Valuation of Double Barrier Options in Heston’s Stochastic Volatility Model

Using Finite Element Methods

Copyright © Changwei Xiong 2010 - 2018

Jan. 2010

last update: June 18, 2018
ABSTRACT

Heston stochastic volatility model has been widely used in financial derivative pricing and risk management. One of the reasons for that is that vanilla options in Heston model have close form solutions. This makes the calibration of the model computationally much efficient and accurate. To understand these closed form formulas, this essay in the first half will introduce the Fourier transform method for option pricing and its application to the Heston model. Additionally, a characteristic function based method is also discussed, which extends the Heston model to have piecewise time-dependent parameters.

Finite element method (FEM) has been developed for decades to solve partial differential equations arise from science and engineering problems. It is well known for requiring a low order storage and for its capability to handle complicated irregular computational domains comparing to finite difference method (FDM). This characteristic advantage makes FEM an ideal numerical method for valuation of exotic options. This essay in the second half will present a detailed implementation of FEM and its application to pricing a double barrier knock-out option with an underlying asset modeled by Heston stochastic volatility process.
# Table of Contents

Abstract ................................................................................................................................. 2  
Table of Contents .................................................................................................................. 3  
1. Kolmogorov Forward and Backward Equations ............................................................. 5  
   1.1. Kolmogorov Forward Equation .................................................................................. 5  
   1.2. Kolmogorov Backward Equation ............................................................................. 8  
2. Heston Model .................................................................................................................. 10  
   2.1. Heston Stochastic Volatility Model ........................................................................ 10  
      2.1.1. Market Price of Risk ......................................................................................... 10  
      2.1.2. Radon-Nikodym Derivative ............................................................................. 15  
      2.1.3. Feller Condition ............................................................................................... 16  
   2.2. Probability Distribution of Spot Returns .................................................................. 16  
      2.2.1. Derivation of the Transition Probability ............................................................. 17  
      2.2.2. Moment Generating Function (in progress...) .................................................... 20  
2.3. Analytical Solution of Vanilla Options ...................................................................... 21  
      2.3.1. Fourier Transform ............................................................................................ 21  
      2.3.2. Levy’s Inversion Formula ................................................................................ 23  
      2.3.3. Characteristic Function .................................................................................... 25  
      2.3.4. Vanilla Option Prices ....................................................................................... 29  
         2.3.4.1. Analogy to Cumulative Density Function ..................................................... 30  
         2.3.4.2. Analogy to Probability Density Function ..................................................... 32  
         2.3.4.3. Heston’s Original Solution ......................................................................... 33  
2.4. Piecewise Time Dependent Heston Model ................................................................ 36  
3. Heston Model: PDE by Finite Element Method .............................................................. 38  
   3.1. The Partial Differential Equation ............................................................................. 39  
   3.2. Numerical Solution of the PDE ............................................................................... 42  
      3.2.1. Temporal Discretization .................................................................................... 42  
      3.2.2. Two-Dimensional Finite Element Method ........................................................ 43  
         3.2.2.1. Weak Formulation ....................................................................................... 43  
         3.2.2.2. Boundary Conditions .................................................................................. 44  
         3.2.2.3. Mesh and Basis Functions ........................................................................... 46  
         3.2.2.4. Affine Transformation ................................................................................ 48  
         3.2.2.5. Stiffness and Mass Matrix .......................................................................... 48  
         3.2.2.6. Computation of Integrals .......................................................................... 51  
   3.2.3. Iterative Linear Solvers ....................................................................................... 53  
      3.2.3.1. Implementation of Sparse Matrix ................................................................. 53  

Changwei Xiong, June 2018
3.2.3.2. Iterative Methods for Asymmetric Sparse Matrix ........................................ 54
3.2.4. Interpolation of Numerical Solution .................................................................. 54
3.3. Solutions and Discussions ......................................................................................... 55
  3.3.1. Vanilla Call ........................................................................................................... 55
  3.3.2. Vanilla Put ........................................................................................................... 58
  3.3.3. Double Barrier Knock-out Call ........................................................................... 59
  3.3.4. Double Barrier Knock-out Put ........................................................................... 61
  3.3.5. Numerical Solutions of Double Barrier Knock-out Options ............................... 63
3.4. Conclusions .................................................................................................................. 64
References .......................................................................................................................... 65
This note is to summarize Heston stochastic volatility model. It was started as my degree essay for the M.S. in Mathematical Finance, which only covered the barrier option pricing by solving PDE using finite element method, i.e. the Chapter 3 in this note. I have been continuously expanding it with more mathematical background, such as the derivation of market price of spot/volatility risk, the Fourier transform method for option pricing, the derivation of characteristic function of the joint spot-variance process, the probability distribution of spot return, the piecewise time dependent Heston parameters, etc.

At first, let’s start from the Kolmogorov forward and backward equations, which fundamentally govern the transition probability density function of a diffusion process.

1. **Kolmogorov Forward and Backward Equations**

The time evolution of the transition probability density function is governed by Kolmogorov forward and backward equations, which will be introduced as follows, without loss of generality, in multi-dimension.

1.1. **Kolmogorov Forward Equation**

Let’s consider the following $m$-dimensional stochastic spot process $S_t \in \mathbb{R}^m$ driven by an $n$-dimensional Brownian motion $W_t$ whose correlation matrix $\rho$ is given by $\rho dt = dW_t dW_t'$

$$dS_t = A(t, S_t) \, dt + B(t, S_t) \, dW_t,$$

We derive the dynamics of $h$, where $h: \mathbb{R}^m \rightarrow \mathbb{R}$ in this case is a scalar-valued Borel-measurable function only on variable $S_t$

$$dh(S_t) = J_h \, dS_t + \frac{1}{2} H_h \, dS_t dS_t = J_h \, Adt + J_h \, BdW_t + \frac{1}{2} dW_t' B' H_h \, BdW_t$$

where $J_h$ is the $1 \times m$ Jacobian (i.e. the same as gradient if $h$ is a scalar-valued function) and $H_h$ the $m \times m$ Hessian (with subscripts of $S$ now denoting the indices of vector components)

$$[J_h]_i = \frac{\partial h}{\partial S_i} \quad \text{and} \quad [H_h]_{ij} = \frac{\partial^2 h}{\partial S_i \partial S_j}$$

Expanding the expression in (2), we have
\[ dh = \sum_{i=1}^{m} \frac{\partial}{\partial S_i} A_i dt + \sum_{i=1}^{m} \frac{\partial}{\partial S_i} \sum_{k=1}^{n} B_{ik} dW_k + \frac{1}{2} \sum_{i,j=1}^{m} \frac{\partial^2 h}{\partial S_i \partial S_j} \sum_{k=1}^{n} B_{ik} \rho_{ij} B_{jk} dt \]

\[ dh = \left( \sum_{i=1}^{m} \frac{\partial}{\partial S_i} A_i + \frac{1}{2} \sum_{i,j=1}^{m} \frac{\partial^2 h}{\partial S_i \partial S_j} \Sigma_{ij} \right) dt + \sum_{i=1}^{m} \frac{\partial h}{\partial S_i} \sum_{k=1}^{n} B_{ik} dW_k \]

where \( \Sigma = B \rho B' \) is the \( m \times m \) instantaneous variance-covariance matrix of \( dS \). Integrating on both sides of (4) from \( t \) to \( T \), we have

\[ h(S_T) - h(S_t) = \int_t^T \left( \sum_{i=1}^{m} \frac{\partial}{\partial S_i} A_i + \frac{1}{2} \sum_{i,j=1}^{m} \frac{\partial^2 h}{\partial S_i \partial S_j} \Sigma_{ij} \right) du + \int_t^T \sum_{i=1}^{m} \frac{\partial h}{\partial S_i} \sum_{k=1}^{n} B_{ik} dW_k \]

(5)

Taking expectation on both sides of (5), we get (using notation \( \mathbb{E}_t[\cdot] = \mathbb{E}[\cdot | F_t] \))

\[ \text{LHS} = \mathbb{E}_t[h(S_T)] - h(S_t) = \int_{\Omega} h_y p_{T,y|t,x} dy - h_x \]

\[ \text{RHS} = \mathbb{E}_t \left[ \int_t^T \left( \sum_{i=1}^{m} \frac{\partial}{\partial S_i} A_i + \frac{1}{2} \sum_{i,j=1}^{m} \frac{\partial^2 h}{\partial S_i \partial S_j} \Sigma_{ij} \right) du \right] + \mathbb{E}_t \left[ \int_t^T \sum_{i=1}^{m} \frac{\partial h}{\partial S_i} \sum_{k=1}^{n} B_{ik} dW_k \right] \]

\[ = \int_t^T \sum_{i=1}^{m} \mathbb{E}_t \left[ \frac{\partial}{\partial S_i} A_i \right] du + \frac{1}{2} \int_t^T \sum_{i,j=1}^{m} \mathbb{E}_t \left[ \frac{\partial^2 h}{\partial S_i \partial S_j} \Sigma_{ij} \right] du \]

(6)

where \( p_{T,y|t,x} \) is the transition probability density having \( S_T = y \) at \( T \) given \( S_t = x \) at \( t \) (i.e. if we solve the equation (1) with the initial condition \( S_t = x \in \mathbb{R}^m \), then the random variable \( S_T = y \in \Omega \) has a density \( p_{T,y|t,x} \) in the \( y \) variable at time \( T \)). Differentiating (6) with respect to \( T \) on both sides, we have

\[ \int_{\Omega} h_y \frac{\partial p_{T,y|t,x}}{\partial T} dy = \sum_{i=1}^{m} \mathbb{E}_t \left[ \frac{\partial}{\partial y_i} A_i \right] + \frac{1}{2} \sum_{i,j=1}^{m} \mathbb{E}_t \left[ \frac{\partial^2 h}{\partial y_i \partial y_j} \Sigma_{ij} \right] \]

\[ = \sum_{i=1}^{m} \int_{\Omega} \frac{\partial h_y}{\partial y_i} A_i p_{T,y|t,x} dy + \frac{1}{2} \sum_{i,j=1}^{m} \int_{\Omega} \frac{\partial^2 h_y}{\partial y_i \partial y_j} \Sigma_{ij} p_{T,y|t,x} dy \]

(7)
If we assume $\Omega \equiv \mathbb{R}^m$ and also assume the probability density $p$ and its first derivatives $\partial p/\partial y_i$ vanish at a higher order of rate than $h$ and $\partial h/\partial y_i$ as $y_i \to \pm \infty \ \forall \ i = 1, \ldots, m$, then we can integrate by parts for the right hand side of (7), once for the first integral and twice for the second

$$
\int_{\Omega} \frac{\partial}{\partial y_i} h y A_i p dy = \int_{\Omega} \left. h y A_i p \right|_{y_i = -\infty}^{+\infty} dy_i - \int_{\Omega} h y \frac{\partial (A_i p)}{\partial y_i} dy \quad \text{and}
$$

$$
\int_{\Omega} \frac{\partial^2}{\partial y_i \partial y_j} \Sigma_{ij} p dy = \int_{\Omega} \left. \frac{\partial}{\partial y_j} \Sigma_{ij} p \right|_{y_i = -\infty}^{+\infty} dy_i - \int_{\Omega} \frac{\partial}{\partial y_j} \left( \frac{\partial (\Sigma_{ij} p)}{\partial y_i} \right) dy
$$

$$
= - \int_{\Omega} h y \left. \frac{\partial (\Sigma_{ij} p)}{\partial y_i} \right|_{y_i = -\infty}^{+\infty} dy_i + \int_{\Omega} h y \left. \frac{\partial^2 (\Sigma_{ij} p)}{\partial y_i \partial y_j} \right|_{y_i = -\infty}^{+\infty} dy
$$

where 

$$
\int_{\Omega} (\cdot) d\bar{y}_i = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} (\cdot) dy_1 \cdots dy_{i-1} dy_{i+1} \cdots dy_m
$$

Plugging the results of (8) into (7), we have

$$
\int_{\Omega} h y \frac{\partial p}{\partial T} dy = - \sum_{i=1}^{m} \int_{\Omega} h y \frac{\partial (A_i p)}{\partial y_i} dy + \frac{1}{2} \sum_{i,j=1}^{m} \int_{\Omega} h y \frac{\partial^2 (\Sigma_{ij} p)}{\partial y_i \partial y_j} dy
$$

$$
\Rightarrow \int_{\Omega} h y \left( \frac{\partial p}{\partial T} + \sum_{i=1}^{m} \frac{\partial (A_i p)}{\partial y_i} - \frac{1}{2} \sum_{i,j=1}^{m} \frac{\partial^2 (\Sigma_{ij} p)}{\partial y_i \partial y_j} \right) dy = 0
$$

By the arbitrariness of $h$, we conclude that for any $y \in \Omega$

$$
\frac{\partial p}{\partial T} + \sum_{i=1}^{m} \frac{\partial (A_i p)}{\partial y_i} - \frac{1}{2} \sum_{i,j=1}^{m} \frac{\partial^2 (\Sigma_{ij} p)}{\partial y_i \partial y_j} = 0, \quad \Sigma = B\rho B'
$$

This is the Multi-dimensional Fokker-Planck Equation (a.k.a. Kolmogorov Forward Equation) [1]. In this equation, the $t$ and $x$ are held constant, while the $T$ and $y$ are variables (called “forward variables”). In the one-dimensional case, it reduces to

$$
\frac{\partial p}{\partial T} + \frac{\partial (Ap)}{\partial y} - \frac{1}{2} \frac{\partial^2 (B^2 p)}{\partial y^2} = 0
$$

where $A = A(T, y)$ and $B = B(T, y)$ are then scalar functions.
1.2. Kolmogorov Backward Equation

Let’s express conditional expectation of \( h(S_t) \) by \( g(t, S_t) = \mathbb{E}_t[h(S_T)] \). Because for \( o \leq t \leq T \) we have

\[
g(o, S_o) = \mathbb{E}_o[h(S_T)] = \mathbb{E}_o[\mathbb{E}_t[h(S_T)]] = \mathbb{E}_o[g(t, S_t)]
\]  

(12)

the \( g(t, S_t) \) is a martingale by the tower rule (i.e. If \( \mathcal{H} \) holds less information than \( \mathcal{G} \), then \( \mathbb{E}[\mathbb{E}[X]|\mathcal{G}] = \mathbb{E}[X|\mathcal{H}] \)). The dynamics of the \( g(t, S_t) \) is given by

\[
dg = \frac{\partial g}{\partial t} dt + J_g dS_t + \frac{1}{2} dS_t^\prime H_g dS_t
\]  

\[
= \frac{\partial g}{\partial t} dt + J_g A dt + J_g B dW_t + \frac{1}{2} dW_t^\prime B^\prime H_g B dW_t
\]  

(13)

where \( J_g \) is the Jacobian (i.e. the same as gradient if \( g \) is a scalar-valued function) and \( H_g \) the Hessian of \( g \) with respect to \( S \) (with subscripts denoting the indices of vector components)

\[
J_g = \left( \frac{\partial g}{\partial S_1}, \ldots, \frac{\partial g}{\partial S_m} \right), \quad H_g = \begin{pmatrix}
\frac{\partial^2 g}{\partial S_1^2} & \cdots & \frac{\partial^2 g}{\partial S_1 \partial S_m} \\
\vdots & \ddots & \vdots \\
\frac{\partial^2 g}{\partial S_m \partial S_1} & \cdots & \frac{\partial^2 g}{\partial S_m^2}
\end{pmatrix}
\]  

(14)

Expanding (13), we have

\[
dg = \left( \frac{\partial g}{\partial t} + \sum_{i=1}^m \frac{\partial g}{\partial S_i} A_i + \frac{1}{2} \sum_{i,j=1}^m \frac{\partial^2 g}{\partial S_i \partial S_j} \Sigma_{ij} \right) dt + \sum_{i=1}^m \frac{\partial g}{\partial S_i} \sum_{k=1}^n B_{ik} dW_k
\]  

(15)

Since \( g(t, S_t) \) is a martingale, the \( dt \)-term must vanish, which gives

\[
\frac{\partial g}{\partial t} + \sum_{i=1}^m \frac{\partial g}{\partial S_i} A_i + \frac{1}{2} \sum_{i,j=1}^m \frac{\partial^2 g}{\partial S_i \partial S_j} \Sigma_{ij} = 0
\]  

(16)

This is the multi-dimensional Feynman-Kac formula\(^1\).

Using the transition probability density \( p_{T,y|t,x} \), we can write the expectation as

\[^{1} \text{https://en.wikipedia.org/wiki/Feynman-Kac_formula}\]
\[ g_{t,x} = \mathbb{E}_t[h(S_T)] = \int_{\Omega} h_y p_{T,y|t,x} dy \] (17)

The formula (16) defines that

\[
\frac{\partial}{\partial t} \int_{\Omega} h_y p dy + \sum_{l=1}^{m} A_l \frac{\partial}{\partial x_i} \int_{\Omega} h_y p dy + \frac{1}{2} \sum_{i,j=1}^{m} \Sigma_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \int_{\Omega} h_y p dy = 0
\]

\[ \Rightarrow \int_{\Omega} h_y \left( \frac{\partial p}{\partial t} + \sum_{l=1}^{m} A_l \frac{\partial p}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{m} \Sigma_{ij} \frac{\partial^2 p}{\partial x_i \partial x_j} \right) dy = 0 \] (18)

By the arbitrariness of \( h \), we have

\[ \frac{\partial p}{\partial t} + \sum_{l=1}^{m} A_l \frac{\partial p}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{m} \Sigma_{ij} \frac{\partial^2 p}{\partial x_i \partial x_j} = 0, \quad \Sigma = B\rho B' \] (19)

This is the multi-dimensional Kolmogorov Backward Equation. In this equation, the \( T \) and \( y \) are held constant, while the \( t \) and \( x \) are variables (called “backward variables”). In the 1-D case, it reduces to

\[ \frac{\partial p}{\partial t} + A \frac{\partial p}{\partial x} + \frac{1}{2} B^2 \frac{\partial^2 p}{\partial x^2} = 0 \] (20)

where \( A = A(t,x) \) and \( B = B(t,x) \) are then scalar functions.
2. **Heston Model**

In this chapter, we will briefly introduce the Heston stochastic volatility model, which has become quite popular in industry to model volatility smiles. One of the reasons for that is that vanilla options in Heston model have close form solutions. This makes the calibration of the model computationally much efficient and accurate. To understand these closed form formulas, we will introduce the Fourier transform method for option pricing and its application to the Heston model. Additionally, a characteristic function based method is also discussed, which extends the Heston model to have piecewise time-dependent parameters.

2.1. **Heston Stochastic Volatility Model**

The stochastic volatility in Heston’s model is a mean-reverting square-root process defined by the following stochastic differential equations (SDE)

\[
\frac{dS_t}{S_t} = (\mu - q)dt + \sqrt{v_t}dW_{1,t} \\
dv_t = \epsilon(\vartheta - v_t)dt + \xi\sqrt{v_t}dW_{2,t} \\
dW_{1,t}dW_{2,t} = \rho dt
\]

where \( t \) denotes the time, \( S_t \) the spot process, \( \mu \) and \( q \) the drift and dividend of the spot, \( v_t \) the variance process, \( \epsilon \) the mean reversion rate, \( \vartheta \) the mean variance, \( \xi \) the volatility of the variance and \( dW_{1,t}, dW_{2,t} \) the two Brownian motions correlated by \( \rho \) under physical measure \( \mathbb{P} \). All the parameters \( \mu, q, \epsilon, \vartheta \) and \( \rho \) are time and state homogenous (invariant).

2.1.1. **Market Price of Risk**

In the Black-Scholes model, a contingent claim is dependent on one or more tradable assets \( S_t \). The randomness in the option value is solely due to the randomness of these tradable assets. Therefore, the option can be hedged by continuously trading the underlying. This makes the market complete (i.e. every contingent claim can be replicated). In a stochastic volatility model, a contingent claim is dependent on both the randomness of the asset \( S_t \) and the randomness associated with the instantaneous volatility \( v_t \).
of the asset. Since the volatility is not a traded asset, this renders the market incomplete and has many implications to the pricing of options.

