Alternating Direction Implicit methods for solutions of the Heston stochastic volatility model

A Dissertation submitted for the degree of

Master of Science

in

Financial Mathematics

by

Magali Calabressi
Student Number 13037211

Supervised by Dr Riaz Ahmad
Department of Mathematics
University College London

September 2014

Abstract

This is a numerical study of ADI schemes applied to problems in finance. The work carried out centers on the Heston stochastic volatility model, which gained its popularity from the existence of closed form solutions. A detailed outline of the analytical procedure for applying the Fourier transform method is presented. ADI schemes Douglas-Rachford, Craig-Sneyd (CS), modified CS and Hundsdorfer-Verwer are studied in detail and the numerics are implemented in Matlab. The option sensitivities (Greeks) are also introduced in theory and some are briefly studied via their closed form solution and finite difference method.

Declaration

I hereby certify that all material in this dissertation which is not my own work has been properly acknowledged.

Acknowledgements

My deepest gratitude goes towards my supervisor and professor Dr Riaz Ahmad, for his guidance and patience throughout this dissertation. His invaluable support and countless suggestions have seen this dissertation be possible. I am also grateful to my classmates who have been a fountain of inspiration and support.

Above all I am indebted to my mum, without whom these opportunities I have had would not have been possible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The existence of a curved implied volatility surface has led to a rise in the popularity of the use of stochastic volatility (SV) models in derivatives pricing and hedging. This phenomenon has highlighted a fundamental drawback of the classic Black-Scholes model; the assumption of constant volatility. In place of the Black-Scholes model, many SV alternatives have been proposed (e.g. Stein and Stein, SABR Heston, Hull and White models). Amongst them, one of the most prominent and widely used is the Heston model. Part of its popularity can be attributed to the availability of a semi-analytical solution [24] in the case of European options, which allows not only for the ability to extensively test numerical methods, but also for calibration of model (see for example [36]). A model that sits closely to the Heston model is the 3/2 model, although the Heston model remains more popular and widely used.

Derivatives pricing via the Heston model can be done via the corresponding partial differential equation (PDE) or via simulation techniques such as Monte Carlo (MC) methods. MC methods require numerical integration of stochastic differential equations (SDEs). As such, they use simulation methods to approximate integrals in the form of expectations. As this requires generating random variables, MC methods are computationally expensive. For this reason, they are more attractive to use when dealing with four or more dimensions. If the problem at hand has three or less dimensions, finite difference methods (FDM) tend to be used instead. With this in mind, as the Heston is a two-dimensional model, alternating direction implicit (ADI) methods are studied. ADI schemes were initially used for problems in fluid mechanics. The vast majority of problems arising in finance can be modeled in terms of parabolic PDEs. It is the existence of fast numerical PDE schemes which motivates the use and attractive features of techniques from the fluid dynamics arena to solving problems in the field of quantitative finance. ADI methods employed in multi-factor modeling has proved particularly powerful.

The implementation developed is based on a European call option and focuses on the performance of different ADI methods. The key idea is to clearly introduce concepts which will aid the understanding of the need of ADI schemes and motivate the use of SV models and in particular the Heston model. The dissertation is also intended to provide a solid understanding of the implementation of ADI methods, which is not easy to grasp.

Chapter 2

General Review

2.1 Motivation

Suppose we use the standard deviation of... possible future returns on a stock... as a measure of its volatility. Is it reasonable to take that volatility as constant over time? I think not.

—Fischer Black, *1973*

The Black-Scholes model of 1973 [3] has provided an essential foundation, which has served as the building blocks of financial engineering over the past forty years. Used to price derivative securities, the purpose of the Black-Scholes model was to form a risk-neutral instrument that could be perfectly hedged against volatility in markets. In their formative paper [3], Black and Scholes derive the pricing partial differential equation (PDE) that governs the price of a security over time. The central idea behind their method was that of constructing a portfolio in which a long position in one option could be hedged by a short position in Δ amount of the underlying. This strategy is also referred to as delta hedging.

However, it has been accepted since by both academics and practitioners, that certain ideal assumptions of the model do not conform with features observed in equity markets. Many papers highlight the weaknesses and limitations of the Black-Scholes model (see for example Haug and Taleb [23]). In this discussion, we concern ourselves with the distributional implications of the Black-Scholes model, although there are further simplifications which have also been criticized.

We remark that one of the most pre-eminent observations is the inability of the Gaussian distribution assumed for log-returns to capture extreme events characterized by the large tails of actual market data distributions. This distributional result is a consequence of modeling asset price dynamics as geometric Brownian Motion (GBM). Contradicting evidence (see for example Gatheral [21] and Mandelbrot [33]) is given by empirical studies which have shown that in practice, the probability density function (PDF) is characterized by a slower decay of tails than the log-normal density predicts ('fat-tails' effect) and high peaks (higher kurtosis, known as Leptokurtic).

A further critical assumption of the Black-Scholes model is that volatility is assumed known and constant over time. Firstly, there is empirical evidence for the tendency of equity returns to be negatively correlated with volatility (known as the 'Leverage Effect') where typically, declines in stock prices are accompanied by an increase in volatility rather than declines in volatility leading to a rise in stock prices. Intuitively this makes sense as drops in asset price would cause an institution to become more leveraged, stock becomes riskier and thus volatility rises. Secondly, the fact that the distribution of stock price returns is fat-tailed and highly peaked indicates a mixture of distributions with different variance (see Kwok [29]). If we plot asset price returns (or log-returns) of actual equity data we often observe the 'volatility-clustering' feature (see Gatheral [21]) which reflects the mean-reversion property of the variance process. These observations have led to the notion of random volatility and incorporation of jumps (we can introduce jumps by relaxing the continuity assumption of the Black-Scholes model). In this dissertation, we concern ourselves with the former of the two notions.

2.1.1 Volatility

The Black-Scholes pricing framework gives rise to option prices that are functions of the input parameters: stock price S, strike price K, risk-free interest rate r, maturity T, current time t and volatility σ. Of these five parameters, σ is the only quantity which is not directly observable. It represents a measure of the uncertainty of asset returns. The difficulties arising in valuing σ are a result of the fact that the value the parameter takes should be the forecast value over the time left until expiration of the option rather than and estimate from past market data (such estimate is known as historical or realized volatility) [29].

Given an observed European call option price $C_{obs}(S,t;K,T)$ for an option contract with strike K and maturity T, the implied volatility σ_{imp} is defined as the value of the volatility parameter that must go into the Black-Scholes formula [3]

$$C_{BS}(S,t;K,T,\sigma) = S_t N(d_1) - Ke^{-r(T-t)}N(d_2),$$

where

$$d_1 = \frac{\ln(\frac{S_t}{K}) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

and

$$d_2 = \frac{\ln(\frac{S_t}{K}) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

with

$$N(x) = \frac{1}{2\pi}\int_{-\infty}^{x} e^{-\frac{1}{2}z^2}\,dz,$$

to match this price

$$C_{BS}(S,t;K,T,\sigma_{imp}) = C^{obs}.$$

We pause to observe that[1]

[1]For further details see Fouque et al. [19].

(1) Noticing that the vega of a European call option is positive

$$\frac{\partial C_{BS}}{\partial \sigma} = \frac{Se^{-\frac{d_1^2}{2}}\sqrt{T-t}}{\sqrt{2\pi}} > 0,$$

we can always find a unique positive implied volatility σ_{imp} given $C^{obs} > C_{BS}(S,t;K,T)$ due to the monotonicity of the Black-Scholes formula in the volatility parameter σ.

(2) The implied volatility σ_{imp} of European call and put options with the same strike K and maturity T are the same because of put-call parity[2].

Implied volatility can be seen as the market's expected future volatility between the present time t and the maturity T of a given option. In particular, we can seek an extensive market perspective by observing several implied volatility values obtained simultaneously from options with differing strike and maturity values written on the same underlying security (such a plot gives the *term structure of implied volatility*). This could be useful to traders who often quote the implied volatility of an option as opposed to its price. The advantage and convenience of doing this is due to the lesser variability in implied volatility than an option price.

One may use implied volatilities to provide further evidence against constant volatility in the Black-Scholes model. Suppose we observe market prices for several European call options with the same maturity written on a common underlying asset. If the assumptions made in the Black-Scholes model were a true representation of the market then we should see flat implied volatilities (i.e. a plot of strike or moneyness[3] against implied volatility would yield a flat horizontal line). Instead, one normally sees a volatility 'smile' where implied volatility is higher at either side of at-the-money (ATM) (i.e. when the spot and strike price are the same). The existence of a smile is evidence that the market prices at a premium out-of-the-money (OTM) (i.e. when the spot exceeds the strike in the case of a put and strike exceeds spot in case of a call) puts and in-the-money (ITM) (i.e. opposite of OTM) calls. By contrast, in a 'Black-Scholes world' unless volatility is increased a lot, chances of ending up ITM are a rare event. In this case, the log-normal model underestimates the probability of large movements in the underlying.

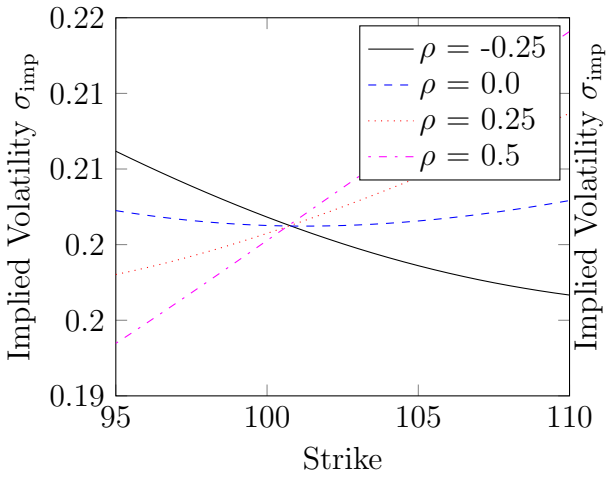
In practice we often observe a 'smirk' (a reverse skew, asymmetry of the distribution). This is because underpricing of OTM puts by Black-Scholes is much stronger. The higher premium the market places on OTM puts may be intuitively explained as follows; the market perceives drops in prices much more likely than hikes and/or investors are more worried about market crashes and thus buy puts for protection.

2.1.2 Heston and the Volatility Smile

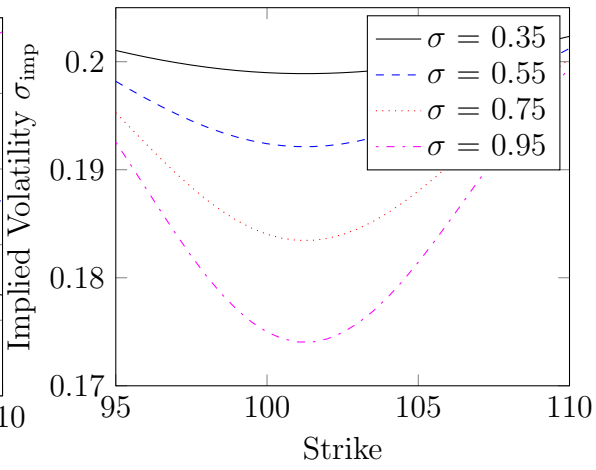
If we look at an implied volatility plot resulting from the stochastic volatility model introduced by Heston (1993) [24] we can observe the phenomenon described above.

[2]The put-call parity formula for European options is $C - P = S - Ke^{-r(T-t)}$, where C and P represent the call and put option price, respectively.

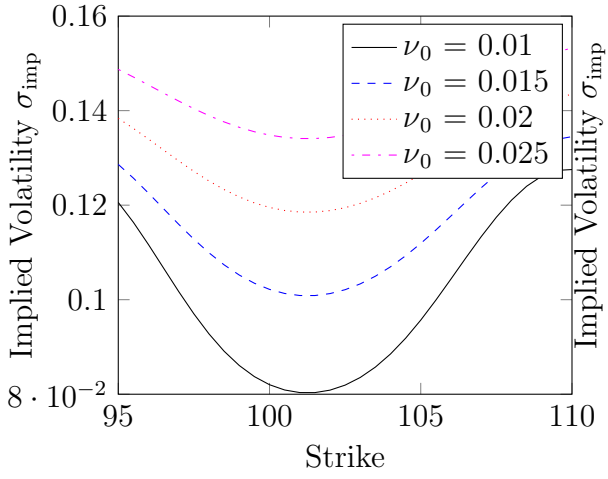
[3]Moneyness can be defined as $m = \frac{\ln(\frac{K}{F})}{\text{ATMV}\sqrt{T}}$ where K is the strike, F is the forward price, T is the maturity and ATMV is the at-the-money implied volatility.



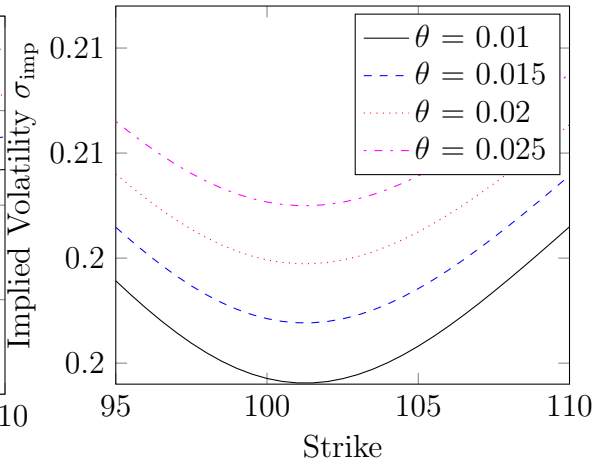
(a) Effect of ρ on Heston implied volatility smile with parameter values $\sigma = 0.25$, $\kappa = 2$, $\theta = 0.01$, $\nu_0 = 0.05$.



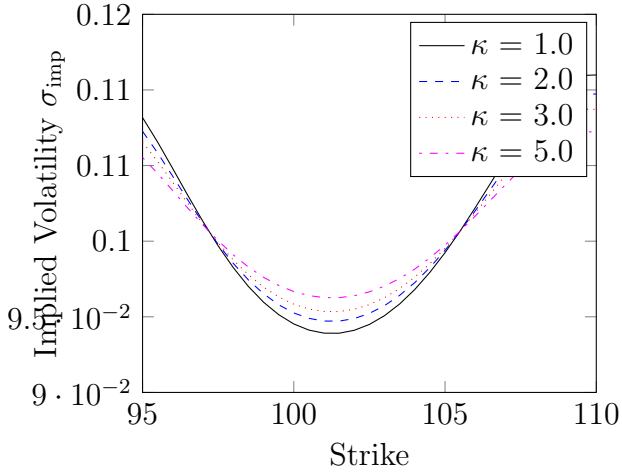
(b) Effect of σ on Heston implied volatility smile with parameter values $\rho = 0$, $\kappa = 2$, $\theta = 0.01$, $\nu_0 = 0.05$.



(c) Effect of ν_0 on Heston implied volatility smile with parameter values $\sigma = 0.5$, $\rho = 0$, $\kappa = 2$, $\theta = 0.01$.



(d) Effect of θ on Heston implied volatility smile with parameter values $\sigma = 0.5$, $\rho = 0$, $\kappa = 2$, $\nu_0 = 0.05$



(e) Effect of κ on Heston implied volatility smile with parameter values $\sigma = 0.2$, $\rho = 0$, $\theta = 0.01$, $\nu_0 = 0.01$.

This presents the first benefit of modeling stochastic volatility in describing features that actual market data display.

The first plot in Figure **??** shows that the direction of the reverse skew is affected by correlation ρ, with a positive slope if $\rho > 0$ and a negative slope if $\rho < 0$. Changing the correlation ρ changes the degree of symmetry with calls being more expensive when $\rho > 0$ and puts more expensive when $\rho < 0$ (see Maslova [34]). In reality, equity options tend to give a negative slope in implied volatility. Indeed, this conforms the earlier observation that when options data are used to calculate implied volatilities its common to see the market requires a premium for OTM put options.

Next, we examine the volatility of the variance σ (vol of vol). Its apparent from Figure 2.1b that an increase in σ results in an increase in the convexity of the smile. As σ approaches zero we expect a deterministic process for the variance which implies a flat horizontal line for implied volatility (this is again the assumption made in the Black-Scholes model).

The initial variance ν_0 allows adjustments in the level of the smile, and has a much lesser influence on the curvature, see Figure 2.1c. The effect of changing the level of mean reversion θ on the volatility smile is similar to that of changing the initial variance. Namely, the level of the smile is adjusted as θ is changed, see Figure 2.1d. Finally, the speed of mean reversion κ has an effect on the curvature of the volatility smile, see Figure 2.1e.

2.2 Local Volatility

If we decide to remain within the framework of a single factor diffusion process but wish to allow volatility to be time or state dependent or both, then we are working with local volatility (also known as actual volatility). Local volatility is essentially a measure of how much randomness is present in an asset's return at a given time. It is therefore a quantity that exists at a particular instant, making it difficult to measure. Through this definition, volatility is only constant 'locally' according to a specific asset price and time to maturity. If the volatility is time-dependent only, it can be shown that $\sigma(T)$ may be derived from available implied volatilities σ_{imp} (see Kwok [29] for an exercise on how to show this). Next, consider a state-time dependent volatility function and suppose that we can compute the implied volatilities σ_{imp} of a series of European option prices at a number of different strikes and maturities. The aim is to find such a volatility function $\sigma(S_t, t)$ that gives the Black-Scholes option prices consistent with observed market option prices under risk-neutrality. The function $\sigma(S_t, t)$ is known as the *local volatility function* (see initial work by Dupire [17] and Derman and Kani [13]).

For the sakes of completeness, we mention that as expected, an empirical study undertaken by Duman, Fleming and Whaley (1998) [16] confirmed the dynamics of the implied volatility surface did not conform the assumption of constant local volatilities.

2.3 A Short Motivational Study of Market Data

The following is an experiment carried out as evidence of some of the PDF characteristics observed in equity markets described earlier in section 2.1.

Daily adjusted close prices for S&P500[4] from January 1950 to June 2014 were downloaded from *Yahoo! Finance* and exported to Excel. In order to calculate the return for a particular day, we use the adjusted price for that day and for the previous day. As we don't have the price for the day preceding the first, we cannot calculate returns for the first day. Starting with the second day, the returns are calculated as follows:

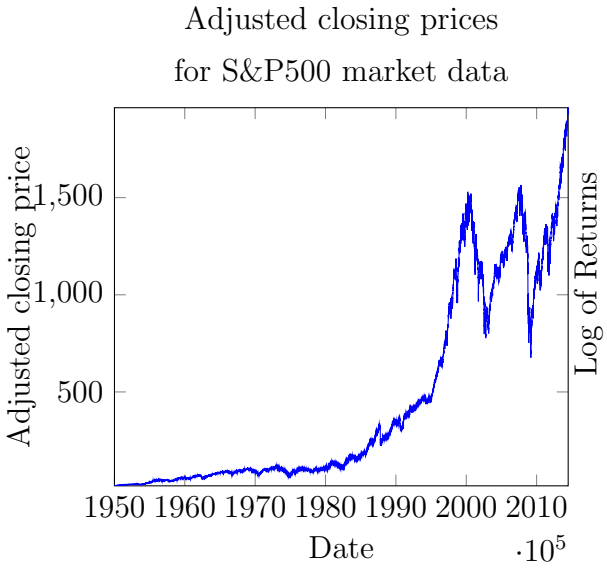
$$r_2 = \frac{p_2 - p_1}{p_1}, \tag{2.1}$$

where p_i denotes the adjusted closing price, in which the effect of cash dividends, stock dividends and stock splits are already incorporated. In order to construct the PDF, we use scaled returns. These are calculated as follows:

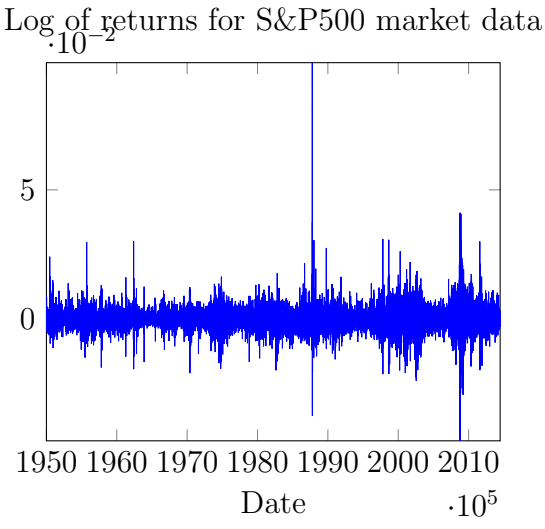
$$\bar{r} = \frac{r - \mu_r}{\sigma_r}, \tag{2.2}$$

where μ_r is the average and σ_r the standard deviation of the returns, respectively.

In Figure 2.2a we plot the raw adjusted closing prices data to observe stock movement. Figure 2.2b is a plot of the log-returns[5] of S&P500 over a 64 year period. We can observe the phenomenon of 'volatility-clustering' described earlier (see [21] and [29]). Such phenomenon implies auto-correlation in variance resulting from the mean-reverting property of the data.



(a) S&P500 scaled returns dynamics over the 64-year period.



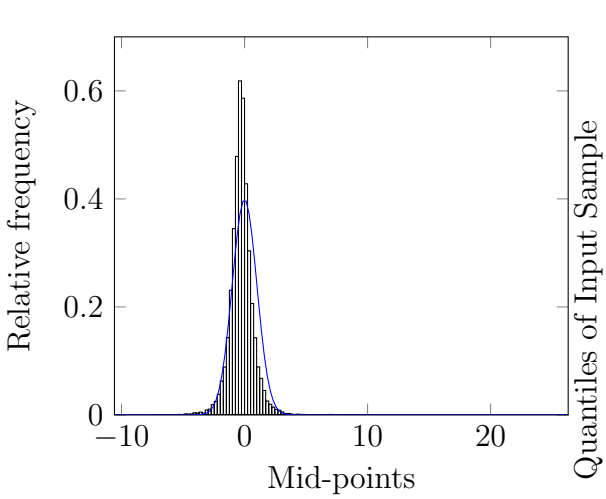
(b) S&P500 daily log-returns over the 64-year period.

Figure 2.2: Empirical study plots of S&P500 returns data.

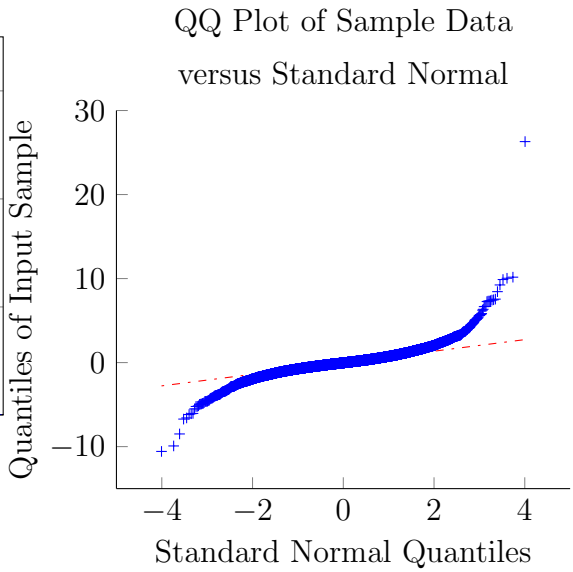
[4]This is Standard & Poor's 500 index, a basket of 500 stocks chosen according to factors such as market size, liquidity and industry grouping.

[5]The log-return of the underlying is evaluated as $R_t = \log(\frac{X_t}{X_{t-1}})$.

In Figure 2.3a, we plot the frequency distribution (note we use relative frequency) of the scaled returns over a 64 year period from 1950 to 2014. We can see that the distribution is heavy tailed and high peaked relative to the normal distribution as similar empirical studies show (see for example Gatheral [21]). The QQ-plot in Figure 2.3b shows deviations from the tails of the empirical distribution indicating that the data inconsistent with a normal distribution. The 'fat'-tails and high peaks indicate a series of mixed distributions with differing variances motivating the need for stochastic volatility models.

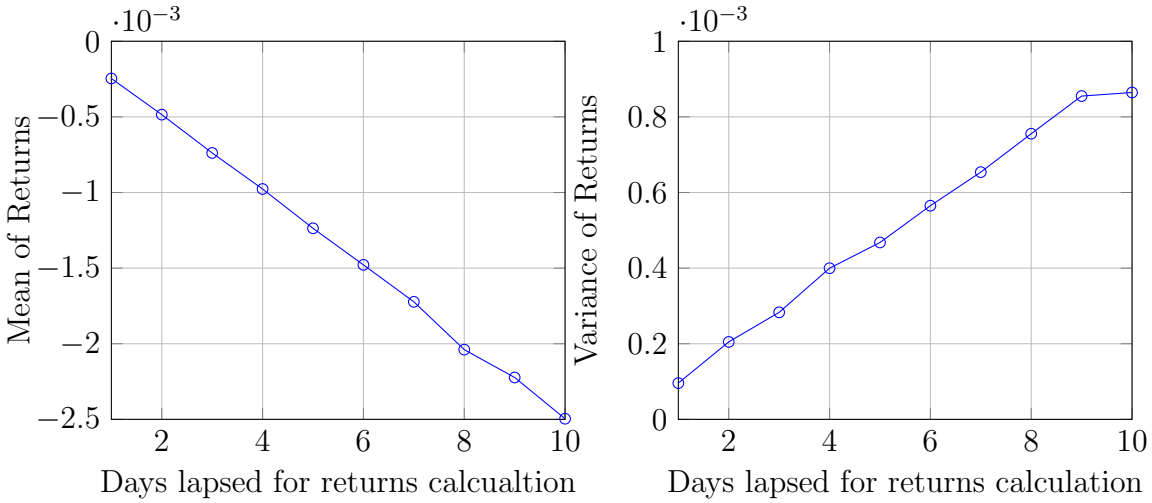


(a) Frequency distribution of S&P500 scaled returns data relative to the normal distribution. It can be seen that empirical distribution has extreme tails and high peaks.



(b) QQ-plot of S&P500 scaled returns relative to the normal distribution. Large deviations from the tails suggest scaled returns are not normal.

Figure 2.3: Plots showing the deviations of the actual distribution of returns from the Gaussian distribution assumed by the theoretical Black-Scholes model.



(a) Plot of the mean of the returns for different day lapses.

(b) Plot of the variance of returns for different day lapses.

Figure 2.4: Plots showing the mean and variance of returns for different days.

Figure 2.4a is a plot of the mean of returns and Figure 2.4b a plot of the variance of returns, where we evaluate returns over 1, 2, 3, 4, 5, 6, 7 and 8 days. Notice that the mean of the returns behave like μdt. Both have a decreasing linear relationship. Also, the standard deviation scales with $\sqrt{dt}$.

We have thus motivated and justified the modeling of variance as a mean-reverting random variable.

Chapter 3

The Heston Model

3.1 Heston Dynamics

The Heston Model (1993) is represented by the system of bi-variate stochastic differential equations (SDE):

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{\nu_t} S_t dW_t^1 \\ d\nu_t &= \kappa(\theta - \nu_t)dt + \sigma\sqrt{\nu_t} dW_t^2 \end{aligned} \tag{3.1}$$

where $\mathbb{E}^{\mathbb{P}}[dW_t^1 dW_t^2] = \rho dt$ with $\mathbb{P}$ the physical measure, $\{S_t\}_{t\geq 0}$ and $\{\nu_t\}_{t\geq 0}$ are the stock price and volatility processes in turn, $\{W_t^1\}_{t\geq 0}$ and $\{W_t^2\}_{t\geq 0}$ are correlated Brownian motions with correlation coefficient $\rho \in [-1, 1]$. The variance follows a mean-reverting square root process, first used by Cox-Ingersoll-Ross in 1985 (see [10]) to capture interest rate dynamics, with $\kappa > 0$ the mean reversion rate, and $\theta > 0$ the mean reversion level. $\sigma > 0$ is known as the volatility-of-volatility. μ is the drift of the stock process.

3.1.1 The Feller Condition

In order for the mean-reverting square root dynamics for the variance to remain positive, there are various analytical results that may be employed. In particular, the Feller condition states that if $2\kappa\theta \geq \sigma^2$ then the variance process can never reach the origin. This means that the volatility cannot become negative. On the contrary, if $2\kappa\theta < \sigma^2$ then the origin is attainable and strongly reflecting so that the variance process may attain zero in finite time, without spending time at this point.

3.2 The Heston Pricing Partial Differential Equation

In this section we derive the Heston PDE. This is a special case of the general derivation for stochastic volatility models, which can be found in Gatheral (2006)

[21].

In the Black-Scholes case volatility is assumed constant. Thus there is only one source of randomness coming from the Brownian Motion used to model stock dynamics, which can be hedged by trading in the underlying stock.

The quantity we are now modeling, stochastic volatility, is not a traded asset. As such, it cannot be hedged by trading the underlying. Instead, to hedge random changes in volatility we trade in two underlying contracts with different maturities. We retain other assumptions of the Black-Scholes model. Mainly, the stock pays no dividends; there are no transaction costs or taxes; short-selling is allowed; and any trading is continuous.

Assume the market is specified by the SDEs (3.1) and the dynamics of the risk free asset are:

$$dB_t = rB_t dt, \tag{3.2}$$

where the risk free rate r is assumed constant. Form a portfolio consisting of one option $U(S, \nu, t)$, $-\Delta$ units of stock, and $-\Delta_1$ units of another option $U_1(S, \nu, t)$ to hedge the volatility. The value of the portfolio is then

$$\Pi = U - \Delta S - \Delta_1 U_1. \tag{3.3}$$

Assuming the portfolio is self-financing, the change in the portfolio across a time step dt is given by

$$d\Pi = dU - \Delta dS - \Delta_1 dU_1. \tag{3.4}$$

3.2.1 Portfolio Dynamics

An application of Itō's Lemma yields

$$\begin{aligned} d\Pi = &\left\{ \frac{\partial U}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 U}{\partial S^2} + \rho\sigma\nu S \frac{\partial^2 U}{\partial \nu \partial S} + \frac{1}{2}\sigma^2 \nu \frac{\partial^2 U}{\partial \nu^2} \right\} dt \\ &- \Delta_1 \left\{ \frac{dU_1}{dt} + \frac{1}{2}\nu S^2 \frac{\partial^2 U_1}{\partial S^2} + \rho\sigma\nu S \frac{\partial^2 U_1}{\partial \nu \partial S} + \frac{1}{2}\sigma^2 \nu \frac{\partial^2 U_1}{\partial \nu^2} \right\} dt \\ &+ \left\{ \frac{\partial U}{\partial S} - \Delta_1 \frac{\partial U_1}{\partial S} - \Delta \right\} dS \\ &+ \left\{ \frac{\partial U}{\partial \nu} - \Delta_1 \frac{\partial U_1}{\partial \nu} \right\} d\nu \end{aligned} \tag{3.5}$$

where we have eliminated the explicit dependence on t of $\{S_t\}_{t\geq 0}$ and $\{\nu_t\}_{t\geq 0}$ for clarity of results.

3.2.2 The Riskless Portfolio

To eliminate the risk set

$$\Delta_1 = \frac{\partial U}{\partial \nu} \Big/ \frac{\partial U_1}{\partial \nu} \tag{3.6a}$$

$$\Delta = \frac{\partial U}{\partial S} - \left(\frac{\partial U}{\partial \nu} \Big/ \frac{\partial U_1}{\partial \nu} \right) \frac{\partial U_1}{\partial S}. \tag{3.6b}$$

By no arbitrage the portfolio must earn the risk free rate. Hence

$$\begin{aligned} d\Pi &= \left\{ \frac{\partial U}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 U}{\partial S^2} + \rho\sigma\nu S \frac{\partial^2 U}{\partial \nu \partial S} + \frac{1}{2}\sigma^2 \nu \frac{\partial^2 U}{\partial \nu^2} \right\} dt \\ &\quad - \Delta_1 \left\{ \frac{dU_1}{dt} + \frac{1}{2}\nu S^2 \frac{\partial^2 U_1}{\partial S^2} + \rho\sigma\nu S \frac{\partial^2 U_1}{\partial \nu \partial S} + \frac{1}{2}\sigma^2 \nu \frac{\partial^2 U_1}{\partial \nu^2} \right\} dt \\ &= r\Pi dt \\ &= r\left(U - \Delta S - \Delta_1 U_1\right) dt \end{aligned} \tag{3.7}$$

Substituting the equations for Δ_1 and Δ given by the equations (3.6a) and (3.6b) into (3.7) we obtain

$$\begin{aligned} &\frac{\left(\frac{\partial U}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 U}{\partial S^2} + \rho\sigma\nu S \frac{\partial^2 U}{\partial \nu \partial S} + \frac{1}{2}\sigma^2 \nu \frac{\partial^2 U}{\partial \nu^2}\right) - rU + rS\frac{\partial U}{\partial S}}{\frac{\partial U}{\partial \nu}} \\ &= \frac{\left(\frac{\partial U_1}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 U_1}{\partial S^2} + \rho\sigma\nu S \frac{\partial^2 U_1}{\partial \nu \partial S} + \frac{1}{2}\sigma^2 \nu \frac{\partial^2 U_1}{\partial \nu^2}\right) - rU_1 + rS\frac{\partial U_1}{\partial S}}{\frac{\partial U_1}{\partial \nu}} \end{aligned} \tag{3.8}$$

Note that the left-hand side is a function of U only whilst the right-hand side is a function of U_1 alone. Therefore both sides must be equal a function f of the independent variables S, ν and t. Following Heston (1993) set

$$f(S, \nu, t) = -\kappa(\theta - \nu) + \Lambda(S, \nu, t)\sigma\sqrt{\nu}, \tag{3.9}$$

leaving us with

$$\begin{aligned} &\frac{\partial U}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 U}{\partial S^2} + \rho\sigma\nu S \frac{\partial^2 U}{\partial \nu \partial S} + \frac{1}{2}\sigma^2 \nu \frac{\partial^2 U}{\partial \nu^2} \\ &+ \left\{\kappa(\theta - \nu) - \Lambda(S, \nu, t)\sigma\sqrt{\nu}\right\} \frac{\partial U}{\partial \nu} - rU + rS\frac{\partial U}{\partial S} = 0, \end{aligned} \tag{3.10}$$

where $\Lambda(S, \nu, t)$ is the market price of volatility risk. The functional form of $\Lambda(S, \nu, t)$ is unknown. Heston motivates the form for the risk premium $\Lambda(S, \nu, t)\sigma\sqrt{\nu}$ by an application of a result by Cox-Ingersoll-Ross (1985) [10], where its shown that for a state variable with dynamics given by variance equation in (3.1), the market price of volatility risk is proportional to volatility (form can also be seen as an application of the equilibrium model posed in Breeden (1979) [5]). So we assume

$$\begin{aligned} \Lambda(S, \nu, t)\sigma\sqrt{\nu} &= \lambda\nu \text{ for some constant } \lambda \\ &= \lambda(S, \nu, t) \end{aligned} \tag{3.11}$$

We denote $\lambda(S, \nu, t)$ as the market price of volatility risk. This parameter needs to be estimated, however this is not an easy task so we will set it to zero throughout the dissertation, unless otherwise stated. Estimation of this parameter is subject to its own research (for example see Bollerslev et al. (2011) [4]).

3.2.3 Heston PDE in Terms of Log Spot

Transform variables by defining the log price $x = \ln(S)$ so we can express the Heston PDE (3.10) in terms of (x, ν, t). By the chain rule

$$\frac{\partial U}{\partial S} = \frac{1}{S}\frac{\partial U}{\partial x} \tag{3.12a}$$

$$\frac{\partial^2 U}{\partial \nu \partial S} = \frac{1}{S}\frac{\partial^2 U}{\partial \nu \partial x} \tag{3.12b}$$

and by the product rule

$$\frac{\partial^2 U}{\partial S^2} = \frac{\partial}{\partial S}\left(\frac{1}{S}\frac{\partial U}{\partial x}\right) = \frac{1}{S^2}\left(\frac{\partial^2 U}{\partial x^2} - \frac{\partial U}{\partial x}\right) \tag{3.13}$$

Substituting into (3.10) gives

$$\begin{aligned}\frac{\partial U}{\partial t} + \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \left(r - \frac{1}{2}\nu\right)\frac{\partial U}{\partial x} + \rho\sigma\nu\frac{\partial^2 U}{\partial \nu \partial x} + \frac{1}{2}\sigma^2\nu\frac{\partial^2 U}{\partial \nu^2} \\ + \{\kappa(\theta - \nu) - \lambda\nu\}\frac{\partial U}{\partial \nu} - rU = 0\end{aligned} \tag{3.14}$$

The Heston PDE is a two-dimensional diffusion-convection reaction equation with a mixed spatial-derivatives term. The convection property arises from the first-order derivative terms, whilst the diffusion arises from the second-order derivative terms. There is the presence of the present valuing term $-rU$ also, which has the effect of discounting value decreases everywhere. This equation is a member of a wider group of linear parabolic equations and such PDE often arise in financial models.

3.3 Risk-Neutral Pricing Approach

We make the important observation that the volatility $\sqrt{\nu_t}$ is not modelled directly rather it is modeled through the variance ν_t. The dynamics of ν_t arise through the Ornstein-Uhlenbeck process for the volatility $h_t = \sqrt{\nu_t}$ given by

$$dh_t = -\beta h_t dt + \delta dW_t^2 \tag{3.15}$$

Apply Itō's Lemma to $\nu_t = h_t^2$ to get

$$d\nu_t = (\delta^2 - 2\beta\nu_t)dt + 2\delta\sqrt{\nu_t}dW_t^2 \tag{3.16}$$

Defining $\kappa := 2\beta$, $\theta := \frac{\delta^2}{2\beta}$ and $\sigma := 2\delta$ we obtain the SDE for the variance given in (3.1).

The SDEs represented by the system (3.1) are given under the physical measure $\mathbb{P}$. For pricing we need the dynamics for the stock and variance processes $\{S_t\}_{t\geq 0}$ and $\{\nu_t\}_{t\geq 0}$ to be under the risk-neutral measure $\mathbb{Q}$. We can achieve a change of measure from the physical measure $\mathbb{P}$ to an equivalent martingale measure (EMM)

$\mathbb{Q}$ via Girsanov's theorem. In particular, the class of EMMs can be considered in terms of the Radon-Nikodym derivatives

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left\{-\left(\int_0^t q_s dW_s^1 + \int_0^t \Lambda(S,\nu,s)dW_s^2\right) - \frac{1}{2}\left(\int_0^t q_s^2 ds + \int_0^t \Lambda^2(S,\nu,s)ds\right)\right\} \tag{3.17}$$

By Girsanov we also have that

$$\begin{aligned} d\widetilde{W}_t^1 &= dW_t^1 + q_t dt \\ &= dW_t^1 + \frac{\mu - r}{\sqrt{\nu}}dt \qquad (3.18a) \\ d\widetilde{W}_t^2 &= dW_t^2 + \Lambda(S,\nu,t)dt \\ &= dW_t^2 + \frac{\lambda(S,\nu,t)}{\sigma\sqrt{\nu}}dt \qquad (3.18b) \end{aligned}$$

are $\mathbb{Q}$-Brownian motions.

