	48.1146	48.1139	48.1139

The results shown in the table above indicate that the approximations obtained by using the CN and ADI methods work. It is quite remarkable that both methods give such similar results. Even on the boundary we obtain very accurate results. It can be readily verified that the error between the analytical solution and one of both numerical methods has order of magnitude 3. This is to be expected since both numerical methods are second order accurate.

4.2 Equity and Bonds

We set the maturity to $T = 1$ and adjust the payoff to $S(T) = 100$ at maturity. We compare the initial payoff at maturity, $S(T)$, to the payoff at time zero, $S(0)$. The figure below shows the differences between these values on the whole grid regarding the CN-method and ADI method.



(a) CN-Method: $S(T = 1) - S(T = 0)$ (b) ADI-Method: $S(T = 1) - S(T = 0)$

Figure 4.1

It can be seen that the differences are very small by looking at the scale of both figures. This result is expected due to the fact that all derivatives with respect to r are equal to zero. Also, the second derivative of V with respect to S is equal to zero and V is approximately equal to $(dV/dS) \cdot S$. Thus, the partial differential equation simplifies to $(dV/dt) = 0$. Consequently, we only expect small errors due to discretizations. This is indeed the case as can be seen from figure 4.1. Besides, it is quite remarkable that the difference for both methods is large in the neighborhood of $r = 0.21$ and $S = 250$. This is exactly a corner in our grid.

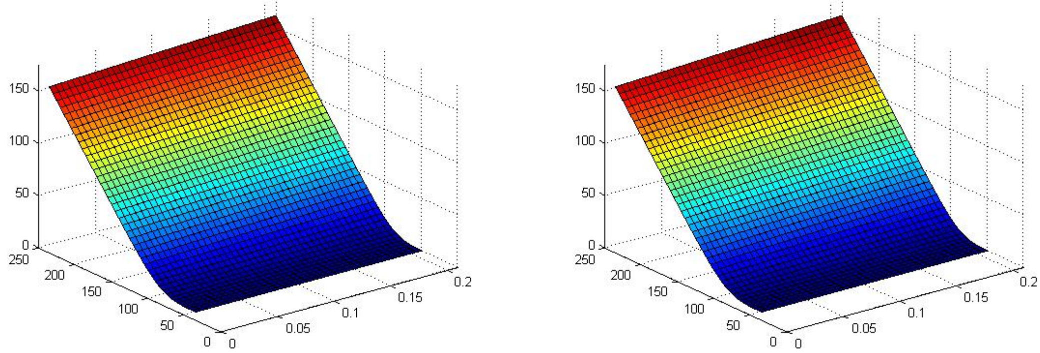
Now we adjust the payoff to $100 + S(T)$ at maturity. Hence, we combine the Zero-Coupon bond of the previous section with equity. This yields the following results at time $t = 0$ for $S = 100$.

	HW	CN	ADI
		T=1	
$r = 0.01$	198.0296	198.0294	198.0294
$r = 0.05$	194.2046	194.2046	194.2046
$r = 0.10$	189.6325	189.6325	189.6325
		T=5	
$r = 0.01$	174.6284	174.6222	174.6222
$r = 0.05$	161.4018	161.4015	161.4015
$r = 0.10$	148.1146	148.1139	148.1139

These results show that the accuracy is really good for maturity $T = 1$ and becomes slightly worse when the maturity is changed to $T = 5$. But taking the second order accuracy of these methods into account, along with the fact that approximations are done on the boundary of our grid, we conclude that the approximations are still acceptable. Moving towards more central laying gridpoints, we obtain better approximations. However, the results indicate that both methods perform equally well when taking only the accuracy of the methods into account. Note that the results presented in this table are also in line with those presented for Zero-Coupon Bonds and figure 4.1.

4.3 European Options

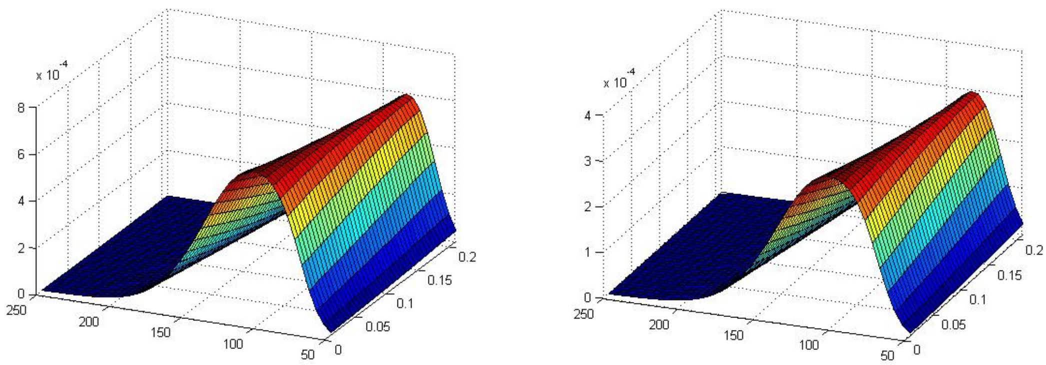
European options are characterized by the fact that the owner has the right, but not the obligation, to buy or sell the underlying stock at maturity. Here we consider a European Call option with strike $K = 100$. Taking the following parameter values for both the ADI as well as the CN method: $N_r = N_S = N_\tau = 40$, $\sigma_1 = 0.3, \sigma_2 = 0.01, \rho = 0.5, S_0 = 50, S_{max} = 250, r_{min} = 0.01, r_{max} = 0.21$. For the Hull White model we take the parameters $\theta = 0.02$ and $\alpha = 0.01$. We obtain the following price surfaces for $T = 0.5$.