Firstly let’s consider two arbitrary derivative securities (presume they are available in the traded markets) whose prices can be written as functions $U(t, S_t, v_t)$ and $V(t, S_t, v_t)$ of variables $t$, $S_t$ and $v_t$ respectively. We construct a self-financing portfolio with a price process $X_t$ by having long one share of $U(t, S_t, v_t)$, short $\Gamma_t$ shares of $V(t, S_t, v_t)$ and short $\Delta_t$ shares of $S_t$, that is

$$X_t = U_t - \Gamma_t V_t - \Delta_t S_t$$

(22)

Following the Ito’s lemma, we can derive the price dynamics of the derivative as

$$dU = \frac{\partial U}{\partial t} dt + \frac{\partial U}{\partial S} dS + \frac{1}{2} \frac{\partial^2 U}{\partial S^2} dS^2 + \frac{\partial U}{\partial v} dv + \frac{1}{2} \frac{\partial^2 U}{\partial v^2} dv^2 + \frac{\partial U}{\partial S} dS \sqrt{v} dW_1 + \frac{\partial U}{\partial v} \epsilon(dv) dt + \frac{\partial U}{\partial v} \xi \sqrt{v} dW_2$$

(23)

and derive the price dynamics of the self-financing portfolio as

$$dX = dU - \Gamma dV - \Delta dS - \Delta q dt$$

(24)

In order to eliminate both spot and volatility risk, we must have

$$\frac{\partial U}{\partial S} - \Delta - \Gamma \frac{\partial V}{\partial S} = 0 \quad \text{and} \quad \frac{\partial U}{\partial v} - \Gamma \frac{\partial V}{\partial v} = 0$$

(25)

and therefore
In this case, the portfolio is riskless and must have a return at risk free rate in order to avoid arbitrage

\[
dX = rXdt = r(U - ΓV - ΔS)dt = rUdt - rΓVdt - r \frac{∂U}{∂S} Sdt + rΓ \frac{∂V}{∂S} Sdt
\]  

(27)

which in turn gives

\[
\frac{∂U}{∂t} + \frac{νS^2}{2} \frac{∂^2 U}{∂S^2} + \frac{νξ^2}{2} \frac{∂^2 U}{∂v^2} + Sνξρ \frac{∂^2 U}{∂v∂S} + \frac{∂U}{∂S}(r - q) - rU = \frac{∂V}{∂v} \]

(28)

\[
\frac{∂V}{∂t} + \frac{νS^2}{2} \frac{∂^2 V}{∂S^2} + \frac{νξ^2}{2} \frac{∂^2 V}{∂v^2} + Sνξρ \frac{∂^2 V}{∂v∂S} + \frac{∂V}{∂S}(r - q) - rV \equiv η
\]

In the above equation, the left-hand side is a function of \( U \) only and the right-hand side is a function of \( V \) only. The only way for the equality to hold is for both sides to equal a common function \( η \) of the independent variables \( t, S_t \) and \( v_t \).

Now let’s consider a delta-neutral portfolio \( Y_t \) by having long one share of \( U(t, S_t, v_t) \) and short \( Δ_t \) shares of \( S_t \)

\[
Y_t = U_t - Δ_t S_t
\]  

(29)

The price dynamics of the portfolio is

\[
dY = dU - ΔdS - ΔSqdt
\]

\[
= \left( \frac{∂U}{∂t} + \frac{∂U}{∂S}(μ - q)S + \frac{νS^2}{2} \frac{∂^2 U}{∂S^2} + \frac{∂U}{∂v} ε(θ - ν) + \frac{νξ^2}{2} \frac{∂^2 U}{∂v^2} + Sνξρ \frac{∂^2 U}{∂v∂S} \right) dt
\]

\[
+ \left( \frac{∂U}{∂S} - Δ \right) S\sqrt{ν}dW_1 + \frac{∂U}{∂v} ξ\sqrt{ν}dW_2 - Δ(μ - q)Sdt - ΔSqdt
\]

(30)

Since delta-neutral implies \( \frac{∂u}{∂s} - Δ = 0 \), we are able to derive the dynamics of the discounted portfolio, a martingale under risk neutral measure, as
\[ \frac{d(D_t Z_t)}{D_t} = \frac{d(D_t Y_t)}{D_t} = dY - rYdt = dU - \Delta S - \DeltaSq dt - rU dt + r\Delta S dt \]

\[ = \left( \frac{\partial U}{\partial v}(\theta - v) + \frac{\partial U}{\partial t} + \frac{vS^2}{2} \frac{\partial^2 U}{\partial S^2} + \frac{v\xi^2}{2} \frac{\partial^2 U}{\partial v^2} + Sv\xi \frac{\partial^2 U}{\partial v \partial S} + \frac{\partial U}{\partial S}(\mu - q) - rU \right) dt \]

\[ + \frac{\partial U}{\partial v} \xi \sqrt{v} dW_2 \]

\[ = \frac{\partial U}{\partial v}(\epsilon(\theta - v) + \eta) dt + \frac{\partial U}{\partial v} \xi \sqrt{v} dW_2 = \frac{\partial U}{\partial v} \xi \sqrt{v} d\tilde{W}_2 \]

by defining

\[ d\tilde{W}_2 = dW_2 + \phi_2 dt \quad \text{and} \quad \phi_2 = \frac{\epsilon(\theta - v) + \eta}{\xi \sqrt{v}} \] (32)

In the above, the \( \tilde{W}_2 \) is a Brownian motion under risk neutral measure \( \mathbb{Q} \) and \( \phi_2 \) is the market price of volatility risk.

We next consider a vega-neutral portfolio \( Z_t \) by having long a share of \( U(t, S_t, v_t) \) and short \( \Gamma_t \) shares of \( V(t, S_t, v_t) \)

\[ Z_t = U_t - \Gamma_t V_t \] (33)

The dynamics of the portfolio reads

\[ dZ = dU - \Gamma dV \]

\[ = \left( \frac{\partial U}{\partial t} + \frac{\partial U}{\partial S}(\mu - q)S + \frac{vS^2}{2} \frac{\partial^2 U}{\partial S^2} + \frac{v\xi^2}{2} \frac{\partial^2 U}{\partial v^2} + Sv\xi \frac{\partial^2 U}{\partial v \partial S} + \frac{\partial U}{\partial S}(\mu - q)S + \frac{v\xi^2}{2} \frac{\partial^2 V}{\partial v^2} + S\xi \frac{\partial^2 V}{\partial v \partial S} \right) dt \]

\[ -\Gamma \left( \frac{\partial V}{\partial t} + \frac{\partial V}{\partial S}(\mu - q)S + \frac{vS^2}{2} \frac{\partial^2 V}{\partial S^2} + \frac{v\xi^2}{2} \frac{\partial^2 V}{\partial v^2} + S\xi \frac{\partial^2 V}{\partial v \partial S} \right) dt \]

\[ + \left( \frac{\partial U}{\partial S} - \Gamma \frac{\partial V}{\partial S} \right) S\sqrt{v} dW_1 + \left( \frac{\partial U}{\partial v} - \Gamma \frac{\partial V}{\partial v} \right) \xi \sqrt{v} d\tilde{W}_2 \] (34)

Since vega-neutral implies \( \frac{\partial U}{\partial v} - \Gamma \frac{\partial V}{\partial v} = 0 \), we can derived the dynamics of the discounted portfolio as

\[ \frac{d(D_t Z_t)}{D_t} = dZ - rZdt = dU - \Gamma dV - rU dt + r\Gamma V dt \] (35)
\[
\begin{aligned}
&= \left( \frac{\partial U}{\partial t} + \frac{\partial U}{\partial S} (\mu - q) S + \frac{v S^2}{2} \frac{\partial^2 U}{\partial S^2} + \frac{v \xi^2}{2} \frac{\partial^2 U}{\partial v^2} + S v \xi \rho \frac{\partial^2 U}{\partial v \partial S} - r U \right) dt \\
&\quad - \Gamma \left( \frac{\partial V}{\partial t} + \frac{\partial V}{\partial S} (\mu - q) S + \frac{v S^2}{2} \frac{\partial^2 V}{\partial S^2} + \frac{v \xi^2}{2} \frac{\partial^2 V}{\partial v^2} + S v \xi \rho \frac{\partial^2 V}{\partial v \partial S} - r V \right) dt \\
&\quad + \left( \frac{\partial U}{\partial S} - \Gamma \frac{\partial V}{\partial S} \right) S \sqrt{v} dW_1 \\
&= \left( \frac{\partial U}{\partial t} + \frac{v S^2}{2} \frac{\partial^2 U}{\partial S^2} + \frac{v \xi^2}{2} \frac{\partial^2 U}{\partial v^2} + S v \xi \rho \frac{\partial^2 U}{\partial v \partial S} - r U \right) dt \\
&\quad - \Gamma \left( \frac{\partial V}{\partial t} + \frac{v S^2}{2} \frac{\partial^2 V}{\partial S^2} + \frac{v \xi^2}{2} \frac{\partial^2 V}{\partial v^2} + S v \xi \rho \frac{\partial^2 V}{\partial v \partial S} - r V \right) dt + \left( \frac{\partial U}{\partial S} - \Gamma \frac{\partial V}{\partial S} \right) dS \\
&= \left( \eta \frac{\partial U}{\partial S} - \frac{\partial U}{\partial S} S (r - q) - \Gamma \eta \frac{\partial V}{\partial v} + \Gamma \frac{\partial V}{\partial S} S (r - q) \right) dt + \left( \frac{\partial U}{\partial S} - \Gamma \frac{\partial V}{\partial S} \right) dS \\
&= \left( \frac{\partial U}{\partial S} - \Gamma \frac{\partial V}{\partial S} \right) (dS - S (r - q) dt) = \left( \frac{\partial U}{\partial S} - \Gamma \frac{\partial V}{\partial S} \right) S \left( (\mu - r) dt + \sqrt{v} dW_1 \right) \\
&= \left( \frac{\partial U}{\partial S} - \Gamma \frac{\partial V}{\partial S} \right) S \sqrt{v} d\tilde{W}_1
\end{aligned}
\]

by defining
\[
d\tilde{W}_1 = dW_1 + \phi_1 dt \quad \text{and} \quad \phi_1 = \frac{\mu - r}{\sqrt{v}} \quad (36)
\]

where \(d\tilde{W}_1\) is a Brownian motion under risk neutral measure \(\mathbb{Q}\) and \(\phi_1\) is the market price of spot risk.

According to (32) and (36), the Heston SDE (21) can be written under risk neutral measure as
\[
\begin{aligned}
\frac{dS_t}{S_t} &= (\mu - q) dt + \sqrt{v_t} (d\tilde{W}_{1,t} - \phi_1 dt) = (r - q) dt + \sqrt{v_t} d\tilde{W}_{1,t} \\
\frac{dv_t}{v_t} &= \epsilon (\theta - v_t) dt + \xi \sqrt{v_t} (d\tilde{W}_{2,t} - \phi_2 dt) = (\epsilon (\theta - v_t) - \phi_2 \xi \sqrt{v_t}) dt + \xi \sqrt{v_t} d\tilde{W}_{2,t}
\end{aligned}
\]

Based on the Consumption-based Capital Asset Pricing Model, Heston [22] assumes that the market price of volatility risk is proportional to volatility, that is
\[
\phi_2 = c \sqrt{v} \quad \text{for a constant} \, \, \, c \quad \Rightarrow \quad \phi_2 \xi \sqrt{v} = c \xi v = \lambda v \quad \text{where} \, \, \, \lambda = c \xi \quad (38)
\]

If we define
\[ \kappa = \epsilon + \lambda \text{ and } \theta = \frac{\epsilon \theta}{\kappa} \] (39)

then the Heston SDE under risk neutral measure \( \mathbb{Q} \) becomes

\[
\frac{dS_t}{S_t} = (r - q)dt + \sqrt{v_t}d\tilde{W}_{1,t}, \quad dv_t = \kappa(\theta - v_t)dt + \xi \sqrt{v_t}d\tilde{W}_{2,t}, \quad d\tilde{W}_{1,t}d\tilde{W}_{2,t} = \rho dt \] (40)

which retains the form of the equation under the transformation from the physical measure \( \mathbb{P} \) to the risk neutral measure \( \mathbb{Q} \).

Since the volatility is not a traded asset, the incompleteness of the market implies the risk neutral measure is not unique and depends on the value of the market price of volatility risk \( \phi_2 \). To estimate the model parameters, one may calibrate the Heston’s model using historical spot data, however the historical calibration does not allow for the estimation of \( \phi_2 \). Instead of using the spot data, one may also calibrate the model to the volatility smile (i.e. prices of vanilla options). In this case, the market price of volatility risk has already been implied in the market smile, and consequently embedded into the calibrated model parameters \( \kappa \) and \( \theta \) through (39).

2.1.2. Radon-Nikodym Derivative

The change of measure from \( \mathbb{P} \) to \( \mathbb{Q} \) is achieved through Radon-Nikodym derivative via multi-dimensional Girsanov’s theorem [2]. To derive this derivative, we may write correlated \( n \)-dimensional Brownian motions as \( dW_t \) and \( d\tilde{W}_t \) under physical measure \( \mathbb{P} \) and risk neutral measure \( \mathbb{Q} \) respectively. The matrix \( \rho \) denotes the instantaneous correlation, e.g. \( dW_t dW_t' = \rho dt \). It should be noted that \( dW_t \) and \( d\tilde{W}_t \) possess the same correlation structure only if each is under its own probability measure, \( \mathbb{P} \) or \( \mathbb{Q} \), otherwise this property does not hold. From (32) and (36), we represent the market price of risk by correlation matrix an \( n \)-dimensional vector \( \phi \) such that

\[
d\tilde{W}_t = dW_t + \rho \phi_t dt
\]

(41)

The Radon-Nikodym derivative is then given for \( t > 0 \) by
\[
\frac{dQ}{d\mathbb{P}} = \exp \left( -\frac{1}{2} \int_0^t \phi_u' \rho \phi_u du - \int_0^t \phi_u' dW_u \right) \quad (42)
\]

To check this, let’s assume under \( \mathbb{Q} \) we have a martingale process

\[
X_t = X_0 \exp \left( -\frac{1}{2} \int_0^t \sigma_u' \rho \sigma_u du + \int_0^t \sigma_u' d\tilde{W}_u \right) \quad (43)
\]

where \( \sigma_u \) is a vector representing an adapted volatility process. According to (41) we have

\[
X_t = X_0 \exp \left( -\frac{1}{2} \int_0^t \sigma_u' \rho \sigma_u du + \int_0^t \sigma_u' dW_u + \int_0^t \sigma_u' \phi_u du \right) \quad \text{and}
\]

\[
X_t \frac{dQ}{d\mathbb{P}} = X_0 \exp \left( -\frac{1}{2} \int_0^t \sigma_u' \rho \sigma_u du + \int_0^t \sigma_u' \phi_u du - \frac{1}{2} \int_0^t \phi_u' \rho \phi_u du + \int_0^t (\sigma_u - \phi_u)' dW_u \right) \quad (44)
\]

\[
= X_0 \exp \left( -\frac{1}{2} \int_0^t (\sigma_u - \phi_u)' \rho (\sigma_u - \phi_u) du + \int_0^t (\sigma_u - \phi_u)' dW_u \right)
\]

The \( X_t \frac{dQ}{d\mathbb{P}} \) is a martingale under \( \mathbb{P} \). Hence we have the following equation

\[
\mathbb{E}_0[X_t] = \mathbb{E}_o \left[ X_t \frac{dQ}{d\mathbb{P}} \right] = X_0 \quad (45)
\]
as expected.

2.1.3. Feller Condition

Feller observed that the variance process \( \nu_t \) in (40) remains strictly positive with probability one for all times \( t > o \), if \( \nu_o > 0 \) and the Feller condition [3] [4] is satisfied

\[
2\kappa \theta \geq \xi^2 \quad (46)
\]

If the condition is not satisfied, i.e. \( 0 < 2\kappa \theta < \xi^2 \), the variance will visit zero recurrently but will not stay at zero, i.e. the zero boundary is strongly reflecting. In typical applications, the Feller condition is often violated due to the convexities of volatility smiles typically encountered in practice. Indeed the process \( \nu_t \) often has a strong affinity for the area around the origin. However, this is not a complete disaster, as the process \( \nu_t \) can only hit zero for an infinitesimally small amount of time.

2.2. Probability Distribution of Spot Returns
In this section, we present a derivation of the distribution of the spot returns in Heston’s model \[5\].

Let’s firstly make a change of variable from \((21)\) to have a centered log-spot for \(t > o\)
\[
x_t = \ln \frac{S_t}{F_{o,t}} = \ln \left(\frac{S_t}{S_o} \exp((r - q)(t - o))\right) = \ln \frac{S_t}{S_o} - (r - q)(t - o)
\]
(47)

The Heston’s model \((40)\) under \(\mathbb{Q}\) can then be converted by Ito’s lemma to the following form
\[
dx_t = -\frac{v_t}{2} dt + \sqrt{v_t} d\tilde{W}_{1,t}, \quad dv_t = \kappa(\theta - v_t) dt + \xi \sqrt{v_t} d\tilde{W}_{2,t}, \quad d\tilde{W}_{1,t} d\tilde{W}_{2,t} = \rho dt
\]
(48)

This defines a 2-D stochastic process characterized by a joint transition probability density function \(p_{t,x,v|v_o}\), which is the probability having log-spot \(x_t\) and instantaneous variance \(v_t\) at time \(t\), conditional on \(v_o\) at time \(t = o\) (independent on \(x_o\) because \(x_o = 0\) almost surely).

2.2.1. Derivation of the Transition Probability

We may rewrite (48) in terms of a 2-D Brownian motion \(d\tilde{W}_t\)
\[
dZ_t = A_t dt + C_t d\tilde{W}_t \quad \text{with} \quad Z_t = \begin{pmatrix} x_t \\ v_t \end{pmatrix}, \quad A_t = \begin{pmatrix} -\frac{v_t}{2} \\ \kappa(\theta - v_t) \end{pmatrix}, \quad C_t = \begin{pmatrix} 0 \\ \xi \sqrt{v_t} \end{pmatrix}, \quad d\tilde{W}_t d\tilde{W}_t' = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} dt
\]
(49)

The instantaneous covariance matrix of \(dZ_t\) becomes
\[
\Sigma_t = C_t \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} C_t = \begin{pmatrix} v_t & \rho \xi v_t \\ \rho \xi v_t & \xi^2 v_t \end{pmatrix}
\]
(50)

The 2-D Markov process is characterized by the transition probability \(p_{t,x,v|x_o,v_o}\). The Fokker-Planck equation that governs the time evolution of the transition probability is given by (10)
\[
\frac{\partial p}{\partial t} - \frac{1}{2} \frac{\partial^2 (vp)}{\partial x^2} + \frac{\partial (\kappa(\theta - v)p)}{\partial v} - \frac{1}{2} \frac{\partial^2 (vp)}{\partial x^2} - \rho \xi \frac{\partial^2 (vp)}{\partial x \partial v} - \frac{\xi^2}{2} \frac{\partial^2 (vp)}{\partial v^2} = 0
\]
\[
\frac{\partial p}{\partial t} - \frac{v}{2} \frac{\partial p}{\partial x} + \kappa(\theta - v) \frac{\partial p}{\partial v} - \frac{v}{2} \frac{\partial^2 p}{\partial x^2} - \rho \xi \frac{\partial^2 (vp)}{\partial x \partial v} - \frac{\xi^2}{2} \frac{\partial^2 (vp)}{\partial v^2} = 0
\]
\[
\frac{\partial^2 (vp)}{\partial x \partial v} = \frac{\partial}{\partial x} \left( p + v \frac{\partial p}{\partial v} \right) = \frac{\partial p}{\partial x} + v \frac{\partial^2 p}{\partial x \partial v}
\]
(51)
with initial condition \( p_{t=0,x,v|v_o} = \delta_{x-x_o} \delta_{v-v_o} = \delta_x \delta_{v-v_o} \), where \( \delta \) is the Dirac delta function. The marginal probability density of the variance alone

\[
\zeta_{t,v|v_o} = \int_{\mathbb{R}} p_{t,x,v|v_o} dx
\]  

(52)

satisfies the following Fokker-Planck equation obtained from (51) by integration over \( x \)

\[
\frac{\partial \zeta}{\partial t} = \frac{\partial (\kappa (v-\theta) \zeta)}{\partial v} + \frac{\xi^2}{2} \frac{\partial^2 (\nu \zeta)}{\partial v^2}
\]  

(53)

Feller has shown that this equation is well defined on the interval \( \nu \in [0, +\infty) \) as long as \( \theta > 0 \). Equation (53) has a stationary solution, which is a Gamma distribution

\[
\zeta^*_{\nu} = \alpha^\alpha \nu^{\alpha-1} \frac{\Gamma(\alpha)}{\theta^\alpha} \exp\left(-\frac{\alpha \nu}{\theta}\right) \quad \text{and} \quad \alpha = \frac{2\kappa \theta}{\xi^2}
\]  

(54)

Since \( x \) appears in (51) only in the derivative operator, it is convenient to take the Fourier transform, such that

\[
\hat{p}_{t,\omega,v|v_o} = \int_{\mathbb{R}} e^{-i\omega x} p_{t,x,v|v_o} dx \quad \text{and} \quad p_{t,x,v|v_o} = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\omega x} \hat{p}_{t,\omega,v|v_o} d\omega
\]  

(55)

Inserting (55) into (51), we have

\[
\frac{\partial \hat{p}}{\partial t} = \frac{\partial (\kappa (v-\theta) \hat{p})}{\partial v} + \frac{i \omega - \omega^2}{2} v \hat{p} + i \rho \xi \omega \frac{\partial (v \hat{p})}{\partial v} + \frac{\xi^2}{2} \frac{\partial^2 (v \hat{p})}{\partial v^2}
\]  

(56)

Since (56) is linear in \( v \) and quadratic in \( \frac{\partial}{\partial v} \), it can be simplified by taking the Laplace transform over \( v \)

\[
\tilde{p}_{t,\omega,\lambda|v_o} = \int_{\mathbb{R}^+} e^{-\lambda \nu} \hat{p}_{t,\omega,v|v_o} d\nu
\]  

(57)

The PDE satisfied by \( \tilde{p}_{t,\omega,\lambda|v_o} \) is of the first order

\[
\frac{\partial \tilde{p}}{\partial t} = \left( \frac{\omega^2 - \xi^2 \lambda^2 - i \omega}{2} - \gamma \lambda \right) \frac{\partial \tilde{p}}{\partial \lambda} - \kappa \theta \lambda \tilde{p}
\]  

(58)

with initial condition \( \tilde{p}_{t=0,\omega,\lambda|v_o} = e^{-\lambda v_o} \), where \( \gamma = \kappa + i \rho \xi \omega \). The solution of this PDE is given by the method of characteristics.
where the function $\lambda_t$ is the solution of the characteristic (ordinary) differential equation
\[ \frac{d\lambda_u}{du} = \gamma \lambda_u + \frac{\xi^2}{2} \lambda_u^2 - \frac{\omega^2 - i\omega}{2} \]
(60)

With a boundary condition $\lambda_t = \lambda$ specified at time $t$, the (60) is a Riccati equation with constant coefficients and its solution is
\[ \lambda_u = \frac{2\Omega}{\xi^2} \Psi e^{\Omega(t-u)} - 1 - \frac{\gamma - \Omega}{\xi^2} \] with $\Omega = \sqrt{\gamma^2 + \xi^2(\omega^2 - i\omega)}$, $\Psi = 1 + \frac{2\Omega}{\xi^2 \lambda + \gamma - \Omega}$ (61)

Plugging (61) into (59), we have
\[ \tilde{p}_{t,\omega,\lambda|v_o} = \exp \left(-\tilde{\lambda}_o v_o - \kappa \theta \int_0^t \tilde{\lambda}_u du \right) \]
(59)

Normally we are interested only in distribution of log-spot $x_t$ and do not care about variance $v_t$.