Under $\mathbb{Q}$, equations in (3.1) become

$$dS_t = rS_t dt + \sqrt{\nu_t}S_t d\widetilde{W}_t^1 \tag{3.19a}$$

$$d\nu_t = \kappa^*(\theta^* - \nu_t)dt + \sigma\sqrt{\nu_t}d\widetilde{W}_t^2 \tag{3.19b}$$

where $\mathbb{E}^{\mathbb{Q}}[d\widetilde{W}_t^1 d\widetilde{W}_t^2] = \rho dt$ and $\kappa^* = \kappa + \lambda$ and $\theta^* = \frac{\kappa\theta}{\kappa+\lambda}$. If we set $\lambda = 0$ then $\kappa^* = \kappa$ and $\theta^* = \theta$ so that the parameters remain the same under physical and risk neutral measures.

3.4 The Heston Closed-Form Solution

3.4.1 The Call Option Price

In this section we would like to solve the Heston PDE (3.10) for a European call option subject to the terminal condition $C(S,\nu,T) = \max(S_T - K, 0)$ and the boundary conditions

$$C(0,\nu,t) = 0, \ (S = 0) \tag{3.20a}$$

$$\frac{\partial C}{\partial S}(\infty,\nu,t) = 1, \ (S = \infty) \tag{3.20b}$$

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} - rC + \kappa\theta\frac{\partial C}{\partial \nu} = 0, \ (\nu = 0) \tag{3.20c}$$

$$C(S,\infty,t) = S, \ (\nu = \infty) \tag{3.20d}$$

We follow the steps in Rouah (2013) [40]. The general risk-neutral approach is taken in which we assume the option price is given by the discounted expected future payoff under an equivalent martingale pricing measure. We know the payoff of a European call option is given by

$$C(S,T) = \max(S_T - K, 0),$$

so we can express the price of the option $C(S,t)$ at any time $t \in [0,T]$ as follows

$$\begin{aligned} C(S,t) &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\text{Payoff}(S)|\mathcal{F}_t] \\ &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\max(S_T-K,0)|\mathcal{F}_t] \\ &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S_T-K)\mathbb{1}_{S_T>K}|\mathcal{F}_t] \\ &= \underbrace{e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[S_T\mathbb{1}_{S_T>K}|\mathcal{F}_t]}_{(1)} - Ke^{-r(T-t)}\underbrace{\mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{S_T>K}|\mathcal{F}_t]}_{(2)} \\ &= S_tP_1 - Ke^{-r(T-t)}P_2 \end{aligned} \tag{3.21}$$

where P_1 and P_2 each represent the probability of the call expiring in-the-money (ITM) (but under different probability measures) conditional on S_t and ν_t. We set $S_t = e^{x_t}$. This means

$$P_j = \Pr(\ln S_T > \ln K|\mathcal{F}_t) \text{ for } j = 1,2. \tag{3.22}$$

To evaluate (2) we just note that this is simply the probability of the call expiring ITM under the risk-neutral measure $\mathbb{Q}$ thus

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{S_T>K}|\mathcal{F}_t] &= \mathbb{Q}(S_T > K|\mathcal{F}_t) \\ &= \mathbb{Q}(\ln S_T > \ln K|\mathcal{F}_t) \\ &= P_2. \end{aligned} \tag{3.23}$$

To evaluate (1) we must perform a change of measure from $\mathbb{Q}$ to $\mathbb{Q}^s$. Consider the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{Q}^s} = \frac{B_T}{B_t}\Big/\frac{S_T}{S_t} = \frac{\mathbb{E}^{\mathbb{Q}}[e^{x_T}]}{e^{x_T}} \tag{3.24}$$

where as usual $B_t = e^{\int_0^t r_s ds} \Rightarrow \frac{B_T}{B_t} = e^{\int_t^T r_s ds}$. Thus

$$\begin{aligned} e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[S_T\mathbb{1}_{S_T>K}|\mathcal{F}_t] &= S_t\mathbb{E}^{\mathbb{Q}}\left[\frac{S_T}{S_t}\Big/\frac{B_T}{B_t}\,\mathbb{1}_{S_T>K}|\mathcal{F}_t\right] \\ &= S_t\mathbb{E}^{\mathbb{Q}^s}\left[\frac{S_T}{S_t}\Big/\frac{B_T}{B_t}\,\mathbb{1}_{S_T>K}\frac{d\mathbb{Q}}{d\mathbb{Q}^s}|\mathcal{F}_t\right] \\ &= S_t\mathbb{E}^{\mathbb{Q}^s}[\mathbb{1}_{S_T>K}|\mathcal{F}_t] \\ &= S_t\mathbb{Q}^s(S_T > K|\mathcal{F}_t) \\ &= S_tP_1. \end{aligned} \tag{3.25}$$

Note that the pricing formula (3.21) is very similar to the Black Scholes equation, where the only difference is the probabilities P_1 and P_2. These probabilities are different and they arise as we model volatility as stochastic, which leads to different distributions.

3.4.2 The Forward Equation

Let the solution take the form given by equation (3.21). Express this solutions in terms of the log price as follows

$$C(e^x,t) = e^xP_1 - Ke^{-r(T-t)}P_2.$$

The call price $C(e^x, t)$ satisfies the Heston PDE (3.14) so we proceed to evaluate the necessary derivatives. We drop the subscripts for notational convenience. We have

$$\frac{\partial U}{\partial t} = e^x \frac{\partial P_1}{\partial t} - Ke^{-r(T-t)} \left[rP_2 + \frac{\partial P_2}{\partial t} \right] \tag{3.26a}$$

$$\frac{\partial U}{\partial x} = e^x \left[P_1 + \frac{\partial P_1}{\partial x} \right] - Ke^{-r(T-t)} \frac{\partial P_2}{\partial x} \tag{3.26b}$$

$$\frac{\partial^2 U}{\partial x^2} = e^x \left[P_1 + 2\frac{\partial P_1}{\partial x} + \frac{\partial^2 P_1}{\partial x^2} \right] - Ke^{-r(T-t)} \frac{\partial^2 P_2}{\partial x^2} \tag{3.26c}$$

$$\frac{\partial U}{\partial \nu} = e^x \frac{\partial P_1}{\partial \nu} - Ke^{-r(T-t)} \frac{\partial P_2}{\partial \nu} \tag{3.26d}$$

$$\frac{\partial^2 U}{\partial \nu^2} = e^x \frac{\partial^2 P_1}{\partial \nu^2} - Ke^{-r(T-t)} \frac{\partial^2 P_2}{\partial \nu^2} \tag{3.26e}$$

$$\frac{\partial^2 U}{\partial x \partial \nu} = e^x \left[\frac{\partial P_1}{\partial \nu} + \frac{\partial^2 P_1}{\partial x \partial \nu} \right] - Ke^{-r(T-t)} \frac{\partial^2 P_2}{\partial x \partial \nu} \tag{3.26f}$$

Substituting terms (3.26a)-(3.26f) into (3.14) gives two equations which may be combined to give the Fokker-Planck forward equation (also known as the forward Kolmogorov equation)

$$\begin{aligned} &\frac{\partial P_j}{\partial t} + \rho\sigma\nu \frac{\partial^2 P_j}{\partial \nu \partial x} + \frac{1}{2}\nu \frac{\partial^2 P_j}{\partial x^2} + \frac{1}{2}\sigma^2 \nu \frac{\partial^2 P_j}{\partial \nu^2} \\ &\qquad + (r + u_j \nu) \frac{\partial P_j}{\partial x} + (a - b_j \nu) \frac{\partial P_j}{\partial \nu} = 0 \end{aligned} \tag{3.27}$$

for $j = 1, 2$ where $u_1 = \frac{1}{2}$, $u_2 = -\frac{1}{2}$, $a = \kappa\theta$, $b_1 = \kappa + \lambda - \rho\sigma$ and $b_2 = \kappa + \lambda$. The probabilities P_j are subject to the terminal condition

$$P_j = \mathbb{1}_{x_T > \ln K}. \tag{3.28}$$

3.4.3 Obtaining the Characteristic Functions

There is no direct closed form solution to find the probabilities P_j. However, they may be recovered from characteristic functions $f_j(\phi; x, \nu)$ which do possess closed form solutions. They can be found via the Gil-Pelaez (1951) (see [22]) inversion formula

$$P_j = \Pr(\ln S_T > \ln K) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left[\frac{e^{-i\phi \ln K} f_j(\phi; x, \nu)}{i\phi} \right] d\phi \tag{3.29}$$

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

Following Heston (1993), assume that the characteristic functions for the log terminal spot price $x_T = \ln(S_T)$ take the log linear form

$$f_j(\phi; x_t, \nu_t) = \exp\left\{ C_j(T-t, \phi) + D_j(T-t, \phi)\nu_t + i\phi x_t \right\} \tag{3.30}$$

where C_j and D_j are coefficients. By the Feynman-Kač theorem, the characteristic functions (3.30) satisfy the forward equation (3.27). By the theorem f has a representation given by

$$f(x_t, \nu_t, t) = \mathbb{E}[f(x_T, \nu_T, T) | \mathcal{F}_t] = \mathbb{E}[e^{i\phi \ln S_T} | x_t, \nu_t]. \tag{3.31}$$

Equation (3.31) is the characteristic function for $x_T = \ln S_T$. Thus the PDE for the characteristic function is

$$\begin{aligned} -\frac{\partial f_j}{\partial \tau} + \rho\sigma\nu\frac{\partial^2 f_j}{\partial \nu \partial x} + \frac{1}{2}\nu\frac{\partial^2 f_j}{\partial x^2} + \frac{1}{2}\sigma^2\nu\frac{\partial^2 f_j}{\partial \nu^2} \\ + (r + u_j\nu)\frac{\partial f_j}{\partial x} + (a - b_j\nu)\frac{\partial f_j}{\partial \nu} = 0 \end{aligned} \tag{3.32}$$

where $\tau = T - t$ is the time to expiry. Next we evaluate the appropriate derivatives

$$\frac{\partial f_j}{\partial \tau} = \left(\frac{\partial C_j}{\partial \tau} + \frac{\partial D_j}{\partial \tau}\nu\right) f_j \tag{3.33a}$$

$$\frac{\partial f_j}{\partial x} = i\phi f_j \tag{3.33b}$$

$$\frac{\partial f_j}{\partial \nu} = D_j f_j \tag{3.33c}$$

$$\frac{\partial^2 f_j}{\partial x^2} = -\phi^2 f_j \tag{3.33d}$$

$$\frac{\partial^2 f_j}{\partial \nu^2} = D_j^2 f_j \tag{3.33e}$$

$$\frac{\partial^2 f_j}{\partial x \partial \nu} = i\phi D_j f_j \tag{3.33f}$$

Substitute (3.33a)-(3.33f) into (3.32) and we obtain a system of two ordinary differential equations (ODEs):

$$\begin{cases} \frac{\partial D_j}{\partial \tau} = \rho\sigma i\phi D_j - \frac{1}{2}\phi^2 + \frac{1}{2}\sigma^2 D_j^2 + i\phi u_j - b_j D_j, & (1) \\ \frac{\partial C_j}{\partial \tau} = ri\phi + aD_j & (2) \end{cases}$$

subject to

$$C_j(0,\phi) = 0,\ D_j(0,\phi) = 0.$$

Equation (1) is a Ricatti equation, while equation (2) is an ordinary derivative for C_j that may be solved by integrating once D_j has been obtained.

3.4.4 Solution to the Heston-Ricatti equation

General Solution to the Ricatti equation

The Ricatti equation for $y(t)$ is given by

$$\frac{dy(t)}{dt} = P(t) + Q(t)y(t) + R(t)y^2(t). \tag{3.34}$$

To solve (3.34), we note that any Ricatti type equation may be reduced to the second order linear ODE for $w(t)$ (see Al Bastami et al. [1])

$$\begin{aligned} w'' - \left[\frac{P'}{P} + Q\right] w' + PRw = 0 \\ \Leftrightarrow w'' + bw' + cw = 0 \end{aligned} \tag{3.35}$$

where $b = -\left[\frac{P'}{P} + Q\right]$, $c = PR$, by use of the substitution

$$y(t) = -\frac{w'(t)}{w(t)}\frac{1}{R(t)}. \tag{3.36}$$

To solve equation (3.35), consider the auxiliary equation $r^2 + br + c = 0$, which has roots

$$\alpha = \frac{-b + \sqrt{b^2 - 4ac}}{2} \tag{3.37a}$$

$$\beta = \frac{-b - \sqrt{b^2 - 4ac}}{2}. \tag{3.37b}$$

The general solution is given by

$$w(t) = Ae^{\alpha t} + Be^{\beta t} \tag{3.38}$$

where A,B are constants. Therefore the general solution to the Ricatti equation (3.34) is

$$y(t) = -\frac{A\alpha e^{\alpha t} + B\beta e^{\beta t}}{Ae^{\alpha t} + Be^{\beta t}}\frac{1}{R(t)}. \tag{3.39}$$

General solution to the Heston-Ricatti equation

We follow Rouah (2013) (see [40]). Rearrange equation (1) to give

$$\frac{\partial D_j}{\partial \tau} = \left[i\phi u_j - \frac{1}{2}\phi^2\right] - \left[b_j - \rho\sigma i\phi\right] D_j + \frac{1}{2}\sigma^2 D_j^2, \tag{3.40}$$

and then rewrite so that

$$\frac{\partial D_j}{\partial \tau} = P_j - Q_j D_j + R D_j^2, \tag{3.41}$$

where

$$P_j = i\phi u_j - \frac{1}{2}\phi^2$$
$$Q_j = b_j - \rho\sigma i\phi$$
$$R = \frac{1}{2}\sigma^2.$$

The corresponding second order linear ODE is

$$w'' + Q_j w' + P_j R w = 0, \tag{3.43}$$

where we have used the substitution

$$D_j = -\frac{1}{R}\frac{w'}{w}. \tag{3.44}$$

The auxiliary equation in this case is $r^2 + Q_j r + P_j r = 0$, which has roots

$$\alpha_j = \frac{-Q_j + \sqrt{Q_j^2 - 4P_j R}}{2} = \frac{-Q_j + d_j}{2} \tag{3.45a}$$

$$\beta_j = \frac{-Q_j - \sqrt{Q_j^2 - 4P_j R}}{2} = \frac{-Q_j - d_j}{2}, \tag{3.45b}$$

where

$$\begin{aligned} d_j = \alpha_j - \beta_j &= \sqrt{Q_j^2 - 4P_j R} \\ &= \sqrt{(\rho\sigma i\phi - b_j)^2 - 4(u_j i\phi - \frac{1}{2}\phi^2)\frac{1}{2}\sigma^2} \\ &= \sqrt{(\rho\sigma i\phi - b_j)^2 - \sigma^2(2i\phi u_j - \phi^2}. \end{aligned}$$

So the general solution is given by

$$D_j = -\frac{1}{R}\frac{w'}{w} = -\frac{1}{R}\left(\frac{A\alpha e^{\alpha\tau} + B\beta e^{\beta\tau}}{Ae^{\alpha\tau} + Be^{\beta\tau}}\right) = -\frac{1}{R}\left(\frac{K\alpha e^{\alpha\tau} + \beta e^{\alpha\tau}}{Ke^{\alpha\tau} + e^{\beta\tau}}\right), \tag{3.46}$$

where $K = \frac{A}{B}$. From the initial condition

$$D_j(0,\phi) = 0 \Rightarrow -\frac{1}{R}\left(\frac{K\alpha + \beta}{K+1}\right) = 0 \Rightarrow K = -\frac{\beta}{\alpha} \tag{3.47}$$

$$\begin{aligned} \Rightarrow D_j &= -\frac{\beta}{R}\left(\frac{-e^{\alpha\tau} + e^{\beta\tau}}{-g_j e^{\alpha\tau} + e^{\beta\tau}}\right) \\ &= -\frac{\beta}{R}\left(\frac{-e^{\alpha\tau}e^{-\beta\tau} + e^{\beta\tau}e^{-\beta\tau}}{-g_j e^{\alpha\tau}e^{-\beta\tau} + e^{\beta\tau}e^{-\beta\tau}}\right) \\ &= -\frac{\beta}{R}\left(\frac{1 - e^{d_j\tau}}{1 - g_j e^{d_j\tau}}\right) = \frac{Q_j + d_j}{2R}\left(\frac{1 - e^{d_j\tau}}{1 - g_j e^{d_j\tau}}\right). \end{aligned}$$

To see the last step, recall that $\beta_j = \frac{-Q_j - d_j}{2}$ so

$$-\frac{\beta_j}{R} = \frac{Q_j + d_j}{2R}. \tag{3.48}$$

We have set

$$g_j = -K = \frac{\beta}{\alpha} = \frac{-b_j + \rho\sigma i\phi - d_j}{-b_j + \rho\sigma i\phi + d_j} = \frac{b_j - \rho\sigma i\phi + d_j}{b_j - \rho\sigma i\phi - d_j} = \frac{Q_j + d_j}{Q_j - d_j} \tag{3.49}$$

$$\Rightarrow D_j(\tau,\phi) = \frac{b_j - \rho\sigma i\phi + d_j}{\sigma^2}\left(\frac{1 - e^{d_j\tau}}{1 - g_j e^{d_j\tau}}\right). \tag{3.50}$$

To find C_j integrate between 0 and τ and we have

$$C_j = \underbrace{\int_0^\tau ri\phi\, dy}_{(1)} + a\left(\frac{Q_j + d_j}{\sigma^2}\right)\underbrace{\int_0^\tau \left(\frac{1 - e^{d_j y}}{1 - g_j e^{d_j y}}\right) dy}_{(2)}. \tag{3.51}$$

Integral (1) can be found straightaway. For integral (2) use a substitution $x = e^{d_j y}$ then $dx = d_j e^{d_j y} dy \Rightarrow dy = \frac{dx}{x d_j}$. To obtain the limits of integration note when $y = 0$, $x = 1$ and when $y = \tau$, $x = e^{d_j \tau}$ so we have

$$C_j = \; ri\phi\tau + \frac{a}{d_j}\left(\frac{Q_j + d_j}{\sigma^2}\right)\int_1^{e^{d_j\tau}} \left(\frac{1-x}{1-g_j x}\right)\frac{1}{x}\,dx + k_1, \tag{3.52}$$

where k_1 is an arbitrary constant of integration. By partial fraction decomposition

$$\frac{1}{x}\left(\frac{1-x}{1-g_j x}\right) = \frac{1}{x} - \frac{1-g_j}{1-g_j x}. \tag{3.53}$$

So (2) becomes

$$\begin{aligned}
\int_1^{e^{d_j\tau}} \left(\frac{1-x}{1-g_j x}\right)\frac{1}{x}\,dx = & \int_1^{e^{d_j\tau}} \left[\frac{1}{x} - \frac{1-g_j}{1-g_j x}\right] dx \\
= & \left[\ln x + \frac{1-g_j}{g_j}\ln(1-g_j x)\right]_{x=1}^{x=e^{d_j\tau}} \\
= & \left[d_j\tau + \frac{1-g_j}{g_j}\ln\left(\frac{1-g_j e^{d_j\tau}}{1-g_j}\right)\right].
\end{aligned}$$

By the initial condition $C_j(0,\phi) = 0$ we have $k_1 = 0$. Therefore we obtain

$$C_j = \; ri\phi\tau + \frac{a}{\sigma^2}\left[(b_j - \rho\sigma i\phi + d_j)\tau - 2\ln\left(\frac{1-g_j e^{d_j\tau}}{1-g_j}\right)\right]. \tag{3.54}$$

In summary the closed-form solution to the Heston model is

$$f_j(\phi; x_t, \nu_t) = exp\left\{C_j(T-t,\phi) + D_j(T-t,\phi)\nu_t + i\phi x_t\right\} \tag{3.55a}$$

$$C_j(T-t,\phi) = ri\phi(T-t) + \frac{a}{\sigma^2}\left[(b_j - \rho\sigma i\phi + d_j)(T-t) - 2\ln\left(\frac{1-g_j e^{d_j(T-t)}}{1-g_j}\right)\right] \tag{3.55b}$$

$$D_j(T-t,\phi) = \frac{b_j - \rho\sigma i\phi + d_j}{\sigma^2}\left(\frac{1-e^{d_j(T-t)}}{1-g_j e^{d_j(T-t)}}\right) \tag{3.55c}$$

$$g_j = \frac{b_j - \rho\sigma i\phi + d_j}{b_j - \rho\sigma i\phi - d_j} \tag{3.55d}$$

$$d_j = \sqrt{(\rho\sigma i\phi)^2 - \sigma^2(2i\phi u_j - \phi^2)} \tag{3.55e}$$

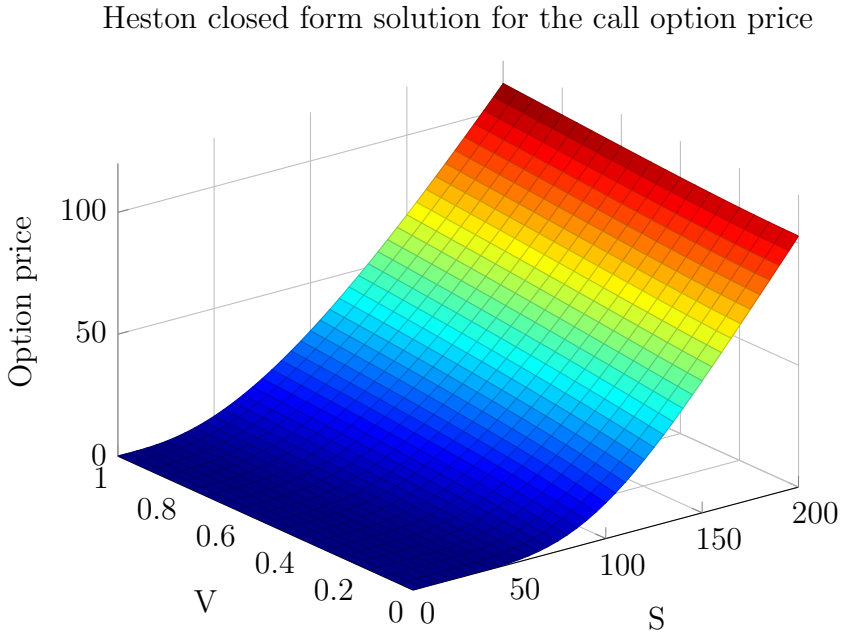


Figure 3.1: European call option analytical solution price grid in the uniform case with parameter values $\kappa = 2$, $\theta = 0.2$, $\sigma = 0.3$, $\rho = 0.8$, $r = 0.03$, $S = 200$, $V = 1$, $T = 1$, $K = 100$, $N_S = 40$ and $N_V = 20$

3.5 Pricing via Fourier Transforms

In this section, the following definitions are adopted.

Definition 3.5.1 (Fourier and Inverse Fourier Transforms)**.** The Fourier transform $\mathcal{F}\{\cdot\}$ and inverse Fourier Transform $\mathcal{F}^{-1}\{\cdot\}$ of a function $f(x)$ are given by, respectively,

$$\mathcal{F}\{f(x)\} \equiv \hat{f}(\xi) = \int_{-\infty}^{\infty} e^{i\xi x} f(x)\, dx \tag{3.56a}$$

$$\mathcal{F}^{-1}\left\{\hat{f}(\xi)\right\} \equiv f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\xi x} \hat{f}(\xi)\, d\xi. \tag{3.56b}$$

□

One important property of the Fourier transform is differentiation. If $0 < n \in \mathbb{Z}$, f^n is piecewise continuously differentiable and each derivative is integrable then we can apply integration by parts iteratively to obtain the Fourier transform of the *n-th* derivative

$$\mathcal{F}\left\{\frac{d^n f(x)}{dx^n}\right\} = (-i\xi)^n \hat{f}(\xi). \tag{3.57}$$

Another useful result is the derivative of the Fourier transform, obtained through

differentiation under the integral sign

$$\frac{d\hat{f}(\xi)}{d\xi} = \int_{-\infty}^{\infty} \frac{d}{d\xi}(e^{i\xi x} f(x))\, dx = \int_{-\infty}^{\infty} xie^{i\xi x} f(x)\, dx = i\hat{g}(x), \tag{3.58}$$

where $g(x) = xf(x)$.

Note further, if we let $f(x)$ represent the density of some random variable X then the Fourier transform is the characteristic function of $f(x)$, i.e.

$$\hat{f}(\xi) = \mathbb{E}\left[e^{i\xi x}\right]. \tag{3.59}$$

This is easily seen by the definition of the expectation operator.

3.5.1 Carr and Madan Approach

Carr and Madan (1999) present an approach to analytically determine the option price by the Fast Fourier tansform (FFT) in terms of the characteristic function and risk-neutral density. The advantages of the approach lie in the reduced computation time and the fast decay of the integrand (faster than that of Heston (1993)). We will see that the FFT cannot be applied directly due to a singular point at the origin. Instead, the method requires us to redefine the call price to incorporate a damping factor. Once we obtain the Fourier transform of the modified call price, we invert it and finally we recover the actual call price by eliminating the damping factor.

Define $x_T := \ln(S_T)$ and $k := \ln(K)$, where K is the strike price. Let $q(x)$ denote the risk-neutral density of the log price x_T. The European call price as a function of k for an option with maturity T is given by (recall (3.21))

$$\begin{aligned} C_T(k) &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\max(S_T - K, 0)\right] \\ &= e^{-r(T-t)} \int_k^{\infty} \left(e^x - e^k\right) q(x)\, dx \\ &= e^{x_t}\mathbb{Q}^s(S_T > e^k) - Ke^{-r(T-t)}\mathbb{Q}(S_T > e^k). \end{aligned} \tag{3.60}$$

In order to find $\mathbb{Q}^s(S_T > e^k)$ and $\mathbb{Q}(S_T > e^k)$ we must invert the characteristic function (recall Gil-Pelaez formula (3.29)). This requires integrability of the characteristic function. Unfortunately, $C_T(k) \nrightarrow 0$ as $k \to -\infty$

$$\begin{aligned} \lim_{k\to-\infty} C_T(k) &= \lim_{k\to-\infty} e^{-r(T-t)} \int_k^{\infty} \left(e^x - e^k\right) q(x)\, dx \\ &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[e^x\right] \\ &= S_t. \end{aligned} \tag{3.61}$$

So the limit does not converge to zero. This means that $C_T(k)$ is not integrable[1] L^1 and therefore its Fourier transform and inverse do not exist. Carr and Madan

[1] A function $f(x)$ is integrable if $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$. This condition must hold for the Fourier transform and its inverse to exist. The space of all integrable functions is L^1 or $L^1(\mathbb{R})$.

(1999) resolve this by introducing a damping factor $e^{\alpha k}$ with $\alpha > 0$ to the call price $C_T(k)$

$$c_T(k) = e^{\alpha k} C_T(k). \tag{3.62}$$

The modified call price (3.62) is an integrable function since

$$\begin{aligned}
\lim_{k\to-\infty} c_T(k) &= \lim_{k\to-\infty} e^{\alpha k} C_T(k) = \lim_{k\to-\infty} e^{-r(T-t)} \int_k^\infty \left(e^{\alpha k+x} - e^{(\alpha+1)k}\right) q(x)\, dx \\
&= \lim_{k\to-\infty} e^{-r(T-t)} \int_k^\infty e^{\alpha k+x} q(x)\, dx - \lim_{k\to-\infty} e^{-r(T-t)+(\alpha+1)k} \int_k^\infty q(x)\, dx \\
&= 0,
\end{aligned} \tag{3.63}$$

such that $c_T(k)$ is integrable L^1. We may thus find the Fourier transform of $c_T(k)$. It is given by

$$\begin{aligned}
\hat{c}_T(\xi) &= \int_{-\infty}^\infty e^{i\xi k} c_T(k)\, dk = \int_{-\infty}^\infty e^{i\xi k} e^{\alpha k} C_T(k)\, dk \\
&= e^{-r(T-t)} \int_{-\infty}^\infty e^{(\alpha+i\xi)k} \left[\int_k^\infty (e^x - e^k) q(x)\right] dk \\
&= e^{-r(T-t)} \int_{-\infty}^\infty q(x) \left[\int_{-\infty}^x \left(e^{(\alpha+i\xi)k+x} - e^{(\alpha+i\xi+1)k}\right)\, dk\right] dx \\
&= e^{-r(T-t)} \int_{-\infty}^\infty q(x) \left[\left(\frac{e^{(\alpha+i\xi)k+x}}{\alpha+i\xi} - \frac{e^{(\alpha+i\xi+1)k}}{\alpha+i\xi+1}\right)_{k=-\infty}^{k=x}\right] dx \\
&= e^{-r(T-t)} \int_{-\infty}^\infty q(x) \left[\frac{e^{(\alpha+i\xi+1)x}}{\alpha^2+\alpha-\xi^2+i(2\alpha+1)\xi}\right] dx \\
&= \frac{e^{-r(T-t)}\psi(\xi-(\alpha+1)i)}{\alpha^2+\alpha-\xi^2+i(2\alpha+1)\xi},
\end{aligned} \tag{3.64}$$

where we have used the fact that having $-\infty < k < \infty$ and $k < x < \infty$ is equivalent to $-\infty < x < \infty$ and $-\infty < k < x$. The call value can be found via the inverse Fourier transform of $c_T(k)$,

$$\begin{aligned}
C_T(k) &= e^{-\alpha k} c_T(k) \\
&= e^{-\alpha k} \frac{1}{2\pi} \int_{-\infty}^\infty e^{-i\xi k} \hat{c}_T(\xi)\, d\xi \\
&= e^{-\alpha k} \frac{1}{\pi} \int_0^\infty e^{-i\xi k} \hat{c}_T(\xi)\, d\xi.
\end{aligned}$$

The last line holds because $C_T(k)$ is real so we need only to consider the real part of the complex integrand $\hat{c}_T(k)$, and the real part is even valued.

The value for a European put option is found similarly.

3.5.2 Lewis' Approach

Generalized Fourier transform The generalized Fourier transform extends our previous discussions to include the complex plane. So far, we have just looked

at the characteristic function defined for real valued transform variables ξ. It is sometimes necessary to extend the contour of integration to that parallel to the real axis, so we now look at transform variables $\xi \in \mathbb{C}$. It is necessary to restrict the plane to where the characteristic function is well defined. It turns out that the expectation in (3.59) is well defined, for a set of values $\xi = \xi_r + i\xi_i$, within a strip of regularity $\alpha < \xi_i < \beta$ parallel to the real z-axis (see Schmelzle (2010) [41]). The inverse of the generalized Fourier transform $\hat{f}(\xi, t)$ is given by

$$f(x,t) = \frac{1}{2\pi} \int_{i\xi_i - \infty}^{i\xi_i + \infty} e^{-i\xi x} \hat{f}(\xi, t)\, d\xi. \tag{3.65}$$

The key idea behind the approach taken by Lewis (2000) (see [30] and [31]) is that of expressing the option price as the convolution of generalized Fourier transforms followed by an application of the Plancherel-Parseval identity. This approach may be applied to many payoff functions.

In this approach, it is not necessary to take the Fourier transform of the option price as in Carr and Madan (1999) [6], only the Fourier transform of the payoff function (call it the payoff transform, see Rouah (2013) [40]). In addition, we require the fundamental transform. Whilst the payoff transform is model independent and independent of the fundamental transform, the fundamental transform is model dependent. We firstly compute the generalized Fourier transform $\hat{f}(\xi, t)$ of a derivative $f(x, t)$ with expiry T

$$\hat{f}(\xi, t) = \int_{-\infty}^{\infty} e^{i\xi x} f(x,t)\, dx. \tag{3.66}$$

Consider the following example of a European call, for which the payoff function is known explicitly.

Example 3.5.1. Recall the payoff of a European call option is given by $C(x,T) = \max(S_T - K, 0)$ such that the payoff transform is

$$\begin{aligned}
\mathcal{F}\{C(x_T,T)\} \equiv \hat{C}(\xi,T) &= \int_{-\infty}^{\infty} e^{i\xi x} \max(e^x - K, 0)\, dx \\
&= \int_{\ln(K)}^{\infty} e^{(i\xi+1)x}\, dx - K \int_{\ln(K)}^{\infty} e^{i\xi x}\, dx \\
&= \left[\frac{1}{i\xi+1} e^{(i\xi+1)x} \right]_{x=\ln(K)}^{x=\infty} - \left[\frac{K}{i\xi} e^{i\xi x} \right]_{x=\ln(K)}^{x=\infty} \\
&= \left(\underbrace{\lim_{x\to\infty} \left[\frac{1}{i\xi+1} e^{(i\xi+1)x} \right]}_{(1)} - \frac{1}{i\xi+1} e^{(i\xi+1)\ln(K)} \right) \\
&\quad - \left(\underbrace{\lim_{x\to\infty} \left[\frac{K}{i\xi} e^{i\xi x} \right]}_{(2)} - \frac{K}{i\xi} e^{i\xi \ln(K)} \right).
\end{aligned} \tag{3.67}$$

We require

$$(1) = \lim_{x\to\infty}\left[\frac{1}{i\xi+1}e^{(i\xi+1)x}\right] = 0 \tag{3.68a}$$

$$(2) = \lim_{x\to\infty}\left[\frac{K}{i\xi}e^{i\xi x}\right] = 0 \tag{3.68b}$$

Express $\xi = \xi_r + i\xi_i$ where ξ_r and ξ_i denote the real and imaginary parts of ξ respectively. In order for equation (3.68a) to hold, we need $\xi_i > 1$. To see why this is true, rewrite the numerator in terms of real and imaginary parts, i.e. $e^{(i\xi+1)x} = e^{(i(\xi_r+i\xi_i)+1)x} = e^{x(i\xi_r-\xi_i+1)}$. Disregarding the real part of ξ, in order for the limit in (1) to tend to zero as $x \to \infty$ we must have

$$-\xi_i + 1 < 0 \Rightarrow \xi_i > 1. \tag{3.69}$$

Similarly, in order for equation (3.68b) we require

$$-\xi_i < 0 \Rightarrow \xi_i > 0. \tag{3.70}$$

We then have

$$\begin{aligned}\hat{C}(\xi,T) &= -\frac{1}{i\xi+1}e^{(i\xi+1)\ln(K)} + \frac{K}{i\xi}e^{i\xi\ln(K)} \\ &= \frac{-i\xi K^{i\xi+1} + (i\xi+1)K^{i\xi+1}}{(i\xi+1)(i\xi)} \\ &= -\frac{K^{i\xi+1}}{\xi^2 - i\xi},\end{aligned} \tag{3.71}$$

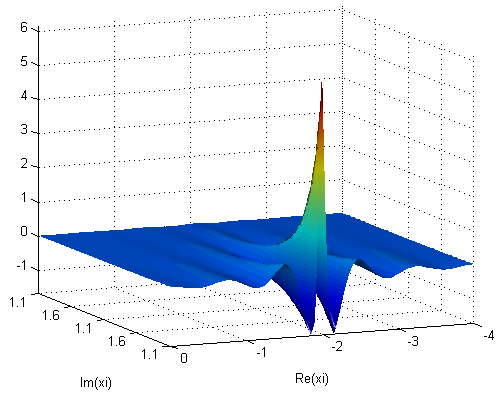
subject to conditions (3.69) and (3.70), which combined imply

$$\xi_i > 1. \tag{3.72}$$

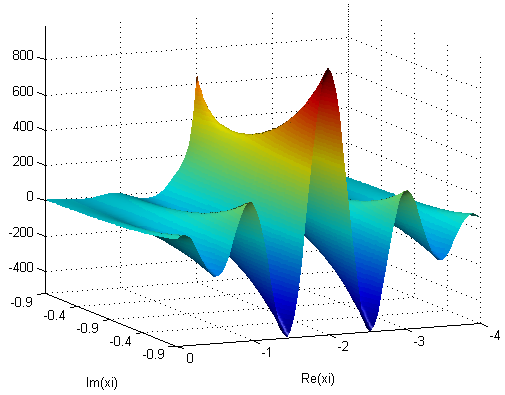
□

It is straightforward to show that the value of a European put option has the same functional form as the call except it is well defined in a different strip of the complex plane $\xi_i < 0$. We can find payoff transforms for other derivatives, for example see Table 2.1 of Lewis (2000) [30] (also Table 1 of Schmelzle (2010) [41] and Table 1 of Cartea (2013) [7]), in particular observe the different restrictions on the complex variable ξ to guarantee existence of the transform.

Figure 3.2a is a plot of the real part of a European call option. The axes are restricted to the strip of regularity for which the Fourier transform exists (recall we need $\xi_i > 1$), set the range of integration following the original presentation in Schmelzle (2010) [41] to $1 < \xi_i < 2$ and $-4 < \xi_r < 4$.



(a) Strip of regularity for European call options $\xi_i > 1$.



(b) Strip of regularity for European put options $\xi_i < 1$.

Similarly, Figure 3.2b is a plot of the real part of a European put option. We plot the real part of the payoff function over the range $-1 < \xi_i < 0$ and $-4 < \xi_r < 4$.

After evaluating the payoff transform, we need to find the fundamental transform. Recall that in the approach taken by Carr and Madan (1999), it is necessary to evaluate the Fourier transform of the option price. This means that we need a separate Fourier transform for each type of derivative for which we have a price. The advantage of the approach taken by Lewis (2000) is that the fundamental transform is model dependent and independent of a particular payoff, so once obtained we need only evaluate the transform of the derivative payoffs. The same fundamental transform can be applied to price different European options, the only requirement is that the payoff transform be obtainable analytically. Payoff transforms are more easily obtained than derivative price transforms, since they are trade data (i.e. specified within a contract) and therefore known explicitly.

