NTNU – Trondheim
Norwegian University of
Science and Technology

Spectral Discretizations of Option Pricing Models for European Put Options

Yngvild Neset

Master of Science in Physics and Mathematics
Submission date: June 2014
Supervisor: Espen Robstad Jakobsen, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Preface

This thesis completes the 5 year Master of Science program in Applied Physics and Mathematics at NTNU. I would like to use this opportunity to thank my supervisor, professor Espen Robstad Jakobsen, for help and guidance throughout the work.

Trondheim, June 2014

Yngvild Neset

Abstract

The aim of this thesis is to solve option pricing models efficiently by using spectral methods. The option pricing models that will be solved are the Black-Scholes model and Heston's stochastic volatility model. We will restrict us to pricing European put options.

We derive the partial differential equations governing the two models and their corresponding weak formulations. The models are then solved using both the spectral Galerkin method and a polynomial collocation method. The numerical solutions are compared to the exact solution. The exact solution is also used to study the numerical convergence. We compare the results from the two numerical methods, and look at the time consumptions of the different methods. Analysis of the methods are also given. This includes coercivity, continuity, stability and convergence estimates.

For Black-Scholes equation, we study both the original equation and the log transformed equation, and we also compare the results to a solution obtained by using a finite element method solver.

Sammendrag

Målet med denne oppgaven er å løse opsjonsprisingsproblemer effektivt ved å bruke spektralmetoder. Vi fokuserer på to ulike opsjonsprisingsmodeller, Black-Scholes modell og Hestons stokastiske volatilitetsmodell. Vi fokuserer på å prise europeiske salgsopsjoner.

Vi utleder modellens partielle differensialligninger og de tilhørende svake formuleringene. Deretter løses ligningene ved å bruke en spektral Galerkinmetode og en polynombasert kollokasjonsmetode. De numeriske løsningnene sammelignes med den eksakte løsningen. Denne brukes også til å studere numerisk konvergens. Resultatene fra de to numeriske metodene sammenlignes og tidsbruk diskuteres. Vi gir også en analyse av metodene med koersivitets-, kontinuitets-, stabiliets- og konvergensestimater.

For Black-Scholes ligning ser vi på både den orginale ligningen og en log-transformert ligning. De numeriske resultatene sammenlignes også med løsningen av ligningen funnet ved å bruke en endelig elementmetode.

Contents

Chapter 1

Introduction

1.1 Background

In this thesis, we will be studying two option pricing models which can be used to price vanilla European put options. European options are options which can be exercised only at maturity. Vanilla options are the simplest types of options, and they do not include any complex financial structures. We will asume that the underlying asset of the options are stocks. With this assumption, a put option is a contract giving the owner the right to sell the stock at a specific date (the maturity) for a specified amount of money (the strike price K). The spot price is the current price of the stock, denoted by S. A European put option will only be exercised if $S \leq K$.

The financial market is assumed to be arbitrage free, meaning that there should not be any possibilities of making risk-free profits. This should hold for options as well, therefore the option must be priced so that this assumption is fulfilled. We also have these two bounds on the price of the put option [15, p 193]:

- The price of an option can not be negative.
- A put option can not be worth more than the strike price K.

The most popular option pricing model is the Black-Scholes model. The model was published by Fischer Black and Myron Scholes in [3]. It models a fincancial market where currencies, bonds and derivatives can be traded. The market is assumed to contain both risky and risk-free assets. The Black-Scholes model is based on the following assumptions:

- The stock does not pay dividens.
- The riskless rate of return is constant.
- The log returns of the stock price follows a Wiener process with positive drift and constant volatility.
- There are no transaction costs when selling or buying stocks.
- One can borrow any amount of money at the risk-free rate.
- One can buy or sell any amount of the stock.

Some of these assumptions does not hold in practice. One issue is that there are transaction costs, but the most important issue is that the volatility is not constant. It can be constant over short periods of time, but in the long run the volatility depends on both the strike price and the expiry date. One way to avoid this problem is to assume

that the volatility is a stochastic function, which is the foundation for the stochastic volatility models. Heston's stochastic volatility model is the most commonly used of these models.

The option pricing models are most commonly solved by using finite differences. In the spesialization project, the Heston stochastic volatility model was solved by using the finite element method. In this thesis, we want to explore the possibility of solving the problems more efficiently by using spectral methods.

1.2 Overview of the thesis

Chapter 2 is devoted to the option pricing models we will use. Section 2.1 concerns Black-Scholes model, where we derive the model, state the analytical solution, discuss boundary conditions and derive the weak formulation. The same topics are covered for Heston's stochastic volatility model in section 2.2.

In chapter 3, we give an introduction to the spectral methods that we will use in this thesis, the spectral Galerkin method and a polynomial collocation method. We also test the methods on the heat equation. The spectral method is new to the author, so this is to learn how to use the method on an easy problem.

Chapter 4 is devoted to solving Black-Scholes model using both the spectral Galerkin method and polynomial collocation. For both methods, we derive the spectral approximation, give numerical results and analysis of the method. Numerical results are given both in terms of S and x, and both convergence and running time is considered. We compare the results to a solution obtained by using finite element methods.

In chapter 5, we try to solve Heston's stochastic volatility model using the spectral Galerkin method and a polynomial collocation method. As in chapter 4, we give the spectral approximation, numerical results and analysis of the method.

Finally, we discuss some problems with the method and draw conclusions in chapter 6, before we give some thoughts about future work on the topic.

Chapter 2

Option pricing models

2.1 Black-Scholes model

2.1.1 Derivation of the model

The derivation of Black-Scholes model is easily found in literature, for example in [15, chapter 8]. The model holds when the market is arbitrage free and complete. Arbitrage free means that there does not exist any risk-free possibilities of earning money, and complete means that all objects in the market can be reconstructed by a replicating portfolio. The model also assumes that the market consists of both risky assets and riskless assets. Riskless assets are typically bank acounts, called bonds.

There are many different approaches to derive the model, we have chosen the PDE approach as in [15, chapter 8.A]. We start by assuming that the stock price S follows a geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S dW_t, \qquad S_0 > 0 \tag{2.1}$$

where μ is a constant which represents the positive drift, σ is the constant volatility and W is a Wiener process. The subscript t means that the equation is evulated at time t. We define V as the value of the option, and assume that the value is a function of S_t and t. If we drop the subscript t and use Itô's Lemma (see $A.1$) on $V = V(S,t)$, we get

$$dV = \left(\mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \frac{\partial V}{\partial S} dW. \tag{2.2}$$

We want to eliminate the Wiener process from the equation. To obtain this, we create a portfolio Π consisting of a short position in V and a fixed amount Δ of S. A short position means that the portfolio will earn money if the market or stock price goes down. At time t, the value of the portfolio will be

$$\Pi = -V + \Delta S. \tag{2.3}$$

At time t_0, the portfolio has value Π_0. Because the market is assumed to be arbitrage free, the portfolio has differential

$$d\Pi = -dV + \Delta dS. \tag{2.4}$$

If we substitute equations (2.1) and (2.2) into (2.4), we get

$$d\Pi = -\left(\mu S\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right)dt - \sigma S\frac{\partial V}{\partial S}dW + \Delta\mu S dt + \Delta\sigma S dW.$$

We observe that if we choose $\Delta = \frac{\partial V}{\partial S}$, the Wiener process cancels out, which is what we wanted to obtain by creating the portfolio. In addition, terms with μ disappears as well. Note that the Δ is still fixed at time t. We are then left with

$$d\Pi = -\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right)dt. \tag{2.5}$$

We observe that the change $d\Pi$ of the portfolio is deterministic. If we use the fact that the market is arbitrage free and the assumptions of the market given in section 1.1, we know that the same amount of money as Π should grow with the risk-free interest rate. At time t, the value should be $\Pi = \Pi_0 e^{-r(t_0-t)}$. From this, we obtain the relation

$$d\Pi = r\Pi dt$$

which is valid with our assumptions on the market. We have two expressions for $d\Pi$, and they must be equal because the market is arbitrage free. If we use this and substitute for the definition of the portfolio in equation (2.3) and equation (2.5), we get

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

By rearranging and dropping dt, we get the Black-Scholes partial differential equation, or simply Black-Scholes equation:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \tag{2.6}$$

which is valid when

$$S > 0, \qquad 0 \leq t \leq T.$$

The initial condition of the problem is the pay-off function at time $t = T$. For a put, this is

$$V(S,T) = \max(K - S, 0). \tag{2.7}$$

Note that this means that the equation is backwards in time, starting at $t = T$.

2.1.2 Black-Scholes formula

There exist expressions for the analytical solution of Black-Scholes equation for both European put and call options. The formula can be found in for example [15, chapter 8.2], and we have the following formula for the option price of a European put:

$$V^{BS} = Ke^{-rT}N(-d_2) - SN(-d_1), \tag{2.8}$$

where $N(\cdot)$ is the distribution function of the normal distribution,

$$d_1 = \frac{\ln(S/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

and

$$d_2 = \frac{\ln(S/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}.$$

The formula will be used to test the numerical methods in chapter 4.

2.1.3 Boundary conditions

We will solve equation (2.6) for a European put option. For this, we need boundary conditions in addition to the inital condition (2.7). When $S = 0$, we have an exact value for the option price, which is the strike price K discounted for the time to maturity:

$$V(0,t) = Ke^{-r(T-t)} \qquad 0 \leq t \leq T. \tag{2.9}$$

This boundary condition is derived by looking at $K-S$, which is the payoff at maturity if the option was free. When $S << K$, we have $K - S \approx K$ and it is very likely that the option will be exercised with pay-off K. Therefore, this should be the price of the option. However, when we move backwards in time from maturity, we have to discount the value to the present value. This gives the boundary condition given, where $T - t$ is the time to maturity.

When we have $S >> K$, the probability of S becoming less than K is very small. This means that the possibility of having $K - S > 0$ is also small. Therefore, the option price goes to zero as the spot price goes to ∞, and we have

$$V(\infty,t) = 0 \qquad 0 \leq t \leq T. \tag{2.10}$$

2.1.4 Log transformation

A log transformation of S can be performed in order to get rid of the S-dependency in (2.6). This will give us an equation without variable coefficients and is often performed in order to obtain an easier equation to work with. The log transformation is also the first step of transforming Black-Scholes equation into the heat equation. We start by introducing the variable $x = \ln(S)$. Then we have

$$\begin{aligned}
\frac{\partial V}{\partial S} &= \frac{\partial V}{\partial x}\frac{\partial x}{\partial S} = \frac{1}{S}\frac{\partial V}{\partial x} \\
\frac{\partial^2 V}{\partial S^2} &= \frac{\partial}{\partial S}\left(\frac{\partial V}{\partial S}\right) = \frac{\partial V}{\partial x}\frac{\partial}{\partial S}\left(\frac{1}{S}\right) + \frac{1}{S}\frac{\partial}{\partial S}\left(\frac{\partial V}{\partial x}\right) = -\frac{1}{S^2}\frac{\partial V}{\partial x} + \frac{1}{S^2}\frac{\partial^2 V}{\partial x^2}.
\end{aligned}$$

If we insert this into (2.6) and denote $v(x,t) = V(e^x,t)$, we get

$$v_t + \frac{1}{2}\sigma^2 v_{xx} + (r - \frac{1}{2}\sigma^2)v_x - rv = 0. \tag{2.11}$$

The spatial domain is now transformed from $S \geq 0$ to $-\infty < x < \infty$. We will denote the transformed domain $\tilde{\Omega}$. The initial condition (2.7) and boundary conditions (2.9) and (2.10) are transformed to

$$v(x,T) = \max\{K - e^x, 0\}, \qquad x \in (-\infty, \infty),$$
$$v(-\infty, t) = Ke^{-r(T-t)}, \qquad 0 \le t \le T$$

and

$$v(\infty, t) = 0, \qquad 0 \le t \le T.$$

2.1.5 Transformation to the heat equation

The Black-Scholes equation (2.6) is a parabolic PDE which can be transformed to the heat equation. We will use this equation to learn how to solve problems using spectral methods in chapter 3.

As for the derivation of the Black-Scholes model, the transformation is easily found in literature. Here, we use the notation and transformations in [20, chapter 4]. We start by performing a log-transformation of S and a transformation of the time variable t:

$$S = Ke^x, \qquad t = T - \frac{\tau}{\sigma^2/2}.$$

If we then introduce the variable $\tilde{w} = \frac{1}{K}V$, we have

$$\tilde{w}(x,\tau) = \frac{1}{K}V\left(Ke^x, T - \frac{\tau}{\sigma^2/2}\right).$$

The domain of $\tilde{w}$ is $-\infty < x < \infty$ and $0 \le \tau \le \frac{\sigma^2}{2}T$. The transformation of t has made the equation forward in time.

If we insert the new variables into equation (2.6) and apply the chain rule, we get

$$\tilde{w}_\tau = \tilde{w}_{xx} + \left(\frac{r}{\sigma^2/2} - 1\right)\tilde{w}_x - \frac{r}{\sigma^2/2}\tilde{w}.$$

To shorten the expression, we define

$$\kappa = \frac{r}{\sigma^2/2}$$

and write

$$\tilde{w}_\tau = \tilde{w}_{xx} + (\kappa - 1)\,\tilde{w}_x - \kappa\tilde{w}. \tag{2.12}$$

To obtain the heat equation, we need to transform $\tilde{w}$ to a new variable w in a way such that the terms with w_x and w disappear. This can be obtained by introducing

$$\gamma = \frac{1}{2}(\kappa - 1) \qquad \text{and} \qquad \beta = \frac{1}{2}(\kappa + 1) = \gamma + 1$$

and define

$$\tilde{w}(x,\tau) = e^{-\gamma x - \beta^2\tau}w(x,\tau).$$

Now, we can calculate the partial derivatives in terms of w

$$\tilde{w}_\tau = e^{-\gamma x - \beta^2\tau}(-\beta^2 w + w_\tau),$$
$$\tilde{w}_x = e^{-\gamma x - \beta^2\tau}(-\gamma w + w_x),$$
$$\tilde{w}_{xx} = e^{-\gamma x - \beta^2\tau}(\gamma^2 w - 2\gamma w_x + w_{xx}).$$

Inserting these expressions into equation (2.12), we get

$$e^{-\gamma x-\beta^2\tau}(-\beta^2 w + w_\tau) = e^{-\gamma x-\beta^2\tau}(2\gamma(-\gamma w + w_x) + \gamma^2 w - 2\gamma w_x + w_{xx} - \kappa w).$$

The exponential term cancels out and we have

$$w_\tau = w_{xx} + (-2\gamma + 2\gamma)w_x + (\beta^2 + \gamma^2 - 2\gamma^2 - \kappa)w.$$

By using $\beta^2 = (\gamma + 1)^2 = \gamma^2 + \kappa$, both the w_x and w terms disappear, and we are left with

$$w_\tau = w_{xx},$$

which is the dimensionless heat equation.

After the transformation, the boundary conditions are

$$w(-\infty, \tau) = e^{-\kappa\tau} \cdot e^{\gamma x+\beta^2\tau} \quad \text{and} \quad w(\infty, \tau) = 0.$$

The initial condition in equation (2.7) must also be transformed:

$$\begin{aligned} w(x, 0) &= \frac{1}{K} e^{\gamma x} V(Ke^x, T) \\ &= \frac{1}{K} e^{\gamma x} \max(K - Ke^x, 0) \\ &= \max(e^{\gamma x} - e^{(\gamma+1)x}, 0). \end{aligned}$$

Together with the transformed boundary conditions, we have a well-posed problem.

As for equation (2.6), we have a closed-form solution to the transformed problem, see for example [20, chapter 4]. For a put, this is given by

$$u(x, \tau) = e^{\gamma x+\gamma^2\tau}\Phi\left(-\frac{x + 2\tau\gamma}{\sqrt{2\pi}}\right) - e^{\beta x+\beta^2\tau}\Phi\left(-\frac{x + 2\tau\beta}{\sqrt{2\pi}}\right),$$

where Φ is a distribution function defined by

$$\Phi(\zeta) = \int_{-\infty}^{\zeta} \phi(\eta)d\eta = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\zeta} e^{-\eta^2/2}d\eta.$$

The closed-form solution to Black-Scholes equation given in equation (2.8) can be derived from the closed-form solution defined above by transforming the variables back to the original variables.

2.1.6 Weak formulation of Black-Scholes equation

When we use the spectral Galerkin method, the weak form of the equation is needed. Therefore, we will derive the weak formulation for both the original equation (2.6) and the log transformed equation (2.11).

We start with equation (2.6), and multiply with a test function $\varphi \in X$ and integrate over the domain Ω:

$$\int_\Omega V_t\varphi + \int_\Omega \frac{1}{2}\sigma^2 S^2 V_{SS}\varphi + \int_\Omega rSV_S\varphi - \int_\Omega rV\varphi = 0.$$

By using integration by parts on the second integral in the above equation, we obtain

$$\int_\Omega \frac{1}{2}\sigma^2 S^2 V_{SS}\varphi = -\int_\Omega \frac{1}{2}\sigma^2 S^2 V_S \varphi_S - \int_\Omega \sigma^2 S V_S \varphi.$$

We have chosen $\varphi(0) = \varphi(\infty) = 0$, which is a reasonable choice when we have Dirichlet boundary conditions. We now have

$$\int_\Omega V_t\varphi - \int_\Omega \frac{1}{2}\sigma^2 S^2 V_S\varphi_S - \int_\Omega (\sigma^2 - r)SV_S\varphi - \int_\Omega rV\varphi = 0. \tag{2.13}$$

We will use the following notation

$$a(V,\varphi) = -\int_\Omega \frac{1}{2}\sigma^2 S^2 V_S\varphi_S - \int_\Omega (\sigma^2 - r)SV_S\varphi - \int_\Omega rV\varphi \tag{2.14}$$

and

$$\frac{d}{dt}(V,\varphi) = \int_\Omega V_t\varphi. \tag{2.15}$$

The weak formulation of the problem can now be stated as

Find $V \in W$ such that $\frac{d}{dt}(V,\varphi) + a(V,\varphi) = 0$ for all $\varphi \in X$.

The spaces X and W are defined as

$$X = \{\varphi \in L^2(\Omega) : S\frac{\partial\varphi}{\partial S} \in L^2(\Omega), \varphi(0) = \varphi(S) = 0\}$$

and

$$W = \{V \in L^2(\Omega) : \frac{\partial V}{\partial t}, S\frac{\partial\varphi}{\partial S} \in L^2(\Omega)\}.$$

Next, we will give a similar weak formulation for the transformed problem in equation (2.11). As earlier, we start by multiplying the equation by a test function φ and itegrate over the domain:

$$\int_{\tilde\Omega} v_t\varphi + \int_{\tilde\Omega} \frac{1}{2}\sigma^2 v_{xx}\varphi + \int_{\tilde\Omega} (r - \frac{1}{2}\sigma^2)v_x\varphi - \int_{\tilde\Omega} rv\varphi = 0.$$

Agian, we apply integration by parts on the second integral and obtain

$$\int_{\tilde\Omega} v_t\varphi - \int_{\tilde\Omega} \frac{1}{2}\sigma^2 v_x\varphi_x + \int_{\tilde\Omega} (r - \frac{1}{2}\sigma^2)v_x\varphi - \int_{\tilde\Omega} rv\varphi = 0,$$

where we have used that $\varphi(\pm\infty) = 0$.
The weak formulation can again be stated as

Find $v \in \tilde W$ such that $\frac{d}{dt}(v,\varphi) + \tilde a(v,\varphi) = 0$ for all $\varphi \in \tilde X$.

Now,

$$\tilde a(v,\varphi) = -\int_{\tilde\Omega} \frac{1}{2}\sigma^2 v_x\varphi_x + \int_{\tilde\Omega} (r - \frac{1}{2}\sigma^2)v_x\varphi - \int_{\tilde\Omega} rv\varphi, \tag{2.16}$$

$$\frac{d}{dt}(v,\varphi) = \int_{\tilde\Omega} v_t\varphi, \tag{2.17}$$

$$\tilde X = \{\varphi \in L^2(\tilde\Omega) : \frac{\partial\varphi}{\partial x} \in L^2(\tilde\Omega), \varphi(-\infty) = \varphi(\infty) = 0\} = H_0^1(\tilde\Omega)$$

and

$$\tilde{W} = \{v \in L^2(\tilde{\Omega}) : \frac{\partial v}{\partial t}, \frac{\partial v}{\partial x} \in L^2(\tilde{\Omega})\}.$$

2.2 Heston's stochastic volatility model

2.2.1 Derivation of the model

Heston's stochastic volatility model was introduced by Steven Heston in 1993 in [9]. The main difference from Black-Scholes equation is that the volatility v is assumed to be a stochastic process. Hence, the option price U also depends on v: $U = U(v, S, t)$. We will derive the model here. The derivation can be found in [13].

Heston starts by assuming that the spot price S can be described by the stochastic differential equation

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_{1,t},$$

where W_i represents a Wiener process and v is the volatility. The subscript t denotes evaluation at time t, but as for the derivation of Black-Scholes, we drop the subscripts. The volatility is assumed to follow an Ornstein-Uhlenbeck process. This is a stochastic process that drifts toward its mean with a constant rate. More details about Ornstein-Uhlenbeck processes can be found in [11]. Heston uses the following Ornstein-Uhlenbeck process

$$d\sqrt{v} = -\beta\sqrt{v}dt + \delta dW_2.$$

The two Wiener processes W_1 and W_2 have correlation ρ. If we apply Ito's lemma $(A.1)$ to the stochastic process with $F(u) = u^2$ and $u = \sqrt{v(t)}$, we obtain

$$dv = (\delta^2 - 2\beta v)dt + 2\delta\sqrt{v}dW_2.$$

It is possible to write this as the Cox, Ingersoll and Ross square-root process in [5]

$$dv = \kappa(\theta - v)dt + \sigma\sqrt{v}dW_2.$$

Now, κ is the mean reversion rate, θ is the long term variance and σ is the volatility of volatility.

For the rest of the derivation, we follow [13]. We create a replicating portfolio Π constisting of one option U, Δ of S and Φ of the option Y. The portfolio is assumed to be self-financing, and we assume that both U and Y are functions of v, S and t. The value of the portfolio is the sum of the contributions,

$$\Pi = U + \Delta S + \Phi Y.$$

The portfolio has differential

$$d\Pi = dU + \Delta dS + \Phi dY + d\Delta S + d\Phi Y,$$

but because the portfolio is self-financing, we have $-d\Delta S = d\Phi Y$, so the differential is

$$d\Pi = dU + \Delta dS + \Phi dY. \tag{2.18}$$

Next, we use Ito's lemma (A.1) on the expression for dU and dY, giving

$$dU = \frac{\partial U}{\partial t}dt + \frac{\partial U}{\partial v}dv + \frac{\partial U}{\partial S}dS + \frac{1}{2}v\left(S^2\frac{\partial^2 U}{\partial S^2} + \sigma^2\frac{\partial^2 U}{\partial v^2}\right)dt + \rho\sigma vS\frac{\partial^2 U}{\partial v \partial S}dt$$

for dU and similar for dY. Inserting this into (2.18) gives

$$\begin{aligned} d\Pi = &\left(\frac{\partial U}{\partial t} + \frac{1}{2}v\left(S^2\frac{\partial^2 U}{\partial S^2} + \sigma^2\frac{\partial^2 U}{\partial v^2}\right)dt + \rho\sigma vS\frac{\partial^2 U}{\partial v \partial S}\right)dt \\ &+ \frac{\partial U}{\partial v}dv + \frac{\partial U}{\partial S}dS + \Delta dS + \Phi\frac{\partial Y}{\partial v}dv + \Phi\frac{\partial Y}{\partial S}dS \\ &+ \Phi\left(\frac{\partial Y}{\partial t} + \frac{1}{2}v\left(S^2\frac{\partial^2 Y}{\partial S^2} + \sigma^2\frac{\partial^2 Y}{\partial v^2}\right)dt + \rho\sigma vS\frac{\partial^2 Y}{\partial v \partial S}\right)dt. \end{aligned} \tag{2.19}$$

The terms constisting dS and dv must be eliminated from the expression in order to create a risk-free portfolio. This is because the expressions for dS and dv involve Brownian motions, and thus involves randomness. Eliminating the terms gives

$$\frac{\partial U}{\partial S} + \Phi\frac{\partial Y}{\partial S} + \Delta = 0 \quad \text{and} \quad \frac{\partial U}{\partial v} + \Phi\frac{\partial Y}{\partial v} = 0,$$

which yields

$$\Delta = -\frac{\partial U}{\partial S} - \Phi\frac{\partial Y}{\partial S} \quad \text{and} \quad \Phi = -\frac{\frac{\partial U}{\partial v}}{\frac{\partial Y}{\partial v}}. \tag{2.20}$$

Now, we have a risk-free portfolio because it is hedged against movements in the stock and the volatility. This means that the portfolio again must earn the risk-free interest rate because of the assumption that the market is arbitrage free. Therefore, we again obtain the relation

$$d\Pi = r\Pi dt, \tag{2.21}$$

where r is the risk-free interest rate.

