Valuation of Options in Heston's Stochastic Volatility Model Using Finite Element Methods [1]

Gunter Winkler

Department of Mathematics
Chemnitz University of Technology
Reichenhainer Strasse 41
09126 Chemnitz
GERMANY
phone: +49 - 371 - 531 - 2154
gunter.winkler@mathematik.tu-chemnitz.de
http://www-user.tu-chemnitz.de/~wgu

Thomas Apel

Department of Mathematics
Chemnitz University of Technology
Reichenhainer Strasse 41
09126 Chemnitz
GERMANY
phone: +49 - 371 - 531 - 2146
thomas.apel@mathematik.tu-chemnitz.de
http://www-user.tu-chemnitz.de/~tap

Uwe Wystup

Commerzbank Treasury and Financial Products
Neue Mainzer Strasse 32–36
60261 Frankfurt am Main
GERMANY
phone: +49 - 69 - 136 - 41067
fax: +49 - 69 - 136 - 48475
wystup@mathfinance.de
http://www.mathfinance.de

October 24 2001

Abstract

We apply finite element methods to determine values of options in HESTON's stochastic volatility model

[1] This article first appeared in *Foreign Exchange Risk*, Risk Publications, London 2001

1 Introduction

Due to the smile observed in options markets numerous authors have suggested different models such as generalized Levy processes, fractional Brownian motion, entropy based models [4], jump diffusions and stochastic volatility models. For vanilla options (put and call options) the dependence of the price on the volatility is monotone, whence using the Black-Scholes formula along with a volatility smile matrix is sufficient. Values of exotic options, however, do not always depend on the volatility in a monotone fashion, whence pricing consistently with the smile requires a more sophisticated model. Therefore, it is important to find efficient ways to calculate exotic option values in exotic models.

Melino and Turnbull showed in [11] that the assumption of stochastic volatility leads to a distribution of the underlying which is closer to empirical observations than the log-normal distribution. Heston's stochastic volatility model [5] can explain the smile observed in foreign exchange vanilla options markets to some extend, particularly well for maturities between one month and one year for liquidly traded currency pairs, and one would hence like to use it to determine values of more exotic options. A closed-form solution, which is available for put and call options, has not been found for more exotic options, not even for the first generation exotics such as barrier- and one-touch options.

In stochastic volatility models the value of an option is usually specified by a partial differential equation. Since closed-form solutions are often unavailable and Monte Carlo Simulations are too slow, one needs to solve a partial differential equation. Hull and White, who contributed numerous articles on stochastic volatility in [6], suggest to use finite difference methods (*FDM*). The FDM is straightforward to implement for rectangular domains, an advantage that is lost, when exotic options are considered. Additionally the terminal and boundary conditions must be sufficiently smooth to guarantee existence and uniqueness of a smooth classical solution. These constraints can be weakened when we use finite element methods (*FEM*) which is based on the weak (variational) formulation of the problem. The difficulties to prove existence and uniqueness of the solution do not disappear, but carry over the area of (weighted) Sobolev spaces. For the finite difference method we would like to mention the work by Kurpiel and Roncalli [10], which illustrates how to apply the Hopscotch method for the valuation of options. This is a rather advanced finite-difference style technique. However, a combination of boundary conditions due to barrier option contract specification and a mixed second derivative term in the partial differential equation due to the correlation of the volatility and the exchange rate process exhibit difficulties. After changing to principal axes the boundary conditions for barrier options usually lead to triangular shaped regions. The solution of partial differential equations specified on such regions is often handled with finite element methods. We take the example of Heston's model to illustrate how to do this in theory and practice.

Heston's partial differential equation is of second order with variable coefficients, which vanish on the border of their domain. In order to show the existence of a solution we need to examine the coefficients in detail. Generally we have[2]

$$u_t - \nabla \cdot A\nabla u + b \cdot \nabla u + \check{c}u \quad = \quad 0, \tag{1}$$

[2] We use the usual notation of the FEM literature, as for example in [12].

which lead after semi-discretization with respect to time to the variational formulation

$$\int_\Omega (\nabla u \cdot A\nabla v + (b \cdot \nabla u)v + cuv) - \int_{\partial\Omega} (A\nabla u \cdot \vec{n})v \quad = \quad \int_\Omega fv\,.$$

For each time step we seek a funktion $u(x, y)$ satisfying this integral equation for all functions $v(x, y) \in V_0$ and the boundary conditions.[3]
The question comes up how to choose V_0. In the case of constant coefficients A, b and c, V_0 is a subspace of the Sobolev space $W^{1,2}(\Omega)$, mostly denoted by $H^1(\Omega)$. This is the space of all functions $v(x, y)$ satisfying $v \in \mathrm{L}_2(\Omega)$ and $\nabla v \in \mathrm{L}_2(\Omega)$. In the case of space dependent coefficients the question becomes much harder. There can be singularities, i.e. the coefficients can become zero. Kufner und Sändig examine *weighted Sobolev spaces* which are adequate for such situations in [9]. They show that a major part of the theory of (regular) Sobolev spaces can be conveyed to weighted spaces. In this environment one can formulate and prove statements about existence and uniqueness, when some parameters are zero on parts of the boundary.
However, in our case where a convection term $b\cdot\nabla v$ is contained in the differential equation (1) and Neumann boundary conditions we would like to use a trace theorem which is beyond the theory developed in Kufner's fundamental book [8].
For this reason we have taken a different approach, namely to modify the region of the domain of the partial differential equation. The price we have to pay is to seek suitable boundary conditions of the smaller region. We suggest a solution, prove its existence and determine a solution numerically. We also discuss how to choose the mesh and how the exactness of the solution depends on the parameters.

2 Heston's stochastic volatility model

2.1 The model

Heston's model is based on the following equations:

$$\begin{aligned}
\mathrm{d}S_t &= S_t \left[\mu\ \mathrm{d}t + \sqrt{v(t)}\,\mathrm{d}W_t^{(1)}\right], \\
\mathrm{d}v_t &= \kappa(\theta - v_t)\ \mathrm{d}t + \xi\sqrt{v(t)}\,\mathrm{d}W_t^{(2)}, \\
\mathbf{Cov}\left[\mathrm{d}W_t^{(1)}, \mathrm{d}W_t^{(2)}\right] &= \rho\ \mathrm{d}t, \\
\lambda(S, v, t) &= \lambda v.
\end{aligned}$$

As usual S denotes the spot, t the time, v the variance, μ the (risk neutral) drift, ξ the volatility of the volatility and W_i two Wiener processes with correlation ρ.
The model for the variance v_t is the same as the one used by Cox, Ingersoll and Ross for the short term interest rate. We think of $\theta > 0$ as the long term variance, of $\kappa > 0$ as the rate of mean-reversion. The quantity $\lambda(S, v, t)$ is called the market price of volatility risk.

[3] In the following sections we will also use the notation ψ for the test functions v.

2.2 The partial differential equation for a general contingent claim

We consider the value function of a general contingent claim $U(t, v, S)$ paying $g(S) = U(T, v, S)$ at time T. HESTON derives the partial differential equation

$$\begin{aligned} &U_t + \frac{1}{2}\xi^2 v U_{vv} + \rho\xi v S U_{vS} + \frac{1}{2} v S^2 U_{SS} \\ &+[\kappa(\theta - v) - \lambda v]U_v + (r_d - r_f)SU_S - r_d U = 0, \end{aligned} \tag{2}$$

which U must satisfy. The numbers r_d and r_f denote the domestic and foreign interest rate respectively.

2.3 Boundary conditions for vanilla options

We obtain a solution to (2) by specifying appropriate boundary conditions. For a European vanilla option these are[4]

$$U(T, v, S) = [\phi(S - K)]^+, \tag{3}$$

$$U(t, v, 0) = \frac{1 - \phi}{2} K e^{-r_d \tau}, \tag{4}$$

$$U_S(t, v, \infty) = \frac{1 + \phi}{2} e^{-r_f \tau}, \tag{5}$$

$$r_d U(t, 0, S) = (r_d - r_f) S U_S(t, 0, S) + \kappa\theta U_v(t, 0, S) + U_t(t, 0, S)\,, \tag{6}$$

$$U(t, \infty, S) = \begin{cases} S e^{-r_f \tau} & (\phi = +1) \\ K e^{-r_d \tau} & (\phi = -1)\,, \end{cases} \tag{7}$$

where K denotes the strike price, $\tau = T - t$ the time to maturity and ϕ denotes the binary variable taking the values $+1$ for a call and -1 for a put.