We proceed to find the fundamental transform. Our attention is restricted to the Heston model wherein lies our interest (for a general derivation see Chapter 2 of [30]). Recall the Heston PDE in equation (3.14) and consider a continuous dividend yield q

$$\begin{aligned} -\frac{\partial U}{\partial t} = & -rU + \left(r - q - \frac{1}{2}\nu\right)\frac{\partial U}{\partial x} + \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} \\ & + \{\kappa(\theta - \nu) - \lambda\nu\}\frac{\partial U}{\partial \nu} + \frac{1}{2}\sigma^2\nu\frac{\partial^2 U}{\partial \nu^2} + \rho\sigma\nu\frac{\partial^2 U}{\partial \nu \partial x}. \end{aligned}$$

Next, translate the PDE (3.73) into a PDE in terms of the Fourier transform

$\hat{U}(\xi,\nu,t)$. This can be easily done using earlier results.

$$\begin{aligned}
\mathcal{F}\left\{-\frac{\partial U}{\partial t}\right\} &= \mathcal{F}\left\{-rU + \left(r-q-\frac{1}{2}\nu\right)\frac{\partial U}{\partial x} + \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \{\kappa(\theta-\nu)-\lambda\nu\}\frac{\partial U}{\partial \nu}\right. \\
&\quad \left. + \frac{1}{2}\sigma^2\nu\frac{\partial^2 U}{\partial \nu^2} + \rho\sigma\nu\frac{\partial^2 U}{\partial \nu \partial x}\right\} \\
&= \int_{-\infty}^{\infty} e^{i\xi x}\left(-rU + \left(r-q-\frac{1}{2}\nu\right)\frac{\partial U}{\partial x} + \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \{\kappa(\theta-\nu)-\lambda\nu\}\frac{\partial U}{\partial \nu}\right. \\
&\quad \left. + \frac{1}{2}\sigma^2\nu\frac{\partial^2 U}{\partial \nu^2} + \rho\sigma\nu\frac{\partial^2 U}{\partial \nu \partial x}\right) dx \\
&= -r\int_{-\infty}^{\infty} e^{i\xi x}U\,dx + \left(r-q-\frac{1}{2}\nu\right)\int_{-\infty}^{\infty} e^{i\xi x}\frac{\partial U}{\partial x}\,dx + \frac{1}{2}\nu\int_{-\infty}^{\infty} e^{i\xi x}\frac{\partial^2 U}{\partial x^2}\,dx \\
&\quad + \{\kappa(\theta-\nu)-\lambda\nu\}\int_{-\infty}^{\infty} e^{i\xi x}\frac{\partial U}{\partial \nu} + \,dx + \frac{1}{2}\sigma^2\nu\int_{-\infty}^{\infty} e^{i\xi x}\frac{\partial^2 U}{\partial \nu^2}\,dx \\
&\quad + \rho\sigma\nu\int_{-\infty}^{\infty} e^{i\xi x}\frac{\partial^2 U}{\partial \nu \partial x}\,dx \\
\Rightarrow \quad -\frac{\partial \hat{U}}{\partial t} &= -r\hat{U} + \left(r-q-\frac{1}{2}\nu\right)(-i\xi)\hat{U} + \frac{1}{2}\nu(i\xi)^2\hat{U} \\
&\quad + \{\kappa(\theta-\nu)-\lambda\nu\}\frac{\partial \hat{U}}{\partial \nu} + \frac{1}{2}\sigma^2\nu\frac{\partial^2 \hat{U}}{\partial \nu^2} + \rho\sigma\nu(-i\xi)\frac{\partial \hat{U}}{\partial \nu}.
\end{aligned} \tag{3.73}$$

Rearranging gives

$$\begin{aligned}
-\frac{\partial \hat{U}}{\partial t} &= \left(-r - i\xi(r-q)\right)\hat{U} - \frac{1}{2}\nu\left(\xi^2 - i\xi\right)\hat{U} \\
&\quad + \{\kappa(\theta-\nu)-\lambda\nu - i\xi\rho\sigma\nu\}\frac{\partial \hat{U}}{\partial \nu} + \frac{1}{2}\sigma^2\nu\frac{\partial^2 \hat{U}}{\partial \nu^2}
\end{aligned}$$

We have used the earlier Fourier differentiation result (3.57). In particular, it's important to realise that we assume boundary terms present during the integration by parts procedure may be neglected. Lewis (2000) explains that typically, within the strip of regularity $\alpha < \xi_i < \beta$ for which the transform is well defined, the boundary terms vanish. He dedicates a paragraph explaining why this is in fact the case.

Following Lewis (2000), define $\tau = T - t$ and write the generalized Fourier transform as follows

$$\hat{U}(\xi,\nu,t) = \exp\left\{(r - i\xi(r-q))\tau\right\}\hat{h}(\xi,\nu,\tau), \tag{3.74}$$

for some function $\hat{h}$. Evaluate the necessary derivatives

$$-\frac{\partial \hat{U}}{\partial t} = \frac{\partial \hat{U}}{\partial \tau} = \{-r - i\xi(r-q)\}\exp\left\{(r - i\xi(r-q))\tau\right\}\hat{h}(\xi,\nu,\tau) \tag{3.75a}$$

$$+ \exp\left\{(r - i\xi(r-q))\tau\right\}\frac{\partial \hat{h}}{\partial \tau}(\xi,\nu,\tau) \tag{3.75b}$$

$$\frac{\partial \hat{U}}{\partial \nu} = \exp\left\{(r - i\xi(r-q))\tau\right\}\frac{\partial \hat{h}}{\partial \nu}(\xi,\nu,\tau) \tag{3.75c}$$

$$\frac{\partial^2 \hat{U}}{\partial \nu^2} = \exp\left\{(r - i\xi(r-q))\tau\right\}\frac{\partial^2 \hat{h}}{\partial \nu^2}(\xi,\nu,\tau). \tag{3.75d}$$

Substituting the above derivatives into the PDE (3.73) gives (cancel common factors)

$$\frac{\partial \hat{h}}{\partial \tau} = \frac{1}{2}\sigma^2\nu\frac{\partial^2 \hat{h}}{\partial \nu^2} + \{\kappa(\theta-\nu) - \lambda\nu - i\xi\rho\sigma\nu\}\frac{\partial \hat{h}}{\partial \nu} - c(\xi)\nu\hat{h}, \tag{3.76}$$

where $c(\xi) = \frac{1}{2}(\xi^2 - i\xi)$.

The payoff function does not depend upon the variance ν. Lewis explains that this is because we have restricted our theory to volatility independent payoffs. He argues that due to this restriction it is sufficient to consider the case $\hat{h}(\xi, \nu, \tau = 0) = 1$, i.e. a payoff of 1. Any other payoff may be dealt with by multiplying the solution for the case in which the payoff is 1, by the corresponding payoff transform. A solution $\hat{h}(\xi, \nu, \tau)$ to the PDE (3.73) satisfying the initial condition $\hat{h}(\xi, \nu, \tau = 0) = 1$ is called a fundamental transform. The PDE (3.73) is regular if there exists a fundamental transform regular within the strip $\alpha < \xi_i < \beta$ where $\xi_i = Im\xi$ and α,$\beta \in \mathbb{R}$.

After having found the fundamental transform, the recipe to evaluate option prices given by Lewis (2000) is as follows

1. Multiply the fundamental transform by the payoff transform.
2. Further multiply by the factor $\exp\{(r - i\xi(r-q))\tau\}$.
3. Evaluate the inverse transform of the result maintaining ξ_i within the strip of regularity for which the transform is well defined.

We proceed to evaluate the fundamental transform for the Heston model. Let $\lambda = 0$ then

$$\frac{\partial \hat{h}}{\partial \tau} = \frac{1}{2}\sigma^2\nu\frac{\partial^2 \hat{h}}{\partial \nu^2} + \{\kappa(\theta-\nu) - i\xi\rho\sigma\nu\}\frac{\partial \hat{h}}{\partial \nu} - c(\xi)\nu\hat{h} \tag{3.77}$$

We now follow the transformation adopted by Lewis (2000) and adapted version for the Heston model presented by Rouah (2013). Let $t = \frac{\sigma^2\tau}{2}$ then

$$\begin{aligned}\frac{\partial \hat{h}}{\partial t} &= \nu\frac{\partial^2 \hat{h}}{\partial \nu^2} + \frac{2}{\sigma^2}\{\kappa(\theta-\nu) - i\xi\rho\sigma\nu\}\frac{\partial \hat{h}}{\partial \nu} - \frac{2}{\sigma^2}c(\xi)\nu\hat{h} \\ &= \nu\frac{\partial^2 \hat{h}}{\partial \nu^2} + \left\{\tilde{\kappa}(\tilde{\theta}-\nu)\right\}\frac{\partial \hat{h}}{\partial \nu} - \tilde{c}(\xi)\nu\hat{h},\end{aligned} \tag{3.78}$$

where

$$\tilde{\kappa} = \frac{2(\kappa + i\xi\rho\sigma)}{\sigma^2} \tag{3.79a}$$

$$\tilde{\theta} = \frac{\kappa\theta}{\kappa + i\xi\rho\sigma} \tag{3.79b}$$

$$\tilde{c}(\xi) = \frac{2c(\xi)}{\sigma^2} = \frac{\xi^2 - i\xi}{\sigma^2}. \tag{3.79c}$$

Equation (3.78) is a parabolic equation, thus the solution takes the form

$$f(\nu, t) = e^{C(t) + D(t)\nu}, \tag{3.80}$$

subject to the initial condition $C(0) = D(0) = 0$. Taking the corresponding derivatives yields

$$\frac{\partial f}{\partial t} = \left(\frac{\partial C}{\partial t} + \frac{\partial D}{\partial t}\right) f \tag{3.81a}$$

$$\frac{\partial f}{\partial \nu} = D(t) f \tag{3.81b}$$

$$\frac{\partial^2 f}{\partial \nu^2} = D^2(t) f. \tag{3.81c}$$

So

$$\left[-\frac{\partial C}{\partial t} + \tilde{\kappa}\tilde{\theta} D\right] + \left[-\frac{\partial D}{\partial t} + D^2 - \tilde{\kappa} D - \tilde{c}\right] \nu = 0. \tag{3.82}$$

We must therefore solve the set of ODEs

$$\begin{cases} \frac{\partial C}{\partial t} = \tilde{\kappa}\tilde{\theta} D, & (1) \\ \frac{\partial D}{\partial t} = D^2 - \tilde{\kappa} D - \tilde{c} & (2) \end{cases}$$

Note that equation (2) is in the form of a Ricatti equation which we have already solved. Thus, proceeding as before the corresponding second order ODE is $w'' + \tilde{\kappa} w' - \tilde{c} w = 0$ and the auxiliary equation is therefore $r^2 + \tilde{\kappa} r - \tilde{c} = 0$. So the roots are given by

$$\alpha = \frac{-\tilde{\kappa} + d}{2} \tag{3.83a}$$

$$\beta = \frac{-\tilde{\kappa} - d}{2} \tag{3.83b}$$

$$d = \sqrt{\tilde{\kappa}^2 + 4\tilde{c}}. \tag{3.83c}$$

From the initial conditions we have $K = -\frac{\alpha}{\beta} = \frac{\tilde{\kappa}+d}{-\tilde{\kappa}+d}$. So

$$D(t) = \frac{\tilde{\kappa} + d}{2}\left(\frac{1 - e^{dt}}{1 - g e^{dt}}\right) \tag{3.84a}$$

$$g = -K \tag{3.84b}$$

$$d = \alpha - \beta. \tag{3.84c}$$

Finally $C(t)$ may be found by integrating the expression for $D(t)$ above

$$\begin{aligned} \int_0^t D(z)dz &= \frac{\tilde{\kappa} + d}{2} \int_0^t \left(\frac{1 - e^{dz}}{1 - g e^{dz}}\right) dz \\ &= \frac{\tilde{\kappa} + d}{2} t - \ln\left(\frac{1 - g e^{dt}}{1 - g}\right) + k_1, \end{aligned} \tag{3.85}$$

where k_1 is an arbitrary constant of integration. From initial conditions $k_1 = 0$. Hence,

$$C(t) = \tilde{\kappa}\tilde{\theta}\left[\frac{\tilde{\kappa} + d}{2} t - \ln\left(\frac{1 - g e^{dt}}{1 - g}\right)\right] \tag{3.86}$$

and we have found the solution to our parabolic equation (3.78).

Chapter 4

Finite Difference Schemes

PDEs are ubiquitous in almost all applied mathematics. Whilst the Heston model admits a closed form solution for branches of European options, in practice most equations do not. For this reason we turn to numerical techniques. Among the vast approaches taken by many authors, one of the most common and perhaps most intuitive one is the finite difference method. In such methods, we use a truncated Taylor series in order to approximate the derivatives appearing in the PDE of interest. The discretized equation is then reduced to some algebraic system which may be solved via techniques known to us from numerical linear algebra. We describe next the finite difference method for approximating the solution to the initial boundary value problem (IBVP) given by the Heston PDE (3.10) subject to a mixture of Dirichlet and Neumann boundary conditions given by equations (3.20a)-(3.20d) for a European call option.

4.1 The IBVP Set-up

Recall from section 3.4.1 that the Heston PDE is subject to the terminal condition $U(S,\nu,T) = \max(S_T - K, 0)$ for a European call. If we are to price a European put then the terminal condition is $U(S,\nu,T) = \max(K - S_T, 0)$. Therefore, as it stands, the PDE given by equation (3.10) is backward marching in time. In order to transform the problem into an IBVP in which we can step forwards in time, we use the change of variable $\tau = T - t$ and note that

$$\frac{\partial U}{\partial t} = \frac{\partial U}{\partial \tau}\frac{\partial \tau}{\partial t} = -\frac{\partial U}{\partial \tau}. \tag{4.1}$$

Our Heston PDE becomes

$$\begin{aligned}\frac{\partial U}{\partial \tau} = \ & \frac{1}{2}\nu S^2\frac{\partial^2 U}{\partial S^2} + \rho\sigma\nu S\frac{\partial^2 U}{\partial \nu \partial S} + \frac{1}{2}\sigma^2\nu\frac{\partial^2 U}{\partial \nu^2} \\ & + \{\kappa(\theta - \nu)\}\frac{\partial U}{\partial \nu} - rU + rS\frac{\partial U}{\partial S},\end{aligned} \tag{4.2}$$

where, as before, we have assumed there is no market price of volatility risk ($\lambda = 0$). Note the terminal condition, $U(S,\nu,T) = \max(S_T - K, 0)$ for a European call, has now become an *initial* condition. We will use t to represent the time left to maturity τ so that we may reserve τ for the truncation error during stability analysis.

4.2 Boundary Conditions

As we have previously discussed, the spatial variables S and ν can theoretically take any value in the interval $[0, \infty)$. However, in order to successfully implement any FD scheme computationally we must impose artificial boundary conditions as $S \to \infty$ and $\nu \to \infty$ as when programming we may only consider a finite number of grid points.

4.2.1 European Call Option

Consider the boundary conditions imposed on a European call option. The boundary conditions we will discuss are well documented in Heston (1993) [24], In'T Hout and Foulon (2010) [27] and Rouah (2013) [40]. The two types of boundary condition considered are Dirichlet and Neumann boundary conditions. The difference between the type of boundary condition specified is as follows; a Dirichlet boundary condition is one where the *value* a solution must take on the boundary of the domain is specified whilst a Neumann boundary condition is one where the value the *derivative* of a solution must take on the boundary of the domain is stated.

There are four boundary conditions to consider as previously stated in Section 3.4.1. At the lower boundary for the asset price spatial variable, we have $S = 0$ therefore the option is worthless. Thus the Dirichlet boundary condition here is

$$U^n_{0,j} = U(0, \nu_j, t_n) = 0 \tag{4.3}$$

for $0 \leq j \leq N_V$ and $0 \leq n \leq N_T$, where N_V and N_T are the number of volatility and time steps, respectively. As S becomes larger, the delta of the call tends towards one (the option price tends towards the asset price as the strike price becomes negligible). So at $S = S_{max}$ we apply the Neumann boundary condition[1]

$$\frac{\delta_S U^n_{N_S,j}}{2\Delta S} = \frac{\partial U}{\partial S}(S_{max}, \nu_j, t_n) = 1 \tag{4.4}$$

for $0 \leq j \leq N_V$ and $0 \leq n \leq N_T$. It is common for the call option value to increase as volatility increases. This is restricted by the growth of the stock price. We thus have at $\nu = \nu_{max}$ the Dirichlet boundary condition

$$U^n_{i,N_V} = U(S_i, \nu_{max}, t_n) = S_i \tag{4.5}$$

for $0 \leq i \leq N_S$ and $0 \leq n \leq N_T$, where N_S is the number of asset steps. Note that this implies $\frac{\partial U}{\partial \nu} = 0$, when U is originally evaluated at the point (S_i, ν_{max}, t_n). The boundary condition at $\nu = \nu_{min}$ is a little more tricky to establish. Note that when $\nu = \nu_{min} = 0$ substituting $\nu = 0$ into the Heston PDE (4.2) yields the following

$$\frac{\partial U}{\partial t}(S, 0, t) = rS\frac{\partial U}{\partial S} + \kappa\theta\frac{\partial U}{\partial \nu} - rU. \tag{4.6}$$

Note that this is not quite a boundary condition as it includes a time derivative. An exact solution to this PDE is not available. Furthermore, there is no second-order derivative present. We thus use a one-sided[2] FD approximation for second-order derivatives.

[1] Note the notation used for the first derivative; this will be introduced in detail in section 4.4.

[2] We will discuss one-sided approximations shortly.

4.3 Construction of a Mesh

4.3.1 Discretization of the Spatial and Time Domains

Let $\Omega \in \mathbb{R}^n$ denote the continuous region on where the PDE of interest is defined. As a first step, this region is discretized and approximated using a regular grid denoted by $\overline{\Omega} \in \Omega$. It is necessary that the domain $\overline{\Omega}$ is either finite or able to be truncated, and rectangular. Within this discretized space, the exact solution $U(\cdot)$ may be approximated by a discrete grid function $\overline{U}(\cdot)$ that is only defined at grid points on the discretized domain. The number of grid points is important as the more grid points there are, the closer to the exact solution will be to the approximation. The finite difference method allows for the approximation of $U(\cdot)$ by $\overline{U}(\cdot)$ through approximations for the differential quotients present in the PDE.

The Heston model is two-dimensional and is spanned by the underlying asset price S and its volatility ν. As these can only take positive values our spatial domain for the IBVP is the semi-infinite Ω: $(0, \infty) \times (0, \infty)$. As we must work with a finite grid for computational reasons, we truncate the domain to $[S_{min}, S_{max}] \times [\nu_{min}, \nu_{max}]$ (where S_{max} and ν_{max} denote the maximum values for the asset S and its volatility ν respectively, and S_{min} and ν_{min} the corresponding minimum values). In effect the upper limits of the bounded domain act as proxies for the limits $S \to \infty$ and $\nu \to \infty$. Since the boundary conditions are needed to substitute for these limits, the values at the boundaries are highly reliant on the quality of the conditions imposed. If the diffusion process[3] in the Heston PDE is dominant, then oscillations caused as a result of poor boundary conditions are minimized. Nevertheless, if a given boundary condition is of low performance then we must extend the truncated domain such that the boundaries lie far from the region of interest. For example, in order to present a robust implementation of the Heston PDE solution with a negligible error, In'T Hout and Foulon (2010) propose the domain $[0, 8K] \times [0, 5]$ where K is the strike price. In this dissertation, we work with the smaller domain $[0, 3K] \times [0, 1.5]$.

The spacing between grid nodes is important. Uniform grids are those whereby grid points are equidistant from each other. The boundary conditions are used as a starting point and the solution $U(\cdot)$ to the IBVP is approximated at each grid point. Therefore the discretization chosen for the truncated bounded domain $\overline{\Omega}$ is important as it is directly related to the size of the error. In regions where approximations are of poor quality it is necessary to employ some technique to refrain errors from growing without bound. In the uniform grid case accuracy may be improved by increasing the number of equidistant grid points[4]. However, this is computationally expensive. Instead, we may choose to employ a non-uniform grid whereby the number of grid points remains constant and thus not affect computational cost. The idea is that we are able to concentrate many grid points near critical points or points of interest by making the mesh finer near these points. In particular, we are able to increase the grid points around the region $S = K$ and

[3]The diffusion process part corresponds to the second order derivatives appearing in the Heston PDE, see In'T Hout and Foulon [27].

[4]This technique is known as *grid refinement*.

$\nu = 0$ where option prices are often needed. Non-uniform grids greatly improve the accuracy of the finite difference discretization compared to uniform grids[5]. When using non-uniform meshes care must be taken however, as accuracy may decay consequently leading to an increase in the error if the mesh if not sufficiently smooth[6].

We first describe the construction steps for a uniform mesh. As before, denote by S_{max} and ν_{max} the maximum value of S and ν on the grid, respectively. Similarly, denote the minimum corresponding values by S_{min} and ν_{min}. Let

$$U_{i,j}^n = U(i\delta S, j\delta\nu, n\delta t) = U(S_i, \nu_j, t_n)$$

denote the value of a European call option at time t_n when underlying has price S_i and volatility ν_j. Here, the values (S_i, ν_j, t_n) represent a particular node. Discretization is composed of $N_S + 1$ nodes in the S direction, $N_V + 1$ nodes in the ν direction, $N_T + 1$ nodes in the t direction.

Express

$$\begin{aligned} S_i &= i\delta S, 0 \leq i \leq N_S \\ \nu_j &= j\delta\nu, 0 \leq j \leq N_V \\ t_n &= n\delta t, 0 \leq t \leq N_T \end{aligned} \tag{4.7}$$

where (δS,$\delta\nu$,δt) are fixed step sizes in (S, ν, t) in turn. The differences are given by $\Delta S = (S_{max} - S_{min})/N_S$, $\Delta\nu = (\nu_{max} - \nu_{min})/N_V$ and $\Delta t = T/N_T$, where T is the maturity of the option. These differences represent the width between two grid points, for example if moving in the S direction ΔS is the difference between the points (S_i, ν_j) and (S_{i+1}, ν_j). We can thus describe the spatial domain as follows

$$\mathcal{G} = \{(S_i, \nu_j) : 0 \leq i \leq N_S, 0 \leq j \leq N_V\}. \tag{4.8}$$

Next we describe a non-uniform grid through coordinate transformations as presented in the work by Tavella and Randall (2000) [43]. The focus is on increasing the density of mesh points around the neighbourhood of $S = K$ and $\nu = 0$. First define the non-uniform mesh in the S-direction $0 = S_0 < S_1 < ... < S_{max} = N_S$ through the transformation

$$S_i = K + c\sinh(\xi_i), i \in [0, N_S] \tag{4.9}$$

where in [27] the parameter $c(> 0)$ controls the quantity of mesh points S_i lying in the region around the strike K and is chosen such that $c = K/5$. The equidistant points ξ_i with $i \in [0, N_S]$ and the respective increments are defined as

$$\xi_i = \sinh^{-1}\left(-\frac{K}{c}\right) + i\Delta\xi \tag{4.10a}$$

$$\Delta\xi = \frac{1}{N_S}\left[\sinh^{-1}\left(\frac{S_{max} - K}{c}\right) - \sinh^{-1}\left(-\frac{K}{c}\right)\right]. \tag{4.10b}$$

[5] In'T Hout and Foulon (2010) [27] explain that this is due to the initial condition $U(S, \nu, 0) = \max(S - K, 0)$ (in the case of a call option) admitting a discontinuity at $S = K$ in the first derivative and for ν approximately zero, the PDE is dominated by the first derivative term.

[6] The mesh as proposed by [27] is smooth. It is explained that the mesh is smooth if there exist $\mathbb{R} \ni c_0$,c_1,$c_2 > 0$ constants such that the increments $\Delta S_i = S_i - S_{i-1}$ satisfy $c_0\Delta\xi \leq \Delta S_i \leq c_1\Delta\xi$ and $|\Delta S_{i+1} - \Delta S_i| \leq c_2(\Delta\xi)^2$.

Similarly, the non-uniform mesh in the ν-direction $0 = \nu_0 < \nu_1 < ... < \nu_{max} = N_V$ is defined through the transformation

$$\nu_j = d \sinh(j\Delta\eta), \, j \in [0, N_V], \tag{4.11}$$

where

$$\Delta\eta = \frac{1}{N_V} \sinh^{-1}\left(\frac{V_{max}}{d}\right). \tag{4.12}$$

In [27] the parameter $d(> 0)$ controls the quantity of mesh points ν_j lying in the region around $\nu = 0$ and is chosen such that $d = V_{max}/500$. An example of a non-uniform mesh is shown in Figure 4.1.

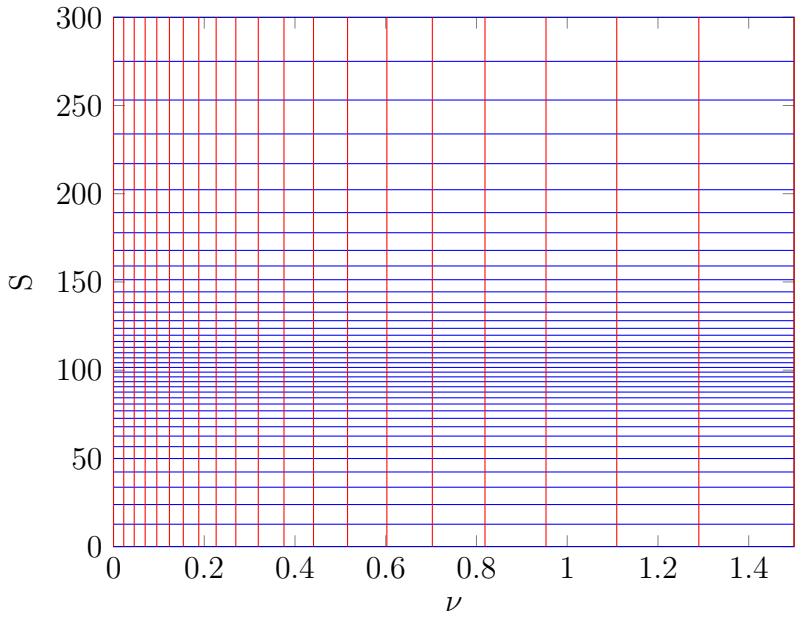


Figure 4.1: Illustration of a non-uniform grid with $S_{max} = 300$ and $\nu_{max} = 1.5$.

For the time discretization technique for the Heston IBVP we let $\Delta t > 0$ be a given time step, where the temporal grid points[7] can be generated by $t_n = n\Delta t$ for $n = 0, 1, 2, ..., N_T$. We progress to describe finite difference methods.

[7] There may be settings where the use of non-equidistant grids in the time direction is beneficial (eg. Bermudan options where the holder has the right to exercise on predetermined dates).

4.4 Finite Difference Operators

4.4.1 Central Differences

In this section we derive approximations to each derivative quotient appearing in the Heston PDE (4.2) by truncating a Taylor series. To this end, we introduce some fundamental notions in a general setting and later apply these to our problem bearing in mind the approximations that are of interest to us. We work with finite difference approximations for internal mesh points in the case that the grid used is uniform. Given a step size Δ, define the finite differences via the compact notation for ease of reference (note that x is used to denote a general spatial variable):

Forward Differences

$$\Delta_{+t}U(x,t) \coloneqq U(x,t+\Delta t) - U(x,t) \tag{4.13a}$$
$$\Delta_{+x}U(x,t) \coloneqq U(x+\Delta x,t) - U(x,t) \tag{4.13b}$$

Backward Differences

$$\Delta_{-t}U(x,t) \coloneqq U(x,t) - U(x,t-\Delta t) \tag{4.14a}$$
$$\Delta_{-x}U(x,t) \coloneqq U(x,t) - U(x-\Delta x,t) \tag{4.14b}$$

Central Differences

$$\delta_t U(x,t) \coloneqq U(x,t+\Delta t) - U(x,t-\Delta t) \tag{4.15a}$$
$$\delta_x U(x,t) \coloneqq U(x+\Delta x,t) - U(x-\Delta x,t) \tag{4.15b}$$

Second Order Central Differences

$$\delta_x^2 U(x,t) \coloneqq\ U(x+\Delta x,t) - 2U(x,t) + U(x-\Delta x,t) \tag{4.16}$$

Note that we are able to derive expressions for the derivative terms present in the discretized form of the Heston PDE (which we write out in full according to a particular scheme in subsequent sections) by considering Taylor series expansions of the forward and backward differences. Assuming that the function being approximated is smooth, Taylor expansions are as follows[8]

$$\begin{aligned}\Delta_{+t}U(x,t) &= U(x,t+\Delta t) - U(x,t)\\ &= \frac{\partial U}{\partial t}\Delta t + \frac{1}{2!}\frac{\partial^2 U}{\partial t^2}\Delta t^2 + \frac{1}{3!}\frac{\partial^3 U}{\partial t^3}\Delta t^3 + h.o.t. \end{aligned}\tag{4.17a}$$
$$\begin{aligned}\Delta_{-t}U(x,t) &= U(x,t) - U(x,t-\Delta t)\\ &= \frac{\partial U}{\partial t}\Delta t - \frac{1}{2!}\frac{\partial^2 U}{\partial t^2}\Delta t^2 + \frac{1}{3!}\frac{\partial^3 U}{\partial t^3}\Delta x^3 + h.o.t. \end{aligned}\tag{4.17b}$$
$$\begin{aligned}\Delta_{+x}U(x,t) &= U(x+\Delta x,t) - U(x,t)\\ &= \frac{\partial U}{\partial x}\Delta x + \frac{1}{2!}\frac{\partial^2 U}{\partial x^2}\Delta x^2 + \frac{1}{3!}\frac{\partial^3 U}{\partial x^3}\Delta x^3 + h.o.t. \end{aligned}\tag{4.17c}$$
$$\begin{aligned}\Delta_{-x}U(x,t) &= U(x,t) - U(x-\Delta x,t)\\ &= \frac{\partial U}{\partial x}\Delta x - \frac{1}{2!}\frac{\partial^2 U}{\partial x^2}\Delta x^2 + \frac{1}{3!}\frac{\partial^3 U}{\partial x^3}\Delta x^3 + h.o.t. \end{aligned}\tag{4.17d}$$

[8]Note that *h.o.t.* stands for *higher order terms.*

From equations (4.17c) and (4.17d) we can obtain a finite difference approximation for the quotient $\frac{\partial U}{\partial x}$. Write

$$\begin{aligned} \Delta_{+x}U(x,t) + \Delta_{-x}U(x,t) &= U(x+\Delta x,t) - U(x-\Delta x,t) \\ &= 2\frac{\partial U}{\partial x}\Delta x + \frac{1}{3}\frac{\partial^3 U}{\partial x^3}\Delta x^3 + h.o.t. \\ \Rightarrow \frac{\partial U}{\partial x} &= \frac{\delta_x U(x,t)}{2\Delta x} + \mathcal{O}(\Delta x^2). \end{aligned} \tag{4.18}$$

Similarly, we can obtain a finite difference approximation for the quotient $\frac{\partial^2 U}{\partial x^2}$ by taking the difference between equations (4.17c) and (4.17d) (this is also directly a Taylor series expansion of the second order central difference $\delta_x^2 U(x,t)$):

$$\begin{aligned} \Delta_{+x}U(x,t) - \Delta_{-x}U(x,t) &= U(x+\Delta x,t) - 2U(x,t) + U(x-\Delta x,t) \\ &= \frac{\partial^2 U}{\partial x^2}\Delta x^2 + \frac{1}{12}\frac{\partial^4 U}{\partial x^4}\Delta x^4 + h.o.t. \\ \Rightarrow \frac{\partial^2 U}{\partial x^2} &= \frac{\delta_x^2 U(x,t)}{\Delta x^2} + \mathcal{O}(\Delta x^2). \end{aligned} \tag{4.19}$$

We use a forward difference for the time derivative. This may be easily obtained by rearranging equation (4.17a):

$$\frac{\partial U}{\partial t} = \frac{\Delta_{+t}U(x,t)}{\Delta t} + \mathcal{O}(\Delta t). \tag{4.20}$$

Applying these approximations to the Heston PDE we can adopt the following notations

$$\frac{\partial U}{\partial t} \approx \frac{\Delta_{+t}U(S,\nu,t)}{\Delta t} \tag{4.21a}$$

$$\frac{\partial U}{\partial S} \approx \frac{\delta_S U(S,\nu,t)}{2\Delta S} \tag{4.21b}$$

$$\frac{\partial U}{\partial \nu} \approx \frac{\delta_\nu U(S,\nu,t)}{2\Delta \nu} \tag{4.21c}$$

$$\frac{\partial^2 U}{\partial S^2} \approx \frac{\delta_S^2 U(S,\nu,t)}{\Delta S^2} \tag{4.21d}$$

$$\frac{\partial^2 U}{\partial \nu^2} \approx \frac{\delta_\nu^2 U(S,\nu,t)}{\Delta \nu^2}. \tag{4.21e}$$

It remains to find the finite difference approximation quotient for the cross derivative term $\frac{\partial^2 U}{\partial S \nu}$ in the Heston PDE arising from the correlation between the asset and variance processes. In order to find such an approximation the following Taylor

expansions are useful

$$\begin{aligned} U(S+\Delta S,\nu+\Delta\nu,t) = U(S,\nu,t) &+ \frac{\partial U}{\partial S}\Delta S + \frac{\partial U}{\partial \nu}\Delta\nu + \frac{1}{2}\frac{\partial^2 U}{\partial S^2}\Delta S^2 \\ &+ \frac{\partial^2 U}{\partial S\partial\nu}\Delta S\Delta\nu + \frac{1}{2}\frac{\partial^2 U}{\partial \nu^2}\Delta\nu^2 + \frac{1}{3!}\frac{\partial^3 U}{\partial S^3}\Delta S^3 \\ &+ \frac{3}{3!}\frac{\partial^3 U}{\partial S^2\partial\nu}\Delta S^2\Delta\nu + \frac{3}{3!}\frac{\partial^3 U}{\partial S\partial\nu^2}\Delta S\Delta\nu^2 + \frac{1}{3!}\frac{\partial^3 U}{\partial \nu^3} + h.o.t. \end{aligned} \tag{4.22a}$$

$$\begin{aligned} U(S-\Delta S,\nu-\Delta\nu,t) = U(S,\nu,t) &- \frac{\partial U}{\partial S}\Delta S - \frac{\partial U}{\partial \nu}\Delta\nu + \frac{1}{2}\frac{\partial^2 U}{\partial S^2}\Delta S^2 \\ &+ \frac{\partial^2 U}{\partial S\partial\nu}\Delta S\Delta\nu + \frac{1}{2}\frac{\partial^2 U}{\partial \nu^2}\Delta\nu^2 - \frac{1}{3!}\frac{\partial^3 U}{\partial S^3}\Delta S^3 \\ &- \frac{3}{3!}\frac{\partial^3 U}{\partial S^2\partial\nu}\Delta S^2\Delta\nu - \frac{3}{3!}\frac{\partial^3 U}{\partial S\partial\nu^2}\Delta S\Delta\nu^2 - \frac{1}{3!}\frac{\partial^3 U}{\partial \nu^3} + h.o.t. \end{aligned} \tag{4.22b}$$

$$\begin{aligned} U(S-\Delta S,\nu+\Delta\nu,t) = U(S,\nu,t) &- \frac{\partial U}{\partial S}\Delta S + \frac{\partial U}{\partial \nu}\Delta\nu + \frac{1}{2}\frac{\partial^2 U}{\partial S^2}\Delta S^2 \\ &- \frac{\partial^2 U}{\partial S\partial\nu}\Delta S\Delta\nu + \frac{1}{2}\frac{\partial^2 U}{\partial \nu^2}\Delta\nu^2 - \frac{1}{3!}\frac{\partial^3 U}{\partial S^3}\Delta S^3 \\ &+ \frac{3}{3!}\frac{\partial^3 U}{\partial S^2\partial\nu}\Delta S^2\Delta\nu - \frac{3}{3!}\frac{\partial^3 U}{\partial S\partial\nu^2}\Delta S\Delta\nu^2 + \frac{1}{3!}\frac{\partial^3 U}{\partial \nu^3} + h.o.t. \end{aligned} \tag{4.22c}$$

$$\begin{aligned} U(S+\Delta S,\nu-\Delta\nu,t) = U(S,\nu,t) &+ \frac{\partial U}{\partial S}\Delta S - \frac{\partial U}{\partial \nu}\Delta\nu + \frac{1}{2}\frac{\partial^2 U}{\partial S^2}\Delta S^2 \\ &- \frac{\partial^2 U}{\partial S\partial\nu}\Delta S\Delta\nu + \frac{1}{2}\frac{\partial^2 U}{\partial \nu^2}\Delta\nu^2 + \frac{1}{3!}\frac{\partial^3 U}{\partial S^3}\Delta S^3 \\ &- \frac{3}{3!}\frac{\partial^3 U}{\partial S^2\partial\nu}\Delta S^2\Delta\nu + \frac{3}{3!}\frac{\partial^3 U}{\partial S\partial\nu^2}\Delta S\Delta\nu^2 - \frac{1}{3!}\frac{\partial^3 U}{\partial \nu^3} + h.o.t. \end{aligned} \tag{4.22d}$$

From these equations we have that

$$\begin{aligned} U(S+\Delta S,\nu+\Delta\nu,t) + U(S-\Delta S,\nu-\Delta\nu,t) & \\ - U(S-\Delta S,\nu+\Delta\nu,t) - U(S+\Delta S,\nu-\Delta\nu,t) &\approx 4\frac{\partial^2 U}{\partial S\partial\nu}\Delta S\Delta\nu, \end{aligned} \tag{4.23}$$

such that

$$\begin{aligned} \frac{\partial^2 U}{\partial S\partial\nu} \approx & \frac{U(S+\Delta S,\nu+\Delta\nu,t) + U(S-\Delta S,\nu-\Delta\nu,t)}{4\Delta S\Delta\nu} \\ & + \frac{-U(S-\Delta S,\nu+\Delta\nu,t) - U(S+\Delta S,\nu-\Delta\nu,t)}{4\Delta S\Delta\nu}. \end{aligned} \tag{4.24}$$

Adopt the following notation for the mixed derivative term

$$\frac{\partial^2 U}{\partial S\partial\nu} \approx \frac{\delta^2_{S\nu} U(S,\nu,t)}{4\Delta S\Delta\nu}. \tag{4.25}$$

Finally, describe the FD approximations at a grid point (S_i, ν_j, t_n) as follows

$$\theta = \frac{\partial U}{\partial t} \approx \frac{U_{i,j}^{n+1} - U_{i,j}^{n}}{\Delta t} \tag{4.26a}$$

$$\Delta_S = \frac{\partial U}{\partial S} \approx \frac{U_{i+1,j}^{n} - U_{i-1,j}^{n}}{2\Delta S} \tag{4.26b}$$

$$\Delta_V = \frac{\partial U}{\partial \nu} \approx \frac{U_{i,j+1}^{n} - U_{i,j-1}^{n}}{2\Delta \nu} \tag{4.26c}$$

$$\Gamma_S = \frac{\partial^2 U}{\partial S^2} \approx \frac{U_{i-1,j}^{n} - 2U_{i,j}^{n} + U_{i+1,j}^{n}}{\Delta S^2} \tag{4.26d}$$

$$\Gamma_V = \frac{\partial^2 U}{\partial \nu^2} \approx \frac{U_{i,j-1}^{n} - 2U_{i,j}^{n} + U_{i,j+1}^{n}}{\Delta \nu^2} \tag{4.26e}$$

$$\frac{\partial^2 U}{\partial S \partial \nu} \approx \frac{U_{i+1,j+1}^{n} + U_{i-1,j-1}^{n} - U_{i-1,j+1}^{n} - U_{i+1,j-1}^{n}}{4\Delta S \Delta \nu}. \tag{4.26f}$$

Remark. We were able to truncate the Taylor series by use of 'big $\mathcal{O}$' notation, which we define below[9].