Next, we note that we can write equation (2.19) on the form $d\Pi = C(U) + \Phi D(Y)$, where C depends only on U and D only on Y. Combining this with equation (2.21) and the expressions in (2.20) gives

$$C(U) - \frac{\frac{\partial U}{\partial v}}{\frac{\partial Y}{\partial v}}D(Y) = r\left(U + (\frac{\frac{\partial U}{\partial v}}{\frac{\partial Y}{\partial v}}\frac{\partial Y}{\partial S} - \frac{\partial U}{\partial S})S - \frac{\frac{\partial U}{\partial v}}{\frac{\partial Y}{\partial v}}Y\right).$$

By rearranging this, we get

$$\frac{C(U) - rU + rS\frac{\partial U}{\partial S}}{\frac{\partial U}{\partial v}} = \frac{D(Y) - rY + rS\frac{\partial Y}{\partial S}}{\frac{\partial Y}{\partial v}}. \tag{2.22}$$

We observe that the left side only depends on U and the right hand side only depends on Y, thus we can write each side as a function f dependent of v, S and t. Heston chose this function to be

$$f(v, S, t) = -\kappa(\theta - v) + \tilde{\lambda}(v, S, t),$$

where $\tilde{\lambda}(v, S, t)$ represents the price of volatility risk. This is by Heston chosen to be λv, for some constant λ. Combining this with the left side of equation (2.22) and the expression for $C(U)$ from equation (2.19), we get

$$U_t + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \rho\sigma Sv\frac{\partial^2 U}{\partial v \partial S} + \frac{1}{2}v\sigma^2\frac{\partial^2 U}{\partial v^2} - rU + rS\frac{\partial U}{\partial S} = (-\kappa(\theta - v) + \lambda v)\frac{\partial U}{\partial v}.$$

This gives the following PDE for Heston's stochastic volatility model:

$$U_t + \frac{1}{2}\sigma^2 v U_{vv} + \rho\sigma v S U_{vS} + \frac{1}{2}vS^2 U_{SS} + (\kappa(\theta - v) - \lambda v)\, U_v + rSU_S - rU = 0, \quad (2.23)$$

which is valid for $S \geq 0$, $v \geq 0$ and $0 \leq t \leq T$. The problem has an initial condition at $t = T$, for a European put option this is

$$U(v, S, T) = \max(K - S, 0).$$

This is the pay-off function for the option at maturity. As Black-Scholes equation, this equation is backwards in time.

By introducing the notation

$$E = \frac{1}{2}v\begin{bmatrix}\sigma^2 & \rho\sigma S \\ \rho\sigma S & S^2\end{bmatrix},$$

$$f = \begin{bmatrix}\kappa(\theta - v) - \lambda v - \frac{1}{2}\rho\sigma v - \frac{1}{2}\sigma^2 \\ rS - \frac{1}{2}\rho\sigma S - vS\end{bmatrix} \quad \text{and } \nabla = \begin{bmatrix}\frac{\partial}{\partial v} \\ \frac{\partial}{\partial S}\end{bmatrix},$$

we can write equation (2.23) as

$$U_t + \nabla \cdot E\nabla U + f \cdot \nabla U - rU = 0. \quad (2.24)$$

2.2.2 Closed-form solution

There exists a closed-form solution for Heston's stochastic volatility model as well. For a European put, this is given by [9]

$$U(v, S, t) = Ke^{-rt}\Pi_1 - Se^{-qt}\Pi_2, \quad (2.25)$$

where

$$\Pi_1 = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{e^{-iw\ln K}\psi(\omega)}{i\omega}\right)d\omega,$$

$$\Pi_1 = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{e^{-iw\ln K}\psi(\omega - i)}{i\omega\psi(-i)}\right)d\omega$$

and

$$\psi(\omega) = E\left[e^{i\omega\ln S(T)}\right].$$

Here, q is the dividend yield.

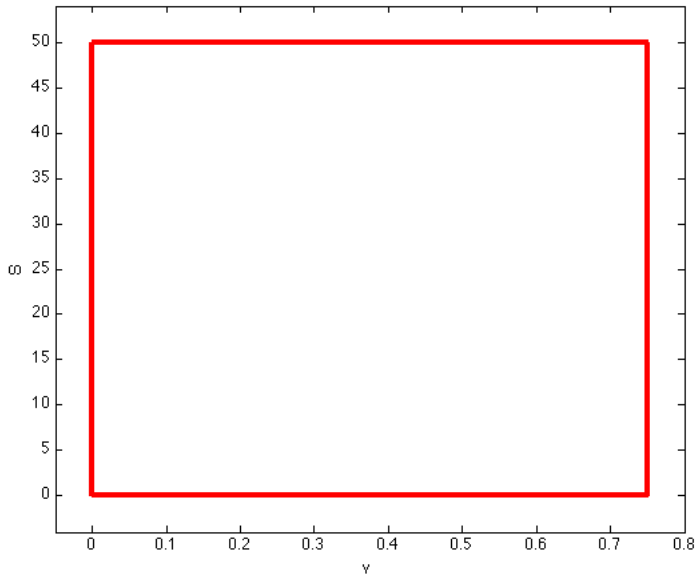


Figure 2.1: The solution domain for Heston's stochastic volatility model.

2.2.3 Boundary conditions

In order to solve the problem, we must choose a reasonable domain. The domain we will use is shown in figure 2.1. There are different boundary conditions on each of the boundaries, and also different possible choices for boundary conditions at each boundary. We start with the boundary conditions from [16], which are

- Γ_1 $S = 0$: $\quad U(v, 0, t) = Ke^{-r(T-t)}$,
- Γ_2 $v = v_{\max}$: $\quad U(v_{\max}, S, t) = Ke^{-r(T-t)}$,
- Γ_3 $S = S_{\max}$: $\quad U_S(v, S_{\max}, t) = 0$ and
- Γ_4 $v = v_{\min}$: $\quad$ Equation (2.23) holds when $v = v_{\min}$.

On Γ_1, we have a Dirichlet boundary condition. This condition gives the present value of the strike price K at maturity. The boundary condition is derived in the same manner as (2.9).

The condition on Γ_2 is also a Dirichlet condition. This condition is valid when $v_{\max}$ is large, but in our case $v_{\max}$ is not large enough to make this boundary conition valid. If we study the solution close to this boundary, we see that the boundary is an outflow boundary for reasonable choices of constants, as argued for in [19]. Therefore, we will replace the boundary condition with a natural boundary condition, and redefine the boundary condition to

$$\Gamma_2 : \quad E\nabla U(v_{\max}, S, t) \cdot \overrightarrow{n} = 0.$$

The third boundary condition is a Neumann condition saying that the solution does not change in the S-direction along Γ_3. For implementations, it is easier to use a natural boundary condition instead. Therefore, we replace it with

$$\Gamma_3 : \quad E\nabla U(v, S_{\max}, t) \cdot \overrightarrow{n} = 0,$$

which essentially says that there is no change in either direction.

On the last boundary, there is no boundary condition. By looking at the characteristics of the equation, we observe that they all point out of the domain. This tells us that we don't need any extra information at the boundary. This is valid when $v_{min} = 0$, and we don't have diffusion at this boundary. However, we will see later that the spectral Galerkin method is not continuous when $v_{min} = 0$. Therefore, we must adjust the lower bound on v to some value $\epsilon > 0$. When we do this, diffusion will be present at the boundary. All though small, it still makes the argument of characteristics invalid. Thus, we need a boundary condition here as well. We choose

$$\Gamma_4 : \quad E\nabla U(v_{\min}, S, t) \cdot \overrightarrow{n} = 0.$$

This condition will be valid when $v_{\min} = 0$ as well, because $E(v = 0) = 0$.

2.2.4 Transformation

As for Black-Scholes equation, we can introduce a log transformation of S and transform equation (2.23). Here, we write $x = \ln(S/K)$ and $u(v, x, t) = U(v, \ln(S/K), t)$. The transformed equation becomes

$$u_t + \frac{1}{2}\sigma^2 v u_{vv} + \rho\sigma v u_{vx} + \frac{1}{2} v u_{xx} + (\kappa(\theta - v) - \lambda v)u_v + (r - \frac{1}{2}v)u_x - ru = 0,$$

with initial condition

$$u(v, x, T) = \max(K - Ke^x, 0).$$

We will see later in the thesis that the log transformation introduces an error to the solution of Black-Scholes equation. Therefore, we will only be working with the original equation for Heston's stochastic volatility model, equation (2.23).

2.2.5 Weak formulation

We start by multiplying equation (2.24) by a test function $\varphi \in X$ and integrate over the domain:

$$\int_\Omega U_t\varphi + \int_\Omega (\nabla \cdot E\nabla U)\,\varphi + \int_\Omega f \cdot \nabla U\varphi + \int_\Omega rU\varphi = 0.$$

We use integration by parts on the second integral to obtain

$$\int_\Omega (\nabla \cdot E\nabla U)\,\varphi = -\int_\Omega E\nabla U \cdot \nabla\varphi + \int_{\partial\Omega} E\nabla U \cdot \overrightarrow{n}\varphi,$$

so we have

$$\int_\Omega U_t\varphi - \int_\Omega E\nabla U \cdot \nabla\varphi + \int_{\partial\Omega} E\nabla U \cdot \overrightarrow{n}\varphi + \int_\Omega f \cdot \nabla U\varphi - \int_\Omega rU\varphi = 0.$$

The boundary integral will disapper on Γ_2, Γ_3 and Γ_4 due to the boundary conditions given. On Γ_1, we have a Dirichlet boundary condition. If we define φ such that they disappear at the Dirichlet boundary, the boundary integral becomes zero at this boundary as well. We are left with

$$\int_\Omega U_t\varphi - \int_\Omega E\nabla U \cdot \nabla\varphi + \int_\Omega f \cdot \nabla U\varphi - \int_\Omega rU\varphi = 0.$$

The weak formulation of the problem can now be stated as

Find $U \in W$ such that $\frac{d}{dt}(U,\varphi) + \hat{a}(U,\varphi) = 0$ for all $\varphi \in X$,

where

$$\hat{a}(U,\varphi) = -\int_\Omega E\nabla U \cdot \nabla\varphi + \int_\Omega f \cdot \nabla U\varphi - \int_\Omega rU\varphi. \tag{2.26}$$

The spaces are defined as

$$X = \{\varphi \in W_0 : \varphi = 0 \text{ on } \Gamma_1\},$$

$$W_0 = \{\varphi : \|\varphi\|_* \leq \infty\}$$

and

$$W = \{\varphi \in W_0 : \varphi \text{ satisfies the boundary condition on } \Gamma_1\}.$$

The norm

$$\|\varphi\|_*^2 = \int_\Omega E\nabla\varphi \cdot \nabla\varphi + \int_\Omega (r + \frac{1}{2}\nabla \cdot f)\varphi^2 \tag{2.27}$$

comes from

$$(U,\varphi)_* = \int_\Omega E\nabla U \cdot \nabla\varphi + \int_\Omega (r + \frac{1}{2}\nabla \cdot f)U\varphi.$$

Note that we can always create an inner product from this by adding $\lambda_c(U,\varphi)_{L^2(\Omega)}$ to $(U,\varphi)_*$. In order for this to be an inner product, we must have

$$\lambda_c + r + \frac{1}{2}\nabla \cdot f \geq 0,$$

which is easily fulfilled by choosing a large enough λ_c. In the analysis of the spectral Galerkin method, we will use this inner product with $\lambda_c = 1/h$.

Chapter 3

Introduction to spectral methods

In this chapter, we will give an introduction to the spectral methods we will use in this thesis. We will also test the methods on the heat equation. This is because the spectral method is new to the author, and an easy problem is good for practice in how to use the method. The lecture notes for the course MA8502 Numerical solution of partial differential equations have been of good use in learning about spectral methods. The lecture notes are written by Einar Rønquist. In addition, the books [4], [10] and [14] have been useful.

Spectral methods are a class of methods for solving partial differential equations. The class of methods are based on basis functions with global support. Spectral methods are not very flexible in terms of domains, but they have very good convergence, especially when we have smooth solutions.

There are many different types of spectral methods. In this thesis, we will look at a spectral Galerkin method and a polynomial collocation method.

3.1 The spectral Galerkin method

The spectral Galerkin method is based on the weak formulation of a problem. We are seeking solutions $w \in W$ of some problem on the form $\frac{d}{dt}(w, v) + a(w, v) = (f, v)$, $\forall v \in W$, and the solutions should be represented by some expansion of polynomials. It is therefore convinient to use basis functions which are polynomials as well. There are many possible choices for polynomials, common choices include Fourier series, Lagrange polynomials and Chebyshev polynomials. For this thesis we have chosen to use the Lagrange polynomials $\ell_j(x)$ as basis functions. They are given by

$$\ell_j(x) = \prod_{i=0, i\neq j}^{N} \frac{x - x_i}{x_j - x_i},$$

for some set of points $\{x_j\}$. The Lagrangian polynomials have the property $\ell_j(x_i) = \delta_{ij}$, which is illustrated in figure 3.1. This property is very important for the spectral Galerkin method, and will be used a lot.

Using this basis, we can represent the solution space W by

$$W = \text{span}\{\ell_0, \ell_1, \ldots, \ell_N\}$$

The integrals in the bilinear form will be evaluated by quadrature. As for the basis functions, there are many possible choices for quadrature. Different types of Gauss

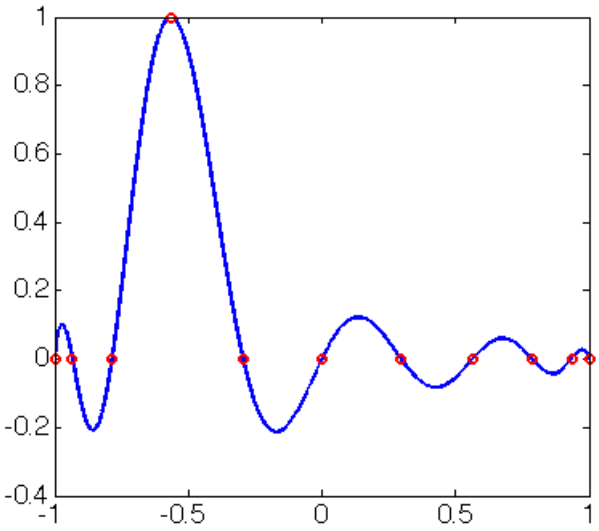


Figure 3.1: The Lagrangian polynomial of order 10 with $i = 3$.

quadrature is common. We use the Gauss-Lobatto-Legendre quadrature (or simply GLL quadrature), with the corresponding nodes ξ_j and weights ρ_j. The nodes are the $N+1$ zeros of $(1-\xi^2)L'_N(\xi)$, where $L_N(\xi)$ are the Legendre polynomials. The weights are

$$\rho_j = \frac{1}{N(N+1)} \frac{1}{L_N(\xi_j)^2}, \quad 0 \leq j \leq N.$$

With this notation, we can approximate an integral on $[-1, 1]$ by

$$\int_{-1}^{1} f(x)dx \approx \sum_{\alpha=0}^{N} \rho_\alpha f(\xi_\alpha).$$

This quadrature is exact for $f(x) \in \mathbb{P}_{2N-1}([-1, 1])$. An example of the quadrature nodes is given in figure 3.2. We observe that the nodes are not equidistant on the interval, but they are symmetric and have higher density close to the boundaries. The placement of the nodes is unique for each N, meaning that the nodes will be placed differently for each value of N.

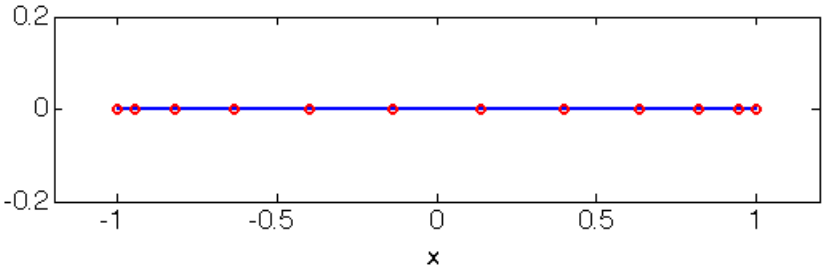


Figure 3.2: Gauss Lobatto Legendre nodes for $N = 10$, giving 11 nodes

3.2 Spectral Galerkin applied to the heat equation

In order to learn the basics of the spectral Galerkin method, we will apply it to

$$\begin{aligned} &w_t = w_{xx}, \qquad x \in [-1,1] \\ &w(-1) = w(1) = 0 \qquad\qquad\qquad w(x,0) = \sin(\pi x). \end{aligned} \tag{3.1}$$

The problem has the exact solution

$$w(x,t) = e^{-\pi^2 t}\sin(\pi x). \tag{3.2}$$

As mentioned, we need the weak formulation of the problem. We multiply (3.1) by a test function $\varphi \in W$ and integrate over $[-1,1]$:

$$\int_{-1}^{1} w_t\varphi - \int_{-1}^{1} w_{xx}\varphi = 0.$$

By applying integration by parts on the second integral and use $w(\pm 1) = 0$, we obtain

$$\int_{-1}^{1} w_t\varphi + \int_{\Omega} w_x\varphi_x = \frac{d}{dt}(w,\varphi) + a(w,\varphi) = 0,$$

where $a(w,\varphi) = \int_\Omega w_x\varphi_x$. This gives the weak formulation

Find $w \in W$ such that $\frac{d}{dt}(w,\varphi) + a(w,\varphi) = 0$ for all $\varphi \in W$.

We define $W = \{w \in H^1(\Omega) : w(\pm 1) = 0\} = H_0^1(\Omega)$.

As basis functions, we choose the $N-1$ Lagrange polynomials $\ell_j(x)$. We only need $N-1$ functions because we can exclude the two endpoints where we have homogeneous Dirichlet conditions. The discrete space W_N can be represented by these basis functions, $W_N = \text{span}\{\ell_1, \ldots, \ell_{N-1}\}$. Now, we express a discrete solution $w_N \in W_N$ to our problem as

$$w_N(x) = \sum_{j=1}^{N-1} w_N(x_j)\ell_j(x).$$

We define $w_N(x_j) = w_j$. Because $\ell_j(x_i) = \delta_{ij}$, the coefficients w_j are equal to w_N at the nodes x_j, so the w_j's are the solution to our problem. We can now write the bilinear formulation as

$$a(\sum_{j=1}^{N-1} w_j\ell_j(x), \varphi).$$

By choosing $\varphi(x) = \ell_i(x)$ with $i = 1, \ldots, N-1$, we get

$$a(\sum_{j=1}^{N-1} w_j\ell_j(x), \ell_i(x)) = \sum_{j=1}^{N-1} a(\ell_j,\ell_i)w_j, \quad i = 1, \ldots, N-1.$$

We can write this as the matrix A where $A_{ij} = a(\ell_j,\ell_i)$. Next, we evaluate the integral in the bilinear form by GLL quadrature

$$\begin{aligned} a(\ell_j,\ell_i) &= \int_{-1}^{1} \ell_j'(x)\ell_i'(x)dx \\ &= \sum_{\alpha=0}^{N} \rho_\alpha \ell_i'(\xi_\alpha)\ell_j'(\xi_\alpha) = a_N(\ell_j,\ell_i) \\ &= \sum_{\alpha=0}^{N} \rho_\alpha D_{\alpha i} D_{\alpha j} = A_{ij}. \end{aligned}$$

Each of the basis functions have degree $N-1$, so $\ell_i\ell_j \in \mathbb{P}_{2N-2}$ and the GLL quadrature is exact.

Similarly, we get that the time derivative leads to a matrix M, where $M_{ij} = (\ell_i, \ell_j)$ and

$$(\ell_i, \ell_j) = \sum_{\alpha=0}^{N} \rho_\alpha \ell_i(\xi_\alpha)\ell_j(\xi_\alpha) = (\ell_i, \ell_j)_N = \rho_i \delta_{ij}.$$

Note that $a_N(\cdot,\cdot)$ and $(\cdot,\cdot)_N$ means that the integrals are evaluated by GLL quadrature.

Now, we can write this as the system $\mathbf{M}\dot{\mathbf{w}} + \mathbf{A}\mathbf{w} = 0$. The matrix formulation will be equal for all methods considered in this report, as long as we use the correct corresponding matrices $\mathbf{M}$ and $\mathbf{A}$, so we will give a general solution method for this system. For a general right hand side $\mathbf{f}$, we have the system

$$\mathbf{M}\dot{\mathbf{w}} + \mathbf{A}\mathbf{w} = \mathbf{f},$$

which can be solved by using an appropriate time discretization. We approximate $\dot{\mathbf{w}}$ by

$$\dot{\mathbf{w}} \approx \frac{\mathbf{w^{k+1}} - \mathbf{w^k}}{h},$$

where h is the time step. By using implicit Euler, we get

$$\mathbf{M}\frac{\mathbf{w^{k+1}} - \mathbf{w^k}}{h} + \mathbf{A}\mathbf{w^{k+1}} = \mathbf{f}^{k+1}.$$

This gives the following expression for $\mathbf{w^{k+1}}$

$$\mathbf{w^{k+1}} = (\mathbf{M} + h\mathbf{A})^{-1}(h\mathbf{f^{k+1}} + \mathbf{M}\mathbf{w^k}).$$

We can also use Crank-Nicolson as discretization scheme. This gives

$$\mathbf{M}\frac{\mathbf{w^{k+1}} - \mathbf{w^k}}{h} + \frac{1}{2}\mathbf{A}\left(\mathbf{w^{k+1}} + \mathbf{w^k}\right) = \frac{1}{2}\left(\mathbf{f}^{k+1} + \mathbf{f}^k\right).$$

The main difference between the two methods is that Crank-Nicolson is quadratic in time, while implicit Euler is linear in time. In addition, Crank-Nicolson requires more calculations per iteration.

Using Crank-Nicolson as discretization, we get the solution given in figure 3.3. We have used $N = 50$ and $h = 1 \cdot 10^{-3}$ in the calculations. To see how good the method is, we study the convergence of the numerical solution. To do this, we will use the GLL quadrature to approximate a norm of the difference between the exact solution and the numerical solution at the last time step. We call this norm $\|\cdot\|_{GLL}$, and define it as

$$\begin{aligned}\|w_{exact}(\cdot,T) - w_N(\cdot,T)\|^2_{GLL} &= \int_{-1}^{1} |w_{exact}(\xi_\alpha,T) - w_N(\xi_\alpha,T)|^2 d\xi \\ &\approx \sum_{\alpha=0}^{N} \rho_\alpha |w_{exact}(\xi_\alpha,T) - w_N(\xi_\alpha,T)|^2, \qquad (3.3)\end{aligned}$$

so $\|\cdot\|_{GLL}$ is the square root of this. This norm is a space norm, and we will use it to study convergence in both space and time. This norm will be used to study the convergence throughout the thesis, and we will call it the GLL-norm. The norm is the

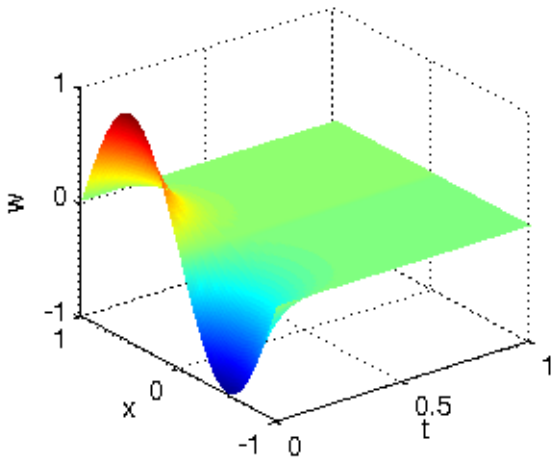


Figure 3.3: Solution of equation (3.1) with $N = 50$ and $h = 10^{-3}$.

best norm we can use for our problem because it is based on the quadrature that we use and the nodes in which we have calculated the solution. We are also interested in the relative norm, which is the above estimate divided by the norm of the exact solution.