Note that $\kappa(\theta - v) - \lambda v = \kappa\theta - (\kappa + \lambda)v = (\kappa + \lambda)\left(\frac{\kappa}{\kappa+\lambda}\theta - v\right)$, hence we can work with two parameters $\kappa' = \kappa + \lambda$ and $\theta' = \frac{\kappa}{\kappa+\lambda}\theta$, instead of three parameters κ, θ and λ, and

$$U(t, v, S; \kappa, \theta, \xi, \rho, \lambda) = U\left(t, v, S; \kappa + \lambda, \frac{\kappa}{\kappa + \lambda}\theta, \xi, \rho, 0\right). \tag{8}$$

This means that when modeling vanilla prices, we can set $\lambda = 0$ by default and just determine the parameters κ, θ, ξ and ρ. Conversely, when doing a Monte Carlo simulation, we can include a market price of volatility risk λ by modifying κ and θ.

[4] Equation (6) can be obtained by inserting $v = 0$ into Equation (2). Thus it is not necessary to state a condition at this boundary.

2.4 Boundary conditions for regular knock-out options

For a regular down-and-out call with strike K, barrier B, rebate R and $0 < B < K$ the boundary conditions are

$$\begin{aligned} U(T,v,S) &= [S-K]^+, && (9)\\ U(t,v,B) &= Re^{-\omega r_d \tau}, && (10)\\ U_S(t,v,\infty) &= e^{-r_f\tau} \text{ or } U(t,v,S_{max}) = S_{max}e^{-r_f\tau} - Ke^{-r_d\tau}, && (11)\\ U(t,0,S) &= [Se^{-r_f\tau} - Ke^{-r_d\tau}]^+, && (12)\\ U(t,\infty,S) &= \left(1-\frac{B}{S}\right)Se^{-r_f\tau}, && (13) \end{aligned}$$

The binary variable ω takes the value 0 if the rebate is paid at hitting time or 1 if the rebate is paid at expiry T. In practice the values 0 and ∞ for the variance v have to be replaced by appropriately chosen v_{min} and v_{max}. Convergence to the stated boundary conditions is slow, whence one has to choose extreme values for the variance. Analogously for an up-and-out put with $0 < K < B$ the boundary conditions are

$$\begin{aligned} U(T,v,S) &= [K-S]^+, && (14)\\ U(t,v,B) &= Re^{-\omega r_d \tau}, && (15)\\ U(t,v,0) &= Ke^{-r_d\tau} \text{ or } U(t,v,S_{min}) = Ke^{-r_d\tau} - S_{min}e^{-r_f\tau}, && (16)\\ U(t,0,S) &= [Ke^{-r_d\tau} - Se^{-r_f\tau}]^+, && (17)\\ U(t,\infty,S) &= \left(1-\frac{S}{B}\right)Ke^{-r_d\tau}. && (18) \end{aligned}$$

Boundary conditions (12), (13), (17) and (18) can be replaced by

$$U_{vv} = 0. \tag{19}$$

3 Finite element method

3.1 The task

In this section we consider the terminal value problem for the function $w = w(t,v,x)$ which follows from the value function $U(t,v,S) = w(t,v,\log S/K)$ after the change of variables $x = \ln(S/K)$,

$$\begin{aligned} &w_t + \frac{1}{2}\xi^2 v w_{vv} + \rho\xi v w_{vx} + \frac{1}{2}v w_{xx} + [\kappa(\theta - v) - \lambda v]w_v \\ &+(r_d - r_f - \frac{1}{2}v)w_x - r_d w = 0, && (20)\\ &w(T,v,x) = g(Ke^x) && (21) \end{aligned}$$

for all $(t,v,x) \in [0,T) \times \Omega_\infty$ with $\Omega_\infty = [0,\infty) \times (-\infty,+\infty)$. Here $g(Ke^x)$ denotes the payoff-function of the option. We reformulate this in matrix notation for the sake of a better overview. The dot $\cdot$ is used to denote a scalar product.

We define

$$A \triangleq \frac{1}{2}v \begin{bmatrix} \xi^2 & \rho\xi \\ \rho\xi & 1 \end{bmatrix}, \tag{22}$$

$$b \triangleq \begin{bmatrix} -\kappa(\theta - v) + \lambda v + \frac{1}{2}\xi^2 \\ -(r_d - r_f) + \frac{1}{2}v + \frac{1}{2}\xi\rho \end{bmatrix} \tag{23}$$

and write the partial differential equation in the form

$$0 = w_t + \nabla \cdot A\nabla w - b \cdot \nabla w - r_d w\,. \tag{24}$$

3.2 How to determine the solution

Three variables The partial differential equation contains derivatives with respect to the time t, the variance v and the logarithm of the spot x. To keep the allocated memory limited we consider a series of terminal value problems in two dimensions rather than a three-dimensional mesh.

Unbounded domain For a solution using finite element methods we must limit Ω_∞ to a bounded domain Ω. For the hence arising new boundaries we need suitable boundary conditions depending on the actual payoff of the option to be valued. We put

$$\Omega = (v_{min}, v_{max}) \times (x_{min}, x_{max}) \tag{25}$$

$$\partial\Omega = \Gamma_1 \cup \Gamma_2\,, \Gamma_1 \cap \Gamma_2 = \emptyset \tag{26}$$

where Γ_1 denotes the part of the boundary of Ω on which Dirichlet-type boundary conditions are set.

Formulation as a variational problem The next step is to solve the obtained boundary value problems for all time steps t^k. For this purpose we multiply it with a test function ψ taken from the function space V_0 and integrate it over the domain Ω. We translate the second derivatives into first derivatives using integration by parts. The resulting integrals over the boundary of the domain can be computed if boundary conditions and test functions are chosen adequately. In particular, we take the test functions to be zero on the part of the boundary, which is defined by Dirichlet-type conditions, making the corresponding integral vanish.

Each solution of the boundary value problem in the classical formulation is a solution of the variational problem. Adversely, if a solution of the variational problem is twice differentiable, then it is a solution of the boundary value problem in the classical solution.

Existence of a solution The question about the existence of a solution strongly depends on which space we are seeking the solution to be in. We need the integrability of the solution, i.e. their norms must be finite, where the norm is derived from the partial differential equation. We draw special attention to the influence of the factor v in front of the second derivatives, since in a neighborhood of zero the derivatives can become arbitrarily large without spoiling the integrability.