Definition 4.4.1. Let $f(x)$ and $g(x)$ be two functions defined on a subset of the real line. Then write

$$f(x) = \quad \mathcal{O}(g(x)) \text{ as } x \to \infty \tag{4.27}$$

if and only if $\exists M > 0$ constant such that for all sufficiently large values of x we have that $f(x)$ will be at most M multiplied by the absolute value of $g(x)$. This means that $f(x) = \quad \mathcal{O}(g(x))$ if and only if $\exists M > 0$ a real constant and $x_0 \in \mathbb{R}$ such that

$$|f(x)| \leq M|g(x)| \ \forall \ x \geq x_0 \tag{4.28}$$

□

Applying this definition, we are able to write equation (4.17c) for example, since the following holds

$$\left| \frac{1}{3!}\frac{\partial^3 U}{\partial S^3}\Delta S^3 + \frac{1}{4!}\frac{\partial^4 U}{\partial S^4}\Delta S^4 + ... \right| \leq M|\Delta S^3| \text{ as } \Delta S \to 0. \tag{4.29}$$

4.4.2 One-sided Differences

The central differences described in the previous section are not applicable on the entire domain. This is because at the boundaries of the domain we fail to find two direct neighbours to approximate the given boundary grid node. So at the boundaries we cannot apply a central difference. For example we may be at the boundary $U_{i,0}^{n}$ corresponding to $\nu = 0$ and at this node the value $U_{i,-1}^{n}$ won't be available to evaluate the second central difference $\delta_\nu^2 U_{i,j}^{n}$. In this case we can apply one-sided FD schemes. For instance at the boundary where $\nu = 0$ the first derivative may be approximated via a one-sided forward difference scheme. Again

[9]Extended definition may we found in Wikipedia http://en.wikipedia.org/wiki/Big˙O˙notation.

we use Taylor series expansions. First seek an expression for the first derivative in terms of $U^n_{i,j-1}$, $U^n_{i,j}$ and $U^n_{i,j+1}$

$$\begin{aligned}
&\eta_0 U(S,\nu,t)+\eta_1 U(S,\nu+\Delta\nu,t)+\eta_2 U(S,\nu+2\Delta\nu)\\
&\quad=\eta_0 U(S,\nu,t)+\eta_1\left(U(S,\nu,t)+\left(\frac{\partial U}{\partial\nu}\right)_{i,j}\Delta\nu+\frac{1}{2}\left(\frac{\partial^2 U}{\partial\nu^2}\right)_{i,j}\Delta\nu^2+\mathcal{O}(\Delta\nu^3)\right)\\
&\qquad+\eta_2\left(U(S,\nu,t)+\left(\frac{\partial U}{\partial\nu}\right)_{i,j}2\Delta\nu+\frac{1}{2}\left(\frac{\partial^2 U}{\partial\nu^2}\right)_{i,j}(2\Delta\nu)^2+\mathcal{O}(\Delta\nu^3)\right)\\
&\quad=(\eta_0+\eta_1+\eta_2)U(S,\nu,t)+(\eta_1\Delta\nu+2\eta_2\Delta\nu)\left(\frac{\partial U}{\partial\nu}\right)_{i,j}\\
&\qquad+\left(\frac{\eta_1\Delta\nu^2}{2}+2\eta_2\Delta\nu^2\right)\left(\frac{\partial^2 U}{\partial\nu^2}\right)_{i,j}+\mathcal{O}(\Delta\nu^3).
\end{aligned}$$

(4.30)

In order to isolate the first derivative term $\left(\frac{\partial U}{\partial\nu}\right)_{i,j}$ we require

$$\left.\begin{array}{l}\eta_0+\eta_1+\eta_2=0\\ \eta_1\Delta\nu+2\eta_2\Delta\nu=1\\ \frac{\eta_1\Delta\nu^2}{2}+2\eta_2\Delta\nu^2=0\end{array}\right\}\Rightarrow\begin{array}{l}\eta_0=-\frac{3}{2\Delta\nu}\\ \eta_1=\frac{4}{2\Delta\nu}\\ \eta_2=-\frac{1}{2\Delta\nu}.\end{array}$$

If we now substitute the expressions for η_0, η_1 and η_2 into equation (4.30) we obtain an approximation to the first derivative

$$\begin{aligned}
\delta_{+\nu}U(S,\nu,t)=&\frac{-3U(S,\nu,t)+4U(S,\nu+\Delta\nu,t)-U(S,\nu+2\Delta\nu,t)}{2\Delta\nu}\\
&-\frac{1}{3!}\left(\frac{\partial^3 U}{\partial\nu^3}\right)_{i,j}\Delta\nu^2+\mathcal{O}(\Delta\nu^3),
\end{aligned}\tag{4.31}$$

or in our favoured notation

$$\delta_{+\nu}U^n_{i,j}\approx\frac{-3U^n_{i,j}+4U^n_{i,j+1}-U^n_{i,j+2}}{2\Delta\nu}.\tag{4.32}$$

This is known as the second-order correct, forward difference approximation to the first derivative. In general, one-sided difference schemes may be used to approximate non-boundary nodes also. As there are some disadvantages (see Section 7 De Graaf (2012) [12] for details) to central differences, a useful one-sided alternative is available for use at interior grid points to replace approximations to the second-order derivative. Furthermore, use of backward difference approximations is convenient when the drift terms of the asset and variance processes in the Heston PDE become negative. This allows the dissipation of the effect of oscillations resulting from the negativity of these coefficients on the solution. For example the drift term for the variance process $\kappa(\theta-\nu)$ can become negative whenever $\theta<\nu$. A backward one-sided difference may be derived as before[10]

$$\delta_{-\nu}\approx\frac{U^n_{i,j-2}-4U^n_{i,j-1}+3U^n_{i,j}}{2\Delta\nu}.\tag{4.33}$$

[10]This improvement is later described in further detail and is known as an *upwinding* technique.

Similarly, we can derive one-sided second order accurate approximations to the second-order derivative, although these are not used in our implementation,

$$\delta^2_{+\nu} U^n_{i,j} \approx \frac{2U^n_{i,j} - 5U^n_{i,j+1} + 4U^n_{i,j+2} - U^n_{i,j+3}}{\Delta\nu^2} \tag{4.34}$$

$$\delta^2_{-\nu} U^n_{i,j} \approx \frac{-U^n_{i,j-3} + 4U^n_{i,j-2} - 5U^n_{i,j-1} + 2U^n_{i,j}}{\Delta\nu^2}. \tag{4.35}$$

4.5 Explicit Finite Difference Scheme

In this section we use the discretized form of the Heston PDE (4.2) and implement the Explicit Finite Difference Method on a uniform grid (equally spaced intervals for both state variables). Substituting the FD approximations into the Heston PDE (3.10) we obtain

$$\begin{aligned}\frac{U^{n+1}_{i,j} - U^n_{i,j}}{\Delta t} = &\frac{1}{2} S_i^2 \nu_j \frac{U^n_{i-1,j} - 2U^n_{i,j} + U^n_{i+1,j}}{\Delta S^2} + \frac{1}{2}\sigma^2 \nu_j \frac{U^n_{i,j-1} - 2U^n_{i,j} + U^n_{i,j+1}}{\Delta\nu^2} \\ &+ rS_i \frac{U^n_{i+1,j} - U^n_{i-1,j}}{2\Delta S} + \kappa(\theta - \nu_j)\frac{U^n_{i,j+1} - U^n_{i,j-1}}{2\Delta\nu} - rU^n_{i,j} \\ &+ \frac{1}{4}\rho\sigma S_i \nu_j \frac{U^n_{i+1,j+1} + U^n_{i-1,j-1} - U^n_{i-1,j+1} - U^n_{i+1,j-1}}{\Delta S \Delta\nu}.\end{aligned} \tag{4.36}$$

Using our more compact notation

$$\begin{aligned}\frac{\Delta_{+t} U^n_{i,j}}{\Delta t} = &\frac{1}{2} S_i^2 \nu_j \frac{\delta^2_S U^n_{i,j}}{\Delta S^2} + \frac{1}{4}\rho\sigma S_i \nu_j \frac{\delta^2_{S\nu} U^n_{i,j}}{\Delta S \Delta\nu} + \frac{1}{2}\sigma^2 \nu_j \frac{\delta^2_\nu U^n_{i,j}}{\Delta\nu^2} + rS_i \frac{\delta_S U^n_{i,j}}{2\Delta S} \\ &+ \kappa(\theta - \nu_j)\frac{\delta_\nu U^n_{i,j}}{2\Delta\nu} - rU^n_{i,j}.\end{aligned} \tag{4.37}$$

where we have used a forward difference for the time derivative. Rearrange this equation grouping common terms to obtain

$$\begin{aligned}U^{n+1}_{i,j} = &\left[1 - \Delta t\left(\frac{S_i^2\nu_j}{\Delta S^2} + \frac{\sigma^2\nu_j}{\Delta\nu^2} + r\right)\right] U^n_{i,j} \\ &+ \left[\frac{1}{2}\Delta t\left(\frac{S_i^2\nu_j}{\Delta S^2} - \frac{rS_i}{\Delta S}\right)\right] U^n_{i-1,j} + \left[\frac{1}{2}\Delta t\left(\frac{S_i^2\nu_j}{\Delta S^2} + \frac{rS_i}{\Delta S}\right)\right] U^n_{i+1,j} \\ &+ \left[\frac{1}{2}\Delta t\left(\frac{\sigma^2\nu_j}{\Delta\nu^2} - \frac{\kappa(\theta - \nu_j)}{\Delta\nu}\right)\right] U^n_{i,j-1} + \left[\frac{1}{2}\Delta t\left(\frac{\sigma^2\nu_j}{\Delta\nu^2} + \frac{\kappa(\theta - \nu_j)}{\Delta\nu}\right)\right] U^n_{i,j+1} \\ &+ \frac{1}{4}\frac{\Delta t\rho\sigma S_i\nu_j}{\Delta S\Delta\nu}\left[U^n_{i+1,j+1} + U^n_{i-1,j-1} - U^n_{i-1,j+1} - U^n_{i+1,j-1}\right].\end{aligned} \tag{4.38}$$

If we now define as previously $S_i = i\Delta S$, $\nu_j = j\Delta\nu$ we can write a the simplified form[11] of the discretized PDE

$$\begin{aligned}\frac{U_{i,j}^{n+1} - U_{i,j}^n}{\Delta t} = &\frac{1}{2}j\Delta\nu i^2\Delta S^2\frac{U_{i-1,j}^n - 2U_{i,j}^n + U_{i+1,j}^n}{\Delta S^2} + \frac{1}{2}\sigma^2 j\Delta\nu\frac{U_{i,j-1}^n - 2U_{i,j}^n + U_{i,j+1}^n}{\Delta\nu^2}\\ &+ \rho\sigma ij\Delta\nu\Delta S\frac{U_{i+1,j+1}^n + U_{i-1,j-1}^n - U_{i-1,j+1}^n - U_{i+1,j-1}^n}{4\Delta S\Delta\nu}\\ &+ ri\Delta S\frac{U_{i+1,j}^n - U_{i-1,j}^n}{2\Delta S} + \kappa(\theta - j\Delta\nu)\frac{U_{i,j+1}^n - U_{i,j-1}^n}{2\Delta\nu} - rU_{i,j}^n.\end{aligned} \tag{4.39}$$

Rewrite this expression via a set of coefficients as follows

$$\begin{aligned}U_{i,j}^{n+1} = &A_{i,j}^n U_{i,j}^n + B_{i,j}^n U_{i-1,j}^n + C_{i,j}^n U_{i+1,j}^n + D_{i,j}^n U_{i,j-1}^n + E_{i,j}^n U_{i,j+1}^n\\ &F_{i,j}^n(U_{i+1,j+1}^n + U_{i-1,j+1}^n + U_{i+1,j-1}^n + U_{i-1,j-1}^n),\end{aligned} \tag{4.40}$$

where

$$A_{i,j}^n = 1 - \Delta t\left(i^2 j\Delta\nu + \frac{\sigma^2 j}{\Delta\nu} + r\right) \tag{4.41a}$$

$$B_{i,j}^n = \frac{1}{2}i\Delta t\left(ij\Delta\nu - r\right) \tag{4.41b}$$

$$C_{i,j}^n = \frac{1}{2}i\Delta t\left(ij\Delta\nu + r\right) \tag{4.41c}$$

$$D_{i,j}^n = \frac{1}{2}\frac{\Delta t}{\Delta\nu}\left(\sigma^2 j - \kappa(\theta - j\Delta\nu)\right) \tag{4.41d}$$

$$E_{i,j}^n = \frac{1}{2}\frac{\Delta t}{\Delta\nu}\left(\sigma^2 j + \kappa(\theta - j\Delta\nu)\right) \tag{4.41e}$$

$$F_{i,j}^n = \frac{1}{4}ij\rho\sigma\Delta t. \tag{4.41f}$$

The explicit scheme is very popular due to its tractability and ease of implementation. Because of its conditional stability however, other methods are often sought in practice.

An illustration of the scheme is given in Figure 4.2. The Heston parameter values chosen are $\kappa = 2$, $\theta = 0.2$, $\sigma = 0.3$, $\rho = 0.8$, $r = 0.03$, $q = 0$, $\lambda = 0$, $K = 100$ and $T = 1$ year. Within the function we have set $S_{max} = 200$, $V_{max} = 1$, number of asset steps $N_S = 40$, number of volatility steps $N_V = 20$ and number of time steps $N_T = 4000$.

[11] Note that when coding in Matlab we use the PDE (4.38) to simplify understanding of the code.

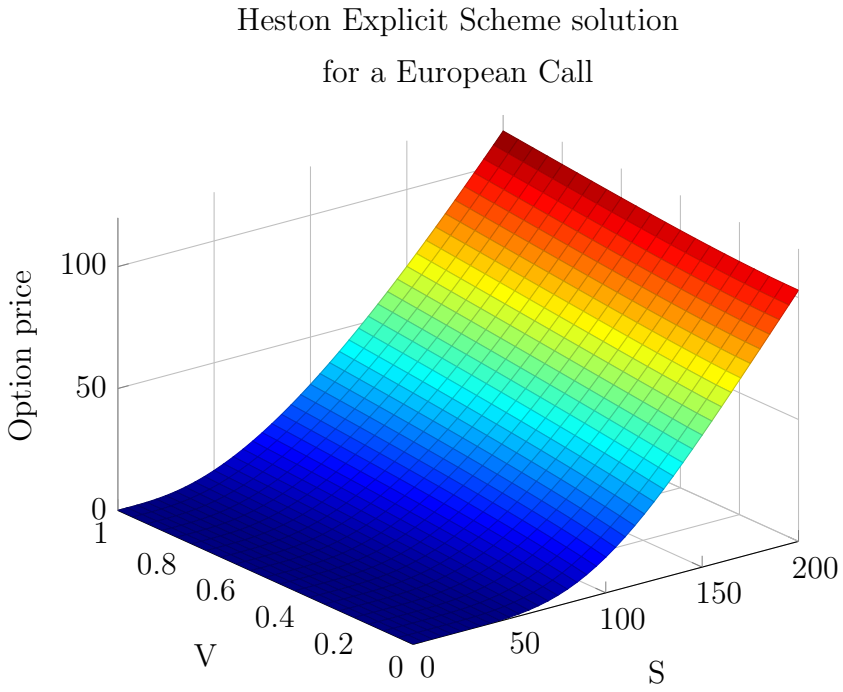


Figure 4.2: Illustration of the Heston explicit scheme for a European call option.

4.6 Implicit Finite Difference Scheme

We now implement the implicit finite difference method. Using a backward difference in time and central differences in space variables S and ν, we have

$$\begin{aligned}\frac{U_{i,j}^{n+1}-U_{i,j}^{n}}{\Delta t} &= \frac{1}{2}S_i^2\nu_j\frac{U_{i-1,j}^{n+1}-2U_{i,j}^{n}+U_{i+1,j}^{n+1}}{\Delta S^2}+\frac{1}{2}\sigma^2\nu_j\frac{U_{i,j-1}^{n+1}-2U_{i,j}^{n+1}+U_{i,j+1}^{n+1}}{\Delta\nu^2}\\ &\quad+rS_i\frac{U_{i+1,j}^{n+1}-U_{i-1,j}^{n+1}}{2\Delta S}+\kappa(\theta-\nu_j)\frac{U_{i,j+1}^{n+1}-U_{i,j-1}^{n+1}}{2\delta\nu}-rU_{i,j}^{n+1}\\ &\quad+\frac{1}{4}\rho\sigma S_i\nu_j\frac{U_{i+1,j+1}^{n+1}+U_{i-1,j-1}^{n+1}-U_{i-1,j+1}^{n+1}-U_{i+1,j-1}^{n+1}}{\Delta S\Delta\nu},\end{aligned} \tag{4.42}$$

which in our more compact notation becomes

$$\begin{aligned}\frac{\Delta_{-t}U_{i,j}^{n+1}}{\Delta t} &= \frac{1}{2}S_i^2\nu_j\frac{\delta_S^2U_{i,j}^{n+1}}{\Delta S^2}+\frac{1}{4}\rho\sigma S_i\nu_j\frac{\delta_{S\nu}^2U_{i,j}^{n+1}}{\Delta S\Delta\nu}+\frac{1}{2}\sigma^2\nu_j\frac{\delta_\nu^2U_{i,j}^{n+1}}{\Delta\nu^2}+rS_i\frac{\delta_SU_{i,j}^{n+1}}{2\Delta S}\\ &\quad+\kappa(\theta-\nu_j)\frac{\delta_\nu U_{i,j}^{n+1}}{2\delta\nu}-rU_{i,j}^{n+1}.\end{aligned} \tag{4.43}$$

Rearrange grouping common terms to obtain

$$\begin{aligned}
&\left[1+\Delta t\left(\frac{S_i^2\nu_j}{\Delta S^2}+\frac{\sigma^2\nu_j}{\Delta\nu^2}+r\right)\right]U_{i,j}^{n+1}\\
&+\left[\frac{1}{2}\Delta t\left(-\frac{S_i^2\nu_j}{\Delta S^2}+\frac{rS_i}{\Delta S}\right)\right]U_{i-1,j}^{n+1}+\left[\frac{1}{2}\Delta t\left(-\frac{S_i^2\nu_j}{\Delta S^2}-\frac{rS_i}{\Delta S}\right)\right]U_{i+1,j}^{n+1}\\
&+\left[\frac{1}{2}\Delta t\left(-\frac{\sigma^2\nu_j}{\Delta\nu^2}+\frac{\kappa(\theta-\nu_j)}{\Delta\nu}\right)\right]U_{i,j-1}^{n+1}+\left[\frac{1}{2}\Delta t\left(-\frac{\sigma^2\nu_j}{\Delta\nu^2}-\frac{\kappa(\theta-\nu_j)}{\Delta\nu}\right)\right]U_{i,j+1}^{n+1}\\
&-\frac{1}{4}\frac{\Delta t\rho\sigma S_i\nu_j}{\Delta S\Delta\nu}\left[U_{i+1,j+1}^{n+1}+U_{i-1,j-1}^{n+1}-U_{i-1,j+1}^{n+1}-U_{i+1,j-1}^{n+1}\right]=\ U_{i,j}^{n}.
\end{aligned}\tag{4.44}$$

Simplified version of the scheme

$$\begin{aligned}
\frac{U_{i,j}^{n+1}-U_{i,j}^{n}}{\Delta t}=&\frac{1}{2}j\Delta\nu i^2\Delta S^2\frac{U_{i-1,j}^{n+1}-2U_{i,j}^{n+1}+U_{i+1,j}^{n+1}}{\Delta S^2}+\frac{1}{2}\sigma^2 j\Delta\nu\frac{U_{i,j-1}^{n+1}-2U_{i,j}^{n+1}+U_{i,j+1}^{n+1}}{\Delta\nu^2}\\
&+\rho\sigma ij\Delta\nu\Delta S\frac{U_{i+1,j+1}^{n+1}+U_{i-1,j-1}^{n+1}-U_{i-1,j+1}^{n+1}-U_{i+1,j-1}^{n+1}}{4\Delta S\Delta\nu}\\
&+ri\Delta S\frac{U_{i+1,j}^{n+1}-U_{i-1,j}^{n+1}}{2\Delta S}+\kappa(\theta-j\Delta\nu)\frac{U_{i,j+1}^{n+1}-U_{i,j-1}^{n+1}}{2\Delta\nu}-rU_{i,j}^{n+1}.
\end{aligned}\tag{4.45}$$

If we define the following coefficients

$$a_{i,j}^{n+1}=1+\Delta t\left(i^2j\Delta\nu+\frac{\sigma^2 j}{\Delta\nu}+r\right)\tag{4.46a}$$

$$b_{i,j}^{n+1}=\frac{1}{2}i\Delta t\left(-ij\Delta\nu+r\right)\tag{4.46b}$$

$$c_{i,j}^{n+1}=\frac{1}{2}i\Delta t\left(-ij\Delta\nu-r\right)\tag{4.46c}$$

$$d_{i,j}^{n+1}=\frac{1}{2}\frac{\Delta t}{\Delta\nu}\left(-\sigma^2 j+\kappa(\theta-j\Delta\nu)\right)\tag{4.46d}$$

$$e_{i,j}^{n+1}=\frac{1}{2}\frac{\delta t}{\Delta\nu}\left(-\sigma^2 j-\kappa(\theta-j\Delta\nu)\right)\tag{4.46e}$$

$$f_{i,j}^{n+1}=\frac{1}{4}ij\rho\sigma\Delta t\tag{4.46f}$$

then we may rewrite the implicit scheme as follows

$$\begin{aligned}
a_{i,j}^{n+1}U_{i,j}^{n+1}+b_{i,j}^{n+1}U_{i-1,j}^{n}+c_{i,j}^{n+1}U_{i+1,j}^{n+1}+d_{i,j}^{n+1}U_{i,j-1}^{n+1}+e_{i,j}^{n+1}U_{i,j+1}^{n+1}\\
+f_{i,j}^{n+1}(U_{i+1,j+1}^{n+1}+U_{i-1,j+1}^{n+1}+U_{i+1,j-1}^{n+1}+U_{i-1,j-1}^{n+1})=\ U_{i,j}^{n}.
\end{aligned}\tag{4.47}$$

The implicit scheme is unconditionally stable.

4.7 The θ'-Scheme

If we take a weighted average[12] of both explicit and implicit schemes we obtain the θ'-Scheme, which is as follows

$$\begin{aligned}\frac{\Delta_{+t}U^n_{i,j}}{\Delta t} = \theta' &\left[\frac{1}{2}S_i^2\nu_j\frac{\delta_S^2 U^{n+1}_{i,j}}{\Delta S^2} + \frac{1}{4}\rho\sigma S_i\nu_j\frac{\delta^2_{S\nu}U^{n+1}_{i,j}}{\Delta S\Delta\nu} + \frac{1}{2}\sigma^2\nu_j\frac{\delta^2_\nu U^{n+1}_{i,j}}{\Delta\nu^2} + rS_i\frac{\delta_S U^{n+1}_{i,j}}{2\Delta S}\right.\\ &\left.+\kappa(\theta-\nu_j)\frac{\delta_\nu U^{n+1}_{i,j}}{2\delta\nu} - rU^{n+1}_{i,j}\right]\\ &+(1-\theta')\left[\frac{1}{2}S_i^2\nu_j\frac{\delta_S^2 U^{n}_{i,j}}{\Delta S^2} + \frac{1}{4}\rho\sigma S_i\nu_j\frac{\delta^2_{S\nu}U^{n}_{i,j}}{\Delta S\Delta\nu} + \frac{1}{2}\sigma^2\nu_j\frac{\delta^2_\nu U^{n}_{i,j}}{\Delta\nu^2} + rS_i\frac{\delta_S U^{n}_{i,j}}{2\Delta S}\right.\\ &\left.+\kappa(\theta-\nu_j)\frac{\delta_\nu U^{n}_{i,j}}{2\Delta\nu} - rU^{n}_{i,j}\right],\end{aligned} \tag{4.48}$$

or in full notation

$$\begin{aligned}\frac{U^{n+1}_{i,j}-U^n_{i,j}}{\Delta t} = \theta' &\left(\frac{1}{2}j\Delta\nu i^2\Delta S^2\frac{U^{n+1}_{i-1,j}-2U^{n+1}_{i,j}+U^{n+1}_{i+1,j}}{\Delta S^2} + \frac{1}{2}\sigma^2 j\Delta\nu\frac{U^{n+1}_{i,j-1}-2U^{n+1}_{i,j}+U^{n+1}_{i,j+1}}{\Delta\nu^2}\right.\\ &+\rho\sigma ij\Delta\nu\Delta S\frac{U^{n+1}_{i+1,j+1}+U^{n+1}_{i-1,j-1}-U^{n+1}_{i-1,j+1}-U^{n+1}_{i+1,j-1}}{4\Delta S\Delta\nu}\\ &\left.+ri\Delta S\frac{U^{n+1}_{i+1,j}-U^{n+1}_{i-1,j}}{2\Delta S} + \kappa(\theta-j\Delta\nu)\frac{U^{n+1}_{i,j+1}-U^{n+1}_{i,j-1}}{2\Delta\nu} - rU^{n+1}_{i,j}\right)\\ &+(1-\theta')\left(\frac{1}{2}j\Delta\nu i^2\Delta S^2\frac{U^{n}_{i-1,j}-2U^{n}_{i,j}+U^{n}_{i+1,j}}{\Delta S^2}\right.\\ &+\sigma^2 j\Delta\nu\frac{1}{2}\frac{U^{n}_{i,j-1}-2U^{n}_{i,j}+U^{n}_{i,j+1}}{\Delta\nu^2}\\ &+\rho\sigma ij\Delta\nu\Delta S\frac{U^{n}_{i+1,j+1}+U^{n}_{i-1,j-1}-U^{n}_{i-1,j+1}-U^{n}_{i+1,j-1}}{4\Delta S\Delta\nu}\\ &\left.+ri\Delta S\frac{U^{n}_{i+1,j}-U^{n}_{i-1,j}}{2\Delta S} + \kappa(\theta-j\Delta\nu)\frac{U^{n}_{i,j+1}-U^{n}_{i,j-1}}{2\Delta\nu} - rU^{n}_{i,j}\right),\end{aligned} \tag{4.49}$$

where $\theta' \in [0,1]$. We present the following important cases:

- $\theta' = 0 \Rightarrow$ explicit Euler finite difference scheme
- $\theta' = \frac{1}{2} \Rightarrow$ Crank-Nicolson scheme[13]
- $\theta' = 1 \Rightarrow$ Euler fully implicit scheme

[12] This is known as the θ'-Scheme. Note the use of a tilde is in order to distinguish from the Heston model parameter θ representing the level of mean reversion for the variance process. Some authors denote this parameter by η.

[13] Such that the Crank-Nicolson scheme is an average of the explicit and implicit schemes.

4.8 ADI Scheme

To solve multi-dimensional initial-value problems (IVPs) (for example the two dimensional Heston model PDE plus its boundary conditions), we must employ techniques such as the θ'-schemes. In particular, we have seen that whilst easy to implement, the explicit scheme is conditionally stable restricting our choice in asset and time steps. On the other hand, implicit schemes such as the Crank-Nicolson, although unconditionally stable, mean that we must perform several iterations in each time step, which is computationally expensive when dealing with two or more dimensions. For this reason, we turn to alternating-direction implicit (ADI) schemes, which are a type of Operator Splitting Scheme (OPS) first developed by Peaceman and Rachford (1955) [37]. ADI schemes involve both explicit and implicit computations. They are essentially a time splitting method which works by breaking down multi-dimensional problems involving large systems of equations into sub problems treating a given spatial direction individually. They require solving a simple tridiagonal system of equations. They are based on a finite difference representation of the PDE we are solving.

We represent the θ'-scheme in matrix form by the following system

$$(\mathbf{I} - \theta'\Delta t\mathbf{A})\mathbf{U}^{n+1} = (\mathbf{I} + (1-\theta')\Delta t\mathbf{A})\mathbf{U}^{n}, \tag{4.50}$$

where $\mathbf{A}$ is the matrix of coefficients and $\mathbf{I}$ is the identity matrix and both matrices are of size $(N_S-1)(N_V-1) \times (N_S-1)(N_V-1)$[14]. Since the initial condition $\mathbf{U}^0$ is known as this is the option payoff at expiry, we can work from expiry by filling in a matrix with this boundary condition and using the matrix A to then obtain $\mathbf{U}^1$, $\mathbf{U}^2$ and so on until the final value $\mathbf{U}^{N_T}$ is reached. This means for every time step Δt we solve the system (4.50). We note that the initial matrix $\mathbf{U}^0$ will contain entries corresponding to the payoff so that in our European call option example the entries at each node will be $S-K$ whenever $S > K$ and 0 otherwise. The order in which the entries appear in the matrix $\mathbf{U}^0$ depends on the positioning of the components. We will see later on that to solve system (4.50) we must invert the matrix appearing on the LHS such that the system becomes

$$\mathbf{U}^{n+1} = (\mathbf{I} - \theta'\Delta t\mathbf{A})^{-1}(\mathbf{I} + (1-\theta')\Delta t\mathbf{A})\mathbf{U}^{n}. \tag{4.51}$$

From the matrix representation of the θ'-scheme (4.50) we can derive as before three particular cases: when $\theta' = 0$ we obtain the fully explicit scheme, when $\theta' = \frac{1}{2}$ the Crank-Nicolson scheme and when $\theta' = 1$ the fully implicit scheme. At this point we are able to exploit the sparse nature of the matrix A in order to increase computational efficiency. The algorithm chosen for our implementation is the Tridiagonal Matrix Algorithm (also known as the Thomas algorithm), however an ample amount of algorithms have been proposed. For example the line Jacobi method may be applied to general difference schemes in two and three dimensions (see Duffy (2006) [15]). Also, other prominent iterative methods include

- Gauss-Seidel relaxation scheme
- Successive over-relaxation (SOR) scheme

[14] This can be clearly seen during the implementation stage.

- Symmetric successive over-relaxation (SSOR) scheme

The three listed methods above are more efficient than Jacobi methods. For further details one may approach Thomas (1999) [44] for example.

In the ADI scheme, each time step is split into sub-steps. In each sub-step one space direction is treated implicitly whilst the other is treated explicitly. In a subsequent sub-step the next direction is treated implicitly whilst the other is treated explicitly. As the implicit scheme is the one that presents a matrix inversion problem, in every sub-step only one matrix needs to be inverted. We decompose the matrix A into sub-matrices according to the classification presented in Craig and Sneyd (1988) [11]

$$A = A_0 + A_1 + A_2, \tag{4.52}$$

where we have chosen the matrix A_0 to represent the mixed spatial derivative term in (4.2) and in line with literature the matrices A_1 and A_2 correspond to all spatial derivatives in the S and ν directions respectively. The term rU in the PDE (4.2) is evenly distributed into the matrices A_1 and A_2. To this end we have

$$A_0 = \frac{1}{4}\rho\sigma\nu S\frac{\Delta t}{\Delta S\Delta\nu}\delta^2_{S\nu} \tag{4.53a}$$

$$A_1 = \frac{1}{2}\nu S^2\frac{\Delta t}{\Delta S^2}\delta^2_S + \frac{1}{2}rS\frac{\Delta t}{\Delta S}\delta_S - \frac{1}{2}r\Delta t \tag{4.53b}$$

$$A_2 = \frac{1}{2}\sigma^2\nu\frac{\Delta t}{\Delta\nu^2}\delta^2_\nu + \frac{1}{2}\kappa(\theta-\nu)\frac{\Delta t}{\Delta\nu}\delta_\nu - \frac{1}{2}r\Delta t, \tag{4.53c}$$

where as before

$$\delta_S U_{i,j} = U_{i+1,j} - U_{i-1,j} \tag{4.54a}$$

$$\delta^2_S U_{i,j} = U_{i-1,j} - 2U_{i,j} + U_{i+1,j} \tag{4.54b}$$

$$\delta^2_{S\nu} U_{i,j} = U_{i+1,j+1} + U_{i-1,j-1} - U_{i-1,j+1} - U_{i+1,j-1} \tag{4.54c}$$

$$\delta_\nu U_{i,j} = U_{i,j+1} - U_{i,j-1} \tag{4.54d}$$

$$\delta^2_\nu U_{i,j} = U_{i,j-1} - 2U_{i,j} + U_{i,j+1}. \tag{4.54e}$$

With this notation we are able to write the θ'-scheme in terms of A_0, A_1 and A_2 with respect to the node (S_i, ν_j, t_n) as follows

$$(1-\theta'\Delta tA_1-\theta'\Delta tA_2)U^{n+1}_{i,j} = \ (1+A_0\Delta t+(1-\theta')\Delta tA_1+(1-\theta'\Delta tA_2)U^n_{i,j}. \tag{4.55}$$

Next we use approximate factorization to reduce the complexity of the linearized scheme. First note that $(1-\theta\Delta tA_1)(1-\theta\Delta tA_2) = 1-\theta\Delta tA_1-\theta\Delta tA_2+\theta^2\Delta t^2A_1A_2$. Then add $\theta^2\Delta t^2A_1A_2U^{n+1}$ to both sides of (4.55) to obtain

$$\begin{aligned}(1 - \theta\Delta tA_1 - \theta\Delta tA_2+\theta^2\Delta t^2A_1A_2)U^{n+1}_{i,j} = \\ (1 + \Delta tA_0 + (1-\theta)\Delta tA_1 + (1-\theta)\Delta tA_2 + \theta^2\Delta t^2A_1A_2)U^n_{i,j} \\ + \underbrace{\theta^2\Delta t^2A_1A_2(U^{n+1}_{i,j} - U^n_{i,j})}_{\mathcal{O}(\Delta t^3)} + \mathcal{O}(\Delta t^3).\end{aligned} \tag{4.56}$$

We can combine the term $\theta^2\Delta t^2A_1A_2(U^{n+1} - U^n) \sim \mathcal{O}(\Delta t^3)$ with the error term maintaining the order of accuracy[15]. Therefore, the scheme is still second order in

[15]Note that $U^{n+1}_{i,j} - U^n_{i,j} \sim \mathcal{O}(\Delta t)$ so that the term $\theta^2\Delta t^2A_1A_2(U^{n+1} - U^n) \sim \mathcal{O}(\Delta t^3)$.

the time and spatial variables. Factorizing gives us

$$\begin{aligned}(1-\theta\Delta tA_1)(1-\theta\Delta tA_2)U_{i,j}^{n+1} = (1+\Delta tA_0+(1-\theta)\Delta tA_1+(1-\theta)\Delta tA_2 \\ +\theta^2\Delta t^2A_1A_2)U_{i,j}^n+\mathcal{O}(\Delta t^3)\end{aligned} \tag{4.57}$$

In practice, the ADI scheme performs a single fully explicit step. It then performs a correction step such that it covers one direction implicitly whilst covering the others explicitly. In a subsequent step, another direction is done implicitly whilst the rest explicitly and so on. So the method is composed of intermediate steps known as the correction steps. The idea is to improve efficiency and computation time whilst retaining the properties of stability and consistency.

In general, ADI schemes were not developed to cope with a mixed derivative term as is present in the Heston PDE and were thus accordingly adapted (see In't Hout and Foulon [27]).

4.8.1 Douglas-Rachford Scheme

The simplest of the ADI schemes is the Douglas-Rachford (DR) scheme[16] described in detail in [14]. The scheme is composed of the following steps[17]

$$\frac{U_{i,j}^*-U_{i,j}^n}{\Delta t} = \theta'A_1U_{i,j}^*+A_0U_{i,j}^n+(1-\theta')A_1U_{i,j}^n+A_2U_{i,j}^n \tag{4.58a}$$

$$\frac{U_{i,j}^{n+1}-U_{i,j}^*}{\Delta t} = \theta'A_2(U_{i,j}^{n+1}-U_{i,j}^n). \tag{4.58b}$$

This may be written in the more common computational form as follows

$$\begin{aligned}(1-\theta'\Delta tA_1)U_{i,j}^* &= (1+\Delta tA_0+(1-\theta')\Delta tA_1+\Delta tA_2)U_{i,j}^n \\ (1-\theta'\Delta tA_2)U_{i,j}^{n+1} &= U_{i,j}^*-\theta'\Delta tA_2U_{i,j}^n.\end{aligned} \tag{4.59}$$

The Douglas-Rachford time splitting scheme is unconditionally stable whenever $\theta' \geq \frac{1}{2}$. The scheme is second order accurate in time and spatial variables.

4.8.2 The Craig and Sneyd Scheme

Further ADI schemes stem from the Douglas scheme with the aim of improving accuracy and stability. The Craig and Sneyd (CS) scheme [11] was developed in 1988 in an attempt to improve accuracy and stability of the DR scheme. The CS iterative scheme is second-order accurate in time and space and is unconditionally stable whenever $\theta' \geq \frac{1}{2}$, as in the Douglas-Rachford scheme. The scheme is shown to be less effective for higher dimensional problems due to the presence of the mixed derivative term.

[16] Or just Douglas scheme.

[17] The DR scheme is commonly written with $U_{i,j}^{n+\frac{1}{2}}$ in place of $U_{i,j}^*$, either option can be understood as a sort of 'half-step'.

Craig and Sneyd:

$$(1-\theta'\Delta tA_1)Y_1 = (1+\Delta tA_0+(1-\theta')\Delta tA_1+\Delta tA_2)U^n \tag{4.60a}$$
$$(1-\theta'\Delta tA_2)Y_2 = Y_1-\theta'\Delta tA_2U^n \tag{4.60b}$$
$$(1-\theta'\Delta tA_1)Y_3 = (1-\theta'\Delta tA_1)Y_1+\frac{1}{2}\Delta tA_0(Y_2-U^n) \tag{4.60c}$$
$$(1-\theta'\Delta tA_2)Y_4 = Y_3-\theta'\Delta tA_2U^n \tag{4.60d}$$
$$U^{n+1} = Y_4. \tag{4.60e}$$

4.8.3 Modified Craig and Sneyd Scheme

The full modified Craig and Sneyd (MCS) scheme step-by-step is outlined first, then the concise Matlab implementation outline is presented.