The convergence results are given in figure 3.4. From the left side of the figure, we observe that the error in space decreases rapidly in the beginning, and then flattens out. From the lecture notes of Rønquist, we have an a priori estimate for the approximation error

$$\|w - w_N\| \leq cN^{1-\sigma}\|w\|_{H^\sigma}. \tag{3.4}$$

We know that the exact solution (3.2) is analytic, and this should give exponential convergence. This is what we see in the first few points in the right side of figure 3.4. We then reach the best approximation possible with this method and time step $h = 10^{-4}$. We observed that we have reached a very small absolute error when the constant convergence is reached.

From the right side of figure 3.4, we see that the convergence in time is quadratic except for the largest values of h. This is expected because the convergence of Crank-Nicolson is quadratic. We are able to achieve relative errors of size 10^{-13}, which is very good.

3.3 The polynomial collocation method

Polynomial collocation methods are based on the principle that the equation should be solved exactly at a set of points called the collocation points. As for the GLL quadrature, there are many possible choices of collocation points. For easy comparison, we use the GLL nodes as collocation points, but most commonly the Chebyshev points are used.

When we use collocation methods, the boundary conditions can be treated explicitly. We enforce the boundary conditions at the boundary nodes and use collocation at the interior nodes.

We will give the details of how to solve a problem using polynomial collocation by solving the heat equation.

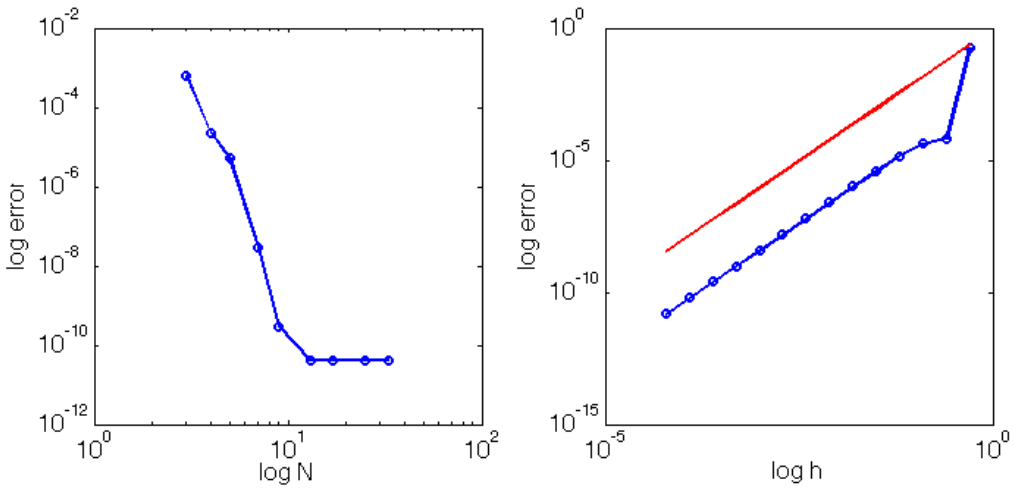


Figure 3.4: Convergence plot. The blue line is the absolute error calculated in the GLL-norm. Left: convergence in space when $h = 10^{-4}$. Right: convergence in time when $N = 50$. The red line is h plotted against h^2.

3.4 Polynomial collocation applied to the heat equation

We study the same problem as in section 3.2, equation (3.1). We use the GLL nodes as collocation points, and require that the approximate solution w_N satisfies

$$w_{N,t}(x) - w_{N,xx}(x)|_{x=x_i} = 0, \qquad i = 1, \ldots, N-1. \tag{3.5}$$

In addition, the boundary conditions require $w_N(x_0) = w_N(x_N) = 0$. Next, we write

$$w_N(x) = \sum_{j=0}^{N} w_N(x_j)\ell_j(x), \tag{3.6}$$

where ℓ_j are the Lagrangian polynomials. It is convenient to use the same polynomials as for the Galerkin method, for easier comparance.

Next, we insert (3.6) into (3.5), giving

$$\sum_{j=0}^{N} \dot{w}_N(x_j)\ell_j(x_i) - \sum_{j=0}^{N} w_N(x_j)\ell_j''(x_i) = 0, \quad i = 1, \ldots, N-1.$$

Then, we write $w_N(x_i)$ as w_i. This gives

$$\dot{w}_i - \sum_{j=0}^{N} w_j D_{ij}^{(2)}, \qquad i = 1, \ldots, N-1.$$

The matrix element $D_{ij}^{(2)}$ is the second derivative of the Lagrangian polynomial ℓ_j evaluated at x_i. We write this system as

$$M\dot{w} + Aw = 0,$$

where

$$M_{ij} = \begin{cases} \delta_{ij} & \text{if } i = 1, \ldots, N-1, \quad j = 0, \ldots, N \\ 0 & \text{otherwise} \end{cases}$$

and

$$A_{ij} = \begin{cases} D^{(2)}_{ij} & \text{if } i = 1, \ldots, N-1, \quad j = 0, \ldots, N \\ 1 & \text{if } i = j = 0 \text{ or } i = j = N \\ 0 & \text{otherwise.} \end{cases}$$

We solve the system by using Crank-Nicolson as in section 3.2. When we use $N = 50$ and $h = 10^{-3}$, we obtain the numerical solution given in figure 3.5. We study the convergence of this method in the same way as for spectral Galerkin. The results are given in figure 3.6, and we see that the results are identical to the ones in figure 3.4. Thus, we conclude that the two methods produce equally good results for this problem. This is expected because we use the same polynomials and set of solution points.

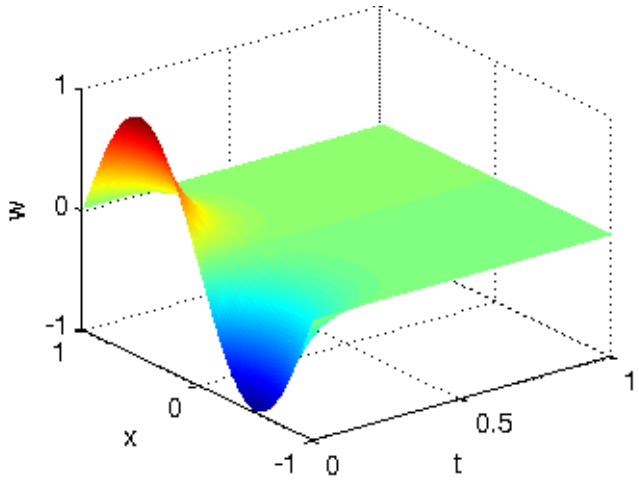


Figure 3.5: Solution of (3.1) by the polynomial collocation method when $N = 50$ and $h = 10^{-3}$.

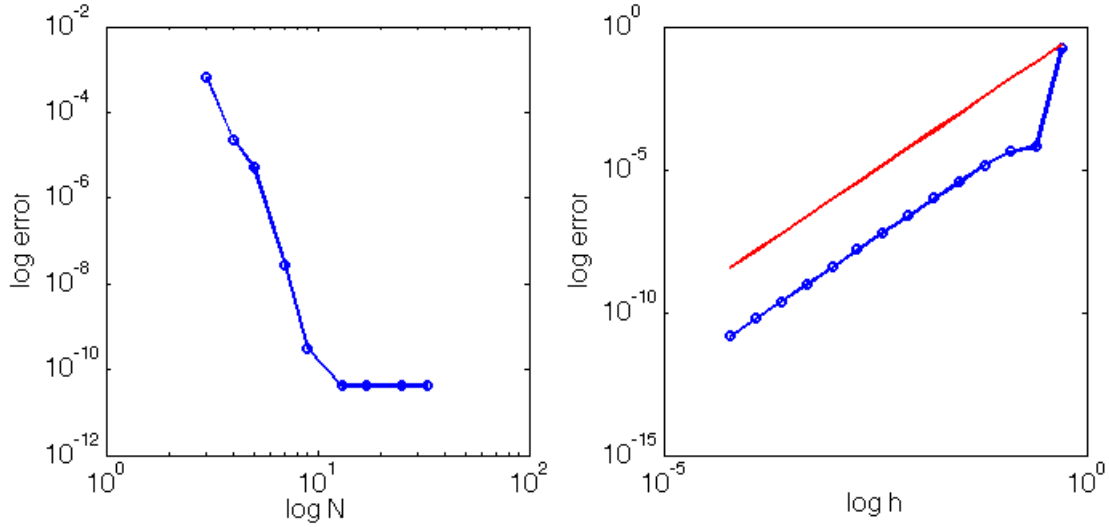


Figure 3.6: Convergence plot. The blue line is the absolute error measured in the GLL-norm. Left: convergence in space when $h = 10^{-4}$. Right: convergence in time when $N = 50$. The red line is h plotted against h^2.

Chapter 4

Solving Black-Scholes equation

In this chapter, we will solve Black-Scholes equation. We will use the spectral Galerkin method and a polynomial collocation method. The two methods will be applied to both equation (2.6) and the transformed equation (2.11) in order to compare the results. We will also solve the original equation using the finite element method.

4.1 The spectral Galerkin Method

4.1.1 Mapping to reference domain

The Lagrangian basis functions that we use for the spectral Galerkin method are constructed on the interval $\xi \in [-1, 1]$, therefore we need transformations from the domains $S \in [0, \infty]$ and $x \in [-\infty, \infty]$ to the reference domain $\hat{\Omega} = [-1, 1]$. An illustration of a mapping $S = \mathcal{F}(\xi)$ is given below.

For the mapping to work in practice, we need to set a limit for S_{max} which is large enough for our purpose. The limit must be such that the error introduced by imposing the boundary condition in the limit is sufficiently small. When $K = 10$, $S_{max} = L_S = 50$ is large enough because it is very unlikely that the spot price will be larger than $4K$ in the time interval that we will use. This is approximately equivalent to $x_{max} = L_x = 4$. The interval $[-4, 4]$ is small compared to $[-\infty, \infty]$, but the upper bound is still large enough when $K = 10$. At the lower bound, we introduce an error because $x_{\min} = -4$ translates back to $S_{\min} = 0.02$ which might not be close enough to 0. We will investigate this further when we solve the transformed equation.

For S, the mapping is given by

$$S(\xi) = \frac{L_S}{2}(\xi + 1), \qquad \frac{\partial S}{\partial \xi} = \frac{L_S}{2} = J_S,$$

while for x it is simply

$$x(\xi) = L_x \xi, \qquad \frac{\partial x}{\partial \xi} = L_x = J_x.$$

The Jacobian of the mapping must be included in the implementations. In our case, the Jacobian is the derivative of the mappings. With the mapping, we can write

$$\frac{\partial V}{\partial S} = \frac{d\xi}{dS}\frac{\partial V}{\partial \xi} = \frac{1}{J_S}\frac{\partial V}{\partial \xi},$$

and similarly for the other derivative. With these definitions, we write out the bilinear forms (2.14) and (2.16):

$$a(V,\varphi) = -\frac{1}{2J_S}\sigma^2 \int_{\hat{\Omega}} S(\xi)^2 V_\xi \varphi_\xi d\xi - (\sigma^2 - r)\int_{\hat{\Omega}} S(\xi) V_\xi \varphi d\xi - rJ_S \int_{\hat{\Omega}} V\varphi d\xi \tag{4.1}$$

and

$$\tilde{a}(v,\varphi) = -\frac{1}{2J_x}\sigma^2 \int_{\hat{\Omega}} v_\xi \varphi_\xi d\xi + r\int_{\hat{\Omega}} v_\xi \varphi d\xi - rJ_x \int_{\hat{\Omega}} v\varphi d\xi. \tag{4.2}$$

The time integral (2.15) is

$$\frac{d}{dt}(V,\varphi) = J_S \int_{\hat{\Omega}} V_t \varphi d\xi,$$

and similarly for the transformed equation, but with J_x and v instead.

4.1.2 Spectral approximation

As in section 3.2, we start with the bilinear form in equation (4.1) and write the discrete solution V_N as

$$V_N = \sum_{j=1}^{N-1} V_j \ell_j(x).$$

And again, we have

$$a\left(\sum_{j=1}^{N-1} V_j \ell_j(x), \ell_i(x)\right) = \sum_{j=1}^{N-1} a(\ell_j, \ell_i) V_j \quad i = 1, \ldots, N-1,$$

where we have excluded the two endpoints where we have Dirichlet conditions. The evaluation of the integrals in (4.1) by GLL quadrature yields

$$\begin{aligned}
a(\ell_j, \ell_i) &= -\frac{1}{2J_S}\sigma^2 \sum_{\alpha=0}^{N} \rho_\alpha \ell_j'(\xi_\alpha)\ell_i'(\xi_\alpha) S_\alpha^2 - (\sigma^2 - r)\sum_{\alpha=0}^{N} \rho_\alpha \ell_j'(\xi_\alpha)\ell_i(\xi_\alpha) S_\alpha \\
&\quad - rJ_S \sum_{\alpha=0}^{N} \rho_\alpha \ell_j(\xi_\alpha)\ell_i(\xi_\alpha) \\
&= -\frac{1}{2J_S}\sigma^2 \sum_{\alpha=0}^{N} \rho_\alpha D_{\alpha i} D_{\alpha j} S_\alpha^2 - (\sigma^2 - r)\sum_{\alpha=0}^{N} \rho_\alpha D_{\alpha j}\ell_i(\xi_\alpha) S_\alpha \\
&\quad - rJ_S \sum_{\alpha=0}^{N} \rho_\alpha \ell_j(\xi_\alpha)\ell_i(\xi_\alpha) \\
&= -\frac{1}{2J_S}\sigma^2 \sum_{\alpha=0}^{N} \rho_\alpha D_{\alpha i} D_{\alpha j} S_\alpha^2 - (\sigma^2 - r)\rho_i D_{ij} S_i - rJ_S \rho_i \delta_{ij} \\
&= A_{ij}, \qquad i = 1, \ldots, N-1, \quad j = 1, \ldots N-1.
\end{aligned}$$

Agian, we have used that $\ell_j(\xi_i) = \delta_{ij}$ and defined $\ell_j'(\xi_i) = D_{ij}$. The GLL quadrature is exact because the degree of the polynomial is $2N - 2$.

Similarly, we can approximate the bilinear form in equation (4.2) by

$$\begin{aligned}\tilde{a}(\ell_j,\ell_i) &= -\frac{1}{2J_x}\sigma^2\sum_{\alpha=0}^{N}\rho_\alpha\ell_j'(\xi_\alpha)\ell_i'(\xi_\alpha) + (r-\frac{1}{2}\sigma^2)\sum_{\alpha=0}^{N}\rho_\alpha\ell_j'(\xi_\alpha)\ell_i(\xi_\alpha)\\ &\quad - rJ_x\sum_{\alpha=0}^{N}\rho_\alpha\ell_j(\xi_\alpha)\ell_i(\xi_\alpha)\\ &= -\frac{1}{2J_x}\sigma^2\sum_{\alpha=0}^{N}\rho_\alpha D_{\alpha i}D_{\alpha j} + (r-\frac{1}{2}\sigma^2)\sum_{\alpha=0}^{N}\rho_\alpha D_{\alpha j}\ell_i(\xi_\alpha) - rJ_x\sum_{\alpha=0}^{N}\rho_\alpha\ell_j(\xi_\alpha)\ell_i(\xi_\alpha)\\ &= r\rho_i D_{ij} - (r-\frac{1}{2}\sigma^2)J_x\rho_i\delta_{ij} - \frac{1}{2J_x}\sigma^2\sum_{\alpha=0}^{N}\rho_\alpha D_{\alpha i}D_{\alpha j}\\ &= \tilde{A}_{ij}, \qquad i=1,\dots,N-1, \quad j=1,\dots N-1.\end{aligned}$$

The time dependent integral is equal in both cases, and is approximated by

$$(\ell_i,\ell_j) = J\sum_{\alpha=0}^{N}\rho_\alpha\ell_j(\xi_\alpha)\ell_i(\xi_\alpha) = \rho_i J\delta_{ij} = M_{ij}.$$

The boundary conditions at $S = 0$ and $x = x_{\min}$ can be imposed directly by setting $V_0 = e^{-r(T-t)}$ and $V_N = 0$, or by adding a right-hand side and changing the corresponding elements in $\mathbf{A}$ and $\mathbf{M}$. We choose the first alternative.

As in section 3.2, we write this as the system $\mathbf{M}\dot{\mathbf{V}} + \mathbf{A}\mathbf{V} = 0$. We use implicit Euler as time discretization.

4.1.3 Numerical results

Original equation

Unless otherwise stated, we have used a time step of size $h = 10^{-3}$ and $N = 60$, giving $N + 1 = 61$ nodes and basis functions. The values used in computations are given in table 4.1.

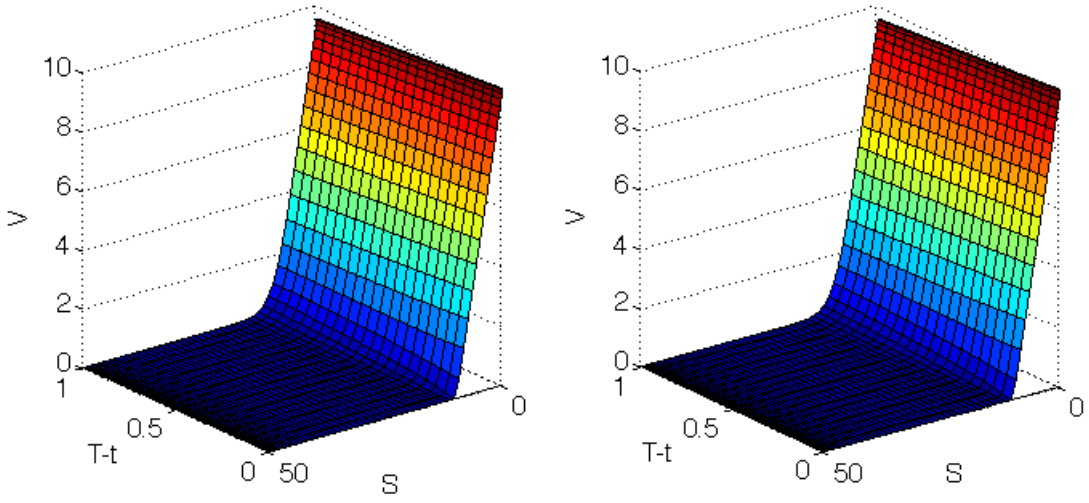


Figure 4.1: Solutions for equation (2.6) when $N = 60$ and $h = 10^{-3}$. Left: the numerical solution. Right: the exact solution.

Constant	Value
T	1
K	10
σ	0.25
r	0.05

Table 4.1: Values of constants used in the calculatuions

N	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
2	0.38675006	0	0.07041806
3	0.41805756	0	0.10762645
4	0.23005799	0.74940399	0.07228541
6	0.17495923	1.25668284	0.05050752
8	0.11975595	0.94190019	0.03530606
12	0.07410233	1.23942794	0.02174223
16	0.03513480	1.76912464	0.01030305
24	0.03407096	1.12097629	0.00999052
32	0.00751120	2.22578589	0.00220248
48	0.00566842	2.58752428	0.00166213
64	0.00031284	4.58553168	0.00009173
96	0.00017004	5.05902414	0.00004986
128	0.00036715	-0.23093662	0.00010766
192	0.00024370	-0.51927920	0.00007146
256	0.00013606	1.43211443	0.00003990
384	0.00004109	2.56819660	0.00001205
512	0.00010336	0.39662536	0.00003031
768	0.00004097	0.00412725	0.00001201

Table 4.2: Values for absolute error and relative error and the corresponding values of α. The time step is kept fixed at $h = 10^{-3}$.

h	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
1/2	0.01336128	0	0.00383226
1/4	0.00797037	0.74533933	0.00230652
1/8	0.00430943	0.88714976	0.00125448
1/16	0.00251610	0.77630871	0.00073498
1/32	0.00209822	0.26202278	0.00061407
1/64	0.00218373	-0.05762556	0.00063972
1/128	0.00230782	-0.07974137	0.00067641

Table 4.3: Values for absolute error and relative error and the corresponding values of α when $N = 60$.

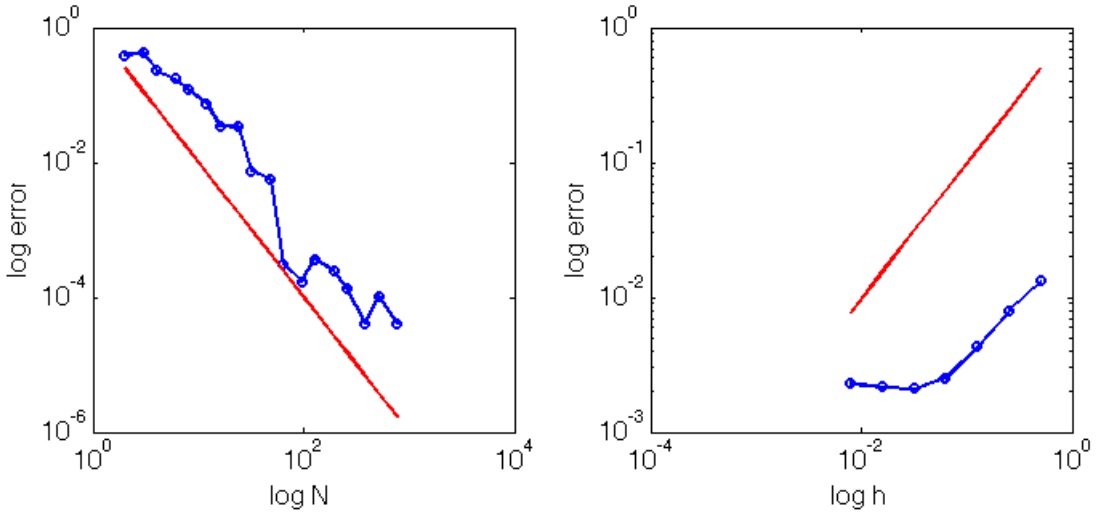


Figure 4.2: Convergence for (2.6). Left: Convergence in space. Log-log plot of the GLL-norm plotted against N when $h = 10^{-3}$. The red line is $1/N^2$ plotted against N. Right: Convergence in time. Log-log plot of the GLL-error plotted against h when $N = 60$. The red line is h against h.

The solution for V is given in figure 4.1. We can see from the figure that the numerical solution and the exact solution looks the same, which is good since we want the numerical solution to converge to the exact solution. We can examine this further by looking at the GLL-norm (3.3). Now, the exact solution is the solution given in equation (2.8). We are interested in both the absolute norm and the relative norm. For the relative norm, we need the norm of the exact solution, which is

$$\|V^{BS}\|_{GLL} \approx \sqrt{\sum_{\alpha=0}^{N} \rho_\alpha V^{BS}(\xi_\alpha, T)^2}.$$

The relative norm will be given in the converence tables. To give expressions for the slope of the convergence, we use the relations

$$\|\cdot(N)\|_{GLL} = CN^{-\alpha}$$

and

$$\|\cdot(2N)\|_{GLL} = C2^{-\alpha}N^{-\alpha},$$

which gives the following expression for α

$$\alpha = \frac{\ln \|\cdot(N)\|_{GLL} - \ln \|\cdot(2N)\|_{GLL}}{\ln 2}.$$

Convergence in space is shown in the left side of figure 4.2, and the corresponding values are given in table 4.2. The figure shows that the norm is approximately quadratic, even though the α-values in the table seems random. We would normally expect better convergence than quadratic, but because the initial data is not smooth, we do not get optimal convergence. The a priori estimate in (3.4) explains this. The solution of (2.6) is in H^1, but the second derivative has a δ-function at $t = 0$. This means that the solution is almost in H^2 and we can obtain quadratic convergence with

a large constant term. Despite the initial condition, we are able to achieve relative errors of size 10^{-5}.