3.3 Semidiscretization in the time direction

The partial differential equation (20) is of parabolic type in three variables. We choose a semidiscretization in time which yields a series of two-dimensional boundary value problems. We take times t^k and consider the difference quotient

$$D_t^+ w(t^k, v, x) = \frac{w(t^{k+1}, v, x) - w(t^k, v, x)}{\tau}, \tag{27}$$

where for simplicity, $\tau = t^{k+1} - t^k$ is constant for all k. We use the usual σ-weighted scheme noting that the computation is working backwards in time, i.e., the initial value is given for $t = t^N = T$, the solution is sought at $t^0 = 0$. We take interpolated values between two successive times t^k und t^{k+1} for the spatial derivatives. For $\sigma = 1$ we obtain a purely implicit method, for $\sigma = 0$ a purely explicit method, for $\sigma = \frac{1}{2}$ the CRANCK-NICOLSON scheme. We obtain

$$\begin{aligned} 0 &= \frac{w(t^{k+1}) - w(t^k)}{\tau} \\ &+ (1-\sigma)\left(\nabla \cdot A\nabla w(t^{k+1}) - b \cdot \nabla w(t^{k+1}) - r_d w(t^{k+1})\right) \\ &+ \sigma\left(\nabla \cdot A\nabla w(t^k) - b \cdot \nabla w(t^k) - r_d w(t^k)\right) . \end{aligned} \tag{28}$$

This setup contains a terminal condition, i.e. in each timestep we can assume we know the values $w(t^{k+1})$ and have to solve a differential equation of the form

$$Lw(t^k, v, x) = f(t^{k+1}, v, x), \tag{29}$$

where L and f are defined by

$$Lw(t, v, x) = -\sigma\nabla \cdot A\nabla w + \sigma b \cdot \nabla w + (\sigma r_d + \frac{1}{\tau})w, \tag{30}$$

$$f(t, v, x) = (1-\sigma)\nabla \cdot A\nabla w - (1-\sigma)b \cdot \nabla w - ((1-\sigma)r_d - \frac{1}{\tau})w. \tag{31}$$

3.4 Boundary conditions for a call

We choose boundary conditions for $w(t, v, x)$ based on section 2.4.

$$\begin{array}{lll} \Gamma_a & v = v_{min} & w(t, v_{min}, x) = Ke^{x-r_f(T-t)}\Phi(d_+) - Ke^{-r_d(T-t)}\Phi(d_-) \\ \Gamma_b & v = v_{max} & w(t, v_{max}, x) = Ke^{x-r_f(T-t)} \\ \Gamma_c & x = x_{min} & w = \lambda w(t, v_{max}, x_{min}) + (1-\lambda)w(t, v_{min}, x_{min}) \\ & & \lambda = \frac{v - v_{min}}{v_{max} - v_{min}} \\ \Gamma_d & x = x_{max} & \frac{\partial}{\partial \nu} w(t, v, x_{max}) \stackrel{\Delta}{=} A\nabla w \cdot \vec{n} = \frac{1}{2} v K e^{x-r_f(T-t)} \end{array}$$

The boundary conditions for $v = v_{min}$ and $v = v_{max}$ follow from the Black-Scholes Equation (33), the values for d_+ and d_- are computed for $\theta = v_{min}$ and $\theta = v_{max}$ respectively. The Black-Scholes Equation (33) itself is a special case of HESTON's model, which we obtain for $\xi = 0$ and initial variance $v(0) = \theta$,

$$\begin{aligned} C &= Se^{-r_f T}\mathcal{N}(d_+) - Ke^{-r_d T}\mathcal{N}(d_-)\,, \\ d_\pm &= \frac{\log(S/K) + (r_d - r_f \pm \theta/2)T}{\sqrt{\theta T}}\,. \end{aligned} \tag{33}$$

The standard normal cumulative distribution function is denoted by $\mathcal{N}$. The Dirichlet boundary $\Gamma_1 = \Gamma_a \cup \Gamma_b \cup \Gamma_c$ consists of three edges. The fourth edge $\Gamma_2 = \Gamma_d$ has a Robin-type boundary condition, which resembles

$$\nabla w = \begin{bmatrix} 0 \\ Ke^{x-r_f(T-t)} \end{bmatrix}. \tag{34}$$

We need to make sure the Dirichlet-type boundary conditions yield the same limits on the corners of the domain, when moving to a corner from two distinct vertices. We need this condition to guarantee existence of a continuous solution on the entire domain Ω.
Assumption. On the boundary Γ_c we interpolate linearly between the values of the right and the left boundary. For sufficiently small x_{min} the error will also be small. The value of a call in this region is nearly zero.

3.5 Type of the partial differential equation

The partial differential equation (29) is a heat equation with a diffusion term, a convection term and a reaction term.
The convection term, i.e. the parameter $b(v, x)$, influences the quality of the solution. The choice of numerical methods depends on whether the problem is diffusion dominated or convection dominated. The method described in this paper works better for diffusion dominated problems. Methods for larger convection terms will be investigated in further research.
Figures 1 and 2 show vector b for two sample data sets, where in the first case $r_d < r_f$, i.e. the drift is upwards, and in the second case $r_d > r_f$. Here we have a change in the direction from down to up of the drift at $v = 2(r_d - r_f) - \xi\rho$, in Figure 2 at $v = 0.04$. The direction and intensity of the drift depends only on v, not on x.
To prove the existence of a unique solution it is necessary to to have an upward flow, that means to have $r_d < r_f$ as in Figure 1. However, we can assume this without loss of generality. We can exchange the currencies using the foreign-domestic symmetry (see the section on vanilaa options of the MathFinance Formula Catalogue [13]),

$$\frac{1}{S}U(t, v, S; K, r_d, r_f, \phi) \quad = \quad KU(t, v, \frac{1}{S}; \frac{1}{K}, r_f, r_d, -\phi). \tag{35}$$

Together with the put-call-parity

$$U(t, v, S; K, r_d, r_f, \phi) - U(t, v, S; K, r_d, r_f, -\phi) \quad = \quad Se^{-r_f\tau} - Ke^{-r_d\tau}, \tag{36}$$

we get an relation between the domestic call and the foreign call ($\phi = 1$),

$$U(t, v, S; K, r_d, r_f, 1) \quad = \quad SKU(t, v, \frac{1}{S}; \frac{1}{K}, r_f, r_d, 1) - Ke^{-r_d\tau} + Se^{-r_f\tau} \tag{37}$$

3.6 Formulation as a variational problem

We define appropriate spaces where we use that the matrix A defined in (22) is positive definite for all $v > 0$.