(a) CN-Method: EU Call option (b) ADI-Method: EU Call option

Figure 4.2

The figures show that the price surfaces at time zero look very similar. These enable us to price a European Call option very fast for different stock prices and interest rates. In practice, one should take the grid as large as possible in order to prevent mis-pricing due to the linear boundary conditions. Let u_{CN} denote the solution obtained by the CN method for this European Call option. Similarly, let u_{ADI} be the solution obtained by the ADI method for this European Call option. In order to emphasize the difference between the results obtained by both methods, we compute the difference for $N_\tau = 40$ and $N_\tau = 80$



(a) The difference $u_{CN} - u_{ADI}$ for $N_\tau = 40$ (b) The difference $u_{CN} - u_{ADI}$ for $N_\tau = 80$

Figure 4.3

Note that the order of magnitude is -4. Hence the difference in results between these methods is very small. However, we see that this difference grows around the discontinuity of the payoff function. This may be caused by the fact that the CN method produces oscillations around the discontinuity $K = 100$. This, was already the case in the one-dimensional problem as shown in figure 3.1. This oscillatory behavior of the

CN method can be controlled and decreases when one uses more gridpoints in the time dimension. This is illustrated, in figure 4.3. For the case, $N_\tau = 40$ we have a maximum difference of $6.4512 \cdot 10^{-4}$ while for the case $N_\tau = 80$ we have a maximum difference of $3.2273 \cdot 10^{-4}$. Hence, this difference $u_{CN} - u_{ADI}$ decreases approximately with a factor of 0.5 whenever we double the number of gridpoints in the time dimension. To emphasize this behavior we present the maximum difference $u_{CN} - u_{ADI}$ for various values of N_τ in the table below.

	$N_\tau = 10$	$N_\tau = 20$	$N_\tau = 40$	$N_\tau = 80$
$max(u_{CN} - u_{ADI})$	0.0026	0.0013	$6.4512 \cdot 10^{-4}$	$3.2273 \cdot 10^{-4}$
	$N_\tau = 160$	$N_\tau = 320$	$N_\tau = 640$	$N_\tau = 1280$
$max(u_{CN} - u_{ADI})$	$1.6141 \cdot 10^{-4}$	$8.0716 \cdot 10^{-5}$	$4.0361 \cdot 10^{-5}$	$2.0181 \cdot 10^{-5}$

Let $N_\tau = 40$. The tables below shows some values of a European Call option for different stock values and interest rates regarding the CN and ADI methods.

CN-Method	$S = 150$	$S = 170$	$S = 200$	$S = 220$
$r = 0.01$	51.0050	70.7981	100.7498	120.7462
$r = 0.05$	52.9054	72.7445	102.7089	122.7064
$r = 0.1$	55.2440	75.1280	105.1040	125.1025

ADI-Method	$S = 150$	$S = 170$	$S = 200$	$S = 220$
$r = 0.01$	51.0047	70.7980	100.7498	120.7462
$r = 0.05$	52.9052	72.7444	102.7089	122.7064
$r = 0.1$	55.2438	75.1280	105.1040	125.1025

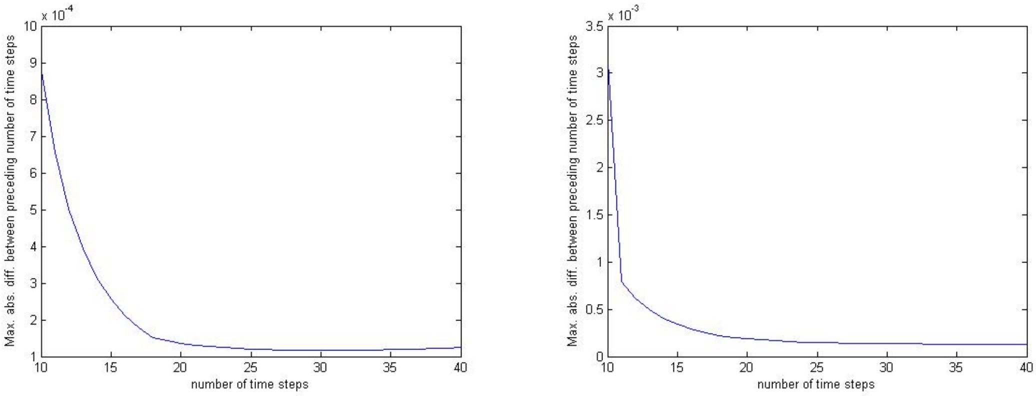
Up till now we only considered a constant drift in the Hull-White model. In chapter 2 we calibrated this parameter to an initial term structure. Since we modeled the methods backwards in time, we have to be careful that the θ-parameter is treated in the right way. Remember, that the θ-parameter is time dependent. Therefore we start with its value at maturity $t = 0.5$ and end at $t = 0$. In addition, we need to make sure that the time step between the data points that make up the calibrated θ-curve equals the time step used in the valuation of the European Call option. Using 2341 gridpoints for the time interval $[0, 30]$ gives the value 0.5 after 40 time steps. Thus, it exactly matches the valuation parameters of the European Call option which tells us that after $N_\tau = 40$ we move 0.5 in time.
Below we present the results obtained by using both methods when this time-dependent parameter is taken into account.