Modified Craig and Sneyd:

$$(1-\theta'\Delta tA_1)Y_1 = (1+\Delta tA_0+(1-\theta')\Delta tA_1+\Delta tA_2)U^n \tag{4.61a}$$
$$(1-\theta'\Delta tA_2)Y_2 = Y_1-\theta'\Delta tA_2U^n \tag{4.61b}$$
$$\begin{aligned} Y_3 &= Y_0+\theta'\Delta tA_0(Y_2-U^n)\\ &= (1-\theta'\Delta tA_1)Y_1+\theta'\Delta tA_1U^n+\theta'\Delta tA_0(Y_2-U^n) \end{aligned} \tag{4.61c}$$
$$Y_4 = Y_3+(\frac{1}{2}-\theta')\Delta tA(Y_2-U^n) \tag{4.61d}$$
$$\begin{aligned} (1-\theta'\Delta tA_1)Y_5 &= Y_4-\theta'\Delta tA_1U^n\\ &= (1-\theta'\Delta tA_1)Y_1+\cancel{\theta'\Delta tA_1U^n}+\theta'\Delta tA_0(Y_2-U^n)\\ &\quad +(\frac{1}{2}-\theta')\Delta tA(Y_2-U^n)-\cancel{\theta'\Delta tA_1U^n} \end{aligned} \tag{4.61e}$$
$$(1-\theta'\Delta tA_2)Y_6 = Y_5-\theta'\Delta tA_2U^n \tag{4.61f}$$
$$U^{n+1} = Y_6. \tag{4.61g}$$

The concise version as computed in Matlab is as follows.

$$(1-\theta'\Delta tA_1)Y_1 = (1+\Delta tA_0+(1-\theta')\Delta tA_1+\Delta tA_2)U^n \tag{4.62a}$$
$$(1-\theta'\Delta tA_2)Y_2 = Y_1-\theta'\Delta tA_2U^n \tag{4.62b}$$
$$(1-\theta'\Delta tA_1)Y_3 = (1-\theta'\Delta tA_1)Y_1+\theta'\Delta tA_0(Y_2-U^n)+(\frac{1}{2}-\theta')\Delta tA(Y_2-U^n) \tag{4.62c}$$
$$(1-\theta'\Delta tA_2)Y_4 = Y_3-\theta'\Delta tA_2U^n \tag{4.62d}$$
$$U^{n+1} = Y_4. \tag{4.62e}$$

Notice that when $\theta' = \frac{1}{2}$ we obtain the Craig and Sneyd scheme.

4.8.4 The Hundsdorfer and Verwer Scheme

The Hundsdorfer and Verwer (HV) scheme was originally developed to obtain a numerical solution to convection-diffusion-reaction equations with applications to atmospheric dispersion problems, see [46] (one can also see a review of the book by Hundsdorfer and Verwer in [26]).

Hundsdorfer and Verwer:

$$(1-\theta'\Delta tA_1)Y_1 = (1+\Delta tA_0+(1-\theta')\Delta tA_1+\Delta tA_2)U^n \tag{4.63a}$$
$$(1-\theta'\Delta tA_2)Y_2 = Y_1-\theta'\Delta tA_2U^n \tag{4.63b}$$
$$(1-\theta'\Delta tA_1)Y_3 = (1-\theta'\Delta tA_1)Y_1+\mu(\Delta tA_0+(1-\theta')\Delta tA_1)(Y_2-U^n) \tag{4.63c}$$
$$(1-\theta'\Delta tA_2)Y_4 = Y_3-\theta'\Delta tA_2U^n \tag{4.63d}$$
$$U^{n+1} = Y_4. \tag{4.63e}$$

When implementing the HV scheme, the solution matrices at expiry (for which the function gives an output) are very similar. However if we type `format long`[18] into the Matlab command line, it can be seen that they are actually different. An illustration of the option price grid at $N_T = 4000$ for different ADI schemes is given next, with $\theta' = \frac{1}{2}$ for DR, CS and MCS schemes and $\theta' = \frac{1}{2}+\frac{\sqrt{3}}{6}$ for the HV scheme.

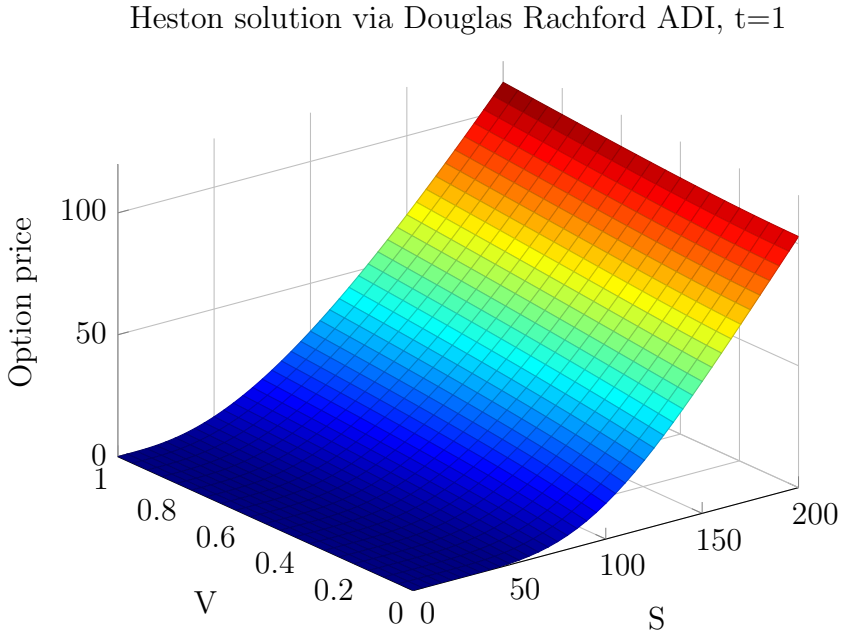


Figure 4.3: Illustration of the Douglas-Rachford ADI scheme applied to the Heston model.

[18] This extends the accuracy up to 15dp.

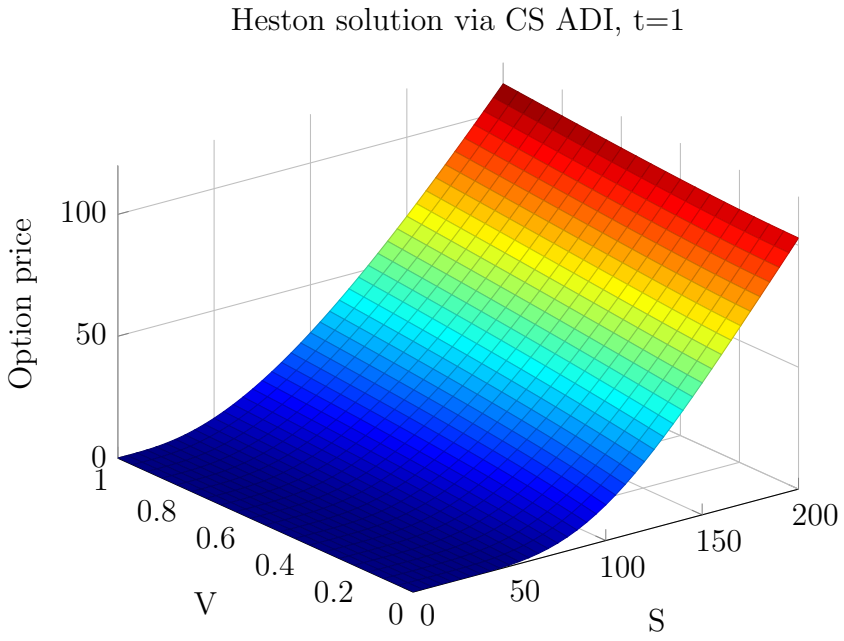


Figure 4.4: Illustration of the Craig-Sneyd ADI scheme applied to the Heston model.

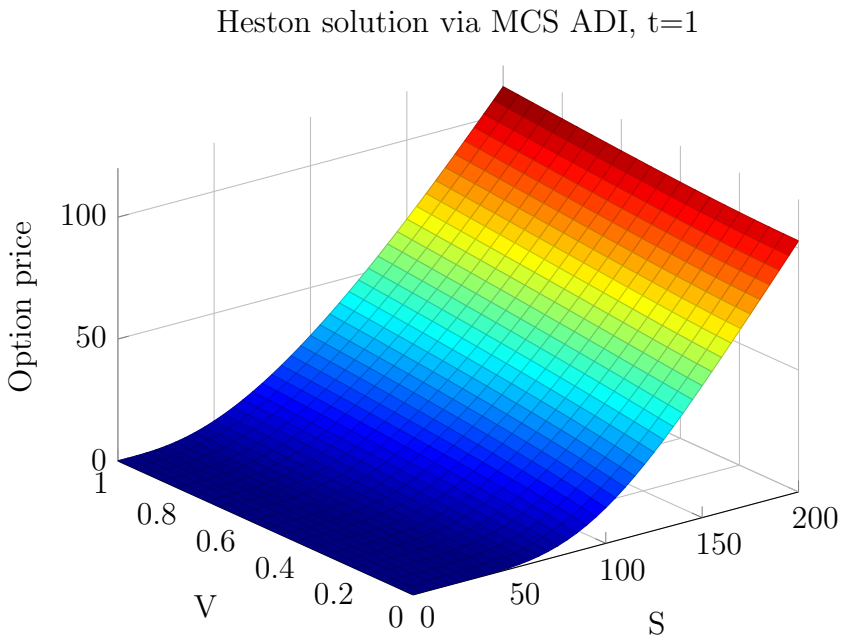


Figure 4.5: Illustration of the Modified Craig-Sneyd ADI scheme applied to the Heston model.

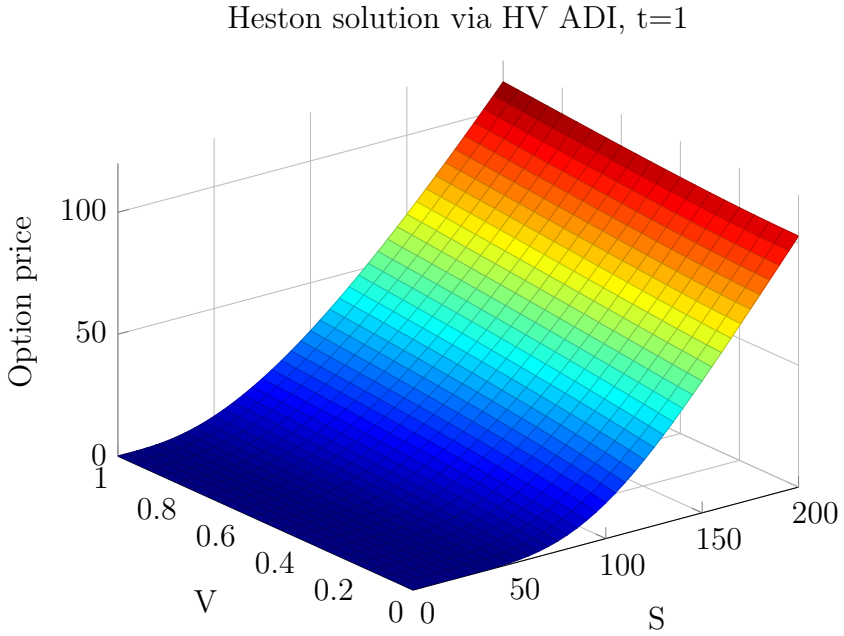


Figure 4.6: Illustration of the Hundsdorfer-Verwer ADI scheme applied to the Heston model.

4.9 ADI Implementation

We describe the procedure for implementing the Douglas-Rachford scheme. The other schemes are similar but with added predictor and corrector steps. In the Matlab computational implementation we maintain common functions for all schemes for construction of the matrices A_0, A_1 and A_2 and matrix inversions. We start by constructing the sub-matrices of the matrix A. Recall equations (4.53a)-(4.53c). The matrix A_0 is the first to be constructed, this is easy as this matrix is only dealt with explicitly.

The first step of the method consists of solving the first equation in the system (4.59). If we substitute (4.53b) into the LHS of that first equation we obtain the following

$$U^*_{i,j} - \theta'\Delta t(\frac{1}{2}\nu_j S_i^2 \frac{\Delta t}{\Delta S^2}\left(U^*_{i+1,j} - 2U^*_{i,j} + U^*_{i-1,j}\right) + \frac{1}{2} r S_i \frac{\Delta t}{\Delta S}\left(U^*_{i+1,j} - U^*_{i-1,j}\right) - \frac{1}{2} r \Delta t U^*_{i,j}) = (1 + \Delta t A_0 + (1-\theta)\Delta t A_1 + \Delta t A_2) U^n_{i,j}, \tag{4.64}$$

and regrouping common terms on the LHS we obtain the coefficients

$$a_{i,j} = -\frac{1}{2}\theta'\nu_j S_i^2 \frac{\Delta t}{\Delta S^2} + \frac{1}{2}\theta' r S_i \frac{\Delta t}{\Delta S} \tag{4.65a}$$

$$b_{i,j} = 1 + \theta'\nu_j S_i^2 \frac{\Delta t}{\Delta S^2} + \frac{1}{2}\theta' r \Delta t \tag{4.65b}$$

$$c_{i,j} = -\frac{1}{2}\theta'\nu_j S_i^2 \frac{\Delta t}{\Delta S^2} - \frac{1}{2}\theta' r S_i \frac{\Delta t}{\Delta S}, \tag{4.65c}$$

such that

$$a_{i,j}U^*_{i-1,j} + b_{i,j}U^*_{i,j} + c_{i,j}U^*_{i+1,j} = (1 + \Delta t A_0 + (1-\theta)\Delta t A_1 + \Delta t A_2) U^n_{i,j}. \tag{4.66}$$

Fix j. Then the first step in the DR scheme (4.59) can be expressed as the system

$$\mathbf{L}_{1,j}U_j^* = \mathbf{q}_j,\ j = 1, ..., N_V - 1, \tag{4.67}$$

where

$$\mathbf{L}_{1,j} = \begin{pmatrix} b_{1,j} & c_{1,j} & 0 & \dots & \dots & 0 \\ a_{2,j} & b_{2,j} & c_{2,j} & & & \\ 0 & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{N_S-2,j} & b_{N_S-2,j} & c_{N_S-2,j} \\ 0 & \dots & \dots & 0 & a_{N_S-1,j} & b_{N_S-1,j} \end{pmatrix}, U_j^* = \begin{pmatrix} U_{1,j}^* \\ U_{2,j}^* \\ \vdots \\ \vdots \\ U_{N_S-2,j}^* \\ U_{N_S-1,j}^* \end{pmatrix},$$

$$\mathbf{q}_j = \begin{pmatrix} (1 + \Delta t A_0 + (1-\theta')\Delta t A_1 + \Delta t A_2)U_{1,j}^n - a_{1,j}U_{0,j}^* \\ (1 + \Delta t A_0 + (1-\theta')\Delta t A_1 + \Delta t A_2)U_{2,j}^n \\ \vdots \\ \vdots \\ (1 + \Delta t A_0 + (1-\theta')\Delta t A_1 + \Delta t A_2)U_{N_S-2,j}^n \\ (1 + \Delta t A_0 + (1-\theta')\Delta t A_1 + \Delta t A_2)U_{N_S-1,j}^n - c_{N_s,j}U_{N_S,j}^* \end{pmatrix}$$

Note that $U_{0,j}^*$ is known as this is the boundary condition at $S = 0$. Similarly. the term $U_{N_S,j}^*$ is also known as this is the boundary condition at $S = S_{max}$. We move these terms onto the right hand side so that the matrix $L_{1,j}$ is tridiagonal, thus presenting us with a tridiagonal matrix inversion for which there are ample algorithms to solve such problems.

Similarly, substituting the matrix A_2 into the LHS of the second step we have

$$U_{i,j}^{n+1} - \theta'\Delta t \left(\frac{1}{2}\sigma^2\nu_j \frac{\Delta t}{\Delta\nu^2}\left(U_{i,j+1}^{n+1} - 2U_{i,j}^{n+1} + U_{i,j-1}^{n+1}\right) + \frac{1}{2}\frac{\Delta t}{\Delta\nu}\kappa(\theta - \nu_j)(U_{i,j+1}^{n+1} - U_{i,j-1}^{n+1}) \right.$$
$$\left. - \frac{1}{2}r\Delta t U_{i,j}^{n+1} \right) = U_{i,j}^* - \theta'\Delta t A_2 U_{i,j}^n, \tag{4.68}$$

where $U_{i,j}^*$ is the matrix resulting from the inversion in step one. The coefficients this time are given by

$$a'_{i,j} = -\frac{1}{2}\theta'\sigma^2\nu_j\frac{\Delta t}{\Delta\nu^2} + \frac{1}{2}\theta'\kappa(\theta - \nu_j)\frac{\Delta t}{\Delta\nu} \tag{4.69a}$$

$$b'_{i,j} = 1 + \theta'\sigma^2\nu_j\frac{\Delta t}{\Delta\nu^2} + \frac{1}{2}\theta' r\Delta t \tag{4.69b}$$

$$c'_{i,j} = -\frac{1}{2}\theta'\sigma^2\nu_j\frac{\Delta t}{\Delta\nu^2} - \frac{1}{2}\theta'\kappa(\theta - \nu_j)\frac{\Delta t}{\Delta\nu}, \tag{4.69c}$$

such that

$$a'_{i,j}U_{i,j-1}^{n+1} + b'_{i,j}U_{i,j}^{n+1} + c'_{i,j}U_{i,j+1}^{n+1} = U_{i,j}^* - \theta'\Delta t U_{i,j}^n. \tag{4.70}$$

Fix i. Then the second step of the DR scheme may be expressed as the system

$$\mathbf{L}_{2,i}U_i^{n+1} = \mathbf{q}_i,\ i = 1, ..., N_S - 1, \tag{4.71}$$

where

$$\mathbf{L}_{2,i} = \begin{pmatrix} b'_{i,1} & c'_{1,1} & 0 & \dots & \dots & 0 \\ a'_{i,2} & b'_{i,2} & c'_{i,2} & & & \\ 0 & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & a'_{i,N_V-2} & b'_{i,N_V-2} & c'_{i,N_V-2} \\ 0 & \dots & \dots & 0 & a'_{i,N_V-1} & b'_{i,N_V-1} \end{pmatrix}, U_i^{n+1} = \begin{pmatrix} U_{i,1}^{n+1} \\ U_{i,2}^{n+1} \\ \vdots \\ \vdots \\ U_{i,N_V-2}^{n+1} \\ U_{i,N_V-1}^{n+1} \end{pmatrix},$$

$$\mathbf{q}_i = \begin{pmatrix} U_{i,1}^* - \theta'\Delta t U_{i,1}^n - a'_{i,1}U_{i,0}^{n+1} \\ U_{i,2}^* - \theta'\Delta t U_{i,2}^n \\ \vdots \\ \vdots \\ U_{i,N_V-2}^* - \theta'\Delta t U_{i,N_V-2}^n \\ U_{i,N_V-1}^* - \theta'\Delta t U_{i,N_V-1}^n - c'_{i,N_V}U_{i,N_V}^{n+1} \end{pmatrix}$$

As before, we have moved the known terms $U_{i,0}^{n+1}$ (boundary condition at $\nu = 0$) and U_{i,N_V}^{n+1} (boundary condition at $\nu = \nu_{max}$) to the right hand side and again we have a tridiagonal matrix inversion problem. We use the tridiagonal matrix algorithm (also known as the Thomas algorithm) to solve each matrix inversion problem. If the coefficients were constant the matrices could be LU-factorized to simplify the algorithm. An outline of the Thomas algorithm is given in Appendix A. In the Matlab implementation we multiply by Δt within the subprogram for the construction of each matrix for simplicity.

4.10 The Greeks

The aim of this section is to introduce the different option price sensitivities, known as the *Greeks*. The Greeks are a measure of the option price changes with a change in one of the input parameters, *ceteris paribus*. They are extremely important in financial risk management as they are hedging tools. In practice, we are interested in how an option price changes with a change in a market parameter such as implied volatility or interest rate. The Greeks are essentially partial derivatives and we can therefore compute them via the finite difference methods developed in this chapter. Table 4.1 shows the different Greeks and their respective meaning.

In this dissertation, we apply the different finite difference schemes to Δ, Γ and *vega*. Note that all the greeks in Table 4.1 are first order derivatives except for $\Gamma = \frac{\partial^2 U}{\partial S^2}$, which does not measure a change in option price but instead a change

Greek	Derivative	Measures change in
Delta	$\Delta = \frac{\partial U}{\partial S}$	the option price when underlying price increases by 1 unit
Gamma	$\Gamma = \frac{\partial \Delta}{\partial S}$	delta when underlying price increases by 1 unit
Vega	$\frac{\partial U}{\partial \sigma}$	the option price when volatility increases by 1%
Theta	$\Theta = \frac{\partial U}{\partial t}$	the option price when time to expiry decreases by 1 day
Rho	$\rho = \frac{\partial U}{\partial r}$	the option price when interest rate increases by 1%
Psi	$\Psi = \frac{\partial U}{\partial q}$	the option price when dividend yield increases by 1%

Table 4.1: Table of option price sensitivities.

in Δ. Out of all the greeks, Θ is the only one in the negative domain as it is a measure of decreasing time.

Delta Delta is defined as the change in the price of the derivative with repect to a change in the underlying asset. The delta of a portfolio of derivatives on the same underlying asset S is the sum of the deltas of the individual options. If a portfolio has a delta of zero then it is said to be *delta-neutral*. In this case, since the portfolio is not exposed to *small* changes in the option price, it is usually used for hedging purposes. For example, one may construct a delta-neutral portfolio by taking a long position in a call option and a short position in Δ amount of the underlying stock. This is called a Δ-hedge and is, for example, the argument composed originally by Black and Scholes to derive the Black-Scholes equation. Since delta tells us how much stock to hold in order to hedge our exposure to risk, it is regarded as one of the most important greeks.

As an example, the central difference used to evaluate delta computationally is as follows

$$\Delta \coloneqq \frac{\partial U_{i,j}^n}{\partial S} = \frac{U_{i+1,j}^n - U_{i-1,j}^n}{2\Delta S}. \tag{4.72}$$

Gamma In practice, its common to refer to the gamma of a portfolio, which is the rate of change of the portfolio's delta with respect to a change in the underlying asset. So we can write gamma as the derivative of delta as in Table 4.1 or we can write it directly as the second derivative of the portfolio Π with respect to the asset S as $\frac{\partial^2 \Pi}{\partial S^2}$. Gamma is a measure of how sensitive delta is to a change in the underlying asset. This means that for a large gamma, our delta-neutral portfolio will need to be re-balanced more often as delta is more sensitive to *small* changes in the underlying. Therefore, gamma is a measure of the degree of risk exposure that a hedged position will have if the current hedge is not adjusted. If gamma is small, then delta is relatively insensitive to option price changes in which case re-balancing will be an infrequent task. So, in effect, gamma tells us how often and how much we need to adjust the delta of a portfolio of derivatives to ensure risk exposure to underling asset price changes is hedged. For large movements in the underlying asset price, delta is not an accurate measure of how the option price changes. This is because of the curvature that is left after delta hedging has been done. The delta hedge is used to eliminate exposure to linear risk and gamma is used to measure the remaining risk.

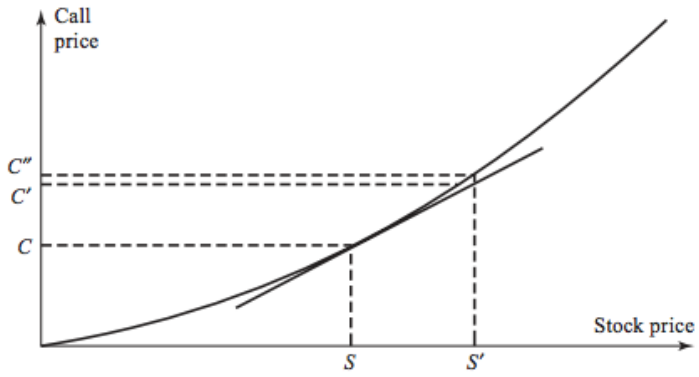


Figure 4.7: Curvature risk left after linear risk is hedged.

Figure 4.7[19] acts as an example of this remaining exposure. As the stock price moves from S to S', delta hedging assumes the price of the derivative moves from C to C', but actually the move is from C to C''. This difference between C and C'' will lead to an error in the hedge. The actual size of this error will be determined by the degree of curvature and this degree is measured by gamma.

Gamma can also be used to find information regarding transaction costs. If there are transaction costs, then by holding delta, a change in delta could mean for example purchase of more stock is required and since gamma is the change in delta, it will indicate how much to buy or sell and consequently the size of transaction costs if present.

Theta The theta of a portfolio of derivatives measures the rate of change of the portfolio's value with respect to time. As mentioned previously, theta tends to be in the negative domain. This is because as time to maturity approaches zero the option loses value. As a matter of fact, theta is in the positive domain only for European puts deep in the money. This is because a put can be worth less than its *intrinsic* value but must eventually increase to match its intrinsic value at expiration.

Vega The vega of a portfolio of derivatives is the rate of change of the portfolio's value with respect to the volatility of the underlying asset. Vega is of course not a greek letter, so in literature it is sometimes referred to by the greek letters ν, λ or κ. A derivative with a large vega is very sensitive to *small* changes in volatility. Vega is typically positive, an observation that coincides with the intuition that a rise in volatility will increase the value of an option.

There is a further approach to solving the Heston model developed by Shaw [42]. The outline of this approach is given in Appendix A, the details are omitted but these are very similar to those in the derivation of the closed form solution as in the original paper by Heston (1993) [24]. The advantage to the approach by Shaw is the computational simplicity of the greeks. For this reason, which is further explained in Appendix A, we consider this approach to evaluate the closed form version of the greeks. The Matlab code used to produce the results is based on

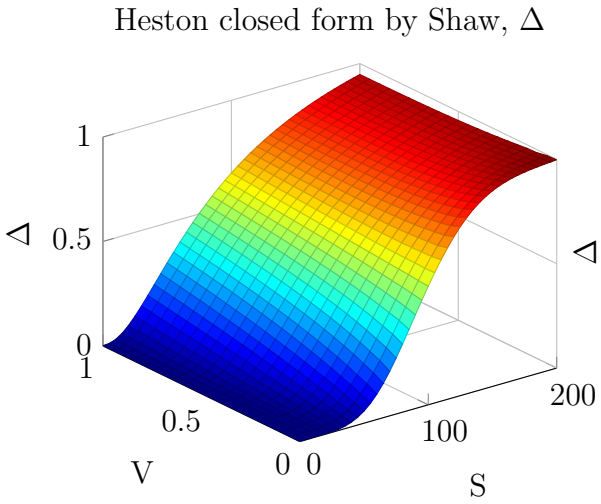
[19]The source of this image is Hull (2012) [25].

the graduate project by Lin [32]. Following are the plots of Δ, Γ and Vega as per Shaw's approach.

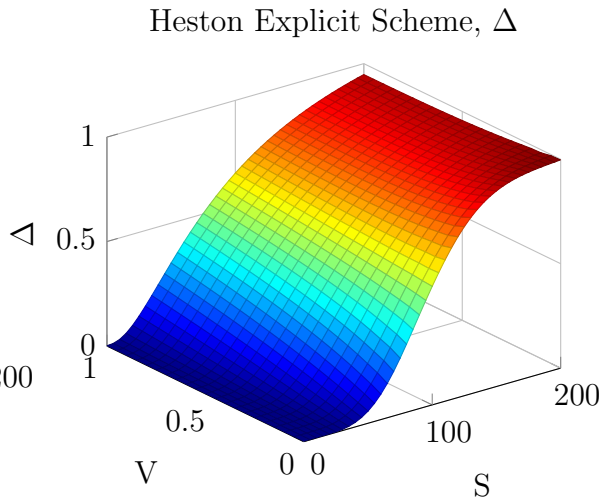
Next we present grids for the Greek sensitivities as evaluated via Shaw's solution [42]. The Greeks evaluated via ADI schemes are very similar (found in the Matlab implementation separately). Below we outline the parameter values used to produce the plots pictured after.

κ	θ	σ	ρ	r	S	V	T	K	N_S	N_V	N_T
2	0.2	0.3	0.8	0.03	200	1	1	100	40	20	4000

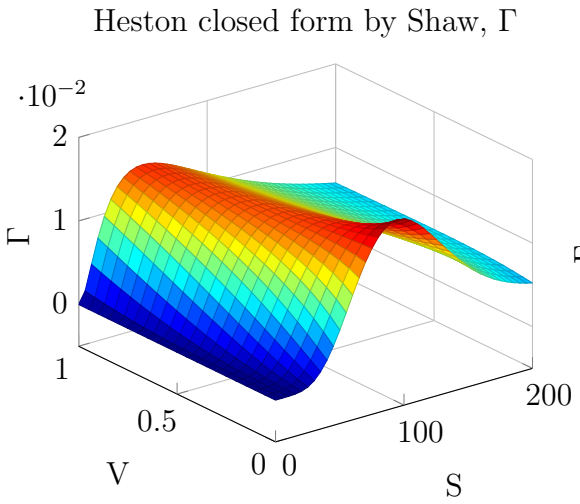
Table 4.2: Table with parameter values.



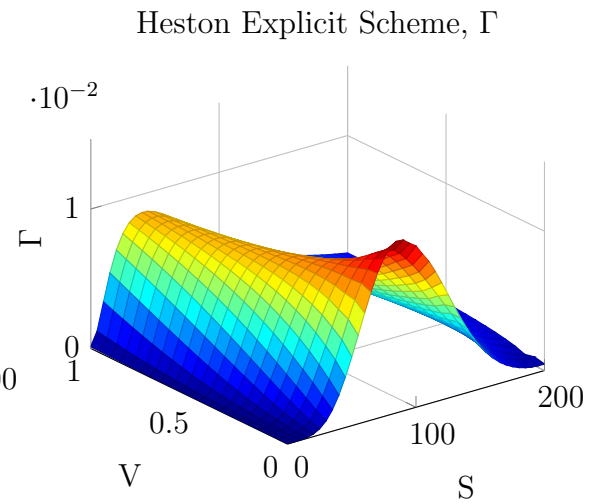
(a) Analytical solution of Δ of the Heston model, evaluated via Shaw's approach.



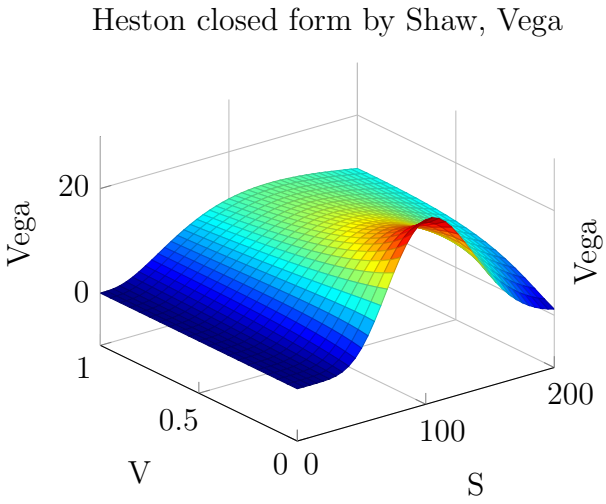
(b) Δ of the Heston model, evaluated with the Explicit scheme.



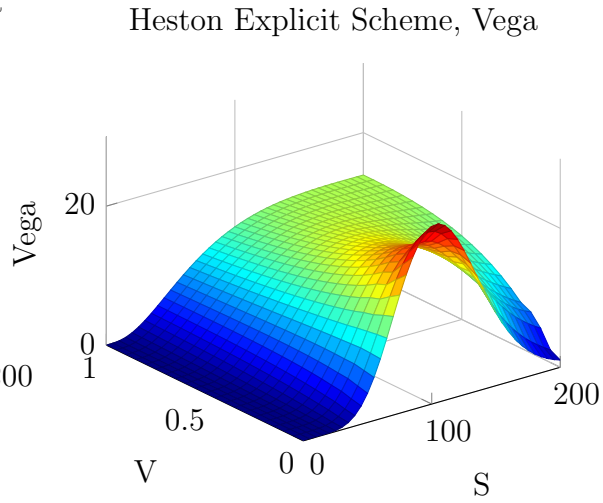
(a) Analytical solution of Γ of the Heston model, evaluated via Shaw's approach.



(b) Γ of the Heston model, evaluated with the Explicit scheme.



(a) Analytical solution of Vega of the Heston model, evaluated via Shaw's approach.



(b) Vega of the Heston model, evaluated with the Explicit scheme.

Following are plots of the difference between the analytical and finite difference approximations to each greek Δ, Γ and Vega.

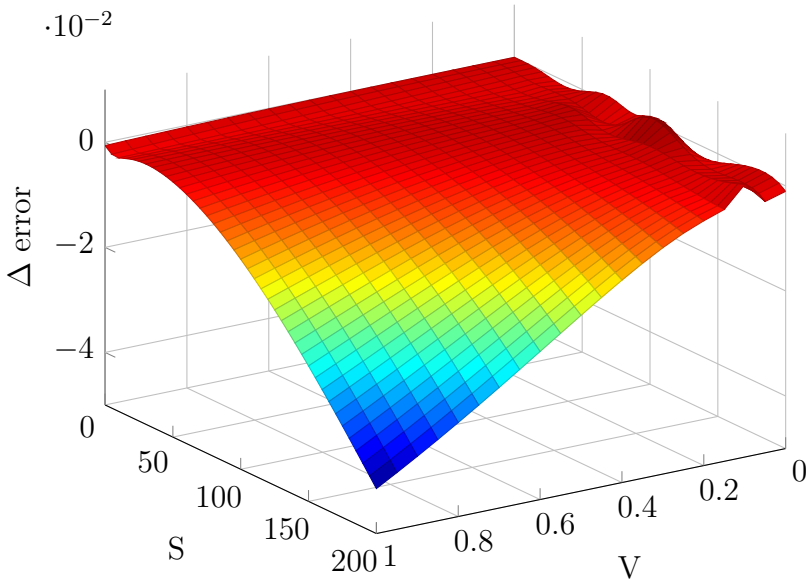


Figure 4.11: Difference between analytical and finite difference approximations to Δ in the Heston model.

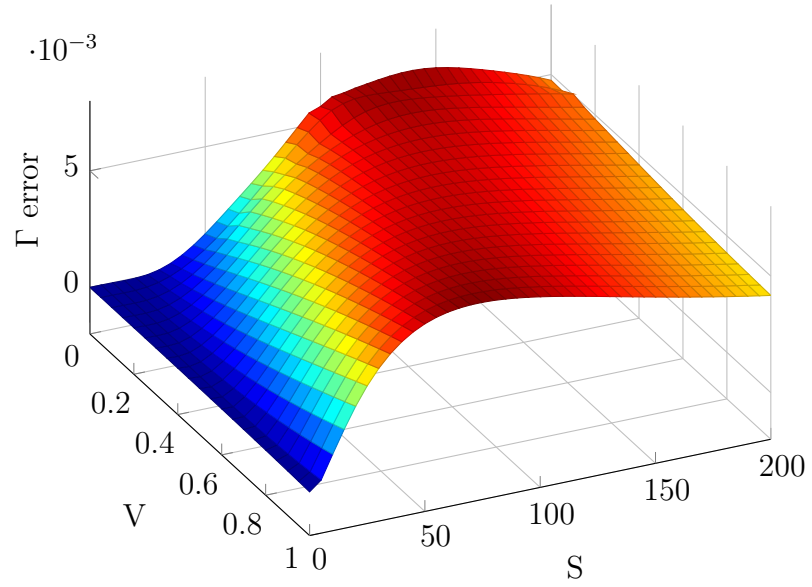


Figure 4.12: Difference between analytical and finite difference approximations to Γ in the Heston model.

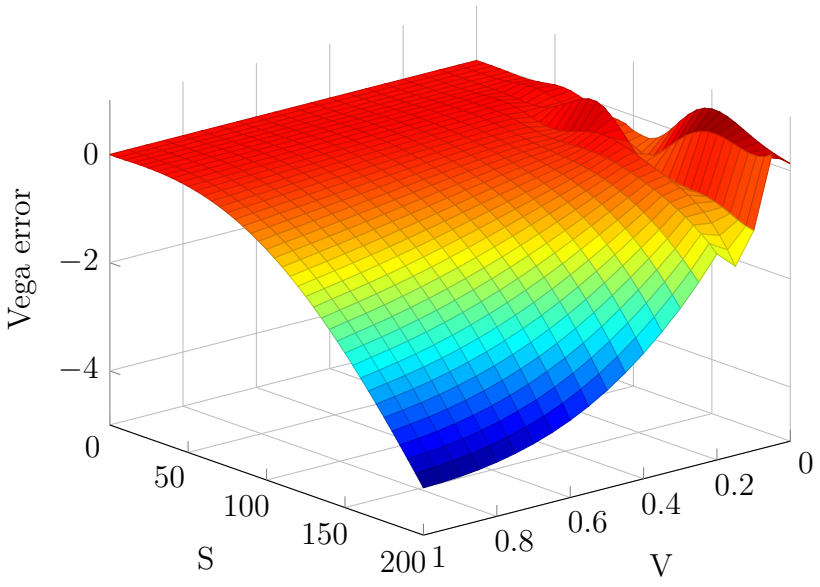


Figure 4.13: Difference between analytical and finite difference approximations to Vega in the Heston model.

4.11 Finite Differences in the Non-Uniform Case

In the non-uniform case the increments in spatial variables are not equidistant. To derive the finite difference scheme we proceed in a similar manner to the approach taken in the uniform case. This time, we leave out the details and present the

backward, central and forward difference schemes directly. These are outlined below for a general spatial variable x.

Backward

$$f'(x_i) \approx \alpha_{i,-2} f(x_{i-2}) + \alpha_{i,-1} f(x_{i-1}) + \alpha_{i,0} f(x_i) \tag{4.73}$$

Central

$$f'(x_i) \approx \beta_{i,-1} f(x_{i-1}) + \beta_{i,0} f(x_i) + \beta_{i,1} f(x_{i+1}) \tag{4.74}$$

Forward

$$f'(x_i) \approx \gamma_{i,0} f(x_i) + \gamma_{i,1} f(x_{i+1}) + \gamma_{i,2} f(x_{i+2}) \tag{4.75}$$

where the coefficients are given by

$$\alpha_{i,-2} = \frac{\Delta x_i}{\Delta x_{i-1}(\Delta x_{i-1} + \Delta x_i)},\ \alpha_{i,-1} = \frac{-\Delta x_{i-1} - \Delta x_i}{\Delta x_{i-1}\Delta x_i},\ \alpha_{i,0} = \frac{\Delta x_{i-1} + 2\Delta x_i}{\Delta x_i(\Delta x_{i-1} + \Delta x_i)}$$

$$\beta_{i,-1} = \frac{-\Delta x_{i+1}}{\Delta x_i(\Delta x_i + \Delta x_{i+1})},\ \beta_{i,0} = \frac{\Delta x_{i+1} - \Delta x_i}{\Delta x_i \Delta x_{i+1}},\ \beta_{i,1} = \frac{\Delta x_i}{\Delta x_{i+1}(\Delta x_i + \Delta x_{i+1})}$$

$$\gamma_{i,0} = \frac{-2\Delta x_{i+1} - \Delta x_{i+2}}{\Delta x_{i+1}(\Delta x_{i+1} + \Delta x_{i+2})},\ \gamma_{i,1} = \frac{\Delta x_{i+1} + \Delta x_{i+2}}{\Delta x_{i+1}\Delta x_{i+2}},\ \gamma_{i,2} = \frac{-\Delta x_{i+1}}{\Delta x_{i+2}(\Delta x_{i+1} + \Delta x_{i+2})}$$

The second order derivative may be approximated by

$$f''(x_i) \approx \delta_{i,-1} f(x_{i-1}) + \delta_{i,0} f(x_i) + \delta_{i,1} f(x_{i+1}), \tag{4.76}$$

where the coefficients are given by

$$\delta_{i,-1} = \frac{2}{\Delta x_i(\Delta x_i + \Delta x_{i+1})},\ \delta_{i,0} = \frac{-2}{\Delta x_i \Delta x_{i+1}},\ \delta_{i,1} = \frac{2}{\Delta x_{i+1}(\Delta x_i + \Delta x_{i+1})}.$$

The mixed derivative term in the Heston PDE is approximated using the following central scheme (composed of nine summands)

$$\begin{aligned}\frac{\partial^2 f}{\partial x \partial y}(x_i, y_j) &\approx \sum_{k,l=-1}^{1} \beta_{i,k}\widehat{\beta}_{j,l} f(x_{i+k}, y_{j+l}) \\ &= \beta_{i,0}\widehat{\beta}_{j,0} f(x_i, y_j) + \beta_{i,-1}\widehat{\beta}_{j,0} f(x_{i-1}, y_j) + \beta_{i,+1}\widehat{\beta}_{j,0} f(x_{i+1}, y_j) \\ &\quad + \beta_{i,0}\widehat{\beta}_{j,-1} f(x_i, y_{j-1}) + \beta_{i,0}\widehat{\beta}_{j,+1} f(x_i, y_{j+1}) + \beta_{i,-1}\widehat{\beta}_{j,-1} f(x_{i-1}, y_{j-1}) \\ &\quad + \beta_{i,-1}\widehat{\beta}_{j,+1} f(x_{i-1}, y_{j+1}) + \beta_{i,+1}\widehat{\beta}_{j,-1} f(x_{i+1}, y_{j-1}) + \beta_{i,+1}\widehat{\beta}_{j,+1} f(x_{i+1}, y_{j+1}),\end{aligned} \tag{4.77}$$

where $\widehat{\beta}_{j,k}$ is the coefficient that is analogous to $\beta_{i,k}$ but in the y-direction (when applied to our problem this will be the variance direction) rather than x-direction.