Therefore we derive the marginal probability density for $x_t$ with $\lambda = 0$

\[ p_{t,x|v_o} = \int_{\mathbb{R}^+} p_{t,x,v|v_o} dv \]
\[ = \frac{1}{2\pi} \int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{i\omega x} \tilde{p}_{t,\omega,v|v_o} d\omega dv = \frac{1}{2\pi} \int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{i\omega x} \tilde{p}_{t,\omega,v|v_o} dv d\omega = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\omega x} \tilde{p}_{t,\omega,0|v_o} d\omega \]
(63)

\[ = \frac{1}{2\pi} \int_{\mathbb{R}} \exp \left( i\omega x - \frac{\omega^2 - i\omega}{\gamma + \Omega \coth \frac{\Omega t}{2}} v_o + \frac{\kappa \theta t}{\xi^2} - \frac{2\kappa \theta}{\xi^2} \ln \left( \cosh \frac{\Omega t}{2} + \frac{\gamma}{\Omega} \sinh \frac{\Omega t}{2} \right) \right) d\omega \]

where the last step comes from substitution of $\tilde{p}_{t,\omega,0|v_o}$ in (59) into (55). The derived density function $p_{t,x|v_o}$ in (63) is still dependent on the unknown initial variance $v_o$. To remove the dependence, the $v_o$ is assumed to have the stationary distribution density as in (54). Thus the unconditional transition density function $p_t(x)$, is derived by averaging (63) over $v_o$ with the weight $\zeta^*$

\[ p_{t,x} = \int_{\mathbb{R}^+} p_{t,x|z} \zeta^* dz \]
(64)
The integral over \( v_o \) is similar to the one of the Gamma function and can be taken explicitly. The final result is the Fourier integral

\[
p_{t,x} = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\omega x + F_{t,\omega}} d\omega, \quad F_{t,\omega} = \frac{\kappa \theta y t}{\xi^2} - 2\kappa \theta - \frac{2\kappa}{\xi^2} \ln \left( \cosh \frac{\Omega t}{2} + \frac{\Omega^2 - \gamma^2 + 2\kappa \gamma}{2\kappa \Omega} \sinh \frac{\Omega t}{2} \right)
\]  

(65)

It is easy to check that \( p_{t,x} \) is real, because \( \Re[F_{t,\omega}] \) is an even function of \( \omega \) and \( \Im[F_{t,\omega}] \) is an odd one.

One can also check that \( F_{t,\omega=0} = 0 \), which implies that \( p_{t,x} \) is correctly normalized at all times.

2.2.2. Moment Generating Function (in progress…)

We can integrate (48) over time from \( o \) to \( t \) to get

\[
x_t = x_o - \frac{1}{2} \int_o^t v_u d\nu_u + \int_o^t v_u d\tilde{W}_{1,u}, \quad v_t = v_o + \kappa \theta t - \kappa \int_o^t v_u d\nu_u + \xi \int_o^t \sqrt{v_u} d\tilde{W}_{2,u}
\]

(66)

where \( \tau = t - o \) and \( d\tilde{W}_{1,t} = \rho d\tilde{W}_{2,t} + \eta dB_t \) for \( \eta = \sqrt{1 - \rho^2} \) and the Brownian motion \( dB_t \) is independent of \( d\tilde{W}_{2,t} \). Defining a function \( h_{t,\alpha,\beta} \) such that

\[
h_{t,\alpha,\beta} = \mathbb{E}_o[\exp(ax_t + \beta v_t)] = \mathbb{E}_o\left[ \exp\left( ax_t + \left( \beta + \frac{\alpha \rho}{\xi} \right) v_t - \frac{\alpha \rho}{\xi} v_t \right) \right]
\]

\[
= \mathbb{E}_o\left[ \exp\left( ax_o - \alpha \frac{v_o + \kappa \theta t}{\xi} + \left( \beta + \frac{\alpha \rho}{\xi} \right) v_t + \left( \frac{\alpha \rho \kappa}{\xi} - \frac{\alpha}{2} \right) \int_o^t v_u d\nu_u + \alpha \eta \int_o^t \sqrt{v_u} dB_u \right) \right]
\]

\[
= \exp\left( ax_o - \alpha \frac{v_o + \kappa \theta t}{\xi} \right) \mathbb{E}_o\left[ \exp\left( \left( \beta + \frac{\alpha \rho}{\xi} \right) v_t + \left( \frac{\alpha \rho \kappa}{\xi} - \frac{\alpha}{2} \right) \int_o^t v_u d\nu_u + \alpha \eta \int_o^t \sqrt{v_u} dB_u \right) \right]
\]

(67)

\[
= e^{ax_o - \alpha \frac{v_o + \kappa \theta t}{\xi}} \mathbb{E}_o\left[ e^{\left( \beta + \frac{\alpha \rho}{\xi} \right) v_t + \left( \frac{\alpha \rho \kappa}{\xi} - \frac{\alpha}{2} \right) \int_o^t v_u d\nu_u + \alpha \eta \int_o^t \sqrt{v_u} dB_u \right] \quad (v_t \text{ is independent of } \tilde{B}_t)
\]

\[
= \exp\left( ax_o - \alpha \frac{v_o + \kappa \theta t}{\xi} \right) \mathbb{E}_o\left[ \exp\left( \left( \beta + \frac{\alpha \rho}{\xi} \right) v_t + \left( \frac{\alpha \rho \kappa}{\xi} - \frac{\alpha}{2} + \frac{\alpha^2 \eta^2}{2} \right) \int_o^t v_u d\nu_u \right) \right]
\]

we can derive the differential of \( \exp(ax_t + \beta v_t) \) as

\[
d\left( e^{ax_t + \beta v_t} \right) = e^{ax_t + \beta v_t} \left[ d(ax_t + \beta v_t) + \frac{1}{2} d(ax_t + \beta v_t) d(ax_t + \beta v_t) \right]
\]

(68)
\[ e^{\alpha x_t + \beta v_t} \left( \left(-\frac{\alpha}{2} + \frac{\alpha^2}{2} + \frac{\beta^2 \xi^2}{2} + \alpha \beta \xi \rho - \beta \kappa \right) v_t dt + \beta \kappa \theta dt + \alpha \sqrt{v_t} d\tilde{W}_{1,t} + \beta \xi \sqrt{v_t} d\tilde{W}_{2,t} \right) \]

Integrating both sides gives

\[
e^{\alpha x_t + \beta v_t} = e^{\alpha x_0 + \beta v_0} + \left(-\frac{\alpha}{2} + \frac{\alpha^2}{2} + \frac{\beta^2 \xi^2}{2} + \alpha \beta \xi \rho - \beta \kappa \right) \int_0^t e^{\alpha x_u + \beta v_u} v_u du + \beta \kappa \theta \int_0^t e^{\alpha x_u + \beta v_u} du + \alpha \int_0^t e^{\alpha x_u + \beta v_u} \sqrt{v_u} d\tilde{W}_{1,u} + \beta \xi \int_0^t e^{\alpha x_u + \beta v_u} \sqrt{v_u} d\tilde{W}_{2,u} \tag{69}\]

Taking expectation, we find

\[
h_{t;\alpha,\beta} = e^{\alpha x_0 + \beta v_0} + \left(-\frac{\alpha}{2} + \frac{\alpha^2}{2} + \frac{\beta^2 \xi^2}{2} + \alpha \beta \xi \rho - \beta \kappa \right) \int_0^t \mathbb{E}_o [v_u e^{\alpha x_u + \beta v_u}] du + \beta \kappa \theta \int_0^t h_{u;\alpha,\beta} du \tag{70}\]

\[\Rightarrow h_{t;\alpha,\beta} = h_{0;\alpha,\beta} + \left(-\frac{\alpha}{2} + \frac{\alpha^2}{2} + \frac{\beta^2 \xi^2}{2} + \alpha \beta \xi \rho - \beta \kappa \right) \int_0^t \frac{\partial h_{u;\alpha,\beta}}{\partial \beta} du + \beta \kappa \theta \int_0^t h_{u;\alpha,\beta} du \]

Differentiating with respect to \( t \), we have the PDE

\[
\frac{\partial h_{t;\alpha,\beta}}{\partial t} - p_{\alpha,\beta} \frac{\partial h_{t;\alpha,\beta}}{\partial \beta} = \beta \kappa \theta h_{t;\alpha,\beta} \tag{71}\]

(To be continued, reference [6])

2.3. Analytical Solution of Vanilla Options

Vanilla option price in Heston model can be computed semi-analytically. Shortly we will see that the spot process in Heston model admits a closed-form characteristic function, which allows us to express the option prices in terms of Fourier-inversion integrals that can be evaluated numerically (e.g. using Gaussian quadratures).

2.3.1. Fourier Transform

In textbooks, Fourier transform and its inverse (in \( n \)-dimensional case) are often depicted as
Forward Transform:  \( \hat{f}_\xi = \int_{\mathbb{R}^n} e^{-2\pi i \xi' x} f_x \, dx \quad \forall \ \omega \in \mathbb{R}^n \)  

(72)

Inverse Transform:  \( f_x = \int_{\mathbb{R}^n} e^{2\pi i \xi' x} \hat{f}_\xi \, d\xi \quad \forall \ x \in \mathbb{R}^n \)

where \( f: \mathbb{R}^n \rightarrow \mathbb{C} \) and \( \hat{f}: \mathbb{R}^n \rightarrow \mathbb{C} \) are integrable functions, and \( \xi \) is the ordinary frequency measured in hertz. This convention defines a unitary transformation on \( L^2(\mathbb{R}^n) \) (i.e. the inner product is preserved before and after the transformation). The transform is equivalent to writing a periodic function \( f_x \) in Fourier series expansion within an interval of \( T = 1/h \) for \( h > 0 \). In 1D, this reads

\[
\begin{align*}
    f_x &= \sum_{k \in \mathbb{Z}} \frac{c_k}{T} e^{2\pi i k \frac{x}{T}} = h \sum_{k \in \mathbb{Z}} c_k e^{2\pi i h k x} \quad \text{and} \quad c_k = \int_{-T/2}^{T/2} f_x e^{-2\pi i h k x} \, dx = \int_{-1/2h}^{1/2h} f_x e^{-2\pi i k x} \, dx \quad (73)
\end{align*}
\]

Letting \( \xi = h k \) and taking the limit \( h \to 0^+ \), we have

\[
\begin{align*}
    \lim_{h \to 0^+} c_k &= \lim_{h \to 0^+} \int_{-1/2h}^{1/2h} f_x e^{-2\pi i h k x} \, dx = \int_{\mathbb{R}} f_x e^{-2\pi i \xi x} \, dx = \hat{f}_\xi \quad \text{and} \\
    f_x &= \lim_{h \to 0^+} h \sum_{k \in \mathbb{Z}} c_k e^{2\pi i h k x} = \int_{\mathbb{R}} \hat{f}_\xi e^{2\pi i \xi x} \, d\xi 
\end{align*}
\]

(74)

The sum over all the integers in the above equation can be regarded as an approximating Riemann sum for the integral [7].

There is no standard definition of a Fourier transform and its inverse. All the definitions are mutually equivalent. The one that we use for characteristic functions follows an angular frequency convention and appears less symmetric, which is defined as follows

Forward:  \( \hat{f}_\omega = \int_{\mathbb{R}^n} e^{i \omega' x} f_x \, dx \quad \forall \ \omega \in \mathbb{R}^n \)

(75)

Inverse:  \( f_x = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i \omega' x} \hat{f}_\omega \, d\omega \quad \forall \ x \in \mathbb{R}^n \)

where the \( \omega \) is the angular frequency measured in radians per second (e.g. \( \omega = 2\pi \xi \)) and the (′) denotes matrix transpose (and hence the \( \omega' x \) denotes a dot product of two column vectors \( \omega \) and \( x \)). The extra term \( 1/(2\pi)^n \) is introduced due to the change of variable \( \omega = 2\pi \xi \) from ordinary frequency to angular
frequency. In fact the normalization factors multiplying the forward and inverse transforms (here 1 and \(1/(2\pi)^n\), respectively) and the signs of the exponents are merely conventions and differ in treatments. The only requirements for these conventions are: 1) the forward and inverse transforms have opposite-sign exponents, and 2) the product of their normalization factors is \(1/(2\pi)^n\). In 1D, the (75) reduces to

Forward:  \[
\hat{f}_\omega = \int_{\mathbb{R}} e^{i\omega x} f_x \, dx \quad \forall \omega \in \mathbb{R}
\]

Inverse:  \[
f_x = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\omega x} \hat{f}_\omega \, d\omega \quad \forall x \in \mathbb{R}
\]

2.3.2. Levy’s Inversion Formula

Characteristic function \(\phi_\omega\) of any random variable \(X\) completely defines its probability distribution. On the real line the characteristic function is given by the following formula

\[
\phi_\omega \equiv \mathbb{E}[e^{i\omega X}] = \int_{\mathbb{R}} e^{i\omega x} p_x \, dx = \int_{\Omega} e^{i\omega x} dP_x \quad \forall \omega \in \mathbb{R}
\]

where the \(p_x\) denotes the probability density function (PDF) and the \(P_x = \int_{-\infty}^{x} p_y \, dy\) is the cumulative density function (CDF). The characteristic function is merely a Fourier transform of the \(p_x\), likewise the \(p_x\) can be recovered from \(\phi_\omega\) through the inverse Fourier transform [8]

\[
p_x = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\omega x} \phi_\omega \, d\omega \quad \forall \omega \in \mathbb{R}
\]

The \(P_x\) can be computed from \(\phi_\omega\) through Levy’s Inversion Formula [9] [10] [11]

\[
P_x = \frac{\phi_0}{2} + \frac{1}{2\pi} \int_{\mathbb{R}^+} \frac{e^{i\omega x} \phi_{-\omega} - e^{-i\omega x} \phi_\omega}{i\omega} \, d\omega
\]

where \(\phi_0 = 1\) if \(\phi_\omega\) is a characteristic function of a random variable. Before proving the formula, we need to find Fourier transform of signum function \(S_x\), which is defined as

\[
S_x = \begin{cases} 
-1 & \text{if } x < 0 \\
1 & \text{if } x > 0 
\end{cases}
\]

Its transform cannot be obtained via direct integration. However we can consider an odd two-sided exponential function \(S_x^h\) with \(h > 0\)
\[ S^h_x = \begin{cases} -e^{hx} & \text{if } x < 0 \\ e^{-hx} & \text{if } x > 0 \end{cases} \] (81)

The \( \hat{S}^h_\omega \), i.e. the Fourier transform of \( S^h_x \), can then be derived as

\[
\hat{S}^h_\omega = \int_{\mathbb{R}} e^{i\omega x} S^h_x dx = -\int_{\mathbb{R}^-} e^{(i\omega + h)x} dx + \int_{\mathbb{R}^+} e^{(i\omega - h)x} dx
\]

\[
= \left. -\frac{e^{(i\omega + h)x}}{i\omega + h} \right|_{x=-\infty}^{0} + \left. \frac{e^{(i\omega - h)x}}{i\omega - h} \right|_{x=0}^{\infty} = -\frac{1}{i\omega + h} - \frac{1}{i\omega - h} = \frac{2i\omega}{\omega^2 + h^2}
\] (82)

The parameter \( h \) controls how rapidly the exponential function decays. As we let \( h \to 0 \), the exponential function resembles more and more closely the signum function. This suggests that

\[
\hat{S}_\omega = \lim_{h \to 0} \hat{S}^h_\omega = \lim_{h \to 0} \frac{2i\omega}{\omega^2 + h^2} = -\frac{2}{i\omega}
\] (83)

Hence the inverse transform of the \( \hat{S}_\omega \) gives

\[
S_x = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\omega x} \hat{S}_\omega d\omega = -\frac{1}{\pi} \int_{\mathbb{R}} \frac{e^{-i\omega x}}{i\omega} d\omega = -\frac{1}{\pi} \int_{\mathbb{R}} \frac{\cos \omega x - i \sin \omega x}{i\omega} d\omega
\]

\[
= -\frac{1}{i\pi} \int_{\mathbb{R}} \frac{\cos \omega x}{\omega} d\omega + \frac{1}{\pi} \int_{\mathbb{R}^+} \frac{\sin \omega x}{\omega} d\omega = \frac{2}{\pi} \int_{\mathbb{R}^+} \frac{\sin \omega x}{\omega} d\omega
\] (84)

With the help of the signum function \( S_x \) in (84), the proof of (79) is given as follows

\[
\int_{\mathbb{R}^+} \frac{e^{i\omega x} \phi_\omega - e^{-i\omega x} \phi_\omega}{i\omega} d\omega = \int_{\mathbb{R}^+} \frac{e^{i\omega x} \int_{\mathbb{R}} e^{-i\omega y} p_y dy - e^{-i\omega x} \int_{\mathbb{R}} e^{i\omega y} p_y dy}{i\omega} d\omega
\]

\[
= \int_{\mathbb{R}^+} \int_{\mathbb{R}} \frac{e^{i\omega (x-y)} - e^{-i\omega (x-y)}}{i\omega} p_y dy d\omega
\]

\[
= \int_{\mathbb{R}} \int_{\mathbb{R}^+} \frac{e^{i\omega (x-y)} - e^{-i\omega (x-y)}}{i\omega} d\omega p_y dy \quad \text{(by Fubini’s theorem)}
\]

\[
= 2 \int_{\mathbb{R}} \int_{\mathbb{R}^+} \frac{\sin \omega (x-y)}{\omega} d\omega p_y dy \quad \text{(by } e^{ix} = \cos x + i \sin x \text{)}
\]

\[
= \pi \int_{\mathbb{R}} S_{x-y} p_y dy = \pi \left( -\int_{-\infty}^{\infty} p_y dy + \int_{-\infty}^{x} p_y dy \right) = \pi \left( -\int p_y dy + 2\int_{-\infty}^{x} p_y dy \right)
\] (85)
\[ = \pi (-\phi_0 + 2P_x) \]

