Forward start options in Heston model

Lund University

Henrik Sandler

August 2021

Contents

Chapter 1

Introduction

This is a master thesis in mathematical statistics department at Lund University, Sweden.

This is survey of different authors methods for forward start options in the Black - Scholes and the Heston model. To get to the forward start options, I first need to introduce some concepts. In chapter two, I give a description of Brownian motion, and the Black - Scholes dynamics, that is the starting point for almost all option pricing models, even though that model has flaws, it is still widely used as a benchmark to other models, also when we are talking about implied volatility, it is to avoid confusion, the Black - Scholes implied volatility that is used. There are three concepts that is being used many times in financial statistics, the change of measure, e.g. Girsanov theorem, Radon - Nikodym density processes, the risk - neutral valuation formula and Feynman–Kac theorem, that provides a link between the stochastic process and the partial differential equation. In chapter three, I look at the diffusion equation, also known as the heat equation. The volatility, the spread of the displacement, is modeled as Einstein modeled the spread of particles in a fluid. I start with the historical expose, to come to the forward and backward Kolmogorov equation. The diffusion equation gives, for me, an understanding of the volatility. In chapter four, I present the idea of local volatility, I am using Bruno Dupire's presentation and follow closely to his article, and I use the Fokker-Plank equation, i.e. the forward Kolmogorov equation. Many modern presentation of Dupire's local volatility uses the backward Kolmogorov, I choosed not to, as the backward Kolmogorov involves the concept of local time, which is beyond the scope of this master thesis. In chapter 5, I enter the huge topic of Fourier transform, since the late 1990:s the theory of Fourier transformation, that rest on the fact, the in probability, the characteristic function is the same as Fourier transformation, and by using the inversion theorem you recover the (transitional) density. A major breakthrough was the work by (Duffie, Pan, & Singleton, 1999). When he showed that affine jump diffusion. If the state-space (the asset price and the variance are linear) in the parameters, you can move to the well established theory of coupled Riccati equations. I only give a brief discussion of that area.

In chapter 6 I discuss the Heston model, it is a model that tries to better explain the non-flat volatility surface. I also derive the characteristic function for it. You can also use the martingale method to prove the Heston model, that was not used by Heston in his derivation. I chose not to include it. There is a good point to go into the log scale of the asset price, because it will make variance affine in the state space. In the last chapter I look at the forward start contracts, under the Black - Scholes and the Heston model. For the Heston model, I write down three different "ansatz" to the problem. I only look at the variance and (log) of the asset price, there are extensions that include random grant time and random interest rate. I choose not to include them in the article.

This would not have been possible without the help, support, and trust of my supervisor Magnus Wiktorsson, Lund University.

Helsingør, Denmark 23rd of August 2021

Chapter 2

Geometric Brownian Motion

2.1 Derivatives, Options

I follow the definition of (Björk, 2009)

Definition 2.1.1 (European Call option)**.** A European call option on the amount of X US dollars, with strike price K $SEK/\$$ and exercise date T, is a contract written at $t = 0$, with the following properties

- The holder of the contract has, exactly at time $t = T$, the right to buy X US dollars at the price K $SEK/\$$
- The holder of the option has no obligation to buy the dollars

This contract is called an option, because it gives the holder of the contract the option, but not the obligation of buying some underlying asset. The prefix European, means that the option can only be exercised at maturity, the exercise date T. There are other types of options that can be exercised at times prior to maturity, T. Option is a derivative asset, in the sense that it is defined in terms of the underlying asset. An option always has a non-negative price when the contract is entered. The price is determined on the existing option market. The value of the option (at T), depends on the future level of the exchange rate, that is, it is stochastic.

2.2 Geometric Brownian motion

Let us consider the population of a country, assume that it grows with a constant rate r, where the unit of time, t, is 1 year. So we have a deterministic model, starting at X_0

$$X_0(1 + r) = X_1 \tag{2.1}$$

and to get the population after the end of t years

$$X_0(1+r)^t = X_t \tag{2.2}$$

Here, x, t are discrete variables, but we will turn them into continuous variables.

Let us consider the number of bacteria, a more dense number, thus making the step to continuous variables more easy. In the previous example we had one compounding per t, (per year), let us here assume that we have f periods, (compounding) per year (unit time), let the rate r be the same as in the previous example. In each compounding period, the numbers of bacteria, grows with $\frac{r}{f}$, for t years, there will be ft periods. We get the following

$$X_t = X_0 \left(1 + \frac{r}{f}\right)^{ft} \tag{2.3}$$

and let the number of compounding period grow, per unit time, that was (year), we get

$$X_t = \lim_{f \to \infty} X_0 \left(1 + \frac{r}{f}\right)^{ft} \tag{2.4}$$

and the above equation, (2.4) is one the definitions of e, using this fact and law of composition of limit we end up with:

$$X_t = X_0 \left(e^r\right)^t = X_0 e^{rt} \tag{2.5}$$

We will interpret the variables as continuous, going back to the population (2.2), and to see what happens to the population in small time step h, we get

$$X_{t+h} = X_0 e^{r(t+h)} \tag{2.6}$$

so the change in the population, over an infinitesimal change in time

$$\lim_{h \to 0} \frac{X_{t+h} - X_t}{h} = \lim_{h \to 0} \frac{X_0 e^{r(t+h)} - X_0 e^{rt}}{h} \tag{2.7}$$

This is nothing than the definition of derivative

$$\frac{dX_t}{dt} = \lim_{h \to 0} \frac{e^{rh} - 1}{h} X_0 e^{rt} \tag{2.8}$$

The Taylor expansion of Euler's constant e, close to zero is

$$e^x = 1 + x + \frac{x^2}{2!} + \dots \tag{2.9}$$

applying (2.9), the Taylor series expansion to (2.8) we get

$$\frac{dX_t}{dt} = \lim_{h \to 0} \frac{\left(1 + rh + \frac{(rh)^2}{2!}\right) - 1}{h} X_0 e^{rt} \tag{2.10}$$

dividing through by h, and canceling the 1 in the denominator we get

$$\frac{dX_t}{dt} = \lim_{h\to 0}\left(r + \frac{(r^2 h)}{2!} + \dots\right) X_0 e^{rt} = rX_0 e^{rt} = rX_t \tag{2.11}$$

This differential equation could also have been solved using the integrating factor, or the separation of variables. This differential equation can be used to model many objects,

$$\frac{dX_t}{dt} = rX_t \tag{2.12}$$

for example the growth of the bank account. Here we assume that r in (2.12) is constant, I will in later chapters, look at the example where r_t is deterministic function of time.

$$dX_t = X_t(r_t\,dt + dR_t) \tag{2.13}$$

When you deposit the money at the bank, you know the interest rate, at the beginning of each period, (risk-free rate), the rate will vary over time, this is a case for the time dependent deterministic interest rate, r_t. Assume that you invest money in a stock, by buying share of the company, than you would expect to earn a return, but this return will be random, since it depends on the future price of the stock, that gives the additional dR_t the random component. This random component will be a stochastic process. There are many random process, but in the Geometric Brownian Motion we assume the following model

$$dX_t = X_t(r_t\,dt + dW_t) \tag{2.14}$$

The increment of a standard Brownian Motion over an interval are normally distributed, with mean equal to 0, and variance equal to the length of interval, t. If you scale the Brownian motion by multiplying it with σ, the volatility.

$$dX_t = X_t(r_t\,dt + \sigma_t dW_t) \tag{2.15}$$

This (2.15) allows you to build and model many as stochastic process, the deterministic r_t allows you to control the mean of the process, and remember the expectation of the integral of a stochastic process with respect to a standard Brownian Motion, is zero, for a fixed time horizon. In the simple model, we assume that r, σ are constant, the process can be written as

$$dX_t = rX_t\,dt + \sigma X_t dW_t \tag{2.16}$$

For the deterministic model, $dX_t = rX_t\,dt$, we have

$$\mathbb{E}[dX_t] = rX_t\,dt \quad Var[dX_t] = 0 \tag{2.17}$$

for the stochastic model $dX_t = rX_t\,dt + \sigma X_t\,dW_t$, we can write it in integral form, and noting that expectation of a stochastic integral is zero.

$$\begin{aligned} \mathbb{E}X_t &= \mathbb{E}X_0 + \int_0^t r\mathbb{E}X_s\,ds \\ &= \mathbb{E}X_T = X_0 e^{rT} \end{aligned} \tag{2.18}$$

and for the second moment we get

$$\mathbb{E}\left[X_t\right]^2 = \mathbb{E}X_0^2 + \int_0^t (2r+\sigma^2)\mathbb{E}X_s^2\, ds$$
$$\mathbb{E}\left[X_t\right]^2 = X_0^2 e^{(2r+\sigma^2)t} \tag{2.19}$$

and the variance becomes

$$Var\left[X_t\right] = X_0^2 e^{2rt}\left(e^{\sigma^2 t} - 1\right) \tag{2.20}$$

In the process with a random process, the infinitesimal change is not deterministic anymore, thus they have a probability distribution and variance refers to the infinitesimal changes, and not the process itself. This variance is the probability weighted average of the displacement squared, but the quadratic variation of the process, this is like the limiting sum of the squares of displacement, when the interval is divided into large number of sub intervals, is also equal to, for the standard Brownian increments, it follows that

$$dW_t^2 = dt \tag{2.21}$$

where dt is supposed to be very small, thus $dt^2,\ dt \cdot dW_t \approx 0$.

2.2.1 Solve the SDE

When trying to solve a differential equation, one usually try to solve an easier variant of the differential equation, and then hope that the this solution also will be possible for the more general form of differential equation. In our example, we start by trying to solve (2.12), so we have

$$dx = rxdt \tag{2.22}$$

collect the x:s and the t on each side of the equation (this technique is called the separation of variables) , and in the next step, we integrate and take the exponent on each side.

$$\frac{dx}{x} = rdt \Leftrightarrow d\log(x) = rdt \Leftrightarrow x = x_0 e^{rt} \tag{2.23}$$

This gives us a hint that $\log(x)$ also should solve the stochastic differential equation, but in the stochastic world, the chain rule, Ito's lemma, is different from the deterministic world.

$$d\log X_t = \frac{\partial \log X_t}{\partial X_t} dX_t + \frac{1}{2}\frac{\partial^2 \log X_t}{\partial X_t^2} dX_t^2 \tag{2.24}$$

putting the first and second derivative of $\log X$ it becomes

$$d\log X_t = \frac{1}{X_t} dX_t - \frac{1}{2}\frac{1}{X_t^2} dX_t^2 \tag{2.25}$$

insert dX_t and dX_t^2, and remember box-algebra

$$d\log X_t = \frac{1}{X_t}\left(rX_t + \sigma X_t dW_t\right) - \frac{1}{2}\frac{1}{X_t^2}\sigma^2 X_t^2 dt \tag{2.26}$$

combine the dt terms, and note that the X_t cancels.

$$d\log X_t = \left(r - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t \tag{2.27}$$

integrate and assume that we are interested in the solution a time T

$$\int_0^T d\log X_t = \int_0^T \left(r - \frac{1}{2}\sigma^2\right)dt + \int_0^T \sigma dW_t \tag{2.28}$$

on the left hand side, the differential and the integral cancels, and taking the constant out of the integrals on the right hand side gives us.

$$\log X_T - \log X_0 = \left(r - \frac{1}{2}\sigma^2\right)\int_0^T dt + \sigma\int_0^T dW_t \tag{2.29}$$

this can be written as

$$\log\left(\frac{X_T}{X_0}\right) = \left(r - \frac{1}{2}\sigma^2\right)T + \sigma W_T \tag{2.30}$$

take the exponent on both side, and noting that $W_0 = 0$, it is in the definition of Brownian motion.

$$X_T = X_0 e^{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma W_T} \tag{2.31}$$

it can also be written as

$$X_T = X_0 e^{\log X_t + \left(r - \frac{1}{2}\sigma^2\right)T + \sigma W_T} \tag{2.32}$$

This makes the probabilistic interpretation easier, as X_T is the exponential of normal random variable. If $Y \sim N(\mu, \sigma^2)$ then the exponential of $X = e^Y \sim LN(\mu, \sigma^2)$ in (2.32), the term in the exponent is normal, and hence X_T is log-normal distributed. The exponent is normally distributed with mean equal to the deterministic part and variance equal the square of the coefficients of the Brownian times the length of the interval

$$Y_T = \log X_t + \left(r - \frac{1}{2}\sigma^2\right)T + \sigma W_T \sim N\left[\log X_0 + (r - \frac{1}{2}\sigma^2)T, \sigma^2 T\right] \tag{2.33}$$

and in our case

$$X_T \sim LN\left[\log X_0 + (r - \frac{1}{2}\sigma^2)T, \sigma^2 T\right] \tag{2.34}$$

We can now generate a sample from the log-normal distribution. We know that $W_T \sim N[0, T]$, it can be represented as standard normal distribution, Z,

$Z \sim N[0,1]$ as $W_T = \sqrt{T}Z$, so we can write (2.32) as, for an arbitrary starting point t

$$X_T = X_t e^{\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma\sqrt{T-t}Z} \tag{2.35}$$

and let Δt be the size of the time step

$$X_{t+\Delta t} = X_t e^{\left(r-\frac{1}{2}\sigma^2\right)\Delta t+\sigma\sqrt{\Delta t}Z} \sim LN\left[\log X_t + \left(r - \frac{1}{2}\sigma^2\right)\Delta t, \sigma^2\Delta t\right] \tag{2.36}$$

2.3 Black Scholes' dynamics

I will derive the dynamics for the Black Scholes stock dynamics, under the risk neutral and the stock measure. In the Black Scholes model the stocks are given by the geometric Brownian motion, which has the following dynamics, where the drift (dt) and the diffusion (dW_t) terms are proportional to the value of the process.

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2.37}$$

and the bank account is assumed to be deterministic. Note that the bank account only has a drift and no diffusion term. The bank account grows at an continuous and constant rate r

$$dB_t = rB_t dt \tag{2.38}$$

This equation can be solved by the separation of variables.

$$B_t = e^{rt} \tag{2.39}$$

We have the stock dynamics under the physical, or the real probability measure $\mathbb{P}$, and we want to know how these dynamics will look like under the risk-neutral measure $\mathbb{Q}$, which is the measure associated by the bank account numeraire and under the measure induced by the stock measure $\mathbb{S}$. The numeraire is an asset which acts as an measure of value, for example money or gold or tulips. There are three key concepts needed to derive the dynamics. The general valuation formula, (RNVF), the theory of Martingales and the Girsanov theorem.