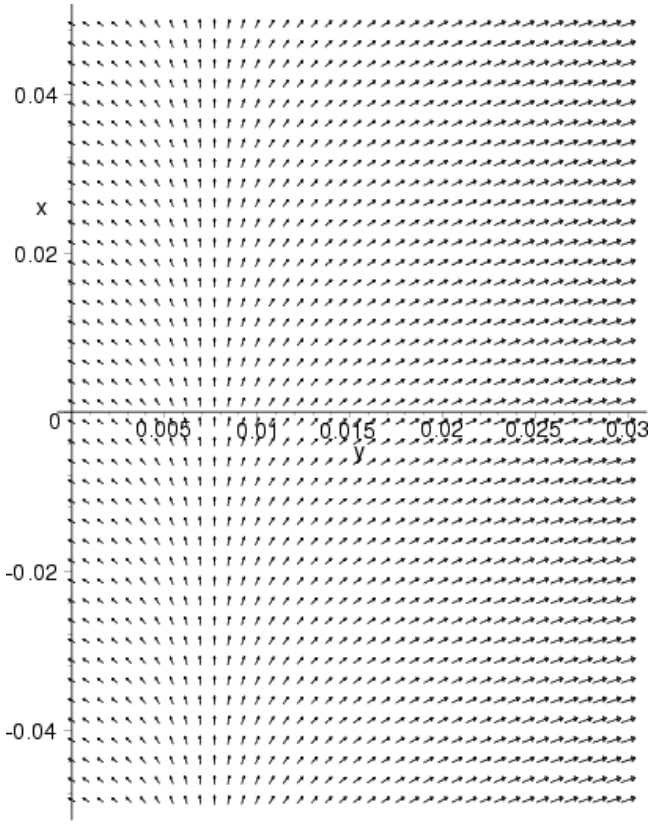

Figure 1: Drift vector b, $r_d < r_f$

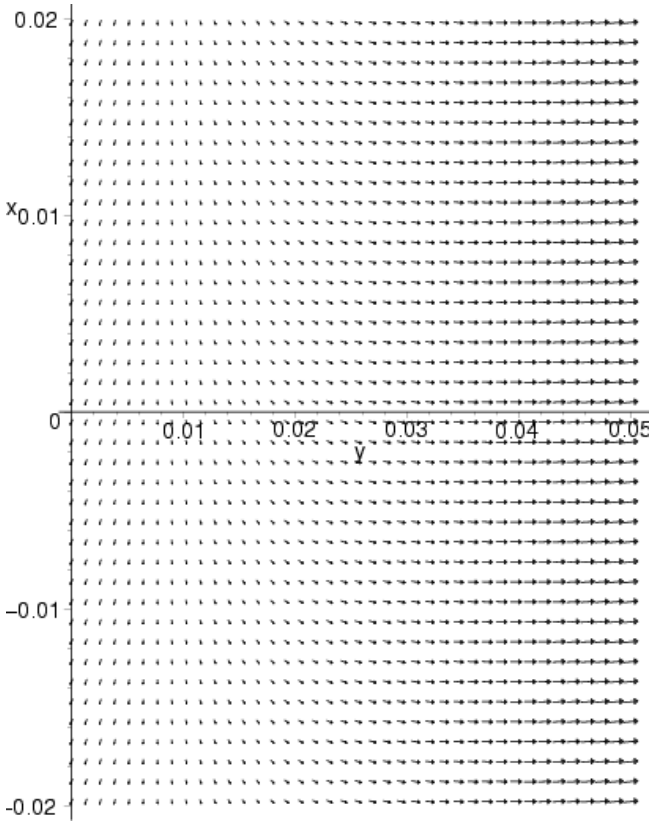

Figure 2: Drift vector b, $r_d > r_f$

Definition 1 *Let $\Omega \subset \mathbf{R}^d$ be a bounded open domain and u and w be two functions. Let the $d \times d$-matrix $A(x)$ be symmetric and positive definite for all $x \in \Omega$ and let the constant $\hat{c}$ be positive. Then we define*

$$(u, w)_A = \int_\Omega A\nabla u \cdot \nabla w + \int_\Omega \hat{c}uw \,, \tag{38}$$

$$\|u\|_A^2 = (u, u)_A = \int_\Omega A\nabla u \cdot \nabla u + \int_\Omega \hat{c}u^2 \,. \tag{39}$$

Lemma 1 *The bilinear form $(u, w)_A$ defines a scalar product and satisfies the* Cauchy-Schwarz *inequality (CSI). $\|u\|_A$ defines a norm.*

The proof is straight forward. Obviously $a(\cdot, \cdot)$ is linear in both components. The matrix A is positive definite, therefore $(u, u)_A = 0$ implies $u = 0$.
Remarks

1. The constant $\hat{c}$ will be taken to be $c - \frac{1}{2}\nabla \cdot b$ in the sequel (see (44)).

2. If $\hat{c} = \tilde{c} - \frac{1}{2}\nabla \cdot b$ we use the symbol $(\cdot, \cdot)_{\tilde{A}}$.

3. Obviously, norms $\|\cdot\|_A$, which only differ through a different choice of the constant $\hat{c} > 0$, are all equivalent.

Definition 2 *We define the spaces*

$$V = \{\psi : \|\psi\|_A < \infty\} \,, \tag{40}$$

$$V_0 = \left\{\psi \in V : \psi = 0 \text{ on } \Gamma_1 \stackrel{\Delta}{=} (\Gamma_a \cup \Gamma_b \cup \Gamma_c)\right\} \,, \tag{41}$$

$$V_* = \left\{\psi \in V : \psi \text{ satisfies the boundary conditions on } \Gamma_1\right\} . \tag{42}$$

Now we are ready to pose the problem in variational formulation.
We multiply both sides of Equation (29) with a test function ψ taken from the space V_0, integrate over the domain Ω and get at the left hand side

$$\int_\Omega Lw\psi = \int_\Omega -\sigma\nabla \cdot A\nabla w\psi + \sigma b \cdot \nabla w\psi + cw\psi, \tag{43}$$

where

$$c = \sigma r_d + \frac{1}{\tau}. \tag{44}$$

We note that only b depends on v. We obtain the right hand side analogously,

$$\int_\Omega f\psi = \int_\Omega (1-\sigma)\nabla \cdot A\nabla w\psi - (1-\sigma)b \cdot \nabla w\psi - \tilde{c}w\psi, \tag{45}$$

$$\tilde{c} = (1-\sigma)r_d - \frac{1}{\tau}. \tag{46}$$

Integration by parts yields

$$\begin{aligned}
\int_\Omega Lw\psi &= \sigma\int_\Omega A\nabla w\cdot\nabla\psi + \frac{\sigma}{2}\int_\Omega (b\cdot\nabla w\psi - wb\cdot\nabla\psi) \\
&+ \int_\Omega (c-\frac{\sigma}{2}\nabla\cdot b)w\psi - \sigma\int_{\partial\Omega} A\nabla w\cdot\vec{n}\psi + \frac{\sigma}{2}\int_{\partial\Omega}(b\cdot\vec{n})w\psi\,, \quad (47)\\
\int_\Omega f\psi &= -(1-\sigma)\int_\Omega A\nabla w\cdot\nabla\psi - \frac{1-\sigma}{2}\int_\Omega (b\cdot\nabla w\psi - wb\cdot\nabla\psi) \quad (48)\\
&- \int_\Omega (\tilde{c}-\frac{1-\sigma}{2}\nabla\cdot b)w\psi \\
&+ (1-\sigma)\int_{\partial\Omega} A\nabla w\cdot\vec{n}\psi - \frac{1-\sigma}{2}\int_{\partial\Omega}(b\cdot\vec{n})w\psi\,.
\end{aligned}$$

Note that the first term is computed at time $t = t^k$ and the second one at time $t = t^{k+1}$, cf (29).
Next we evaluate the boundary integrals using the given boundary conditions. Plugging this into Equation (47) leads to

$$\int_{\partial\Omega} A\nabla w\cdot\vec{n}\psi = \int_{\Gamma_d} A\nabla w\cdot\begin{bmatrix}0\\1\end{bmatrix}\psi\,. \quad (49)$$

Along the boundary Γ_a we have $v = v_{min}$; if $v_{min} = 0$ we obtain the boundary condition by plugging $v = 0$ into the differential equation. However, in this case we do not need to define a boundary condition, because $A = 0$ and all boundary integrals vanish. Along the boundaries Γ_a, Γ_b and Γ_c we have $\psi = 0$.
On the boundary Γ_d the conormal derivative is known, whence this integral can be moved to the right hand side of the variational equality. The convection dependent boundary integral yields

$$\int_{\partial\Omega}(b\cdot\vec{n})w\psi = \int_{\Gamma_d}\left[\frac{1}{2}v - (r_d - r_f) + \frac{1}{2}\xi\rho\right]w\psi\,. \quad (50)$$

3.7 Existence and uniqueness

After all preparations we are finally able to formulate the problem:
Task: Search for each $k = N-1, N-2, \ldots, 0$ a function $w^k = w(t^k, v, x) \in V_*$ such that for all $\psi \in V_0$

$$a(w^k,\psi) = \langle F,\psi\rangle \quad (51)$$

holds, where

$$\begin{aligned}
a(w^k,\psi) &= \int_\Omega A\nabla w^k\cdot\nabla\psi + \frac{1}{2}\int_\Omega (b\cdot\nabla w^k\psi - w^k b\cdot\nabla\psi) \\
&+ \int_\Omega (c-\frac{1}{2}\nabla\cdot b)w^k\psi\frac{1}{2} + \int_{\Gamma_d}(\frac{1}{2}v - (r_d-r_f) + \frac{1}{2}\xi\rho)w^k\psi\,. \quad (52)\\
\langle F^k,\psi\rangle &= f(t^{k+1};\psi) + \int_{\Gamma_d}\underbrace{A\nabla w^k\cdot\begin{bmatrix}0\\1\end{bmatrix}}_{g_2}\psi\,, \quad (53)\\
f(t^{k+1};\psi) &= -\int_\Omega A\nabla w^{k+1}\cdot\nabla\psi - \int_\Omega (b\cdot\nabla w^{k+1})\psi - \int_\Omega(\tilde{c}-\frac{1}{2}\nabla\cdot b)w^{k+1}\psi\,.
\end{aligned}$$

Here we use the constant $\tilde{c}$ from Equation (46). We remark

1. A is positive definite for all $v > 0$.
2. If $v_{min} > 0$, then A is positive definite on $\overline{\Omega}$ and hence regular and we have $V = H^1(\Omega)$.
3. If $v_{min} = 0$, then A becomes a zero matrix on the boundary Γ_a. Therefore, unbounded functions also belong to V which are not contained in $H^1(\Omega)$. This means that V is a weighted Sobolev space.