The Levy Inversion Formula can also be written in the following form

\[ P_x = \frac{\phi_0}{2} - \frac{1}{2\pi} \int_{\mathbb{R}} \frac{e^{-i\omega x} \phi_{\omega}}{i\omega} d\omega \]  

which can be proved as

\[ \int_{\mathbb{R}} \frac{e^{-i\omega x}}{i\omega} d\omega = \int_{\mathbb{R}^{-}} \frac{e^{-i\omega x}}{i\omega} d\omega + \int_{\mathbb{R}^{+}} \frac{e^{-i\omega x}}{i\omega} d\omega = \int_{\mathbb{R}^{+}} \frac{e^{-i\omega x}}{-i\omega} d\omega + \int_{\mathbb{R}^{+}} \frac{e^{-i\omega x}}{i\omega} d\omega \]

\[ \int_{\mathbb{R}^{+}} \frac{e^{i\omega x} - e^{-i\omega x}}{i\omega} d\omega = 2\int_{\mathbb{R}^{+}} \frac{\sin \omega x}{\omega} d\omega = \pi S_x \]

Note that if \( p_x \) is real-valued function (i.e. probability density function), its Fourier transform \( \phi_\omega \) is then even in its real part and odd in its imaginary part [12], we therefore have \( \phi_\omega = \overline{\phi_{-\omega}} \), the inversion formula (86) becomes identical to (79), which can be further reduced to

\[ P_x = \frac{\phi_0}{2} - \frac{1}{2\pi} \int_{\mathbb{R}^{+}} \frac{e^{i\omega x} \phi_{-\omega} - e^{-i\omega x} \phi_{\omega}}{i\omega} d\omega = \frac{\phi_0}{2} + \frac{1}{2\pi} \int_{\mathbb{R}^{+}} \frac{e^{-i\omega x} \phi_{\omega} - e^{-i\omega x} \phi_{\omega}}{i\omega} d\omega \]

\[ = \frac{\phi_0}{2} - \frac{1}{\pi} \int_{\mathbb{R}^{+}} \mathcal{R} \left[ \frac{e^{-i\omega x} \phi_{\omega}}{i\omega} \right] d\omega = \frac{\phi_0}{2} + \frac{1}{\pi} \int_{\mathbb{R}^{+}} \mathcal{I} \left[ \frac{e^{-i\omega x} \phi_{\omega}}{i\omega} \right] d\omega \]  

(88)

For \( P_x^C = \int_{x}^{\infty} p_y dy \), the complementary of \( P_x \), we can use the simple relation \( P_x^C = \phi_0 - P_x \)

\[ P_x^C = \frac{\phi_0}{2} - \frac{1}{2\pi} \int_{\mathbb{R}^{+}} \frac{e^{i\omega x} \phi_{-\omega} - e^{-i\omega x} \phi_{\omega}}{i\omega} d\omega = \frac{\phi_0}{2} + \frac{1}{\pi} \int_{\mathbb{R}^{+}} \mathcal{R} \left[ \frac{e^{-i\omega x} \phi_{\omega}}{i\omega} \right] d\omega \]

\[ = \frac{\phi_0}{2} + \frac{1}{\pi} \int_{\mathbb{R}^{+}} \mathcal{I} \left[ \frac{e^{-i\omega x} \phi_{\omega}}{i\omega} \right] d\omega \]

(89)

2.3.3. Characteristic Function
In this section, we present a derivation of the closed form characteristic function of the spot in Heston model [13]. The definition of the Heston joint process in (48) will be used. Suppose there exists a payoff function \( g_{x,T,v_T} \) on \( x_T \) and \( v_T \), we may calculate risk neutral expectation of the payoff function as

\[
h_t = \mathbb{E}_t[g_{x_T,v_T}]
\]  

(90)

For example, the characteristic function of the joint distribution of \( x_T \) and \( v_T \) would be given by

\[
\phi_{t,\alpha,\beta} = \mathbb{E}_t[g_{x_T,v_T;\alpha,\beta}] \quad \text{where} \quad g_{x_T,v_T;\alpha,\beta} = \exp(i\alpha x_T + i\beta v_T)
\]

(91)

The risk neutral expectation is actually a martingale because for \( o < t < T \)

\[
h_o = \mathbb{E}_o[g_{x_T,v_T}] = \mathbb{E}_o[\mathbb{E}_t[g_{x_T,v_T}]] = \mathbb{E}_o[h_t]
\]

(92)

Applying Ito’s lemma to \( h \) and forcing the drift to be zero (martingale property), we end up with a PDE

\[
\frac{\partial h}{\partial t} - \frac{v}{2} \frac{\partial h}{\partial x} + \kappa(\theta - v) \frac{\partial h}{\partial v} + \frac{\xi^2 v}{2} \frac{\partial^2 h}{\partial x^2} + \frac{\theta}{2} \frac{\partial^2 h}{\partial v^2} + \xi \rho v \frac{\partial^2 h}{\partial x \partial v} = 0
\]

(93)

To determine the solution of (93), the terminal condition \( h_T = \mathbb{E}_T[g_{x_T,v_T}] = g_{x_T,v_T} \) at time \( T \) must be specified. The terminal payoff function that we will consider has the form \( g_{x_T,v_T} = e^{\gamma + \delta v_T + i\alpha x_T} \). If \( \gamma = 0 \) and \( \delta = i\beta \), the resulting payoff function becomes \( g_{x_T,v_T;\alpha,\beta} \) in (91), corresponding to the characteristic function of the joint distribution.

Heston [22] guessed a solution that has the form

\[
h_t = \mathbb{E}_t[g_{x_T,v_T}] = e^{C + Dv_T + i\alpha x_T} \quad \text{where} \quad C = C_{t,\alpha,\gamma,\delta}, \quad D = D_{t,\alpha,\gamma,\delta}, \quad \tau = T - t
\]

(94)

Substituting the tentative solution (94) in (93) yields

\[
\frac{\partial C}{\partial t} + v \frac{\partial D}{\partial t} - \frac{i\alpha v}{2} + \kappa(\theta - v)D - \frac{\alpha^2 v}{2} + \frac{\xi^2 v}{2} D^2 + i\alpha D \xi \rho v = 0
\]

(95)

\[
\Rightarrow \frac{\partial C}{\partial t} + \kappa \theta D + \left( \frac{\partial D}{\partial t} + \frac{\xi^2 v}{2} D^2 - m D - \frac{\alpha(i + \alpha)}{2} \right) v = 0 \quad \text{where} \quad m = \kappa - \alpha \xi \rho i
\]

As the \( v \) is an independent variable, (95) is zero only if

\[
\frac{\partial C}{\partial t} + \kappa \theta D = 0 \quad \text{and} \quad \frac{\partial D}{\partial t} + \frac{\xi^2 v}{2} D^2 - m D - \frac{\alpha(i + \alpha)}{2} = 0
\]

(96)
Changing the variable \( t \) to \( \tau = T - t \), we have

\[
\frac{\partial C}{\partial \tau} - \kappa \theta D = 0 \quad \text{and} \quad \frac{\partial D}{\partial \tau} - \frac{\xi^2}{2} D^2 + mD + \frac{\alpha(i + \alpha)}{2} = 0
\]  

(97)

The terminal condition for this system of equations is given by \( C_{\tau = 0} = \gamma \) and \( D_{\tau = 0} = \delta \). The ODE for \( D \) is a Riccati equation that only depends on \( D \). This Riccati equation can be turned into an ODE through the change of variable, \( Z = (D - \bar{D})^{-1} \), where \( \bar{D} \) is a particular solution to the second equation in (97)

\[
\frac{\partial Z}{\partial \tau} = -\frac{1}{(D - \bar{D})^2} \frac{\partial (D - \bar{D})}{\partial \tau} = -Z^2 \left( \frac{\xi^2}{2} D^2 - mD - \frac{\xi^2}{2} \bar{D}^2 + m\bar{D} \right)
\]

\[= -Z \left( \frac{\xi^2}{2} (D + \bar{D}) - m \right) = -\frac{\xi^2}{2} \left( (D - \bar{D} + 2\bar{D}) \right) + mZ = -\left( \frac{\xi^2}{2} \bar{D} - m \right) - \frac{\xi^2}{2} \]

(98)

\[\Rightarrow \frac{\partial Z}{\partial \tau} + (\frac{\xi^2}{2} \bar{D} - m)Z + \frac{\xi^2}{2} = 0 \quad \text{or} \]

\[\frac{\partial Z}{\partial \tau} + BZ + A = 0 \quad \text{where} \quad A = \frac{\xi^2}{2}, \quad B = \frac{\xi^2}{2} \bar{D} - m \]

The solution to (98) is given by

\[Z = -\frac{A}{B} + \left( Z_0 + \frac{A}{B} \right) e^{-B \tau} \quad \text{where} \quad Z_0 = \frac{1}{D_0 - \bar{D}} = \frac{1}{\delta - \bar{D}} \]

(99)

\[\Rightarrow D = \frac{1}{-\frac{A}{B} + \left( Z_0 + \frac{A}{B} \right) e^{-B \tau}} + \bar{D} \]

The particular solution \( \bar{D} \) can be as simple as a constant, which implies it would be the solution of the quadratic equation

\[-\frac{\xi^2}{2} \bar{D}^2 + m\bar{D} + \frac{\alpha(i + \alpha)}{2} = 0 \Rightarrow \bar{D} = \frac{m + d}{\xi^2} \quad \text{where} \quad d = \pm \sqrt{m^2 + \xi^2 \alpha(i + \alpha)} \]

(100)

This particular solution for \( \bar{D} \) makes \( B = d \). The solution for \( D \) in (99) can then be derived

\[D = \frac{1}{-\frac{\xi^2}{2d} + \left( \frac{-\xi^2}{m + d - \delta \xi^2 + \frac{\xi^2}{2d}} \right) e^{-d \tau}} + \frac{m + d}{\xi^2} \]

(101)
$$\Rightarrow \xi^2 D = -\frac{2d(m + d - \delta \xi^2)}{2e^{-d\tau}d + (m + d - \delta \xi^2)(1 - e^{-d\tau})} + m + d$$

$$= -\frac{2d(m + d - \delta \xi^2)}{(m + d - \delta \xi^2) - (m - d - \delta \xi^2)e^{-d\tau}} + m + d$$

$$= -\frac{2d}{1 - \hat{g}e^{-d\tau}} + m + d \quad \text{(by defining } \hat{g} = \frac{m - d - \delta \xi^2}{m + d - \delta \xi^2})$$

$$= \frac{(m + d)(1 - \hat{g}e^{-d\tau}) - 2d}{1 - \hat{g}e^{-d\tau}} = \frac{m - d - (m + d)\hat{g}e^{-d\tau}}{1 - \hat{g}e^{-d\tau}}$$

$$= \frac{(m + d)g - \hat{g}e^{-d\tau}}{1 - \hat{g}e^{-d\tau}} \quad \text{(by defining } g = \frac{m - d}{m + d})$$

$$\Rightarrow D = \frac{m + d}{\xi^2} \frac{g - \hat{g}e^{-d\tau}}{1 - \hat{g}e^{-d\tau}}$$

The $D$ can then be plugged into the ODE for $C$ in (97)

$$\frac{\partial C}{\partial \tau} - \kappa \theta \frac{m + d}{\xi^2} \frac{g - \hat{g}e^{-d\tau}}{1 - \hat{g}e^{-d\tau}} = 0 \quad \Rightarrow \quad C = \kappa \theta \frac{m + d}{\xi^2} \int \frac{g - \hat{g}e^{-d\tau}}{1 - \hat{g}e^{-d\tau}} d\tau + K_c$$

(102)

where $K_c$ is a constant to be fixed by terminal condition. The indefinite integral in (102) can be solved through change of variable $u = e^{-d\tau}$ where $\frac{\partial u}{\partial \tau} = -ud$

$$\int \frac{g - \hat{g}e^{-d\tau}}{1 - \hat{g}e^{-d\tau}} d\tau = -\int \frac{(g - \hat{g}u)}{1 - \hat{g}u} \frac{1}{ud\tau} = -\frac{1}{d} \int \left( \frac{g - \hat{g}}{1 - \hat{g}u} \right) \partial u$$

$$= -\frac{1}{d} \int \left( \frac{g - \hat{g} - \hat{g}u(1 - \hat{g}u)}{1 - \hat{g}u} + \frac{g}{u} \right) \partial u = -\frac{1}{d} \int \left( \frac{(g - 1)\hat{g}}{1 - \hat{g}u} + \frac{g}{u} \right) \partial u$$

(103)

$$= \frac{g - 1}{d} \ln(1 - \hat{g}u) - \frac{g}{d} \ln u = \frac{g - 1}{d} \ln(1 - \hat{g}e^{-d\tau}) + g\tau$$

which gives
\[ C = \kappa \theta \frac{m + d}{\xi^2} \left( \frac{g - 1}{d} \ln(1 - \hat{g} e^{-d\tau}) + g\tau \right) + K_c \]

\[
= \frac{\kappa \theta}{\xi^2} \left( -2 \ln(1 - \hat{g} e^{-d\tau}) + (m - d)\tau \right) + K_c
\]

We then fix \( K_C \) by \( C_0 = \gamma \)

\[
C_0 = -2 \frac{\kappa \theta}{\xi^2} \ln(1 - \hat{g}) + K_C = \gamma \Rightarrow K_C = 2 \frac{\kappa \theta}{\xi^2} \ln(1 - \hat{g}) + \gamma
\]

Therefore we have the solution for \( C \)

\[
C = \frac{\kappa \theta}{\xi^2} \left( 2 \ln \frac{1 - \hat{g}}{1 - \hat{g} e^{-d\tau}} + (m - d)\tau \right) + \gamma
\]

Combining the solutions in (101) and (106), the \( h_t \) is of the following form

\[ h_t = \mathbb{E}_t [g_{x_T, v_T}] = \exp(C + D v_t + i\alpha x_t) \quad \text{where} \]

\[
m = \kappa - \alpha \xi \rho i, \quad d = \pm \sqrt{m^2 + \xi^2 \alpha (i + \alpha)}, \quad g = \frac{m - d}{m + d}, \quad \hat{g} = \frac{m - d - \delta \xi^2}{m + d - \delta \xi^2}
\]

\[
C = \frac{\kappa \theta}{\xi^2} \left( 2 \ln \frac{1 - \hat{g}}{1 - \hat{g} e^{-d\tau}} + (m - d)\tau \right) + \gamma, \quad D = \frac{m + d \ g - \hat{g} e^{-d\tau}}{\xi^2 \ 1 - \hat{g} e^{-d\tau}}
\]

The marginal characteristic function of \( x_T \) is given by \( \gamma = 0 \) and \( \delta = 0 \) where \( \hat{g} = g \)

\[
\phi^{x_T}_\alpha = \exp \left( \frac{\kappa \theta}{\xi^2} \left( 2 \ln \frac{1 - g}{1 - g e^{-d\tau}} + (m - d)\tau \right) + \frac{m - d}{\xi^2} \frac{1 - e^{-d\tau}}{1 - g e^{-d\tau}} v_t + i\alpha x_t \right)
\]

\[ \text{where} \quad m = \kappa - \alpha \xi \rho i, \quad d = \pm \sqrt{m^2 + \xi^2 \alpha (i + \alpha)}, \quad g = \frac{m - d}{m + d} \]

2.3.4. Vanilla Option Prices

Once we know the analytical form of the characteristic function \( \phi^{x_T}_\alpha \) of the centered log-spot \( x_T \), we are able to compute the vanilla option prices using inversion methods. Here, we are going to discuss two methods, which treat the option price function analogous to the cumulative density function or the probability density function, respectively. In addition, we also summarize the original Heston’s method [22].
2.3.4.1. Analogy to Cumulative Density Function

For $t \leq T$, using a European call option as an example we change the variable to $x_t$ as in (47)

$$C_{T,K} = \mathbb{E}_t \left[ \frac{M_t}{M_T} (S_T - K)^+ \right] = e^{-rt} F_{t,T} \mathbb{E}_t [(e^{x_T} - e^K)^+]$$

with $\tau = T - t$, $\mu = r - q$, $F_{t,T} = S_t e^{\mu_T}$, $x_T = \ln \frac{S_T}{F_{t,T}}$, $K = \ln \frac{K}{F_{t,T}}$

where the cash account $M_T = M_t \exp \left( \int_t^T r_u du \right)$ and the interest rate $r$ and dividend yield $q$ are assumed constant. We then define an option forward price in percentage of the underlying forward as a function of moneyness $K$ defined as the log strike over forward

$$C_K = \frac{C_{T,K}}{e^{-rt} F_{t,T}} = \mathbb{E}_t [(e^{x_T} - e^K)^+] = \mathbb{E}_t [(e^{x_T} - e^K) 1\{x_T > K\}]$$

with $1\{x > X\} = \begin{cases} 1 & \text{if } x > X \\ 0 & \text{otherwise} \end{cases}$

Given we know the characteristic function of the log-spot process $x_T$, we can derive the Fourier transform of the call option, then use numerical inversion to obtain option prices directly [14]. Since $C_{T,K} \in [0, e^{-rt} F_{0,T}]$, we have $C_K \in [0,1]$, which can treated as a cumulative density function on $K$, the Fourier transform of the option price is then given by

$$\chi^C = \int_{K \in \mathbb{R}} e^{i\omega K} dC_K = e^{i\omega K} C_K \bigg|_{K=\infty} - e^{i\omega K} C_K dK$$

$$= -e^{-i\omega \infty} - \int_{\mathbb{R}} i\omega e^{i\omega K} C_K dK$$

(by $C_\infty = 0$ and $C_{-\infty} = 1$)

$$= -e^{-i\omega \infty} - \int_{\mathbb{R}} i\omega e^{i\omega K} \int_{\Omega} (e^{x} - e^{K}) 1\{x > K\} dP_x^{x_T} dK$$

(by Fubini’s theorem)

$$= -e^{-i\omega \infty} - \int_{\Omega} \int_{\mathbb{R}} i\omega (e^{i\omega K + x} - e^{(i\omega + 1)K}) 1\{x > K\} dK dP_x^{x_T}$$
The \( \mathcal{X}_\omega \) is the characteristic function of the option forward price \( \mathcal{C}_\mathcal{K} \), which is treated just as a cumulative density function. The price is then given by the Levy’s Inversion Formula (79) \[15\]

\[
\mathcal{C}_\mathcal{K} = \frac{X_0}{2} + \frac{1}{2\pi} \int_{\mathbb{R}^+} \frac{e^{i\omega \mathcal{C}} - e^{-i\omega \mathcal{C}}}{i\omega} d\omega
\]

\[
= \frac{1}{2} + \frac{1}{2\pi} \int_{\mathbb{R}^+} \frac{e^{i\omega \mathcal{C}} \phi_{\omega-i}^x + e^{-i\omega \mathcal{C}} \phi_{\omega-i}^x}{i\omega} d\omega
\]

(112)

In the integrand of (112), the \( \phi_{\omega-i}^x \) for \( u = \pm 1 \) can be derived from the characteristic function \( \phi_{\omega}^x \) in (108) and given as follows (note that \( \mathcal{X}_t = 0 \))

\[
\phi_{\omega-i}^x = \exp \left( \frac{\kappa \theta}{\xi^2} \left( 2 \ln \frac{1 - g}{1 - ge^{-d\tau}} + (m - d)\tau \right) + \frac{m - d}{\xi^2} \frac{1 - e^{-d\tau}}{1 - ge^{-d\tau}} v_t \right)
\]

(113)

with \( \tau = T - t, \quad m = \kappa - u\omega \xi p - \xi \rho, \quad d = \pm \sqrt{m^2 + \xi^2 \omega^2 (\omega - ui)}, \quad g = \frac{m - d}{m + d} \)
The inversion formula in (112) involves evaluation of the \( \phi_{xT}^{\omega-i} \) function twice, which is less efficient.