The right side of figure 4.2 shows the convergence in time. We see that the convergence is linear in the beginning, before it reaches a constan convergence. The linear convergence is expected, because this is the convergence of the time discretization we have used. Because the convergence in space is approximately quadratic, we could expect to obtain the same in time. Therefore, we try to use Crank-Nicolson as discretization scheme instead. The discrtized system is then

$$\left(\mathbf{M} + \frac{h}{2}\mathbf{A}\right) V^{k+1} = \frac{h}{2}\left(f^k + f^{k+1} - \mathbf{A}V^k\right) + \mathbf{M}V^k.$$

This expression is used to obtain the convergence results in figure 4.3. We observe that the convergence in space is approximately the same and that the convergence in time reaches the constant convergence much faster than earlier. From table 4.4, we see that the values are almost the same as in table 4.3 when $h \leq 1/16$.

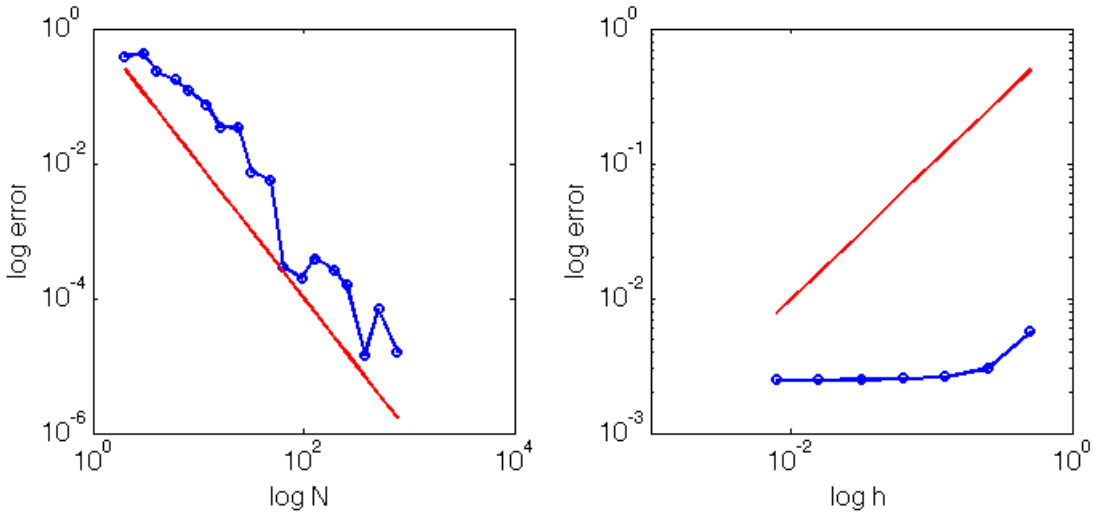


Figure 4.3: Convergence for (2.6) when Crank-Nicolson is used as time discretization. Left: Convergence in space. Log-log plot of the GLL-norm plotted against N when $h = 10^{-3}$. The red line is $1/N^2$ plotted against N. Right: Convergence in time. Log-log plot of the GLL-error plotted against h when $N = 60$. The red line is h against h.

To study the impact of the "bad" initial data, we study the difference between the exact solution from Black-Scholes formula and the numerical solution, in other words the difference between the two plots in figure 4.1. This is given in figure 4.4. We see that the difference is largest at $K = S$ when we are close to the initial condition at $t = T$, when $T - t$ is small. The error then decreases with time. This indicates that the spectral method does not handle the sharp edge of the initial data very well.

In figure 4.5, we have increased T from 1 to 5 in order to see that the error continues to decrease with time. To avoid introducing errors because L_S is too small, we increased it to $L_S = 100$. For both figures, the largest difference is approximately $1.5 \cdot 10^{-2}$, which is small compared to the solution. It also seems to be the case that most of the error is caused by the initial data. When we reach the constant convergence in time, this difference is constant.

h	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
1/2	0.00571293	0	0.00163857
1/4	0.00304873	0.90601989	0.00088226
1/8	0.00263321	0.21138982	0.00076653
1/16	0.00254081	0.05153433	0.00074219
1/32	0.00250727	0.01916958	0.00073378
1/64	0.00249287	0.00830960	0.00073028
1/128	0.00248624	0.00383845	0.00072871

Table 4.4: Values for absolute and relative error and the corresponding values of α when $N = 60$ and Crank-Nicolson is used as time discretization.

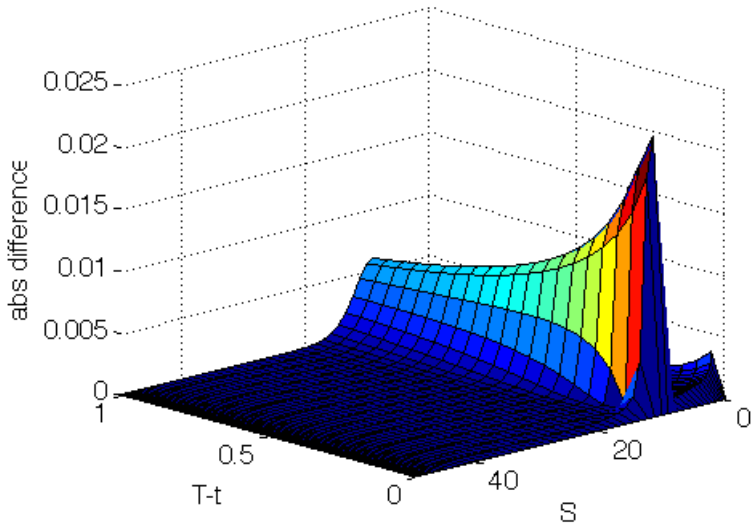


Figure 4.4: Difference between the exact solution and the numerical solution for $N = 60$ and $h = 10^{-3}$.

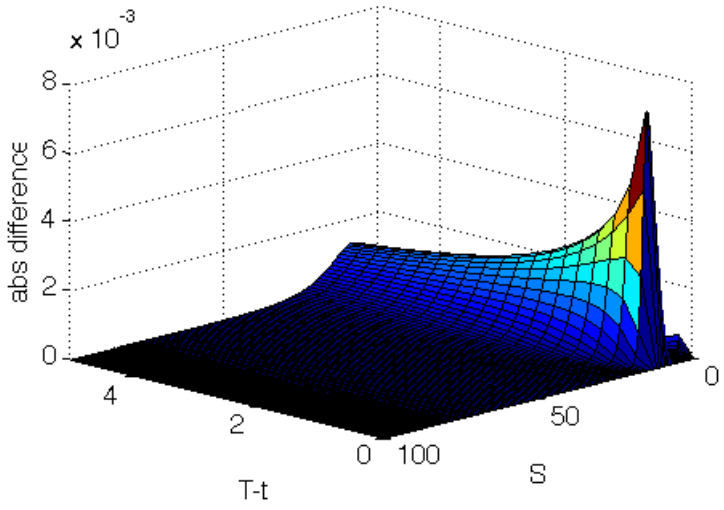


Figure 4.5: Difference between the exact solution and the numerical solution on a larger domain where $T = 5$, $L_S = 100$, $N = 120$ and $h = 10^{-3}$.

To get better convergence and get rid of the error due to initial data, it could be tempting to try to approximate the initial data by a smooth function. In order to get a good approximation, we would need many nodes around the point $K = S$. This means that we would need a large N to obtain sufficiently accuracy on the domain, and the computation time would probably increase more than the correctness. In addition, the solution would converge to the wrong solution. When we then reach small enough errors, the error due to the approximation of the initial data would dominate. Thus, this is not a good idea if we want to reduce the error.

Transformed equation

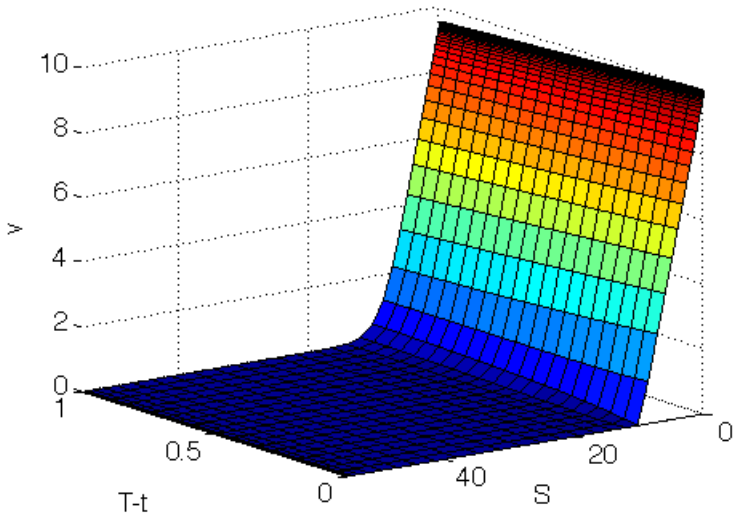


Figure 4.6: Numerical solution for the transformed equation (2.11) when $N = 60$ and $h = 10^{-3}$.

The numerical solution for the transformed equation (2.11) is given in figure 4.6. From the figure, we see that the solution looks like the solutions in figure 4.1. However, the convergence plots in figure 4.7 shows that the convergence is not as good as for the original equation. From the left side of figure 4.7, we see that the error is approximately quadratic in the beginning before it reaches a constant error of size $3 \cdot 10^{-3}$, but the error of the original equation continues to decrease. This is reflected in the values in table 4.5. It also seems that the solution of the transformed equation is not very stable for small N

The right side of figure 4.7 and table 4.6 shows that the convergence in time behaves as the convergence of the original equation, but is a little bit larger.

From figure 4.8, we see that the difference between the exact and numerical solution behaves almost in the same way as for the original equation, but we also have an error along the boundary $x = x_{\min}$ because $e^{-4} \neq 0$. The error along this boundary does not decrease as we increase N because the endpoint value does not change. Therefore, this can be one of the reasons that we have a larger error for the transformed equation.

From the results in this section, it looks like it is better to solve the original equation when we want to obtain good numerical convergence.

N	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
2	0.01697580	0	0.00150831
3	0.21477835	0	0.02165946
4	0.06293417	-1.89036352	0.00602614
6	0.14856888	0.53171665	0.01440888
8	0.05674811	0.14927103	0.00545679
12	0.10092918	0.55778863	0.00972871
16	0.03479542	0.70567512	0.00335268
24	0.01707888	2.56305793	0.00164555
32	0.01757554	0.98532797	0.00169340
48	0.01010895	0.75658020	0.00097400
64	0.00567725	1.63030467	0.00054700
96	0.00310244	1.70415895	0.00029892
128	0.00324821	0.80554765	0.00031296
192	0.00297616	0.05994806	0.00028675
256	0.00295461	0.13667779	0.00028468

Table 4.5: Values for absolute error and relative error and the corresponding values of α. The time step is kept fixed at $h = 10^{-3}$.

h	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
1/2	0.02424988	0	0.00229451
1/4	0.01480968	0.71143664	0.00141152
1/8	0.00790935	0.90490974	0.00075751
1/16	0.00402587	0.97425983	0.00038667
1/32	0.00270893	0.57157665	0.00026059
1/64	0.00276022	-0.02706003	0.00026574
1/128	0.00301833	-0.12897001	0.00029071

Table 4.6: Values for absolute error and relative error and the corresponding values of α when $N = 60$.

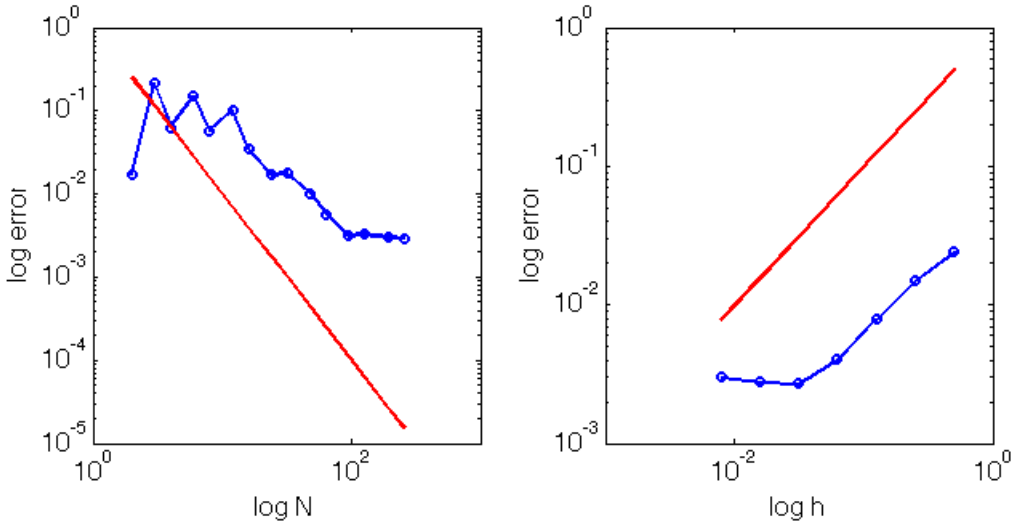


Figure 4.7: Convergence for (2.11). Left: Convergence in space. Log-log plot of the GLL-norm plotted against N when $h = 10^{-3}$. The red line is $1/N^2$ plotted against N. Right: Convergence in time. Log-log plot of the GLL-error plotted against h when $N = 60$. The red line is h against h.

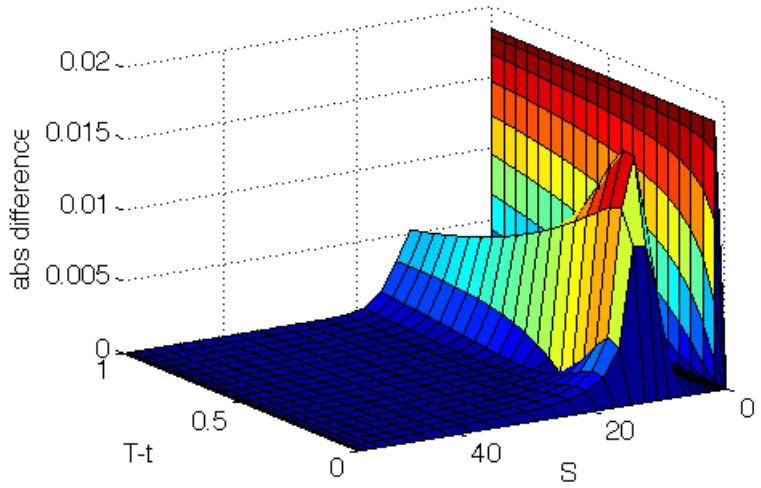


Figure 4.8: Difference between the exact solution and the numerical solution for the transformed equation.

4.1.4 Analysis of the method

The book by Achdou and Pironneau [1] gives the analysis of the weak formulation of Black-Scholes. The analysis in this section is similar to the book, but slightly changed to adapt it to our problem. To avoid repeating results, the analysis is performed only for the original problem, equation (2.6). We assume that we have transformed the equation to the domain $[-1, 1]$ and that we have homogeneous Dirichlet boundary conditions.

Before we start the analysis, we introduce the following energy norm

$$\|\varphi\|_E^2 = \|S\varphi_S\|_{L^2(\Omega)}^2 + \|\varphi\|_{L^2(\Omega)}^2, \tag{4.3}$$

the seminorm

$$|\varphi|_E^2 = \|S\varphi_S\|_{L^2(\Omega)}^2 \tag{4.4}$$

and the corresponding space

$$E = \left\{\varphi \in L^2(\Omega) : \|\varphi\|_E < \infty\right\}.$$

We will also use the following version of Poincaré's inequality

$$\|\varphi\|_{L^2(\Omega)} \leq 2\|S\varphi_S\|_{L^2(\Omega)} \qquad \forall \varphi \in E. \tag{4.5}$$

Next, we state a lemma showing that the seminom is a norm.

Lemma 1. The seminorm (4.4) is a norm equivalent to (4.3).

Proof.

$$\begin{aligned} |\varphi|_E^2 &= \|S\varphi_S\|_{L^2(\Omega)}^2 \\ &\leq \|\varphi\|_E^2 = \|S\varphi_S\|_{L^2(\Omega)}^2 + \|\varphi\|_{L^2(\Omega)}^2 \\ &\leq \|S\varphi_S\|_{L^2(\Omega)}^2 + 4\|S\varphi_S\|_{L^2(\Omega)}^2 = 5|\varphi|_E^2. \end{aligned}$$

To obtain the last inequality, we used Poincaré's inequality. □

For easier analysis, we will make the equation forward in time. We let $t = T - t$, and write

$$a(V, \varphi) = \int_\Omega \frac{1}{2}\sigma^2 S^2 V_S \varphi_S + \int_\Omega (\sigma^2 - r) S V_S \varphi + \int_\Omega r V \varphi. \tag{4.6}$$

The sign of the time derivative remains unchanged.

Gårding's inequality

In stead of the coercivity of the bilinear form, we give an estimate for the lower bound of the bilinear form through Gårding's inequality. The estimate is valid at each time step, which is enough to know that it holds for the method.

Lemma 2.

$$a(\varphi, \varphi) \geq \frac{1}{4}\sigma^2 |\varphi|_E^2 - K_0 \|\varphi\|_{L^2(\Omega)}^2. \tag{4.7}$$

Proof. Starting with the bilinear form (4.6), we have

$$\begin{aligned} a(\varphi, \varphi) &= \frac{1}{2}\sigma^2 |\varphi|_E^2 + r\|\varphi\|_{L^2(\Omega)}^2 + (\sigma^2 - r)\int_\Omega S\varphi_S\varphi \\ &\geq \frac{1}{2}\sigma^2 |\varphi|_E^2 + r\|\varphi\|_{L^2(\Omega)}^2 - |\sigma^2 - r|\int_\Omega (\frac{1}{2}\epsilon S^2 \varphi_S^2 + \frac{1}{2\epsilon}\varphi^2) \\ &\geq (\frac{1}{2}\sigma^2 - \frac{1}{2}|\sigma^2 - r|\epsilon)|\varphi|_E^2 - \left(-r + \frac{|\sigma^2 - r|}{2\epsilon}\right)\|\varphi\|_{L^2(\Omega)}^2. \end{aligned}$$

Choosing

$$\epsilon = \frac{\sigma^2}{2|\sigma^2 - r|}$$

gives

$$a(\varphi,\varphi) \geq \frac{1}{4}\sigma^2|\varphi|_E^2 - \left(-r + \frac{|\sigma^2 - r|^2}{\sigma^2}\right)\|\varphi\|_{L^2(\Omega)}^2,$$

which is equal to (4.7) with

$$K_0 = -r + \frac{|\sigma^2 - r|^2}{\sigma^2}.$$

□

By using the Poincaré inequality, we can obtain an estimate of the form $a(\varphi,\varphi) \geq \mathcal{K}|\varphi|_E r$, but the lower bound obtained form Gårding's inequality gives a better estimate and is therefore preferred.

Continuity

We state a lemma concerning the continuity of the bilinear form.

Lemma 3.

$$|a(\varphi,\psi)| \leq C|\varphi|_E|\psi|_E. \tag{4.8}$$

Proof. By using the Poincaré inequality, we get the following estimates for each of the integrals in (4.6)

$$\begin{aligned}
&|\int_\Omega \frac{1}{2}\sigma^2 S^2 \varphi_S \psi_S| \leq \frac{1}{2}\sigma^2|\varphi|_E|\psi|_E,\\
&|\int_\Omega (-r+\sigma^2) S\varphi_S\psi| \leq (r+\sigma^2)|\varphi|_E\|\psi\|_{L^2(\Omega)} \leq 2(r+\sigma^2)|\varphi|_E|\psi|_E \qquad \text{and}\\
&r|\int_\Omega \varphi\psi| \leq 4r|\varphi|_E|\psi|_E.
\end{aligned}$$

Combining the above, we get

$$|a(\varphi,\psi)| \leq \left(\frac{5}{2}\sigma^2 + 6r\right)|\varphi|_E|\psi|_E$$

which shows that the bilinear form is continuous with constant $C = \frac{5}{2}\sigma^2 + 6r$. □

Now, we have showed that the bilinear form is both "coercive" and continuous at each time t. If we then apply a time discretization to the weak formulation, Lax-Milgram's lemma $A.2$ ensures that a solution exists and that the solution will be unique.

Stability

We will give an energy estimate to show that the solution is stable.

Lemma 4. The equation is stable with energy estimate

$$e^{-2K_0 t}\|\varphi(t)\|_{L^2(\Omega)}^2 + \frac{\sigma^2}{2}\int_0^t e^{-2K_0 t}|\varphi(t)|_E^2 dt \leq \|\varphi(0)\|_{L^2(\Omega)}^2. \tag{4.9}$$

Proof. We start with the weak equation (2.13) and use the test function $\varphi e^{-2K_0 t}$:

$$\int_0^t \int_\Omega e^{-2K_0 t}\varphi_t \varphi dt\, dx + \int_0^t e^{-2K_0 t} a(\varphi, \varphi) dt = 0. \tag{4.10}$$

Note that

$$\frac{\partial}{\partial t}\left(e^{-2K_0 t}\varphi^2\right) = -2K_0 e^{-2K_0 t}\varphi^2 + 2e^{-2K_0 t}\varphi\varphi_t,$$

so we can write (4.10) as

$$\int_0^t \frac{1}{2}\frac{\partial}{\partial t}\left(e^{-2K_0 t}\|\varphi\|^2_{L^2(\Omega)}\right) dt + K_0 \int_0^t e^{-2K_0 t}\|\varphi\|^2_{L^2(\Omega)} dt$$
$$+ \frac{\sigma^2}{4}\int_0^t e^{-2K_0 t}|\varphi|^2_E dt - K_0 \int_0^t e^{-2K_0 t}\|\varphi\|^2_{L^2(\Omega)} dt \leq 0.$$

This is equal to

$$\int_0^t \frac{1}{2}\frac{\partial}{\partial t}\left(e^{-2K_0 t}\|\varphi\|^2_{L^2(\Omega)}\right) dt + \frac{\sigma^2}{4}\int_0^t e^{-2K_0 t}|\varphi|^2_E dt \leq 0.$$

By applying the fundamental theorem of calculus to the above equation, we get

$$\frac{1}{2}e^{-2K_0 t}\|\varphi(t)\|^2_{L^2(\Omega)} - \frac{1}{2}\|\varphi(0)\|^2_{L^2(\Omega)} + \frac{\sigma^2}{4}\int_0^t e^{-2K_0 t}|\varphi|^2_E dt \leq 0,$$

which is equal to (4.9) when we multiply by 2 and rearrange. □

Convergence

The last part of the analysis concerns the convergence of the method. We want the discrete solution V_h to converge to the continuous solution V. A similar proof is found in [12, chapter 5.3].