CN-Method	$S = 150$	$S = 170$	$S = 200$	$S = 220$
$r = 0.01$	50.7845	70.5716	100.5217	120.5179
$r = 0.05$	52.6877	72.5220	102.4851	122.4825
$r = 0.1$	55.0302	74.9106	104.8857	124.8840

ADI-Method	$S = 150$	$S = 170$	$S = 200$	$S = 220$
$r = 0.01$	50.7842	70.5715	100.5217	120.5179
$r = 0.05$	52.6875	72.5219	102.4851	122.4825
$r = 0.1$	55.0300	74.9105	104.8857	124.8840

The results differ strongly from the case in which we assumed a constant θ. This suggests that it is important to calibrate the time-dependent parameter. Otherwise, it is possible to obtain bad results as can be readily verified from the results presented in the table. In addition, we notice that the differences between the two methods have order of magnitude 3. Hence, we obtain very reliable results for both methods. Since $r_{min} = 0.01$ is at the boundary of our grid, it is also quite remarkable that both methods lead to

similar values. This indicates that linear boundary conditions do not pose any problem for both methods.
Another important factor is the time it takes in order to get a sufficient accurate solution. To this end, we compare the maximum absolute difference between the value of a European Call option obtained by taking N_τ time steps and $N_\tau - 1$ time steps. In addition, we fix the parameters $N_S = N_r = 40$. The following figure shows the result.



(a) CN method (b) ADI method

Figure 4.4

I want to stress out that the scales in figure 4.4 (a) and figure 4.4 (b) are different. From figure 4.4 it is clear that for the CN method the difference starts at a very small value relative to that of the ADI method. It can also be seen that this difference drops faster for the ADI method than for the CN method. The values at $N_\tau = 30$ is given by $1.17 \cdot 10^{-4}$ for the CN method while it is $1 \cdot 10^{-4}$ for the ADI method. Hence, the difference drops to a lower value for the ADI method.

In order to obtain a suitable comparison, we plot the logarithm of the maximum absolute error obtained by both methods against the computation time. Since we do not have an analytical solution we take $N_S = N_r$ and N_τ sufficiently high in order to obtain a reference solution. However, when we vary N_S and N_r we obtain grids that differ in size. Hence, we need a reference grid at which we compare the numerical approximation with the reference solution to obtain the error. Here, the reference grid is defined by $N_S = N_r = 10$. To provide a fair comparison we vary N_S, N_r and N_τ such that

$$\Delta S = \Delta r \quad \text{and} \quad \frac{\Delta t}{2\left(\Delta S\right)^2} = 5 \cdot 10^{-5}$$

The following figure shows the result.

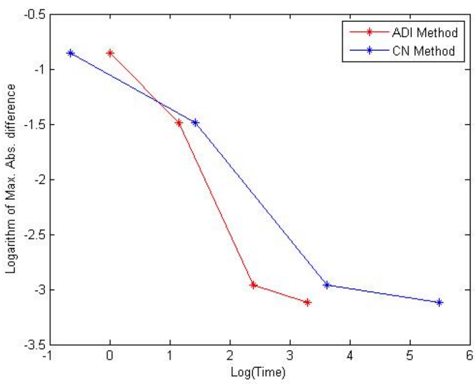
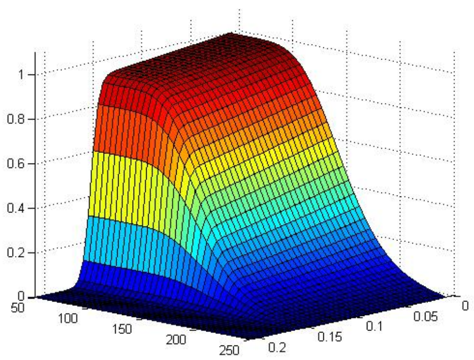


Figure 4.5: log error against time

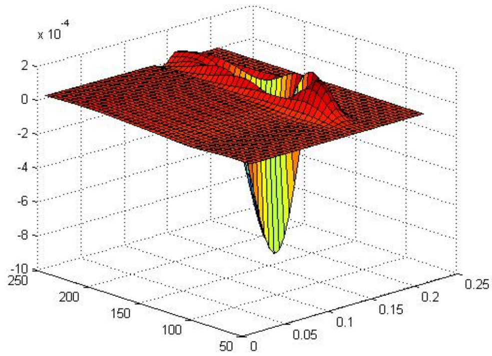
We measured the error for the four cases: $N_S = N_r = 10$, $N_S = N_r = 20$, $N_S = N_r = 30$ and $N_S = N_r = 40$, where $N_\tau = 50$, $N_\tau = 200$, $N_\tau = 450$ and $N_\tau = 800$ respectively. This implies that we need to adjust for each case the time step between the data points that make up the calibrated θ-curve in such a way that it equals the time step used in the valuation of the European Call option. Figure 4.5 shows that the CN method provides us very fast with a solution for small grids in which the error is quite big. For more accurate solutions, we see that the computational time increases significantly for the CN method in comparison to the ADI method.