Since \( \mathcal{C}_\mathcal{K} \) is real-valued, we may use (88) to perform the inversion

\[
\mathcal{C}_\mathcal{K} = \frac{X_0}{2} - \frac{1}{\pi} \int_{\mathbb{R}^+} \mathfrak{R} \left[ \frac{e^{-i\omega \mathcal{C}_\mathcal{K} \omega}}{\omega} \right] d\omega = \frac{1}{2} + \frac{1}{\pi} \int_{\mathbb{R}^+} \mathfrak{R} \left[ \frac{e^{-i\omega \mathcal{C}_\mathcal{K} \omega}}{i\omega^2 + \omega} \right] d\omega
\]

(114)

where \( \phi_{xT}^{\omega-i} \) is given in (113) with \( u = 1 \) and the integral can be estimated numerically by Gauss-Laguerre quadrature. Once we have \( \mathcal{C}_\mathcal{K} \) for log-moneyness, the call option price can be computed by (110)

\[
\mathcal{C}_{T,K} = \mathcal{C}_\mathcal{K} e^{-rT}S_t e^{\mu T} = \mathcal{C}_\mathcal{K} S_t e^{-qT}
\]

(115)

and the put option price by call-put parity

\[
\mathcal{C}_{T,K} - P_{T,K} = e^{-rT}(F_{t,T} - K) \Rightarrow P_{T,K} = \mathcal{C}_{T,K} - S_t e^{-qT} + e^{-rT}K
\]

(116)

2.3.4.2. Analogy to Probability Density Function

In this section, we treat \( \mathcal{C}_\mathcal{K} \) analogous to a probability density [16] [17] and define in terms of generalized Fourier transform with \( z = z_r - iz_l, \ z_r \in \mathbb{R}, \ z_l \in \mathcal{D} \subseteq \mathbb{R}^+ \)

\[
\chi^p_z = \int_{\mathbb{R}} e^{iz \mathcal{K}} \mathcal{C}_\mathcal{K} d\mathcal{K} = \int_{\mathbb{R}} e^{iz \mathcal{K}} \mathfrak{E}_t[(e^{x_T} - e^{\mathcal{K}})\mathbb{1}[x_T > \mathcal{K}]] d\mathcal{K}
\]

\[
= \mathfrak{E}_t \left[ \int_{\mathbb{R}} e^{iz \mathcal{K}} (e^{x_T} - e^{\mathcal{K}})\mathbb{1}[x_T > \mathcal{K}] d\mathcal{K} \right] = \mathfrak{E}_t \left[ \int_{-\infty}^{x_T} e^{iz \mathcal{K}} (e^{x_T} - e^{\mathcal{K}}) d\mathcal{K} \right]
\]

\[
= \mathfrak{E}_t \left[ \frac{e^{iz \mathcal{K} + x_T} - e^{(iz + 1) \mathcal{K}}}{iz} \bigg|_{\mathcal{K} = -\infty}^{x_T} \right] = \mathfrak{E}_t \left[ \frac{e^{(iz + 1) x_T}}{iz - z^2} \right] \text{ (by } \lim_{\mathcal{K} \rightarrow -\infty} e^{iz \mathcal{K}} = 0) \]

\[
= \mathfrak{E}_t \left[ \frac{e^{i(z - i) x_T}}{iz - z^2} \right] = \phi_{z-i}^{x_T} \frac{\phi_{xT}^{z-i}}{iz - z^2}
\]

(117)

For some return distributions, the return transform \( \phi_{z-i}^{x_T} \) is well-defined only when \( z_l \) is in a subset of the real line. We use \( \mathcal{D} \subseteq \mathbb{R}^+ \) to denote the subset that both guarantees the convergence of \( e^{iz \mathcal{K}} \) and \( e^{i(z - i) \mathcal{K}} \) as \( \mathcal{K} \rightarrow -\infty \), and assures the finiteness of the transform \( \phi_{z-i}^{x_T} \).

The option forward price \( \mathcal{C}_\mathcal{K} \) is then given by the inverse Fourier transform
The last equality holds because $C_{\mathcal{K}}$ is real, which implies that the function $\chi_{x}^{p}$ is odd in its imaginary part and even in its real part. The functional form of $\phi_{z-i}^{x_t}$ can be derived again from (108) in a similar way

$$\phi_{z-i}^{x_t} = \exp \left( \frac{k\theta}{\xi^2} \left( 2 \ln \frac{1 - g}{1 - ge^{-d\tau}} + (m - d)\tau \right) + \frac{m - d}{\xi^2} \frac{1 - e^{-d\tau}}{1 - ge^{-d\tau}} \psi_t \right) \quad \text{with}$$

$$m = \kappa - (z_i + 1)\xi p - z_i\xi p_i, \quad d = \pm \sqrt{m^2 + \xi^2(z_r - i(z_i + 1))(z_r - iz_i)}, \quad g = \frac{m - d}{m + d}$$

**(2.3.4.3. Heston’s Original Solution)**

At time $t$, the price of a European call on a stock (assuming non-dividend-bearing to qualify as a numeraire) with spot $S_t$ and strike $K$ is given by the no-arbitrage formula

$$C_{t,T,S_t} = \mathbb{E}_t \left[ \frac{M_t}{M_T} (S_T - K)^+ \right] = \mathbb{E}_t \left[ \frac{M_t}{M_T} S_T \mathbb{1}\{S_T > K\} \right] - K \mathbb{E}_t \left[ \frac{M_t}{M_T} \mathbb{1}\{S_T > K\} \right]$$

$$= \mathbb{E}_t \left[ \frac{S_t}{S_T} S_T \mathbb{1}\{S_T > K\} \right] - K \mathbb{E}_t \left[ \frac{B_{t,T}}{B_{T,T}} \mathbb{1}\{S_T > K\} \right] \quad \text{change numeraire } M_t \to S_t, \quad M_t \to B_{t,T}$$

$$= S_t \mathbb{P}_t^S [S_T > K] - K B_{t,T} \mathbb{P}_t^T [S_T > K]$$

where $\mathbb{P}_t^S [S_T > K]$ and $\mathbb{P}_t^T [S_T > K]$ are both conditional probabilities of spot finishing in-the-money at maturity. The $\mathbb{P}_t^S [S_T > K]$ is computed under the measure associated with the stock as the numeraire, whereas the $\mathbb{P}_t^T [S_T > K]$ is computed under $T$-forward measure associated with zero coupon bond $B_{t,T}$ as the numeraire [18]. In Black-Scholes model

$$C_{t,T,S_t}^{BS} = S_t \Phi(d_+) - K B_{t,T} \Phi(d_-)$$

the $\mathbb{P}_t^S [S_T > K]$ and $\mathbb{P}_t^T [S_T > K]$ are computed as $\Phi(d_+)$ and $\Phi(d_-)$ respectively. Since the drift adjustment due to change of numeraire is
\[
\frac{dW_t^N}{\text{Under } N} = \frac{d\tilde{W}_t}{\text{Under } Q} - \sigma_N dt
\] (122)

where \( N \) denotes the measure associated with numeraire \( N \) and \( Q \) the risk neutral measure. The stock process under the measure with itself as the numeraire is given by

\[
\frac{dS_t}{S_t} = r dt + \sigma d\tilde{W}_t = (r + \sigma^2) dt + \sigma dW_t^S
\] (123)

The total drift adjustment \( \sigma^2 \tau \) for period \( \tau = T - t \) is then normalized by the total volatility \( \sigma \sqrt{\tau} \) of the stock to give a shift term \( \sigma \sqrt{\tau} \) as the difference between \( d_+ \) and \( d_- \) in the classic Black-Scholes formula.

Suppose we use the definition in (109) for a call option, the two conditional probabilities can be expressed as

\[
C_{T,K} = e^{-rT} F_{t,T} \hat{E}_t \left[ (e^{x_T} - e^K)^+ \right] = e^{-rT} F_{t,T} \left( \hat{E}_t e^{x_T 1\{x_T > K\}} - e^K \hat{E}_t [1\{x_T > K\}] \right)
\]

\[
= e^{-rT} F_{t,T} P^+_K - e^{-rT} K P^-_K
\]

where \( P^+_K = \hat{E}_t e^{x_T 1\{x_T > K\}}, \quad P^-_K = \hat{E}_t [1\{x_T > K\}] \) (124)

Because in Heston model the \( P^+_K \) and \( P^-_K \) are not available in closed form, Heston [22] firstly derived the characteristic functions (i.e. the Fourier transforms) of \( P^+_K \) and \( P^-_K \) by solving the PDE (93) for each of them, and then found inverse of the two characteristic functions to obtain the option price. Since we have already derived the characteristic function of \( x_T \) as in (108), we can easily derive those for \( P^+_K \) and \( P^-_K \).

For example, the \( P^+_K \) can be treated as a CDF and then the characteristic function of \( P^+_K \) becomes

\[
\chi_\omega^+ = \int_{\mathbb{R}} e^{i\omega \mathcal{K}} dP^+_\mathcal{K} = e^{i\omega \mathcal{K}} P^+_\mathcal{K}\bigg|_{\mathcal{K} = -\infty}^{\mathcal{K} = \infty} - \int_{\mathbb{R}} i\omega e^{i\omega \mathcal{K}} P^+_\mathcal{K} d\mathcal{K}
\]

\[
= -e^{-i\omega \infty} - \int_{\mathbb{R}} i\omega e^{i\omega \mathcal{K}} P^+_\mathcal{K} d\mathcal{K} \quad \text{(by } P^-_\infty = 0 \text{ and } P^+_\infty = \hat{E}_t [e^{x_T}] = 1\text{)}
\]

\[
= -e^{-i\omega \infty} - \int_{\mathbb{R}} i\omega e^{i\omega \mathcal{K}} \int_{\Omega} e^\mathcal{K} 1\{\mathcal{K} = \mathcal{K} \} dP^x_{\mathcal{K}} d\mathcal{K}
\]

\[
= -e^{-i\omega \infty} - \int_{\Omega} \int_{\mathbb{R}} i\omega e^{i\omega \mathcal{K} + x} 1\{x > \mathcal{K} \} d\mathcal{K} dP^x_{\mathcal{K}} d\mathcal{K} \quad \text{(by Fubini’s theorem)}
\]

\[
= -e^{-i\omega \infty} - \int_{\Omega} \int_{-\infty}^{\mathcal{X}} i\omega e^{i\omega \mathcal{K} + x} d\mathcal{K} dP^x_{\mathcal{K}} = -e^{-i\omega \infty} - \int_{\Omega} e^{i\omega \mathcal{K} + x} dP^x_{\mathcal{K}}
\]

\[
\text{ where } \mathcal{X} = \mathcal{K} - x
\]
Similarly, the characteristic function of \( P_\mathcal{K}^- \) is given by

\[
\chi^- = \int_{\mathcal{K} \in \mathbb{R}} e^{i \omega \mathcal{K}} dP_\mathcal{K}^- = e^{i \omega \mathcal{K}} P_\mathcal{K}^- \bigg|_{\mathcal{K} = -\infty} - \int_{\mathbb{R}} i \omega e^{i \omega \mathcal{K}} P_\mathcal{K}^- d\mathcal{K}
\]

\[
= -e^{-i \omega \infty} - \int_{\mathbb{R}} i \omega e^{i \omega \mathcal{K}} P_\mathcal{K}^- d\mathcal{K} \quad \text{(by } P_\infty = 0 \text{ and } P_-^\infty = 1) \]

\[
= -e^{-i \omega \infty} - \int_{\mathbb{R}} i \omega e^{i \omega \mathcal{K}} \int_{\mathcal{K}} 1 \{x > \mathcal{K}\} dP_{xT}^\mathcal{K} d\mathcal{K} = -e^{-i \omega \infty} - \int_{\mathbb{R}} i \omega e^{i \omega \mathcal{K}} \int_{\mathcal{K}} 1 \{x > \mathcal{K}\} d\mathcal{K} dP_{xT}^\mathcal{K}
\]

\[
= -e^{-i \omega \infty} - \int_{\mathbb{R}} \int_{-\infty}^x i \omega e^{i \omega \mathcal{K}} d\mathcal{K} dP_{xT}^\mathcal{K} = -e^{-i \omega \infty} - \int_{\mathbb{R}} i \omega e^{i \omega \mathcal{K}} \int_{-\infty}^x dP_{xT}^\mathcal{K}
\]

\[
= -e^{-i \omega \infty} - \int_{\Omega} \left(e^{i \omega x} - e^{-i \omega \infty}\right) dP_{xT}^\mathcal{K} = -e^{-i \omega \infty} + e^{-i \omega \infty} \int_{\Omega} dP_{xT}^\mathcal{K} - \int_{\Omega} e^{i \omega x} dP_{x}
\]

\[
= -\phi_{xT}^\omega
\]

The \( P_+^\mathcal{K} \) and \( P_-^\mathcal{K} \) are then obtained through the inverse of \( \chi^+ \) and \( \chi^- \) respectively. Given that \( \chi^+_0 = \chi^-_0 = 1 \), the Heston’s vanilla option price formula can be derived from \( \phi_{xT}^\omega \) in (108) and the inverse formula in (88) directly. Here we summarize the solution as follows

\[
\text{HestonVanilla}(\kappa, \theta, \xi, \rho, r, q, v_t, S_t, K, \tau, \eta) = e^{-\eta \tau} S_t(P_+ - \eta) - e^{-\eta \tau} K(P_- - \eta)
\]

with \( \eta = \begin{cases} 0 & \text{for call} \\ 1 & \text{for put} \end{cases} \)

\( m = \begin{cases} \kappa - \rho \xi \omega_i - \xi \rho & \text{for } P_+ \\ \kappa - \rho \xi \omega & \text{for } P_- \end{cases} \)

\( u = \begin{cases} 1 & \text{for } P_+ \\ -1 & \text{for } P_- \end{cases} \)

\( g = \frac{m - d}{m + d}, \quad d = \pm \sqrt{m^2 + \xi^2 \omega (\omega - ui)}, \quad C = \frac{\kappa \theta}{\xi^2} \left(2 \ln \frac{1 - g}{1 - ge^{-d \tau}} + (m - d) \tau\right)\)
\[
D = \frac{m - d}{\xi^2} \frac{1 - e^{-\Delta t}}{1 - e^{-\Delta t}}, \quad 
\chi_\omega = \exp(C + D \nu_t + i \omega \ln S_t e^{(r-q)\tau})
\]

\[
P = \frac{1}{2} + \frac{1}{\pi} \int_{\mathbb{R}^+} \Im \left[ \frac{e^{-i \omega \ln K} \chi_\omega}{\omega} \right] d\omega
\]

where the \( \eta \) denotes the option type with \( \eta = 1 \) for calls and \( \eta = -1 \) for puts. In Heston’s original formula [22], the negative solution of \( d \) is used, which makes the calculation of the complex logarithm prone to numerical instabilities. This is because taking the principal value of the logarithm causes \( C \) to jump discontinuously each time the imaginary part of the argument of the logarithm crosses the negative real axis (i.e. discontinuity due to branch cut of complex numbers), especially for long maturities. Albrecher et al. [19] presents an extensive study proving that both solutions are completely equivalent from a theoretical point of view, it is also mentioned that rather than using the negative solution of \( d \), the positive \( d \) guarantees numerical stability of the resulting formula under a full dimensional and unrestricted parameter space.

2.4. Piecewise Time Dependent Heston Model

The time-dependent Heston model we present here was proposed by Elices [20] in 2008. The model relies on the characteristic function of the two dimensional Markov process, which we have derived in (107). It bootstraps a series of piecewise constant Heston parameters starting from the earliest maturity, each parameter set for a period of time, which allows the model to fit to a term structure of the implied volatility surfaces.

Consider a Markov \( n \)-dimensional stochastic process \( Y \) (i.e. \( Y = \left( x_t \right) \) in Heston model), we define \( p_{y_t}^{s,t} \) the transition probability density function for having \( Y = y_t \) at time \( t \) conditional on \( Y = y_s \) at time \( s \) where \( o \leq s \leq t \leq T \). Its characteristic function \( \phi_{\omega}^{s,t} \) is the multi-dimensional Fourier transform of the transition density \( p_{y_t}^{s,t} \) such that

\[
\phi_{\omega}^{o,t} = \int_{\mathbb{R}^n} e^{i \omega^T y_t} p_{y_t}^{o,t} dy_t = \int_{\mathbb{R}^n} e^{i \omega^T y_t} \int_{\mathbb{R}^n} p_{y_t}^{s,t} p_{y_s}^{o,s} dy_s dy_t = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i \omega^T y_t} p_{y_t}^{s,t} p_{y_s}^{o,s} dy_s dy_t \quad (128)
\]
Consider a family of exponential characteristic functions with exponent linear in the state $y_s$ at time $s$

$$
\phi^s_t = \exp(C^{s,t}(\omega) + D^{s,t}(\omega)')y_s)
$$

we have

$$
\phi^{s,t}_\omega = \int_{\mathbb{R}^n} \exp(C^{s,t}(\omega) + D^{s,t}(\omega)'y_s) p_{y_s}^{s,t} dy_s = \int_{\mathbb{R}^n} \exp(C^{s,t}(\omega)) \int_{\mathbb{R}^n} \exp(D^{s,t}(\omega)'y_s) p_{y_s}^{s,t} dy_s
$$

Identifying terms between the (129) for $s = 0$ and the (130), we find that

$$
C^{0,t}(\omega) = C^{s,t}(\omega) + C^{o,s}(-iD^{s,t}(\omega)) \quad \text{and} \quad D^{0,t}(\omega) = D^{o,s}(-iD^{s,t}(\omega))
$$

Formula (131) gives a recursive definition of $C^{o,t}(\omega)$ and $D^{o,t}(\omega)$, which can be used to bootstrap the Heston model parameters for each time period starting from the earliest maturity. For illustrative purpose, the first 3 periods are presented below

<table>
<thead>
<tr>
<th>Period</th>
<th>$C$</th>
<th>$D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_0 \to t_1$</td>
<td>$C^{0,1}(\omega)$</td>
<td>$D^{0,1}(\omega)$</td>
</tr>
<tr>
<td>$t_1 \to t_2$</td>
<td>$C^{1,2}(\omega) + C^{0,1}(-iD^{1,2}(\omega))$</td>
<td>$D^{0,1}(-iD^{1,2}(\omega))$</td>
</tr>
<tr>
<td>$t_2 \to t_3$</td>
<td>$C^{2,3}(\omega) + C^{1,2}(-iD^{2,3}(\omega)) + C^{0,1}(-iD^{1,2}(-iD^{2,3}(\omega)))$</td>
<td>$D^{0,1}(-iD^{1,2}(-iD^{2,3}(\omega)))$</td>
</tr>
</tbody>
</table>
3. **Heston Model: PDE by Finite Element Method**

Valuation of financial derivatives has become a very important subject in modern financial theory and practice. In 1973, Black and Scholes introduced a simple formula [21] to price European-style options under a few strong assumptions. It is the first successful attempt to provide an arbitrage free valuation of financial derivatives. However due to limitations, it fails to capture some critical features observed in financial markets, such as heavy tails of return, skewness and smile in implied volatility, clustering and autocorrelation in volatility, etc. Many approaches have been proposed to address these issues. One of them is to employ stochastic volatility models to describe the price dynamics in financial markets, in which the volatility of the underlying is assumed to be stochastic. Heston [22] proposed a stochastic volatility model in 1993. It extends the Black-Scholes model and includes it as a special case. One major advantage of the Heston model is that it can be solved in closed-form for vanilla options. For exotic options, closed-form solutions are generally not available, numerical methods must be considered for option pricing. Monte Carlo simulations, though easy to implement, are not very suitable for computing risk sensitivities. It turns out that solving partial differential equation (PDE) is not only more numerically efficient but also more stable in risk sensitivity estimation. Traditionally, the PDE’s that arise from option pricing are solved dominantly by finite difference method (FDM). FDM is straightforward to implement, but meanwhile it also imposes many constraints, such as, sufficiently smooth terminal and boundary conditions, rectangular domains, etc. To relax these constraints, it becomes the primary motivation of this essay to develop an efficient finite element methods (FEM) for solving the PDE’s. Many studies have been focused on this topic. Topper [23] provides an excellent introduction to the FEM in the context of financial engineering applications. Previous work by Winkler et al. [24] illustrates an application of FEM to valuation of vanilla option in Heston stochastic volatility model. Achdou and Tchou [25] Propose a finite element analysis for Black-Scholes equation with stochastic volatility process. A recent publication by Miglio and Sgarra [26] discusses an application of finite element method for option pricing in a stochastic volatility model with jumps, known as the Bates model.
In this chapter, we will present an application of finite element method in the Heston model to price exotic options (specifically, the double barrier knock-out options). It presents a detailed implementation of the finite element method and its application in option pricing. The numerical results and conclusions are outlined in the end.