Lemma 5.

$$\|V(t) - V_h(t)\|^2_{L^2(\Omega)} \leq \frac{h(t)}{N^2} e^{(K_0 + \frac{1}{2})t}.$$

Proof. Note that we can write

$$a(V - V_h, V - V_h) = a(V - V_h, V - \varphi_h) + a(V - V_h, \varphi_h - V_h).$$

By using the weak formulation of the equation and subtracting the same formulation for the discrete solution, we get

$$\left(\frac{\partial}{\partial t}(V - V_h), \varphi_h - V_h\right) + a(V - V_h, \varphi_h - V_h) = 0. \tag{4.11}$$

By using the estimate from Gårding inequality (4.7) and the two above equations, we get

$$\frac{1}{4}\sigma^2|V - V_h|^2_E - K_0\|V - V_h\|^2_{L^2(\Omega)} \leq a(V - V_h, V - \varphi_h) - \left(\frac{\partial}{\partial t}(V - V_h), \varphi_h - V_h\right). \tag{4.12}$$

Next, we apply the continuity estimate and Young's inequality

$$\begin{aligned} a(V - V_h, V - \varphi_h) &\leq C|V - V_h|_E |V - \varphi_h|_E \\ &\leq \epsilon |V - V_h|_E^2 + \frac{C^2}{4\epsilon}|V - \varphi_h|_E^2 \\ &\leq \frac{1}{4}\sigma^2 |V - V_h|_E^2 + \frac{C^2}{\sigma^2}|V - \varphi_h|_E^2, \end{aligned} \tag{4.13}$$

where we have chosen $\epsilon = \frac{1}{4}\sigma^2$. Next, we use $\varphi_h - V_h = (\varphi_h - V) + (V - V_h)$ to write

$$-\left(\frac{\partial}{\partial t}(V - V_h), \varphi_h - V_h\right) = \left(\frac{\partial}{\partial t}(V - V_h), V - \varphi_h\right) - \frac{1}{2}\frac{\partial}{\partial t}\|V - V_h\|_{L^2(\Omega)}^2. \tag{4.14}$$

Inserting (4.13) and (4.14) into (4.12), we get

$$\begin{aligned} &\frac{1}{4}\sigma^2 |V - V_h|_E^2 - K_0 \|V - V_h\|_{L^2(\Omega)}^2 \\ &\leq \frac{1}{4}\sigma^2 |V - V_h|_E^2 + \frac{C^2}{\sigma^2}|V - \varphi_h|_E^2 + \left(\frac{\partial}{\partial t}(V - V_h), V - \varphi_h\right) - \frac{1}{2}\frac{\partial}{\partial t}\|V - V_h\|_{L^2(\Omega)}^2. \end{aligned}$$

Rearranging, this gives

$$\frac{1}{2}\frac{\partial}{\partial t}\|V - V_h\|_{L^2(\Omega)}^2 \leq \frac{C^2}{\sigma^2}|V - \varphi_h|_E^2 + K_0\|V - V_h\|_{L^2(\Omega)}^2 + \left(\frac{\partial}{\partial t}(V - V_h), V - \varphi_h\right).$$

Now, we multiply the inequality by 2 and integrate over t which results in

$$\begin{aligned} \|(V - V_h)(t)\|_{L^2(\Omega)}^2 \leq \|(V - V_h)(0)\|_{L^2(\Omega)}^2 &+ \frac{2C^2}{\sigma^2}\int_0^t |V(t) - \varphi_h|_E^2 dt \\ &K_0 \int_0^t \|(V - V_h)(t)\|_{L^2(\Omega)}^2 dt + 2\int_0^t \left(\frac{\partial}{\partial t}(V - V_h)(t), V(t) - \varphi_h\right) dt. \end{aligned} \tag{4.15}$$

The last integral in (4.15) can be written as

$$\begin{aligned} &\int_0^t \left(\frac{\partial}{\partial t}(V - V_h)(t), V(t) - \varphi_h\right) dt = -\int_0^t \left((V - V_h)(t), \frac{\partial}{\partial t}(V(t) - \varphi_h)\right) dt \\ &+ ((V - V_h)(t), (V - \varphi_h)(t)) - ((V - V_h)(0), (V - \varphi_h)(0)) \\ &\leq \int_0^t \|\frac{\partial}{\partial t}(V(t) - \varphi_h)\|_{L^2(\Omega)}^2 dt + \frac{1}{4}\int_0^t \|(V - V_h)(t)\|_{L^2(\Omega)}^2 dt + \frac{1}{4}\|\frac{\partial}{\partial t}(V - V_h)(t)\|_{L^2(\Omega)}^2 \\ &+ \|(V - \varphi_h)(t)\|_{L^2(\Omega)}^2 + \|(V - V_h)(0)\|_{L^2(\Omega)}\|(V - \varphi_h)(0)\|_{L^2(\Omega)}. \end{aligned}$$

Inserting this into (4.15), we get

$$\begin{aligned} \|(V - V_h)(t)\|_{L^2(\Omega)}^2 \leq \|(V - V_h)(0)\|_{L^2(\Omega)}^2 &+ \frac{2C^2}{\sigma^2}\int_0^t |V(t) - \varphi_h|_E^2 dt \\ &+ K_0 \int_0^t \|(V - V_h)(t)\|_{L^2(\Omega)}^2 dt + 2\int_0^t \|\frac{\partial}{\partial t}(V(t) - \varphi_h)\|_{L^2(\Omega)}^2 dt \\ &+ \frac{1}{2}\int_0^t \|\frac{\partial}{\partial t}(V - V_h)(t)\|_{L^2(\Omega)}^2 dt + \frac{1}{2}\|(V - V_h)(t)\|_{L^2(\Omega)}^2 \\ &+ 2\|(V - \varphi_h)(t)\|_{L^2(\Omega)}^2 + 2\|(V - V_h)(0)\|_{L^2(\Omega)}\|(V - \varphi_h)(0)\|_{L^2(\Omega)}. \end{aligned}$$

Finally, we can write this as

$$\begin{aligned}\frac{1}{2}\|(V-V_h)(t)\|^2_{L^2(\Omega)} \leq \|(V-V_h)(0)\|^2_{L^2(\Omega)} + \frac{2C^2}{\sigma^2}\int_0^t |V(t)-\varphi_h|^2_E dt \\ + (K_0+\frac{1}{2})\int_0^t \|(V-V_h)(t)\|^2_{L^2(\Omega)}dt + 2\int_0^t \|\frac{\partial}{\partial t}(V(t)-\varphi_h)\|^2_{L^2(\Omega)}dt \\ + 2\|(V-\varphi_h)(t)\|^2_{L^2(\Omega)} + 2\|(V-V_h)(0)\|_{L^2(\Omega)}\|(V-\varphi_h)(0)\|_{L^2(\Omega)}.\end{aligned}$$

To bound the integrals, we will use $\varphi_h = \Pi_N^{GLL}V$, together with $V_h(0) = \varphi_h(0)$,

$$\|V - \Pi_N^{GLL}V\|_{L^2(\Omega)} \leq \tilde{C}\frac{1}{N}\|V_S\|_{L^2(\Omega)}$$

and

$$|V - \Pi_N^{GLL}V|_E \leq L_S|V - \Pi_N^{GLL}V|_{H^1(\Omega)} \leq \tilde{C}_S\frac{1}{N}\|V_{SS}\|_{L^2(\Omega)},$$

where $\tilde{C}_S$ depends on L_S. This gives

$$\begin{aligned}E_1 &= \frac{2C^2}{\sigma^2}\int_0^t |V(t)-\Pi_N^{GLL}V|^2_E dt \leq \tilde{C}_1\frac{1}{N^2}\int_0^t \|V_{SS}\|^2_{L^2(\Omega)} \leq C_1\frac{1}{N^2}C_0(t) \\ E_2 &= 2\int_0^t \|\frac{\partial}{\partial t}(V(t)-\Pi_N^{GLL}V)\|^2_{L^2(\Omega)}dt \leq \tilde{C}_2\frac{1}{N^2}\int_0^t \|\frac{\partial}{\partial t}V_S\|^2_{L^2(\Omega)} \leq C_2\frac{1}{N^2}C_0(t) \\ E_3 &= 2\|(V-\Pi_N^{GLL}V)(t)\|^2_{L^2(\Omega)} \leq \tilde{C}_3\frac{1}{N^2}\|V_S\|^2_{L^2(\Omega)} \leq C_3\frac{1}{N^2}C_0(t) \\ E_4 &= \|(V-V_h)(0)\|^2_{L^2(\Omega)} + 2\|(V-V_h)(0)\|_{L^2(\Omega)}\|(V-\Pi_N^{GLL}V)(0)\|_{L^2(\Omega)} \\ &\leq \tilde{C}_4\frac{1}{N^2}\|V_S(0)\|^2_{L^2(\Omega)} \leq C_4\frac{1}{N^2}C_0(0).\end{aligned}$$

Now, we have

$$E_1 + E_2 + E_3 + E_4 \leq C_5\frac{1}{N^2}g(V)$$

for some suitable function $g(V)$. The coercivity bound can then be written as

$$\begin{aligned}&\|V(t)-V_h(t)\|^2_{L^2(\Omega)} \\ &\leq C_5\frac{1}{N^2}g(V) + (K_0+\frac{1}{2})\int_0^t \|V(t)-V_h(t)\|^2_{L^2(\Omega)}dt.\end{aligned}$$

If we use Gronwall's lemma $A.3$ to this, we obtain

$$\|V(t)-V_h(t)\|^2_{L^2(\Omega)} \leq \frac{h(t)}{N^2}e^{(K_0+\frac{1}{2})t},$$

where $h(t)$ is a suitable function involving $g(V)$. □

Other convergence estimates are also possible. For example, if we choose $\epsilon = \frac{1}{8}\sigma^2$, the norms $|V-V_h|^2_E$ will not cancel out and we get a better estimate containing $|V-V_h|^2_E$.

4.2 The spectral collocation method

4.2.1 Mapping to reference domain

As for the Galerkin method, we must use a mapping from the domains $S \in [0, S]$ and $x \in [-\infty, \infty]$ to the reference domain $\hat{\Omega} = [-1, 1]$. The mappings are the same as before, $S(\xi) = \frac{L_S}{2}(\xi + 1)$ and $x(\xi) = L_x\xi$. When we change domain, the partial derivatives change to

$$\frac{\partial V}{\partial S} = \frac{\partial V}{\partial \xi}\frac{d\xi}{dS} = \frac{2}{L_S}\frac{\partial V}{\partial \xi} = \frac{1}{J_S}\frac{\partial V}{\partial \xi}$$
$$\frac{\partial^2 V}{\partial S^2} = \frac{\partial}{\partial S}\left(\frac{\partial V}{\partial S}\right) = \frac{\partial}{\partial \xi}\frac{d\xi}{dS}\left(\frac{2}{L_S}\frac{\partial V}{\partial \xi}\right) = \frac{4}{L_S^2}\frac{\partial^2 V}{\partial \xi^2} = \frac{1}{J_S^2}\frac{\partial^2 V}{\partial \xi^2},$$

and similarly for the derivatives of v.

4.2.2 Spectral approximation

The first thing we do is to require that the approximate solution V_N solves the equation at the interior collocation points

$$V_{N,t} + \frac{1}{2}\sigma^2 S^2 V_{N,SS} + rSV_{N,S} - rV_N\Big|_{S=S_i} = 0 \qquad i = 1, \ldots, N-1.$$

The boundary points are excluded because we have Dirichlet boundaries. At the refernce domain, this is

$$V_{N,t} + \frac{1}{2}\sigma^2 S(\xi)^2 V_{N,\xi\xi} + rS(\xi)V_{N,\xi} - rV_N\Big|_{\xi=\xi_i} = 0 \qquad i = 1, \ldots, N-1.$$

As earlier, we use the GLL nodes as collocation points. If we write $V_N(S,t) = (V_N(S_1,t), \ldots, V_N(S_{N-1},t))^T$, we can approximate $V_N(S,t)$ by

$$V_N(S,t) \approx \sum_{j=1}^{N-1} V_j \ell_j(S),$$

where we have used $V_N(S_j,t) = V_j$. With this approximation, we can write

$$\begin{aligned}
&\sum_{j=1}^{N-1} \dot{V}_j \ell_j(\xi_i) + \frac{1}{2J_S^2}\sigma^2 \sum_{j=1}^{N-1} V_j S_i^2 \ell_j''(\xi_i) + r\frac{1}{J_S}\sum_{j=1}^{N-1} V_j S_i \ell_j'(\xi_i) - r\sum_{j=1}^{N-1} V_j \ell_j(\xi_i) \\
=&\sum_{j=1}^{N-1} \dot{V}_j \delta_{ij} + \frac{1}{2J_S^2}\sigma^2 \sum_{j=1}^{N-1} V_j S_i^2 \ell_j''(\xi_i) + r\frac{1}{J_S}\sum_{j=1}^{N-1} V_j S_i \ell_j'(\xi_i) - r\sum_{j=1}^{N-1} V_j \delta_{ij} \\
=&\dot{V}_i + \frac{1}{2J_S^2}\sigma^2 \sum_{j=1}^{N-1} S_i^2 D_{ij}^{(2)} V_j + r\frac{1}{J_S}\sum_{j=1}^{N-1} D_{ij} S_i V_j - rV_i = 0, \quad i = 1, \ldots, N-1,
\end{aligned}$$

where $D_{ij} = \ell_j'(\xi_i)$ and $D_{ij}^{(2)} = \ell_j''(\xi_i)$.

We now have a system of equations to solve for V, which we can write in matrix form with $M_{ij} = \delta_{ij}$ and

$$A_{ij} = \frac{1}{2J_S^2}\sigma^2 S_i^2 D_{ij}^{(2)} + r\frac{1}{J_S} D_{ij} S_i - r\delta_{ij}.$$

By the same method, we obatin the following stiffness matrix for the transformed equation

$$\tilde{A}_{ij} = \frac{1}{2J_x^2}\sigma^2 D_{ij}^{(2)} + (r - \frac{1}{2}\sigma^2)\frac{1}{J_x}D_{ij} - r\delta_{ij}.$$

We agian obtain the system $\mathbf{M}\dot{\mathbf{V}} + \mathbf{A}\mathbf{V} = 0$, which we solve using implicit Euler as in section 3.2. Similarly as for the Galerkin method, the boundary conditions are enforced directly.

4.2.3 Numerical results

Original equation

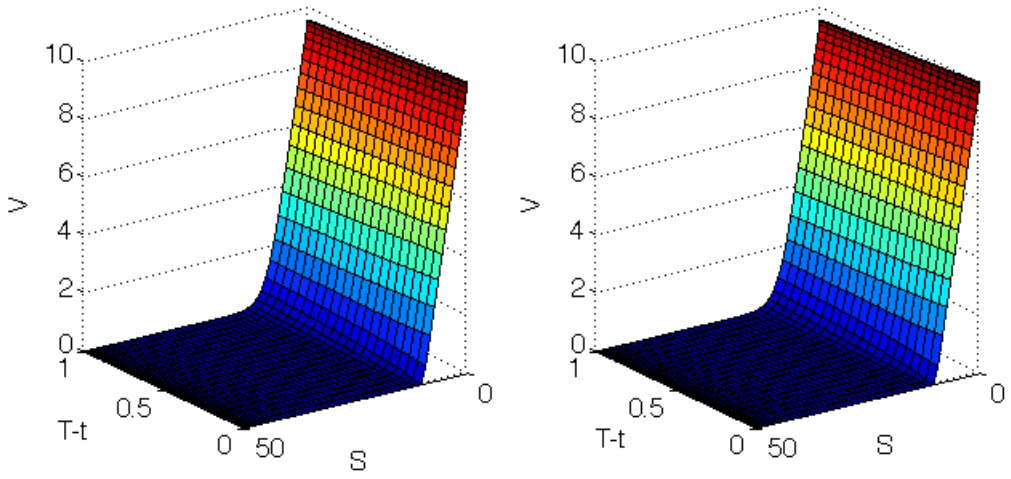


Figure 4.9: Solution for equation (2.6) when polynomial collocation is used, $N = 60$ and $h = 10^{-3}$. Left: the numerical solution. Right: the exact solution.

We are still using $N = 60$, $h = 10^{-3}$ and the values in table 4.1 in the calculations. The solution for V when we have used polynomial collocation to solve equation (2.6) is given in figure 4.9. Again, we see that the solution looks the same as the exact solution and it also looks equal to the solution in figure 4.1. This is good, as the two method should produce the same result.

We use equation (3.3) to estimate the GLL-error for the convergence plots. We look at both the absolute error and the relative error as earlier. The error plots are given in figure 4.10, with corresponding values in tables 4.7 and 4.8. We see that the convergence in both space and time behaves in the same way as in figure 4.2. In particular, the error values in table 4.3 is almost identical and when $N \geq 12$, the values in tables 4.2 and 4.7 are similar as well. We can therfore assume that we will get the same improved convergence result if we try Crank-Nicolson as time discretization here as well. This is what we see in figure 4.11 and table 4.9.

We study the difference between the exact solution and the numerical solution for this method as well, which is shown in figure 4.12. We have the same error trend here as for the spectral Galerkin method, where the error is largest close to $K = S$ for t close to T and decreasing with time.

N	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
2	0.06642501	0	0.01209443
3	0.05027341	0	0.01294259
4	0.08774829	-0.40164460	0.02757097
6	0.07104010	-0.49883813	0.02050798
8	0.07641986	0.19942344	0.02252985
12	0.07055524	0.00988049	0.02070149
16	0.03548788	1.10662109	0.01040658
24	0.03307573	1.09298009	0.00969869
32	0.00744941	2.25212869	0.00218437
48	0.00564706	2.55020040	0.00165587
64	0.00029235	4.67136395	0.00008572
96	0.00016792	5.07163657	0.00004924
128	0.00036710	-0.32849581	0.00010764
192	0.00024370	-0.53733280	0.00007146
256	0.00013606	1.43192556	0.00003990
384	0.00004109	2.56821455	0.00001205
512	0.00010336	0.39662572	0.00003031
768	0.00004097	0.00410489	0.00001201

Table 4.7: Values for absolute error and relative error and the corresponding values of α. The time step is kept fixed at $h = 10^{-3}$.

h	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
1/2	0.01336123	0	0.00383225
1/4	0.00797038	0.74533250	0.00230652
1/8	0.00430937	0.88717265	0.00125446
1/16	0.00251587	0.77641923	0.00073491
1/32	0.00209786	0.26214009	0.00061396
1/64	0.00218333	-0.05761517	0.00063960
1/128	0.00230743	-0.07975490	0.00067630

Table 4.8: Values for absolute error and relative error and the corresponding values of α when $N = 60$.

h	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
1/2	0.00555973	0	0.00159463
1/4	0.00291599	0.93102904	0.00084385
1/8	0.00262636	0.15092316	0.00076453
1/16	0.00254022	0.04811137	0.00074202
1/32	0.00250686	0.01906969	0.00073366
1/64	0.00249247	0.00830387	0.00073017
1/128	0.00248585	0.00383678	0.00072859

Table 4.9: Values for absolute error and relative error and the corresponding values of α when $N = 60$ and Crank-Nicolson is used as time discretization.

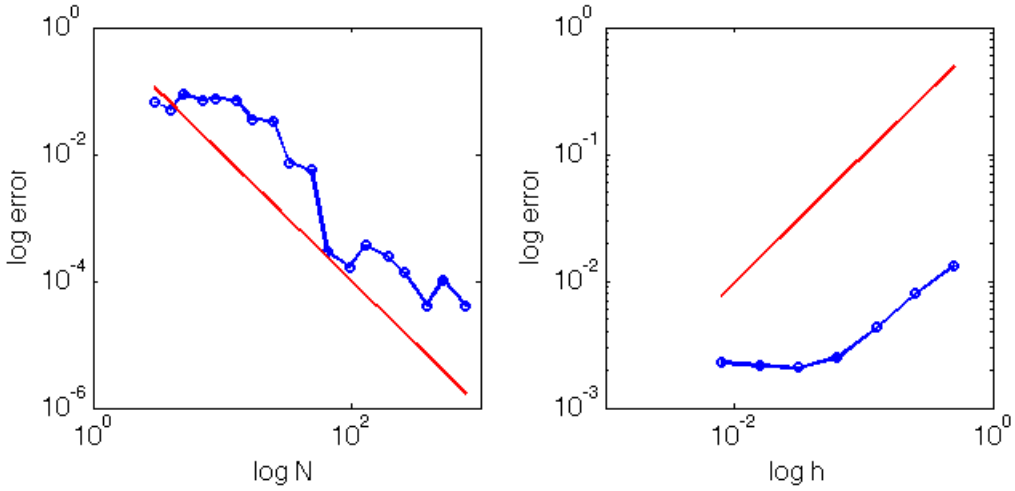


Figure 4.10: Convergence for (2.6). Left: Convergence in space. Log-log plot of the GLL-norm plotted against N when $h = 10^{-3}$. The red line is $1/N^2$ plotted against N. Right: Convergence in time. Log-log plot of the GLL-error plotted against h when $N = 60$. The red line is h against h.

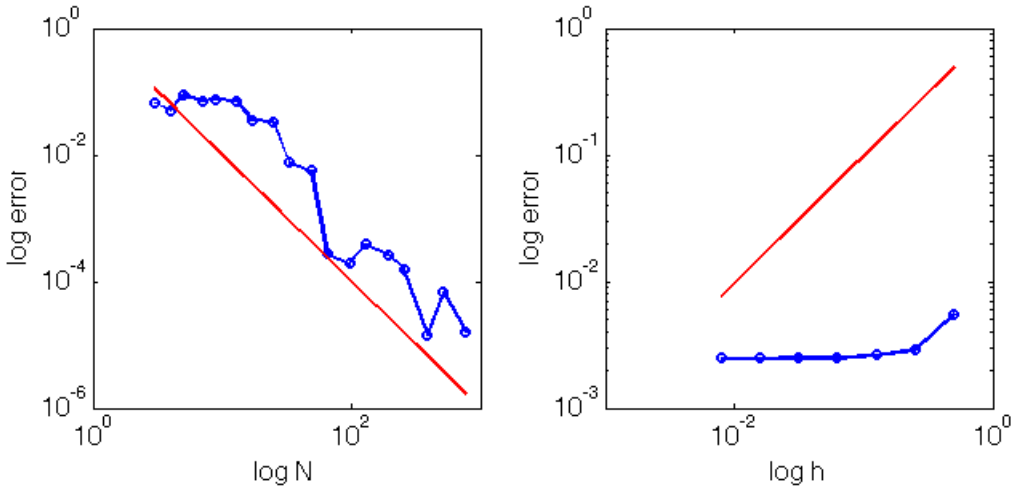


Figure 4.11: Convergence for (2.6) when Crank-Nicolson is used as time discretization. Left: Convergence in space. Log-log plot of the GLL-norm plotted against N when $h = 10^{-3}$. The red line is $1/N^2$ plotted against N. Right: Convergence in time. Log-log plot of the GLL-error plotted against h when $N = 60$. The red line is h against h

Transformed equation

The transformed solution is given in figure 4.13. Again, we observe that the solution looks like the other solutions. The convergence plots in figure 4.14 shows the same trend as in figure 4.7. The left side of figure 4.14 shows the convergence in space is approximately quadratic in the beginning before it becomes constant, and the right side of the figure shows that the convergence in time is linear in the beginning, and then becomes constant. The convergence values in tables 4.10 and 4.11 is identical to the values in table 4.5 and 4.6, which confirms that the two methods are equivalent.

Figure 4.15 shows the absolute difference between the exact and numerical solution.