The FD formulas (4.74), (4.76) and (4.77) represent central schemes whereas formulas (4.73) and (4.75) are *upwind* schemes. Upwind schemes are one-sided discretizations which may be used on boundaries or for instance, to correct instabilities that sometimes arise as a result of central schemes.

j \ i	-1	0	1
-1	$\frac{-\Delta S_{i+1}}{\Delta S_i(\Delta S_i+\Delta S_{i+1})}$ $\frac{-\Delta\nu_{j+1}}{\Delta\nu_j(\Delta\nu_j+\Delta\nu_{j+1})}$	$\frac{-\Delta S_i+\Delta S_{i+1}}{\Delta S_i\Delta S_{i+1}}$ $\frac{-\Delta\nu_{j+1}}{\Delta\nu_j(\Delta\nu_j+\Delta\nu_{j+1})}$	$\frac{\Delta S_i}{\Delta S_{i+1}(\Delta S_i+\Delta S_{i+1})}$ $\frac{-\Delta\nu_{j+1}}{\Delta\nu_j(\Delta\nu_j+\Delta\nu_{j+1})}$
0	$\frac{-\Delta S_{i+1}}{\Delta S_i(\Delta S_i+\Delta S_{i+1})}$ $\frac{-\Delta\nu_j+\Delta\nu_{j+1}}{\Delta\nu_j\Delta\nu_{j+1}}$	$\frac{-\Delta S_i+\Delta S_{i+1}}{\Delta S_i\Delta S_{i+1}}$ $\frac{-\Delta\nu_{j+1}}{\Delta\nu_j(\Delta\nu_j+\Delta\nu_{j+1})}$	$\frac{\Delta S_i}{\Delta S_{i+1}(\Delta S_i+\Delta S_{i+1})}$ $\frac{-\Delta\nu_{j+1}}{\Delta\nu_j(\Delta\nu_j+\Delta\nu_{j+1})}$
1	$\frac{-\Delta S_{i+1}}{\Delta S_i(\Delta S_i+\Delta S_{i+1})}$ $\frac{\Delta\nu_j}{\Delta\nu_{j+1}(\Delta\nu_j+\Delta\nu_{j+1})}$	$\frac{-\Delta S_i+\Delta S_{i+1}}{\Delta S_i\Delta S_{i+1}}$ $\frac{\Delta\nu_j}{\Delta\nu_{j+1}(\Delta\nu_j+\Delta\nu_{j+1})}$	$\frac{\Delta S_i}{\Delta S_{i+1}(\Delta S_i+\Delta S_{i+1})}$ $\frac{\Delta\nu_j}{\Delta\nu_{j+1}(\Delta\nu_j+\Delta\nu_{j+1})}$

Table 4.3: Table of central derivatives for mixed spatial derivatives term.

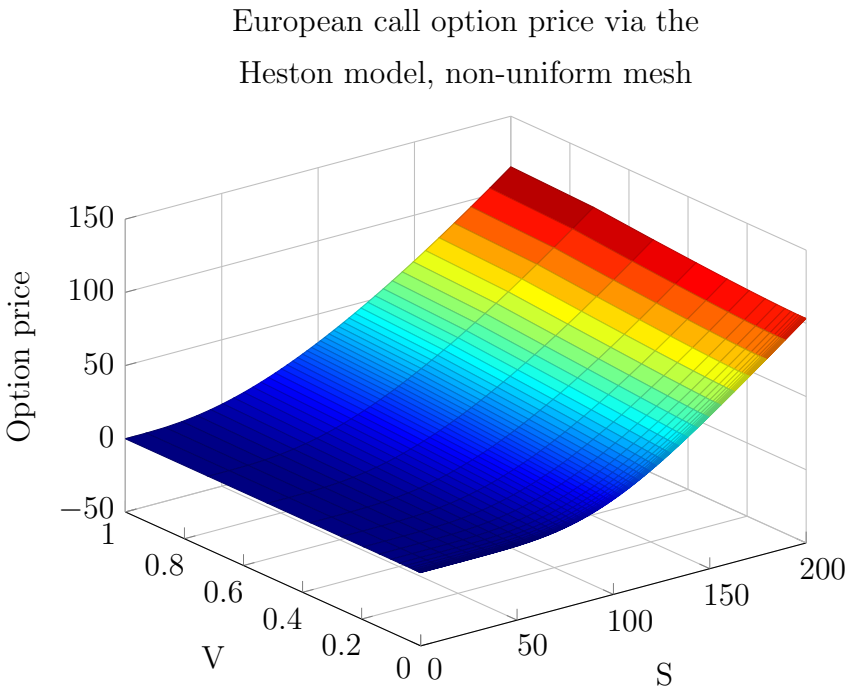


Figure 4.14: European call option price grid in the non-uniform case with parameter values $\theta' = \frac{1}{2}$, $\kappa = 2$, $\theta = 0.2$, $\sigma = 0.3$, $\rho = 0.8$, $r = 0.03$, $S = 200$, $V = 1$, $T = 1$, $K = 100$, $N_S = 40$, $N_V = 20$ and $N_T = 4000$.

Chapter 5

Analysis

5.1 Error Analysis

5.1.1 Convergence, Consistency and Stability

The purpose of this section is to familiarize the reader with some important concepts encountered when dealing with finite difference schemes. With this in mind, the following definitions serve as a base from which stems our analysis.

Denote as before the approximation to the Heston problem at a grid point (S_i, ν_j, t_n) by $U^n_{i,j}$ and let $u(S_i, \nu_j, t_n)$ denote the exact solution[1]. In order to keep the problem as simple as possible define an operator L as follows

$$L = \frac{1}{2}\nu S^2\frac{\partial^2}{\partial S^2} + \frac{1}{2}\sigma^2\nu\frac{\partial^2}{\partial \nu^2} + \rho\sigma\nu S\frac{\partial^2}{\partial S\partial\nu} + rS\frac{\partial}{\partial S} + \kappa(\theta-\nu)\frac{\partial}{\partial\nu} - r, \tag{5.1}$$

such that our Heston PDE can be written

$$\frac{\partial U}{\partial t} = LU. \tag{5.2}$$

If we define the following approximation operator

$$D_{\Delta S\Delta\nu}U^n_{i,j} \coloneqq \frac{1}{2}\nu S^2\frac{\delta_S^2 U^n_{i,j}}{\Delta S^2} + \frac{1}{2}\sigma^2\nu\frac{\delta_\nu^2 U^n_{i,j}}{\Delta\nu^2} + \rho\sigma\nu S\frac{\delta^2_{S\nu}U^n_{i,j}}{4\Delta S\Delta\nu} + rS\frac{\delta_S U^n_{i,j}}{2\Delta S} + \kappa(\theta-\nu)\frac{\delta_\nu U^n_{i,j}}{2\Delta\nu} - rU^n_{i,j}, \tag{5.3}$$

we obtain the corresponding finite difference scheme

$$\frac{\Delta_{+t}U^n_{i,j}}{\Delta t} = D_{\Delta S\Delta\nu}U^n_{i,j}, \tag{5.4}$$

or in concise matrix notation

$$\mathbf{L}\mathbf{U}^{n+1} = \mathbf{Q}\mathbf{U}^n,\ n \in [0, N_T - 1] \tag{5.5}$$

[1]Note that this is the only time we highlight this difference and in previous sections we use the two interchangeably.

where the matrices $\mathbf{L}$ and $\mathbf{Q}$ are of size $(N_S-1)(N_V-1)\times(N_S-1)(N_V-1)$ and $\mathbf{U}^q$ with $q=n,n+1$ are vectors of size $(N_S-1)(NV-1)$ with arrangement as follows (see Rouah [40]):

$$\left(U^q_{1,1},...,U^q_{1,N_V-1},...,U^q_{N_S,1},...,U^q_{N_S-1,N_V-1}\right)^T. \tag{5.6}$$

We are now ready to introduce some important notions.

Definition 5.1.1 (Truncation error)**.** The truncation error is defined as the difference between the exact differential equation and its finite difference representation. In mathematical terms the truncation error is

$$\tau(S,\nu,t) \coloneqq \frac{\Delta_{+t}u(S,\nu,t)}{\Delta t} - Lu(S,\nu,t), \tag{5.7}$$

where u denotes the exact solution. If the truncation error is given at a point in time and space then it is called a local truncation error. □

Definition 5.1.2 (Convergence)**.** A numerical scheme is convergent if at a grid node (S_i,ν_j,t_n) the numerical solution tends to the exact solution as the mesh length tends to zero. Equivalently, if any grid point $(S,\nu,t)\times(S_{min},S_{max})\times(\nu_{min},\nu_{max})\times(0,T]$ with $S_i\to S$, $\nu_j\to\nu$, $t_n\to t \Rightarrow U^n_{i,j}\to U(S_i,\nu_j,t_n)$. □

Definition 5.1.3 (Consistency)**.** A numerical scheme is consistent if it converges to the solution of the PDE that is being discretized. This means that as the increments in time and space tend to zero, the truncation error tends to zero also. Therefore, the finite difference scheme (5.4) is consistent with the PDE (4.2) if the following relationship holds

$$\left(\frac{\partial}{\partial t}-L\right)U(S_i,\nu_j,t_n) - \left[\frac{\Delta_{+t}U^n_{i,j}}{\Delta t} - D_{\Delta S\Delta\nu}U^n_{i,j}\right] \to 0 \text{ as } \Delta S,\ \Delta\nu,\ \Delta t\to 0. \tag{5.8}$$

The numerical scheme is inconsistent if the numerical solution converges to the solution of a *different* partial differential equation. □

Definition 5.1.4 (Stability)**.** The difference scheme (5.4) is said to be stable with respect to the norm $\|\cdot\|$ if there exist positive constants ΔS_0, $\Delta\nu_0$ and Δt_0 and two non-negative constants K and β such that

$$\left\|\mathbf{U}^{n+1}\right\| \le Ke^{\beta t}\left\|\mathbf{U}^n\right\| \tag{5.9}$$

for $0<\Delta S<\Delta S_0$, $0<\Delta\nu<\Delta\nu_0$, $0<\Delta t<\Delta t_0$. □

Definition 5.1.5 (Oder of accuracy)**.** A difference scheme is said to be accurate up to order (p,q,r) in ΔS, $\Delta\nu$ and Δt respectively, with respect to the PDE of interest if for some constant C

$$\tau(S,\nu,t)\le C(\Delta S^p+\Delta\nu^q+\Delta t^r) \tag{5.10}$$

as ΔS, $\Delta\nu$, and Δt tend to zero. □

Definition 5.1.6 (Von-Neumann Stability)**.** Let $\bar{a}$ be the *amplification factor* associated with the difference scheme of interest. The scheme is Von-Neumann stable if there exist positive constants ΔS_0, $\Delta\nu_0$ and Δt_0 such that

$$|\bar{a}| \leq 1 + C\Delta t \tag{5.11}$$

for $0 < \Delta S < \Delta S_0$, $0 < \Delta\nu < \Delta\nu_0$ and $0 < \Delta t < \Delta t_0$. If the *amplification factor* $\bar{a}$ is independent of ΔS, $\Delta\nu$ and Δt, the stability condition can be replaced with the restricted stability condition

$$|\bar{a}| \leq 1. \tag{5.12}$$

□

Proving convergence directly from its definition can be quite difficult, in general. Instead, it may be easier to use the *Lax equivalence theorem*, which connects the definitions of convergence, consistency and stability. It asserts that as long as the scheme is consistent, then convergence is synonymous with stability. We give a definition and then the formal statement of the theorem.

Definition 5.1.7 (Well-posedness)**.** A problem is well-posed if

1. it has a solution
2. the solution is unique[2]
3. the solution depends continuously on the initial and boundary conditions[3].

□

Theorem 1 (Lax equivalence theorem)**.** A consistent difference scheme for a well-posed linear initial value problem is convergent if and only if it is stable. □

A formal proof of this theorem is beyond the scope of this dissertation, but can be found in Richtmyer and Morton (1994) [39]. As mentioned, the theorem is very useful as we need only show consistency and stability to obtain convergence.

5.2 The Explicit Scheme

5.2.1 Truncation Error and its Consistency Implications

In this section we derive an expression for the truncation error in the explicit scheme. The work hereon for stability considerations in the Explicit FD scheme

[2]It could be the case that sometimes a solution is "unique within a certain class of functions". For example, a problem may have many solutions and only one of them is bounded. Then the solution is unique within the space of bounded functions.

[3]A solution depends continuously on the initial and boundary conditions if *small* changes in the initial or boundary conditions and in parameter values result in *small* changes in the solution (with respect to the appropriate norm).

follows closely the work by Mitchell and Griffiths [35] and Engan [18].

$$\begin{aligned}\tau(S,\nu,t) :=& \frac{\Delta_{+t}u(S,\nu,t)}{\Delta t} - Lu(S,\nu,t) \\ =& \frac{\Delta_{+t}u(S,\nu,t)}{\Delta t} - \frac{1}{2}\nu S^2\frac{\delta_S^2 u(S,\nu,t)}{\Delta S^2} + \frac{1}{2}\sigma^2\nu\frac{\delta_\nu^2 u(S,\nu,t)}{\Delta\nu^2} \\ &+ \rho\sigma\nu S\frac{\delta_{S\nu}^2 u(S,\nu,t)}{4\Delta S\Delta\nu} + rS\frac{\delta_S u(S,\nu,t)}{2\Delta S} + \kappa(\theta-\nu)\frac{\delta_\nu u(S,\nu,t)}{2\Delta\nu} - ru(S,\nu,t).\end{aligned} \tag{5.13}$$

Using our earlier Taylor series expansions we have

$$\begin{aligned}\tau =& \frac{u_t\Delta t + \frac{1}{2}u_{tt}\Delta t^2 + \frac{1}{6}u_{ttt}\Delta t^3 + \mathcal{O}(\Delta t^4)}{\Delta t} - \frac{1}{2}\left[\frac{u_{SS}\Delta S^2 + \frac{1}{12}u_{SSSS}\Delta S^4 + \mathcal{O}(\Delta S^6)}{\Delta S^2}\right] \\ &+ \rho\sigma\nu S\left[\frac{4u_{S\nu}\Delta S\Delta\nu + \mathcal{O}((\Delta S\Delta\nu)^2)}{4\Delta S\Delta\nu}\right] + \frac{1}{2}\sigma^2\nu\left[\frac{u_{\nu\nu}\Delta\nu^2 + \frac{1}{12}u_{\nu\nu\nu\nu}\Delta\nu^4 + \mathcal{O}(\Delta\nu^6)}{\Delta\nu^2}\right] \\ &+ \kappa(\theta-\nu)\left[\frac{u_\nu\Delta\nu + \frac{1}{6}u_{\nu\nu\nu}\Delta\nu^3 + \mathcal{O}(\Delta\nu^5)}{\Delta\nu}\right] + rS\left[\frac{u_S\Delta S + \frac{1}{6}u_{SSS}\Delta S^3 + \mathcal{O}(\Delta S^5)}{\Delta S}\right].\end{aligned} \tag{5.14}$$

However, since u is the exact solution, it identically satisfies the original PDE at all grid nodes (S_i, ν_j, t_n), thus

$$\frac{\partial u}{\partial t} - Lu = 0. \tag{5.15}$$

Substituting (5.15) into (5.14) gives the truncation error

$$\tau(S,\nu,t) = \tau_{i,j}^n = \mathcal{O}(\Delta S^2 + \Delta\nu^2 + \Delta t). \tag{5.16}$$

Thus the classical explicit solution to the Heston PDE is $\mathcal{O}(\Delta t)$ accurate in time and $\mathcal{O}(\Delta S^2)$ and $\mathcal{O}(\Delta\nu^2)$ accurate in space. This means that the explicit Euler scheme is first-order accurate in time and second-order accurate in space. Moreover, since $\tau(S,\nu,t) \to 0$ as ΔS, $\Delta\nu$, $\Delta t \to 0$ the explicit Euler scheme is also consistent.

5.2.2 Convergence

The error of the explicit scheme at the grid node (S_i, ν_j, t_n) is given by

$$e_{i,j}^n = U_{i,j}^n - u(S_i,\nu_j,t_n), \tag{5.17}$$

where as usual, $U_{i,j}^n$ denotes the approximation to the exact solution $u(S_i,\nu_j,t_n)$. Since $U_{i,j}^n$ is the solution to the difference equation it has no truncation error, however $u_{i,j}^n$ leaves the residual truncation error $\tau_{i,j}^n$. So if we apply the truncation formula to the error we have

$$\begin{aligned}\tau e_{i,j}^n =& -\tau(S_i,\nu_j,t_n) \\ =& \frac{\Delta_{+t}e_{i,j}^n}{\Delta t} - Le_{i,j}^n\end{aligned} \tag{5.18}$$

$\Rightarrow$

$$\begin{aligned}
e_{i,j}^{n+1} = & e_{i,j}^{n}\left(1 - r\Delta t - \frac{\Delta t}{\Delta S^2}\nu_j S_i^2 - \frac{\Delta t}{\Delta \nu^2}\sigma^2\nu_j\right) - \Delta t\tau_{i,j}^{n} \\
& + e_{i-1,j}^{n}\left(\frac{\Delta t}{2\Delta S^2}\nu_j S_i^2 - \frac{\Delta t}{2\Delta S}rS_i\right) + e_{i+1,j}^{n}\left(\frac{\Delta t}{2\Delta S^2}\nu_j S_i^2 + \frac{\Delta t}{2\Delta S}rS_i\right) \\
& + e_{i,j-1}^{n}\left(\frac{\Delta t}{2\Delta \nu^2}\sigma^2\nu_j - \kappa(\theta - \nu_j)\frac{\Delta t}{2\Delta \nu}\right) + e_{i,j+1}^{n}\left(\frac{\Delta t}{2\Delta \nu^2}\sigma^2\nu_j + \kappa(\theta - \nu_j)\frac{\Delta t}{2\Delta \nu}\right) \\
& + \rho\sigma\nu_j S_i \frac{\Delta t}{4\Delta S \Delta \nu}\left(e_{i+1,j+1}^{n} + e_{i-1,j-1}^{n} + e_{i-1,j+1}^{n} + e_{i+1,j-1}^{n}\right).
\end{aligned} \tag{5.19}$$

So that the coefficients preceding the error terms are non-negative we need to impose the following conditions

$$\frac{\frac{\Delta t}{2\Delta S}|rS_i|}{\frac{\Delta t}{2\Delta S^2}(\nu_j S_i^2)} \leq 1,\ \frac{\frac{\Delta t}{2\Delta \nu}|\kappa(\theta - \nu_j)|}{\frac{\Delta t}{2\Delta \nu^2}(\sigma^2\nu_j)} \leq 1,\ r\Delta t + \frac{\Delta t}{\Delta S^2}\nu_j S_i^2 + \frac{\Delta t}{\Delta \nu^2}\sigma^2\nu_j \leq 1$$

for $i = 0, 1, ..., N_S$, $j = 0, 1, ..., N_V$. We must also have

$$\rho\sigma\nu_j S_i \geq 0 \Leftrightarrow \rho > 0,\ \sigma > 0 \text{ or } \rho < 0,\ \sigma < 0$$

where we note that the asset price S_i and variance ν_j must be positive processes. Note that

$$\begin{aligned}
|e_{i,j}^{n+1}| \leq (1+r)\max\big\{&|e_{i-1,j}^{n}|, |e_{i+1,j}^{n}|, |e_{i,j-1}^{n}|, |e_{i,j+1}^{n}|, |e_{i,j}^{n}|, |e_{i+1,j-1}^{n}|, |e_{i-1,j+1}^{n}|, |e_{i-1,j+1}^{n}|, \\
&|e_{i+1,j+1}^{n}|, |e_{i-1,j-1}^{n}|\big\} + \Delta t M(\Delta t + \Delta S^2 + \Delta \nu^2),
\end{aligned} \tag{5.20}$$

and define

$$E^n = \max_{0 \leq i \leq N_S, 0 \leq j \leq N_V} |e_{i,j}^{n}|, \tag{5.21}$$

such that

$$E^{n+1} \leq E^n(1 + r\Delta t) + L(\Delta t + \Delta S^2 + \Delta \nu^2). \tag{5.22}$$

By induction

$$\begin{aligned}
E^n &\leq E^0(1 + r\Delta t)^n + \Delta t L(\Delta t + \Delta S^2 + \Delta \nu^2)\sum_{k=0}^{n-1}(1 + r\Delta t)^k \\
&= \frac{\Delta t L(\Delta t + \Delta S^2 + \Delta \nu^2)((1 + r\Delta t)^n - 1)}{r\Delta t} \\
&\leq e^{rn\Delta t}\tilde{L}(\Delta t + \Delta S^2 + \Delta \nu^2)
\end{aligned} \tag{5.23}$$

$\Rightarrow$ convergence.

5.2.3 Von-Neumann Stability

There are different ways of determining the stability of a scheme[4]. A common approach to determining any stability conditions on a scheme is the Von-Neumann

[4]Further techniques with examples can be found in Duffy (2006) [15].

stability approach. Strictly speaking, this technique is only valid for linear initial value problems with constant coefficients and periodic boundary conditions. Regardless, it is often applied in the literature to general problems as a necessary stability condition. When dealing with non-constant coefficients, stability is dealt with as a local condition, where we *freeze* the coefficients and determine the stability at a given grid point.

Bearing this in mind, suppress the indices from the spatial variables and let $i = \sqrt{-1}$ be the imaginary unit and p,q be the spatial indices. Write

$$U_{p,q}^{n} = \bar{a}^{n} e^{i(pk_S\Delta S + qk_\nu \Delta\nu)}, \tag{5.24}$$

and use this representation in the finite difference PDE (5.4) to obtain

$$\begin{aligned}
\bar{a} = & \left(1 - \frac{\Delta t}{\Delta S^2} S^2\nu - \frac{\Delta t}{\Delta \nu^2}\sigma^2\nu - r\Delta t\right) \\
& + \left(\frac{1}{2}\frac{\Delta t}{\Delta S^2} S^2\nu - \frac{1}{2}\frac{\Delta t}{\Delta S} rS\right) e^{-ik_S\Delta S} + \left(\frac{1}{2}\frac{\Delta t}{\Delta S^2} S^2\nu + \frac{1}{2}\frac{\Delta t}{\Delta S} rS\right) e^{ik_S\Delta S} \\
& + \left(\frac{1}{2}\frac{\Delta t}{\Delta \nu^2}\sigma^2\nu - \frac{1}{2}\frac{\Delta t}{\Delta \nu}\kappa(\theta-\nu)\right) e^{-ik_\nu\Delta\nu} + \left(\frac{1}{2}\frac{\Delta t}{\Delta \nu^2}\sigma^2\nu + \frac{1}{2}\frac{\Delta t}{\Delta \nu}\kappa(\theta-\nu)\right) e^{ik_\nu\Delta\nu} \\
& + \frac{1}{4}\frac{\Delta t}{\Delta S \Delta\nu}\rho\sigma S\nu \left[e^{ik_S\Delta S + ik_\nu\Delta\nu} + e^{-ik_S\Delta S - ik_\nu\Delta\nu} - e^{-ik_S\Delta S + ik_\nu\Delta\nu} - e^{ik_S\Delta S - ik_\nu\Delta\nu}\right]
\end{aligned} \tag{5.25}$$

$$\Updownarrow$$

$$\begin{aligned}
\bar{a} - 1 = & \frac{1}{2}\frac{\Delta t}{\Delta S^2} S^2\nu \left[e^{-ik_S\Delta S} - 2 + e^{ik_s\Delta S}\right] + \frac{1}{2}\frac{\Delta t}{\Delta \nu^2}\sigma^2\nu \left[e^{-ik_\nu\Delta\nu} - 2 + e^{ik_\nu\Delta\nu}\right] \\
& + \frac{1}{2}\frac{\Delta t}{\Delta S} rS \left[e^{ik_S\Delta S} - e^{-ik_S\Delta S}\right] + \frac{1}{2}\frac{\Delta t}{\Delta \nu}\kappa(\theta-\nu)\left[e^{ik_\nu\Delta\nu} - e^{-ik_\nu\Delta\nu}\right] \\
& + \frac{1}{4}\frac{\Delta t}{\Delta S\Delta\nu}\rho\sigma S\nu \left[e^{ik_S\Delta S + ik_\nu\Delta\nu} + e^{-ik_S\Delta S - ik_\nu\Delta\nu} - e^{-ik_S\Delta S + ik_\nu\Delta\nu} - e^{ik_S\Delta S - ik_\nu\Delta\nu}\right] \\
& - r\Delta t.
\end{aligned} \tag{5.26}$$

Recall the following identities which allows us to express the sine and cosine functions in terms of the exponential

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i},\ \cos(x) = \frac{e^{ix} + e^{-ix}}{2}. \tag{5.27}$$

Also recall the double angle formula

$$\cos(2x) = 1 - 2\sin^2(x). \tag{5.28}$$

So we have

$$e^{ik_S\Delta S} + e^{-ik_S\Delta S} = 2i\sin(k_S\Delta S), \tag{5.29}$$

and

$$\frac{e^{ik_S\Delta S} + e^{-ik_S\Delta S}}{2} = 1 - 2\sin^2\left(\frac{1}{2}k_S\Delta S\right) \tag{5.30}$$

$\Rightarrow$

$$e^{ik_S\Delta S} + e^{-ik_S\Delta S} - 2 = -4\sin^2\left(\frac{1}{2}k_S\Delta S\right). \tag{5.31}$$

Similarly, we can express the mixed derivative term as follows

$$\begin{aligned} & e^{ik_S\Delta S}(e^{ik_\nu\Delta\nu} - e^{-ik_\nu\Delta\nu} - e^{-ik_S\Delta S}(e^{ik_\nu\Delta\nu} - e^{-ik_\nu\Delta\nu}) \\ & \quad = 2ie^{ik_S\Delta S}\sin(k_\nu\Delta\nu) - 2ie^{-ik_S\Delta S}\sin k_\nu\Delta\nu \\ & \quad = 4i^2\sin(k_\nu\Delta\nu)\sin(k_S\Delta S) \\ & \quad = -4\sin(k_\nu\Delta\nu)\sin(k_S\Delta S) \end{aligned} \tag{5.32}$$

$\Rightarrow$

$$\begin{aligned} \bar{a} = 1 - r\Delta t - 4\left[\frac{1}{2}\frac{\Delta t}{\Delta S^2}S^2\nu\sin^2\left(\frac{1}{2}k_S\Delta S\right) + \frac{1}{2}\frac{\Delta t}{\Delta\nu^2}\sigma^2\nu\sin^2\left(\frac{1}{2}k_\nu\Delta\nu\right)\right] \\ + i\left[\frac{\Delta t}{\Delta S}rS\sin(k_S\Delta S) + \frac{\Delta t}{\Delta\nu}\kappa(\theta-\nu)\sin(k_\nu\Delta\nu)\right] - \frac{\Delta t}{\Delta S\Delta\nu}\rho\sigma S\nu\sin(k_\nu\Delta\nu)\sin(k_S\Delta S). \end{aligned} \tag{5.33}$$

The Von-Neumann condition for stability is

$$G = |\bar{a}| \leq 1,$$

where we recall that G is the *amplification factor*.

We therefore see that the explicit scheme is conditionally stable.

5.2.4 Non-rigorous Stability Derivation

This section follows closely the work on stability analysis for the Explicit FD scheme done by Lin [32]. The first step is to observe that the following relationship holds, as long as $r > 0$,

$$A^n_{i,j} + B^n_{i,j} + C^n_{i,j} + D^n_{i,j} + E^n_{i,j} < 1 \tag{5.34}$$

as it implies that

$$1 - r\Delta t < 1 \tag{5.35}$$

where Δt, $r > 0$. In order to maintain the above relationship we should ascertain that each individual bracketed term is positive. We thus first obtain

$$\begin{aligned} 1 - \Delta t(i^2 j\Delta\nu + \frac{\sigma^2 j}{\Delta\nu} + r) \geq 0 \\ \Rightarrow \Delta t \leq \frac{1}{i^2 j\Delta\nu + \frac{\sigma^2 j}{\Delta\nu} + r} \end{aligned} \tag{5.36}$$

and noting that $V = N_V\Delta\nu \Rightarrow \frac{N_V}{V} = \frac{1}{\Delta\nu}$[5] we can write

$$\Delta t \leq \frac{1}{i^2 j\Delta\nu + \frac{N_V^2\sigma^2 j\Delta\nu}{V^2} + r}. \tag{5.37}$$

Similarly,

$$i \geq \frac{r}{j\Delta\nu} \tag{5.38}$$

[5]Note here that $V = V_{max}$.

$$j \geq \frac{\kappa(\theta - j\Delta\nu)}{\sigma^2} \tag{5.39}$$

It's equivalent to write equation (5.37) as

$$\Delta t \leq \frac{1}{N_S^2 V + \frac{N_V^2 \sigma^2}{V} + r}. \tag{5.40}$$

5.3 Numerical Experiments

This section presents a series of numerical experiments carried out to test properties of the ADI methods previously introduced. Thus far, in any illustration given we have been working with a single set of parameters as given in Table 4.2. In this section we extend the analysis to investigate different parameter scenarios, so that we can make comparisons and draw conclusions on the effect of changing a particular parameter value. In Table 5.1, we present our original parameter set (case 1) and four further sets which are taken from In'T Hout and Foulon [27]. Note that case 3 is important as σ is close to zero which implies that the Heston PDE is dominated by the convection term in the ν-direction. Furthermore, case 4 is important as here the Feller condition $2\kappa\theta' \geq \sigma^2$ is only just met.

	Case 1	Case 2	Case 3	Case 4	Case 5
κ	2	1.5	3	0.6067	2.5
θ	0.2	0.04	0.12	0.0707	0.06
σ	0.3	0.3	0.04	0.2928	0.5
ρ	0.8	-0.9	0.6	-0.7571	-0.1
r	0.03	0.025	0.01	0.03	0.0507
T	1	1	1	3	0.25
K	100	100	100	100	100

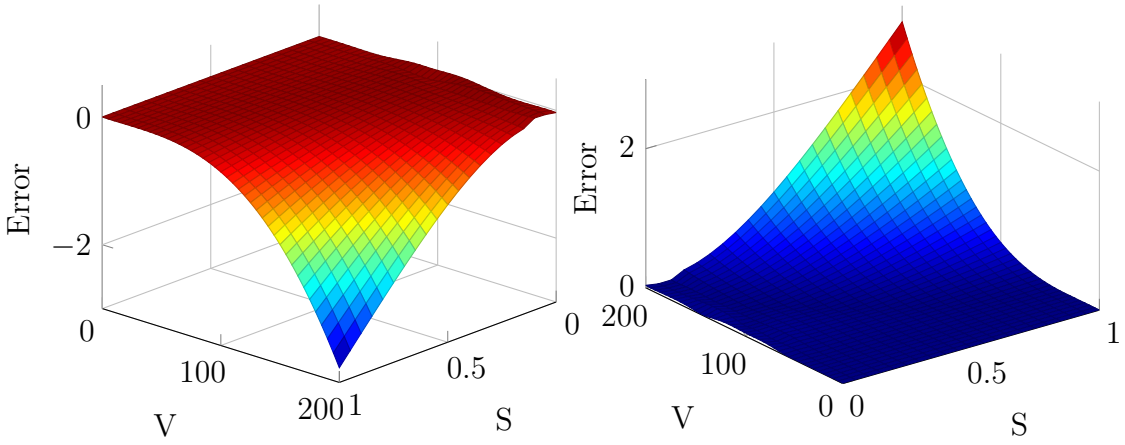
Table 5.1: Sets of parameter values for the Heston model used to price a European call option.

5.3.1 The Explicit Scheme

This section pays brief attention to condition (5.40) on Δt. The parameters used are those in case 1 with $N_S = 40$, $N_V = 20$ and $N_T = 4000$. In this case $\Delta t = \frac{T}{N_T} = \frac{1}{4000} = 0.00025$, and looking at inequality (5.40), the requirement for stability is $\Delta t \leq \frac{1}{40^2+20^2 0.3^2+0.03} = \frac{1}{1636.03} \approx 0.00061124$. Thus in this case the requirement is met.

Figure 5.1a is a plot of the error surface in the explicit FD scheme. It is calculated as the difference between the exact solution and its finite difference approximation. In turn, Figure 5.1b is a plot of the absolute of this error quantity. It is evident from both plots that the error increases towards the maximum values of asset price and volatility. Note that by setting $V = 1$ the stability condition (5.40) breaks down to $\Delta t \leq \frac{1}{N_S^2+\sigma^2 N_V^2}$, in which case increasing the number of asset steps N_S and volatility steps N_V will require an increase in the number of time steps N_T to maintain stability. For example, if we increase the number of asset steps to $N_S = 60$ and the number of volatility steps $N_V = 30$ then we need for instance $N_T = 6000$ to ensure condition (5.40) is satisfied[6].

[6]The derivation of the stability condition on Δt is non-rigorous and as such we may find the scheme is unstable if the condition is only just met for example.



(a) Plot of the error in the explicit FD scheme.

(b) Plot of the absolute error in the explicit FD scheme.

Figure 5.1: Explicit FD scheme error and absolute error.

Figure 5.2 is a plot of the number of time steps against the error in the Explicit FD scheme. It can be seen that as the number of time steps is increased (i.e. Δt is decreased), the error decreases, as expected. If we were to plot the error for a Δt value violating the condition (5.40) then the error would grow without bound and we would expect to see the error for such Δt suddenly shoot off.

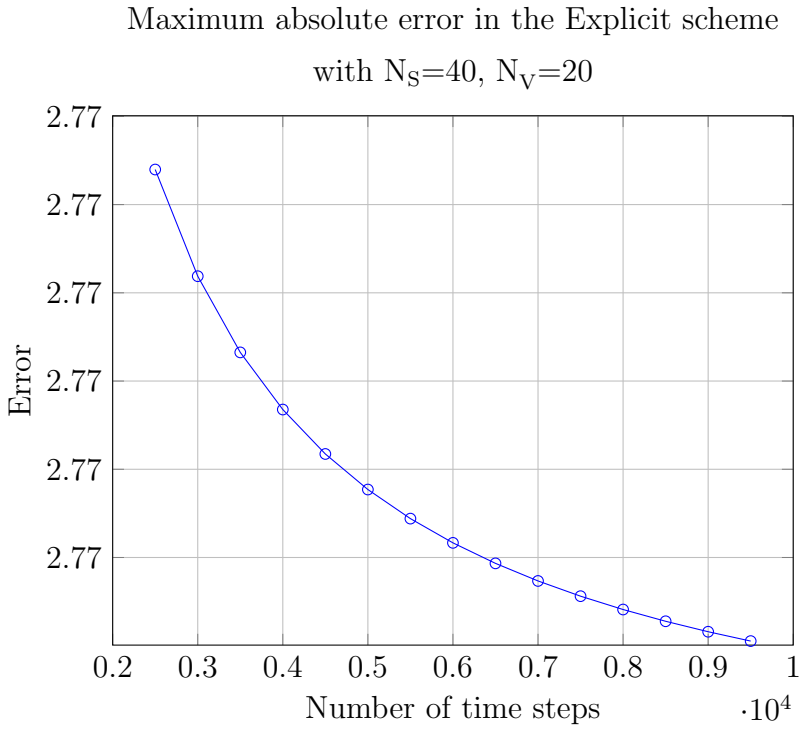


Figure 5.2: Plot of the error vs. the number of time steps in the Explicit FD scheme.

ADI Methods

To examine global convergence of a particular scheme, we must observe changes in the error across all mesh nodes as the number of grid points is increased. Due to the availability of a closed form solution of the Heston PDE for European options, the error can be easily defined as the difference between the exact solution $u(S_i, \nu_j, t_n)$ and numerical approximation $U^n_{i,j}$ at a given grid point. This is only valid for a uniform mesh however, as in the non-uniform case it is not possible to monitor a single node as its position changes non-linearly. Instead, using the infinity norm one is able to capture the largest error across the entire mesh. Formally, we define the *maximum absolute error* (also termed *global discretization error* in [27]) which is valid for both uniform and non-uniform mesh discretizations as follows[7]

$$e^n_{i,j} \coloneqq ||u(S_i, \nu_j, t_n) - U^n_{i,j}||_\infty = \max_{0\leq i\leq N_S, 0\leq j\leq N_V} |u(S_i, \nu_j, t_n) - U^n_{i,j}| \tag{5.41}$$

We remark that in the case that the exact price of a derivative is not known (eg. barrier options[8]), one must use alternative ways of evaluating the size of the error. For instance, one may use Richardson extrapolation where the error in the option value is obtained by subtracting an infinite grid result from the calculated present option value (see Tavella and Randall [43]).