3.1. The Partial Differential Equation

Let \( U(t, S_t, \nu_t) = \frac{1}{\partial_t} \mathbb{E}_t[D_tU(T, S_T, \nu_T)] \) be the price of a contingent payoff \( U(T, S_T, \nu_T) \) that occur at maturity \( T \). Assuming deterministic interest rate, we have the dynamics of \( D_tU(t, S_t, \nu_t) \) in Heston model (21) under risk neutral measure \( \mathbb{Q} \), that is

\[
\frac{1}{D_t} d(D_tU(t, S_t, \nu_t)) = dU - rU dt
\]

\[
\frac{\partial U}{\partial t} dt + \frac{\partial U}{\partial S} dS + \frac{1}{2} \frac{\partial^2 U}{\partial S^2} dS dS + \frac{\partial U}{\partial \nu} d\nu + \frac{1}{2} \frac{\partial^2 U}{\partial \nu^2} d\nu d\nu + \frac{\partial^2 U}{\partial \nu \partial S} d\nu dS - rU dt
\]

\[
= \frac{\partial U}{\partial t} dt + \frac{\partial U}{\partial S} (r - q) S dt + \frac{\partial U}{\partial S} S \sqrt{\nu} d\tilde{W}_1 + \frac{\nu S^2}{2} \frac{\partial^2 U}{\partial S^2} dt + \frac{\partial U}{\partial \nu} \kappa (\theta - \nu) dt + \frac{\partial U}{\partial \nu} \xi \sqrt{\nu} d\tilde{W}_2
\]

\[
+ \frac{\nu \xi^2}{2} \frac{\partial^2 U}{\partial \nu^2} dt + S \nu \xi \rho \frac{\partial^2 U}{\partial \nu \partial S} dt - rU dt
\]

Since the \( D_tU(t, S_t, \nu_t) \) is a martingale under \( \mathbb{Q} \), the \( dt \)-term in (132) must vanish, which defines a PDE that the derivative price \( U \) must follow

\[
\frac{\partial U}{\partial t} + \frac{\nu \xi^2 \partial^2 U}{2 \partial \nu^2} + S \nu \xi \rho \frac{\partial^2 U}{\partial \nu \partial S} + \frac{\nu S^2}{2} \frac{\partial^2 U}{\partial S^2} + \kappa (\theta - \nu) \frac{\partial U}{\partial \nu} + (r - q) S \frac{\partial U}{\partial S} - rU = 0
\]

(133)

Prices of different types of derivative securities can be obtained by solving the PDE (133) under different terminal and boundary conditions. We firstly define a rectangle domain

\[
\Omega := \{(S, \nu) : S \in (S_{\min}, S_{\max}), \ \nu \in (\nu_{\min}, \nu_{\max})\}
\]

(134)

The domain boundaries are then taken as

\[
\Gamma_1 := \{(S, \nu) : S \in (S_{\min}, S_{\max}), \ \nu = \nu_{\min}\}
\]

(135)

\[
\Gamma_2 := \{(S, \nu) : S \in (S_{\min}, S_{\max}), \ \nu = \nu_{\max}\}
\]


\[ \Gamma_3 := \{ (S, v) : S = S_{\text{min}}, \ v \in [v_{\text{min}}, v_{\text{max}}] \} \]

\[ \Gamma_4 := \{ (S, v) : S = S_{\text{max}}, \ v \in [v_{\text{min}}, v_{\text{max}}] \} \]

The terminal condition is defined at option maturity. A vanilla call ($\eta = 1$) or put ($\eta = -1$) that matures in $\tau = T - t$ would have a terminal condition defined by the maturity payoff function

\[ U(T, S_T, v_T) = (\eta(S_T - K))^+ \]  \hfill (136)

Given a strike $K$ and a properly defined min/max range for $S$ and $v$, the Dirichlet boundary conditions for a vanilla option can be defined as follows (here we choose Dirichlet boundary condition because it is easier to handle. Ideally a zero gamma boundary condition should be used, however the choice of boundary condition becomes insignificant when the domain for $S$ and $v$ gets large enough)

\[ U(t, S_t, v_{\text{min}}) = (\eta(S_t e^{-q \tau} - K e^{-r \tau}))^+ \quad \forall \ (S, v) \in \Gamma_1 \]

\[ U(t, S_t, v_{\text{max}}) = \frac{1 + \eta}{2} S_t e^{-q \tau} + \frac{1 - \eta}{2} K e^{-r \tau} \quad \forall \ (S, v) \in \Gamma_2 \]

\[ U(t, S_{\text{min}}, v_t) = \frac{1 - \eta}{2} (\eta(S_{\text{min}} e^{-q \tau} - K e^{-r \tau}))^+ \quad \forall \ (S, v) \in \Gamma_3 \]

\[ U(t, S_{\text{max}}, v_t) = \frac{1 + \eta}{2} (\eta(S_{\text{max}} e^{-q \tau} - K e^{-r \tau}))^+ \quad \forall \ (S, v) \in \Gamma_4 \]  \hfill (137)

For a European up-and-out-down-and-out double barrier call/put option, the terminal condition is of the same form. The Dirichlet boundary conditions change a bit as the min/max value for $S$ is limited to the up and down barriers

\[ U(t, S_t, v_{\text{min}}) = (\eta(S_t e^{-q \tau} - K e^{-r \tau}))^+ \quad \forall \ (S, v) \in \Gamma_1 \]

\[ U(t, S_t, v_{\text{max}}) = 0 \quad \forall \ (S, v) \in \Gamma_2 \]

\[ U(t, S_{\text{min}}, v_t) = 0 \quad \forall \ (S, v) \in \Gamma_3 \]  \hfill (138)

\[ U(t, S_{\text{max}}, v_t) = 0 \quad \forall \ (S, v) \in \Gamma_4 \]

Equation (133) can be further simplified by change of variable, which yields
\[
\frac{\partial w}{\partial t} + \xi^2 v \frac{\partial^2 w}{\partial v^2} + \rho \xi \nu \frac{\partial^2 w}{\partial v \partial y} + \frac{v \partial^2 w}{2 \partial y^2} + \kappa (\theta - v) \frac{\partial w}{\partial v} + \left( r - q - \frac{v}{2} \right) \frac{\partial w}{\partial y} - rw = 0
\]  

(139)

where we define \( y_t = \ln \frac{S_t}{K} \) and \( w(t, y_t, v_t) = U(t, S_t, v_t) \). The PDE (139) is actually identical to our previously derived (93) with variables

\[
w(t, y_t, v_t) = e^{rt} h(t, x_t, v_t), \quad x_t = \ln \frac{S_t}{S_0} - (r - q) t
\]  

(140)

This can be shown by using the chain rule to derive partial derivatives after change of variables

\[
\frac{\partial w}{\partial t} = \left( \frac{\partial}{\partial t} + \frac{\partial x}{\partial \tau} \frac{\partial}{\partial x} \right) (e^{rt} h) = e^{rt} \left( rh + \frac{\partial h}{\partial \tau} + \frac{\partial x \partial h}{\partial x} \right) = e^{rt} \left( rh + \frac{\partial h}{\partial \tau} - (r - q) \frac{\partial h}{\partial x} \right)
\]

\[
\kappa (\theta - v) \frac{\partial w}{\partial v} = e^{rt} \kappa (\theta - v) \frac{\partial h}{\partial v}, \quad \frac{\xi^2 v \partial^2 w}{2 \partial v^2} = e^{rt} \frac{\xi^2 v \partial^2 h}{2 \partial v^2}
\]

\[
\left( r - q - \frac{v}{2} \right) \frac{\partial w}{\partial y} = \left( r - q - \frac{v}{2} \right) \frac{\partial (e^{rt} h)}{\partial x} \frac{\partial x}{\partial y} = e^{rt} \left( r - q - \frac{v}{2} \right) \frac{\partial h}{\partial x}
\]  

(141)

\[
\frac{v \partial^2 w}{2 \partial y^2} = e^{rt} \frac{v \partial}{\partial y} \left( \frac{\partial h}{\partial \tau} + \frac{\partial h}{\partial \tau} + \frac{\partial x \partial h}{\partial x} \right) = e^{rt} \frac{v \partial^2 h}{2 \partial x^2}
\]

\[
\rho \xi \nu \frac{\partial^2 w}{\partial v \partial y} = \rho \xi \nu \frac{\partial}{\partial v} \left( \frac{\partial (e^{rt} h)}{\partial x} \frac{\partial x}{\partial y} \right) = e^{rt} \rho \xi \nu \frac{\partial^2 h}{\partial v \partial x}
\]

For simplicity, the PDE (139) can be written in terms of the gradient and divergence operator

\[
(\partial_t + \nabla \cdot \mathbf{A} \nabla - \mathbf{b} \cdot \nabla - r)w = 0, \quad \mathbf{A} = \frac{v}{2} \begin{pmatrix} \xi^2 & \rho \xi \\ \rho \xi & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \frac{\xi^2}{2} - \kappa (\theta - v) \\ \frac{v + \rho \xi}{2} - (r - q) \end{pmatrix}
\]  

(142)

where \( \partial_t = \partial / \partial t \), the gradient operator \( \nabla = \begin{pmatrix} \partial / \partial x \\ \partial / \partial y \end{pmatrix} \), and the divergence operator \( \nabla \cdot = \begin{pmatrix} \partial / \partial x \\ \partial / \partial y \end{pmatrix} \). Note that we have the following identities
\[ \nabla \cdot \mathbf{A} \nabla = \left( \frac{\partial v}{\partial y} \right)^T \frac{\xi}{2} \left( \begin{array}{cc} \frac{\xi^2}{\rho \xi} & \rho \xi \\ \rho \xi & 1 \end{array} \right) \left( \frac{\partial v}{\partial y} \right) = \left( \frac{\partial v}{\partial y} \right)^T \frac{\xi}{2} \left( \frac{\xi^2 \partial_v + \rho \xi \partial_y}{\rho \xi \partial_v + \partial_y} \right) \]

\[ = \frac{\xi^2 \partial_v + \rho \xi \partial_y}{2} + \frac{\nu}{2} \left( \frac{\xi^2 \partial_v \partial_v + \rho \xi \partial_v \partial_y}{\rho \xi \partial_v \partial_v + \partial_y} \right) + \frac{\nu}{2} \left( \rho \xi \partial_v \partial_y + \partial_y \partial_y \right) \]

\[ = \frac{\xi^2}{2} \partial_v + \rho \xi \frac{\partial_y}{2} + \rho \xi \nu \partial_v \partial_y + \frac{\nu \xi^2}{2} \partial_v \partial_y + \frac{\nu}{2} \partial_y \partial_y \]

(143)

\[ \mathbf{b} \cdot \nabla = \left( \frac{\xi^2}{\nu + \rho \xi} \right)^T \left( \begin{array}{c} \partial_v \\ \partial_y \end{array} \right) = \frac{\xi^2}{2} \partial_v - \kappa (\theta - v) \partial_v + \frac{\nu + \rho \xi}{2} \partial_y - (r - q) \partial_y \]

Henceforth, we will use (\cdot) to denote dot product. If we think of column vectors as the standard vectors, the divergence operator can be regarded as transpose of the gradient operator (a bit abuse of notation).

### 3.2. Numerical Solution of the PDE

The PDE (142) shows a 2-D dynamic problem. To solve this problem, a semi-discretization in time is applied, which yields a series of 2-D static boundary value problems. These 2-D static boundary value problems are then solved numerically by a 2-D finite element method as time advances.

#### 3.2.1. Temporal Discretization

It should be noted that the time advances backwards, such that the initial value is given at the terminal \( t^N = T \) and the solution is sought at \( t^0 = 0 \). Note that the superscript is used to denote the index of time-step. The partial derivative with respect to time is approximated by finite difference method.

According to (142), the \( \sigma \)-weighted finite difference scheme for time is given as

\[ \frac{w^{k+1} - w^k}{t^{k+1} - t^k} + (1 - \sigma)(\nabla \cdot \mathbf{A} \nabla - \mathbf{b} \cdot \nabla - r)w^{k+1} + \sigma(\nabla \cdot \mathbf{A} \nabla - \mathbf{b} \cdot \nabla - r)w^k = 0 \]  

(144)

In the above equation, the scheme becomes purely explicit when \( \sigma = 0 \), purely implicit when \( \sigma = 1 \), and becomes Crank-Nicolson scheme when \( \sigma = 0.5 \). The computation starts backwards from \( t^N = T \) when \( w^N \) is known which is given by initial condition. At each time-step, provided that the \( w^{k+1} \) is computed
from previous step, the $w^k$ can be solved from (144) recursively. Rearrangement of (144) gives a new equation below, in which all the $w^{k+1}$ terms are collected on the right hand side (RHS) and all the $w^k$ terms on the left hand side (LHS)

$$Lw^k = Rw^{k+1}$$

where

$$L = \sigma (\nabla \cdot A\nabla - b \cdot \nabla) - c$$

and

$$R = (\sigma - 1)(\nabla \cdot A\nabla - b \cdot \nabla) - (c - r)$$

(145)

$$c = \sigma r + \frac{1}{t^{k+1} - t^k}$$

3.2.2. Two-Dimensional Finite Element Method

Finite element method is based on weak formulation. It is ideal for solving PDE when the solution lacks smoothness and when the domain is irregular, dynamically changing, and/or unevenly spaced. Valuation of complex exotic options can often exhibit these features, which makes FEM an ideal tool for this type of applications.

3.2.2.1. Weak Formulation

Equation (145) can be transformed to a weak formulation by multiplying both sides with a scalar-valued test function $\psi$ (in the subsequent sections, we will see that the $\psi$ can be constructed as a linear combination of basis functions on the 2-D domain)

$$\int_{\Omega} \psi L w^k d\Omega = \int_{\Omega} \psi R w^{k+1} d\Omega$$

where

$$\text{LHS} = \sigma \int_{\Omega} \psi \nabla \cdot A\nabla w^k d\Omega - \sigma \int_{\Omega} \psi b \cdot \nabla w^k d\Omega - c \int_{\Omega} \psi w^k d\Omega$$

(146)

$$\text{RHS} = (\sigma - 1) \int_{\Omega} \psi \nabla \cdot A\nabla w^{k+1} d\Omega - (\sigma - 1) \int_{\Omega} \psi b \cdot \nabla w^{k+1} d\Omega - (c - r) \int_{\Omega} \psi w^{k+1} d\Omega$$

where $\Omega$ is the open domain given in (134).

Assuming that $u$ is a scalar function and $\mathbf{v}$ a vector-valued function, both are continuously differentiable, then the integration by parts in multi-dimension is given by
\[
\int_{\Omega} u \nabla \cdot v \, d\Omega = \int_{\Gamma} u v \cdot n \, d\Gamma - \int_{\Omega} \nabla u \cdot v \, d\Omega
\]  
(147)

where \( \Gamma \) denotes the piecewise smooth boundary, and \( n \) denotes the outward unit surface normal to \( \Gamma \).

Applying the transformation (147) to (146) yields (here, it becomes the first Green’s identity)

\[
\int_{\Omega} \psi L w^k \, d\Omega = \int_{\Omega} \psi R w^{k+1} \, d\Omega \quad \text{where}
\]

\[
\text{LHS} = \sigma \int_{\Gamma} \psi A \nabla w^k \cdot n \, d\Gamma - \sigma \int_{\Omega} \nabla \psi \cdot A \nabla w^k \, d\Omega - \sigma \int_{\Omega} \psi b \cdot \nabla w^k \, d\Omega - c \int_{\Omega} \psi w^k \, d\Omega
\]

\[
\text{RHS} = (\sigma - 1) \int_{\Gamma} \psi A \nabla w^{k+1} \cdot n \, d\Gamma - (\sigma - 1) \int_{\Omega} \nabla \psi \cdot A \nabla w^{k+1} \, d\Omega
\]

\[
- (\sigma - 1) \int_{\Omega} \psi b \cdot \nabla w^{k+1} \, d\Omega - (c - r) \int_{\Omega} \psi w^{k+1} \, d\Omega
\]

(148)

Now let’s define a few functional spaces for the trial solution \( w \) and the test function \( \psi \)

\[
L^2(\Omega) := \{ f : \| f \|_{L^2} < \infty \} \quad \text{where} \quad \| f \|_{L^p} = \left( \int_{\Omega} |f|^p \, d\Omega \right)^{1/p}
\]

\[
H^1(\Omega) := \{ f \in L^2(\Omega) : D^1 f \in L^2(\Omega) \}
\]

\[
H^1_0(\Omega) := \{ f \in H^1(\Omega) : f = 0 \text{ on } \Gamma \}
\]

\[
H^1_\delta(\Omega) := \{ f \in H^1(\Omega) : f = g \text{ on } \Gamma \}
\]

(149)

where \( D^1 f \) denotes the first order weak partial derivatives of function \( f \). The \( L^2(\Omega) \) is the Lebesgue space with Euclidean norm which coincides with Hilbert space in this case. The \( H^1(\Omega) \), \( H^1_0(\Omega) \) and \( H^1_\delta(\Omega) \) are all Sobolev space. The variational form (148) is also called the weak form. The requirements of the solution \( w \) in (148) have been considerably weakened over the strong form in (142). In the weak formulation of the problem, the trial solution \( w \) and the test function \( \psi \) must belong to \( H^1(\Omega) \), it is not necessary that all functions and derivatives be continuous.

3.2.2.2. **Boundary Conditions**

Given different types of boundary conditions, (148) can be further transformed as follows
3.2.2.2.1. **Homogeneous Dirichlet Boundary Condition** \((w = 0 \text{ on } \Gamma)\)

In this case, \(w \in H^1_0(\Omega)\) and \(\psi \in H^1_0(\Omega)\). The boundary integral \(\oint_{\Gamma} \psi \mathbf{A} \nabla w \cdot \mathbf{n} d\Gamma\) in (148) vanishes because \(\psi\) is zero on \(\Gamma\). We gather the known terms on the right hand side, and the unknowns on the left hand side, which yields the equation below

\[
\sigma \int_{\Omega} \nabla \psi \cdot \nabla w^k d\Omega + \sigma \int_{\Omega} \psi \mathbf{b} \cdot \nabla w^k d\Omega + c \int_{\Omega} \psi w^k d\Omega
\]

\[
= (\sigma - 1) \int_{\Omega} \nabla \psi \cdot \nabla w^{k+1} d\Omega + (\sigma - 1) \int_{\Omega} \psi \mathbf{b} \cdot \nabla w^{k+1} d\Omega + (c - r) \int_{\Omega} \psi w^{k+1} d\Omega
\]

3.2.2.2.2. **Inhomogeneous Dirichlet Boundary Condition** \((w = g \text{ on } \Gamma)\)

Since \(g\) is a function defined on \(\Gamma\), we assume that there is a function \(G \in H^1(\Omega)\) such that \(G = g\) on \(\Gamma\). It turns out that \(\psi \in H^1_0(\Omega)\) still holds in this case. The solution has the form \(w = u + G\) and \(w \in H^1_0(\Omega)\), where \(G\) is assumed to be known and \(u \in H^1_0(\Omega)\) is unknown. The boundary integral \(\oint_{\Gamma} \psi \mathbf{A} \nabla w \cdot \mathbf{n} d\Gamma\) in (148) vanishes because \(\psi\) is zero on \(\Gamma\), which yields

\[
\sigma \int_{\Omega} \nabla \psi \cdot \nabla w^k d\Omega + \sigma \int_{\Omega} \psi \mathbf{b} \cdot \nabla w^k d\Omega + c \int_{\Omega} \psi w^k d\Omega
\]

\[
= (\sigma - 1) \int_{\Omega} \nabla \psi \cdot \nabla w^{k+1} d\Omega + (\sigma - 1) \int_{\Omega} \psi \mathbf{b} \cdot \nabla w^{k+1} d\Omega + (c - r) \int_{\Omega} \psi w^{k+1} d\Omega
\]

Writing \(w^k\) in terms of sum of function \(u^k\) and \(G^k\), we have

\[
\sigma \int_{\Omega} \nabla \psi \cdot \nabla u^k d\Omega + \sigma \int_{\Omega} \psi \mathbf{b} \cdot \nabla u^k d\Omega + c \int_{\Omega} \psi u^k d\Omega
\]

\[
= (\sigma - 1) \int_{\Omega} \nabla \psi \cdot \nabla w^{k+1} d\Omega + (\sigma - 1) \int_{\Omega} \psi \mathbf{b} \cdot \nabla w^{k+1} d\Omega + (c - r) \int_{\Omega} \psi w^{k+1} d\Omega
\]

\[
- \sigma \int_{\Omega} \nabla \psi \cdot \nabla G^k d\Omega - \sigma \int_{\Omega} \psi \mathbf{b} \cdot \nabla G^k d\Omega - c \int_{\Omega} \psi G^k d\Omega
\]
3.2.2.2.3. **Homogeneous Neumann Boundary Condition** $(\nabla w = 0 \text{ on } \Gamma)$

In this case, $w \in H^1(\Omega)$ and $\psi \in H^1(\Omega)$. The boundary integral $\oint_{\Gamma} \psi A \nabla w \cdot n d\Gamma$ in (148) vanishes because $\nabla w \cdot n$ is zero on $\Gamma$ (Note that this is different from that of Dirichlet boundary condition, in which it was the test function $\psi = 0$ on $\Gamma$ that causes the boundary integral to vanish), which yields

$$
\sigma \int_{\Omega} \nabla \psi \cdot A \nabla w^k d\Omega + \sigma \int_{\Omega} \psi b \cdot \nabla w^k d\Omega + c \int_{\Omega} \psi w^k d\Omega
$$

$$
= (\sigma - 1) \int_{\Omega} \nabla \psi \cdot A \nabla w^{k+1} d\Omega + (\sigma - 1) \int_{\Omega} \psi b \cdot \nabla w^{k+1} d\Omega + (c - r) \int_{\Omega} \psi w^{k+1} d\Omega
$$

(153)

It should be noted that the weak form is the same as that of the homogeneous Dirichlet boundary condition; however the trial and test function spaces are different.