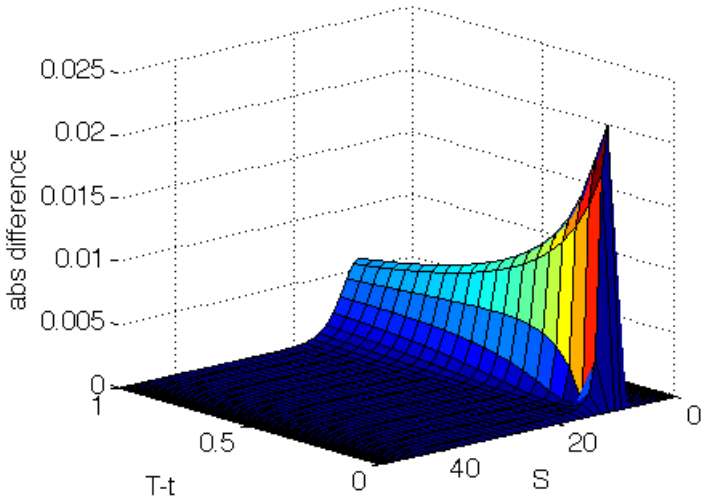


Figure 4.12: Difference between the exact solution and the numerical solution

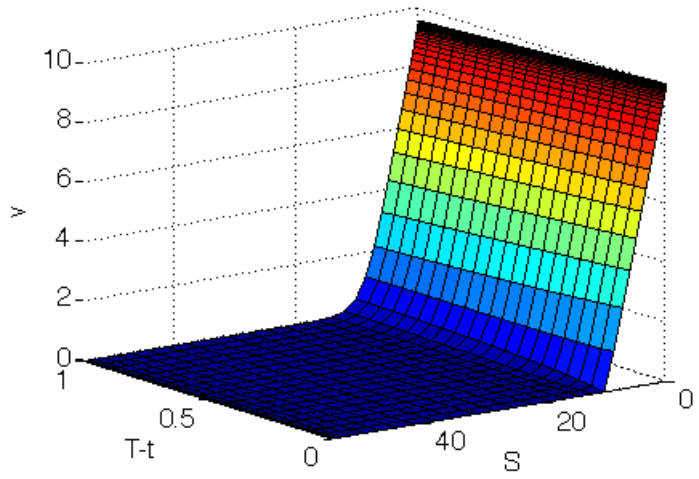


Figure 4.13: Numerical solution for equation (2.11) when polynomial collocation is used, $N = 60$ and $h = 10^{-3}$.

The difference is similar to the result in figure 4.8, a decreasing error due to the initial condition and a constant error from the boundary condition at $x = x_{\min}$.

Again, we see that the original equation gives better convergence results and better solution than the transformed equation, confirming that the transformation is not necessary to solve the equations numerically.

4.2.4 Analysis of the method

In this section, we will show that the collocation method is mathematically equivalent to the spectral Galerkin method that we used. This implies that the analysis is the same as in section 4.1.4, and no further analysis is required. We will show that the two methods are equivalent at each time t. If we use the same time discretization for the two methods, they should be equivalent. The numerical results obtained confirms

N	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
2	0.01697580	0	0.00150831
3	0.21477835	0	0.02165946
4	0.06293417	-1.89036352	0.00602614
6	0.14856888	0.53171665	0.01440888
8	0.05674811	0.14927103	0.00545679
12	0.10092918	0.55778863	0.00972871
16	0.03479542	0.70567512	0.00335268
24	0.01707888	2.56305793	0.00164555
32	0.01757554	0.98532797	0.00169340
48	0.01010895	0.75658020	0.00097400
64	0.00567725	1.63030467	0.00054700
96	0.00310244	1.70415895	0.00029892
128	0.00324821	0.80554765	0.00031296
192	0.00297616	0.05994806	0.00028675
256	0.00295461	0.13667779	0.00028468

Table 4.10: Values for absolute error and relative error and the corresponding values of α. The time step is kept fixed at $h = 10^{-3}$.

h	$\|V^{BS} - V\|_{GLL}$	α	$\frac{\|V^{BS}-V\|_{GLL}}{\|V^{BS}\|_{GLL}}$
1/2	0.02424988	0	0.00229451
1/4	0.01480968	0.71143664	0.00141152
1/8	0.00790935	0.90490974	0.00075751
1/16	0.00402587	0.97425983	0.00038667
1/32	0.00270893	0.57157665	0.00026059
1/64	0.00276022	-0.02706003	0.00026574
1/128	0.00301833	-0.12897001	0.00029071

Table 4.11: Values for absolute error and relative error and the corresponding values of α when $N = 60$.

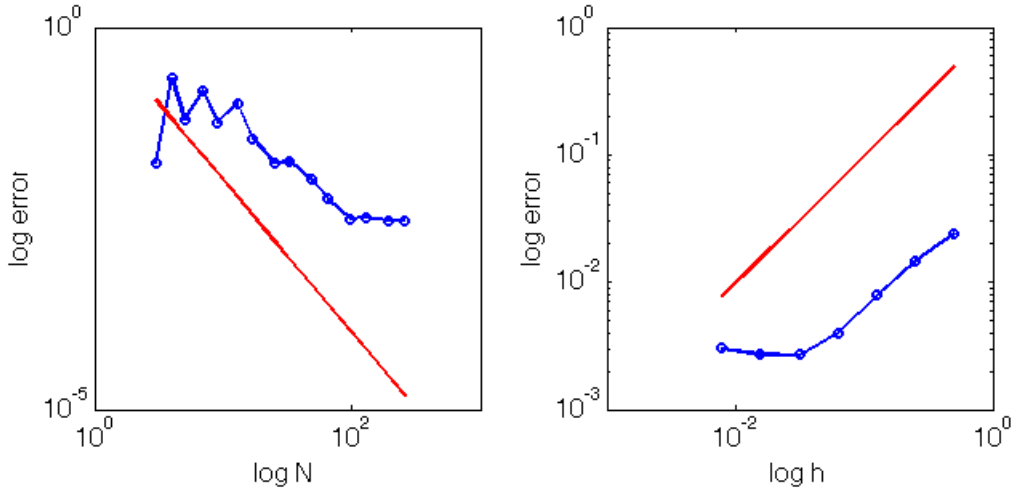


Figure 4.14: Convergence for (2.11). Left: Convergence in space. Log-log plot of the GLL-norm plotted against N when $h = 10^{-3}$. The red line is $1/N^2$ plotted against N. Right: Convergence in time. Log-log plot of the GLL-error plotted against h when $N = 60$. The red line is h against h.

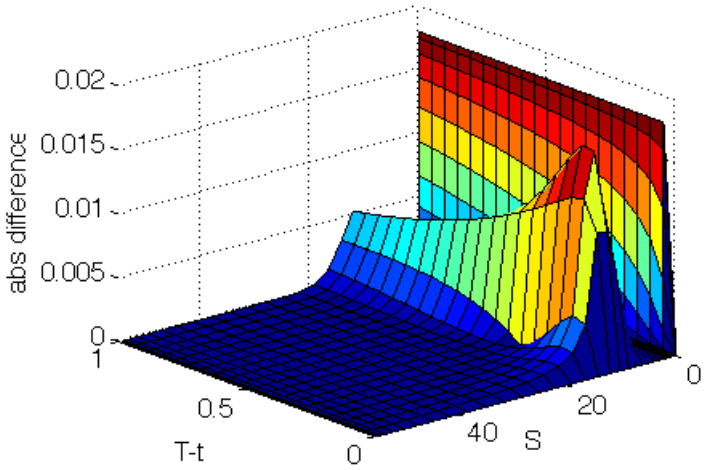


Figure 4.15: Difference between the exact solution and the numerical solution for the transformed equation

this. The proof can be found in [12, Chapter 10.3].

We start by defining the operator Π^{GLL} as

$$\Pi^{GLL}V(S) = \sum_{i=1}^{N-1} V(S_i)\ell_i(S),$$

and the discrete inner product

$$(f, g)_N = \sum_{i=1}^{N-1} \rho_i f(S_i)g(S_i),$$

where ρ_i is the GLL weight corresponding to node i. This is the GLL quadrature. Remeber from section 3.1 that the quadrature is exact when $fg \in \mathbb{P}_{2N-1}$, and thus

$(f,g)_N = (f,g)$. In addition, we have

$$\begin{aligned}(\Pi^{GLL} f, g)_N &= \sum_{i=1}^{N-1} \rho_i \Pi^{GLL} f(S_i) g(S_i) \\ &= \sum_{i=1}^{N-1} \rho_i \sum_{j=1}^{N-1} f(S_i) g(S_i) \ell_j(S_i) \\ &= \sum_{i=1}^{N-1} \rho_i f(S_i) g(S_i) = (f,g)_N.\end{aligned}$$

Now, we write the bilinear form (2.14) as the first step in the derivation in section 2.1.6:

$$\frac{1}{2}\sigma^2(\Pi^{GLL} S^2 V_{SS}, \varphi) + r(\Pi^{GLL} S V_S, \varphi) - r(\Pi^{GLL} V, \varphi).$$

By including the time derivative, we obtain the weak formulation

$$\text{Find } V \in X_N \text{ such that } (\Pi^{GLL} V_t, \varphi_i) + (\mathcal{L}V, \varphi_i)_N = 0 \quad i = 1, 2, \ldots, N-1,$$

where we have defined

$$\mathcal{L}V = \frac{1}{2}\sigma^2 \Pi^{GLL} S^2 V_{SS} + r\Pi^{GLL} S V_S - r\Pi^{GLL} V.$$

Looking at the i-th equation, we get for each of the three terms in $\mathcal{L}V$

$$\begin{aligned}&\frac{1}{2}\sigma^2(\Pi^{GLL} S^2 V_{SS}, \varphi_i)_N = \frac{1}{2}\sigma^2 \rho_i \Pi^{GLL} S_i^2 V_{SS}(S_i), \\ &r(\Pi^{GLL} S V_S, \varphi_i)_N = r\rho_i \Pi^{GLL} S_i V_S(S_i) \text{ and} \\ &r(\Pi^{GLL} V, \varphi_i)_N = r\rho_i \Pi^{GLL} V(S_i),\end{aligned}$$

and the time derivative

$$(\Pi^{GLL} V_t, \varphi_i)_N = \rho_i \Pi^{GLL} \dot{V}(S_i)$$

Dividing by ρ_i and collecting the terms, we get

$$\Pi^{GLL} \dot{V}(S_i) + \frac{1}{2}\sigma^2 \Pi^{GLL} S_i^2 V_{SS}(S_i) + r\Pi^{GLL} S_i V_S(S_i) + r\Pi^{GLL} V$$

This is equal to

$$\sum_{j=1}^{N-1} \dot{V}_i \ell_j(S_i) + \frac{1}{2}\sigma^2 \sum_{j=1}^{N-1} S_i^2 V_j \ell_j''(S_i) + r \sum_{j=1}^{N-1} S_i V_i \ell_j'(S_i) + r \sum_{j=1}^{N-1} V_i \ell_j(S_i),$$

where we have chosen the basis functions φ_i to be the Lagrange polynomials ℓ_i. This expression is the same as the approximation in section 4.2.2. Therefore, the analysis of stability and coercivity in section 4.1.4 holds for the collocation case as well.

Generally, we can not say that the two methods are equivalent. It holds in our case because both boundary conditions are of Dirichlet type. If we had other types of boundary conditions, both i, j would be $0 \leq i, j \leq N$. The boundary conditions would be treated differently for the two methods, making the two methods different at the boundaries.

4.3 Solution using the finite element method

In order to have something to compare the results with, we solve Black-Scholes equation using the finite element method as well. The finite element method is similar to the spectral Galerkin method, but the basis functions are no longer global. In stead, we use local basis functions at each element.

We do not give all the details of the derivation of this method, but we will point out some similarities and differences from the spectral Galerkin method.

For the finite element method, we use a grid with equally spaced nodes on the domain. An element is then defined as the interval between two neighbouring points on the grid: $[S_j, S_{j+1}]$. All elements will be of the same size $h_N = S_{j+1} - S_j$. The calculations are performed by looping over all the elements.

As for the spectral Galerkin method, the calculations are performed on a reference domain. Here, this means that the calculations are perfomed on a reference element, and we need a transformation from an arbitrary element to the reference element. We will use the element $\xi \in [-1, 1]$ as reference element. The transformation from an element $[S_j, S_{j+1}]$ is then given by

$$S(\xi) = \frac{S_{j+1} + S_j}{2} + \frac{(S_{j+1} - S_j)\xi}{2}, \qquad \frac{dS}{d\xi} = \frac{S_{j+1} - S_j}{2}.$$

The mapping is performed elementwise when we loop over the elements. We define the basis functions on the reference element as

$$\hat{\varphi}_1(\xi) = \frac{1-\xi}{2} \quad \text{and} \quad \hat{\varphi}_2(\xi) = \frac{1+\xi}{2}.$$

As earlier, we write the approximate solution $V_N(S)$ as

$$V_N(S) = \sum_{j=1}^{N-1} V_j \varphi_j(S).$$

We choose $\varphi = \hat{\varphi_{1,2}}$ and get an expression for each element. From [1, chapter 4.3] we have expressions for the matrix elements in the stiffness matrix A

$$A_{ij} = \begin{cases} -S_i^2\sigma^2/2h_N + rS_i/2 & \text{if } j = i-1, \quad i = 2, \ldots, N+1 \\ S_i^2\sigma^2/h_N + rh_N & \text{if } i = j, \quad i = 2, \ldots, N+1 \\ -S_i^2\sigma^2/2h_N - rS_i/2 & \text{if } j = i+1, \quad i = 1, \ldots, N \\ rh_N/2 & \text{if } i = j = 1 \\ 0 & \text{otherwise.} \end{cases}$$

and the mass matrix M

$$M_{ij} = \begin{cases} h/6 & \text{if } j = i-1, \quad i = 2, \ldots, N+1 \\ 2h/3 & \text{if } i = j, \quad i = 2, \ldots, N+1 \\ h/6 & \text{if } j = i+1, \quad i = 1, \ldots, N \\ h/3 & \text{if } i = j = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Again, we solve the system using implict Euler. The numerical solution when $N = 60$ and $h = 10^{-3}$ is shown in figure 4.16. Again, we get the same solution as for the other methods.

To study the convergence in the GLL-norm, we must interpolate the solution from the equidistant grid to a GLL-grid. This is done by using built-in functions in Matlab. The numerical convergence is displayed in figure 4.17. We see that the convergence is slower than for both the spectral methods, the convergence in space is almost linear and the convergence in time is linear at first, than grows a little before it becomes constant.

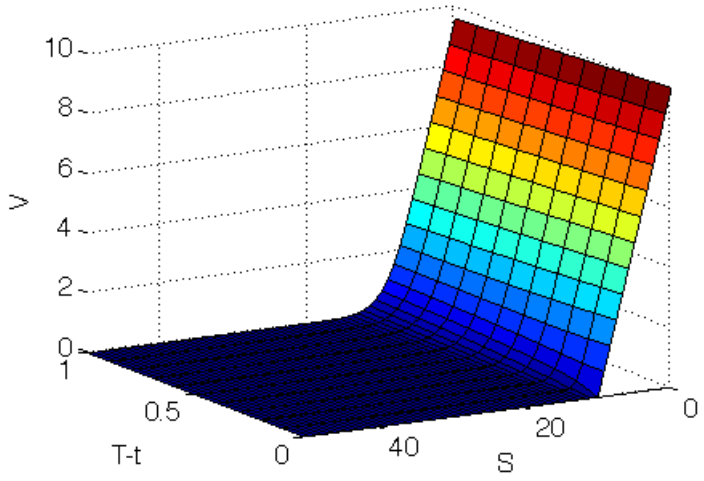


Figure 4.16: Solution of (2.6) by the finite element method

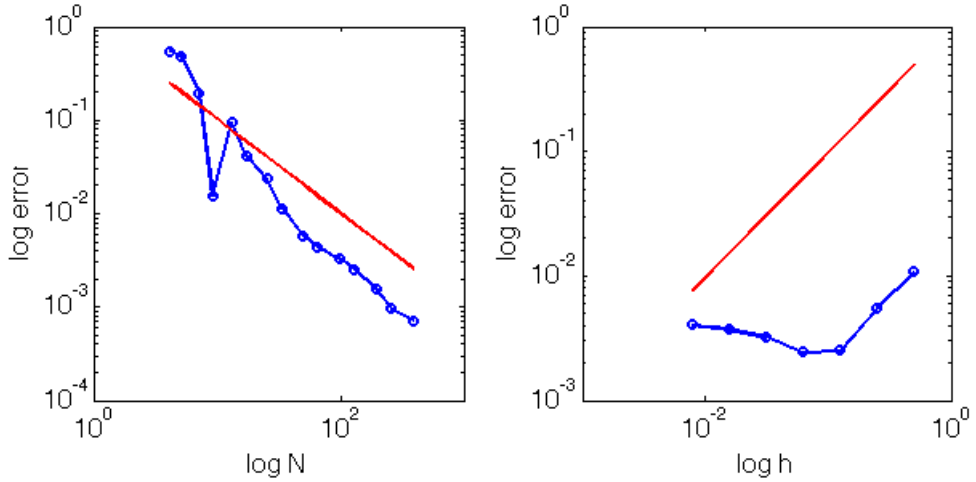


Figure 4.17: Convergence for (2.6) when the finite element method is used, $N = 60$ and $h = 10^{-3}$. The blue line is the GLL-norm. Left: Convergence in space. Log-log plot of the GLL-error plotted against N. The red line is $1/N$ plotted against N. Right: Convergence in time. Log-log plot of the GLL-error plotted against h. The red line is h against h.

4.4 Timing

We would like to know something about how fast we can reach a certain tolerance to compare the two spectral methods and the finite element solver. We will only look at

	$\|V^{BS} - V\|_{GLL}$			
Method	10^{-1}	10^{-2}	10^{-3}	10^{-4}
Galerkin	$N = 9$	$N = 25$	$N = 47$	$N = 108$
Collocation	$N = 5$	$N = 25$	$N = 47$	$N = 108$
FEM	$N = 9$	$N = 38$	$N = 240$	$N = 2112$

Table 4.12: Number of nodes N needed to obtain the desired accuracy for the different methods when h is kept fixed at 10^{-3}.

	$\|V^{BS} - V\|_{GLL}$			
Method	10^{-1}	10^{-2}	10^{-3}	10^{-4}
Galerkin	$h = 1/5$	$h = 1/5$	$h = 1/46$	$h = 1/698$
Collocation	$h = 1/5$	$h = 1/5$	$h = 1/46$	$h = 1/666$
FEM	$h = 1/5$	$h = 1/5$	–	–

Table 4.13: Size of time step h needed to obtain the desired accuracy for the different methods when N is kept fixed at 108.

the solutions of (2.6), and not the transformed equation because the solution has slow convergence compared to the solution of the original equation.

To obtain the timing results, we first check which values of N and h we must have to reach a certain accuracy. We keep h fixed at $1/1000$ and vary N. The results are given in table 4.12, where we give the values of N that we need to obtain an error measured in the GLL-norm of size 10^{-1}, 10^{-2}, 10^{-3} and 10^{-4}. All calculations are performed in the absolute norm. We observe that the values for spectral Galerkin and polynomial collocation are approximately equal in every case. This is explained by the fact that the two methods are mathematically equivalent in our case, as shown in section 4.2.4.

From table 4.12 we observed that by choosing $N = 108$, we can obtain good accuracy. Therefore, we use this value to check which values of h that we need. This is shown in table 4.13. Again, the values for the two spectral methods are almost equal.

We also included the corresponding values obtained when we use the the finite element method to solve the equation. Figure 4.17 shows that the convergence of the finite element method is much slower than for both the spectral methods that we used. This is also reflected in table 4.12 and 4.13, as we found that we need a lot more nodes for the finite element method. With $N = 108$, we are not able to reach $\|V^{BS} - V\|_{GLL} < 10^{-3}$ as we can for the two spectral methods.

In addition, we want to check if the number of nodes in table 4.12 can be improved by choosing a smaller time step h. The number of nodes needed is shown in table 4.14. We see that the values of N needed to reach the different accuracies are the same as before, but the computation time will be about 10 times larger than when $h = 10^{-3}$. As figure 4.18 illustrates, the placement of the nodes (which depends on N) has a large impact on the error. The placement of the nodes that yield high accuracies are still good when we decrease h, therefore it is sufficient to use $h = 10^{-3}$ to find the values of N which give good accuracy of the solution.

We can use the values from table 4.12 and 4.13 to test the computation time needed to reach the desired accuracy. For each of the combinations in the two tables, we run the program 10 times, and calculate the average of the 10 runs. The results are given in tables 4.15 and 4.16. We observe that the fastest way to obtain the desired accuracy is by finding a good value of N, then keep it fixed and change h. We also observe

	$\|V^{BS} - V\|_{GLL}$			
Method	10^{-1}	10^{-2}	10^{-3}	10^{-4}
Galerkin	$N = 9$	$N = 25$	$N = 47$	$N = 108$
Collocation	$N = 5$	$N = 25$	$N = 47$	$N = 108$

Table 4.14: Number of nodes N needed to obtain the desired accuracy for the different methods when h is kept fixed at 10^{-4}.

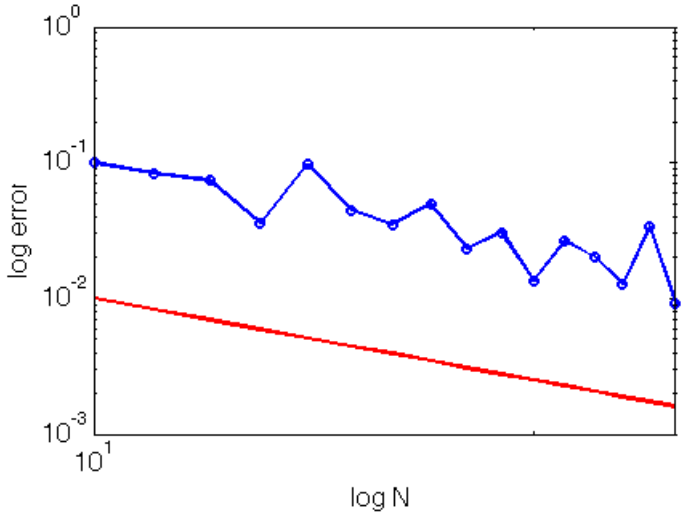


Figure 4.18: Convergence for (2.6) when spectral Galerkin is used. The values of N is in the range $[10, 25]$ and $h = 10^{-3}$.

that polynomial collocation is a bit faster than the spectral Galerkin method. This is because there is an extra for-loop in the stiffness matrix coming from the GLL quadrature in the spectral Galerkin method. The polynomial collocation method does not have this for-loop, and it is therefore a bit faster. As expected, the finite element method is significantly slower than the two spectral methods when we are interested in errors of size 10^{-3}.

The results show that the two spectral methods are faster than the finite element method in terms of reaching a desired accuracy, even though the inital data is not smooth.

	$\|V^{BS} - V\|_{GLL}$			
Method	10^{-1}	10^{-2}	10^{-3}	10^{-4}
Galerkin	0.3601 s	0.9243 s	1.7129 s	4.3014 s
Collocation	0.1869 s	0.910 s	1.6757 s	4.0489 s
FEM	0.3548 s	1.3502 s	8.8078 s	85.2170 s

Table 4.15: The time needed to calculate the solution when $h = 10^{-3}$ and we use values for N in table 4.12

	$\|V^{BS} - V\|_{GLL}$			
Method	10^{-1}	10^{-2}	10^{-3}	10^{-4}
Galerkin	0.3025 s	0.3025 s	0.4674 s	3.0896 s
Collocation	0.0465 s	0.0465 s	0.2109 s	2.7217 s

Table 4.16: The time needed to calculate the solution when $N = 60$ and we use values for h in table 4.13

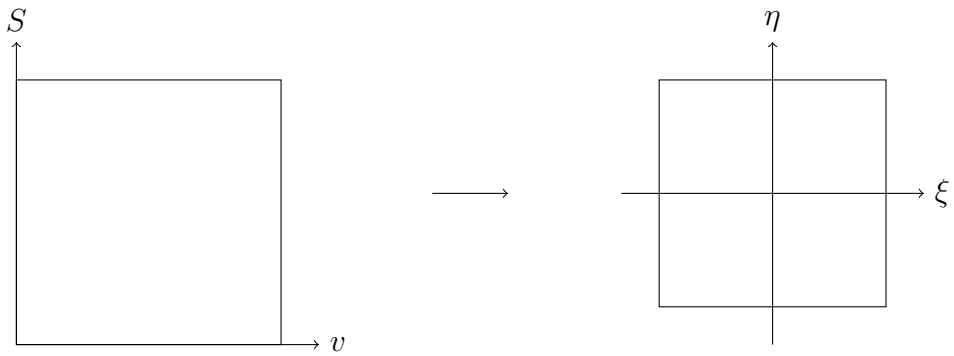
Chapter 5

Solving the Heston model

5.1 The spectral Galerkin method

5.1.1 Mapping to reference domain

The domain of this problem is $v \in [0, 0.75]$, $S \in [0, 50]$. The reference domain $\hat{\Omega}$ is the two dimensional version of the reference domain in chapter 3, namely $\xi, \eta \in [-1, 1]$. The domains are displayed below.