Hence, without loss of generality we assume that $\hat{c} := (c - \frac{1}{2}\nabla \cdot b)$ is equal to zero.

Theorem 1 *The equation*

$$a(w, \psi) = \langle F, \psi \rangle \qquad \forall \psi \in V_0 \tag{54}$$

along with the bilinear functional a taken from Equation (52) and the linear functional F taken from Equation (53) has a unique solution $w \in V_*$, if $r_d \leq r_f + \frac{1}{2}\xi\rho$ and $c - \frac{1}{2}\nabla \cdot b > 0$.

Proof. The Lemma of Lax and Milgram implies the existence and the uniqueness of Equation (54), provided $a(\cdot, \cdot)$ is a V_0-elliptic and V_0-bounded operator, $\langle F, \psi \rangle \leq \|F\|_{V_0^*} \|\psi\|_A \, \forall \, \psi \in V_0$ and additionally there exists a continuous extension of the boundary contitions on the entire domain. We will now show these properties.
We consider first $a(\psi, \psi)$ for $\psi \in V_0$,

$$a(\psi, \psi) \quad = \quad \int_\Omega A\nabla\psi \cdot \nabla\psi + \frac{1}{2}\int_{\Gamma_d} (\frac{1}{2}v - (r_d - r_f) + \frac{1}{2}\xi\rho)\psi^2. \tag{55}$$

We examine the case $r_d \leq r_f + \frac{1}{2}\xi\rho$, i.e. $-(r_d - r_f) + \frac{1}{2}\xi\rho > 0$.
The matrix A is positive definite for $v > 0$ and the constant c is above zero. Thus

$$a(\psi, \psi) \quad \geq \quad C\|\psi\|_A^2. \tag{56}$$

This shows the V_0-ellipticity. We now show the V_0-boundedness, which is a consequence of the Cauchy-Schwarz inequality.

$$\begin{aligned} a(w, \psi) &= \int_\Omega A\nabla w \cdot \nabla\psi + \int_\Omega (b \cdot \nabla w)\psi \\ &= (w, \psi)_A + \int_\Omega (A^{-1/2} b \cdot A^{1/2}\nabla w)\psi \\ &\leq \|w\|_A \|\psi\|_A + \left(\int_\Omega |A^{-1/2} b \cdot A^{1/2}\nabla w|^2 \int_\Omega \psi^2 \right)^{1/2} \\ &\leq \|w\|_A \|\psi\|_A + \left(\|A^{-1/2} b\|_{0,\infty,\Omega}^2 \int_\Omega |A^{1/2}\nabla w|^2 \int_\Omega \psi^2 \right)^{1/2} \end{aligned}$$

Since A is regular for all $v \geq v_{min} > 0$ and $v < v_{max}$, the norm is finite. Both the integrals are bounded by the norm $\|\cdot\|_A$. Therefore,

$$\begin{aligned} a(w, \psi) &\leq C\|w\|_A \|\psi\|_A \\ C &= 1 + C_A \|A^{-1/2} b\|_{0,\infty,\Omega}. \end{aligned} \tag{57}$$

The constant C_A depends on c. This shows the V_0-boundedness.
We now consider the right hand side of the partial differential equation.

$$\begin{aligned}
\langle F, \psi \rangle &= -(w^{k+1}, \psi)_{\tilde{A}} - \int_\Omega (b \cdot \nabla w^{k+1})\psi + \int_{\Gamma_d} g_2 \psi \qquad (58)\\
&\leq \|w^{k+1}\|_{\tilde{A}} \|\psi\|_{\tilde{A}} + \|A^{-1/2} b\|_{0,\infty,\Omega} \|w^{k+1}\|_{\tilde{A}} \|\psi\|_A + \int_{\Gamma_d} g_2 \psi \\
&\leq \|w^{k+1}\|_{\tilde{A}} \|\psi\|_{\tilde{A}} + \|A^{-1/2} b\|_{0,\infty,\Omega} \|w^{k+1}\|_{\tilde{A}} \|\psi\|_A + \|g_2\|_{0,\Gamma_d} \|\psi\|_{0,\Gamma_d} \\
&\leq \|w^{k+1}\|_{\tilde{A}} \|\psi\|_{\tilde{A}} + \|A^{-1/2} b\|_{0,\infty,\Omega} \|w^{k+1}\|_{\tilde{A}} \|\psi\|_A + C\|g_2\|_{0,\Gamma_d} \|\psi\|_A \\
&\leq C\|\psi\|_A
\end{aligned}$$

The constant C is now different from the previous one.
While choosing the boundary conditions we already made sure the Dirichlet-type boundary conditions are continuous along the corners of the domain. Since the domain is a rectangle, its boundary is piecewise smooth, whence a continuous extension of the boundary values to the entire domain Ω is possible. This completes the proof.

We note the following remarks.

1. If the interest rate condition is not fullfilled we can exchange the currencies.
2. If c is too small we can either decrease the time step τ or transform the variable, because the solvability of our problem does not change if we substitute $w(t, v, x)$ by $z(t, v, x) = w(t, v, x)e^{\alpha t}$. Hence, without loss of generality we can assume that $\hat{c} \triangleq (c - \frac{1}{2}\nabla \cdot b)$ is equal to zero.
3. Equation (57) shows that the constant of boundedness depends on the size of the drift b, particularly on the speed of 'mean reversion', and on the lower bound of the variance v. For large κ we have roughly

$$\|A^{-1/2} b\|^2_{0,\infty,\Omega} \approx \operatorname{ess\,sup} \left\{ \frac{1}{\sqrt{v}} (\kappa + \lambda) v \right\}^2 = (\kappa + \lambda)^2 v_{max} \,. \qquad (59)$$

4 Numerical solution

In this section we discuss our implementation of the finite element method for the variational problem (51) of Section 3. We switch to the standard notation in the literature, i.e., for points in $\mathbf{R}^2$ we state (x, y) instead of (v, x). Further information for implementation of finite element methods can be found in [7].

4.1 The basic idea of the finite element method

The finite element method usually starts with the variational formulation (51), i.e., in each time step we have to solve a problem of the form

$$a(w, v) = \langle F, v \rangle \qquad \forall v \in V_0 \,. \qquad (60)$$

Thus, we are looking for a function $w \in V$ satisfying the given integral equation for arbitrary test functions $v \in V_0$.

The region will be decomposed into a mesh $\mathcal{T}_h$ of triangles[5] or quadrilaterals. Adjacent triangles must share a common edge and the mesh must cover the region completely. Along with the triangles we choose nodes, whose number determines the degree of the polynomial basis functions. For linear functions three nodes per triangle suffice, namely the corners of the triangle, for quadratic functions one needs six nodes for each triangle, namely the corners and the midpoints of the edges. On this mesh we define a basis function for each node, which is piecewise polynomial and takes the value *one* at this node and the value *zero* on all the other nodes. The basis functions form a basis of the space V_h. The space V_{0h} is then defined by $V_{0h} = V_h \cap V_0$, that means the basis functions corresponding to nodes at the Dirichlet boundary are omitted. A finite element solution of the homogeneous task[6] is a function $w_h \in V_{0h}$ which satisfies the integral equation for all $v \in V_{0h}$. Together with a continuous extension of the Dirichlet-type boundary conditions on the entire region, one can state the solution of the task.

4.2 Mesh and basis functions

We decompose the region $\Omega = (v_{min}, v_{max}) \times (x_{min}, x_{max})$ into a mesh $\mathcal{T}_h$ consisting of (open) triangles T_i, $i = 1, \ldots, N_T$ (*triangle*). We label the resulting vertices by v_i, $i = 1, \ldots, N_v$ (*vertex*). Then we define the basis functions φ_i, $i = 1, \ldots, N_v$, which are piecewise linear, take the value 1 at vertex v_i and 0 at all other vertices. These functions form a basis of V_h. In order to simplify the computation we state for each triangle T_i a transformation $F : \mathbf{R}^2 \to \mathbf{R}^2$ on the unit triangle $\hat{T} = \{(\hat{x}, \hat{y}) \in \mathbf{R}^2 : 0 \le \hat{x} \le 1 \text{ and } 0 \le \hat{y} \le 1 - \hat{x}\}$. All integral computations will be reduced to integrals of the transformed basis functions $\hat{\varphi}_i$.