3.2.2.2.4. **Inhomogeneous Neumann Boundary Condition** $(\nabla w = h \text{ on } \Gamma)$

In this case, the $w \in H^1(\Omega)$ and $\psi \in H^1(\Omega)$. The boundary integral $\oint_{\Gamma} \psi A \nabla w \cdot n d\Gamma$ is known, because $\nabla w \cdot n$ on $\Gamma$ is given. Thus we have

$$
\sigma \int_{\Omega} \nabla \psi \cdot A \nabla w^k d\Omega + \sigma \int_{\Omega} \psi b \cdot \nabla w^k d\Omega + c \int_{\Omega} \psi w^k d\Omega
$$

$$
= (\sigma - 1) \int_{\Omega} \nabla \psi \cdot A \nabla w^{k+1} d\Omega + (\sigma - 1) \int_{\Omega} \psi b \cdot \nabla w^{k+1} d\Omega + (c - r) \int_{\Omega} \psi w^{k+1} d\Omega
$$

$$
- (\sigma - 1) \oint_{\Gamma} \psi A h^{k+1} \cdot n d\Gamma + \sigma \oint_{\Gamma} \psi A h^k \cdot n d\Gamma
$$

(154)

Boundaries involve a combination of these boundary conditions will be treated accordingly in a similar manner [27]. The problem studied in this essay is a 2D dynamic problem with inhomogeneous Dirichlet boundary condition. Hence we will primarily focus on (152).

3.2.2.3. **Mesh and Basis Functions**

A triangular mesh (Figure 1) is used for the computation. The mesh can be either either structured or unstructured. We denote the closed domain $\Omega = \Omega \cup \Gamma = [v_{min}, v_{max}] \times [y_{min}, y_{max}]$. The $\Omega$ can be decomposed into a mesh consisting of triangles $T_i$, $i = 1, \cdots, N^T$ and their nodes $D_i$, $i = 1, \cdots, N^D$. A mesh generator provides the mesh information by two arrays. The first array indexes all the nodes with
their corresponding coordinates and boundary ID’s. The boundary ID indicates the boundary information of a node (for example, whether an interior (free) node or a boundary (constrained) node, which boundary it belongs to, etc.). The second array indexes all the triangular elements with their three vortices. Figure 1 illustrates a simple evenly spaced structured mesh used in the computation. It is obvious that the mesh consists of total $I \times J$ squares, $N^T = 2IJ$ triangles and $N^D = (I + 1)(J + 1)$ nodes, where $I$ and $J$ are the spatial resolution for variable $v$ and $y$, respectively. The basis functions $\phi_i$, $i = 1, \cdots, N^D$ are defined as piecewise linear functions that take the value 1 at node $D_i$ and 0 at all other nodes. This basis is usually called Lagrange basis or nodal basis.

![Structured Triangular Mesh](image)

**Figure 1. A structured triangular mesh**

To simplify the computation of integrals over the triangles, each triangle is mapped by affine transformation to a reference triangle $\hat{T} = \{ (\hat{v}, \hat{y}) \in \mathbb{R}^2 : \hat{v} \in [0, 1], \hat{y} \in [0, 1 - \hat{v}] \}$, where a “hat” is used to denote the variables defined on the reference triangle. The reference triangle is shown in Figure 2. The three vortices $\hat{D}_i, i = 1, 2, 3$ are $(0, 0), (1, 0)$ and $(0, 1)$, respectively. The open circle denotes $\hat{\phi}_i = 0$ at node $\hat{D}_j$ if $i \neq j$ while the full circle denotes $\hat{\phi}_i = 1$ at node $\hat{D}_j$ if $i = j$. Hence the three basis functions on the reference triangle are

$$
\hat{\phi}_0(\hat{v}, \hat{y}) = 1 - \hat{v} - \hat{y}, \quad \hat{\phi}_1(\hat{v}, \hat{y}) = \hat{v} \quad \text{and} \quad \hat{\phi}_2(\hat{v}, \hat{y}) = \hat{y}
$$  \hspace{1cm} (155)
3.2.2.4. **Affine Transformation**

The mapping is achieved by affine transformation. Considering an arbitrary triangle $T$ with vortices $D_0 = (v_0, y_0), D_1 = (v_1, y_1)$ and $D_2 = (v_2, y_2)$, the transformation maps the reference triangle $\hat{T}$ onto the triangle $T$ by

\[
(v, y) = F(\hat{v}, \hat{y}) = \mathbf{J}(\hat{v}/\hat{y}) + (v_0/y_0) \quad \text{and} \quad \mathbf{J} = \begin{bmatrix}
    v_1 - v_0 & v_2 - v_0 \\
    y_1 - y_0 & y_2 - y_0
\end{bmatrix}
\]

(156)

3.2.2.5. **Stiffness and Mass Matrix**

Our problem involves inhomogeneous Dirichlet boundary condition, e.g. $w = g$ on $\Gamma$. To incorporate such boundaries, we need to extend the function $g$ to any function $G$ on the closed domain $\tilde{\Omega}$ such that $G \in H^1(\Omega)$ and $G = g$ on $\Gamma$. It is not easy to find such $G$ function exactly, but it is easy to define a function $G$ that approximately satisfies the boundary condition. Indeed, following the same definition of the basis function $\phi$, the continuous piecewise linear function $G$ defined by

\[
G(D_i) = \begin{cases}
    0, & \text{if } D_i \not\in \Gamma \\
    g_i \phi_i, & \text{if } D_i \in \Gamma
\end{cases}
\]

(157)

agrees with $g$ at the endpoints of every constrained edge. The linear combination

\[
G = \sum_{D_i \in \Gamma} g_i \phi_i
\]

(158)

is a sufficiently good approximation to $g$ for the purpose of the finite element method. As discussed previously, the trial solution has the form $w = u + G$ where $G \in H^1(\Omega)$ is known, $u \in H^1_0(\Omega)$ is unknown and $w \in H^1_g(\Omega)$. In the discrete case, at time $t^k$, the trial solution $w^k$ can be expressed as

Figure 2. The reference triangle and the basis functions
\[ w^k = G^k + u^k \quad \text{where} \quad G^k = \sum_{D_i \in \Gamma} g_i^k \phi_i \quad \text{and} \quad u^k = \sum_{D_i \in \Omega} u_i^k \phi_i \]

The test function \( \psi \in H^1_0(\Omega) \) can be chosen as a linear combination of the basis functions \( \psi = \sum_{D_i \in \Omega} b_i \phi_i \). Although \( b_i \) can be arbitrarily chosen, it has no impact because eventually it will be cancelled out when we multiply the test function to both side of the PDE. As such, it allows us to simplify the test function to \( \psi = \sum_{D_i \in \Omega} \phi_i \).

According to (148), we have

\[
\int_{\Omega} \left( \sum_{D_i \in \Omega} \phi_i \right) \mathcal{L} \left( \sum_{D_i \in \Omega} g_i^k \phi_i + \sum_{D_i \in \Omega} u_i^k \phi_i \right) d\Omega = \int_{\Omega} \left( \sum_{D_i \in \Omega} \phi_i \right) \mathcal{R} \left( \sum_{D_i \in \Omega} w_i^{k+1} \phi_i \right) d\Omega
\]

and we will try to solve the following linear system

\[
\int_{\Omega} \left( \sum_{D_i \in \Omega} \phi_i \right) \mathcal{L} \left( \sum_{D_i \in \Omega} u_i^k \phi_i \right) d\Omega = \int_{\Omega} \left( \sum_{D_i \in \Omega} \phi_i \right) \mathcal{R} \left( \sum_{D_i \in \Omega} w_i^{k+1} \phi_i \right) d\Omega - \int_{\Omega} \left( \sum_{D_i \in \Omega} \phi_i \right) \mathcal{L} \left( \sum_{D_i \in \Gamma} g_i^k \phi_i \right) d\Omega
\]

This is consistent with (152), and can be subsequently translated into the following linear system

\[
S_{ij} u^k_{f \times f} = M_{ij} w^{k+1} - P_{ij} g^k_{f \times c}
\]

where

\[
S_{ij} = \int_{\Omega} \phi_i \mathcal{L} \phi_j d\Omega \quad \forall \ D_i, D_j \in \Omega, \quad M_{ij} = \int_{\Omega} \phi_i \mathcal{R} \phi_j d\Omega \quad \forall \ D_i \in \Omega, \ D_j \in \Omega
\]

\[
P_{ij} = \int_{\Omega} \phi_i \mathcal{L} \phi_j d\Omega \quad \forall \ D_i \in \Omega, \ D_j \in \Gamma
\]

\( S, M \) and \( P \) are matrices of size \( f \times f, f \times n \) and \( f \times c \) respectively, where the \( f, n \) and \( c \) denote the number of free (interior) nodes, total nodes and constrained (Dirichlet boundary) nodes, respectively. Note that these matrices can be evaluated from the basis functions, and because the \( w^{k+1} \) and \( g^k \) are known, the \( u^k \) can be computed using a linear system solver.
Because constructing the matrix $S$ and $P$ needs a 1-to-1 mapping of the matrix entries onto the node indices, the process is complicated and error prone. Alternatively, we consider another and yet simpler approach for the implementation. This approach can be easily adapted to other boundary conditions with minor modifications. The steps are described as follows:

1. We first construct a matrix $S$ such that $S_{ij} = \int_{\Omega} \phi_i \mathcal{L} \phi_j d\Omega \ \forall \ D_i, D_j \in \Omega$ (rather than $D_i, D_j \in \Omega$ originally for $S_{ij}$) and a matrix $M$ such that $M_{ij} = \int_{\Omega} \phi_i \mathcal{R} \phi_j d\Omega \ \forall \ D_i, D_j \in \Omega$ (rather than $D_i \in \Omega$ and $D_j \in \Omega$ originally for $M_{ij}$). Both $S$ and $M$ are $(n \times n)$ matrix, where $n = N^D$ is the total number of nodes in the closed domain $\Omega$. The matrix $S$ can be subdivided into four sub-matrices $S = (S_{ff} \ S_{fc} \ S_{cf} \ S_{cc})$, where the subscript $f$ and $c$ represent a sub-matrix $S_{ij}$ corresponding to integrals between the free and constrained nodes. The sub-matrix $S_{fc}$, for example, is defined as $S_{ij} = \int_{\Omega} \phi_i \mathcal{L} \phi_j d\Omega$ with $D_i \in \Omega$ (free nodes) and $D_j \in \Gamma$ (constrained nodes). Similarly, we have $M = (M_{ff} \ M_{fc} \ M_{cf} \ M_{cc})$.

2. If $D_i \in \Gamma$, the value of $w_i^k = G_i^k = g^k(D_i)$ is known. Hence we create an $(n \times 1)$ vector $\omega^k$ such that $\omega_i^k = \begin{cases} 0, & \text{if } D_i \in \Omega \\ G_i^k, & \text{if } D_i \in \Gamma \end{cases}$. In other words, $\omega^k = \left( \begin{array}{c} 0 \\ w_i^k \end{array} \right)$ if we subdivide the vector $w^k = \left( \begin{array}{c} w_f^k \\ w_c^k \end{array} \right)$ by following the same definition as in step 1.

3. We create a vector $h$ as a right hand side term such that

$$
\mathbf{h}_{n \times 1} = \mathbf{M}_{n \times n} \mathbf{w}_{n \times 1}^{k+1} - S_{n \times n} \omega_{n \times 1}^k = (M_{ff} \ M_{fc} \ M_{cf} \ M_{cc}) (w_f^{k+1} \ w_c^{k+1}) - (S_{ff} \ S_{fc} \ S_{cf} \ S_{cc}) (0 \ w_c^k) = (M_{ff}w_f^{k+1} + M_{fc}w_c^{k+1} - S_{fc}w_c^k) \\
(M_{cf}w_f^{k+1} + M_{cc}w_c^{k+1} - S_{cc}w_c^k)
$$

then manually assign $h_i = G_i^k \ \forall \ D_i \in \Gamma$, that is
\[ h = \left( \frac{M_{ff}w_{f}^{k+1} + M_{fc}w_{c}^{k+1} - S_{fc}w_{c}^{k}}{w_{c}^{k}} \right) \] (164)

4. Because \( w_{c}^{k} \) is known, computation for this part is not necessary. We further process matrix \( S \) by replacing its sub-matrix \( S_{cc} \) with an identity matrix and replacing \( S_{fc}, S_{cf} \) with zero matrices. We end up with \( S = \begin{pmatrix} S_{ff} & 0 \\ 0 & I \end{pmatrix} \).

5. Finally, we have an augmented linear system.

\[
S_{n \times n \times 1} w^{k} = h \Rightarrow \begin{pmatrix} S_{ff} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} w_{f}^{k} \\ w_{c}^{k} \end{pmatrix} = h \Rightarrow \begin{pmatrix} S_{ff}w_{f}^{k} \\ w_{c}^{k} \end{pmatrix} = \begin{pmatrix} M_{ff}w_{f}^{k+1} + M_{fc}w_{c}^{k+1} - S_{fc}w_{c}^{k} \end{pmatrix} \] (165)

\[ \Rightarrow S_{ff}w_{f}^{k} = M_{ff}w_{f}^{k+1} + M_{fc}w_{c}^{k+1} - S_{fc}w_{c}^{k} \]

The last equation is equivalent to (162). The solution \( w^{k} = S^{-1}h \) can be solved by an iterative linear solver.

3.2.2.6. Computation of Integrals

Based on previous discussion, we know that the construction of the matrix \( S \) and \( M \) is directly related to the computation of the integral \( \int_{\Omega} \phi_{i}L\phi_{j} \, d\Omega \) and \( \int_{\Omega} \phi_{i}R\phi_{j} \, d\Omega \)

\[ S_{ij} = \int_{\Omega} \phi_{i}L\phi_{j} \, d\Omega \quad \forall \, D_{i}, D_{j} \in \tilde{\Omega} \quad \text{and} \quad M_{ij} = \int_{\Omega} \phi_{i}R\phi_{j} \, d\Omega \quad \forall \, D_{i}, D_{j} \in \tilde{\Omega} \] (166)

The computation of the two integrals shares a great similarity. Each can be decompose into three parts

\[
\int_{\Omega} \phi_{i}L\phi_{j} \, d\Omega = \sigma \int_{\Omega} \nabla \phi_{i} \cdot A \nabla \phi_{j} \, d\Omega + \sigma \int_{\Omega} \phi_{i} b \cdot \nabla \phi_{j} \, d\Omega + c \int_{\Omega} \phi_{i} \phi_{j} \, d\Omega \quad \text{and} \quad \int_{\Omega} \phi_{i}R\phi_{j} \, d\Omega = (\sigma - 1) \int_{\Omega} \nabla \phi_{i} \cdot A \nabla \phi_{j} \, d\Omega + (\sigma - 1) \int_{\Omega} \phi_{i} b \cdot \nabla \phi_{j} \, d\Omega + (c - r) \int_{\Omega} \phi_{i} \phi_{j} \, d\Omega \] (167)

Instead of directly computing the integral over the whole domain \( \Omega \), we can compute the integral element by element and then sum all the pieces together. Given a triangular element \( T \), it is obvious that there are locally only three of the basis functions (corresponding to the three nodes) are nonzero. The reference triangle \( \tilde{T} \) is mapped onto the triangular element \( T \) and the integral evaluated over \( \tilde{T} \) is then transformed back to \( T \).
The first integral appearing in the right hand side of (167) is denoted by \( P_1 \). Over a given triangular element \( T \), the \( P_1 \) is a \( 3 \times 3 \) matrix and is evaluated as

\[
P_1 = \left[ \int_{T} \mathbf{V} \phi_i \cdot \mathbf{A} \mathbf{V} \phi_j dT \right]_{i,j=0,1,2} = \left[ \int_{\hat{T}} (\mathbf{J}^{-T} \mathbf{V} \hat{\phi}_i)^T (\mathbf{A} \mathbf{J}^{-T} \mathbf{V} \hat{\phi}_j) \left[ \frac{\partial F(\hat{\mathbf{v}}, \hat{\mathbf{y}})}{\partial (\mathbf{v}, \mathbf{y})} \right] d\hat{T} \right]_{i,j=0,1,2}
\]

\[
= |\text{det} \mathbf{J}| \left[ \int_{\hat{T}} (\mathbf{J}^{-T} \mathbf{V} \hat{\phi}_i)^T (\mathbf{A} \mathbf{J}^{-T} \mathbf{V} \hat{\phi}_j) d\hat{T} \right]_{i,j=0,1,2} = |\text{det} \mathbf{J}| \left( \mathbf{J}^{-T} \hat{\mathbf{G}} \right)^T \left( \int_{\hat{T}} \mathbf{A} d\hat{T} \right) \hat{\mathbf{G}}^{-T}
\]

\[
\approx \frac{1}{2} |\text{det} \mathbf{J}| \hat{\mathbf{G}}^T \hat{\mathbf{A}}^{-1} \hat{\mathbf{J}} \hat{\mathbf{G}}^{-T}
\]

where \( \hat{\mathbf{G}} = [\mathbf{V} \hat{\phi}_j]_{j=0,1,2} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \) and

\[
\int_{\hat{T}} \mathbf{A} d\hat{T} \approx \hat{\mathbf{A}} \int_{\hat{T}} d\hat{T} = \left( \text{average of } \mathbf{A} \text{ over triangular element} \right) \times \left( \text{area over reference triangle} \right) = \frac{\sum_{n=0}^{2} \mathbf{A}_n}{3} \cdot \frac{1}{2}
\]

In (168), the subscript \( i \) and \( j \) denote the matrix row and column index respectively, the \( |\text{det} \mathbf{J}| \) is the determinant absolute of matrix \( \mathbf{J} \), the \( \hat{\mathbf{G}} \) is the gradient matrix in \( \hat{T} \) for the three local basis functions defined in (155). Since \( \hat{\mathbf{G}} \) is constant, it can be moved out of the integral. The \( \mathbf{A}_n \), \( n = 0,1,2 \) are the matrix \( \mathbf{A} \) evaluated at the three vortices of the elemental triangle \( T \). To compute the integral \( \int_{\hat{T}} \mathbf{A} d\hat{T} \), matrix \( \mathbf{A} \) is assumed to be constant within \( T \) and takes an average over the three vortices.

The second integral denoted by \( P_2 \) is computed as

\[
P_2 = \left[ \int_{T} \phi_i \mathbf{b} \cdot \mathbf{V} \phi_j dT \right]_{i,j=0,1,2} \approx |\text{det} \mathbf{J}| \left[ \int_{\hat{T}} \phi_i d\hat{T} \right]_{i=0,1,2} \hat{\mathbf{b}}^T \hat{\mathbf{G}}^{-T}
\]

\[
= |\text{det} \mathbf{J}| \frac{1}{6} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \hat{\mathbf{b}}^T \hat{\mathbf{G}}^{-T} \quad \text{where} \quad \hat{\mathbf{b}} = \frac{1}{3} \sum_{n=0}^{2} \mathbf{b}_n
\]

In (169), the \( \left[ \int_{\hat{T}} \phi_i d\hat{T} \right]_{i=0,1,2} \) is simply the volume below the basis functions integrated over the reference triangle \( \hat{T} \). Similarly as with matrix \( \mathbf{A} \), The vector \( \mathbf{b} \) is treated constant within the element triangle and takes an average over the three vortices. Hence it can be moved out of the integral along with the \( \hat{\mathbf{G}} \) matrix.