2.3.1 Concepts needed for the derivation of B-S

The risk neutral valuation formula (RNVF) states that the value of an asset expressed in the units of another asset

$$\frac{V_t}{X_t} = \mathbb{E}^{\mathbb{X}}\left[\frac{V_T}{X_T} \,\middle|\, \mathcal{F}_t\right] \tag{2.40}$$

is a martingale under some probability measure $\mathbb{X}$. The value of an asset expressed with the bank account as numeraire

$$\frac{V_t}{B_t} = \mathbb{E}^{\mathbb{Q}}\left[\frac{V_T}{B_T} \,\middle|\, \mathcal{F}_t\right] \tag{2.41}$$

under the risk neutral measure $\mathbb{Q}$. The value of the unit under the stock measure $\mathbb{S}$ will also be a martingale.

$$\frac{V_t}{S_t} = \mathbb{E}^{\mathbb{S}}\left[\frac{V_T}{S_T} \,\Big|\, \mathcal{F}_t\right] \tag{2.42}$$

To derive the dynamics of the stock under the risk neutral measure, we write it in terms of the price of the stock expressed in the units of the bank account. The stock price scaled by the value of the bank account, will be a martingale under the measure $\mathbb{Q}$

$$\frac{V_t}{B_t} = \mathbb{E}^{\mathbb{Q}}\left[\frac{V_T}{B_T} \,\Big|\, \mathcal{F}_t\right] \tag{2.43}$$

Let the ratio between the stock price and the bank account be

$$Z_t = \frac{S_t}{B_t} \tag{2.44}$$

then it follows from the theory of Martingales that the SDE of Z_t, i.e. the dynamics of the stock asset will have zero drift under the measure induced by the bank account ($\mathbb{Q}$).

$$dZ_t = \sigma Z_t dW_t^{\mathbb{Q}} \tag{2.45}$$

where $W^{\mathbb{Q}}$ is the standard Brownian motion under the $\mathbb{Q}$ measure. Use Ito's lemma for ratio Z_t

$$\begin{aligned}
d\left(\frac{S_t}{B_t}\right) &= \frac{dS_t}{B_t} + S_t d\left(\frac{1}{B_t}\right) \\
&= \frac{dS_t}{B_t} + S_t\left[d(e^{-rt})\right] \\
&= \frac{dS_t}{B_t} + S_t(-re^{-rt}) \\
&= \frac{dS_t}{B_t} + S_t\left(-r\frac{1}{B_t}dt\right) \\
&= \frac{dS_t}{B_t} - r\frac{S_t}{B_t}dt \\
&= \frac{\mu S_t dt + \sigma S_t dW_t}{B_t} - r\frac{S_t}{B_t}dt \\
&= \sigma\frac{S_t}{B_t}\left(\frac{\mu - r}{\sigma}dt + dW_t\right) \\
dZ_t &= \sigma Z_t\left(\frac{\mu - r}{\sigma}dt + dW_t\right)
\end{aligned} \tag{2.46}$$

So we have two equation that describe the dynamics of Z_t and they therefore must be equal

$$dZ_t = \sigma Z_t dW_t^{\mathbb{Q}} \tag{2.47}$$

$$dZ_t = \sigma Z_t \left(\frac{\mu - r}{\sigma} dt + dW_t \right) \tag{2.48}$$

The Brownian motion under the original measure $\mathbb{P}$ and the Brownian motion under the risk neutral measure are linked as follows

$$dW_t^{\mathbb{Q}} = \frac{\mu - r}{\sigma} dt + dW_t \tag{2.49}$$

Substitute the new Brownian, under $\mathbb{Q}$ to get the orginal Brownian under the physical measure $\mathbb{P}$ to get

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t \left(dW_t^{\mathbb{Q}} - \frac{\mu - r}{\sigma} dt \right) = \\ &= r S_t dt + \sigma S_t dW_t^{\mathbb{Q}} \end{aligned} \tag{2.50}$$

The dynamics under the stock measure, here I will use the Girsanov theorem. We know that the value of an asset under the bank account as the numeraire is a martingale

$$\frac{V_0}{B_0} = \mathbb{E}^{\mathbb{Q}} \left[\frac{V_t}{B_t} \,\middle|\, \mathcal{F}_t \right] \tag{2.51}$$

and the value of an asset under the stock numeraire will be a martingale under the numeraire induced by the denominator

$$\frac{V_0}{S_0} = \mathbb{E}^{\mathbb{S}} \left[\frac{V_t}{S_t} \,\middle|\, \mathcal{F}_t \right] \tag{2.52}$$

B_0 and S_0 are known at the filtration by $\mathcal{F}_0$ so we can put them inside our expectation.

$$V_0 = \mathbb{E}^{\mathbb{Q}} \left[\frac{B_0}{B_t} V_t \,\middle|\, \mathcal{F}_t \right] \tag{2.53}$$

and for the stock as numeraire.

$$V_0 = \mathbb{E}^{\mathbb{S}} \left[\frac{S_0}{S_t} V_t \,\middle|\, \mathcal{F}_t \right] \tag{2.54}$$

as both expression represents the price of the same asset, and as it holds for any asset, it means that the terms inside the expectation must be equal.

$$\begin{aligned} &\left[\frac{B_0}{B_t} \right] d\mathbb{Q} = \left[\frac{S_0}{S_t} \right] d\mathbb{P}^{\mathbb{S}} \Leftrightarrow \\ &\left[\frac{d\mathbb{P}^{\mathbb{S}}}{d\mathbb{Q}} \right] = \frac{B_0}{B_t} \frac{S_t}{S_0} \end{aligned} \tag{2.55}$$

The solution for the dynamics for the stock price SDE is, under the risk neutral measure $\mathbb{Q}$

$$S_t = S_0 \exp \left\{ rt - \frac{\sigma^2}{2} t + \sigma W_t^{\mathbb{Q}} \right\} \tag{2.56}$$

put into the equation and also for the ratio of the bank account and it gives us

$$\left[\frac{d\mathbb{P}^{\mathbb{S}}}{d\mathbb{Q}}\right] = e^{-rt}e^{rt-\frac{1}{2}\sigma^2 t+\sigma W_t^{\mathbb{Q}}} \leftrightarrow e^{-\frac{1}{2}\sigma^2 t+\sigma W_t^{\mathbb{Q}}} \tag{2.57}$$

The Girsanov theorem states that if $W_t^{\mathbb{Q}}$ is a Brownian motion under the measure $\mathbb{Q}$ and if we shift the process by $Y(t) = \int_0^t y_u \, du$ than the shifted process $W_t^{\mathbb{S}} = W_t^{\mathbb{Q}} - \int_0^t y_u \, du$ will be a Brownian motion under the measure $P^{\mathbb{S}}$ that can be identified of its density

$$\left[\frac{d\mathbb{P}^{\mathbb{S}}}{d\mathbb{Q}}\right] = \exp\left\{-\frac{1}{2}\int_0^t y_u^2 \, du + \int_0^t y_u dW_u^{\mathbb{Q}}\right\} \tag{2.58}$$

The relationship between the two measure would in differential form be

$$dW_t^{\mathbb{S}} = dW_t^{\mathbb{Q}} - y_t \, dt \tag{2.59}$$

In our example $y_t = \sigma$, i.e. a constant, and the relationship between the two Brownian Motion would be

$$dW_t^{\mathbb{S}} = dW_t^{\mathbb{Q}} - \sigma \, dt \tag{2.60}$$

so the dynamics under the stock measure as numeraire

$$\begin{aligned} dS_t =& rS_t \, dt + \sigma S_t dW_t^{\mathbb{S}} = \\ =& rS_t \, dt + \sigma S_t + \left(dW_t^{\mathbb{S}} + \sigma dt\right) \\ dS_t =& \left(r + \sigma^2\right) S_t \, dt + \sigma S_t + dW_t^{\mathbb{S}} \end{aligned} \tag{2.61}$$

2.4 Black Scholes price for an option

In the Black Scholes model the dynamics of the underlying asset is

$$dS_t = rS_t \, dt + \sigma S_t dW_t \tag{2.62}$$

has the solution

$$S_t = S_0 \exp\left\{\sigma W_t + rt - \frac{\sigma^2 t}{2}\right\} \tag{2.63}$$

It can be verified by applying Ito's lemma to $y = \log(S_t)$ and integrating the resulting equation over $[0, t]$. Let us look what distribution S_t has, notice that $y = \log S_t$ is normally distributed, and that $S_t \geq 0$, we can calculate the expectation and variance for y

$$\mathbb{E}[y] = \mathbb{E}\left[\log(S_0) + \sigma W_t + rt - \frac{\sigma^2 t}{2}\right] \Leftrightarrow \log(S_0) + rt - \frac{\sigma^2 t}{2} \tag{2.64}$$

$$\text{Var}[y] = \sigma^2 \int_0^t du = \sigma^2 t \tag{2.65}$$

and the distribution for S_t can be found by variable transformation.

$$f_{S_t}(S) = \frac{1}{\sqrt{2\pi \mathrm{Var}[\log S]}} \exp\left\{-\frac{1}{2}\left(\frac{\log S - \mathbb{E}[\log S]}{\sqrt{\mathrm{Var}[\log S]}}\right)^2\right\} \frac{1}{S} \leftrightarrow \\ \frac{1}{S\sqrt{2\pi\sigma^2 t}} \exp\left\{-\frac{1}{2}\left(\frac{\log S - \log S_0 - rt + \frac{\sigma^2 t}{2}}{\sqrt{\sigma^2 t}}\right)^2\right\} \tag{2.66}$$

The price for a call option, it pays at maturity of the option the difference between the price of the underlying and the strike if the difference is positive.

$$\text{Payoff} \quad (S_T - K)^+ \tag{2.67}$$

If you want to know the price at time 0 that is before maturity

$$\text{Payoff}_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+\right] = \\ e^{-rT}\left(\mathbb{E}^{\mathbb{Q}}\left[S_T \mathbb{1}_{S_T > K}\right] - \mathbb{E}^{\mathbb{Q}}\left[K \mathbb{1}_{S_T > K}\right]\right) \tag{2.68}$$

We can simplify the two expression, i.e. solve analytically, and we end up with

$$\text{Payoff}_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+\right] = \\ e^{-rT}\left(S_0 e^{rT} N\left[\frac{\log\frac{S_0}{K} + rT + \frac{\sigma^2 T}{2}}{\sigma\sqrt{T}}\right] - KN\left[\frac{\log\frac{S_0}{K} + rT - \frac{\sigma^2 T}{2}}{\sigma\sqrt{T}}\right]\right) \\ S_0 N[d_1] - Ke^{rT} N[d_2] \tag{2.69}$$

and the

$$d(1) = \frac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T - t)}{\sigma\sqrt{T - t}} \\ d(2) = \frac{\log\left(\frac{S_t}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)(T - t)}{\sigma\sqrt{T - t}} \tag{2.70}$$

2.4.1 Derivation of Black Scholes PDE

I am using the Delta hedging argument to derive the Black - Scholes PDE. We will get the answer

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{2.71}$$

Let the price of the option be V. V is a function that depends on $V = (T - t, S_t; r, \sigma, K)$ and assume that r, σ, K, the risk - free rate, the volatility and the strike price are all constant, than the price of the option will depend on

$V = (T - t, S)$, the time to maturity and the price of the underlying. We next use Ito's lemma on the differential of V[1]

$$dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} dS^2 \tag{2.72}$$

insert dS from (2.37) and dB from (2.38) and using box algebra, and collecting the dt terms we end up with

$$dV = \left[\frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right] dt + \sigma S \frac{\partial V}{\partial S} dW_t \tag{2.73}$$

Equation (2.73) looks as a SDE, with a drift term and a diffusion term, the later is driven by a geometric Brownian motion. Use the Delta argument to eliminate the stochastic component, that is done by trading in the underlying, i.e. the stock. Hedging the risk of the option by trading in the underlying stock. As both are driven by the same Brownian, we almost eliminate our exposure to the gBm in the option price by trading the underlying. We need to know how many units of the underlying to buy or sell, when to buy and sell, or for how long should we keep the position hedged. Assume that we are hedging a short position in a call option, our strategy will involve buying stocks, buying stocks will require funding possibility. Assume that we have unlimited access to a bank account, and we need to pay the bank interest rate when we borrow, and we must be able to repay our debt. Reversely the bank bank will pay us interest rate, when we have excess cash. Assume that we bought Δ units of the stock and borrowed α units, of the currency. Δ, α can be negative or positive. That gives us

$$\Pi = \Delta S + \alpha B \tag{2.74}$$

and the differential of the portfolio

$$d\Pi = \Delta dS + \alpha dB \tag{2.75}$$

insert from (2.37) and (2.38)

$$d\Pi = \Delta \left(\mu S_t \, dt + \sigma S_t dW_t \right) + \alpha r B dt \tag{2.76}$$

combining the dt terms and multiplying in Δ gives

$$d\Pi = (\Delta \mu S_t dt + \alpha r B) + \Delta \sigma S_t dW_t \tag{2.77}$$

This is a SDE for the portfolio, since the Brownian motion is the same in both (2.73) and (2.77) we can set the stochastic terms equal.

$$\sigma S \frac{\partial V}{\partial S} = -\Delta \sigma S_t \tag{2.78}$$

[1] the reason for not adding $\frac{\partial^2 V}{\partial t^2} dt^2$ is box - algebra, and the reason for not adding the cross term is the same, so we need only the second partial to the asset price

isolating the $\frac{\partial V}{\partial S}$, and canceling terms, we find that

$$\Delta = -\frac{\partial V}{\partial S} \tag{2.79}$$

i.e. the derivative of the option price w.r.t. the stock price. Then the combined portfolio is

$$dV + d\Pi = \left[\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} - \mu S\frac{\partial V}{\partial S} + \alpha r B\right] dt \tag{2.80}$$

we only have the drift terms left, and some cancellation gives us.

$$d\left(V + \Pi\right) = \left[\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \alpha r B\right] dt \tag{2.81}$$

The total portfolio, has only a deterministic component and must, to avoid arbitrage grow at the risk free rate

$$d\left(V + \Pi\right) = \left(V + \Pi\right) r dt \tag{2.82}$$

use that

$$\Pi = -\frac{\partial V}{\partial S} S + \alpha B \tag{2.83}$$

substitute for Π in (2.82) we get

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \alpha r B = rV - r\frac{\partial V}{\partial S} S + \alpha R B \tag{2.84}$$

and cancellation gives us

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{2.85}$$

If you shift the last terms to the right hand side, you will see that

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} = r\left(V - S\frac{\partial V}{\partial S}\right) \tag{2.86}$$

then the r.h.s is the return of the bank account, that is equal to the option premium minus the amount that we borrowed to finance the delta units of the stocks, is equal to the balance times interest rate, in an infinitesimal time period. The left hand side represents how the Delta hedged option changes in an infinitesimal time. The first term captures the shortening of the maturity, the second term, the gamma-impact, the risk that remains after the Delta is hedged. This is almost the backward diffusion equation. You can also solve this problem by introducing a replicated portfolio.