4.4 Binary Options

A Binary option pays out a certain amount whenever the underlying as well as the interest rate is below a predetermined threshold. In this section we consider the Cash-Or-Nothing Binary option. This option is characterized by the fact that it pays out a fixed amount if the underlying as well as the interest rate is below a predetermined threshold and it pays out nothing if either the underlying or the interest rate is above this threshold. Note that this option has a payoff function that has a discontinuity along the interest rate threshold and a discontinuity along the asset price threshold. It is known that some ADI methods blow up along these discontinuities in the non-smooth payoff function [21]. In order to investigate if this is the case for the ADI method from chapter 2 we take the following parameter set: $N_r = N_S = N_\tau = 40$, $\sigma_1 = 0.3$, $\sigma_2 = 0.01$, $\rho = 0.5$, $S_0 = 50$, $S_{max} = 250$, $r_{min} = 0.01$, $r_{max} = 0.21$, $\theta = 0.02$, $\alpha = 0.01$ and $T = 0.5$. The threshold regarding the interest rate is given by $r_{thres} = 0.15$ and the threshold regarding the asset price is given by $S_{thres} = 150$. The option pays out an amount of one if $S \leq S_{thres}$ and $r \leq r_{thres}$. The figure below shows the result



(a) Cash-Or-Nothing option by CN method

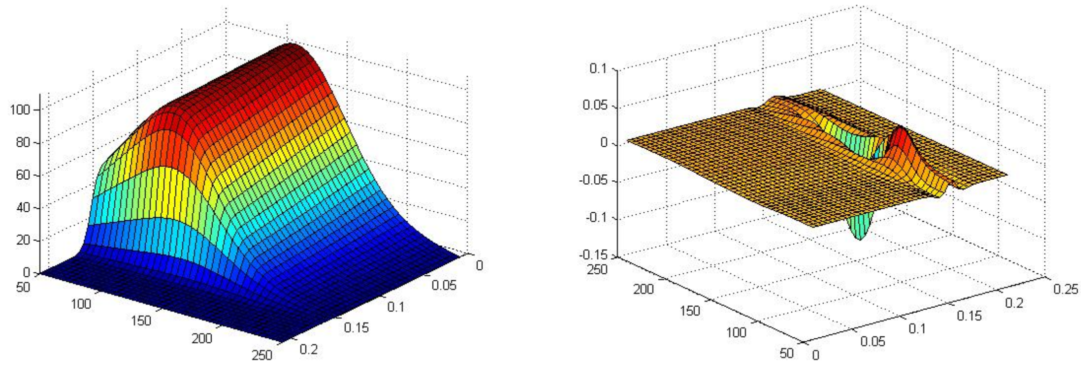


(b) The difference $u_{CN} - u_{ADI}$

Figure 4.6

Figure 4.6 (a) provides the solution obtained by the CN method in which the θ-parameter regarding the HW model is calibrated. This is done in exactly the same way as for the European Call option of section 4.3. Namely, we take the number of gridpoints in the time dimension such that the timestep in the cubic spline approximation equals that of the discretization in the time dimension. Note that the CN method doesn't show any signs of blowup behavior around discontinuities in the payoff function. From figure 4.6 (b), we conclude that this is also the case for the ADI method. It can be seen that there is a difference around $r = 0.15$. However, this difference is very small as the scale of the figure is given by 10^{-4}.

Besides the Cash-Or-Nothing option, there is also the Asset-Or-Nothing option. This option is characterized by the fact that it pays the value of the underlying asset whenever the asset price is below a predetermined threshold and the interest rate is below a predetermined threshold. Taking the same parameters as for the the Cash-Or-Nothing option along with the same thresholds we find the following result.



(a) Asset-Or-Nothing option by CN method (b) The difference $u_{CN} - u_{ADI}$

Figure 4.7

We observe a bump in the price surface in figure 4.7 (a). This bump is due to the fact that the asset price increases and therefore the value of the option increases, until the moment at which the threshold is exceeded and the option becomes worthless. The price surface of figure 4.7 (a) doesn't show any blowup behavior. However, the difference between both solutions is not as small as desired. In order to obtain a smaller difference we need to increase the number of time steps. If we significantly increase N_τ, say we take $N_\tau = 800$ instead of $N_\tau = 40$ then we obtain the following result.

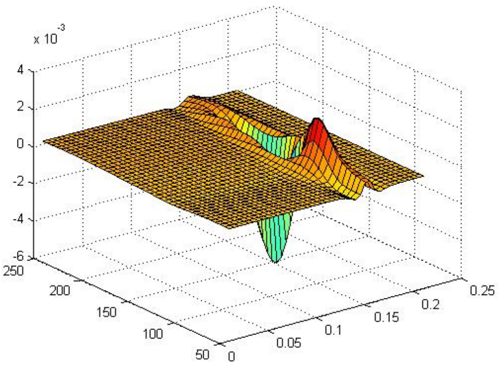
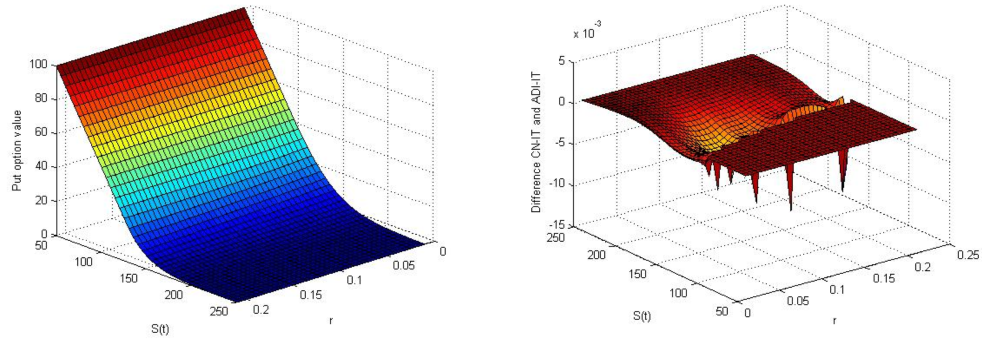


Figure 4.8: The difference $u_{CN} - u_{ADI}$ for $N_\tau = 800$

Figure 4.8 shows that the difference between the two solutions is similar as for the case in which we took $N_S = N_r = N_\tau = 40$, see figure 4.7 (b). However, the order of magnitude is much smaller, namely -3. The need for such an increase in the number of time steps is due to discontinuity along the interest rate threshold. The interest rate threshold leads to a steep surface as can be seen in figure 4.7 (a). Here, the CN method shows oscillatory behavior which can luckily be controlled.
Moreover, we do not have any blow up behavior in the solution obtained by both the CN method as well as the ADI method. This indicates that both methods work for both the Cash-Or-Nothing option and the Asset-Or-Nothing option.