4.2.1 Transformations

Given the triangle T with vertices $v_0 = (x_0, y_0)$, $v_1 = (x_1, y_1)$ and $v_2 = (x_2, y_2)$. The transformation

$$F(\hat{x}, \hat{y}) = M \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}, \tag{61}$$

$$M = \begin{bmatrix} x_1 - x_0 & x_2 - x_0 \\ y_1 - y_0 & y_2 - y_0 \end{bmatrix}, \tag{62}$$

maps the unit triangle $\hat{T}$ onto the given triangle. The (affine) mapping F ist invertible. In the sequel we denote functions defined on the unit triangle with a hat.
Thus we have $f(x, y) = \hat{f}(\hat{x}, \hat{y})$, whenever $(x, y) = F(\hat{x}, \hat{y})$. In particular, the following three equations hold.

$$\hat{f}(0, 0) = f(x_0, y_0)\,, \tag{63}$$

$$\hat{f}(1, 0) = f(x_1, y_1)\,, \tag{64}$$

$$\hat{f}(0, 1) = f(x_2, y_2)\,. \tag{65}$$

[5] For higher dimensional regions one uses tetrahedra etc.
[6] All Dirichlet-type boundary conditions are homogeneous.

Now we examine the integral of a function $f(x,y)$ on the given triangle. We have

$$\begin{aligned}\int_T f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_{\hat{T}} f(F(\hat{x},\hat{y}))\left|\frac{\partial F(\hat{x},\hat{y})}{\partial(x,y)}\right|\,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y}\\ &= |\det M|\int_{\hat{T}}\hat{f}(\hat{x},\hat{y})\,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y}\,. \end{aligned} \tag{66}$$

For the integral of the gradient of f we have

$$\begin{aligned}&\int_T b(x,y)\cdot\nabla f(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_{\hat{T}} b(F(\hat{x},\hat{y}))\cdot(\,\mathrm{J}F(\hat{x},\hat{y}))^{-T}\hat{\nabla}f(F(\hat{x},\hat{y}))\left|\frac{\partial F(\hat{x},\hat{y})}{\partial(x,y)}\right|\,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y}\\ &= |\det M|\int_{\hat{T}}\hat{b}(\hat{x},\hat{y})\cdot M^{-T}\hat{\nabla}\hat{f}(\hat{x},\hat{y})\,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y}\,. \end{aligned} \tag{67}$$

The symbol $\hat{\nabla}$ denotes $\begin{bmatrix}\partial/\partial\hat{x}\\ \partial/\partial\hat{y}\end{bmatrix}$. The Jacobi matrix of the function $F(\hat{x},\hat{y})$ will be denoted by $\mathrm{J}F(\hat{x},\hat{y})$. Analogously we obtain

$$\begin{aligned}&\int_T \nabla g(x,y)\cdot A(x,y)\nabla f(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_{\hat{T}}(\mathrm{J}F(\hat{x},\hat{y}))^{-T}\hat{\nabla}g(F(\hat{x},\hat{y}))\\ &\quad\cdot A(F(\hat{x},\hat{y}))(\mathrm{J}F(\hat{x},\hat{y}))^{-T}\hat{\nabla}f(F(\hat{x},\hat{y}))\left|\frac{\partial F(\hat{x},\hat{y})}{\partial(x,y)}\right|\,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y}\\ &= |\det M|\int_{\hat{T}}(M^{-T}\hat{\nabla}\hat{g})\cdot(\hat{A}M^{-T}\hat{\nabla}\hat{f})\,\mathrm{d}\hat{x}\,\mathrm{d}\hat{y}\,. \end{aligned} \tag{68}$$

4.2.2 Notation

We define the following basis functions on the unit triangle

$$\hat{\varphi}_0(\hat{x},\hat{y}) = 1-\hat{x}-\hat{y}, \tag{69}$$
$$\hat{\varphi}_1(\hat{x},\hat{y}) = \hat{x}, \tag{70}$$
$$\hat{\varphi}_2(\hat{x},\hat{y}) = \hat{y}. \tag{71}$$

The basis functions on the region Ω can be built from these. Let $I(v_j)$ be the set of indices of the triangles T_i, which have v_j as a vertex,

$$I(v_j) = \{i\in\{1,..,N_T\} : v_j\in\overline{T_i}\}\,. \tag{72}$$

Let F_i, $i\in I(v_j)$, be the affine transformations, which map the unit triangle onto T_i, where the point $(0,0)$ should be mapped onto v_j. Then we have

$$\varphi_j(x,y) = \begin{cases}\hat{\varphi}_0(F_i^{-1}(x,y)) \text{ if } (x,y)\in\overline{T_i},\ i\in I(v_j)\\ 0 \text{ otherwise}\end{cases}. \tag{73}$$

The hence defined functions φ_j form a basis of the space V_h of all basis functions on Ω.

4.2.3 Integration of linear functions

In this section we present formulae which help to compute integrals of linear functions as well as integrals of their products on the unit triangle. We will use these formulae in the sequel to compute the arising integrals on each triangle T_i of the finite element mesh $\mathcal{T}_h$.
Let an affine linear function $g : \mathbf{R}^2 \to \mathbf{R}$ be given. The integral of the product of f and g over the unit triangle is

$$\begin{aligned}
&\int_{\hat{T}} f(x,y)g(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{x=0}^{x=1}\int_{y=0}^{1-x} (ax+by+c(1-x-y))\,(\alpha x+\beta y+\gamma(1-x-y))\,\mathrm{d}y\,\mathrm{d}x \\
&= \frac{1}{24}\begin{bmatrix} a & b & c \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}. \qquad (74)
\end{aligned}$$

We see that the computation of the integrals is essentially a multiplication of matrices. The other integration formulae needed for computing $a(\cdot,\cdot)$ and $\langle F,\cdot\rangle$ are similar.

4.3 The finite element system of equations

4.3.1 Matrix notation

We plug the basis functions of the previous section into the variational equation (60) and solve the resulting system of linear equations.
Task: Find a function $w_h : \Omega \to \mathbf{R}$ satisfying

$$a(w_h,\psi_h) = \langle F,\psi_h\rangle \quad \forall \psi_h \in V_{0h} \qquad (75)$$

We write w_h in the form $w_h = \sum_{j=1}^{N_v} w_j\varphi_j$. Since the set $\{\varphi_i, i = 1\ldots N_v\}$[7] forms a basis of V_{0h} it suffices to solve the system

$$a\left(\sum_{j=1}^{N_v} w_j\varphi_j,\ \varphi_i\right) = \langle F,\varphi_i\rangle, \quad i = 1\ldots N_v. \qquad (76)$$

In matrix notation it looks like

$$\begin{aligned}
K_h\underline{w_h} &= \underline{f_h}, && (77)\\
K_h &= [a(\varphi_j,\varphi_i)]_{i,j=1\ldots N_v}, && (78)\\
\underline{w_h} &= [w_j]_{j=1\ldots N_v}, && (79)\\
\underline{f_h} &= [\langle F,\varphi_i\rangle]_{i=1\ldots N_v}. && (80)
\end{aligned}$$

When computing the elements of the stiffness matrix K_h we must ensure to use the correct indices, because we have $K_{ij} = a(\varphi_j,\varphi_i)$.

[7] Here N_v denotes the number of vertices excluding all vertices from Dirichlet-type boundaries.