2.5 Bibliographical notes

A good starting point is (Björk, 2009). There are many books in this area, a bit more applied is (Lindström, Madsen, & Nielsen, 2015), a more hands on with many coding exercises are (Iacus, 2011), a nice introduction to Brownian motion, stochastic integral and the existence of it is (Evans, 2012). There are also many lecture notes in the internet, I have used, http://www.frouah.com/pages/finmath.html for some clarification, and from Rolf Poulsen notes Copenhagen University http://web.math.ku.dk/ rolf/teaching/ctff03/ . A good place with derivations and videos is https://quantpie.co.uk/

Chapter 3

Diffusion Equation

3.1 Diffusion

This chapter will deal with the diffusion equation. I have included it because, it gives a better understanding of variance, (volatility). It will help to get a better picture of Dupire's local volatility model, and in the Heston model, where the variance is assumed also to be a volatile process, not only as in the Black - Scholes model, the underlying stock price process. This derivation of the diffusion follows Einstein's solution, using probability to the problem. Assume that we have suspended particles in a liquid, we will take the 1-dimensional view of Einstein solution to the diffusion equation. See figure (3.1), at time t, there will be $f(x,t) \cdot dx$ particles in the left rectangle. I make an area, dx around the generic x As time moves on to $t + \tau t$, from the upper to the lower subplot in figure (3.1) there will be another number of particles in the same area, keeping the generic x on the x-axis fixed. Assume that τ the time step is small, but big enough to assume that the two figures in figure (3.1) are independent. Look at the distance a particle in the upper figure need to move, from the right rectangle to the left rectangle, during a time interval τ, the length of the movement, or displacement has a probabilistic interpretation, let Δ is a random variable. Einstein assumed that most particles will have a small displacement, and the probability of a large movement is small. Notice that I have chosen a displacement to the left. We assume that there is no influx of particles. The number of particles in the rectangle Δ away from the original, at time t, will be a rectangles $f(x+\Delta,t) \cdot dx$. Let $\phi(\Delta)$ be the probability that a particle has a displacement equal to Δ, the number of particles in the new rectangle, after a time step equal to τ that will move to the original rectangle will be $dxf(x+\Delta,t)\phi(\Delta)$[1]. This will hold for every rectangle to the right from the original rectangle, as Δ is the distance, the movement from the left would be $dxf(x-\Delta,t)\phi(-\Delta)$. I assume that ϕ is symmetric around the generic x i.e. $\phi(\Delta) = \phi(-\Delta)$. I further assume that the possibility that a particle will make

[1] This is nothing but the expected number, $\mathbb{E}(X) = \int f(x+\Delta,t)\phi(\Delta)\,dx$

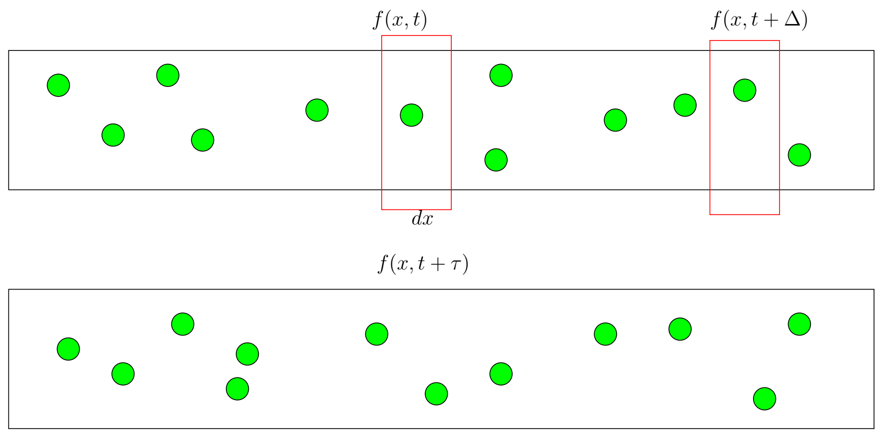

Figure 3.1: Particles in a fluid. The upper is at time t and the lower is at time $t+\tau$

two movements in a small time interval, τ, is zero. Integrate across the x-axis, we get the number in particles in x at a later time $t+\tau$. There is no influx of new particles.

$$\begin{aligned} f(x,t+\tau)\,dx &= dx\int_{\Delta=-\infty}^{\infty} f(x+\Delta,t)\phi(\Delta)\,d\Delta \\ f(x,t+\tau) &= \int_{\Delta=-\infty}^{\infty} f(x+\Delta,t)\phi(\Delta)\,d\Delta \end{aligned} \tag{3.1}$$

Expend the left hand side in equation (3.1), using a Taylor expansion

$$f(x,t+\tau) = f(x,t) + \frac{\partial f}{\partial t}\tau \tag{3.2}$$

and the right hand side in equation (3.1), using a Taylor expansion

$$f(x+\Delta,t) = f(x,t) + \frac{\partial f}{\partial x}\Delta + \frac{1}{2!}\frac{\partial^2 f}{\partial x^2}\Delta^2 \tag{3.3}$$

substitute these in equation (3.1) we get

$$f(x,t) + \frac{\partial f}{\partial t}\tau = \int_{\Delta=-\infty}^{\infty}\left(f(x,t) + \frac{\partial f}{\partial x}\Delta + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\Delta^2\right)\phi(\Delta)\,d\Delta \tag{3.4}$$

expanding the right hand side in equation (3.4) gives us

$$
\begin{aligned}
&\int_{\Delta=-\infty}^{\infty} \left(f(x,t) + \frac{\partial f}{\partial x}\Delta + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\Delta^2 \right) \phi(\Delta)\, d\Delta \\
&= \int_{\Delta=-\infty}^{\infty} f(x,t)\phi(\Delta)\, d\Delta \\
&+ \int_{\Delta=-\infty}^{\infty} \frac{\partial f}{\partial x}\Delta\phi(\Delta)\, d\Delta \\
&+ \int_{\Delta=-\infty}^{\infty} \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\Delta^2\phi(\Delta)\, d\Delta
\end{aligned}
\tag{3.5}
$$

The total probability is equal to one, so the first integral after the equal sign is equal to $f(x,t)$. The number of particles in a rectangle $f(x,t)$ does not depend on Δ, the displacement, and Δ is symmetric around zero, makes the second integral to zero. So we end up with

$$
\begin{aligned}
\frac{\partial f}{\partial t}\tau &= \int_{\Delta=-\infty}^{\infty} \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\Delta^2\phi(\Delta)\, d\Delta \\
\frac{\partial f}{\partial t} &= \frac{1}{2\tau}\frac{\partial^2 f}{\partial x^2}\int_{\Delta=-\infty}^{\infty} \Delta^2\phi(\Delta)\, d\Delta \\
\frac{\partial f}{\partial t} &= D\frac{\partial^2 f}{\partial x^2} \text{ where } \quad D = \frac{1}{2\tau}\int_{\Delta=-\infty}^{\infty} \Delta^2\phi(\Delta)\, d\Delta
\end{aligned}
\tag{3.6}
$$

Let D be the diffusion coefficient. It is the average of the displacement square. The larger the D the faster the particles will move.

3.1.1 Solution to the diffusion equation

You can solve the diffusion equation in many ways, I choose to use the fundamental solution method, also known as the solution with the diffusion kernel. We have the the initial condition of a point at a known position. The heat equation, also know as the diffusion equation is

$$
\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} \tag{3.7}
$$

and we want to find $f(x,t)$ that would solve equation (3.7) You can use the similarity principle. We have a PDE, the key is to find an invariant transformation, of the variables x, t, to a set with less variables, that also solves the heat equation. We then proceed to solve the simpler equation, and hope it is easier. It turns out that for the diffusion equation the following transformation reduces the number of parameters (x,t) to $v = \lambda x$, $u = \lambda^2 t$ for a new set of variables (u,v), so $f(v,u) = f(\lambda x, \lambda^2 t)$. The diffusion equation under the new transformed variables becomes

$$\begin{aligned}
\frac{\partial f(v,u)}{\partial t} &= \frac{\partial f(v,u)}{\partial u}\frac{\partial u}{\partial t} = \lambda^2 \frac{\partial f(v,u)}{\partial u} \\
\frac{\partial f(v,u)}{\partial x} &= \frac{\partial f(v,u)}{\partial v}\frac{\partial v}{\partial x} = \lambda \frac{\partial f(v,u)}{\partial v} \\
\frac{\partial^2 f(v,u)}{\partial x^2} &= \lambda \frac{\partial}{\partial x}\left(\frac{\partial f(v,u)}{\partial v}\right) = \lambda^2 \frac{\partial^2 f(v,u)}{\partial v^2}
\end{aligned} \tag{3.8}$$

The invariant transformation satisfies the diffusion equation, notice that $\frac{1}{\lambda^2}$ cancels

$$\frac{\partial f(\lambda x, \lambda^2 t)}{\partial t} = D\frac{\partial^2 f(\lambda x, \lambda^2 t)}{\partial x^2} \tag{3.9}$$

The question is now, how to find the function $f(x,t)$. I will not go through the steps, a brief outline if you set $\lambda = \frac{1}{\sqrt{t}}$ the then the transformed heat equation becomes $f(x,t) = \lambda f(\lambda x, \lambda^2)$ becomes $\frac{1}{\sqrt{t}} f\left(\frac{x}{\sqrt{t}}, 1\right)$ we also want the result to be dimension less, remember that x is the distance in the horizontal axis and t is time

$$\frac{1}{\sqrt{t}} = \frac{\text{length}}{\sqrt{\text{time}}} \tag{3.10}$$

putting this into the diffusion equation, and let f be the number of particles.

$$\begin{aligned}
\frac{\partial f}{\partial t} &= D\frac{\partial^2 f}{\partial x^2} \\
\frac{\text{particle}}{\text{time}} &= D\frac{\text{Particle}}{\text{area}}
\end{aligned} \tag{3.11}$$

it makes D to have the dimension $\frac{\text{Area}}{\text{time}}$, that makes the square root of D to have the the same dimension $\sqrt{D} = \frac{\text{length}}{\sqrt{\text{time}}}$. So it is dimension-less.

$$f(x,t) = \frac{m}{\sqrt{4\pi D t}} \exp^{-\frac{1}{4}\frac{x^2}{Dt}}\ dz \tag{3.12}$$

at time 0, the number of particles will be at location 0 $f(0,0) = \psi(0)$, every particle will be concentrated in its initial value, so for a generic point x on the x-axis, we have $f(x,0) = \psi(x)$. If we want to see how the particle spread over the x-axis and time

$$f(x,t) = \int \psi(z) \frac{1}{\sqrt{4\pi D t}} \exp^{-\frac{1}{4}\frac{(x-z)^2}{Dt}}\ dz \tag{3.13}$$

where $f(x,0) = \psi(x)$ the initial distribution, or the initial number of particles, is also called the impulse function. The exponent in equation (3.13) is Green's function, and is the response to the impulse.

3.1.2 Diffusion equation with drift

We have the diffusion equation

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} \tag{3.14}$$

and the diffusion equation with drift is

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} - \mu\frac{\partial f}{\partial x} \tag{3.15}$$

The particles will move as in the standard diffusion equation, but now they will have a force acting upon them, think gravity or current, that make them move in a preferred direction. In the diffusion equation, we were only interested in the size of the displacement, as it was assumed to be symmetric around the current value x, that made the second integral in (3.5) to become zero, but this will no longer be the case. A particle will move from the right, from $x+\Delta$ if it experience a displacement of $(-\Delta)$, that will be $\phi(-\Delta)$ and movement from the left of x will then be characterized as $x-\Delta$ as $\phi(\Delta)$. So the number of particles in x an instant (τ) later, is thus

$$\begin{aligned}
f(x,t+\tau)dx &= dx\int_{\Delta=\infty}^{\infty} f(x+\Delta,t)\phi(-\Delta)\,d\Delta \\
f(x,t+\tau) &= \int_{\Delta=\infty}^{\infty} f(x+\Delta,t)\phi(-\Delta)\,d\Delta \\
f(x,t)+\frac{\partial f}{\partial t}\tau &= \int_{\Delta=-\infty}^{\infty}\left(f(x,t)+\frac{\partial f}{\partial x}\Delta+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}\Delta^2\right)\phi(-\Delta)\,d\Delta \\
&= f(x,t)+\frac{\partial f}{\partial x}\int_{-\infty}^{\infty}\Delta\phi(-\Delta)\,d\Delta+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}\int_{\infty}^{\infty}\Delta^2\phi(-\Delta),d\Delta \\
\frac{\partial f}{\partial t} &= \frac{\partial f}{\partial x}\frac{1}{\tau}\int_{-\infty}^{\infty}\Delta\phi(-\Delta)\,d\Delta+\frac{1}{2\tau}\frac{\partial^2 f}{\partial x^2}\int_{\infty}^{\infty}\Delta^2\phi(-\Delta),d\Delta \\
\frac{\partial f}{\partial t} &= -\frac{\partial f}{\partial x}\frac{1}{\tau}\int_{-\infty}^{\infty}\Delta\phi(\Delta)\,d\Delta+\frac{1}{2\tau}\frac{\partial^2 f}{\partial x^2}\int_{\infty}^{\infty}\Delta^2\phi(\Delta),d\Delta
\end{aligned} \tag{3.16}$$

We made a Taylor expansion in the last line, just as in equation (3.5), expanding the parenthesis on the right hand side, noting that the first integral will $f(x,t)$, due to the fact that the total probability is equal to one, the second, will however not cancel, as in equation (3.5) as it has a drift. I also change Δ to $-\Delta$, remember that $(-1)^2 = 1$ in the second partial derivative. So the diffusion coefficient becomes

$$D = \frac{1}{2\tau}\int_{-\infty}^{\infty}\Delta^2\phi(\Delta)\,d\Delta, \quad \mu = \frac{1}{\tau}\int_{-\infty}^{\infty}\Delta\phi(\Delta),d\Delta \tag{3.17}$$

The left hand side in equation (3.15) tells us how the particle change under a small interval, keeping the position constant. The D tells us, just as the heat equation, that if you are in a location where the number of particles are lower than the surrounding, remember that we are in a 1-dimensional setting, more particles will enter that point, and if the number of particles are higher than the surrounding, more particles will move out. The drift component tells us that if the location is on the left, more particle will move into the area of x.

3.1.3 Solution to the Diffusion equation with drift

The diffusion equation with drift is equation (3.15). I re-state it here

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} - \mu\frac{\partial f}{\partial x} \tag{3.18}$$

You can view the drift term as the current in a stream, and the diffusion term as the spread, displacement of the particles. It will have different speeds in different fluids. Under an interval of time t the a particle dropped in the water will travel a distance μt. As we have seen before, the diffusion equation has the following solution.