4.5 American Options

The holder of an American option has the right to exercise the option before and at maturity. In section 3.3 we already saw that this extra right could lead to higher option values. Here we present the solution of an American Put option obtained by the CN-IT and ADI-IT methods. To this end, we use the following parameters set: $N_S = N_r = N_\tau = 40$, $\sigma_1 = 0.3$, $\sigma_2 = 0.01$, $\rho = 0.5$, $S_0 = 50$, $S_{max} = 250$, $r_0 = 0.01$, $r_{max} = 0.21$, $K = 150$ and $T = 0.5$. The parameters regarding the Hull-White model is given by $\alpha = 0.01$. Below, we depicted the solutions obtained by both methods.



(a) CN-IT Method: American Put option (b) The difference $u_{CN-IT} - u_{ADI-IT}$

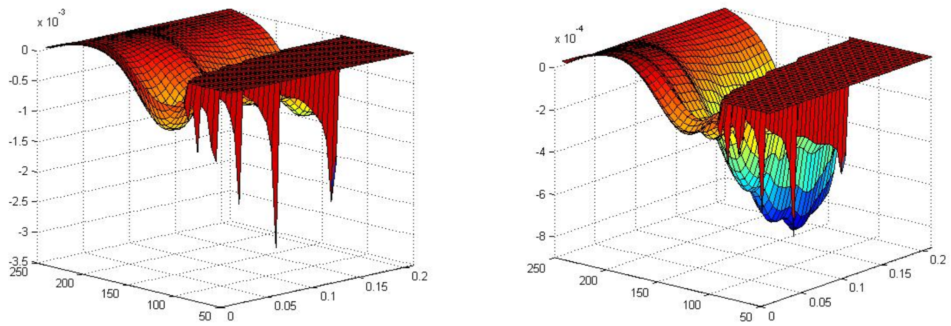
Figure 4.9

Note that the difference between the two solutions has order of magnitude -3. Hence both methods lead to similar results. For completeness we provide a table containing American Put option values for various values of r and S regarding both methods.

CN-IT Method	$S = 150$	$S = 170$	$S = 200$	$S = 220$
$r = 0.01$	12.2890	5.5723	1.4116	0.5043
$r = 0.05$	11.0600	4.8086	1.1493	0.3960
$r = 0.1$	9.7865	4.0129	0.8871	0.2916

ADI-IT Method	$S = 150$	$S = 170$	$S = 200$	$S = 220$
$r = 0.01$	12.2894	5.5719	1.4111	0.5039
$r = 0.05$	11.0623	4.8100	1.1494	0.3960
$r = 0.1$	9.7889	4.0146	0.8876	0.2917

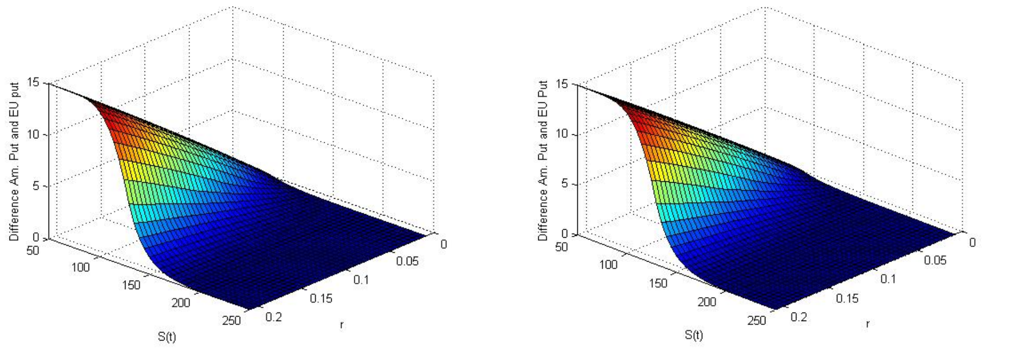
The results in these tables are obtained after calibration of the time-dependent parameter of the HW model. In section 4.3 we saw that calibration has a significant effect on the option values. Therefore, we chose to display the results after calibration.



(a) $u_{CN-IT} - u_{ADI-IT}$ for $N_\tau = 80$ (b) $u_{CN-IT} - u_{ADI-IT}$ for $N_\tau = 160$

Figure 4.10

Figure 4.10 shows that doubling the number of time grid points significantly reduces the difference between the two solutions. To see the difference between the solution of a European Put option u_{EU} and American Put option u_{AM} we present the following figure.