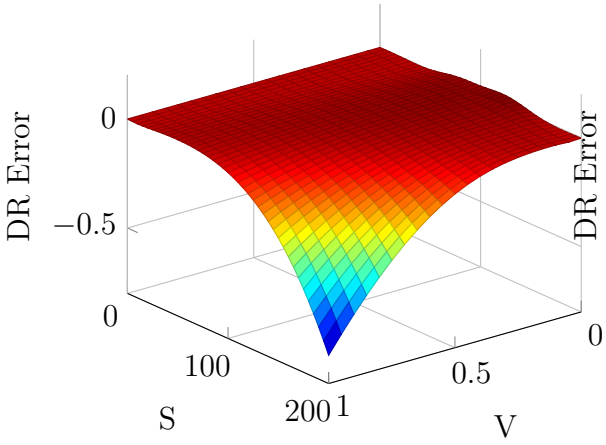
The first treatment presented is on ADI schemes possessing a uniform mesh. It is expected that errors are higher and behave in a more erratic manner than in the non-uniform case. Such conclusion may be attributed to the relative position of the point $S = K$ to the mesh in the S-direction (see further issues in Tavella and Randall [43]). Begin by performing an experiment on the DR ADI scheme with $\theta' = \frac{1}{2}$ and all other parameters as in case 1. Recall that for this value of θ' the DR scheme is consistent and unconditionally stable thus by the Lax equivalence theorem the scheme converges and the error should approach zero as the number of grid-points is increased.

Figures 5.3a and 5.3b show the error and absolute error surfaces as functions of both asset and volatility spatial variables, respectively. It can be seen that the error is high around the critical point $S = K$. This is a *quantization* error and is a result of discontinuities in the payoff function. To reduce this error one must ensure that the present value of the option is dependent to a minimum on the location of the strike relative to the mesh grid. Furthermore, the error seems to be larger at the maximum grid-points S_{max} and ν_{max}, as was the case in the explicit scheme. This is likely to be related to the treatment of the boundary conditions, in particular those at S_{max} and ν_{max}. If we increase the number of grid points (currently $N_S = 40$ and $N_V = 20$) we expect the error to decrease. Recall that since we are dealing with a uniform grid, results around influential points, such as the strike $K = 100$ are not emphasized and errors around these areas may be larger, as we see in Figure 5.3b. The benefit of using a non-uniform grid is that we may concentrate the number of grid points around critical regions of the domain Ω whilst keeping computation time unchanged (as the total number of grid points remains the same). Note that throughout we maintain the relationship $N_S = 2N_V$ as, from an efficiency point of

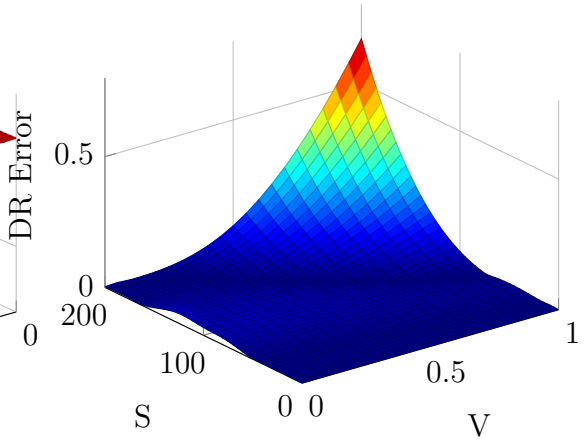
[7] In the literature, often the error is observed over a subset of the region of interest in order to minimize noise from the boundaries.

[8] For a brief description of barrier options see Appendix A.

view, we are able to use less points in the ν-direction than in the S-direction, see [27].



(a) Douglas-Rachford error plot for entire uniform grid.

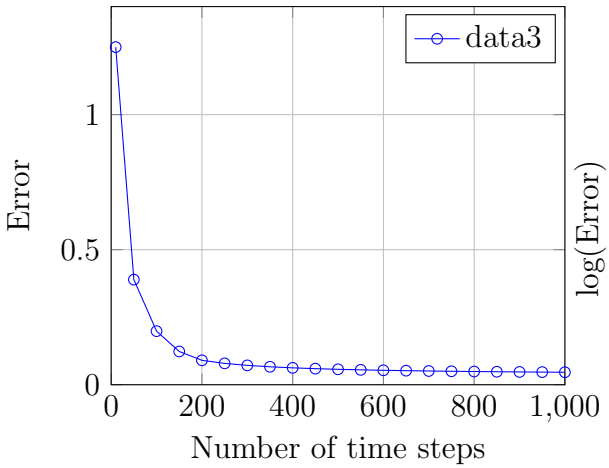


(b) Douglas-Rachford absolute error plot for entire uniform grid.

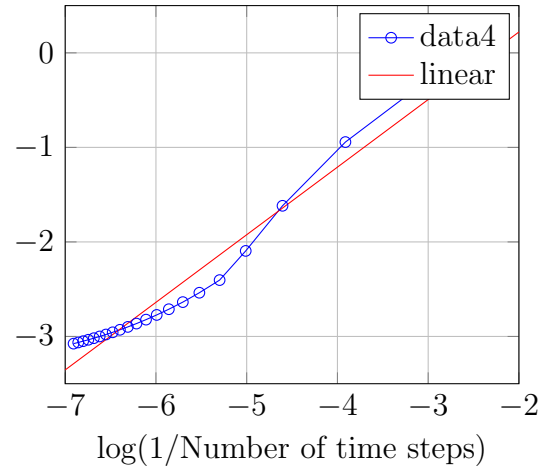
Figure 5.3: DR scheme error taken first as the difference between the exact solution and the approximation and then the absolute value of this difference. Here we take $\Delta t = \frac{1}{1000}$.

Figure 5.4a is a plot of the maximum absolute error as a function of the number of time steps where $N_S = 40$ and $N_V = 20$. As expected, it is clear that as the number of time steps is increased, the error decreases. It is clear that the error values remain within a reasonable bound and decrease monotonically with N_T. This favourable result indicates that the DR scheme converges as the number of time steps is increased (i.e. as Δt is decreased). This result is significant since the unconditional stability of the DR scheme in the time direction does not follow from Von Neumann stability analysis. Figure 5.4b shows that the DR scheme is approx. 1 with respect to time, i.e. $\mathcal{O}(\Delta t)$[9].

[9]Deducing the order of a particular method is described in further detail later on.



(a) Difference between analytical and DR FD approximations (MAE) as number of time steps is increased.



(b) DR scheme, log of error against $\log \frac{1}{N_T}$, with $N_S = 40$ and $N_V = 20$.

Figure 5.4: DR scheme error plots.

Next we present the results from the plots above in table format. Table 5.2 shows the relationship between the number of time steps and an improvement in the error. The percentage improvement changes quickly after the first 300 time steps but then seems to progress at a lower rate. It may be reasonable to select around 900 time steps as an *optimal* choice, yet it remains a trade-off between computational time and accuracy.

No. of grid points in S-direction	Maximum absolute error	% improvement in error
100	0.198321195556034	Null
200	0.090374174644612	26.56%
300	0.071596160224036	9.44%
400	0.062509157872594	5.86%
500	0.057055346895524	4.07%
600	0.053418721771813	3.00%
700	0.050820736810159	2.30%
800	0.048872020517759	1.83%
900	0.047356212289699	1.48%
1000	0.046143474927018	1.23%

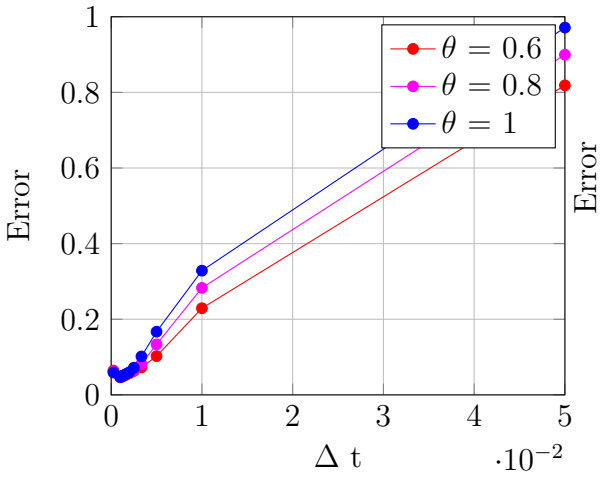
Table 5.2: Douglas scheme for varying values of N_S with $N_V = \frac{N_S}{2}$ and $\Delta t = \frac{1}{1000}$ fixed.

Consider now the effect of varying the weighing parameter θ' on the error across different ADI schemes. We test the validity of (or consistency with) the conditions for unconditional stability under each alternating-direction implicit method as presented in [27]. Recall that the DR and CS schemes are unconditionally stable whenever $\theta' \geq \frac{1}{2}$, the MCS scheme is unconditionally stable whenever $\theta' \geq \frac{1}{3}$ and

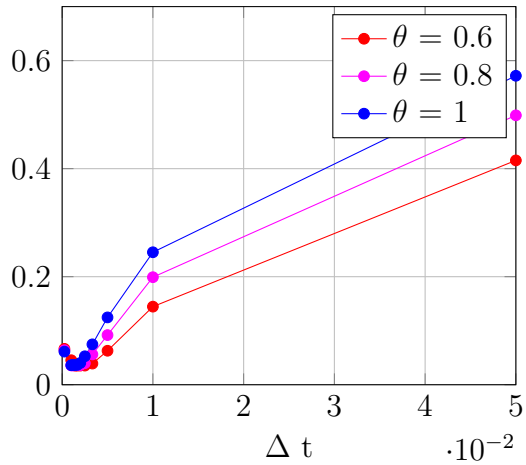
the HV scheme is unconditionally stable whenever $\theta' \geq \frac{1}{2}+\frac{\sqrt{3}}{6}$ for a two-dimensional convection-diffusion equation (as is the case with the Heston PDE) or $\theta' \geq 1 - \frac{\sqrt{2}}{2}$ for a two-dimensional pure diffusion equation.

Using the error definition (5.41), plot a series of graphs to observe the error under each ADI schemed as Δt varies, where we use the data in case 1. In Figures 5.5a-5.5d, the maximum value of Δt tested is $\Delta t = \frac{1}{20} = 0.05$ where $N_T = 20$ is the number of time steps and a uniform grid is used throughout. Note that the parameters used are those in case 1, and the number of time steps varies from $N_T = 20$ to 4000.

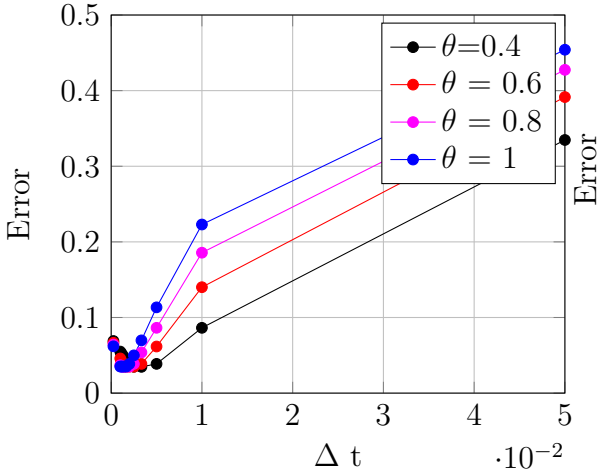
The DR scheme is unstable for $\theta' < \frac{1}{2}$ as evidenced by Figure 5.5a (any smaller values of θ' have led to an unstable error). Similarly, the CS scheme is stable in the case that $\theta' \geq \frac{1}{2}$. On the other hand, the MCS scheme is stable for $\theta' \geq \frac{1}{3}$ and this can be seen from Figure 5.5c as the scheme is stable for $\theta' = 0.4$. With the HV scheme, there are limited sources providing conclusions on unconditional stability in the case of two-dimensional equation with a convection term (as is the case with the Heston PDE). One conjecture mentioned in [27] is that the HV scheme is claimed unconditionally stable when applied to two-dimensional convection-diffusion equation in presence of a mixed derivative term whenever $\theta' \geq \frac{1}{2} + \frac{\sqrt{3}}{6}$. It can be seen from Figure 5.5d that in fact, in the case of the chosen parameters, the error converges for $\theta' = 1$, 0.8, 0.6, and 0.4, rather than just for 1 and 0.8 as the conjecture would suggest. This could be due to the parameter set used, as in the case where $\theta' < \frac{1}{2} + \frac{\sqrt{3}}{6}$ there may be additional parameter constraints necessary for stability, which happen to be met with this particular parameter set.



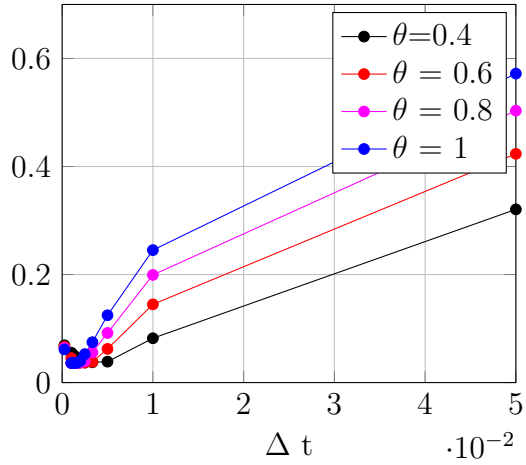
(a) Plot of the error as Δt is decreased in Douglas ADI scheme. For $\theta' < \frac{1}{2}$ the scheme is unstable.



(b) Plot of the error as Δt is decreased in the CS ADI scheme. As with Douglas, for $\theta' < \frac{1}{2}$ the scheme is unstable.



(c) Plot of the error as Δt is decreased in the MCS ADI scheme. In this case, the scheme is unstable whenever $\theta' < \frac{1}{3}$. This is evidenced by the figure, as we can now see a convergent error for $\theta' = 0.4$.



(d) Plot of the error as Δt is decreased in the HV ADI scheme. In this case, the scheme is unconditionally stable whenever $\theta' \geq \frac{1}{2} + \frac{\sqrt{3}}{6}$. Observe in this case stability is obtained (bounded error) for $\theta' \geq 1 - \frac{\sqrt{2}}{2}$.

Figure 5.5: Plots to show the maximum absolute error in different ADI schemes as Δt is decreased. Note that wherever a value of *theta'* leads to unstable unbounded errors, these have not been plotted.

Next we compare accuracy of different ADI schemes. Table 5.3 shows the option price at node $(S, \nu) = (70, 0.12)$ across different ADI schemes, where $N_T = 5000$. This node has been selected to give a minimum error as it is away from the critical areas $S = K$ and $\nu = 0$. We have set $\theta' = \frac{1}{2}$ in schemes DR, CS and MCS, and $\theta' = 1 - \frac{\sqrt{2}}{2}$ for the HV scheme. Note that when $\theta' = \frac{1}{2}$ the CS and MCS schemes are indistinguishable[10]. This is evidenced by equal option price values for both CS

[10]This can be easily verified by inserting $\theta' = \frac{1}{2}$ into the MCS scheme steps.

and MCS schemes. Note that when $\theta' = \frac{1}{3}$ the MCS scheme performs better in the sense of accuracy, however this difference can only be seen by observing more than 4 decimal places (use `format long` in Matlab to see up to 15 dp). Notice that the HV scheme is not different (to 4dp) to the CS or MCS schemes. It performs better as a result of damping methods, or could perform better if tested on other areas of the grid. When θ' was varied, there was not a large change between option price values for the different schemes. This is because Δt is very small. If we pay attention to the section where θ' was varied, we can see that for small Δt, the error is approx. constant across ADI schemes.

Mesh size	ADI call option price				Absolute error			
$(N_S \times N_V)$	DR	CS	MCS	HV	DR	CS	MCS	HV
(21×11)	4.5816	4.5832	4.5832	4.5832	0.0756	0.0740	0.0740	0.0740
(41×21)	4.6350	4.6366	4.6366	4.6366	0.0222	0.0206	0.0206	0.0206
(61×31)	4.6467	4.6483	4.6483	4.6483	0.0105	0.0089	0.0089	0.0089
(81×41)	4.6508	4.6524	4.6524	4.6524	0.0064	0.0048	0.0048	0.0048
(101×51)	4.6528	4.6544	4.6544	4.6544	0.0044	0.0028	0.0028	0.0028
(121×61)	4.6534	4.6550	4.6550	4.6550	0.0038	0.0022	0.0022	0.0022

Table 5.3: ADI schemes applied with a uniform mesh, with the option price value given at point $(S, \nu) = (70, 0.12)$.

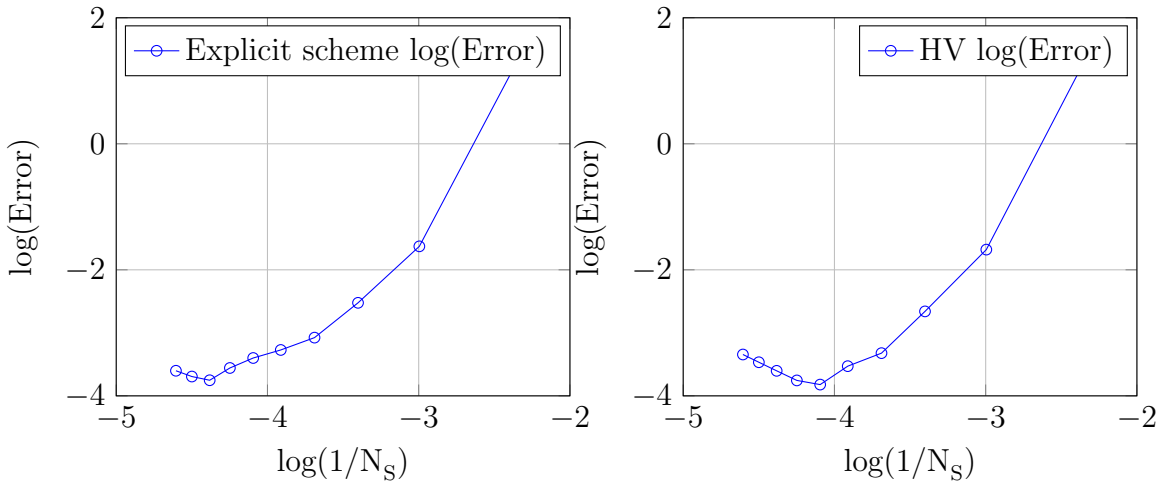
We next study the order of accuracy with respect to spatial variables for different ADI schemes applying parameter set in case 1. In order to establish the order of accuracy with respect to the number of asset and volatility steps of a particular scheme, the common approach is to impose the following relationships

$$\begin{aligned} \epsilon &\sim \Delta S^p \\ \epsilon &\sim \Delta \nu^q \end{aligned} \tag{5.42}$$

where ϵ denotes the error. Then by use of the laws of logarithms the following relationships are obtained

$$\begin{aligned} \log \epsilon &\sim p \log \Delta S \\ \log \epsilon &\sim q \log \Delta \nu \end{aligned} \tag{5.43}$$

This implies that the gradient given by the plots of $\log \epsilon$ vs. $\log \Delta S$ and $\log \epsilon$ vs. $\log \Delta \nu$ will show the order of dependence p on ΔS and q on $\Delta \nu$, respectively. We plot logarithms of the error $e(2N_V, N_V)$ against $\frac{1}{N_S}$ for $N_S = 10, 20, ..., 100$. This is done similarly to In'T Hout and Foulon [27]. This is shown in Figure 5.6a for the DR scheme. The slope is approximately 2, which suggests that the order of the Douglas-Rachford scheme with respect to the asset price is 2 (i.e. $\epsilon \sim \mathcal{O}(\Delta S^2)$). Figure 5.6b is an equivalent plot for the HV scheme with $\theta = \frac{1}{2} + \frac{1}{6}\sqrt{3}$. The order of convergence with respect to the asset price is also approximately two again, as expected. Remaining ADI schemes, CS and MCS, have order of convergence approx. 2 also.



(a) Log of $\frac{1}{N_S}$ against log of the error for DR scheme.

(b) Log of $\frac{1}{N_S}$ against log of the error for HV scheme.

Figure 5.6: Log of inverse number of time steps against log of the error for both Douglas and HV schemes.

In addition to observing the behaviour of the global discretization error, it is of interest to observe individual errors across the grid. In the literature, often the region of interest is cut into a subset in order to avoid unnecessary noise from the boundaries. For example, some papers confine the error to be contained within a subset of the region of interest as follows

$$e^n_{i,j} = \max\left\{|u(S_i,\nu_j,t_n) - U^n_{i,j}|\colon \frac{1}{2}K < S_i < \frac{3}{2}K,\ \nu_{min} < \nu_j < \nu_{max}\right\} \tag{5.44}$$

Figure 5.7 shows four plots representing the error surface in the MCS scheme with $\theta' = \frac{1}{3}$[11], using a uniform mesh, for different cases. Here, the full region of interest is shown. Case 3 has been left out as the error in this case behaves in an unstable manner; this is most likely due to the value of σ in this case, which lies close to zero. In turn, this means that the problem is convection dominated in the ν-direction and such domination is associated with oscillations and numerical errors. Case 1 seems to result in the best performance by the chosen method, MCS. In case 2 there are visible spurious oscillations around the critical region of $\nu_{min} = 0$, with the larger oscillations appearing around the area of $S = K$ where the payoff discontinuity occurs. These oscillations may be a result of the strong negative correlation between the asset price and its volatility, which also explains the growth in the wave amplitude as the asset price grows. Case 4 shows brief oscillations around the critical region of $S = K$. This may again be related to the negative correlation between spatial variables. The large error at the maximum values of S and ν may be due to the extended maturity in this case of 3 years. Finally, case 5 has a short option life span, such that the aforementioned error is not apparent in this case. Again, the small oscillations may be a result of a weak negative correlation

[11] In the paper by [27] the MCS scheme with $\theta' = \frac{1}{3}$ used with damping at $t = 0$ is found to be preferred over other schemes in terms of providing a fast, accurate and robust numerical solution to the Heston PDE with an arbitrary correlation $\rho \in [-1, 1]$.

between spatial variables. Other ADI schemes behave similarly (depending on value of θ' chosen).

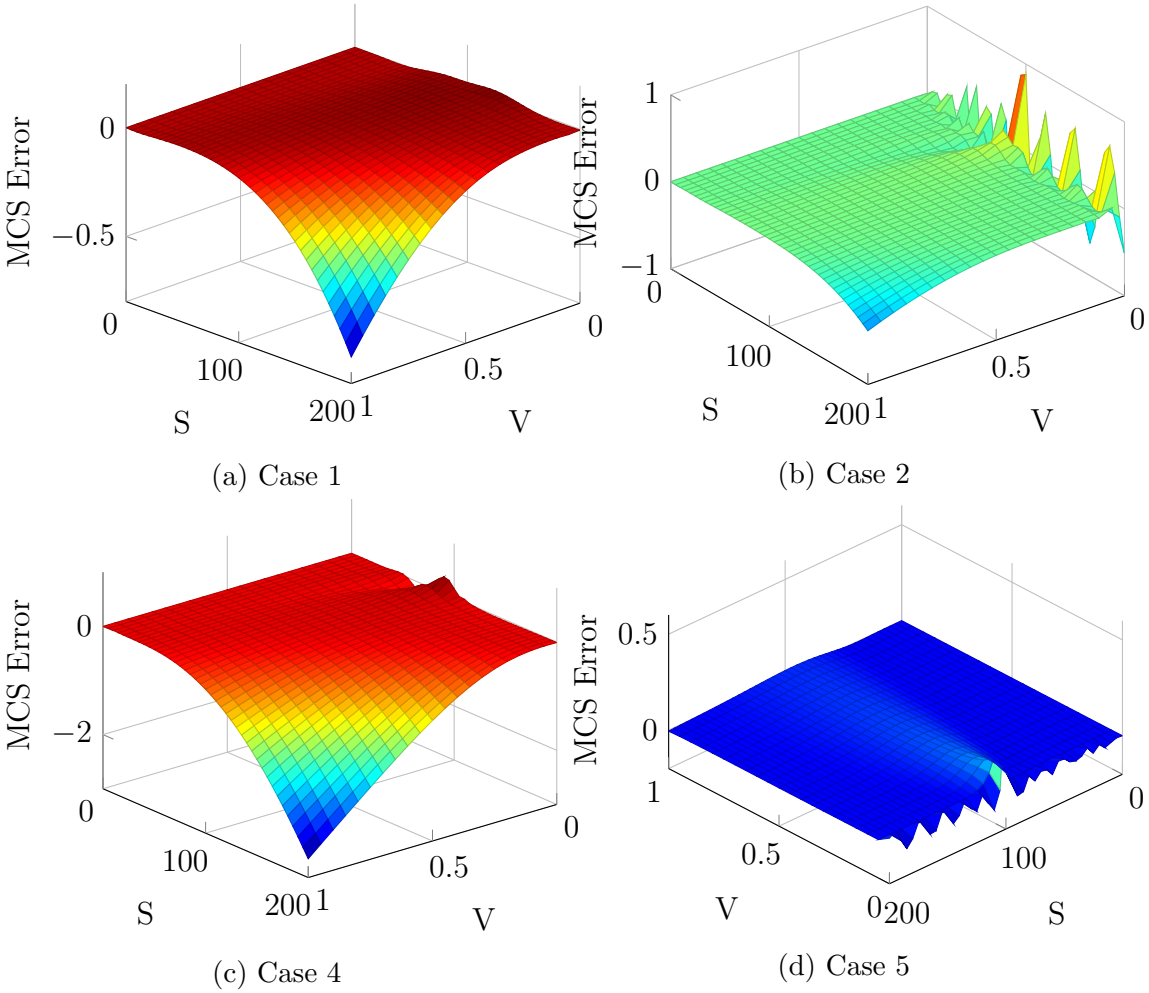


Figure 5.7: Error in MCS scheme with $N_S = 40$, $N_V = 20$ and $\Delta t = \frac{1}{4000}$.

The oscillatory behaviour perceived in the error surfaces may be remedied by either applying damping procedures or an upwinding technique. One such damping procedure involves initially replacing at $t = 0$ the time step Δt by two backward Euler time-steps of size $\frac{\Delta t}{2}$ and continue after $t = \Delta t$ with the time-stepping scheme currently used. This was originally proposed by Rannacher (1984) where the time-stepping in the CN scheme should be preceded by a finite number of implicit steps. The idea is that high frequency error terms are dampened by the implicit steps resulting in a smooth convergence. In this method, it's expected that the second order time-stepping convergence is maintained as only a finite number of time-steps are completed. The reduction in spurious oscillations also simplifies hedging of the underlying. This technique is expensive in terms of computational cost when applied to the Heston model. It has been suggested in [38] that a similar strategy would be to use the Douglas-Rachford scheme with $\theta' = 1$. The scheme becomes

$$\begin{aligned}(1 - \Delta t A_1)U^*_{i,j} &= (1 + \Delta t A_0 + \Delta t A_2)U^n_{i,j} \\ (1 - \Delta t A_2)U^{n+1}_{i,j} &= U^*_{i,j} - \Delta t A_2 U^n_{i,j}.\end{aligned} \tag{5.45}$$

The difference between the DR scheme with $\theta' = \frac{1}{2}$ and $\theta' = 1$ is shown in Figure 5.8a and 5.8b for cases 1 and 2, respectively. The damping procedure looks to be effective for high values of stochastic volatility. The solution, in general for all cases, remains slightly oscillatory however around the critical region of $S = K$. This is more pronounced in cases 1 and 4. To further improve results, one could apply an upwind scheme (one-sided scheme) for a section of inner points in the grid where the diffusion coefficient in the ν-direction is negative and maintain a central scheme whenever it is positive. A combination of both schemes may lead to reductions in the spatial error. De Graaf (2012) [12] a proportion of $\lambda = 0.8$, where $\lambda = 1$ represents an entirely central scheme, yields the lowest error.

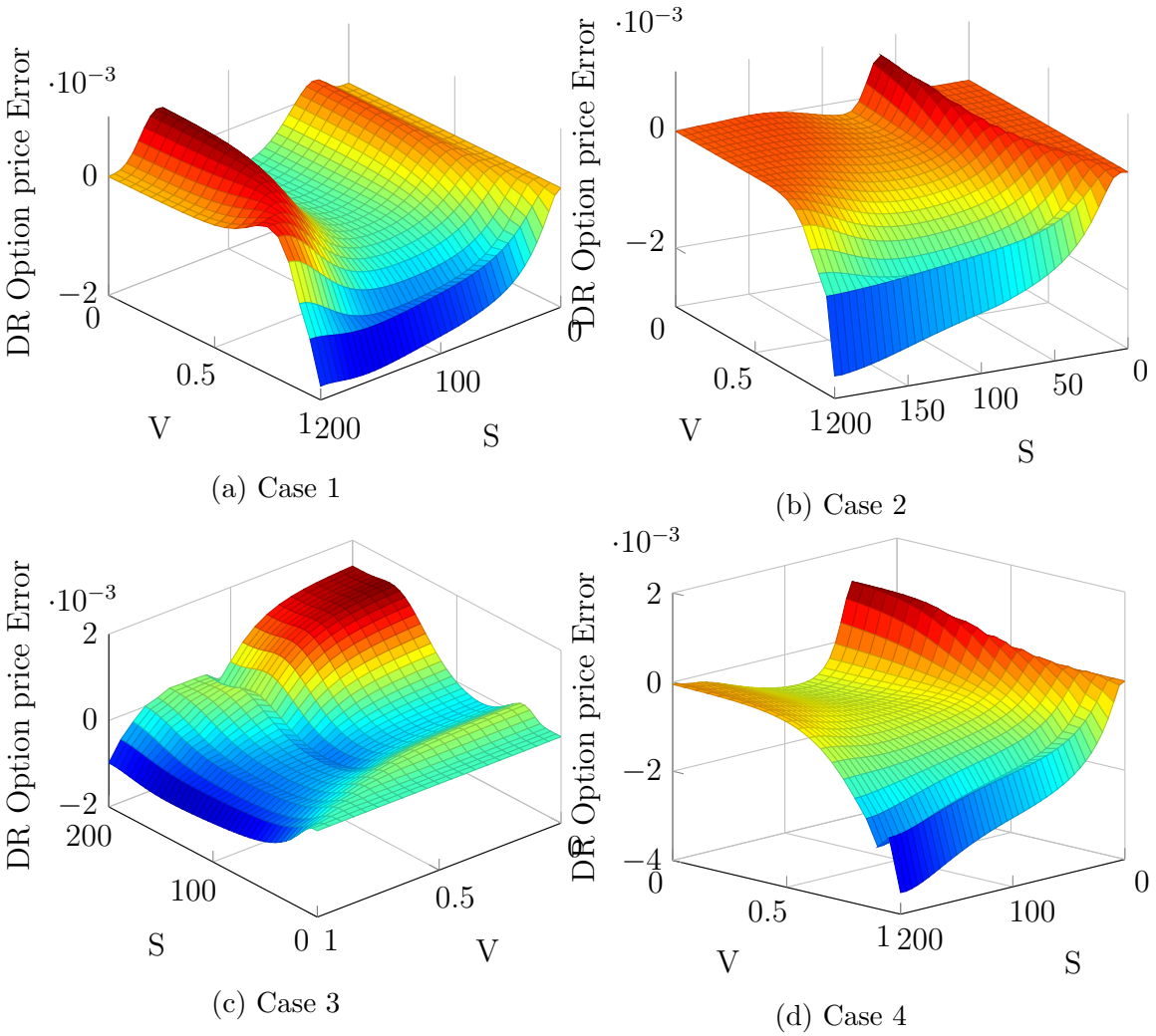


Figure 5.8: Difference between DR scheme with $\theta' = \frac{1}{2}$ and $\theta = 1$ with a non-uniform grid applied.

Some further methods to deal with payoff discontinuities have been proposed. These include an averaging technique outlined in Heston and Zhou (2000) (previously discussed by Thomée and Wahlbin (1974)) adapted to help improve convergence, translating the grid such that critical points have specific grid positions, for example exactly on a grid point or mid-way between adjacent grid points, to help improve accuracy as described in Tavella and Randall (2000) [43] and pro-

jecting initial conditions under L_2 onto a space spanned by a predetermined set of basis functions (Wahlbin (1980)).

A reduction in the spatial error itself may be achieved by applying a non-uniform grid. Table 5.4 presents results for the European call option price under the Heston model for different grid sizes when applying the DR scheme using a non-uniform mesh. The option price is evaluated at the point $(S, \nu) = (70, 0.15)$ as treatment of the boundaries in the non-uniform case needs to be improved and leads to inaccurate error valuations. If this point doesn't lie on a specific grid point then a Matlab interpolation function is used. The error is considerably less than in the uniform case.

Mesh size $(N_S \times N_V)$	DR call option price		Absolute error	
	Uniform	Non-uniform	Uniform	Non-uniform
(21×11)	4.5816	4.7184	0.0756	0.0612
(41×21)	4.6350	4.6657	0.0222	0.0085
(61×31)	4.6467	4.6634	0.0105	0.0062
(81×41)	4.6508	4.6611	0.0064	0.0039
(101×51)	4.6528	4.6594	0.0044	0.0022
(121×61)	4.6534	4.6582	0.0034	0.0010

Table 5.4: DR scheme applied with both uniform and non-uniform grids, option price value given at point $(S, \nu) = (70, 0.15)$.

Chapter 6

Conclusion

This study has shed light on finite difference methods used to approximate derivatives in the Heston PDE pricing European options. The focus has been on the use of different ADI schemes to price a European call option. The advantage of ADI schemes over traditional finite difference methods is the increased accuracy as well as the combination of stability properties attained from the implicit scheme whilst maintaining low computational times by limiting its use of matrix inversion.

We have introduced the topic by first motivating the idea of modeling volatility as a stochastic variable and providing background theory on the derivation of the Heston PDE and its semi-analytical solution. Furthermore, for completion we have presented two further pricing methods that use Fourier transforms. Finite difference theory was then introduced, followed by non-ADI methods including the Explicit, Implicit and Crank-Nicolson schemes.

The Explicit FD scheme was found to be conditionally stable and a non-rigorous stability condition for the time-step Δt was proposed. This condition was proved numerically as number of asset or volatility steps which violated this stability condition were found to produce instability and error growth in the output grid of the scheme. We saw the shape of the error and its slight increase around the critical region of $S = K$ where the payoff presents a discontinuity.

The Greeks were also evaluated and plotted for the explicit FD scheme. The difference grids showing the variation between a given greek and the exact solution were easily attainable as the implementation of exact Greek values followed without much effort from the Heston exact option price solution presented by Shaw [42].

ADI methods were then introduced and this guided us onto numerical experiments. The errors were observed to converge in all four ADI schemes presented and the order of convergence was approx. 2 for all schemes when a uniform grid was applied. The following task was to introduce techniques to improve the present form of the schemes. A simplified damping technique was applied and it was seen to be effective for large values of volatility, however was detrimental to the scheme around the critical area of $\nu = 0$ where it performed worse than without applying the technique. Further techniques that could be tried are upwinding and grid shifting (see Tavella and Randall [43] for the latter), especially to try and reduce the spurious oscillations appearing around $S = K$.

Applying a non-uniform grid allowed us to reduce the error significantly whilst keeping computation time unchanged. It is of interest to implement further splitting schemes to the DR method under a non-uniform grid to see if it is the case that all perform better than in the uniform case. Also, it is worth applying any correction techniques mentioned above to the same ADI schemes but with a non-uniform grid. This will improve accuracy as well as stability. The way the non-uniform grid was generated is as described in Tavella and Randall [43]. Applying a non-uniform mesh is a non-trivial task and as such it may be rewarding to spend time researching an effective and efficient coordinate transformation technique.

One may also try to apply Monte Carlo (MC) methods for reference. Although for European options an exact solution is available in the case of the Heston model, it is worth expanding to a wider range of options once the procedures for simple European options are refined. For example barrier options, briefly described in Appendix A, are a class of exotic options without a closed-form solution. In this case we may use MC simulations where we rely on a risk neutral valuation and obtain the price as the expected value of discounted payoffs. Further, one can apply the COS method or Milstein's method. This may be applied on other exotic options also.

Finally, a further improvement that one may carry out on the Heston model itself is the addition of jumps. The presence of jumps has been shown to account for the higher than expected short term volatility observed in market data. When considering jumps, one needs to make further assumptions such as the frequency and size of the jump. In particular, jumps are necessary to describe short expiry option prices and often resolve problems encountered when calibrating for the vol-of-vol parameter. Incorporating jumps in stochastic volatility models is currently receiving much interest, but it is not any form of focus in this study.

Case 4 presented problems due to having a small value for the vol-of-vol parameter σ implying a convection-dominated Heston PDE. Duffy [15] claims in this case poor performance by use of ADI schemes as was discovered during the analysis. It is of interest to find an alternative scheme in this case, perhaps iterative FD or pure splitting methods, and also compare results with those obtained via ADI schemes presented.

Bibliography

[1] Al Bastami, A., Belić, M.R. and Petrović, N.Z. (2010) *Special Solutions of the Ricatti Equation with Applications to the Gross-Pitaevskii Nonlinear PDE*, Electronic Journal of Differential Equations, Vol. 2010, No. 66, pp. 1-10.

[2] Albrecher, H., Mayer, P., Schoutens, W. and Tistaert, J. (2007) *Wilmott Mag.*, pp. 83-92.

[3] Black, F. and Scholes, M. (1973) *The Pricing of Options and Corporate Liabilities*, The Journal of Political Economy, Vol. 81, No. 3, pp. 637-654.

[4] Bollerslev, T., Gibson, M. and Zhou, H. (2011) *Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities*, Journal of Econometrics, pp. 235-245.

[5] Breeden, D.T. (1979) *An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities*, Journal of Financial Economics, Vol. 7, pp. 265-296.

[6] Carr, P. and Madan, D.B. (1999) *Option Valuation Using the Fast Fourier Transform*, Journal of Computational Finance, Vol. 2(4), pp. 61-73.

[7] Cartea, A. (2013) *Applied Lévy Processes in Finance*, University College London.

[8] Cartea, A. (2013) *Implied and Stochastic Volatility*, University College London.

[9] Cont, R. (2001) *Empirical properties of asset returns: stylized facts and statistical issues*, Quantitative Finance, Vol. 1, pp. 223-236.

[10] Cox, J.C., Ingersoll, Jr., and Ross, S.A. (1985) *A Theory of the Term Structure of Interest Rates*, Econometrica, Vol. 53, Issue 2, pp. 385-408.

[11] Craig, I.J.D. and Sneyd, A.D. (1988) *And alternating-direction implicit scheme for parabolic equations*, Computers & mathematics with applications, Vol. 6, Issue 4, pp. 341-350.

[12] De Graaf, C.S.L. (2012) *Finite Difference Methods in Derivatives Pricing under Stochastic Volatility Models*, Mathematisch Instituut, Universiteit Leiden.

[13] Derman, E. and Kani, I. (1994) *Riding on a smile*, Risk 7, pp. 32-39.

[14] Douglas, J. and Rachford, H.H. (1995) *On the numerical solution of the heat conduction problems in two and three space variables*, Trans. Amer. Math. Soc. 82, pp. 421-439.

[15] Duffy, D.J. (2006) *Finite difference methods in financial engineering: A Partial Differential Equation Approach*, Chichester, West Sussex: John Wiley & Sons Ltd.

[16] Dumas, B., Fleming, J. Whaley, R.E. (1998) *Implied volatility functions: Empirical tests*, The Journal of Finance, Vol. 53.