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} \quad f(x,t) = \frac{m}{\sqrt{4\pi Dt}}e^{-\frac{1}{4}\frac{x^2}{Dt}} \tag{3.19}$$

and now we want to find the solution for equation (3.18), we want to find $f(x,t)$. I will do with a transformation, using the similarity method, using the variable y

$$y = x - \mu t \quad \tilde{f}(y,t) =_{df} f(x,t) \tag{3.20}$$

We need the derivatives of y w.r.t. x and t, when we are using the chain rule from calculus. The derivative w.r.t. t forces you to use the total derivative, as f is a function of t and of y, and y depends on t also.

$$\begin{aligned}\frac{\partial y}{\partial x} &= \frac{\partial \tilde{f}}{\partial y}\frac{\partial y}{\partial x} = \frac{\partial \tilde{f}}{\partial y}\\ \frac{\partial^2 y}{\partial x^2} &= \frac{\partial \tilde{f}}{\partial y}\\ \frac{\partial \tilde{f}}{\partial t} &= \frac{\partial \tilde{f}}{\partial t} + \frac{\partial \tilde{f}}{\partial y}\frac{\partial y}{\partial t} = \frac{\partial \tilde{f}}{\partial t} - \mu\frac{\partial \tilde{f}}{\partial y}\end{aligned} \tag{3.21}$$

We can now plug it into the diffusion - convection equation (3.18) and we get

$$\frac{\partial \tilde{f}}{\partial t} - \mu\frac{\partial \tilde{f}}{\partial y} = D\frac{\partial^2 f}{\partial y^2} - \mu\frac{\partial \tilde{f}}{\partial y} \tag{3.22}$$

The drift terms cancels, and we have the famous diffusion equation. We need to transform the initial conditions for $f(x,t)$ when we transform it to $\tilde{f}(y,t)$, but it will be the same at time equal zero, $f(x,0) = f(\tilde{y},0)$, and the solution will be for y

$$\tilde{f}(y,t) = \frac{m}{\sqrt{4\pi Dt}}e^{-\frac{1}{4}\frac{y^2}{Dt}} \tag{3.23}$$

substitute back we get

$$f(x,t) = \frac{m}{\sqrt{4\pi Dt}} e^{-\frac{1}{4}\frac{(x-\mu t)^2}{Dt}} \tag{3.24}$$

Splitting the square in the exponent in the above equation gives us

$$\begin{aligned} f(x,t) &= e^{-\frac{1}{4}\frac{\mu^2 t^2 - 2x\mu t}{Dt}} \frac{m}{\sqrt{4\pi Dt}} e^{-\frac{1}{4}\frac{x^2}{Dt}} \\ f(x,t) &= e^{-\frac{\mu}{2D}\left(x-\frac{\mu}{t}\right)} \frac{m}{\sqrt{4\pi Dt}} e^{-\frac{1}{4}\frac{x^2}{Dt}} \end{aligned} \tag{3.25}$$

Here we can make a change measure, use the Girsanov theorem. Setting $D = \frac{1}{2}$, then it becomes

$$\begin{aligned} f(x,t) &= e^{-\frac{1}{2}\frac{\mu^2 t^2 - 2x\mu t}{t}} \frac{m}{\sqrt{2\pi t}} e^{-\frac{1}{2}\frac{x^2}{t}} \\ f(x,t) &= e^{-\frac{1}{2}\mu^2 t + \mu x} \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}\frac{x^2}{t}} \end{aligned} \tag{3.26}$$

I replaced m with 1, meaning that we start with 1 particle that behaves as a Brownian Motion and use the fact the Radon- Nikodym density process can be written as

$$\frac{d\mathbb{Q}'}{d\mathbb{Q}} = e^{-\frac{1}{2}\mu^2 t + \mu \tilde{W}_t} \tag{3.27}$$

Then the new process, W'_t is equal to the old process minus the drift, $\tilde{W}_t - \mu t$

3.1.4 Fokker - Plank equation

This equation is also known as the Kolmogorov forward equation, I will use the SDE approach to derive the Fokker - Plank for various stochastic differential equations. Wiener used the path of a particle to build the theory for Brownian motion. The Wiener process starts at zero and has independent and stationary increment, that are normally distributed.

$$\begin{aligned} &\mathbb{P}[W_0 = 0] = 1 \\ &W_t - W_s \sim N(0, t-s) \end{aligned} \tag{3.28}$$

I start with a simple SDE, $X_t = W_t$. To apply Ito's lemma, I need a function, $f()$, let it be arbitrary, and be in C^2, and has compact support, and also f' has compact support. The differential of f thus becomes

$$\begin{aligned} df &= \frac{\partial f}{\partial x} dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} dX_t^2 \\ \mathbb{E}[df] &= \mathbb{E}\left[\frac{\partial f}{\partial x} dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} dX_t^2\right] \\ \mathbb{E}[df] &= \frac{1}{2}\mathbb{E}\left[\frac{\partial^2 f}{\partial x^2}\right] dt \\ \frac{d}{dt}\mathbb{E}[f] &= \frac{1}{2}\mathbb{E}\left[\frac{\partial^2 f}{\partial x^2}\right] \end{aligned} \tag{3.29}$$

I used the fact expectation removes the random part of the Brownian motion and put it in the expectation, since expectation means integral w.r.t. probability. I have also interchanged expectation and derivative. The Fokker - Plank contains probability density in place of Brownian motion. Let the probability density of $p(x,t)$ for a fixed x and a fixed t[2], so the expectation can be written as

$$\frac{d}{dt}\int_{-\infty}^{\infty} f(x)p(x,t)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} f_{XX}(x)p(t,x)\,dx \tag{3.30}$$

Look at the rhs of equation (3.30), we preform integration by parts twice. We want to express our answer in terms of p, $f()$ is just an arbitrary function that we will get dispense at the end. Integration by parts twice gives us

$$\begin{aligned} &\frac{d}{dt}\int_{-\infty}^{\infty} f(x)p(x,t)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} f(x)\frac{\partial^2 p(x,t)}{\partial x^2}\,dx \\ &\int_{-\infty}^{\infty} f(x)\left(\frac{\partial p(x,t)}{\partial t} - \frac{1}{2}\frac{\partial^2 p(x,t)}{\partial x^2}\right)\,dx = 0 \end{aligned} \tag{3.31}$$

and the only way for equation (3.31) to equal zero is if the parenthesis is equal to zero, we said that the function $f()$ was arbitrary,

$$\frac{\partial p(x,t)}{\partial t} = \frac{1}{2}\frac{\partial^2 p(x,t)}{\partial x^2} \tag{3.32}$$

This is the Fokker - Plank equation, it is similar to the diffusion equation for Brownian motion.

3.1.5 A general SDE

Let

$$dX_t = \mu(X_t,t)\,dt + \sigma(X_t,t)\,dW_t \tag{3.33}$$

where the there is a drift and diffusion term, that are not constants. The procedure is as before, let $f()$ be an arbitrary function, with compact support, and also assume its derivative has compact support. Let $f \in C^2$, be twice

[2] It is actually $f(x,t|x_0,t_0)$, it is a Markovian projection

differentiable. We use Ito's lemma.

$$
\begin{aligned}
df &= \frac{\partial f}{\partial x} dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} dX_t^2 \\
&= f_X(\mu dt + \sigma dW_t) + \frac{1}{2} f_{XX} dt \\
&= \left(\mu f_X + \frac{1}{2}\sigma^2 f_{XX}\right) + f_X \sigma dW_t \\
\mathbb{E}[df] &= \mathbb{E}\left[\left(\mu f_X + \frac{1}{2}\sigma^2 f_{XX}\right) + f_X \sigma dW_t\right] \\
\mathbb{E}[df] &= \mathbb{E}\left[\left(\mu f_X + \frac{1}{2}\sigma^2 f_{XX}\right)\right] \\
\frac{d}{dt}\mathbb{E}[f] &= \mathbb{E}\left[\left(\mu f_X + \frac{1}{2}\sigma^2 f_{XX}\right)\right]
\end{aligned}
\tag{3.34}
$$

Let $p(x,t)$ be the probability density function for (forward) for an object starting in position x_0 at time t_0, it is more correct to write the function p as $p(x, t|x_0, t_0)$. This also reveals that it is a forward equation, hence the name Kolmogorov forward equation.

$$
\frac{d}{dt}\int_{-\infty}^{\infty} f(x)p(x,t)\,dx = \int_{-\infty}^{\infty}\left(\mu(x,t)f_X(x) + \frac{1}{2}\sigma^2(x,t)f_{XX}(x)\right)p(t,x)\,dx
\tag{3.35}
$$

Which has the following solution

$$
\int_{-\infty}^{\infty} f(x)\frac{\partial p(x,t)}{\partial t}dx = -\int_{-\infty}^{\infty} f(x)\frac{\partial}{\partial x}\left(\mu(x,t)p(x,t)\right)\,dx + \frac{1}{2}\int_{-\infty}^{\infty} f(x)\frac{\partial^2}{\partial x^2}\left(\sigma^2(x,t)p(x,t)\right)\,dx
\tag{3.36}
$$

and as f is just an arbitrary function we get

$$
\frac{\partial p(x,t)}{\partial t} + \frac{\partial}{\partial x}\left(\mu(x,t)p(x,t)\right) - \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(x,t)p(x,t)\right) = 0
\tag{3.37}
$$

Some examples for the Fokker - Plank equation

We have the general SDE, see equation (3.33) which I reproduce here

$$
dX_t = \mu(X_t,t)\,dt + \sigma(X_t,t)\,dW_t
\tag{3.38}
$$

has the general solution (Fokker- Plank)

$$
\frac{\partial p(x,t)}{\partial t} + \frac{\partial}{\partial x}\left(\mu(x,t)p(x,t)\right) - \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(x,t)p(x,t)\right) = 0
\tag{3.39}
$$

for the Ornstein–Uhlenbeck process, which has the following SDE,

$$
dX_t = -\kappa X_t dt + \sigma dW_t
\tag{3.40}
$$

looking at equation (3.33) and substitute the constant σ for σ and κX for the μ taking the constants out of the derivative and we get

$$\frac{\partial p(x,t)}{\partial t} - \kappa \frac{\partial}{\partial x}\left(xp(x,t)\right) - \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}\left(p(x,t)\right) = 0 \tag{3.41}$$

For the geometric Brownian SDE,

which can be written as

$$dX_t = \mu X_t dt + \sigma dW_t \tag{3.42}$$

we substitute the σX for σ and μX for the μ and arrive at

$$\frac{\partial p(x,t)}{\partial t} + \mu \frac{\partial}{\partial x}\left(xp(x,t)\right) - \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}\left(x^2 p(x,t)\right) = 0 \tag{3.43}$$

we get the Fokker - Plank equation for the Geometric Brownian motion.

Infinitesimal generator

Avoiding the dependency on time, which otherwise that can complicate things for a Markov process, we have

$$\frac{\partial p(x,t)}{\partial t} = \left(-\frac{\partial}{\partial x}\mu(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\sigma^2(x)\right) p(x,t) \tag{3.44}$$

Taking the finite approximation of the derivative as

$$\lim_{\epsilon \to 0} \frac{P(x,t+\epsilon) - P(x,t)}{\epsilon} = L(x)(p(x,t)) \tag{3.45}$$

where $L(x)$ is the linear differential operator, also known as the infinitesimal generator.

$$L(x) = -\frac{\partial}{\partial x}\mu(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\sigma^2(x) \tag{3.46}$$

The infinitesimal generator is usually just the local Taylor series expansion. This concept was also used to Delta-hedge the Black-Scholes PDE. The infintesimal generator can also describe the transition densities in a Markov process, it gives the probability to move from state (i) to state (j) in a short time interval, but this can be expended to work for a process that can take infinitely many states.

3.1.6 Kolmogorov Backward equation

We have the general SDE, see equation (3.33) which I reproduce here

$$dX_t = \mu(X_t,t)\,dt + \sigma(X_t,t)\,dW_t \tag{3.47}$$

This process can used to model many stochastic processes, by changing the drift and the diffusion coefficients, we can arrive at many types of process, e.g.

GBM, the O-U process. The Kolmogorov Forward equation, also known as the Fokker-Plank equation, can be written as.

$$\frac{\partial p(x,t)}{\partial t} - \frac{\partial}{\partial x}\left(\mu(x,t)p(x,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(x,t)p(x,t)\right) = 0 \tag{3.48}$$

This density process $p(x,t)$ is dependent on its conditional start $p(x,t|x_0,t_0)$. Here x_0 and t_0 are fixed. The left hand side tells us what happens when t changes. Here t is a forward variable. The right hand side in equation (3.48) is with respect to the *forward* variable x. In the *backward* Kolmogorov equation, we describe the conditional probability density with respect to the initial time t_0, and the derivatives on the right hand side is with respect to the initial position x_0. The Kolmogorov backward equation has thus the following 3 derivatives.

$$\begin{gathered}\frac{\partial p(x,t|x_0,t_0)}{\partial t_0}\\ \frac{\partial p(x,t|x_0,t_0)}{\partial x_0}\\ \frac{\partial^2 p(x,t|x_0,t_0)}{\partial x_0^2}\end{gathered} \tag{3.49}$$

Let the conditional probability $\mathbb{P}(A,t|x_0,t_0) = \mathbb{P}(X_i \in A|X_0 = x_0)$, if x is a real number it can be viewed as

$$\mathbb{P}(x,t|x_0,t_0) = \mathbb{P}(X_t \leq x|X_0 = x_0) = \int_{-\infty}^{x} p(z,t|x_0,t_0)\,dz = \int_{-\infty}^{x} \mathbb{P}(dz,t|x_0,t_0) \tag{3.50}$$

The Chapman- Kolmogorov equation states that the probability to go from x_0 to x is the same as to go through an intermediate step y, summing (or integrating) over all y in the system.

$$\mathbb{P}(x,t|x_0,t_0) = \int \mathbb{P}(x,t|y,t_1)\mathbb{P}(dy,t_1|x_0,t_0) \tag{3.51}$$

The backward equation is about the variable t, that is going backward. The finite difference approximation is, remember that we are going backward in time and assume $h > 0$

$$\frac{\mathbb{P}(x,t|x_0,t_0) - \mathbb{P}(x,t|x_0,t_0-h)}{h} \tag{3.52}$$

Using the equations (3.51) (3.52) I will re-write $\mathbb{P}(x,t|x_0,t_0-h)$. It becomes, I also change x_0 to x_{-1}

$$\mathbb{P}(x,t|x_{-1},t_0-h) = \int \mathbb{P}(x,t|y,t_0)\mathbb{P}(dy,t_0|x_{-1},t_0-h) \tag{3.53}$$

but we to bring it back to x_0, because the starting values in equation (3.52). We are free to choose whatever initial value as we like.

$$\mathbb{P}(x,t|x_0,t_0-h) = \int \mathbb{P}(x,t|y,t_0)\mathbb{P}(dy,t_0|x_0,t_0-h) \tag{3.54}$$

what equation (3.54) says it that the probability to go x_0 to x in a time interval from time $t_0 - h$ to t, (the left hand side on the above equation) to go from x_0 at time $t_0 - h$ through an intermediate point y and from point y to the value x at time t, the integral is for all $y \leq A$. We can return to equation (3.52) and use the derivative approximation formula.

$$\begin{aligned} &\frac{\int \mathbb{P}(x,t|y,t_0)\mathbb{P}(dy,t_0|x_0,t_0-h) - \mathbb{P}(x,t|x_0,t_0)}{h} \\ &\frac{\int \left(\mathbb{P}(x,t|y,t_0) - \mathbb{P}(x,t|x_0,t_0)\right)\mathbb{P}(dy,t_0|x_0,t_0-h)}{h} \end{aligned} \tag{3.55}$$