4.3.2 Computation of the stiffness matrix

The basis function introduced in Section 4.2.2 will be used to compute the bilinear form $a(\cdot,\cdot)$. In the process we determine the nine combinations of i and j in one step for each triangle T,

$$\int_T A\nabla\varphi_j \cdot \nabla\varphi_i \quad = \quad |\det M| \int_{\hat{T}} (\hat{A}M^{-T}\hat{\nabla}\hat{\varphi}_j)^T (M^{-T}\hat{\nabla}\hat{\varphi}_i). \tag{81}$$

The gradients of the local basis functions $\hat{\varphi}_i$, $i = 1, 2, 3$, are constants, whence they can be moved in front of the integral. Writing the three gradients into a matrix

$$G \quad = \quad \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \tag{82}$$

we can represent the nine numbers in matrix notation as

$$\left[\int_T A\nabla\varphi_j \cdot \nabla\varphi_i\right]_{i,j=1,2,3} \quad = \quad |\det M| G^T M^{-1} \left(\int_{\hat{T}} \hat{A}\right) M^{-T} G. \tag{83}$$

The middle integral is computed componentwise. Since each element of $\hat{A}$ is a affine lineare function, we can compute the integral and obtain

$$\int_{\hat{T}} \hat{A} \quad = \quad \frac{1}{6}(A(x_0, y_0) + A(x_1, y_1) + A(x_2, y_2)). \tag{84}$$

Now we compute the next integral of the bilinear form using Equation (67). Since the gradient of $\hat{\varphi}_j$ is constant it can be moved in front of the integral. The convection b and $\hat{\varphi}_j$ are affine linear, so we can envoke Equation (74). Finally, we obtain a 3×3 - matrix

$$\begin{aligned} B \quad &\triangleq \quad \left[\int_{\hat{T}} (\hat{b}^T M^{-T}\hat{\nabla}\hat{\varphi}_j)\hat{\varphi}_i\right]_{i,j=1,2,3} \\ &= \quad \frac{1}{24} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} b^T(x_0, y_0) \\ b^T(x_1, y_1) \\ b^T(x_2, y_2) \end{bmatrix} M^{-T} G. \end{aligned} \tag{85}$$

Therefore, we have

$$\left[\int_T (b^T\nabla\varphi_j)\varphi_i - \varphi_j(b^T\nabla\varphi_i)\right]_{i,j=1,2,3} \quad = \quad |\det M|(B - B^T). \tag{86}$$

The third part of $a(\varphi_j, \varphi_i)$ can also be deduced from Equation (74).

5 Selected solutions

5.1 Results

In this section we discuss some numerical effects which arose with our implementation of the Finite Element Method.

Figures 3 and 4 show a solution, Table 1 shows the output of the program of the three meshes, where the third is the finest. The horizontal axes of the figures label the square v of the volatility and the logarithm of the ratio of spot and strike. This way, the prices for at-the-money calls are located exactly at $x = 0$. The third dimension shows the price.

Choosing a non-uniform mesh improves the results significantly, since the transformation of the v-axis with a function of the kind $\frac{1}{v}$, exactly $\hat{v} = \frac{\alpha(1+\alpha)}{v+\alpha} - \alpha$, $\alpha = \frac{1}{2}$ compensates the drift growing proportionally to v fairly well. Some experiments with different transformations showed that this particular function works best.

In Figure 3 we show the solution with a coarse mesh of 17×65 points. Notice the mark in the middle indicating negative values. The actual figures show a global minimum of -0.00091798. This is a good indication of the size of the numerical error one can expect, namely in the third digit after the decimal point.

The closed-form solution in [5] yields a price of a call of 0.044943966 for an initial variance $v_0 = 0.05225$, the solution from Figure 3 and the second column of Table 1 is 0.0414271. This difference is not within the bid-ask spread, but the result can be improved using a finer mesh. The actual numbers in Figure 4 and the third column of Table 1 show a price of 0.0439985 which is significantly closer to the target value computed with the closed form solution. Looking at the global minima we observe that one more digit can be considered reliable.

Table 2 summarizes the results for the price of a call using different methods and the corresponding calculation times.

The computation time using FEM grows roughly by a factor of 8, when we refine the mesh once. Here we used MATLAB, version 6, release 12 on a Pentium III with 933 MHz under Linux. Converting the code into a compiler language such as C/C++ normally allows a speed-up by a factor around 10–20. Using adaptive meshes or existing optimized FEM solvers will allow taking fewer triangles without diminishing the precision.

5.2 Discussion of the errors

As we have seen in our solution we need to make the meshes finer to gain more precise solutions. Particularly the region, where the drift is small and changes the direction (see Figures 1 and 2) requires many gridlines. This can be achieved either using an a-priori set mesh as in our implementation or – more effectively – using adaptive meshes. We noticed furthermore that there should be significantly more nodes on the spot axis than on the variance axis, because the drift in spot direction is small compared to the one in the variance direction and because the domain in spot direction is much larger than in volatility direction. This effect is – to some extend – compensated by the diffusion, but must not be neglected specially in the case of small diffusion (ξ).

The negative values for the price of the option also vanish, if the mesh at the respective zone is made finer. It is particularly possible to take more nodes along the spot axis in this zone or to choose a completely different approach that preserves properties like positivity.

We observed large errors specially for very small values of ξ, which spread wave-like over the entire solution. To limit such an effect one can possibly use a more refined finite element method, which seek a solution for a slightly modified

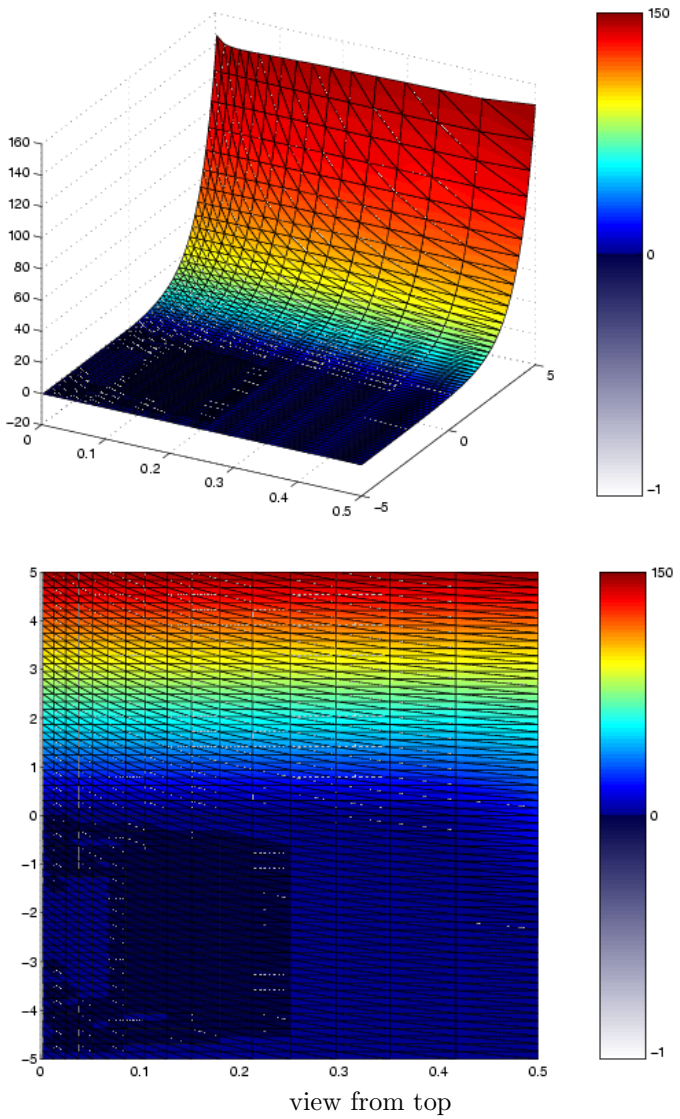

Figure 3: Solution for: $\xi = 0.5$, $\rho = -0.1$, $\lambda = 0.0$, $\kappa = 2.5$, $\theta = 0.06$, $r_f = \log(1.048)$, $r_d = \log(1.052)$, $K = 1$, $T = 0.25$ in the domain $[0.0025, 0.5] \times [-5, 5]$ with discretization $\sigma = 0.5$, $\tau = 0.025$, $n = 17 \times 65$
colors: negative (white) ... zero (black/blue) ... positive (red)