(a) CN-IT Method: Difference $u_{AM} - u_{EU}$ (b) ADI-IT Method: Difference $u_{AM} - u_{EU}$

Figure 4.11

One readily verifies that the difference between the American Put option and the European Put option significantly grows whenever the interest rate increases. This is due to the fact that a European Put option decreases in value whenever interest rates increase. Hence the difference between the American Put option value and the European Put option value becomes larger. In order to compare both methods, we plot the logarithm of the maximum absolute error obtained by both methods against the logarithm of the computation time. Moreover, we apply the same procedure as we did for the European Call option. Hence, we take $N_S = N_r$ and N_τ sufficiently high in order to obtain a reference solution. We measure the error at the gridpoints as defined by the reference grid, $N_S = N_r = 10$. In addition we vary N_S, N_r and N_τ such that

$$\Delta S = \Delta r \quad \text{and} \quad \frac{\Delta t}{2(\Delta S)^2} = 5 \cdot 10^{-5}$$

The result is presented in the plot below.

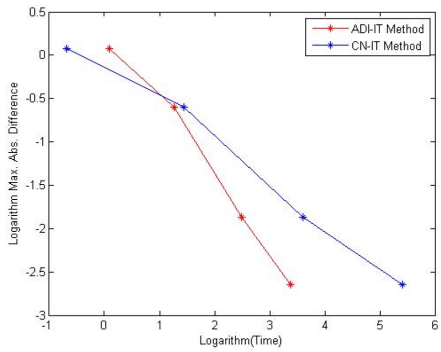
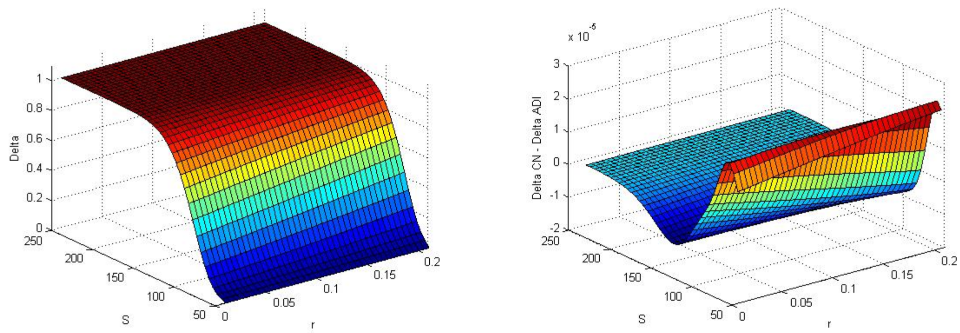


Figure 4.12: Logarithm of Error for Am. Put option

Figure 4.10 clearly shows that the CN-IT method provides us very fast with a solution. However, this approximation still has a very large error. In order to obtain a more accurate solution it is readily verified that the ADI-IT method provides us faster with such a solution.

4.6 Greeks

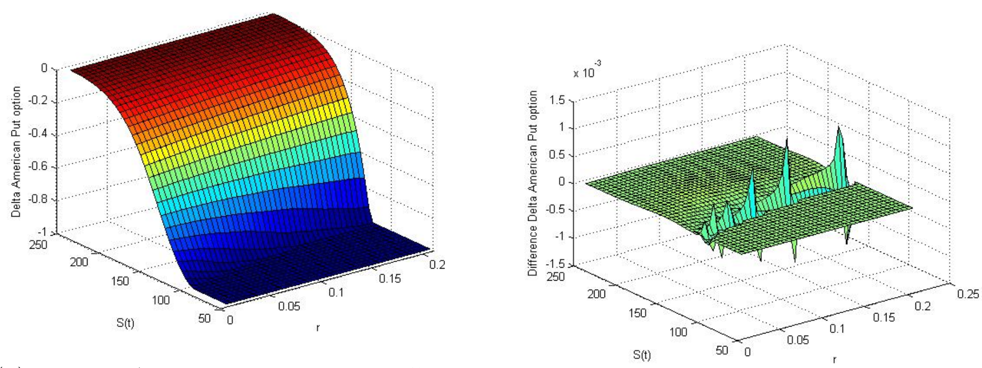
In this section we take $K = 100$, $N_r = N_S = N_\tau = 40$, $\sigma_1 = 0.3$, $\sigma_2 = 0.01$, $\rho = 0.5$, $S_0 = 50$, $S_{max} = 250$, $r_{min} = 0.01$, $r_{max} = 0.21$, $\alpha = 0.01$ and $T = 0.5$. Earlier on we determined the price of a European Call option. Besides the importance of knowing the right price of a European Call Option, one is also interested in knowing the sensitivity of this price with respect to a change in one of the parameters. These quantities are known as Greeks as they are assigned to a Greek letter. To this end, we first compute the sensitivity with respect to a change of 1 unit in the underlying. This quantity is called the Delta and is formally written as $\Delta = \frac{\partial u}{\partial S}$. We approximate this derivative by using a central difference scheme in order to obtain second order accuracy.



(a) Delta obtained by CN Method (b) The difference $\Delta u_{CN} - \Delta u_{ADI}$