The computation of the third integral \( P_3 \) is even simpler
\[
P_3 = \left[ \int_T \phi_i \phi_j dT \right]_{i,j=0,1,2} = |\text{det } J| \left[ \int_T \hat{\phi}_i \hat{\phi}_j d\hat{T} \right]_{i,j=0,1,2} = |\text{det } J| \frac{1}{24} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}
\]

(170)

It is obvious that \(P_1, P_2\) and \(P_3\) are all \(3 \times 3\) matrix and just \(P_2\) appears asymmetric. Within each element triangle, we have

\[
\left[ \int_T \phi_i \mathcal{L} \phi_j dT \right]_{i,j=0,1,2} = \sigma(P_1 + P_2) + cP_3 \quad \text{and}
\]

\[
\left[ \int_T \phi_i \mathcal{R} \phi_j dT \right]_{i,j=0,1,2} = (\sigma - 1)(P_1 + P_2) + (c - r)P_3
\]

(171)

Aggregating integrals over all the triangular elements in the domain, the global stiffness matrix \(\mathcal{S}\) and mass matrix \(\mathcal{M}\) can then be assembled.

### 3.2.3. Iterative Linear Solvers

#### 3.2.3.1. Implementation of Sparse Matrix

As one can see, due to the use of localized basis functions, each node can interact with at most, including itself, seven nodes. The stiffness and mass matrix can therefore be sparse. Since this is a 2-D problem, as the resolution increases, the storage for the matrix increases quadratically. To avoid the unnecessary memory storage for the zeros, a sparse matrix implementation is developed. The rows of a sparse matrix are implemented as an array. Each entry of the array consists of a map, which maps the column index to the matrix entry. Each map stores all the nonzero elements of a row vector of the sparse matrix. Given a matrix index \(i\) and \(j\), it first finds the \(i\)-th row through the map array, and then locate the \(j\)-th column entry through the \(i\)-th map. Using this implementation one can easily and efficiently access the matrix members. This turns out to be very handy for constructing the sparse matrices.

Since the matrix is generally very large, for example, a \(100 \times 100\) domain will generate (roughly) a \(10000 \times 10000\) matrix. In order to apply the iterative methods for the linear system, the matrix-vector multiplication operation must be extremely efficient. To address this issue, the fully constructed sparse matrix is then converted to a compressed row storage format, which consists of only three arrays: two
arrays store the matrix indices $i$ and $j$ respectively, and one array stores the corresponding matrix entries. Although adding or removing the matrix entries is no longer as straightforward as before, the matrix multiplications can be done very efficiently.

3.2.3.2. *Iterative Methods for Asymmetric Sparse Matrix*

The linear systems resulted from this problem are usually very large. Traditional direct method and simple iterative methods are no longer appropriate for this type of applications. Furthermore, as one can see that the $P_2$ in (169) is asymmetric, and so is the global stiffness matrix $\mathcal{S}$. In this case, the widely used Conjugate Gradient method is no longer applicable. Instead, the Conjugate Gradient Squared (CGS) and/or Generalized Minimal Residual (GMRES) method must be used to solve the linear systems. Since the two methods have been widely documented, the details of the algorithm will not be described here. The implementation of the two methods is based on Kelley’s book [28]. In general, CGS method performs faster than GMRES method as it has much fewer floating point operations in each iteration, however in some cases CGS method does not guarantee a convergence. In our application, the tolerance for convergence is set to $1.0 \times 10^{-8}$.

3.2.4. *Interpolation of Numerical Solution*

Numerical solution to the PDE is provided as a vector $w$ with values at each of the prescribed nodes. To obtain a solution for an arbitrary position in the domain, interpolation must be performed. Since the mesh can be either structured or unstructured, bilinear interpolation is not always appropriate. In this case, the point coordinate is converted to a barycentric coordinate in an element triangle to perform the linear interpolation. Firstly, it is desired to determine the triangle where the point locates. Given a point coordinate $(v, y)$, the barycentric coordinates $\lambda_0, \lambda_1,$ and $\lambda_2$ at this point can be computed based upon the three triangle vortices $((v_0, y_0), (v_1, y_1), (v_2, y_2))$

$$
\begin{pmatrix}
\lambda_0 \\
\lambda_1
\end{pmatrix} = \begin{pmatrix}
v_0 - v_2 & v_1 - v_2 \\
v_0 - y_2 & y_1 - y_2
\end{pmatrix}^{-1} \begin{pmatrix}
v - v_2 \\
y - y_2
\end{pmatrix}, \quad \lambda_2 = 1 - \lambda_0 - \lambda_1
$$

(172)

If $\lambda_i \in [0,1] \forall i = 0,1,2$, then the point is inside the triangle and the interpolated solution is
\[ w(v,y) = \lambda_0 w(v_0,y_0) + \lambda_1 w(v_1,y_1) + \lambda_2 w(v_2,y_2) \]  

(173)

3.3. Solutions and Discussions

The correctness of the algorithm is examined against the solutions obtained from closed form formulas for special cases. All the tests are based on the default model parameters listed in Table 1 unless otherwise stated explicitly. The CGS is used as the default linear solver. Crank-Nicolson scheme is assumed for the temporal discretization.

<table>
<thead>
<tr>
<th>Table 1. default model parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Spot Price ( S_0 ) 100</td>
</tr>
<tr>
<td>Initial Variance ( v_0 ) 0.25</td>
</tr>
<tr>
<td>Mean Variance ( \theta ) 0.09</td>
</tr>
<tr>
<td>Mean Reversion Rate ( \kappa ) 1.0</td>
</tr>
<tr>
<td>Volatility of Variance ( \xi ) 0.4</td>
</tr>
<tr>
<td>Correlation ( \rho ) -0.7</td>
</tr>
<tr>
<td>Spatial Resolution 150 ( \times ) 150</td>
</tr>
<tr>
<td>Temporal Resolution 150/year</td>
</tr>
</tbody>
</table>

3.3.1. Vanilla Call

<table>
<thead>
<tr>
<th>Table 2. vanilla call with different temporal and spatial resolutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution in ( v )</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>50</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>150</td>
</tr>
<tr>
<td>200</td>
</tr>
<tr>
<td>400</td>
</tr>
</tbody>
</table>

The first test is to price a vanilla call option. Winkler’s implementation [29] of closed form formula for call options is used as a benchmark. In this test, the domain is set to \( v \in [0, 3] \) and \( S \in [25, 400] \). The computational time is estimated for a fully optimized program running on an Intel Core 2 Duo 1.86 GHz processor. Varying the resolution in \( v, S \) and \( t \), a series of prices can be computed and the results are shown in Table 2. It shows that the solutions are quite accurate when the resolution is at or above 100. As
one can see, the CGS solver performs very well. It takes only 5.5 minutes to solve a problem with resolution as large as $400 \times 400 \times 400$.

To examine the convergence property with respect to resolution in time, the spatial resolution is fixed at $150 \times 150$, the temporal resolution varies from 5 to 200. The results are listed in Table 3. It shows that the result is still acceptable even with a temporal resolution down to 5. The accuracy is improved rapidly as the temporal resolution increases.

<table>
<thead>
<tr>
<th>temporal resolution</th>
<th>Computational Time (seconds)</th>
<th>Closed-Form</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>4.217</td>
<td></td>
<td>13.761879</td>
</tr>
<tr>
<td>10</td>
<td>4.883</td>
<td>13.853335</td>
<td>13.814776</td>
</tr>
<tr>
<td>20</td>
<td>5.879</td>
<td>13.835519</td>
<td>13.849429</td>
</tr>
<tr>
<td>50</td>
<td>8.220</td>
<td>13.854338</td>
<td>13.856847</td>
</tr>
<tr>
<td>100</td>
<td>11.414</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>16.155</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The vanilla call prices are also evaluated at different strikes and shown in Table 4. The computed prices are in excellent agreement with the closed form solutions. It should be noted that the computational domain must be adjusted accordingly for the strike of 200 due to its large value. Since the call option with a strike of 200 is far out of the money, it is expected that there is a relative large error in the computed option price. This can be seen from the last row of the Table 4.

<table>
<thead>
<tr>
<th>Strike $K$</th>
<th>Closed-Form</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>15.935692</td>
<td>15.939498</td>
</tr>
<tr>
<td>110</td>
<td>13.853335</td>
<td>13.856007</td>
</tr>
<tr>
<td>115</td>
<td>11.975267</td>
<td>11.980188</td>
</tr>
<tr>
<td>130</td>
<td>7.4758044</td>
<td>7.4790939</td>
</tr>
<tr>
<td>150</td>
<td>3.6874922</td>
<td>3.6916604</td>
</tr>
<tr>
<td>200</td>
<td>0.42266949</td>
<td>0.43570695</td>
</tr>
</tbody>
</table>
Table 5. vanilla call with different maturities

<table>
<thead>
<tr>
<th>Maturity Time</th>
<th>Closed-Form</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/12</td>
<td>2.180542</td>
<td>2.1765933</td>
</tr>
<tr>
<td>1/4</td>
<td>5.7929266</td>
<td>5.7890237</td>
</tr>
<tr>
<td>1/2</td>
<td>9.3177983</td>
<td>9.3154635</td>
</tr>
<tr>
<td>1</td>
<td>13.853335</td>
<td>13.856007</td>
</tr>
<tr>
<td>2</td>
<td>19.56252</td>
<td>19.44822</td>
</tr>
<tr>
<td>5</td>
<td>29.918653</td>
<td>28.500867</td>
</tr>
<tr>
<td>10</td>
<td>N/A</td>
<td>36.782293</td>
</tr>
</tbody>
</table>

Resolutions increase from 150 to 300:

| 5             | 29.918653   | 28.938853          |

Figure 3. A typical solution surface for a vanilla call

Table 5 shows the computed option price with different maturity times. As compared with the closed form solutions, the agreement is excellent except for those long term maturities (5 years and up). Larger domain and higher resolutions can improve the accuracy. This has been shown for the case of 5
years maturity time. A typical solution surface for the vanilla call options is shown in Figure 3. Note that the axis for stock process has been log transformed to $y = \ln \frac{s}{K}$.

3.3.2. **Vanilla Put**

The second test is to price a vanilla put option. In order to compare the results with closed form solutions, the dividend rate $q = 0$ is assumed. Rouah-Vainbergh’s Excel spreadsheet [30] is used as a benchmark, which analytically solves for the put option price. In this test, the domain is also set as $v \in [0, 3]$ and $S \in [25, 400]$. The put price is evaluated at different strikes. The results are shown in the Table 6. The computed prices are in excellent agreement with the closed form solutions, except for the strike of 20 due to the extreme out-of-the-money. It should be noted that the computational domain must be adjusted accordingly for small strikes.

<table>
<thead>
<tr>
<th>Strike $K$</th>
<th>Rouah-Vainbergh’s Spreadsheet</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>95</td>
<td>N/A</td>
<td>11.816262</td>
</tr>
<tr>
<td>90</td>
<td>9.8286519</td>
<td>9.8259365</td>
</tr>
<tr>
<td>85</td>
<td>8.0667635</td>
<td>8.0587285</td>
</tr>
<tr>
<td>70</td>
<td>3.9702575</td>
<td>3.9744305</td>
</tr>
<tr>
<td>50</td>
<td>1.0635452</td>
<td>1.0610897</td>
</tr>
<tr>
<td>20</td>
<td>0.0118353</td>
<td>0.018820474</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maturity Time $T$</th>
<th>Rouah-Vainbergh’s Spreadsheet</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/12</td>
<td>1.8136404</td>
<td>1.8153803</td>
</tr>
<tr>
<td>1/4</td>
<td>4.6993203</td>
<td>4.7001984</td>
</tr>
<tr>
<td>1/2</td>
<td>7.2118815</td>
<td>7.2117123</td>
</tr>
<tr>
<td>1</td>
<td>9.8286519</td>
<td>9.8259365</td>
</tr>
<tr>
<td>2</td>
<td>N/A</td>
<td>11.796426</td>
</tr>
</tbody>
</table>

Table 7 shows the computed put prices with different maturities. As compared with the benchmark values, the computed prices are very close to the closed form solutions. A typical solution surface for the vanilla put options is shown in Figure 4.
3.3.3. **Double Barrier Knock-out Call**

Table 8. double barrier knock-out call with different barriers and strikes

<table>
<thead>
<tr>
<th>Down Barrier</th>
<th>Up Barrier</th>
<th>Strike K</th>
<th>Faulhaber’s Spreadsheet</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>200</td>
<td>110</td>
<td>5.99628132</td>
<td>5.986791</td>
</tr>
<tr>
<td>50</td>
<td>150</td>
<td>110</td>
<td>1.296335529</td>
<td>1.2971659</td>
</tr>
<tr>
<td>80</td>
<td>120</td>
<td>110</td>
<td>0.006384768</td>
<td>0.0066368712</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>100</td>
<td>8.812262221</td>
<td>8.8034578</td>
</tr>
<tr>
<td>50</td>
<td>150</td>
<td>100</td>
<td>2.655277021</td>
<td>2.6582231</td>
</tr>
<tr>
<td>80</td>
<td>120</td>
<td>100</td>
<td>0.051849246</td>
<td>0.053989178</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>90</td>
<td>12.51875619</td>
<td>12.510153</td>
</tr>
<tr>
<td>50</td>
<td>150</td>
<td>90</td>
<td>4.766805909</td>
<td>4.7734182</td>
</tr>
<tr>
<td>80</td>
<td>120</td>
<td>90</td>
<td>0.166447999</td>
<td>0.17333944</td>
</tr>
</tbody>
</table>

Resolutions increase from 150 to 300:

| 80           | 120        | 90       | 0.166447999             | 0.16993712        |

Figure 4. A typical solution surface for a vanilla put
The third test is to price double barrier knock-out call options. To allow the existence of closed form solutions, the correlation $\rho = 0$ and the dividend rate, $q = r$, the risk free interest rate are assumed. Faulhaber’s Excel spreadsheet [31] is used as a benchmark, which analytically solve for the double barrier option prices. In this test, the domain is assumed as $v \in [0, 3]$ and $S \in [B_d, B_u]$. The double barrier call price is evaluated with different strikes and barriers. The results are shown in Table 8. The differences between the computed prices and the closed form solutions are quite small, except for those close-to-zero prices, in which the problems are ill-conditioned and are very difficult for the solver to compute a solution accurately. As shown in the table, higher resolution can improve the accuracy.

Table 9 shows the computed double barrier knock-out call prices with different maturity times. To avoid the close-to-zero prices, the down barrier $B_d = 50$ and the up barrier $B_u = 150$ are assumed. As compared with the benchmark values, the numerical solutions are accurate, especially for the short term options. A typical solution surface for the double barrier knock-out call options is shown in Figure 5. To avoid the out-of-the-money, the barriers have been assume to be $B_d = 50$ and $B_u = 150$.

<table>
<thead>
<tr>
<th>Maturity Time $T$</th>
<th>Rouah-Vainbergh Spreadsheet</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/12</td>
<td>2.073151427</td>
<td>2.056223</td>
</tr>
<tr>
<td>1/4</td>
<td>2.748504846</td>
<td>2.7439708</td>
</tr>
<tr>
<td>1/2</td>
<td>2.048575635</td>
<td>2.049062</td>
</tr>
<tr>
<td>1</td>
<td>1.296335529</td>
<td>1.2971659</td>
</tr>
<tr>
<td>2</td>
<td>0.744555858</td>
<td>0.77808438</td>
</tr>
</tbody>
</table>
3.3.4. Double Barrier Knock-out Put

Table 10. double barrier knock-out put with different barriers and strikes

<table>
<thead>
<tr>
<th>Down Barrier</th>
<th>Up Barrier</th>
<th>Strike K</th>
<th>Faulhaber’s Spreadsheet</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>200</td>
<td>110</td>
<td>21.87927024</td>
<td>21.872058</td>
</tr>
<tr>
<td>50</td>
<td>150</td>
<td>110</td>
<td>12.46280153</td>
<td>12.452305</td>
</tr>
<tr>
<td>80</td>
<td>120</td>
<td>110</td>
<td>0.235744502</td>
<td>0.24543624</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>100</td>
<td>15.93365322</td>
<td>15.927314</td>
</tr>
<tr>
<td>50</td>
<td>150</td>
<td>100</td>
<td>8.307725123</td>
<td>8.2976528</td>
</tr>
<tr>
<td>80</td>
<td>120</td>
<td>100</td>
<td>0.090320836</td>
<td>0.094025866</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>90</td>
<td>10.87854926</td>
<td>10.872598</td>
</tr>
<tr>
<td>50</td>
<td>150</td>
<td>90</td>
<td>4.905236115</td>
<td>4.8971387</td>
</tr>
<tr>
<td>80</td>
<td>120</td>
<td>90</td>
<td>0.014031445</td>
<td>0.014613467</td>
</tr>
</tbody>
</table>

The fourth test is to price a double barrier knock-out put option. In this case, the correlation $\rho = 0$ and the dividend rate, $q = r$, the risk free interest rate are assumed. Faulhaber’s excel spreadsheet is once again used as a benchmark. In this test, the domain is $v \in [0, 3]$ and $S \in [B_d, B_u]$. The double barrier put
price is evaluated with different strikes and barriers. The results are shown in Table 10 and Table 11. The computed prices are in good agreement with the closed form solutions. As usual, higher resolution leads to a better accuracy.

Table 11. double barrier knock-out put with different strikes

<table>
<thead>
<tr>
<th>Down Barrier</th>
<th>Up Barrier</th>
<th>Strike $K$</th>
<th>Faulhaber’s Spreadsheet</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>120</td>
<td>110</td>
<td>0.235744502</td>
<td>0.24543624</td>
</tr>
<tr>
<td>80</td>
<td>120</td>
<td>100</td>
<td>0.090320836</td>
<td>0.094025866</td>
</tr>
<tr>
<td>80</td>
<td>120</td>
<td>90</td>
<td>0.014031445</td>
<td>0.014613467</td>
</tr>
</tbody>
</table>

Table 12 shows the computed double barrier knock-out put prices with different maturity times. To avoid the close-to-zero prices, the down barrier $B_d = 50$ and up barrier $B_u = 150$ are assumed. A typical solution surface for the double barrier double barrier knock-out put options is shown in Figure 6. As compared with the benchmark values, the numerical solutions are accurate, especially for the short term options.

Table 12. double barrier knock-out put with different maturities

<table>
<thead>
<tr>
<th>Maturity Time $T$</th>
<th>Rouah-Vainbergh Spreadsheet</th>
<th>Crank-Nicolson FEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/12</td>
<td>1.790372434</td>
<td>1.7880031</td>
</tr>
<tr>
<td>1/4</td>
<td>4.564088509</td>
<td>4.5554437</td>
</tr>
<tr>
<td>1/2</td>
<td>5.606674031</td>
<td>5.5991206</td>
</tr>
<tr>
<td>1</td>
<td>4.905236115</td>
<td>4.8971387</td>
</tr>
<tr>
<td>2</td>
<td>3.368687428</td>
<td>3.3595535</td>
</tr>
</tbody>
</table>
3.3.5. Numerical Solutions of Double Barrier Knock-out Options

The Crank-Nicolson finite element solver is used to price the double barrier options for our problem. Specifically, the double barrier options are assumed to follow Heston stochastic volatility model and take the values for the parameters as listed in Table 1. The computation is performed on a domain defined as $v \in [0, 3]$ and $S \in [B_d, B_u]$. The spatial and temporal resolutions are set to various levels to gain an overview about the accuracy and convergence trend. The solutions are shown in Table 13. Since these two option prices are very close to zero, the results computed at lower resolutions may have large errors. As the resolution improves, we do see the convergence of the solution. Another observation is that at the same resolution the computational time for double barrier options is much longer than that of plain vanilla options. This is partially because that double barrier options usually have more dynamic features than plain vanillas and therefore they are more computationally challenging.
3.4. Conclusions

It has been illustrated how to price a double barrier knock-out option (the knock-in type barrier options can be priced in a similar manner, except that the vanilla option prices are used for Dirichlet boundaries) in a stochastic volatility model using finite element method. The numerical solutions of the testing cases are in excellent agreement with the closed form values. This strongly suggests the accuracy and efficiency advantages of this method, as compared with finite difference method and Monte Carlo simulations. With appropriate boundary condition definitions, this method can be easily extended to price more complicated exotic options. It can also be further improved by applying adaptive mesh refinement and/or higher order basis functions.
REFERENCES