This gives us the following mappings to the reference domain

$$v(\xi) = \frac{L_v}{2}(\xi + 1) \quad \text{and} \quad S(\eta) = \frac{L_S}{2}(\eta + 1)$$

with derivatives

$$\frac{dv}{d\xi} = \frac{L_v}{2} = J_v \quad \text{and} \quad \frac{dS}{d\eta} = \frac{L_S}{2} = J_S.$$

The partial derivatives in equation (2.23) are

$$\frac{\partial U}{\partial v} = \frac{\partial \hat{U}}{\partial \xi}\frac{d\xi}{dv} = \frac{1}{J_v}\frac{\partial \hat{U}}{\partial \xi}, \quad \frac{\partial \varphi}{\partial v} = \frac{\partial \hat{\varphi}}{\partial \xi}\frac{d\xi}{dv} = \frac{1}{J_v}\frac{\partial \hat{\varphi}}{\partial \xi}$$

and so on. Now, we can write equation (2.23) as

$$U_t + \frac{1}{2J_v^2}\sigma^2 v U_{\xi\xi} + \frac{1}{J_v J_S}\rho\sigma v S U_{\xi\eta} + \frac{1}{2J_S^2} v S^2 U_{\eta\eta} + (\kappa(\theta - v) - \lambda v)U_\xi + rSU_\eta - rU = 0. \quad (5.1)$$

Using the same method as in section 2.2.5, we obtain the bilinear form of this equation

$$\begin{aligned}
&\int_{\hat{\Omega}} J_v J_S U_t \varphi d\xi d\eta - \frac{1}{2}\int_{\hat{\Omega}} \sigma^2 v \frac{J_S}{J_v} U_\xi \varphi_\xi d\xi d\eta - \frac{1}{2}\int_{\hat{\Omega}} \rho\sigma v S U_\eta \varphi_\xi d\xi d\eta \\
&- \frac{1}{2}\int_{\hat{\Omega}} \rho\sigma v S U_\xi \varphi_\eta d\xi d\eta - \frac{1}{2}\int_{\hat{\Omega}} v S^2 \frac{J_v}{J_S} U_\eta \varphi_\eta d\xi d\eta \\
&+ \int_{\hat{\Omega}} \left((\kappa\theta - \frac{1}{2}\sigma^2) - (\kappa + \lambda + \frac{1}{2}\rho\sigma) v \right) J_S U_\xi \varphi d\xi d\eta + \int_{\hat{\Omega}} (r - v - \frac{1}{2}\rho\sigma) S J_v U_\eta \varphi d\xi d\eta \\
&- \int_{\hat{\Omega}} r J_v J_S U \varphi d\xi d\eta = 0.
\end{aligned} \tag{5.2}$$

This equation will be used in the next section to find the spectral approximation.

5.1.2 Spectral approximation

We start by writing the approximate solution $U_N(\xi, \eta)$ as

$$U_N(\xi, \eta) = \sum_{m=0}^{N}\sum_{n=0}^{N} U_{mn} \ell_m(\xi) \ell_n(\eta),$$

where $\ell_i(\xi)$ represents the one-dimensional Lagrangian basis function and U_{ij} is a two dimensional nodal basis. This is a tensor-product extension of the one-dimensional basis we used to solve Black-Scholes equation. We choose the Lagrangian polynomials as test functions as well, giving $\varphi = \ell_i(\xi)\ell_j(\eta)$.

Using this basis, we can write the first integral in (5.2) as

$$\begin{aligned}
\int_{\hat{\Omega}} J_v J_S U_t \varphi d\xi d\eta &= \int_{\hat{\Omega}} J_v J_S \left(\sum_{m=0}^{N}\sum_{n=0}^{N} \dot{U}_{mn} \ell_m(\xi)\ell_n(\eta) \right) \ell_i(\xi)\ell_j(\eta) \\
&= J_v J_S \sum_{m=0}^{N}\sum_{n=0}^{N} \dot{U}_{mn} \left(\int_{-1}^{1} \ell_i(\xi)\ell_m(\xi) d\xi \right) \left(\int_{-1}^{1} \ell_j(\eta)\ell_n(\eta) d\eta \right) \\
&= J_v J_S \sum_{m=0}^{N}\sum_{n=0}^{N} \dot{U}_{mn} (\ell_i(\xi), \ell_m(\xi))^1 \cdot (\ell_j(\eta), \ell_n(\eta))^1 \\
&\approx J_v J_S \sum_{m=0}^{N}\sum_{n=0}^{N} \dot{U}_{mn} (\ell_i(\xi), \ell_m(\xi))^1_N \cdot (\ell_j(\eta), \ell_n(\eta))^1_N \\
&= J_v J_S \sum_{m=0}^{N}\sum_{n=0}^{N} \dot{U}_{mn} M^1_{im} M^1_{jn}, \quad 0 \le i, j \le N,
\end{aligned}$$

where $(\cdot,\cdot)^1_N$ represents an one-dimensional inner product evaluated by GLL quadrature and M^1_{im} is the one dimensional mass matrix. Using the same approach for the other integrals gives

$$\frac{1}{2}\int_{\hat{\Omega}} \sigma^2 v \frac{J_S}{J_v} U_\xi \varphi_\xi d\xi d\eta \approx \frac{1}{2}\sigma^2 \frac{J_S}{J_v} \sum_{m=0}^{N}\sum_{n=0}^{N} v_n U_{mn} A^1_{im} M^1_{jn},$$

$$\frac{1}{2}\int_{\hat{\Omega}} \rho\sigma v S U_\eta \varphi_\xi d\xi d\eta \approx \frac{1}{2}\rho\sigma \sum_{m=0}^{N}\sum_{n=0}^{N} v_n S_m U_{mn} D^1_{im} C^1_{jn},$$

$$\frac{1}{2}\int_{\hat{\Omega}} \rho\sigma v S U_\xi \varphi_\eta d\xi d\eta \approx \frac{1}{2}\rho\sigma \sum_{m=0}^{N}\sum_{n=0}^{N} v_n S_m U_{mn} C^1_{im} D^1_{jn},$$

$$\frac{1}{2}\int_{\hat{\Omega}} vS^2\frac{J_v}{J_S}U_\eta\varphi_\eta d\xi d\eta \approx \frac{1}{2}\frac{J_v}{J_S}\sum_{m=0}^{N}\sum_{n=0}^{N} v_n S_m^2 U_{mn} M_{im}^1 A_{jn}^1,$$

$$\begin{aligned}&\int_{\hat{\Omega}}\left((\kappa\theta-\frac{1}{2}\sigma^2)-(\kappa+\lambda+\frac{1}{2}\rho\sigma)v\right)J_S U_\xi\varphi d\xi d\eta\\ &\approx J_S\sum_{m=0}^{N}\sum_{n=0}^{N}\left((\kappa\theta-\frac{1}{2}\sigma^2)-(\kappa+\lambda+\frac{1}{2}\rho\sigma)v_n\right)U_{mn}C_{im}^1 M_{jn}^1,\end{aligned}$$

$$\int_{\hat{\Omega}}(r-v-\frac{1}{2}\rho\sigma)SJ_v U_\eta\varphi d\xi d\eta \approx J_v\sum_{m=0}^{N}\sum_{n=0}^{N}(r-v-\frac{1}{2}\rho\sigma)S_m U_{mn}M_{im}^1 C_{jn}^1$$

and

$$\int_{\hat{\Omega}} rJ_vJ_S U\varphi d\xi d\eta \approx J_vJ_S r\sum_{m=0}^{N}\sum_{n=0}^{N}U_{mn}M_{im}^1 M_{jn}^1.$$

The matrix elements A_{im}^1, C_{im}^1 and D_{im}^1 are the corresponding elements in the matrices coming from the GLL evaluation of the one-dimensional inner products $(\ell_i',\ell_m')_N^1$, $(\ell_i,\ell_m')_N^1$ and $(\ell_i',\ell_m)_N^1$, respectively.

As usual, we combine the evaluation of the integrals into two matrices, the mass matrix $\mathbf{M}$ the stiffness matrix $\mathbf{A}$. To include the boundary condition at $S=0$, we change the corresponding matrix elements. This gives a system of equations with $(N+1)^2$ unknows, which we will write as

$$\begin{bmatrix}0 & \cdots & & & 0\\ & \ddots & & & \\ & & 0 & & \\ & & & M_{N+2,N+2} & \\ \vdots & & & \ddots & \vdots\\ & & & & \\ 0 & \cdots & & & M_{(N+1)^2,(N+1)^2}\end{bmatrix}\begin{bmatrix}\dot{U}_{0,0}\\ \vdots\\ \dot{U}_{0,N+1}\\ \vdots\\ \dot{U}_{m,n}\\ \vdots\\ \dot{U}_{N+1,N+1}\end{bmatrix}$$

$$+\begin{bmatrix}1 & 0 & & \cdots & & 0\\ 0 & \ddots & 0 & \ddots & & 0\\ 0 & & 1 & 0 & \cdots & 0\\ A_{N+2,0} & A_{N+2,1} & & \cdots & & A_{N+2,(N+1)^2}\\ \vdots & & \cdots & & & \vdots\\ & & & & & \\ A_{(N+1)^2,0} & A_{(N+1)^2,1} & & \cdots & & A_{(N+1)^2,(N+1)^2}\end{bmatrix}\begin{bmatrix}U_{0,0}\\ \vdots\\ U_{0,N+1}\\ \vdots\\ U_{m,n}\\ \vdots\\ U_{N+1,N+1}\end{bmatrix}=\begin{bmatrix}F_{0,0}\\ \vdots\\ F_{0,N+1}\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix},$$

where $F_{0,n}=e^{-r(T-t)}$. This system can then be solved using an appropriate time discretization. We have used Crank-Nicolson.

5.1.3 Numerical results

Figure 5.1 shows the numerical solution obtained when we solve the system of equations given in the previous section, and the exact solution is given in figure 5.2. The values used in the calculations are given in table 5.1, and we have used $N=60$ in both directions and $h=10^{-3}$. We see that the numerical solution does not look like the

Constant	Value
T	1
K	10
σ	0.5
ρ	0.1
λ	0
κ	2.5
θ	0.06
r	0.05

Table 5.1: Values of constants used in the calculatuions

exact solution. The difference between the numerical and exact solution is shown in figure 5.3. We observe that the error is smallest close to $v = v_{\min}$, but it is not particularly small here either. From figure 5.1, it can look like the sign of the diffusion term or the convection term might be wrong, but this is not the case. We have also tried using smaller time steps without success.

There are some possible reasons for why we do not get the correct answer. The most obvious explanation is that the approximation in the previous section is wrong, but we can't see that this is the case. An other option is that the equation is not solvable by the spectral Galerkin method, or at least not without using any special tricks. In [19], Kopriva and Zhu solves the Heston model by using the spectral element method. By using the spectral element method one might avoid the problems with the inital data by using $S = K$ as a boundary between elements. This is exactly what is done in [19], and might be necessary to obtain correct solutions.

We have not studied the convergence of the numerical method in this case because the solution converges to the wrong solution.

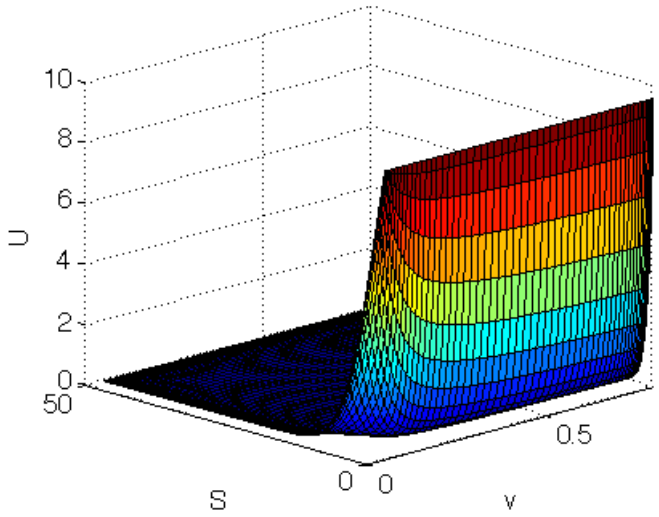


Figure 5.1: Numerical solution for equation (2.23) when $N = 60$ and $h = 10^{-3}$.

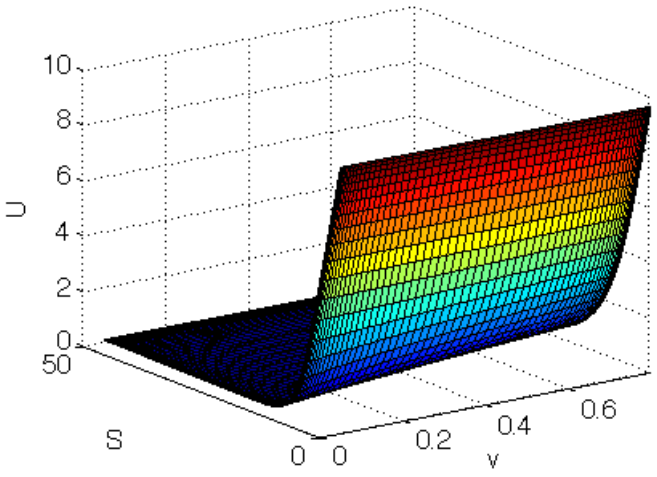


Figure 5.2: Exact solution for equation (2.23) when $N = 60$ and $h = 10^{-3}$.

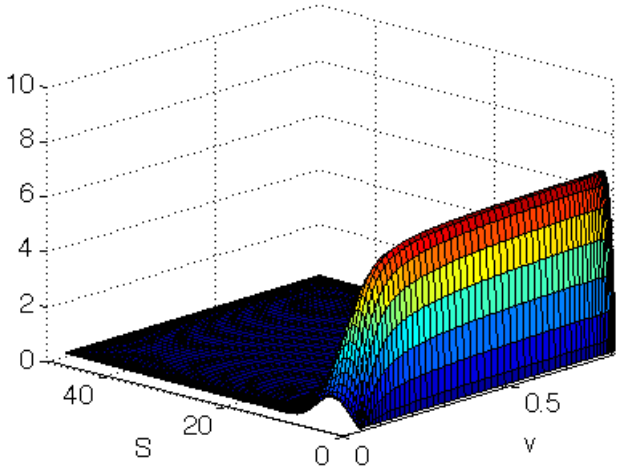


Figure 5.3: Difference between the exact and numerical solution when $N = 60$ and $h = 10^{-3}$.

5.1.4 Analysis of the method

The analysis of this problem is not straight forward. We can either follow the analysis by Winkler et al in [16], but this is not valid for the continuity when $v_{min} = 0$. The next option is to follow the analysis by Achdou and Pironneau in [1], but this analysis is for the Stein-Stein stochastic volatility model. The third option is to follow Daskalopoulos in [6], but the analysis in this article is very advanced. Understanding all parts of the analysis of Daskalopoulos is out of scope of this thesis, therefore we will follow Winkler in stead.

We need some definitions before we can start the analysis. Winkler uses the time discretized version of the bilinear form. We use the same time discretization as we have done throughout the thesis, so we can write

$$\int_\Omega \frac{U^{k+1} - U^k}{h}\varphi + a(U^{k+1}, \varphi) = 0,$$

which equals

$$\int_\Omega \frac{U^{k+1}}{h}\varphi + a(U^{k+1},\varphi) = \int_\Omega \frac{U^k}{h}\varphi.$$

Next, we state the following lemma.

Lemma 6.

$$\int_\Omega (f\cdot\nabla U)\varphi = \frac{1}{2}\int_\Omega (f\cdot\nabla U\varphi - Uf\cdot\nabla\varphi) - \frac{1}{2}\int_\Omega \nabla\cdot fU\varphi + \frac{1}{2}\int_\Omega (f\cdot\overrightarrow{n})U\varphi.$$

Proof. Write

$$\int_\Omega (f\cdot\nabla U)\varphi = \frac{1}{2}\int_\Omega (f\cdot\nabla U)\varphi + \frac{1}{2}\int_\Omega (f\cdot\nabla U)\varphi,$$

and use integration by parts to obtain the result. □

We know that equation (2.24) is backwards in time, but in order to make the analysis simpler, we make it forward in time be writing $t = T - t$. This means that everything except the time derivative changes sign. If we in addition include the time derivative in the bilinear form and use the result from lemma 6, we get

$$\begin{aligned} a(U^{k+1},\varphi) = &\int_\Omega E\nabla U^{k+1}\cdot\nabla\varphi - \frac{1}{2}\int_\Omega (f\cdot\nabla U^{k+1}\varphi - U^{k+1}f\cdot\nabla\varphi) \\ &+ \int_\Omega (r + \frac{1}{h} + \frac{1}{2}\nabla\cdot f)U^{k+1}\varphi - \frac{1}{2}\int_{\Gamma_2}(\kappa(\theta - v) - \lambda v - \frac{1}{2}\rho\sigma v - \frac{1}{2}\sigma^2)U^{k+1}\varphi \\ &- \frac{1}{2}\int_{\Gamma_3}(rS - \frac{1}{2}\rho\sigma S - vS)U^{k+1}\varphi - \frac{1}{2}\int_{\Gamma_4}(\frac{1}{2}\sigma^2 - \kappa\theta)U^{k+1}\varphi. \end{aligned} \tag{5.3}$$

We have written out the boundary integral in lemma 6 and inserted $S = 0$ and $v = 0$ at the appropriate boundaries.

We will also need the inner product

$$(U,\varphi)_* = \int_\Omega E\nabla U\cdot\nabla\varphi + \int_\Omega (r + \frac{1}{h} + \frac{1}{2}\nabla\cdot f)U\varphi \tag{5.4}$$

and the corresponding norm

$$\|U\|_*^2 = (U,U)_*.$$

This is the same norm as we defined in section 2.2.5 with $\lambda_c = \frac{1}{h}$. In order for (5.4) to be an inner product, we need

$$r + \frac{1}{h} + \frac{1}{2}\nabla\cdot f \geq 0.$$

Throughout the analysis, we require that h fulfilles this condition.

We will also use that

$$f\cdot\nabla U = f^T I\cdot\nabla U = f^T(E^{-1/2}E^{1/2})\cdot\nabla U = (E^{-T/2}f)^T(E^{1/2}\nabla U).$$

Where we have used that $E^{-T/2} = E^{-1/2}$, which comes form the fact that E is symmetric. We now have everything we need to perform the analysis of the problem.

Coercivity

We start by showing that the bilinear form is coercive.

Lemma 7. Assuming that Feller's condition holds, the bilinear form is coercive with lower bound

$$a(U^{k+1}, U^{k+1}) \geq \|U^{k+1}\|_*^2.$$

Proof. We start by inserting U^{k+1} into (5.3), yielding

$$\begin{aligned} a(U^{k+1}, U^{k+1}) = \|U^{k+1}\|_*^2 &- \frac{1}{2}\int_\Omega (f \cdot \nabla U^{k+1} U^{k+1} - U^{k+1} f \cdot \nabla U^{k+1}) \\ &- \frac{1}{2}\int_{\Gamma_2} (\kappa(\theta - v) - \lambda v - \frac{1}{2}\rho\sigma v - \frac{1}{2}\sigma^2) U^{k+1} U^{k+1} \\ &- \frac{1}{2}\int_{\Gamma_3} (rS - \frac{1}{2}\rho\sigma S - vS) U^{k+1} U^{k+1} - \frac{1}{2}\int_{\Gamma_4} (\frac{1}{2}\sigma^2 - \kappa\theta) U^{k+1} U^{k+1}. \end{aligned}$$

We have

$$\int_\Omega (f \cdot \nabla U^{k+1} U^{k+1} - U^{k+1} f \cdot \nabla U^{k+1}) = \int_\Omega ((f \cdot \nabla U^{k+1}) U^{k+1} - U^{k+1}(f \cdot \nabla U^{k+1})) = 0.$$

Thus,

$$a(U^{k+1}, U^{k+1}) \geq \|U^{k+1}\|_*^2$$

if the constants satisfy

$$\begin{aligned} \kappa(\theta - v) - \lambda v - \frac{1}{2}\rho\sigma v - \frac{1}{2}\sigma^2 \leq 0 \qquad &\text{on } \Gamma_2 \\ rS - \frac{1}{2}\rho\sigma S - vS \leq 0 \qquad &\text{on } \Gamma_3 \\ \frac{1}{2}\sigma^2 - \kappa\theta \leq 0 \qquad &\text{on } \Gamma_4. \end{aligned}$$

□

We have some comments on the constraints in the proof. The first constraint is harmless and tells us that

$$\kappa\theta - \frac{1}{2}\sigma^2 \leq (\kappa + \lambda + \frac{1}{2}\rho\sigma) v_{\max},$$

which is easily fulfilled. The second constraint says that

$$r \leq \frac{1}{2}\rho\sigma + v.$$

Here, we must be more carefull to ensure that the constraint is fulfilled, but it is for our choice of constants. The last constraint is known as Feller's condition. This constant ensures that the CIR process for v is positive and is fulfilled by our choice of constants.

Continuity

To prove that the bilinear form is continuous, we need to find an upper bound for the bilinear form. We look at the bilinear form without boundary conditions, and the result is given in the following lemma.

Lemma 8. Assuming that $v_{\min} = \epsilon > 0$,

$$a(U^{k+1}, \varphi) \leq C\|U^{k+1}\|_*\|\varphi\|_*.$$

Proof. We start with equation (5.3) without the boundary conditions:

$$\begin{aligned}
a(U^{k+1}, \varphi) &= \int_\Omega E\nabla U^{k+1} \cdot \nabla\varphi - \frac{1}{2}\int_\Omega \Big(f \cdot \nabla U^{k+1}\varphi - U^{k+1} f \cdot \nabla\varphi\Big) \\
&+ \int_\Omega (r + \frac{1}{h} + \frac{1}{2\nabla \cdot f})U^{k+1}\varphi \\
&= (U^{k+1}, \varphi)_* - \frac{1}{2}\int_\Omega \Big((E^{-1/2} f \cdot E^{1/2}\nabla U^{k+1})\varphi - U^{k+1}(E^{-1/2} f \cdot E^{1/2}\nabla\varphi)\Big) \\
&\leq \|U^{k+1}\|_*\|\varphi\|_* + \frac{1}{2}\|E^{-1/2} f\varphi\|_{L^2(\Omega)}\|E^{1/2}\nabla U^{k+1}\|_{L^2(\Omega)} \\
&+ \frac{1}{2}\|E^{-1/2} f U^{k+1}\|_{L^2(\Omega)}\|E^{1/2}\nabla\varphi\|_{L^2(\Omega)} \\
&\leq \|U^{k+1}\|_*\|\varphi\|_* + C_1\|\varphi\|_{L^2(\Omega)}\|U^{k+1}\|_* + C_2\|U^{k+1}\|_{L^2(\Omega)}\|\varphi\|_* \\
&\leq \|U^{k+1}\|_*\|\varphi\|_* + C_3\|\varphi\|_*\|U^{k+1}\|_* + C_4\|U^{k+1}\|_*\|\varphi\|_* \\
&\leq C\|U^{k+1}\|_*\|\varphi\|_*.
\end{aligned}$$

□

We have some comments on the proof of the lemma. First, note that the matrix E is diagonalizable

$$E = \frac{1}{2}v\begin{bmatrix} \sigma^2 & \rho\sigma S \\ \rho\sigma S & S^2 \end{bmatrix} = \frac{1}{2}vPDP^{-1},$$

where D is a diagonal matrix. With this diagonalization, we can write

$$E^{-1/2} = \frac{1}{\sqrt{2v}}PD^{-1/2}P^{-1} \leq \frac{1}{\sqrt{2\epsilon}}PD^{-1/2}P^{-1}.$$

This expression is used to obtain the second inequality in the proof. Therefore, the constant in the proof depends on $1/\sqrt{\epsilon}$, where $\epsilon = v_{\min} > 0$. This the reason why we can't have $v_{\min} = 0$.