Let us focus on last term in the above expression. If we scale it by h it would represent the probability per unit of time.

$$\begin{aligned} &\frac{\mathbb{P}(dy,t_0|x_0,t_0-h)}{h} \\ &\frac{\mathbb{P}(y,t_0|x_0,t_0-h)}{h} \end{aligned} \tag{3.56}$$

if you are at x_0 you would expect that the change over a small interval of time h could be made arbitrarily small, say δ $||y - x_0|| < \delta$ is *diffusion* part, and when δ $||y - x_0|| > \delta$ we have a process with *jumps*. Since this thesis is not about jump -process, I will assume that $||y - x_0|| < \delta$. So the equation (3.55) becomes

$$\int \left(\mathbb{P}(x,t|y,t_0) - \mathbb{P}(x,t|x_0,t_0)\right)||y - x_0|| \tag{3.57}$$

We make a Taylor series expansion of (3.57) in the difference of the probability keeping only terms up to the second order.

$$\begin{aligned} &\mathbb{P}(x,t|y,t_0) - \mathbb{P}(x,t|x_0,t_0)|_{||y-x_0||<\delta} \\ &= \frac{\partial P(x,t|x_0,t_0}{\partial x_0}(y-x_0) + \frac{1}{2}\frac{\partial^2 P(x,t|x_0,t_0}{\partial x_0^2}(y-x_0)^2 \end{aligned} \tag{3.58}$$

note that t is the same in the above expression, the variable that changes is x, we have reduced a problem in 2-dimensions to a 1-dimensional problem. Going back to (3.52) we get

$$\begin{aligned} &\frac{\mathbb{P}(x,t|x_0,t_0-h) - \mathbb{P}(x,t|x_0,t_0}{h} \\ &= \frac{1}{h}\int \frac{\partial P(x,t|x_0,t_0)}{\partial x_0}(y-x_0)\mathbb{P}(dy,t_0|x_0,t_0-h) \\ &+ \frac{1}{h}\int \frac{1}{2}\frac{\partial^2 P(x,t|x_0,t_0)}{\partial x_0^2}(y-x_0)^2\mathbb{P}(dy,t_0|x_0,t_0-h) \end{aligned} \tag{3.59}$$

taking the limit as $h \to \infty$, and noting that the integration is with respect to y, we can take the differential out of the integral. we get

$$
\begin{aligned}
&-\frac{\partial P(x,t|x_0,t_0}{\partial t_0} \\
&=\mu(x_0,t_0)\frac{\partial P(x,t|x_0,t_0)}{\partial x_0}+\frac{1}{2}\sigma^2(x_0,t_0)\frac{\partial^2 P(x,t|x_0,t_0)}{\partial x_0^2}
\end{aligned} \tag{3.60}
$$

Where I have used the fact that the drift is the average displacement per unit time, and the variance i is the squared displacement over a unit time. The minus sign in the first derivative is due to the fact that we are using the finite difference approximation for a value to the left of t, we are going *backward* in time. Stating equation (3.60) in terms of probability density functions we get, it is with respect to the forward variable x Here is the Kolmogorov backward equation

$$
\begin{aligned}
&-\frac{\partial p(x,t|x_0,t_0}{\partial t_0} \\
&=\mu(x_0,t_0)\frac{\partial p(x,t|x_0,t_0)}{\partial x_0}+\frac{1}{2}\sigma^2(x_0,t_0)\frac{\partial^2 p(x,t|x_0,t_0)}{\partial x_0^2}
\end{aligned} \tag{3.61}
$$

and the Kolmogorov forward equation

$$
\begin{aligned}
&\frac{\partial p(x,t|x_0,t_0}{\partial t_0}= \\
&-\mu(x_0,t_0)\frac{\partial p(x,t|x_0,t_0)}{\partial x_0}+\frac{1}{2}\sigma^2(x_0,t_0)\frac{\partial^2 p(x,t|x_0,t_0)}{\partial x_0^2}
\end{aligned} \tag{3.62}
$$

3.2 Bibliographical notes

The diffusion equation begins with (Einstein, 1905) by Einstein. It then describes the diffusion equation, a place to read is (Björk, 2009) and there places in the internet, that provides you with a clear understanding, I used https://quantpie.co.uk to get the intuition behind the ideas.

Chapter 4

Dupire local volatility

4.1 Dupire's local volatility model

This model modifies the Black-Scholes' model, in the Black - Scholes model we assumed a constant volatility, Dupire[1] made a change to the model, assuming that the volatility is a deterministic function of time and stock price $\sigma(t, S_t)$ In the Black - Scholes' model we assume

- that the stock price follows a geometric Brownian motion, which implies that the log returns of the underlying are normally distributed
- constant volatility
- the dynamics of the stock price is, under the risk - neutral measure $\mathbb{Q}$ is $dS_t = rS_t dt + \sigma S_t dW_t$

But when you look at the log returns of financial assets, you often notice, that the return distribution does not look Gaussian, it has higher peak and fatter tails than would be expected if it were Gaussian. Another way of saying that is that very high and very low values occur more often than if it were Gaussian. Another fact from reality is that volatility changes over time, hence it does not appear constant. Supply and demand for different options are also different at different times. You will notice that when you are pricing an option, i.e. the market price of an option. When you plot the implied volatility as a function of strike and maturity, you get a volatility surface which is not flat as the Black - Scholes' model would suggest. The strike dimension of the surface for a given maturity is called the *smile curve*. The maturity, time, dimension of the surface is called the *term structure*. This anomaly would lead you to revisit the Black - Scholes' assumptions, and see which conditions that you need to change. In Dupire's model we change the volatility from being a constant to being a deterministic function of time and asset price $dS_t = rS_t dt + \sigma(t, S_t) S_t dW_t$. This modification

[1] At the same year Derman-Kani also introduced in the same way the local volatility for a binomial setting. I call this the Dupire model, without making any preferences.

will produce the current market prices of options, but the implied dynamics for future times are not good. If the smile curve of the volatility is very steep for the shorter maturities and flattens as the maturities increases. But supply and demand will give a steeper smile curve as we are nearing the time for the longer maturities. It will most likely behave, get a steeper slope just as the shorter maturity. Local volatility only uses today's prices and makes no assumption over the behavior over time, this must be viewed as a weakness of the model. It provides a perfect fit for today's data, but when the data changes, you will need to refit.

4.1.1 Derivation of the Dupire PDE

I will follow Dupire's orginal derivation, (Dupire et al., 1994). The Dupire PDE makes the volatility a function of time and stock price, because we are matching the volatility surface that has maturity dimension as the term structure and smile dimension, where the volatility varies by strike. The Dupire PDE has the following form

$$\sigma^2(T,K) = \frac{\frac{\partial C_{K,T}}{\partial T} + rK\frac{\partial C_{K,T}}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C_{K,T}}{\partial K^2}} \tag{4.1}$$

If you use (4.1) to calculate the local volatility and than use the calculated volatility to price the option, via PDE or Monte Carlo, you will reproduce the initial input prices, that we already know. The advantage is that we can price non-vanilla options, and options that are not quoted in a consistent way. The derivation is similar to the Fokker - Plank derivation forward in time. It uses the Markovian properties, and write the transition from x_0 at time t_0 to move to x at time t we can write it as $p(x,t|x_0,t_0)$, conditional on its initial values, which is known, The call option price has a similar representation it is a function of maturity and strike. So we write $C(K,T|S_t,t)$ conditioning on the the current value of the stock S_t and the current time t, this the standard terminology in the option pricing. The Fokker Plank is its simpler form, for the scaled Brownian motion the Fokker - Plank equation is

$$dX_t = dW_t \tag{4.2}$$

and the solution is the diffusion equation.

$$\frac{\partial p(x,t)}{\partial t} - \frac{1}{2}\sigma^2\frac{\partial^2 p(x,t)}{\partial x^2} = 0 \tag{4.3}$$

and if we isolate the σ we get

$$\sigma^2 = \frac{\frac{\partial p(x,t)}{\partial t}}{\frac{1}{2}\frac{\partial^2 p(x,t)}{\partial x^2}} \tag{4.4}$$

which is very similar to (4.1). Estimating the local volatility is like estimating the diffusion equation. Assume that you know the density at two different times, it is the *inverse problem*. In the diffusion equation we interpret the diffusion coefficients in terms of the distribution of particles over time. If you have the distribution of particles over time, let us say that you know the distribution at two different times, at time t and at time $t + \delta t$, we could find the diffusion coefficient by some finite approximation method. In the Dupire PDE, we replace the distribution function $p(x, t|x_0, t_0)$ with the call option price, and the diffusion term, that is the local volatility function, that is a function of K and T, given the call option prices we can estimate the volatility values at different level of strike and time to maturity. I will follow the steps that Dupire did when he derived the local volatility, he used the Fokker - Plank equation, in many presentation they use the backward diffusion, but I will follow his original presentation.

4.1.2 Dupire's derivation

The derivation of the Dupire PDE has been similar to the forward Kolmogorov equation, or the Fokker-Plank equation. If a process is defined by a SDE

$$dX_t = \mu(X_t, t)\, dt + \sigma(X_t, t)\, dW_t \tag{4.5}$$

than the Fokker - Plank equation is

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial(\mu(x,t)p(x,t))}{\partial x} + \frac{1}{2}\frac{\partial^2(\sigma^2(x,t)p(x,t))}{\partial x^2} \tag{4.6}$$

and the local volatility SDE becoms

$$dS_t = rS_t\, dt + \sigma(S_t, t)S_t\, dW_t \tag{4.7}$$

inserting this into (4.6) it becomes

$$\frac{\partial p(S,T)}{\partial T} = -r\frac{\partial(Sp(S,T))}{\partial S} + \frac{1}{2}\frac{\partial^2(\sigma^2(S,Tp(S,T))}{\partial S^2} \tag{4.8}$$

suppressing the argument in (4.8) to simplify the presentation

$$\frac{\partial p}{\partial t} = -r\frac{\partial(Sp}{\partial S} + \frac{1}{2}\frac{\partial^2(\sigma^2 S^2 p)}{\partial S^2} \tag{4.9}$$

The undiscounted price of a call option as the expectation of the payoff, that is integration for all values of S, the stock price where it is greater than K

$$C^u_{K,T} = \int_{S=K}^{\infty} p(S,T)(S_T - K)\, dS \tag{4.10}$$

let us take its derivative with respect to the maturity T

$$\frac{\partial C^u_{K,T}}{\partial T} = \int_{S=K}^{\infty} \frac{\partial p(S,T)}{\partial T}(S_T - K)\, dS \tag{4.11}$$

The left hand side in the above equation is the Dupire PDE, and the derivative inside the integration is the Fokker- Plank equation, it states the propagation over time for $p(S,T)$

$$
\begin{aligned}
\frac{\partial C^u_{K,T}}{\partial T} &= \int_{S=K}^{\infty} \left(-r\frac{\partial (Sp)}{\partial S} + \frac{1}{2}\frac{\partial^2(\sigma^2 S^2 p)}{\partial S^2} \right) (S_T - K)\, dS \\
\frac{\partial C^u_{K,T}}{\partial T} &= r\int_{S=K}^{\infty} \frac{\partial (Sp)}{\partial S}\, dS + \frac{1}{2}\int_{S=K}^{\infty} \frac{\partial^2(\sigma^2 S^2 p}{\partial S^2}(S_T - K)\, dS
\end{aligned}
\tag{4.12}
$$

Let us look at the two integrals on the right hand side of equation (4.12), I start with the first, I will be using the integration by parts, remember that K is a constant.

$$
\begin{aligned}
&\int_{S=K}^{\infty} \frac{\partial (Sp)}{\partial S}\, dS \\
&= [(S,T)(S-K)]_{S=K}^{\infty} - \int_{S=K}^{\infty} Sp(S,T)\, dS \\
&= -\int_{S=K}^{\infty} Sp(S,T)\, dS
\end{aligned}
\tag{4.13}
$$

Let's at the second integral on the right hand side of equation (4.12), here we need to perform the integration by parts two times.

$$
\begin{aligned}
&= \int_{S=K}^{\infty} \frac{\partial^2(\sigma^2 S^2 p)}{\partial S^2}(S_T - K)\, dS \\
&= \left[\frac{\partial(\sigma^2 S^2 p)}{\partial S}(S-K) \right]_{S=K}^{\infty} - \int_{S=K}^{\infty} \frac{\partial(\sigma^2 S^2 p)}{\partial S}\, dS \\
&= -\int_{S=K}^{\infty} \frac{\partial(\sigma^2 S^2 p)}{\partial S}\, dS \\
&= \left[-(\sigma^2(S,T)S^2 p(S,T) \right]_{S=K}^{\infty} \\
&= \sigma^2(K,T)K^2 p(K,T)
\end{aligned}
\tag{4.14}
$$

So the main expression (4.12), we have thus

$$
\begin{aligned}
\frac{\partial C^u_{K,T}}{\partial T} &= \int_{S=K}^{\infty} \left(-r\frac{\partial (Sp}{\partial S} + \frac{1}{2}\frac{\partial^2(\sigma^2 S^2 p}{\partial S^2} \right) (S_T - K)\, dS \\
&= r\int_{S=K}^{\infty} Sp(S,T)\, dS + \frac{1}{2}\sigma^2(K,T)K^2 p(K,T)
\end{aligned}
\tag{4.15}
$$

Remember that the call option price, expressed with the density is

$$C^u_{K,T} = \int_{S=K}^{\infty} p(S,T)(S-K)\,dS$$
$$C^u_{K,T} = \int_{S=K}^{\infty} Sp(S,T)\,dS - K\int_{S=K}^{\infty} p(S,T)(\,dS$$
$$\int_{S=K}^{\infty} Sp(S,T)\,dS = C^u_{K,T} + K\int_{S=K}^{\infty} p(S,T)(\,dS$$
$$\frac{\partial C^u_{K,T}}{\partial T} = rC^u_{K,T} + rK\int_{S=K}^{\infty} p(S,T)(\,dS + \frac{1}{2}\sigma^2(K,T)K^2p(K,T) \tag{4.16}$$

The undiscounted price for a call option is

$$C^u_{K,T} = \int_{S=K}^{\infty} p(S,T)(S_T-K)\,dS \tag{4.17}$$

taking the derivative with respect to K, applying Leibniz rule of integration we get.

$$\begin{aligned}\frac{\partial}{\partial K}C^u_{K,T} &= -\int_{S=K}^{\infty} p(S,T)\,dS - p(K,T)(K-K)\\ &= -\int_{S=K}^{\infty} p(S,T)\,dS\end{aligned} \tag{4.18}$$

and differentiating again, we get

$$\frac{\partial^2}{\partial K^2}C^u_{K,T} = p(S,T) \tag{4.19}$$

substituting into (4.16) we get

$$\frac{\partial C^u_{K,T}}{\partial T} = rC^u_{K,T} - rK\frac{\partial C^u_{K,T}}{\partial K} + \frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 C^u_{K,T}}{\partial K^2} \tag{4.20}$$

That is the Dupire PDE.