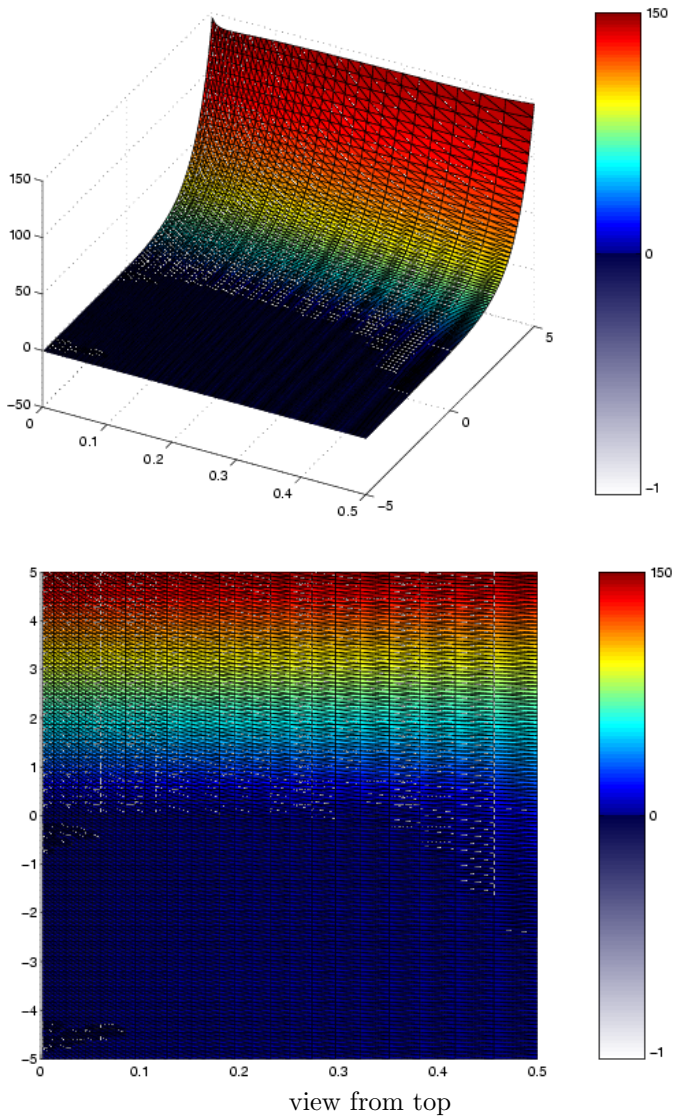

Figure 4: Solution for: $\xi = 0.5$, $\rho = -0.1$, $\lambda = 0.0$, $\kappa = 2.5$, $\theta = 0.06$, $r_f = \log(1.048)$, $r_d = \log(1.052)$, $K = 1$, $T = 0.25$ in the domain $[0.0025, 0.5] \times [-5, 5]$ with discretization $\sigma = 0.5$, $\tau = 0.025$, $n = 33 \times 129$
colors: negative (white) ... zero (black/blue) ... positive (red)

	mesh 1	mesh 2	mesh 3
global max	146.684	146.684	146.684
global min	-0.00091798	-0.000108395	-1.35879e-17
calculation time	65.4466 sec	326.928 sec	2521.6 sec
v=0.0025, K=S	0.01033	0.01033	0.01033
v=0.00511842, K=S			0.0174284
v=0.00779255, K=S		0.0207066	0.021529
v=0.0105242, K=S			0.0245046
v=0.0133152, K=S	0.0235529	0.02603	0.0268957
v=0.0161676, K=S			0.0289403
v=0.0190833, K=S		0.0299348	0.0307599
v=0.0220646, K=S			0.0324243
v=0.0251136, K=S	0.0304918	0.0332053	0.0339769
v=0.0282328, K=S			0.0354468
v=0.0314244, K=S		0.0361357	0.036854
v=0.0346912, K=S			0.0382132
v=0.0380357, K=S	0.0361731	0.038865	0.0395352
v=0.0414608, K=S			0.0408283
v=0.0449695, K=S		0.0414701	0.0420988
v=0.0485648, K=S			0.0433519
v=0.05225, K=S	0.0414271	0.0439985	0.0445917
v=0.0560285, K=S			0.0458217
v=0.0599038, K=S		0.0464822	0.0470446
v=0.0638799, K=S			0.0482629
v=0.0679605, K=S	0.0465361	0.0489438	0.0494787

Table 1: Solution for: $\xi = 0.5$, $\rho = -0.1$, $\lambda = 0.0$, $\kappa = 2.5$, $\theta = 0.06$, $r_f = \log(1.048)$, $r_d = \log(1.052)$, $K = 1$, $T = 0.25$ in the domain $[0.0025, 0.5] \times [-5, 5]$ with discretization $\sigma = 0.5$, $\tau = 0.025$

Method	Result	Computation Time
Black-Scholes ($\sigma = \sqrt{v_0}$)	0.04549	0
Heston, analytic	0.04494	0
FEM, Mesh 1	0.04142	65s
FEM, Mesh 2	0.04400	327s
FEM, Mesh 3	0.04459	2522s

Table 2: Comparison of call prices in Heston's model using different meshes

variational equality.

5.3 Exotic options

We have seen how to price a call option in a stochastic volatility model using finite elements. Extensions to other path-independent options are now straightforward. A put option for example can be computed using the put-call-parity or by switching the condition on the boundaries Γ_c and Γ_d.

The method also allows a simple extension to price barrier options. For instance, to price a *down-and-out call* it will require only a change in the condition of the lower boundary Γ_c, namely setting the value equal to the respective rebate. This is even easier than a vanilla call, because we know where the boundary should be chosen and we know the value of the solution a-priori. However, we need to adapt the boundary conditions on Γ_a and Γ_b in such a way that there are no jumps in the corners of the region, lest it will be difficult to guarantee the existence of a solution. A continuous extension on the boundary of the region might then cease to exist.

Options with discountinuous payoffs, such as digital options or reverse knock-out options, are problematic, because the boundary condition at expiry $t = T$ is not contained in H^1. The smoothing effect of the diffusion term of the partial differential equation guarantees the solution to be contained in H^1 for all times $t < T$.

References

[1] Belledin, M. and Schlag, C. (1999). *An Empirical Comparison of Alternative Stochastic Volatility Models*. Working paper, Goethe-University Frankfurt, Department of Economics and Business Administration.

[2] Cox, J.C., Ingersoll, J.E. and Ross, S.A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica* **53**, 385-407.

[3] Flannery, B., Press, W., Teukolsky, S. and Vetterling, W. (1992). *Numerical Recipes in C*. . Cambridge University Press.

[4] Gulko, L. (1999). The Entropy Theory of Stock Option Pricing. *International Journal of Theoretical and Applied Finance*, Vol. 2, No. 3. pp 331-355.

[5] Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, Vol. 6, No. 2.

[6] Hull, J. and White, A. *Hull-White on Derivatives*. Risk Publications, London

[7] MacPhedran, I. and Yound, R. (2001). *Internet Finite Element Resources*. http://www.engr.usask.ca/~macphed/finite/fe_resources/fe_resources.html

[8] Kufner, A. (1980). *Weighted Sobolev Spaces*. Teubner, Leipzig.

[9] Kufner, A. and Sändig, A.-M. (1987). *Some Applications of Weighted Sobolev Spaces*. Teubner, Leipzig.

[10] Kurpiel, A. and Roncalli, T. (2000). Hopscoch Methods for Two-State Financial Models. *The Journal of Computational Finance*. Vol 3, Number 2, pp. 53-89.

[11] Melino, A. and Turnbull, S. M. (1990). Pricing Foreign Currency Options with Stochastic Volatility. *Journal of Econometrics*, 45, pp. 239-265.

[12] Roos, H.-G., Stynes, M. and Tobiska, L. (1996). *Numerical Methods for Singularly Perturbed Differential Equations*. Springer.

[13] Wystup, U. et al. *MathFinance*. http://www.mathfinance.de