[17] Dupire, B. (1994) *Pricing with a smile*, Risk 7, pp. 18-20.

[18] Engan, A. (2013) *Numerical Option Pricing in CARMA Models*, Faculty of Mathematics and Natural Sciences, University of Oslo.

[19] Fouque, J.P., Papanicolaou, G. and Sircar, K.R. (2000) *Derivatives in financial markets with stochastic volatility*, Cambridge University Press.

[20] Galiotos, V. (2008) *Stochastic Volatility and the Volatility Smile*, Department of Mathematics, Uppsala University.

[21] Gatheral, J. (2006) *The Volatility Surface: A Practitioner's Guide*, New York, NY: John Wiley & Sons Ltd.

[22] Gil-Pelaez, J. (1951) *Note on the inversion theorem*, Biometrika, Vol. 38(3/4), pp. 481-482.

[23] Haug, E.G. and Taleb, N.N. (2010) *Options traders use (very) sophisticated heuristics, never the Black-Scholes-Merton formula*

[24] Heston, S.L. (1993) *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, The Review of Financial Studies, Vol. 6, Issue 2, pp. 327-343.

[25] Hull, J.C. (2012) *Options, Futures, And Other Derivatives*, Eighth Edition, Essex: Pearson Education Limited.

[26] Hundsdorfer, W. and Verwer, J.G. (2004) *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, SIAM Review, Vol. 46, No. 3, pp. 581-583.

[27] In'T Hout, K.J., and Foulon, S. (2010) *ADI Finite Difference Schemes for Option Pricing in the Heston Model with Correlation*, International Journal of Numerical Analysis and Modeling, Vol. 7, Issue 2, pp. 303-320.

[28] Karlsson, P. *The Heston Model Stochastic Volatility and Approximation*, Department of Economics, Lund University.

[29] Kwok, Y.K. (2008) *Mathematical Models of Financial Derivatives*, Second Edition, Berlin: Springer.

[30] Lewis, A.L. (2000) *Option Voluation Under Stochastic Volatility: With Mathematica Code*, Finance Press.

[31] Lewis, A.L. (2001) *A Simple Option Formula for General Jump-Diffusion and Other Exponential Levy Processes*

[32] Lin, S. (2008) *Finite Difference Schemes for Heston Model*, The University of Oxford.

[33] Mandelbrot, B. (1963) *The Variation of Certain Speculative Prices*, The Journal of Business, Vol. 36, No. 4, pp. 394-419.

[34] Maslova, M. (2008) *Calibration of parameters for the Heston model in the high volatility period of markets*, School of Information Science, Computer and Electrical Engineering, Halmstad University.

[35] Mitchell, A.R. and Griffiths, D.F. (1980) *The Finite Difference Method in Partial Differential Equations*, John Wiley & Sons Ltd.

[36] Moodley, N. (2005) *The Heston Model: A Pratical Approach with Matlab Code*, Faculty of Science, University of Witwatersrand.

[37] Peaceman, D.W. and Rachford, H.H. (1955) *The Numerical Solution of Parabolic and Elliptic Differential Equations*, Journal of the Society for Industrial and Applied Mathematics, Vol. 3, No. 1 (Mar., 1955), pp. 28-41.

[38] Rasmussen, L. (2011) *Pricing the European call option in the Heston model - using finite difference methods*, Faculty of Science, University of Copenhagen.

[39] Richtmyer, R.D. and Morton, K.W. (1994) *Difference Methods for Initial-Value Problems*, Second Revised Edition, Florida: Krieger Publishing Company.

[40] Rouah, F.D. (2013) *The Heston Model and Its Extensions in Matlab and C#*, New Jersey: John Wiley & Sons Ltd.

[41] Schmelzle, M. (2010) *Option Pricing Formulae using Fourier Transform: Theory and Application*

[42] Shaw, W. *Stochastic Volatility Models of Heston Type*, King's College London.

[43] Tavella, D. and Randall, C. (2000) *Pricing Financial Instruments: The Finite Difference Method*, First Edition, New York: John Wiley & Sons Ltd.

[44] Thomas, J.W. (1999) *Numerical Partial Differential Equations, Volume II. Conversion Laws and Elliptic Equations*, New York: Springer.

[45] Tong, Z. (2012) *Option Pricing with Long Memory Stochastic Volatility Models*, Faculty of Science, University of Ottawa.

[46] Verwer, J.G., Spee, E.J., Blom, J.G. and Hundsdorfer, W. (1999) *A second-order Rosenbrock method applied to photochemical dispersion problems*, SIAM J. Sci. Comp., 20, pp. 1456-1480.

[47] Wang, J. (2007) *Convexity of option prices in the Heston model*, Department of Mathematics, Uppsala University.

[48] Wilmott, P. *Paul Wilmott introduces Quantitative Finance*, Second Edition, West Sussex: John Wiley & Sons Ltd.

[49] Xiong, C. (2010) *Valuation of Double Barrier European Options in Heston's Stochastic Volatility Model Using Finite Element Methods*, Department of Mathematics, Rutgers University.

Appendix A

Appendix A

A.1 Details of ADI Scheme

We describe in detail the steps leading to the implementation of the Douglas-Rachford scheme. In the main text we keep Δt in the scheme, but for simplicity during computational implementation we keep Δt as part of the A matrices.

First note how the entries in the tridiagonal matrices A_1 and A_2 are split. We have

lower diagonal:

$$\frac{1}{2}\nu_j S_i^2 \frac{\Delta t}{\Delta S^2} - \frac{1}{2} r S_i \frac{\Delta t}{\Delta S} \tag{A.1}$$

main diagonal:

$$-\nu_j S_i^2 \frac{\Delta t}{\Delta S^2} - \frac{1}{2} r \Delta t \tag{A.2}$$

upper diagonal:

$$\frac{1}{2}\nu_j S_i^2 \frac{\Delta t}{\Delta S^2} + \frac{1}{2} r S_i \frac{\Delta t}{\Delta S} \tag{A.3}$$

where the lower, main and upper diagonals correspond to $U_{i-1,j}$, $U_{i,j}$ and $U_{i+1,j}$ respectively. Similarly the matrix A_2 is split as follows

lower diagonal:

$$\frac{1}{2}\sigma^2 \nu_j \frac{\Delta t}{\Delta \nu^2} - \frac{1}{2}\kappa(\theta - \nu_j)\frac{\Delta t}{\Delta \nu} \tag{A.4}$$

main diagonal:

$$-\sigma^2 \nu_j \frac{\Delta t}{\Delta \nu^2} - \frac{1}{2} r \Delta t \tag{A.5}$$

upper diagonal:

$$\frac{1}{2}\sigma^2 \nu_j \frac{\Delta t}{\Delta \nu^2} + \frac{1}{2}\kappa(\theta - \nu_j)\frac{\Delta t}{\Delta \nu} \tag{A.6}$$

Then the first step of the two step scheme is

$$(1 - \theta_1' A_1) U_{i,j}^* = \ (1 + A_0 + (1 - \theta') A_1 + A_2) U_{i,j}^n \tag{A.7}$$

$$\Rightarrow U^*_{i,j} - \theta' \left(\frac{1}{2}\nu_j S_i^2 \frac{\Delta t}{\Delta S^2} \left(U^*_{i+1,j} - 2U^*_{i,j} + U^*_{i-1,j} \right) + \frac{1}{2} r S_i \frac{\Delta t}{\Delta S} \left(U^*_{i+1,j} - U^*_{i-1,j} \right) - \frac{1}{2} r \Delta t U^*_{i,j} \right) = \left(1 + A_0 + (1-\theta')A_1 + A_2\right) U^n_{i,j} \tag{A.8}$$

Rearranging the terms on the LHS we can form coefficients for $U^*_{i-1,j}$, $U^*_{i,j}$ and $U^*_{i+1,j}$. These are given by

$$a_{1,j} = -\frac{\theta'}{2}\nu_j S_i^2 \frac{\Delta t}{\Delta S^2} + \frac{\theta'}{2} r S_i \frac{\Delta t}{\Delta S} \tag{A.9a}$$

$$b_{1,j} = 1 + \theta' \nu_j S_i^2 \frac{\Delta t}{\Delta S^2} + \frac{\theta'}{2} r \Delta t \tag{A.9b}$$

$$c_{1,j} = -\frac{\theta'}{2}\nu_j S_i^2 \frac{\Delta t}{\Delta S^2} - \frac{\theta'}{2} r S_i \frac{\Delta t}{\Delta S} \tag{A.9c}$$

such that we have

$$a_{1,j} U^*_{i-1,j} + b_{1,j} U^*_{i,j} + c_{1,j} U^*_{i+1,j} = \left(1 + A_0 + (1-\theta')A_1 + A_2\right) U^n_{i,j} \tag{A.10}$$

The system to be solved is then as given in the ADI implementation section.

A.2 Tridiagonal Matrix Algorithm (TDMA)

The Tridiagonal Matrix Algorithm (also known as the Thomas algorithm, named after Llevellyn Thomas) is a form of Gaussian elimination used to solve tridiagonal systems of equations. The algorithm is composed of two steps; the first consists of a forward elimination sweep and the second is a result of applying backward substitution. Consider the following tridiagonal system:

$$a_i x_{i-1} + b_i x_i + c_i x_{i+1} = d_i,\ i = 1, ..., n \tag{A.11}$$

which can be written in matrix form, with $a_1 = 0$ and $c_n = 0$, as follows

$$\begin{pmatrix} b_1 & c_1 & 0 & \dots & \dots & 0 \\ a_2 & b_2 & c_2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & \dots & \dots & 0 & a_n & b_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_{n-1} \\ d_n \end{pmatrix} \tag{A.12}$$

Now perform the following row operations to system A.12. First subtract row two multiplied by a_2 from row one multiplied by b_1. We have

$$b_1(a_2 x_1 + b_2 x_2 + c_2 x_3) - a_2(b_1 x_1 + c_1 x_2) = b_1 d_2 - a_2 d_1 \tag{A.13}$$

Simplifying yields

$$(b_1 b_2 - a_2 c_1) x_2 + b_1 c_2 x_3 = b_1 d_2 - a_2 d_1 \tag{A.14}$$

such that x_1 has been eliminated from the second equation. Similarly, we are able to eliminate x_2 using the modified second equation and the third equation. This gives

$$\begin{aligned}(b_1b_2 - a_2c_1)(a_3x_2 + b_3x_3 + c_3x_4) - a_3((b_1b_2 - a_2c_1)x_2 + b_1c_2x_3)\\ = (b_1b_2 - a_2c_1)d_3 - a_3(b_1d_2 - a_2d_1)\\ \Rightarrow((b_1b_2 - a_2c_1)b_3 - a_3b_1c_2)x_3 + (b_1b_2 - a_2c_1)c_3x_4\\ = (b_1b_2 - a_2c_1)d_3 - a_3(b_1d_2 - a_2d_1)\end{aligned}$$

Repeating this procedure on all n rows, one eventually arrives at the last equation which involves the unknown x_n only. The unknown x_n will be embedded in a final function which may be used to solve the modified equation for $i = n - 1$ etc. until all x_i are found via back subsitution. In effect we are left with the upper diagonal system

$$\begin{pmatrix} b'_1 & c'_1 & 0 & \dots & \dots & 0 \\ 0 & b'_2 & c'_2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & 0 & b'_{n-1} & c'_{n-1} \\ 0 & \dots & \dots & 0 & 0 & b'_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} d'_1 \\ d'_2 \\ \vdots \\ \vdots \\ d'_{n-1} \\ d'_n \end{pmatrix} \tag{A.15}$$

where

$$b'_i = 1,\ i = 1, ..., n \tag{A.16a}$$

$$c'_i = \begin{cases} \frac{c_i}{b_i}, & i = 1 \\ \frac{c_i}{b_i - a_i c'_{i-1}}, & i = 2, 3, ..., n-1 \end{cases} \tag{A.16b}$$

$$d'_i = \begin{cases} \frac{d_i}{b_i}, & i = 1 \\ \frac{d_i - a_i d'_{i-1}}{b_i - a_i c'_{i-1}}, & i = 2, 3, ..., n-1 \end{cases} \tag{A.16c}$$

where the solution may then be obtained by back substitution

$$\begin{aligned} x_n &= d'_n \\ x_i &= d'_i - c'_i x_{i+1},\ i = n-1, n-2, ..., 1 \end{aligned} \tag{A.17}$$

A.3 Shaw's Approach to a Closed-Form Solution to the Heston Model

This section provides a short summary of the approach taken, to compute the semi-analytical solution to the Heston PDE, by Shaw [42].

A.3.1 Change of Variables

The first step is to perform the following change of variables (assume there are no dividends): $\tau = T - t$, $x = \log S + r\tau$ and $U = We^{-r\tau}$ (where $W = W(x, \nu, \tau)$).

By the chain rule

$$\frac{\partial U}{\partial t} = W\frac{\partial}{\partial \tau}(e^{-r\tau})\frac{\partial \tau}{\partial t} + e^{-r\tau}\left(\frac{\partial W}{\partial \tau}\frac{\partial \tau}{\partial t} + \frac{\partial W}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial W}{\partial \nu}\frac{\partial \nu}{\partial t}\right) \tag{A.18a}$$

$$\frac{\partial U}{\partial S} = e^{-r\tau}\left(\frac{\partial W}{\partial x}\frac{\partial x}{\partial S} + \frac{\partial W}{\partial \nu}\frac{\partial \nu}{\partial x} + \frac{\partial W}{\partial \tau}\frac{\partial \tau}{\partial x}\right) = e^{-r\tau}\frac{1}{S}\frac{\partial W}{\partial x} \tag{A.18b}$$

$$\frac{\partial^2 U}{\partial S^2} = e^{-r\tau}\frac{\partial}{\partial S}\left(\frac{1}{S}\frac{\partial W}{\partial x}\right) = e^{-r\tau}\frac{1}{S^2}\left(\frac{\partial^2 W}{\partial x^2} - \frac{\partial W}{\partial x}\right) \tag{A.18c}$$

$$\frac{\partial U}{\partial \nu} = e^{-r\tau}\frac{\partial W}{\partial \nu} \tag{A.18d}$$

$$\frac{\partial^2 U}{\partial \nu^2} = e^{-r\tau}\frac{\partial^2 W}{\partial \nu^2} \tag{A.18e}$$

$$\frac{\partial^2 U}{\partial \nu \partial S} = \frac{\partial}{\partial \nu}\left(\frac{\partial U}{\partial S}\right) = \frac{\partial}{\partial \nu}\left(e^{-r\tau}\frac{1}{S}\frac{\partial W}{\partial x}\right) = e^{-r\tau}\frac{1}{S}\frac{\partial^2 W}{\partial \nu \partial x} \tag{A.18f}$$

Substituting into the Heston PDE gives us a PDE for W

$$\frac{\partial W}{\partial \tau} = \frac{1}{2}\nu\left(\frac{\partial^2 W}{\partial x^2} - \frac{\partial W}{\partial x}\right) + \rho\sigma\nu\frac{\partial^2 W}{\partial \nu \partial x} + \frac{1}{2}\sigma^2\nu\frac{\partial^2 W}{\partial \nu^2} + (\kappa(\theta - \nu) - \lambda\nu)\frac{\partial W}{\partial \nu} \tag{A.19}$$

A.3.2 The Fourier Transform

The following step is to introduce an inverse Fourier Transform as follows

$$W(x, \nu, \tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iwx}\widetilde{W}(w, \nu, \tau)dw \tag{A.20}$$

where the Fourier Transform is

$$\mathcal{F}\{W(x, \nu, \tau)\} = \widetilde{W}(w, \nu, \tau) = \int_{-\infty}^{\infty} e^{iwx}W(x, \nu, \tau)dx \tag{A.21}$$

Notice that at expiry, where $\tau = 0$, we have

$$\begin{aligned}
\widetilde{W}(w, \nu, 0) &= \int_{-\infty}^{\infty} e^{iwx}W(x, \nu, 0)dx \\
&= \int_{-\infty}^{\infty} e^{iwx}U(x, \nu, 0)dx \\
&= \int_{-\infty}^{\infty} e^{iwx}max(e^x - K, 0)dx \\
&= \int_{\log K}^{\infty} e^{iwx}(e^x - K)dx \\
&= \left[\frac{1}{iw+1}e^{(iw+1)x} - \frac{K}{iw}e^{iwx}\right]_{\log K}^{\infty} \\
&= -\frac{K^{iw+1}}{w^2 - iw}
\end{aligned} \tag{A.22}$$

where the derivation is identical to that in Section 3.5. Recall that it is necessary to check convergence of the integral as w could be any complex number. We don't want unbounded growth of the exponential terms so look for them to decay as x increases. Recall for a call option this will happen whenever $Im(w) > 1$.

A.3.3 The Transformed PDE

Differentiating with respect to x becomes multiplication by $-iw$ in the Fourier transform. Note the following

$$\frac{\partial W}{\partial \tau} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iwx}\frac{\partial \widetilde{W}}{\partial \tau}dw \tag{A.23a}$$

$$\frac{\partial W}{\partial x} = -\frac{iw}{2\pi}\int_{-\infty}^{\infty} e^{-iwx}\widetilde{W}dw \tag{A.23b}$$

$$\frac{\partial^2 W}{\partial x^2} = -\frac{w^2}{2\pi}\int_{-\infty}^{\infty} e^{-iwx}\widetilde{W}dw \tag{A.23c}$$

$$\frac{\partial W}{\partial \nu} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iwx}\frac{\partial \widetilde{W}}{\partial \nu}dw \tag{A.23d}$$

$$\frac{\partial}{\partial \nu}\left(\frac{\partial W}{\partial x}\right) = -iw\frac{\partial}{\partial \nu}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iwx}\widetilde{W}dw\right) = -\frac{iw}{2\pi}\int_{-\infty}^{\infty} e^{-iwx}\frac{\partial \widetilde{W}}{\partial \nu}dw \tag{A.23e}$$

$$\begin{aligned}\Rightarrow \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iwx}\frac{\partial \widetilde{W}}{\partial \tau}dw = &-\frac{\nu}{4\pi}(w^2 - iw)\int_{-\infty}^{\infty} e^{-iwx}\widetilde{W}dw + (\kappa(\theta-\nu)\\ &-\lambda\nu - iw\rho\sigma\nu)\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iwx}\frac{\partial \widetilde{W}}{\partial \nu}dw\\ &+\frac{\sigma^2\nu}{4\pi}\int_{-\infty}^{\infty} e^{iwx}\frac{\partial^2 \widetilde{W}}{\partial \nu^2}dw\end{aligned} \tag{A.24}$$

Take the Fourier transform and we have that

$$\frac{\partial \widetilde{W}}{\partial \tau} = -\frac{1}{2}\nu(w^2 - iw)\widetilde{W} + (\kappa(\theta-\nu) - \lambda\nu - iw\rho\sigma\nu)\frac{\partial \widetilde{W}}{\partial \nu} + \frac{1}{2}\sigma^2\nu\frac{\partial^2 \widetilde{W}}{\partial \nu^2} \tag{A.25}$$

(where we use that $\mathcal{F}\left\{\frac{\partial W}{\partial \tau}\right\} = \frac{\partial \widetilde{W}}{\partial \tau}$).

A.3.4 The Fundamental Solution

Assume there exists a solution $G(w,\nu,\tau)$ of PDE (A.25). Assume further that the solution has the property $G(w,\nu,0) = 1$. The solution to the transformed PDE with payoff condition $\widetilde{W}(w,\nu,0)$ is then the product of the payoff condition with $G(w,\nu,\tau)$. The solution to the original PDE is then the discounted value of this product with any coordinate transformations undone.

$$U = \frac{1}{2\pi}e^{-r(T-t)}\int_{iw_i-\infty}^{iw_i+\infty} e^{-iwx}\widetilde{W}(w,\nu,0)G(w,\nu,T-t)dw \tag{A.26}$$

where $x = \log S + r(T-t)$.

The next step is to find G. Lewis (2000) [30] studies the process of finding G with more general market price of risk functions. As we are interested in the Heston model, we find G for this specific model.

$$\frac{\partial G}{\partial \tau} = -\frac{1}{2}\nu(w^2 - iw)G + (\kappa(\theta-\nu) - \lambda\nu - iw\rho\sigma\nu)\frac{\partial G}{\partial \nu} + \frac{1}{2}\sigma^2\nu\frac{\partial^2 G}{\partial \nu^2} \tag{A.27}$$

such that $G(w,\nu,0)=1$. Ansatz is $G=e^{C(\tau,w)+\nu D(\tau,w)}$ (we use this form for the solution for example when dealing with bond pricing equations in affine models also). From the payoff condition we require $C(0,w)=0=D(0,w)$ in order to satisfy $G=1$ at expiry. Substituting the ansatz into the PDE (A.27) gives

$$\left(\frac{\partial C}{\partial \tau}+\nu\frac{\partial D}{\partial \tau}\right)G = \ -\frac{1}{2}\nu(w^2-iw)G+(\kappa(\theta-\nu)-\lambda\nu-iw\rho\sigma\nu)DG+\frac{1}{2}\sigma^2\nu D^2G \tag{A.28}$$

which must hold for all ν.

$$\Rightarrow \left[\frac{\partial C}{\partial \tau}-\theta\kappa D\right]+\left[\frac{\partial D}{\partial \tau}+\frac{1}{2}(w^2-iw)+\left[\kappa+\lambda+iw\rho\sigma\right]D-\frac{1}{2}\sigma^2D^2\right]\nu=0 \tag{A.29}$$

This yields the following two ODE's

$$\frac{\partial C}{\partial \tau}=\theta\kappa D \tag{A.30a}$$

$$\frac{\partial D}{\partial \tau}=\frac{1}{2}\sigma^2D^2-D(\kappa+\lambda+iw\rho\sigma)-\frac{1}{2}(w^2-iw) \tag{A.30b}$$

Solve the second equation first for D, then the first for C. Express the solution in terms of auxiliary functions d, g defined as follows

$$d=\sqrt{(w^2-iw)\sigma^2+(\kappa+\lambda+iw\rho\sigma)^2} \tag{A.31a}$$

$$g=\frac{\kappa+\lambda+iw\rho\sigma+d}{\kappa+\lambda+iw\rho\sigma-d} \tag{A.31b}$$

Then

$$D=\frac{\kappa+\lambda+iw\rho\sigma+d}{\sigma^2}\frac{1-e^{d\tau}}{1-ge^{d\tau}} \tag{A.32a}$$

$$C=\frac{\kappa\theta}{\sigma^2}\left\{(\kappa+\lambda+iw\rho\sigma+d)\tau-2\ln\left(\frac{1-ge^{d\tau}}{1-g}\right)\right\} \tag{A.32b}$$

Note that it is better to perform numerical integration of the ODE for C directly as then one can avoid any branch cut difficulties which are a result of the branch of the complex logarithm. With this we conclude the solution of the Heston model. Recall that we take the market price of risk $\lambda=0$ in our computational experiments. In order to move on to price instruments, all that is left is to compute the inverse transform integrals.

A.3.5 The Greeks

The reason we present the solution by Shaw [42] is the simplicity in computation of the Greeks which follows its representation. We can easily obtain delta and gamma by multiplying the integrand in (A.26) by $-\frac{iw}{S}$ and $-\frac{w^2}{S^2}$, respectively.

A.4 General Stability Approach in the Explicit Scheme

A.4.1 Von Neumann Stability

Consider the general IBVP in two space dimensions with time independent coefficients and boundary conditions given as follows

$$\begin{aligned}\frac{\partial U}{\partial t} &= b_1(x,y)\frac{\partial U}{\partial x} + b_2(x,y)\frac{\partial U}{\partial y} + a_1(x,y)\frac{\partial^2 U}{\partial x^2} + a_2(x,y)\frac{\partial^2 U}{\partial y^2} - c(x,y)U, \\ &\quad \forall (x,y,t) \in (0,T]\times\Omega \\ U(x,y;0) &= g(x,y),\ \forall (x,y)\in\overline{\Omega} \\ U(x,y;t) &= 0,\ \forall (x,y)\in(0,T]\times\delta\Omega\end{aligned} \tag{A.33}$$

where c is non-negative, and a_1, a_2 are striclty positive. We remark that the spatial domain is defined as in the main text, namely we consider $\Omega = (X_{min}, X_{max}) \times (Y_{min}, Y_{max})$. Note that the discretized form of equation A.33 is

$$\begin{aligned}U_{i,j}^{n+1} = \ U_{i,j}^n + \Delta t \Bigg(& b_1(x,y)\frac{U_{i+1,j}^n - U_{i-1,j}^n}{2\Delta x} + b_2(x,y)\frac{U_{i,j+1}^n - U_{i,j-1}^n}{2\Delta y} \\ & + a_1(x,y)\frac{U_{i+1,j}^n - 2U_{i,j}^n + U_{i-1,j}^n}{\Delta x^2} + a_2(x,y)\frac{U_{i,j+1}^n - 2U_{i,j}^n + U_{i,j-1}^n}{\Delta y^2} - c(x,y)U_{i,j}^n \Bigg)\end{aligned} \tag{A.34}$$

Following the Von Neumann stability approach, let p and q denote spatial indices (instead of i,j as we reserve i to denote the imaginary unit $i = \sqrt{-1}$) and let $U_{p,q}^n = \bar{a}^n e^{i(pk_x\Delta x + qk_y\Delta y)}$. Substituting this into equation (A.34) gives (dropping subscripts in coefficients for notational convenience)

$$\begin{aligned}\bar{a} = & \left[1 - \frac{2\Delta t}{\Delta x^2}a_1 - \frac{2\Delta t}{\Delta y^2}a_2 - c\Delta t\right] + \left[a_1\frac{\Delta t}{\Delta x^2} - b_1\frac{\Delta t}{2\Delta x}\right]e^{-ik_x\Delta x} \\ & + \left[b_1\frac{\Delta t}{2\Delta x} + a_1\frac{\Delta t}{\Delta x^2}\right]e^{ik_x\Delta x} + \left[a_2\frac{\Delta t}{\Delta y^2} - b_2\frac{\Delta t}{2\Delta y}\right]e^{-ik_y\Delta y} \\ & + \left[b_2\frac{\Delta t}{2\Delta y} + a_2\frac{\Delta t}{\Delta y^2}\right]\end{aligned} \tag{A.35}$$

$$\Updownarrow$$

$$\begin{aligned}\bar{a} - 1 = & \ a_1\frac{\Delta t}{\Delta x^2}\left[e^{-ik_x\Delta x} - 2 + e^{ik_x\Delta x}\right] + a_2\frac{\Delta t}{\Delta y^2}\left[e^{-ik_y\Delta y} - 2 + e^{ik_y\Delta y}\right] \\ & + b_1\frac{\Delta t}{2\Delta x}\left[e^{ik_x\Delta x} - e^{-ik_x\Delta x}\right] + b_2\frac{\Delta t}{2\Delta y}\left[e^{ik_y\Delta y} - e^{-ik_y\Delta y}\right] - c\Delta t\end{aligned} \tag{A.36}$$

Using the following useful result

$$e^{ik_x\Delta x} + e^{-ik_x\Delta x} - 2 = \ -4\sin^2\left(\frac{1}{2}k_x\Delta x\right) \tag{A.37}$$

we have

$$\begin{aligned}\bar{a} = 1 - c\Delta t - 4\left[\frac{\Delta t}{\Delta x^2}a_1 \sin^2\left(\frac{1}{2}k_x\Delta x\right) + \frac{\Delta t}{\Delta y^2}a_2 \sin^2\left(\frac{1}{2}k_y\Delta y\right)\right] \\ + i\left[b_1\frac{\Delta t}{\Delta x}\sin\left(k_x\Delta x\right) + b_2\frac{\Delta t}{\Delta y}\sin\left(k_y\Delta y\right)\right]\end{aligned} \tag{A.38}$$

Note that since the amplification is a complex number, the finite difference approximation will affect both the amplitude and phase of the oscillatory solution. Now, the absolute value of $\bar{a}$ is the square root of the product of $\bar{a}$ with its complex conjugate. This implies

$$\begin{aligned}|\bar{a}|^2 = \left(1 - c\Delta t - 4\left[\frac{\Delta t}{\Delta x^2}a_1 \sin^2\left(\frac{1}{2}k_x\Delta x\right) + \frac{\Delta t}{\Delta y^2}a_2 \sin^2\left(\frac{1}{2}k_y\Delta y\right)\right]\right)^2 \\ + \left[b_1\frac{\Delta t}{\Delta x}\sin\left(k_x\Delta x\right) + b_2\frac{\Delta t}{\Delta y}\sin\left(k_y\Delta y\right)\right]^2\end{aligned} \tag{A.39}$$

The necessary and sufficient Von Nuemann condition for stability is $|\bar{a}| \leq 1$. Recall that c is non-negative. Further assume that a_1, $a_2 > 0$. Then the first term in the equation above is ≤ 1, and noting further that the maximum value sine squared function can take with the given arguments is 1, we have

$$\begin{aligned}|\bar{a}|^2 &\leq 1 + \Delta t^2\left(\frac{b_1}{\Delta x} + \frac{b_2}{\Delta y}\right)^2 \\ &\Rightarrow |\bar{a}| \leq 1 + \frac{1}{2}\Delta t^2\left(\frac{b_1}{\Delta x} + \frac{b_2}{\Delta y}\right)^2 + \mathcal{O}(\Delta t^2) \\ &\Rightarrow |\bar{a}| \leq 1 + \mathcal{O}(\Delta t^2)\end{aligned} \tag{A.40}$$

where the last line is a result of TSE. So the Explicit FD scheme is conditionally stable.

A.4.2 The Truncation Error and its Consistency Implications

The truncation error in the Explicit FD scheme is given by

$$\begin{aligned}\tau(x,y,t) &:= \frac{\Delta_{+t}u(x,y,t)}{\Delta t} - Lu(x,y,t) \\ &= \frac{\Delta_{+t}u(x,y,t)}{\Delta t} - b_1(x,y)\frac{\delta_x u(x,y,t)}{2\Delta x} - b_2(x,y)\frac{\delta_y u(x,y,t)}{2\Delta y} \\ &- a_1(x,y)\frac{\delta_x^2 u(x,y,t)}{\Delta x^2} - a_2(x,y)\frac{\delta_y^2 u(x,y,t)}{\Delta y^2} + c(x,y)u(x,y,t)\end{aligned} \tag{A.41}$$

We now use TSE to obtain

$$\begin{aligned}\tau(x,y,t) = &\frac{u_t\Delta t + \frac{1}{2}u_{tt}\Delta t^2 + \frac{1}{6}u_{ttt}\Delta t^3 + \mathcal{O}(\Delta t^4)}{\Delta t} - b_1(x,y)\frac{u_x\Delta x + \frac{1}{6}u_{xxx}\Delta x^3 + \mathcal{O}(\Delta x^5)}{\Delta x}\\ &- b_2(x,y)\frac{u_y\Delta y + \frac{1}{6}u_{yyy}\Delta y^3 + \mathcal{O}(\Delta y^5)}{\Delta y} - a_1(x,y)\frac{u_{xx}\Delta x^2 + \frac{1}{12}u_{xxxx}\Delta x^4 + \mathcal{O}(\Delta x^6)}{\Delta x^2}\\ &- a_2(x,y)\frac{u_{yy}\Delta y^2 + \frac{1}{12}u_{yyyy}\Delta y^4 + \mathcal{O}(\Delta y^6)}{\Delta y^2} + c(x,y)u(x,y,t)\\ = &\, u_t - [b_1(x,y)u_x + b_2(x,y)u_y + a_1(x,y)u_{xx} + a_2(x,y)u_{yy} - c(x,y)u(x,y,t)]\\ &+ \mathcal{O}(\Delta t + \Delta x^2 + \Delta y^2)\end{aligned} \tag{A.42}$$

where u_t, u_{tt} and u_{ttt} are the first, second and third derivatives of u with respect to time t, respectively. Similar definitions apply for spatial variables x and y. It's clear that the truncation error is first order accurate in time, and second order accurate in space. In addition, since $\tau(x,y,t) \to 0$ as Δt, Δx, $\Delta y \to 0$, the Explicit FD scheme is consistent.

A.4.3 Convergence

The error of the Explicit FD scheme at the node (x_i, y_j, t_n) is given by

$$e^n_{i,j} = \; U^n_{i,j} - u(x_i, y_j, t_n) \tag{A.43}$$

where $U^n_{i,j}$ denotes the approximation to the exact solution $u(x_i, y_j, t_n)$. Since $U^n_{i,j}$ is the solution to the difference equation it has no truncation error, however $u^n_{i,j}$ leaves the residual truncation $\tau^n_{i,j}$. Applying the truncation formula to the error as defined above, we have

$$\begin{aligned}\tau e^n_{i,j} = &\; -\tau(x_i, y_j, t_n)\\ = &\; \frac{\Delta_{+t}e^n_{i,j}}{\Delta t} - Le^n_{i,j}\end{aligned} \tag{A.44}$$

$\Rightarrow$

$$\begin{aligned}e^{n+1}_{i,j} = &\, e^n_{i,j}\left(1 - \Delta t c(x,y) - 2a_1(x,y)\frac{\Delta t}{\Delta x^2} - 2a_2(x,y)\frac{\Delta t}{\Delta y^2}\right) - \Delta t\tau(x_i, y_j, t_n)\\ &+ e^n_{i-1,j}\left(a_1(x,y)\frac{\Delta t}{\Delta x^2} - b_1(x,y)\frac{\Delta t}{2\Delta x}\right) + e^n_{i+1,j}\left(a_1(x,y)\frac{\Delta t}{\Delta x^2} + b_1(x,y)\frac{\Delta t}{2\Delta x}\right)\\ &+ e^n_{i,j-1}\left(a_2(x,y)\frac{\Delta t}{\Delta y^2} - b_2(x,y)\frac{\Delta t}{2\Delta y}\right) + e^n_{i,j+1}\left(a_2(x,y)\frac{\Delta t}{\Delta y^2} + b_2(x,y)\frac{\Delta t}{2\Delta y}\right)\end{aligned} \tag{A.45}$$

For numerical stability, we require all coefficients linearly associated with the error terms e to be non-negative. As such, the followsing conditions are required

$$\frac{|b_1(x,y)}{\frac{1}{\Delta x}a_1(x,y)} \le 2,\; \frac{|b_2(x,y)|}{\frac{1}{\Delta y}a_2(x,y)} \le 2,\; \Delta t c(x,y) + 2a_1(x,y)\frac{\Delta t}{\Delta x^2} + 2a_2(x,y)\frac{\Delta t}{\Delta y^2}$$

for $i = 0, 1, ..., N_S$, $j = 0, 1, ..., N_V$. Now note that

$$\begin{aligned}|e^{n+1}_{i,j}| \le &\,(1 - c(x,y))\max\left\{|e^n_{i-1,j}|, |e^n_{i+1,j}|, |e^n_{i,j-1}|, |e^n_{i,j+1}|, |e^n_{i,j}|\right\}\\ &+ \Delta t L(\Delta t + \Delta S^2 + \Delta\nu^2)\end{aligned} \tag{A.46}$$

and define

$$E^n = \max_{0\leq i\leq N_S, 0\leq j\leq N_V} |e^n_{i,j}| \tag{A.47a}$$

$$c = \max_{0\leq i\leq N_S, 0\leq j\leq N_V} |c(x,y)| \tag{A.47b}$$

such that

$$E^{n+1} \leq E^n(1+\Delta tc) + L(\Delta t + \Delta x^2 + \Delta y^2) \tag{A.48}$$

By induction

$$\begin{aligned} E^n &\leq E^0(1+\Delta tc)^n + \Delta tL(\Delta t + \Delta x^2 + \Delta y^2)\sum_{k=0}^{n-1}(1+\Delta tc)^k \\ &= \frac{\Delta tL(\Delta t + \Delta x^2 + \Delta y^2)((1+\Delta tc)^n - 1)}{\Delta tc} \\ &\leq e^{n\Delta tc}\tilde{L}(\Delta t + \Delta x^2 + \Delta y^2) \end{aligned} \tag{A.49}$$

$\Rightarrow$ convergence.

A.5 Barrier Options

Barrier options are European vanilla options with the added feature of a *barrier*, which is a set value of the underlying that when crossed alters the contract. In the case of European options the payoff depends only on the value of the underlying at maturity and thus these options are known as *path-independent*. Barrier options are *path-dependent* as their value can be affected during the life-cycle of the option.

There are two types of barrier options; knock-out and knock-in. If the barrier is crossed a knock-out option means the option ceases to exist whereas a knock-in option would bring the option into existence. Barrier options are cheaper than vanilla options and thus serve as a way of hedging when the trader believes the stock will not make very large movements. Each type of barrier option can be sub-divided into two further categories; knock-out options can be either down-and-out or up-and-out and knock-in options can be either down-and-in or up-and-in. This section focuses on knock-out call options.

Down-and-out call: The barrier would be placed below the current spot price and if it is triggered during the lifetime of the option then the option becomes worthless. If maturity is reached without triggering the barrier then the payoff is the same as a vanilla European call. We thus have

$$\text{Payoff} = \begin{cases} \max S_t - K & \text{if } S_t > B, \forall 0 \leq t \leq T, \\ 0 & \text{otherwise} \end{cases}$$

where S_t denotes the price of the underlying at time t, K the strike and B is the barrier level as agreed on the initial contract.

Up-and-out call: The barrier in this case is placed above the current spot price of the underlying. If it is triggered during the lifetime of the option then the option becomes worthless. In this case the payoff is as follows

$$\text{Payoff} = \begin{cases} \max S_t - K & \text{if } S_t < B, \forall 0 \leq t \leq T, \\ 0 & \text{otherwise} \end{cases}$$

We now focus on the down-and-out call option. The spatial non-uniform FD discretization can be adapted to cope with barrier options. Let $B \in (0, K)$ denote the barrier as before, then the region of interest becomes $[B, S_{max}] \times [0, \nu_{max}]$. The boundary conditions (4.3) and (4.5) change to $U(B, \nu, t) = 0$, $(0 \leq t \leq T)$, and $U(S, \nu_{max}, t) = S - B$, $(0 \leq t \leq T)$, respectively. The equidistant points ψ_i in the S-direction are now calculated as follows

$$\xi_i = \sinh^{-1}\left(\frac{B-K}{c}\right) + i\Delta\xi \tag{A.50}$$

where $0 \leq i \leq N_S$. Each step is the evaluated via

$$\Delta\xi = \frac{1}{N_S}\left(\sinh^{-1}\left(\frac{S-K}{c}\right) - \sinh^{-1}\left(\frac{B-K}{c}\right)\right) \tag{A.51}$$

The points in the ν-direction are as before. Only a handful of small changes are required to implement a down-and-out call in place of the vanilla call option. There exists a semi-analytical formula derived by Lipton (2001) for the price of double[1] barrier options in the Heston model as long as the correlation between the underlying and its volatility is zero and the domestic rate is the same as the foreign rate (which for our purposes has been the case throughout the dissertation).

[1] In this case there are two distinct barriers within which the option may fluctuate.