Figure 4.13

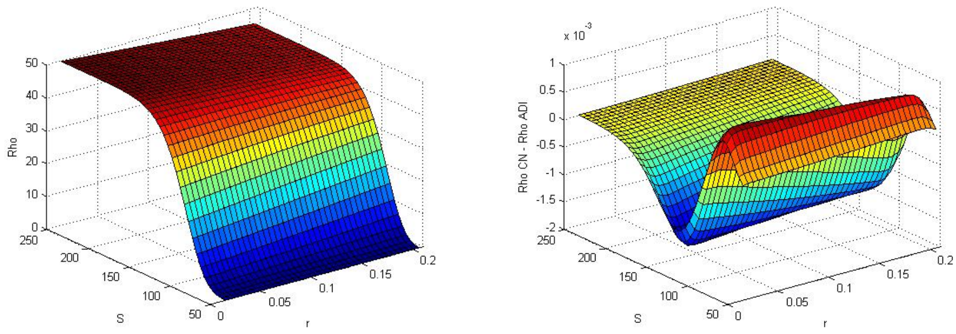
The difference of the Δ obtained by the CN method and the Δ obtained by the ADI method is very small. Note that the Δ itself ranges between 0 and 1. This is always the case for a European Call option. We also observe that the Δ increases quite fast to 1. This is explained by the fact that the strike price is $K = 100$ and the asset dimension is captured by $[50, 250]$, hence the Call option is for the most part in the money. For the American Put option we take the same parameter set as in the previous section: $N_S = N_r = N_\tau = 40$, $\sigma_1 = 0.3$, $\sigma_2 = 0.01$, $\rho = 0.5$, $S_0 = 50$, $S_{max} = 250$, $r_0 = 0.01$, $r_{max} = 0.21$, $K = 150$, $\alpha = 0.01$ and $T = 0.5$. The figure below presents the Δ of this American Put option.



(a) Delta Am. Put option by CN-IT method (b) The difference $\Delta u_{CN-IT} - \Delta u_{ADI-IT}$

Figure 4.14

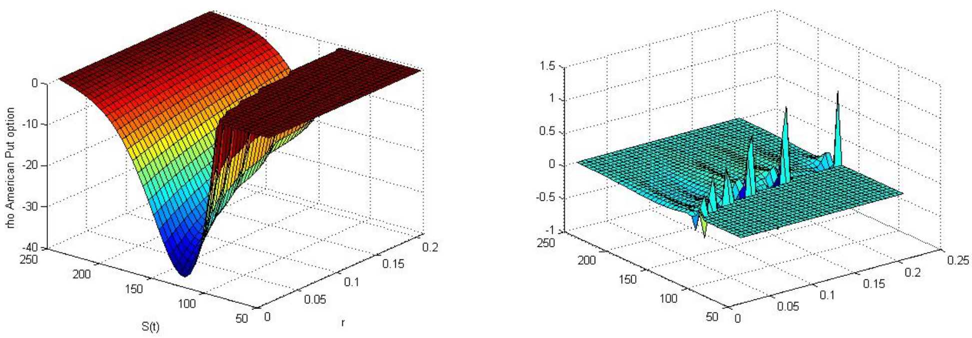
Notice that the Δ ranges between 0 and -1. This is always the case for Put options. Moreover, the difference between the Δ's show that both methods work. Another important factor is the sensitivity of the European Call option with respect to a change in interest rate. This quantity is called the Rho and is mathematically expressed by, $\rho = \frac{\partial u}{\partial r}$. Again, we approximate this quantity by a second order difference scheme. The result is given by the figure below.



(a) Rho obtained by CN Method (b) The difference $\rho u_{CN} - \rho u_{ADI}$

Figure 4.15

The ρ is often high when stock prices are high too and the time until maturity is long. Figure 4.15 shows that the ρ increases quite fast at first but this increase slows down and achieves almost 50. Again, we obtained a very small difference between both methods. This indicates that both methods provide similar results. For the American Put option we have the following figures.



(a) ρ by CN-IT method (b) The difference $\rho u_{CN-IT} - \rho u_{ADI-IT}$

Figure 4.16

We observe that the difference between the solutions of both methods is not as small as desired. Changing $N_\tau = 40$ into $N_\tau = 800$ yields a significant improvement.

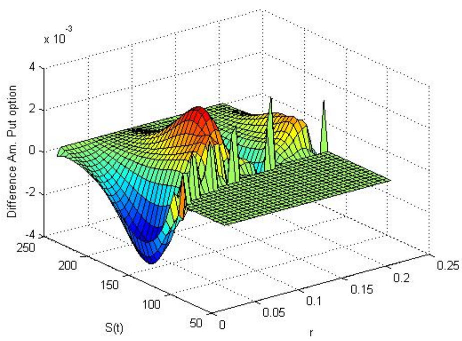


Figure 4.17: The difference $\rho u_{CN-IT} - \rho u_{ADI-IT}$ for $N_\tau = 800$

Chapter 5

Conclusion

In this thesis we investigated the valuation of European options by means of two numerical methods. In contrast to common practice in which stochastic volatility of the underlying is taken into account, we instead analyzed the cause of stochastic interest rates. To this end, we used the famous Hull White model which we calibrated by using an initial term structure. We priced several zero-coupon bonds along with European Call options, Binary options and American Put options. For these options we provided the results for both the Crank Nicolson method and the Alternating Direction Implicit method. This analysis steers us back to the research question:

For which options does the ADI method outperform the CN method in a model with stochastic interest rates?

Recall, that the Crank-Nicolson method is very practical for the one-dimensional problem of section 3.3. It leads to reliable and accurate solutions. However, towards more realistic problems including the Black Scholes equation extended with stochastic interest rates, we enter a domain of problems containing more than one dimension. The treatment of such problems is computationally inefficient when a CN method is used. In addition, we saw that the CN method gives rise to bad damping properties around the discontinuity of the payoff function, along with a large computational time in order to obtain a sufficient accurate solution. The ADI methods provides a practical alternative. First of all, it is based on a brief formulation of the problem, by partitioning it into subproblems along to the different dimensions. Secondly, the ADI method gives rise to an inversion problem of tridiagonal matrices. These type of matrices are easy to handle and there are several well known algorithms to, efficiently inverting them. Thirdly, it has better damping properties around the discontinuity of the payoff function along with a much faster computational speed for large dimensions. At last, it is easier to implement in a software package such as Matlab, which makes it more practical.