4.1.3 Dupire PDE in moneyness

The strike is not very meaningful concept for the estimation of the local volatility, a strike of 10 for a stock worth 10, is not the same as a strike of 10, for a stock worth 1000, thus we usually write the Dupire PDE in terms as some scaled value, e.g. *moneyness*. Let us define y as the discounted value of the strike divided by the strike price, it can also be written as the strike divided by the forward price[2]

$$y = \log\left(\frac{Ke^{-rT}}{S_0}\right) = \log\left(\frac{K}{F}\right) \tag{4.21}$$

[2] bit unsure what is the forward price, is it the terminal price, at $t = T$, or is the forward with the same maturity as the option, normally the moneyness is defined the other way around, $y = \frac{F}{K}$, the smile curve is usually shown as a function of strike prices

S_0, is a fixed number, the price of today, and it is the strike that gives the option prices. Let us reproduce, the Dupire PDE for the strike

$$\frac{\partial C^u_{K,T}}{\partial T} = rC^u_{K,T} - rK\frac{\partial C^u_{K,T}}{\partial K} + \frac{1}{2}K^2\sigma^2(T,K)\frac{\partial^2 C^u_{K,T}}{\partial K^2} \tag{4.22}$$

and the Dupire PDE, in terms of moneyness, y

$$\frac{\partial \tilde{C}^u_{K,T}}{\partial T} = r\tilde{C}^u_{K,T} + \frac{1}{2}\sigma^2(T,y)\left(\frac{\partial^2 \tilde{C}^u_{y,T}}{\partial y^2} - \frac{\partial \tilde{C}^u_{y,T}}{\partial y}\right) \tag{4.23}$$

You can also write the Dupire PDE in terms of the BS implied volatility

4.2 Bibliographical notes

The starting point in this chapter is (Dupire et al., 1994). In Dupire's derivation he used the forward Kolmogorov equation, in many modern presentations of his study, local volatility, they use the backward Kolmogorov equation. I have chosen not to include it, as it deals with local time, and sub-martingales, which I, at the present time, don't master. You can find treatments of this at https://frouah.com/pages/finmath.html and in the wonderful lecture notes of (Gatheral, 2011). At the year 1994 also (DERMAN & Kani, 1994)

Chapter 5

Fourier Transformation

5.1 Fourier

Here I will follow (Matsuda, 2004) notes. Many of the option pricing models assumes that the stock follows an exponential (geometric) Lévy process.

$$S_t = S_0 e^{L_t} \tag{5.1}$$

where $\{L_t; 0 \le t \le T\}$. In the classic Black- Scholes method, which is a standard Brownian motion with a drift, is the only continuous Lévy process with a risk - neutral Lévy process

$$L_t = \left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t \tag{5.2}$$

That leads to normally distributed conditional risk-neutral log return density

$$\mathbb{Q}\left((\log(\frac{S_T}{S_0}) \middle| \mathcal{F}_0 \right) = \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left(- \frac{\left((\log \frac{S_T}{S_0}) - \left(r - \frac{1}{2}\sigma^2\right) T\right)^2}{2\sigma^2 T} \right) \tag{5.3}$$

The Black - Scholes price is the calculated discounted value of the expected terminal payoff, under the risk - neutral measure $\mathbb{Q}$

$$C(S_0, T) = e^{-rT} \int_K^{\infty} (S_T - K) \mathbb{Q}(S_T | \mathcal{F}_0) \, dS_T \tag{5.4}$$

where $\mathbb{Q}(S_T|\mathcal{F}_0)$ is a log normal density

$$\mathbb{Q}(S_T|\mathcal{F}_0) = \frac{1}{S_T\sqrt{2\pi\sigma^2 T}} \exp\left[- \frac{\left(\log S_T - \left(\log S_0 + (r - \frac{1}{2}\sigma^2)T\right)\right)^2}{2\sigma^2 T} \right] \tag{5.5}$$

It was a well known fact, even before the publication of the famous Black - Scholes paper that the empirical log return density is not a normal distribution, it has excess kurtosis and skewness. The models after Black - Scholes have tried to capture this deviation. (Carr & Madan, 1999) did a re-write of ((5.4)), in terms of a CF of the conditional log terminal stock price $\phi(\log S_T|\mathcal{F}_0)$

$$C(T,k) = \frac{e^{-\alpha k}}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega k}\frac{e^{-rT}\phi_T(\omega-(\alpha+1)i)}{\alpha^2+\alpha-\omega^2+i(2\alpha+1)\omega}\,d\omega \tag{5.6}$$

The option pricing with Fourier transforms is simple, and will work if the CF of the conditional log terminal stock price $S_T|\mathcal{F}_0$ is obtained in closed form.

5.1.1 Definition of the Fourier Transform, and Characteristic function

Assume there is a function $g(t)$ from a time domain t, into an angular frequency domain ω and let $\omega = 2\pi f$, where f is the frequency. In the most general setting, the Fourier transform is a function $g(t)$ to a function $\mathscr{G}(\omega)$, going from the time to the frequency domain, is defined using two constants a and b, the FT parameters.

$$\mathscr{G}(\omega) \equiv \mathscr{F}[g(t)](\omega) \equiv \sqrt{\frac{|b|}{(2\pi)^{1-a}}}\int_{-\infty}^{\infty} e^{ib\omega t}g(t)\,dt \tag{5.7}$$

and the inverse Fourier transform will be

$$g(t) \equiv \mathscr{F}^{-1}[\mathscr{G}(\omega](t) \equiv \sqrt{\frac{|b|}{(2\pi)^{1+a}}}\int_{-\infty}^{\infty} e^{-ib\omega t}\mathscr{G}(\omega)\,d\omega \tag{5.8}$$

In order to calculate the characteristic function, set the parameters $(a,b) = (1,1)$ [1] and thus (5.7) and (5.8) becomes

$$\begin{aligned} \mathscr{G}(\omega) &\equiv \mathscr{F}[g(t)](\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega t}g(t)\,dt \\ g(t) &\equiv \mathscr{F}^{-1}[\mathscr{G}(\omega](t) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega t}\mathscr{G}(\omega)\,d\omega \end{aligned} \tag{5.9}$$

Let X be a random variable with the probability density function $\mathbb{P}(x)$, a characteristic function $\phi(\omega)$ with $\omega \in \mathbb{R}$ is defined as the Fourier transform of the probability density function $\mathbb{P}(x)$, using Fourier transform parameters $(a,b) = (1,1)$ From the definition (5.9)

$$\phi(\omega) \equiv \mathscr{F}[\mathbb{P}(x)] \equiv \int_{-\infty}^{\infty} e^{i\omega x}\mathbb{P}(x)\,dx \equiv \mathbb{E}\left[e^{i\omega x}\right] \tag{5.10}$$

The probability density function can be obtained by the inverse Fourier transform of the characteristic function.

$$\mathbb{P}(x) \equiv \mathscr{F}^{-1}[\phi(\omega)] \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega x}\phi(\omega)\,d\omega \equiv \mathbb{E}\left[e^{i\omega x}\right] \tag{5.11}$$

[1] in pure mathematics the pair $(a,b) = (1,-1)$ is used, in modern physics the pair $(a,b) = (0,1)$ is used, and in signal processing the pairing $(a,b) = (0,-2\pi)$ is used.

5.1.2 Derivation, Carr - Madan (1999), of the call price with Fourier transform

Let $\mathbb{Q}(S_T|\mathcal{F}_t)$ be the pdf of the terminal asset price S_T under the risk neutral measure $\mathbb{Q}$ conditional on the information at $\mathcal{F}_t$. The call price will thus be

$$\begin{aligned} C(t,S_t) &= e^{-r(T-t)}\left[\int_K^\infty (S_T-K)\mathbb{Q}(S_T|\mathcal{F}_t\, dS_T + \int_0^K (0)\mathbb{Q}(S_T|\mathcal{F}_t\, dS_T\right] \\ C(t,S_t) &= e^{-r(T-t)}\int_K^\infty (S_T-K)\mathbb{Q}(S_T|\mathcal{F}_t\, dS_T \end{aligned} \tag{5.12}$$

for the further derivation, assume $t=0$, change the stock asset variable to its logarithm $S_T = \log(S_T) \equiv s_T$ and do the same for the strike $K = \log K \equiv k$. We can rewrite (5.12) as

$$C(T,k) = e^{-rT}\int_k^\infty (e^{s_T}-e^k)\mathbb{Q}(s_T|\mathcal{F}_0)ds_T \tag{5.13}$$

The characteristic function of s_T is a Fourier transform of its density function $\mathbb{Q}(s_T)$

$$\phi_T(\omega) \equiv \mathscr{F}[\mathbb{Q}(s_T)](\omega) = \int_{-\infty}^{\infty} e^{i\omega s_T}\mathbb{Q}(s_T)\, ds_T \tag{5.14}$$

In the Black - Scholes model, the log of the terminal stock price has the following density, (5.4), combine that with (5.14), we get the characteristic function

$$\phi_T(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega s_T}\mathbb{Q}(s_T)\, ds_T = \exp\left[i\left\{s_0 + (r-\frac{1}{2}\sigma^2)T\right\}\omega - \frac{(\sigma^2 T)\omega^2}{2}\right] \tag{5.15}$$

When the call price is expressed under the logarithm, $k \equiv \log(K)$ in (5.4) we get

$$\begin{aligned} C(T,k) &= e^{-rT}\int_{-\infty}^{\infty}\left[e^{s_T}-e^{-\infty}\right]\mathbb{Q}(s_T|\mathcal{F}_0)\, ds_T = e^{-rT}\int_{-\infty}^{\infty} e^{s_T}\mathbb{Q}(s_T|\mathcal{F}_0)\, ds_T \\ C(T,k) &= e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[e^{s_T}\,|\,\mathcal{F}_0\right] \end{aligned} \tag{5.16}$$

under the equivalent martingale measure

$$\mathbb{E}^{\mathbb{Q}}\left[S_T \equiv e^{s_T}\,|\,\mathcal{F}_0\right] = S_0 e^{rT} \tag{5.17}$$

So we end up with

$$C(T,k) = S_0 \tag{5.18}$$

A sufficient, but not necessary condition for the Fourier transform and its inverse, is, for a given function $g(t)$

$$\int_{-\infty}^{\infty} |g(t)|^2\, dt < \infty \tag{5.19}$$

So the (5.18) can't have a Fourier transform. Carr - Madan, defined a modified call price

$$C_{mod}(T,k) = e^{\alpha k} C(T,k) \tag{5.20}$$

and by carefully choosing $\alpha > 0$ we get

$$\int_{-\infty}^{\infty} |C_{mod}(T,k)| \, dt < \infty \tag{5.21}$$

Using the definition of the Fourier transform we have

$$\psi_T(\omega) = \int_{-\infty}^{\infty} e^{i\omega k} C_{mod}(T,k) \, dk \tag{5.22}$$

as it is the call price that we want, we use the definition of the inverse Fourier transform

$$\begin{aligned} C_{mod}(T,k) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \psi_T(\omega) \, d\omega \\ e^{\alpha k} C(T,k) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \psi_T(\omega) \, d\omega \\ C(T,k) &= \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \psi_T(\omega) \, d\omega \end{aligned} \tag{5.23}$$

Carr Madan than derived an analytic expression for $\psi_T(\omega)$ in terms of the characteristic function, and the end result is

$$\psi_T(\omega) = \frac{e^{-rT} \phi_T(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega} \tag{5.24}$$

So the call option price becomes

$$C(T,k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega} \, d\omega \tag{5.25}$$

5.1.3 Characteristic function

The characteristic function (CF) for any random variable X, completely defines its probability distribution. On the real line it is defined as

$$\phi_X(u) \coloneqq \mathbb{E}\left[e^{iux}\right] = \int_{-\infty}^{\infty} e^{iux} f_X(x) \, dx = \int_{\Omega} e^{iux} \, dF(x) \tag{5.26}$$

where $u \in \mathbb{R}$. For option prices we extend the definition to the complex plane, $u \in \mathcal{D} \subseteq \mathbb{C}$, where $\mathcal{D}$ denotes the subset of the complex plane on which the expectation is well defined. $\phi_X(u)$ under the extended definition is called the generalized Fourier transform. Note that the generalized Fourier transform includes as a special case the Laplace transform, when ($\Im(u) > 0$) and the cumulative generating function, when ($\Im(u) < 0$), if they are well defined.

Example 5.1.1. Under B-S model, the log security return is given by

$$s_T := \log\left(\frac{S_t}{S_0}\right) = \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t \tag{5.27}$$

the return is normally distributed with mean $\left(\mu - \frac{1}{2}\sigma^2\right) t$ and variance $\sigma^2 t$. The pdf for a normal distributed r.v. Z is

$$\frac{1}{\sqrt{2\pi(Var(Z)}} \exp\left\{-\frac{1}{2}\left(\frac{z - \mathbb{E}(\mathbb{X})}{\sqrt{VarZ}}\right)^2\right\} \tag{5.28}$$

and for the above mean and variance the PDF for s will be

$$f_s(x) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left\{-\frac{1}{2}\left(\frac{x - \left(\mu - \frac{1}{2}\sigma^2\right) t}{\sqrt{\sigma^2 t}}\right)^2\right\} \tag{5.29}$$

The characteristic function, (CF) will be

$$\phi_s(u) = \mathbb{E}[e^{ius_t}] = e^{iu\mathbb{E}s_t + \frac{1}{2}(iu)^2 Var(s_t)} = \exp\left\{iu\mu t - \frac{1}{2}\sigma^2(iu + u^2)t\right\} \tag{5.30}$$

5.1.4 Inversion of the CF to the CDF/PDF

There is a bijection between the CDF and CF, that means that two different probability distribution never share the same CF. Given a CF, ϕ it is possible to reconstruct the corresponding CDF. It can be done in different ways

$$F_X(y) - F_X(x) = \lim \tau \to \infty \frac{1}{2\pi} \int_{-\tau}^{+\tau} \frac{e^{-iux} - e^{-iuy}}{iu} \phi_X(u)\, du \tag{5.31}$$

Another form of inversion is

$$F_X(x) = \frac{1}{2} + \frac{1}{2\pi} \int_{u=0}^{\infty} \frac{e^{iux}\phi_X(-u) - e^{-iux}\phi_X(u)}{iu}\, du \tag{5.32}$$

and the inversion for PDF, remember that $f(x) = F'(x)$ is

$$f_X(x) = \frac{1}{2\pi} \int_{u=-\infty}^{\infty} e^{-iux}\phi_X(u)\, du = \frac{1}{\pi} \int_{u=0}^{\infty} \Re\left(e^{-iux}\phi_X(u)\right)\, du \tag{5.33}$$

The integrals are to be taken as the principal value.