With continuity and coercivity in place, we can again use Lax-Milgram's Lemma *A*.2 that a unique solution exist at each time t when all constraints are fulfilled and $v_{\min} > 0$.

Stability

The article by Zhu and Kopriva [19] gives a stability estimate which we will present in this section.

Lemma 9.

$$\|U(T)\|_{L^2(\Omega)} \leq e^{bT}\|U_0\|_{L^2(\Omega)}.$$

Proof. We start with the forward weak form

$$\int_\Omega U_t\varphi = \int_{\partial\Omega} \varphi(E\nabla U)\cdot\overrightarrow{n} - \int_\Omega \nabla\varphi\cdot\nabla U + \int_\Omega f\cdot\nabla U\varphi - r\int_\Omega U\varphi.$$

If we choose $\varphi = U$ and use the boundary conditions in section 2.1.3 to make the boundary integral disappear, we have

$$\int_\Omega U_t U d\Omega = -\int_\Omega \nabla U\cdot\nabla U d\Omega + \int_\Omega Uf\cdot\nabla U d\Omega - r\int_\Omega U^2 d\Omega.$$

By using the L^2-norm and the fact that $E \geq 0$, we can write this as

$$\frac{1}{2}\frac{\partial}{\partial t}\|U\|^2_{L^2(\Omega)} \leq \int_\Omega Uf\cdot\nabla U d\Omega - r\|U\|^2_{L^2(\Omega)}.$$

The integral in this expression can be written as

$$\begin{aligned}\int_\Omega Uf\cdot\nabla U d\Omega &= \frac{1}{2}\int_\Omega (r - v - \frac{1}{2}\rho\sigma)SU_S U d\Omega \\ &+ \frac{1}{2}\int_\Omega (\kappa\theta - \frac{1}{2}\sigma^2)U_v U d\Omega \\ &- \frac{1}{2}\int_\Omega (\kappa + \lambda + \frac{1}{2}\rho\sigma)vU_v U d\Omega \\ &= \frac{1}{2}(I_1 + I_2 - I_3).\end{aligned}$$

We use integration by parts on I_1, I_2 and I_3 and obtain

$$\begin{aligned}\int_\Omega Uf\cdot\nabla U d\Omega = &\frac{L_S}{2}\int_0^{L_v}(r - v - \frac{1}{2}\rho\sigma)U^2 dv \mid_{S=L_S} \\ &+ \frac{1}{2}\left[(\kappa\theta - \frac{1}{2}\sigma^2) - (\kappa + \lambda + \frac{1}{2}\rho\sigma)L_v\right]\int_0^{L_S} U^2 dS \mid_{v=L_v} \\ &- \frac{1}{2}(\kappa\theta - \frac{1}{2}\sigma^2)\int_0^{L_S} U^2 dS \mid_{v=0} \\ &+ \frac{1}{2}(\kappa + \lambda + \rho\sigma - r)\|U\|^2_{L^2(\Omega)} + \frac{1}{2}\int_\Omega vU^2 d\Omega.\end{aligned}$$

We need all this integrals to be zero or negative. We then obtain some conditions that need to be fulfilled. First, we note that again we need Feller's condition to hold. We must also have

$$(\kappa\theta - \frac{1}{2}\sigma^2) - (\kappa + \lambda + \frac{1}{2}\rho\sigma)L_v \leq 0.$$

The first integral is equal to zero or negative if we change the boundary condition on Γ_3 to

$$\begin{cases} U = 0 & \text{if } r - v - \frac{1}{2}\rho\sigma > 0 \\ E\nabla U\cdot\overrightarrow{n} = 0 & \text{if } r - v - \frac{1}{2}\rho\sigma \leq 0. \end{cases}$$

If we also use

$$\int_\Omega vU^2 d\Omega \leq L_v\|U\|^2_{L^2(\Omega)},$$

we obtain

$$\frac{1}{2}\frac{\partial}{\partial t}\|U\|^2_{L^2(\Omega)} \leq \frac{1}{2}(\kappa + \lambda + \rho\sigma - r + L_v)\|U\|^2_{L^2(\Omega)} - r\|U\|^2_{L^2(\Omega)},$$

or

$$\frac{\partial}{\partial t}\|U\|^2_{L^2(\Omega)} \leq 2\gamma\|U\|^2_{L^2(\Omega)},$$

where $2\gamma = (\kappa + \lambda + \rho\sigma - 3r + L_v)$. Using Duhamel's formula then gives

$$\|U(T)\|^2_{L^2(\Omega)} \leq e^{\gamma T}\|U(0)\|^2_{L^2(\Omega)}.$$

□

Convergence

When coercivity, convergence and stability is assured, we have the following convergence estimate from [4]. This estimate is valid at each time t for the time discretized problem.

Lemma 10.

$$\|U - U_N\|_* \leq (1 + C)\|U - \Pi_N U\|_*,$$

where C is the continuity constant.

Proof. Assume that there exists a subspace $\mathcal{W} \subset W$ with an interpolation operator

$$\Pi_N : \mathcal{W} \to X$$

such that when $N \to \infty$,

$$\|U - \Pi_N U\|_* \to 0, \quad \forall U \in \mathcal{W}.$$

Next, we write $e = U_N - \Pi_N U$ such that

$$(\mathcal{L}e, \varphi) = (\mathcal{L}(U - \Pi_N U), \varphi), \quad \forall \varphi \in X.$$

The operator $\mathcal{L}$ is the linear operator corresponding to writing equation (2.23) as $\mathcal{L}U = 0$. Then, from the coercivity and convergence estimates, we have

$$\|e\|_* \leq C\|U - \Pi_N U\|_*.$$

If we use $U - U_N = U - \Pi_N U - e$, we obtain

$$\|U - U_N\|_* \leq (1 + C)\|U - \Pi_N U\|_*.$$

□

5.2 The polynomial collocation method

5.2.1 Spectral approximation

The mapping is the same as in section 5.1, giving us the following equation for the numerical solution U_N of (5.1) at the collocation points ξ_m, η_n:

$$\begin{aligned} &U_{N,t} + \frac{1}{2J_v^2}\sigma^2 v U_{N,\xi\xi} + \frac{1}{J_v J_S}\rho\sigma v S U_{N,\eta\xi} + \frac{1}{2J_S^2} v S^2 U_{N,\eta\eta} \\ &+ (\kappa(\theta - v) - \lambda v)\frac{1}{J_v}U_{N,\xi} + \frac{1}{J_S} r S U_{N,\eta} - r U_N \mid_{\xi,\eta=\xi_i,\eta_j} = 0. \end{aligned} \tag{5.5}$$

Then, we write U_N as

$$U_N = \sum_{m=0}^{N}\sum_{n=0}^{N} U_{mn}\ell_m(\xi)\ell_n(\eta).$$

As earlier, the Lagrangian polynomials represents a nodal basis. Inserting this into 5.5 gives

$$\begin{aligned}
&\sum_{m=0}^{N}\sum_{n=0}^{N}\dot{U}_{mn}\ell_m(\xi_i)\ell_n(\eta_j) + \frac{1}{2}\sigma^2\frac{1}{J_v^2}\sum_{m=0}^{N}\sum_{n=0}^{N}U_{mn}v_n\ell_m''(\xi_i)\ell_n(\eta_j)\\
&+\rho\sigma\frac{1}{J_vJ_S}\sum_{m=0}^{N}\sum_{n=0}^{N}v_nS_mU_{mn}\ell_m'(\xi_i)\ell_n'(\eta_j) + \frac{1}{2J_S^2}\sum_{m=0}^{N}\sum_{n=0}^{N}v_nS_m^2U_{mn}\ell_m(\xi_i)\ell_n''(\eta_j)\\
&+\frac{1}{J_v}\sum_{m=0}^{N}\sum_{n=0}^{N}(\kappa(\theta-v_n)-\lambda v_n)U_{mn}\ell_m'(\xi_i)\ell_n(\eta_j) + r\frac{1}{J_S}\sum_{m=0}^{N}\sum_{n=0}^{N}S_mU_{mn}\ell_m(\xi_i)\ell_n'(\eta_j)\\
&-r\sum_{m=0}^{N}\sum_{n=0}^{N}\ell_m(\xi_i)\ell_n(\eta_j)\\
&=\sum_{m=0}^{N}\sum_{n=0}^{N}\dot{U}_{mn}\delta_{im}\delta_{jn} + \frac{1}{2J_v^2}\sigma^2\sum_{m=0}^{N}\sum_{n=0}^{N}v_nU_{mn}D_{im}^{(2)}\delta_{jn}\\
&+\rho\sigma\frac{1}{J_vJ_S}\sum_{m=0}^{N}\sum_{n=0}^{N}v_nS_mU_{mn}D_{im}D_{jn} + \frac{1}{2J_S^2}\sum_{m=0}^{N}\sum_{n=0}^{N}v_nS_m^2U_{mn}D_{jn}^{(2)}\delta_{im}\\
&+\frac{1}{J_v}\sum_{m=0}^{N}\sum_{n=0}^{N}(\kappa(\theta-v_n)-\lambda v_n)U_{mn}D_{im}\delta_{jn} + r\frac{1}{J_S}\sum_{m=0}^{N}\sum_{n=0}^{N}S_mU_{mn}D_{jn}\delta_{mi}\\
&-r\sum_{m=0}^{N}\sum_{n=0}^{N}U_{mn}\delta_{im}\delta_{jn}, \quad 0\le i,j\le N.
\end{aligned}$$

In the same manner as for the spectral Galerkin method, this can be written as a system

$$\mathbf{M}\dot{\mathbf{U}} + \mathbf{A}\mathbf{U} = \mathbf{F},$$

where A_{ij} and M_{ij} are defined in the same manner as for the spectral Galerkin method and $\mathbf{F}$ contains the boundary values.

5.2.2 Implementation of boundary conditions

In order for the boundary conditions to be fulfilled, we must include them in the system. The Dirichlet condition on Γ_1 is included in the same manner as for the spectral Galerkin method.

On Γ_2, we have $E\nabla U\cdot\overrightarrow{n} = 0$, which is equal to $\rho SU_S + \sigma U_v = 0$. Approximating this gives

$$\frac{1}{J_S}\rho\sum_{m=0}^{N-1}S_mU_{mN}\delta_{im}D_{jN} + \sigma\frac{1}{J_v}\sum_{m=0}^{N-1}U_{mN}D_{im}\delta_{jN}.$$

Using the same approach for the condition on Γ_3 gives

$$\frac{1}{J_S}L_S\sum_{n=0}^{N}U_{Nn}\delta_{iN}D_{jn} + \rho\sigma\frac{1}{J_v}\sum_{n=0}^{N}U_{Nn}D_{iN}\delta_{jn}.$$

The stiffness matrix $\mathbf{A}$ must be updated to include these contributions as well.

5.2.3 Numerical results

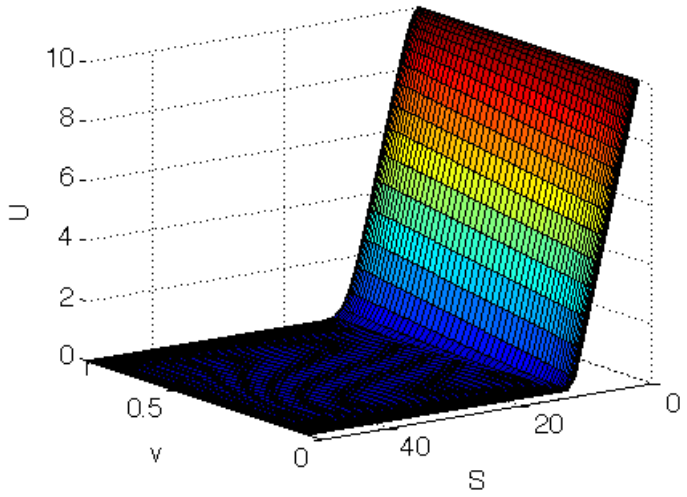


Figure 5.4: Numerical solution after 10 time steps for equation (2.23) when $N = 60$ and $h = 10^{-3}$.

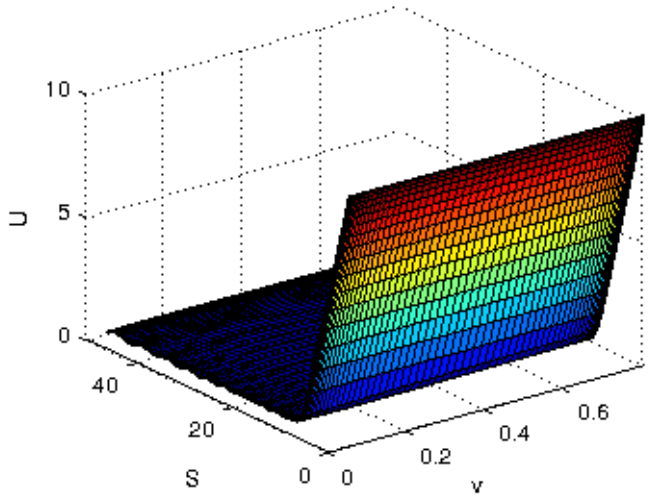


Figure 5.5: Exact solution after 10 time steps for equation (2.23) when $N = 60$ and $h = 10^{-3}$.

The solution after 10 time steps when $N = 60$ and $h = 10^{-3}$ is shown in figure 5.4, and the exact solution at the same time is shown in figure 5.5. The difference between the exact and numerical solution is shown in figure 5.6. As for the spectral Galerkin method, we do not get the correct solution. We see from figure 5.6 that the error is large for $S \in (0, K]$ and increases as v increases. In addition, we have an error at $(v_{\max}, S_{\max})$ which increases at each time step. After 50 time steps, this error has blown up to 10^7 and completely dominates the solution. This is why we present solutions after only 10 of 1000 time steps.

The oscillations in figure 5.6 comes from the exact solution in figure 5.5. The exact solution is calculated for a European call option and transformed to a put option by

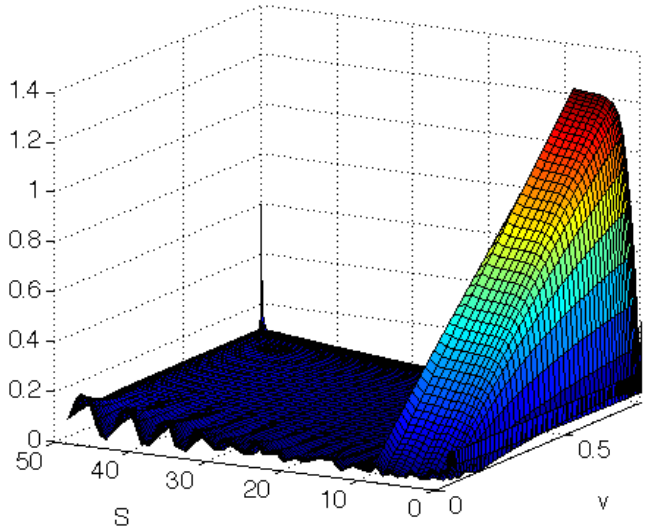


Figure 5.6: Difference between the exact and numerical solution after 10 time steps when $N = 60$ and $h = 10^{-3}$.

the put-call-parity

$$V_{\text{put}} = V_{\text{call}} - S + Ke^{-r(T-t)},$$

and this is unstable when $T - t$ is small.

In [17], the Heston model is used to price American perpetual put options. Perpetual means that the expiry date is infinite, so the problem that is solved is time-independent. The model is solved using the spectral collocation method. Making the problem time-independent can be the difference needed to solve the problem.

The solution may be incorrect because we have done something wrong in the approximation or implementation. This is even more likely in this case than for the spectral Galerkin method because the error at $(v_{\max}, S_{\max})$ probably comes from wrong boundary conditions or wrong implementation of the boundary conditions. If the boundary conditions are incorrect, then it is the case for the spectral Galerkin method as well and might be the reason why we don't get the correct solution.

5.2.4 Analysis of the method

The analysis of the polynomial collocation method involves using the discrete inner product

$$(U, \varphi)_N = \sum_{k=0}^{N} \sum_{l=0}^{N} U(v_k, S_l)\varphi(v_k, S_l)\rho_k\rho_l$$

with the discrete norm

$$\|U\|_N^2 = (U, U)_N$$

to write the collocation solution as

$$(\mathcal{L}_N U_N, \varphi_{kl}) = 0, \quad 0 \le k, l \le N,$$

and then proving that the operator $\mathcal{L}_N$ is coercieve and continuous. Then, the same theory as for the spectral Galerkin method can be applied to prove stability and co-

ercivity. The results will therefore be similar to section 5.1.4, and we will not repeat them here.

Chapter 6

Conclusions and further work

6.1 Conclusion

The goal of this thesis was to solve option pricing models efficiently by using spectral methods. We have solved two option pricing models, Black-Scholes model and Heston's stochastic volatility model. Both models have been solved using two spectral methods, the spectral Galerkin method and a polynomial collocation method. All methods have been based on the Gauss-Lobatto-Legendre nodes and the Lagrange polynomials.

Black-Scholes model was solved with the finite element method as well. We observed that the two spectral methods required a much smaller number of unknowns than the finite element method, which also made the computation time smaller. However, we did not achieve spectral convergence for the two method because the inital data is not analytic. The computational results from the two spectral methods were almost identical, but the polynomial collocation method was a bit faster than the spectral Galerkin method. We also observed that solving the original equation gives better convergence and smaller errors than solving the transformed equation.

For the Heston stochastic volatility model, we were not able to get any good solutions neither by the spectral Galerkin method nor the polynomial collocation method. There are two possibilities for why this did not work: either we have done something wrong in the spectral approximation or implementation of both methods, or the Heston model is not solvable by the spectral method with our time discretization schemes. We have found one article [19] which solves the Heston model by using a spectral element method, and one article [17] which solves the Heston model for American perpetual options by using a polynomial collocation method, but no articles where the Heston model for European option is solved with either of the methods we have tried. Therefore, there is a possibility that the methods do not work for this problem.

The thesis shows that even though the spectral method is best for solving problems with smooth solutions and inital data, we get good results when we solve the one-dimensional option pricing problems as well. However, the two-dimensional stochastic volatility models are probably better solved by using other numerical methods than the spectral method.

6.2 Further work

Some problems have accured while solving the option pricing models with spectral methods. The main issue is the inital condition which is not smooth. This introduces

the Gibbs phenomenom, which is oscillations of the solution in the interior of the domain. Together with the variable coefficients of the equations, this causes errors to our solutions. However, as suggested in [8], this can be avoided by using various techniques such as using filters or introducing a different basis with very special properties.

As mentioned earlier, there are many possible choices of basis functions and quadrature. We have used Gauss-Lobatto-Legendre nodes and quadrature and Lagrange polynomials as basis functions. Other common choices include using Fourier series instead of Lagrange polynomials, and one could try to solve the models using this approach. The Chebyshev points are also commonly used as nodes for the methods. We have observed that the placement of the nodes impacts the solution, so we could obtain better results by changing to other sets of nodes.

We could also try to use the spectral element method instead. This combines the advantages of the spectral method and the finite element method. The method is used in [18] to solve Black-Scholes for two assets and in [19] to solve the Heston model. Because we know that this method gives correct solutions, we could use this method instead.

A third possibility is to try to approximate the inital condition by a smooth function. As mentioned earlier, this would give convergence towards the wrong solution and it is therefore not the best approach to obtain better convergence.

To obtain faster solvers, one could parallelize the codes in order to solve the problems on supercomputers or clusters using multiple processors or threads.

Appendix A

Theorems and lemmas

A.1 Ito's Lemma in multiple dimensions

The lemma can be found in for example [2] or in continuous form in [7]. We start by assuming that $x_1, x_2, \ldots, x_N$ are solutions to

$$dx_i = a_i(t, x_1, x_2, \ldots, x_N)dt + \sum_{p=1}^{M} \sigma_{ip}(t, x_1, x_2, \ldots, x_N)dW_p,$$

where $W_1, \ldots, W_M$ are Wiener processes with correlation structure defined in the $N \times N$ matrix Q, where $Q_{ij}dt = \mathbb{E}[dW_i dW_j]$ and σ_{ip} is an $N \times M$ array.

Let $F(t, x_1, \ldots, x_N)$ be a function. Then Ito's lemma states

Lemma 11.

$$dF = \left[\frac{\partial F}{\partial t} + \sum_{j=1}^{N} \frac{\partial F}{\partial x_j} a_j + \frac{1}{2} \sum_{j,k=1}^{N} \frac{\partial^2 F}{\partial x_j \partial x_k} g_{jk}\right] dt + \sum_{j=1}^{N} \sum_{p=1}^{M} \frac{\partial F}{\partial x_j} \sigma_{jp} dW_p,$$

where

$$g_{jk} = \sum_{p=1}^{N} \sum_{q=1}^{M} \sigma_{jp}\sigma_{kq}Q_{pq}.$$

A.2 Lax-Milgram's Lemma

From [12], we have the following Lemma:

Lemma 12. Let V be a Hilbert space, $a(\cdot, \cdot) : V \times V \to \mathbb{R}$ a continuous and coercive bilinear form, $F(\cdot) : V \to \mathbb{R}$ a linear and continouos functional. Then there exists one unique solution to the problem

$$\text{find } u \in V : \quad a(u, v) = F(v) \quad \forall v \in V.$$

A.3 Gronwall's Lemma on integral form

Lemma 13. If $u(t)$ satisfies

$$u(t) \leq \alpha(t) + \int_a^t \beta(s)u(s)ds$$

for some non-negative β and non-decreasing α, then

$$u(t) \leq \alpha(t)\exp\left(\int_a^t \beta(s)ds\right).$$

The lemma can be found in [12, chapter 2.7].

Bibliography

[1] Y. Achdou and O. Pironneau. *Computational Methods for Option Pricing*. SIAM Society for Industrial and Applied Mathematics, 2005.

[2] J. Akahori. Multi-dimensional discrete stochastic calculus and its application to a problem in mathematical finance, 2010.

[3] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[4] C. Canuto, M. Y. Hussani, A. Quarteroni, and T. A. Zang. *Spectral Methods. Fundamentals in Single Domains*. Springer Verlag, 2006.

[5] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structures of interest rates. *Econometrica*, 53:385–407, 1985.

[6] P. Daskalopoulos and P. M. N. Feehan. Existence, uniqueness and global regularity for degenerate elliptic obstacle problems in mathematichal finance. 2011.

[7] R. J. Elliott and E. Kopp. *Mathematics of Financial Markets*. Springer Verlag, 2005.

[8] D. Gottlieb and J. S. Hesthaven. Spectral methods for hyperbolic problems. *Journal of Computational and Applied Mathematics*, 128:83–131, 2000.

[9] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6:327–343, 1993.

[10] D. A. Kopriva. *Implementing Spectral Methods for Partial Differential Equations*. Springer Verlag, 2009.

[11] L. Ornstein and G. E. Uhlenbeck. On the theory of the brownian motion. *Physical Review*, 36:823–841, 1930.

[12] A. Quarteroni. *Numerical Models for Differential Problems*. Springer-Verlag Italia, 2009.

[13] F. D. Rouah. Simplified derivation of the heston model, 2013.

[14] J. Shen, T. Tang, and L.-L. Wang. *Spectral Methods. Algorithms, Analysis and Applications*. Springer Verlag, 2011.

[15] N. V. D. Wijst. *Finance - A Quantitative Introduction*. Cambridge University Press, 2013.

[16] G. Winkler, T. Apel, and U. Wystup. Valuation of options in heston's stochastic volatility model using finite element methods. 2001.

[17] S.-P. Zhu and W.-T. Chen. A spectral-collocation method for pricing perpetual american puts with stochastic volatility. *Applied Mathematics and Computation*, 217:9033—-9040, 2011.

[18] W. Zhu and D. A. Kopriva. A spectral element approximation to price european options. ii. the black-scholes model with two underlying assets. *Journal of Scientific Computing*, 39:323–339, 2009.

[19] W. Zhu and D. A. Kopriva. A spectral element approximation to price european options with one asset and stochastic volatility. *Journal of Scientific Computing*, 42:426–446, 2010.

[20] Ömür Ugur. *An Introduction to Computational Finance*. Imperial College Press, 2009.