For the valuation of zero-coupon bonds we were able to compare the results with those obtained by a closed form solution. These results were very close for both methods. The difference between the results obtained by both methods for the valuation of zero-coupon bonds, European options, Binary options and American options is negligible. However, due to bad damping properties of the CN method one needs a finer grid in comparison to the ADI method. At first, we only considered the case in which the time-dependent parameter θ of the Hull White model was constant. Afterwards, we took the calibrated θ and this led to more realistic results. Both methods still provided us with consistent results. In conclusion, it is safe to say that the ADI method is suitable for the valuation of European, Binary and American options under stochastic interest rates. It provides us faster with a sufficient accurate result than the CN method, see for instance figure 4.5 and figure 4.12. We saw that this is not only desired for obtaining a right price for an option but also in computing the Greeks. Especially the ρ needs a very fine grid on order to overcome the possibility of inaccurate results due to oscillations of the CN method.

5.1 Recommendation

The ADI method is often used in solving heat conduction or diffusion problems. The difficulty in the partial differential equation we focused on presented itself in the way the cross derivative term needed to be handled. This term leads to the introduction of an extra operator besides the two operators corresponding to the different space dimensions. As far as I know there isn't any other way to treat this term. On the other hand, the CN method faces some serious difficulties for problems involving multiple dimensions. The most important factors involve the computational time until one obtains a sufficient accurate solution, the practicality of implementing the scheme and the bad damping properties around a discontinuity. These shortcommings have not only a significant impact on the valuation of options but also on the Greeks. To this end, I would recommend to use the CN(-IT) method whenever one wants to obtain a solution very fast but the approximation doesn't need to be accurate. On the other hand, I would recommend to use the ADI(-IT) method whenever an accurate solution is needed.

5.2 Future Research

Taking the deficiencies of the ADI method into account, I recommend to be careful when applying this method to a partial differential equation with many cross derivative terms as it may lead to many operators. To this end, it is interesting to know how such problems can be solved in a way such that the ADI scheme doesn't lose too much of its attractive properties. On the other hand, it would also be interesting to know how a property as bad damping around a discontinuity could be treated to make the CN method more attractive. In addition, an implementation that enhances the computational speed of the CN method could make the method practically more attractive. Besides, we didn't apply the ADI method to exotic options under the Hull White model. This would be an interesting topic for future research. Also, a possible extension to the two-factor Hull White model along with a comparison to the one-factor Hull White model would be very interesting.

Bibliography

[1] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of political economy*, vol. 81, no. 3, pp. 637–654, 1973.

[2] R. C. Merton, "Theory of rational option pricing," *The Bell Journal of economics and management science*, pp. 141–183, 1973.

[3] S. L. Heston, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *The review of financial studies*, vol. 6, no. 2, pp. 327–343, 1993.

[4] L. A. Grzelak and C. W. Oosterlee, "On the Heston model with stochastic interest rates," *SIAM Journal on Financial Mathematics*, vol. 2, no. 1, pp. 255–286, 2011.

[5] P. Wilmott, S. Howison, and J. Dewynne, *The mathematics of financial derivatives: a student introduction.* Cambridge University Press, 1995.

[6] R.-A. Dana and M. Jeanblanc, *Financial markets in continuous time.* Springer Science & Business Media, 2007.

[7] R. Seydel and R. Seydel, *Tools for computational finance*, vol. 3. Springer, 2006.

[8] C. W. Cryer, "The solution of a quadratic programming problem using systematic overrelaxation," *SIAM Journal on Control*, vol. 9, no. 3, pp. 385–392, 1971.

[9] C. Schwab, N. Hilber, and C. Winter, "Computational methods for quantitative finance," *Lecture notes and exercises, ETH Zurich*, 2007.

[10] N. Hilber, O. Reichmann, C. Schwab, and C. Winter, "Finite Element methods for Parabolic Problems," in *Computational Methods for Quantitative Finance*, Springer, 2013.

[11] J. Hull and A. White, "Pricing interest-rate-derivative securities," *The Review of Financial Studies*, vol. 3, no. 4, pp. 573–592, 1990.

[12] D. Brigo and F. Mercurio, "Interest rate models theory and practice," 2001.

[13] N. Rom-Poulsen, "Computations in the Hull-White Model," 2005.

[14] J. van Kan, A. Segal, and F. Vermolen, *Numerical methods in scientific computing.* VSSD, 2005.

[15] D. J. Higham, *An introduction to financial option valuation: mathematics, stochastics and computation*, vol. 13. Cambridge University Press, 2004.

[16] J. Douglas Jr and S. Kim, "On accuracy of Alternating Direction Implicit methods for parabolic equations," *Preprint*, 1999.

[17] J. D. DANIEL, "Finite Difference Methods in Financial Engineering," *J. Duffy.-Hoboken, NJ: Wiley*, 2006.

[18] N. N. Janenko, *The Method of Fractional Steps.* Springer, 1971.

[19] K. in ’t Hout and R. Valkov, “Numerical study of splitting methods for American option valuation,” *ArXiv e-prints*, Oct. 2016.

[20] T. Haentjens, K. int Hout, and K. Volders, “ADI Schemes with Ikonen-Toivanen Splitting for pricing American Put options in the Heston Model,” in *AIP Conference Proceedings*, vol. 1281, pp. 231–234, AIP, 2010.

[21] D. Jeong and J. Kim, “A comparison study of ADI and operator splitting methods on option pricing models,” *Journal of Computational and Applied Mathematics*, vol. 247, pp. 162–171, 2013.