5.1.5 Proofs and facts for the Fourier inversion

- $e^{iux} = \cos(ux) - i\sin(ux)$
- $\frac{1}{\pi}\int_{u=-\infty}^{\infty} \frac{e^{-iu\zeta}}{iu})\, du = \frac{1}{\pi}\int_{u=-\infty}^{\infty} \frac{\sin(u\zeta)}{u}\, du = sgn(\zeta)$

- $\frac{2}{\pi}\int_{u=0}^{\infty}\frac{\sin(u\zeta)}{u}\,du = sgn(\zeta)$
- $\int_{y=-\infty}^{\infty} sgn(y-x)\,dF(y) = -\int_{y=-\infty}^{x} dF(y) + \int_{y=x}^{\infty} dF(y) = 1-2F(x)$
- $\phi(u)$ and $\phi(-u)$ are complex conjugate

Here comes a proof of the inversion formula, for a more detailed proof, see for example (Kendall et al., 1946)

$$\begin{aligned}
I = & \int_{u=0}^{\infty}\frac{e^{iux}\phi_X(-u)-e^{-iux}\phi_X(u)}{iu}\,du \\
& \int_{u=0}^{\infty}\int_{z=-\infty}^{\infty}\frac{e^{iux}e^{-iuz}-e^{-iux}e^{iuz}}{iu}\,dF(z)\,du \\
& \int_{u=0}^{\infty}\int_{z=-\infty}^{\infty}\frac{2\sin(u(x-z))}{u}\,dF(z)\,du \\
& \int_{z=-\infty}^{\infty}\int_{u=0}^{\infty}\frac{2\sin(u(x-z))}{u}\,du\,dF(z) \\
& \int_{z=-\infty}^{\infty}\pi sgn(x-z)\,dF(z) = \pi\left(2F(x)-1\right)) \\
& \text{hence } F(x) = \frac{1}{2}+\frac{1}{2\pi}I \quad \square
\end{aligned} \tag{5.34}$$

and the PDF inversion

$$f(x) = F'(x) = \frac{1}{2\pi}\int_{u=0}^{\infty}\frac{e^{iux}\phi_X(-u)-e^{-iux}\phi_X(u)}{iu}\,du = \frac{1}{\pi}\int_{u=0}^{\infty}e^{-iux}\phi_X(u)\,du \tag{5.35}$$

5.1.6 Inversion of an option

Take a European call option, preform the following rescaling and change of variable

$$c(k) = e^{rt}\frac{c(K,t)}{F_0} = \mathbb{E}_0^{\mathbb{Q}}\left[(e^{S_t}-e^k)\mathbb{1}_{s_t>k}\right] \tag{5.36}$$

with $s_t = \log\left(\frac{F_t}{F_0}\right)$ and $k = \log\left(\frac{K}{F_0}\right)$ where $c(k)$ is the option forward price in percentage of the underlying forward as a function of moneyness, defined as the log strike over forward k. We can now derive the Fourier transform of the call option in terms of the Fourier Transform, of the log-returns $\log\left(\frac{F_t}{F_0}\right)$. If we know the CF of the returns, we would know the transform of the option, then we can use numerical inversion to obtain the option price directly. The idea is to treat the call option $c(k)$ as a CDF, that idea was proposed by (Duffie et al., 1999), and (Singleton, 2001). The setup is as follows.

$$c(k) = \mathbb{E}_0^{\mathbb{Q}}\left[(e^{S_t}-e^k)\mathbb{1}_{s_t>k}\right] = \int_{s=-\infty}^{\infty}\left(e^{s_t}-e^k\right)\mathbb{1}_{s_t>k}\,dF(s) \tag{5.37}$$

The option transform will be

$$\chi_c'(u) = \int_{k=-\infty}^{\infty} e^{iuk}\, dc(k) = -\frac{\phi_s(u-i)}{iu+1}, u \in \mathbb{R} \tag{5.38}$$

and the inversion formula will be

$$c(x) = \frac{1}{2} + \frac{1}{2\pi}\int_{u=0}^{\infty} \frac{e^{iux}\chi_c'(-u) - e^{-iux}\chi_c'(u)}{iu}\, du \tag{5.39}$$

Proof of the option transform

Proof 5.1.1. Proof of the option transform. Using integration by parts

$$\chi_c'(u) = \int_{k=-\infty}^{\infty} e^{iuk}\, dc(k) = \left[e^{iuk}c(k)\right]_{k=-\infty}^{\infty} - \int_{k=-\infty}^{\infty} c(k)iue^{iuk}\, dk \tag{5.40}$$

the boundary conditions give that $c(\infty) = 0$, when strike is infinity and $c(-\infty) = 1$ when the strike is zero, remember we have been scaling the option parameters, hence $e^{iu\infty} = 0$

$$\begin{aligned}
\chi_c'(u) = e^{-iu\infty} \int_{k=-\infty}^{\infty} & iue^{iuk}\, dk \\
e^{-iu\infty} - iu \int_{k=-\infty}^{\infty} & \left[\int_{s=-\infty}^{\infty} \left(e^{s_t} - e^k\right) \mathbb{1}_{s_t > k}\, dF(s)\right] e^{iuk}\, dk \\
e^{-iu\infty} - iu \int_{s=-\infty}^{\infty} & \left[\int_{k=-\infty}^{\infty} \left(e^{s_t} - e^k\right) \mathbb{1}_{s_t > k} e^{iuk}\, dk\right] dF(s) \\
e^{-iu\infty} - iu \int_{s=-\infty}^{\infty} & \left[\int_{k=-\infty}^{s_t} \left(e^{iuk+s_t} - e^{(iu+1)k}\right) dk\right] dF(s) \\
e^{-iu\infty} - iu \int_{s=-\infty}^{\infty} & \left[e^{s_t}\frac{e^{iuk}}{iu} - \frac{e^{(iu+1)k}}{iu+1}\Bigg|_{k=\infty}^{k=s_t} \right] dF(s)
\end{aligned} \tag{5.41}$$

Another check on the boundary conditions $\lim_{k\to-\infty} e^{(iu+1)k} = 0$ given the real component $e^{-\infty}$, the other boundary is non-convergent $e^{s_t}e^{-iu\infty}$, which we pull out and take the expectation to have

$$iu \int_{s=-\infty}^{\infty} \frac{e^{s_t}e^{-iu\infty}}{iu}\, dF(s) = e^{-iu\infty} \tag{5.42}$$

which cancel with the other non-convergent term

$$\begin{aligned}
\chi_c'(u) = -iu \int_{s=-\infty}^{\infty} \left[\frac{e^{(iu+1)s_t}}{iu} - \frac{e^{(iu+1)s_t}}{iu+1}\right] dF(s) = \\
-\int_{s=-\infty}^{\infty} \left[\frac{e^{(iu+1)s_t}}{iu+1}\right] dF(s) = -\frac{\phi(u-i}{iu+1}
\end{aligned} \tag{5.43}$$

The scaled version of the $c(k)$ behaves as a CDF, in particular it has $c(\infty) = 0$, when strike is infinity and $c(-\infty) = 1$ when strike is zero.

$$
\begin{aligned}
I = \int_{u=0}^{\infty} \frac{e^{iux}\chi(-u) - e^{-iux}\chi(u)}{iu}\, du = \\
\int_{z=-\infty}^{\infty} \pi sgn(x-z)\, dF(z) = -\pi(1-2c(x)) \\
\text{thus } c(x) = \frac{1}{2} + \frac{1}{2\pi} I
\end{aligned}
\tag{5.44}
$$

Treat c(k) as a PDF, than the option transform is, for more information please look at (Carr & Wu, 2004)

$$
\chi_c''(z) = \int_{k=-\infty}^{\infty} e^{izk} c(k)\, dk = \frac{\phi_x(z-i)}{(iz)(iz+1)} \tag{5.45}
$$

with $z = u - i\alpha, \alpha \in \mathcal{D} \subseteq \mathbb{R}^+$ and the inversion is analogous to that for a PDF

$$
c(k) = \frac{1}{2\pi} \int_{z=-ia-\infty}^{-ia+\infty} e^{-izk} \chi_c''(z)\, dz = \frac{e^{-\alpha k}}{\pi} \int_{u=0}^{\infty} e^{-izk} \chi_c''(u - i\alpha)\, du \tag{5.46}
$$

Proof 5.1.2. Proof of the inversion transformation for the option prices.

$$
\begin{aligned}
\chi_c''(z) &= \int_{k=-\infty}^{\infty} e^{izk} c(k)\, dk \\
& \int_{k=-\infty}^{\infty} \left[\int_{s=-\infty}^{\infty} \left(e^{s_t} - e^k\right) \mathbb{1}_{s_t > k}\, dF(s) \right] e^{izk}\, dk \\
&= \int_{s=-\infty}^{\infty} \left[\int_{k=-\infty}^{\infty} \left(e^{s_t} - e^k\right) \mathbb{1}_{s_t > k}\, e^{izk}\, dk \right] dF(s) \\
&= \int_{s=-\infty}^{\infty} \left[\int_{k=-\infty}^{s_t} \left(e^{izk+s_t} - e^{(iz+1)k}\right) dk \right] dF(s) \\
&= \int_{s=-\infty}^{\infty} \left[e^{s_t} \frac{e^{izk}}{iz} - \frac{e^{(iz+1)k}}{iz+1} \Bigg|_{k=\infty}^{k=s_t} \right] dF(s)
\end{aligned}
\tag{5.47}
$$

We need again to consider the boundary conditions at $k = -\infty, \lim_{k\to-\infty} e^{(iz+1)k} = 0$, as long the real component of iz is greater than -1,and $\lim_{k\to-\infty} e^{(iz+1)k} = 0$ as long as the real component is greater than 0. We need $\alpha > 0$ for the boundary conditions to converge. Given that $u_i > 0$ we have

$$
\begin{aligned}
\chi_c''(z) = \int_{s=-\infty}^{\infty} \left[\frac{e^{(iz+1)s_t}}{iz} - \frac{e^{(iz+1)s_t}}{iz+1} \right] dF(s) = \\
\int_{s=-\infty}^{\infty} \left[\frac{e^{(iz+1)s_t}}{(iz)(iz+1)} \right] dF(s) = \frac{\phi(z-1)}{iz(iz-1)}
\end{aligned}
\tag{5.48}
$$

5.2 Bibliographical notes

The theory of Fourier transformation is huge. A good starting point is a course from Stanford University, EE261 - The Fourier Transform and its Applications, you can find it at https://see.stanford.edu/Course/EE261 a good note from a financial emphasis is (Matsuda, 2004). The financial application began with (Carr & Madan, 1999) another good introduction is (Černý, 2004). Also the PPT -presation by Liuren Wu [2] has been used. A financial view of Fourier transformation can be found in (Pascucci, 2011). A standard text book is (Kendall et al., 1946). To see the connection between the option price and the characteristic function, look at (Carr & Wu, 2004)

[2] `http://faculty.baruch.cuny.edu/lwu/890/ADP_Transform.pdf`, retr. 2021-08-04

Chapter 6

Heston model

6.1 Heston Model

I will derive the pricing PDE for the Heston stochastic volatility model. I will follow the same steps as in the Black - Scholes derivation of the PDE, using the delta-hedging argument, but for the Heston model, there are two sources of randomness, the stock price, and the variance-process. The stock price is assumed to follow the geometric Brownian motion, with volatility, that was a constant in the Black - Scholes model, is now a stochastic process[1], that follows a square root diffusion process, this is the same as the CIR - interest rate model.

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dZ_1 \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dZ_2 \end{aligned} \tag{6.1}$$

There is a problem with the classic Black Scholes model, the model does not fit the observations, the prices observed by the market. These observations became even more at odds with the Black Scholes model for the data after the black Monday, in October 1987 Heston assumed that the variance follows a square root diffusion process, it is the same process that is in the CIR- interest rate modeling. with a mean-reversion. The two Brownian motion in (6.1) are assumed to be correlated with a constant correlation, ρ

$$\mathbb{E}\left[dZ_1, dZ_2\right] = \rho dt \tag{6.2}$$

It follows from Ito calculus, (6.2), that $dt^2 = 0$, and $dt \cdot dZ_1 = 0$ There is also a need for the Bank account, as in the Black - Scholes model, which has a deterministic growth at a constant rate r.

$$dB_t = rB_t dt \tag{6.3}$$

The problem with Black - Scholes model, is that the constant variance gives normally distributed log-returns, but the market does not follow this behavior.

[1]sometimes v is variance, and sometimes in other books volatility, the square root of the variance.

This is called the stylized facts, in the financial mathematical literature. When observing the log-returns, you see higher peaks and fatter tails, there is also the smile curve in the observed log-returns, were the Black - Scholes model to be correct, you would see a flat surface over the maturity dimension and the strikes. The correlation ρ gives control over the relationship between the volatility and the stock price. For example, the volatility is usually higher when prices are depressed. Setting $\rho < 0$ gives you that feature in Heston's model. The correlation also effect the *skew* of the volatility surface. Remember that the variance is positive, that is why there is square root in the variance of dv_t, and mean-reverting. As there are only a limited number of parameters, it makes the calibration complicated, but it is still an attractive model, since it capture more of the dynamics of the stocks. It does not capture the skew seen in the short tenors, for that the jump process can be an attractive model, but then the Heston model would compete with other models, for example Bates modeling, local stochastic volatility models with jumps. The Heston model belongs to the class of *affine jump diffusion* AJD, but without jumps, the J_t are set to zero.

Let us start with deriving the pricing PDE for the Heston model. The price of the option V will depend on the time to maturity, and the two diffusion process, S_t, v_t, I will suppress, the (r, K, T), as they are constant in the Heston model.

$$V = V(t, S_t, v_t; r, K, T) \sim V(t, S_t, v_t) \tag{6.4}$$

The differences to the Black Scholes pricing PDE and the Heston model pricing PDE

BS	Heston
One Brownian motion	Two Brownian motion
One source of randomness	Two sources of randomness
Use Delta hedging	Use Delta and Sigma hedging
Complete market	Volatility is not traded, hence incomplete market
Unique Martingale measure	Many Martingale measure

To start solving the Heston PDE, we need the 2-dimensional version of Ito's lemma, with time dependency. I will avoid writing the subscript t, to the two random processes, but they are always there, to make the presentation clearer. We begin with the a small change in the value process. I will use the two-dimensional version of Ito's lemma, with time -dependency.

$$dV = \frac{\partial V}{\partial t}dt + \left(\frac{\partial V}{\partial S}dS + \frac{\partial^2 V}{\partial S^2}dS^2\right) + \left(\frac{\partial V}{\partial v}dv + \frac{\partial^2 V}{\partial v^2}dv^2\right) + \frac{\partial^2 V}{\partial v \partial S}dvdS$$

leave the dS, dv, unchanged and subsitute for $dS^2\, dv^2$, and the cross-term, $dSdv$

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

combine the dt terms

$$\left(\frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2} + \rho\sigma v S\frac{\partial^2 V}{\partial v \partial S}\right)dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv \tag{6.5}$$

Use the differential operator, and denote it by L

$$L = \frac{\partial}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2}{\partial S^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2}{\partial v^2} + \rho\sigma vS\frac{\partial^2}{\partial v\partial S} \tag{6.6}$$

and the drift term can be written as (LV) and the arguments then by using the differential operator (6.6) in (6.5) gives us

$$dV = (LV)(t,s,v)dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv \tag{6.7}$$

To get the two randomnesses in Heston's model, we need two options with different maturities T_1 and T_2, let the value of the option at the shorter maturity, T_1 be $V = V(t, S_t, v_t; r, K, T_1)$, we need two assets, one is the stock price as in BS, and for the other randomness, assume that we know the option value at another time T_2 and denote it by $U = U(t, S_t, v_t; r, K, T_2)$, where $T_1 < T_2$. We need a longer maturity to be able to hedge the option with the shorter maturity all to its maturity. The changes for the second option, dU is given by Ito differential, for two processes and time dependency.

$$dU = (LU)(t,s,v)dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv \tag{6.8}$$

We are hedging the risk of the option with maturity T_1 using the stock price and longer maturity option, T_2. We are constructing a portfolio, with delta unit of the stock and sigma units of the second option, with the longer maturity, and alpha units from the bank account.

$$V = \Delta S + \Sigma U + \alpha B \tag{6.9}$$

Using the self-financing concept, we can find the change in the value of the option, by[2]

$$dV = \Delta dS + \Sigma dU + \alpha dB \tag{6.10}$$

we have (6.7) and (6.8) and (6.3) is only a deterministic growth. Than (6.10) becomes

$$\begin{aligned} dV &= \Delta dS + \Sigma dU + \alpha dB \\ &= \Delta dS + \Sigma\left((LU)(t,s,v)dt + \frac{\partial U}{\partial S}dS + \frac{\partial U}{\partial v}dv\right) + arBdt \\ &\text{combine the dS-terms} \\ &= \Sigma(LU)(t,s,v)dt + \left(\Delta + \Sigma\frac{\partial U}{\partial S}\right)dS + \Sigma\frac{\partial U}{\partial v}dv + arBdt \end{aligned} \tag{6.11}$$

the stochastic terms, and the source of randomness are dS, dv, our aim is to remove them. Equating the dS in the first two lines in (6.11) gives us

$$\begin{aligned} \frac{\partial V}{\partial S}dS &= \Delta dS + \Sigma\frac{\partial U}{\partial S}dS \\ \frac{\partial V}{\partial S} &= \Delta + \Sigma\frac{\partial U}{\partial S} \end{aligned} \tag{6.12}$$

[2]note that no cross-term is needed, e.g. (?, ?), rebalancing after each short step.

and for the dv

$$\frac{\partial V}{\partial v} = \Sigma \frac{\partial U}{\partial v} \tag{6.13}$$

that gives that

$$\Sigma = \frac{\frac{\partial V}{\partial v}}{\frac{\partial U}{\partial v}} \tag{6.14}$$

isolate Δ in (6.12)

$$\Delta = \frac{\partial V}{\partial S} - \Sigma \frac{\partial U}{\partial S} \tag{6.15}$$

after some cancellation we get

$$(LV)(t,s,v)dt \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial v} dv = \Sigma(LU)(t,s,v)dt + \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial v} dv + arBdt \tag{6.16}$$

Cancellation of the stochastic terms gives

$$(LV)(t,s,v)dt = \Sigma(LU)(t,s,v)dt + arBdt \tag{6.17}$$

Use the replicating portfolio to get rid of the bank account. Remember that $\alpha B = V - \Delta S - \Sigma U$, just a re-write of (6.9), and using $\Delta = \frac{\partial V}{\partial S}$

$$\alpha B = V - \frac{\partial V}{\partial S} S + \Sigma \frac{\partial U}{\partial S} - \Sigma U \tag{6.18}$$

put it in the (6.16)

$$(LV)(t,s,v)dt = \Sigma(LU)(t,s,v)dt + r\left(V - \frac{\partial V}{\partial S} S + \Sigma \frac{\partial U}{\partial S} - \Sigma U\right) dt \tag{6.19}$$

every term has dt so you can get rid off it, put all the terms containing V which PDE we are after, remember that it is the option with shorter maturity, to the left hand side.

$$(LV)(t,s,v)dt - rV + r\frac{\partial V}{\partial S} S = \Sigma(LU)(t,s,v) + r\Sigma \frac{\partial U}{\partial S} - r\Sigma U \tag{6.20}$$

and we end up after some more straight forward calculation

$$\frac{(LV)(t,s,v) - rV + r\frac{\partial V}{\partial S} S}{\frac{\partial V}{\partial v}} = \frac{(LU)(t,s,v) - rU + r\frac{\partial U}{\partial S} S}{\frac{\partial U}{\partial v}} \tag{6.21}$$

This means that the fraction must not depend on V, U, but most depend on the parameters, I ignore the normalization.

$$\frac{(LV)(t,s,v) - rV + r\frac{\partial V}{\partial S} S}{\frac{\partial V}{\partial v}} = -f(t,s,v) \tag{6.22}$$

So the option for V is

$$(LV)(t,s,v) - rV + rS\frac{\partial V}{\partial S} = -f(t,s,v)\frac{\partial V}{\partial v} \tag{6.23}$$

A more revealing rewritten is

$$(LV)(t, s, v) - rV = -rS\frac{\partial V}{\partial S} - f(t, s, v)\frac{\partial V}{\partial v} \tag{6.24}$$

Remember that rS is the drift of the stock price S so f must also be some drift of the SDE.

6.1.1 Intuition of Heston model

The dynamics of the stock price is given by

$$dS = \mu S dt + \sqrt{v} S dZ_1 \tag{6.25}$$

and the variance is given by

$$dv = \kappa(\theta - v)dt + \sigma\sqrt{v}dZ_2 \tag{6.26}$$

If a risk neutral measure were to exist, than we know that the drift of the stock price is rS

$$dS = rSdt + \sqrt{v}SdZ_1^{\mathbb{Q}} \tag{6.27}$$

The connection of the drift under the physical and risk - neutral measure can be seen by the following Girsanov theorem.

$$\frac{dS}{S} = \left(\mu - \frac{\mu - r}{\sqrt{v}}\sqrt{v}\right) dt + \sqrt{v}dZ_1^{\mathbb{Q}} \tag{6.28}$$

the fraction in the above formula is the excess return divided by the volatility, is used in the CAPM, and market price of risk, usually denoted by λ. Use this in the variance SDE, and call it λ

$$dv = \left(\kappa(\theta - v) - \lambda\sigma\sqrt{v}\right) dt + \sigma\sqrt{v}dZ_2^{\mathbb{Q}} \tag{6.29}$$

because the volatility is not a traded asset, we need another specification in dynamics of the variance. So the

$$f(t, s, v) = \kappa(\theta - v) - \lambda\sigma\sqrt{v} \tag{6.30}$$

Going back to (6.24) we get

$$(LV)(t, s, v) - rV = -rS\frac{\partial V}{\partial S} - \kappa(\theta - v) - \lambda\sigma\sqrt{v}\frac{\partial V}{\partial v} \tag{6.31}$$

And going to definition of the linear operator (6.6)

$$L = \frac{\partial}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2}{\partial S^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2}{\partial v^2} + \rho\sigma v S\frac{\partial^2}{\partial v \partial S} \tag{6.32}$$

we get the Heston PDE

$$\begin{aligned} \frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2} + \rho\sigma v S\frac{\partial^2 V}{\partial v \partial S} - rV = \\ - rS\frac{\partial V}{\partial S} - \kappa(\theta - v) - \lambda\sigma\sqrt{v}\frac{\partial V}{\partial v} \end{aligned} \tag{6.33}$$

I will now show that $f(t, s, v) = \kappa(\theta - v) - \lambda\sigma\sqrt{v}$ is the market price of risk.

6.1.2 Another specification of the Heston model

We have the following system of the Browian motions see equation (6.1), which I reproduce here

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dZ_1 \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dZ_2 \end{aligned} \tag{6.34}$$

We know that the two Brownian motions are correlated

$$\mathbb{E}\left[dZ_1, dZ_2\right] = \rho\, dt \tag{6.35}$$

if we try to express them as uncorrelated Brownian motion, where we want

$$\mathbb{E}\left[dW_1, dW_2\right] = 0 \tag{6.36}$$

we can proceed as follows, let the second Brownian motion, be the same, in the new specification.

$$dZ_2 = dW_2 \tag{6.37}$$

and the other as linear combination of the, correlated Brownian motion, the old Brownian, as in equation (6.1)

$$dZ_1 = \sqrt{1-\rho^2} dW_1 + \rho dW_2 \tag{6.38}$$

remember that Z_1, Z_2 are correlated. It is obvious that dZ_2 is a Brownian motion, and to see that dZ_1 is Brownian motion, note that

$$\mathbb{E}[dZ_1] = 0,\ \mathrm{Var}[dZ_1] = \mathrm{Var}\left[\sqrt{1-\rho^2} dW_1 + \rho dW_2\right] = Var[dW] \tag{6.39}$$

and the correlation, $< dW_1, dW_2 >= 0$ So we can write Heston's model in terms of independent Brownian motions

$$\begin{aligned} dS &= \mu S dt + \sqrt{v} S \left(\sqrt{1-\rho^2} dW_1 + \rho dW_2\right) \\ dv &= \kappa(\theta - v)dt + \sigma\sqrt{v} dW_2 \end{aligned} \tag{6.40}$$

rearrange the stock SDE

$$\frac{dS}{S} = \mu dt + \sqrt{1-\rho^2}\sqrt{v} dW_1 + \rho\sqrt{v} dW_2 \tag{6.41}$$

In the Black Scholes -setting, the dynamics of the stock is

$$\frac{dS}{S} = \mu dt + \sigma dW \tag{6.42}$$

with a constant σ, under the physical measure, and under the risk - neutral measure

$$\frac{dS}{S} = r dt + \sigma dW^{\mathbb{Q}} \tag{6.43}$$

and the Radon - Nikodym, density process, or the Girsanov theorem gives the link between the physical and the risk-neutral measure

$$dW^{\mathbb{Q}} = \frac{\mu - r}{\sigma} dt + dW \tag{6.44}$$

that under the risk neutral measure can be written as

$$\begin{aligned} \frac{dS}{S} &= \mu dt + \sigma \left(dW^{\mathbb{Q}} - \frac{\mu - r}{\sigma} dt \right) \\ &= \mu dt + \sigma dW^{\mathbb{Q}} - \mu dt + r dt \\ &= \mu dt + \sigma dW^{\mathbb{Q}} \\ &= (\mu - \lambda\sigma) dt + \sigma dW^{\mathbb{Q}} \end{aligned} \tag{6.45}$$

In this specification, $(\mu - \lambda\sigma) = r$ and the ratio $\frac{\mu - r}{\sigma} = \lambda$ is the market price of risk. this λ is used when markets are not complete, or the assets are not traded, which is the case for the variance process. Our risk neutral measure Q depends on λ and for different λ you get different risk neutral measures. For stocks, which is a traded asset, we can write the relationship using

$$dS = \mu S dt + \sqrt{v} S \left(\sqrt{1 - \rho^2} dW_1 + \rho dW_2 \right) \tag{6.46}$$

where the stock price SDE is the second Brownian, the uncorrelated, which is not traded.

$$\begin{aligned} dW_1^{\mathbb{Q}_\lambda} &= dW_1 + \frac{\mu - r - \lambda\rho\sqrt{v}}{\sqrt{1 - \rho^2}\sqrt{v}} \\ dW_2^{\mathbb{Q}_\lambda} &= dW_2 + \lambda dt \end{aligned} \tag{6.47}$$

then the drift of the stock will be equal to r, the risk free rate. into (6.45) we get

$$\frac{dS}{S} = \mu dt + \sqrt{1 - \rho^2}\sqrt{v} \left[dW_1^{\mathbb{Q}_\lambda} - \frac{\mu - r - \lambda\rho\sqrt{v}}{\sqrt{1 - \rho^2}\sqrt{v}} dt \right] + \rho\sqrt{v}(dW_2^{\mathbb{Q}_\lambda} - \lambda dt) \tag{6.48}$$

combining the dt terms

$$\begin{aligned} \frac{dS}{S} &= \left(\mu - \mu + r + \lambda\rho\sqrt{v} - \lambda\rho\sqrt{v} \right) dt + \sqrt{1 - \rho^2}\sqrt{v} dW_1^{\mathbb{Q}_\lambda} + \rho\sqrt{v} dW_2^{\mathbb{Q}_\lambda} \\ \frac{dS}{S} &= r dt + \sqrt{1 - \rho^2}\sqrt{v} dW_1^{\mathbb{Q}_\lambda} + \rho\sqrt{v} dW_2^{\mathbb{Q}_\lambda} \end{aligned} \tag{6.49}$$

So the drift is r. For the variance SDE, insert dW_2 see (6.40)

$$\begin{aligned} dv &= \kappa(\theta - v) dt + \sigma\sqrt{v} dB_2 = \\ &= \kappa(\theta - v) dt + \sigma\sqrt{v} \left(dW_2^{\mathbb{Q}_\lambda} - \lambda dt \right) \\ &= \left(\kappa(\theta - v) - \lambda\sigma\sqrt{v} \right) dt + \sigma\sqrt{v} dW_2^{\mathbb{Q}_\lambda} \end{aligned} \tag{6.50}$$

going back to Z_1, and Z_2, that was a linear combination of dW_1 and dW_2, we get the system in the original form

$$\begin{aligned}\frac{dS}{S} &= rdt + \sqrt{v}dZ_1^{\mathbb{Q}_\lambda} \\ dv &= \left(\kappa(\theta - v) - \lambda\sigma\sqrt{v}\right)dt + \sigma + \sqrt{v}dZ_2^{\mathbb{Q}_\lambda}\end{aligned} \tag{6.51}$$

conditioning on λ we are in the risk - neutral world and we can use the theory of Martingale, as we did in the Black - Scholes setting, the stock price process is a Martingale under the risk neutral measure, than we can write the price of the option as the discounted expected value. That can be done as follows

$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}_\lambda}\left[h(S_T)|S_0, v_0\right] \tag{6.52}$$

and the PDE will be given by the 2-dimensional version of the Feynman-Kac theorem. The $h(S_T, *)$ is the payoff function.

$$\begin{aligned}0 = \frac{\partial V}{\partial t} &+ \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV \\ &+ \frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2} + \left(\kappa(\theta - v) - \lambda\sigma\sqrt{v}\right)\frac{\partial V}{\partial v} + \rho\sigma vS\frac{\partial V^2}{\partial v\partial S}\end{aligned} \tag{6.53}$$

where the second line corresponds to the second dimension in the Feynman - Kac, and the last term is the cross-term. Rearranging the terms gives us

$$\begin{aligned}0 = \frac{\partial V}{\partial t} + \frac{1}{2}\left(vS^2\frac{\partial^2 V}{\partial S^2} + \sigma^2 v\frac{\partial^2 V}{\partial v^2} + 2\rho\sigma vS\frac{\partial V^2}{\partial v\partial S}\right) + \\ rS\frac{\partial V}{\partial S} + \left(\kappa(\theta - v) - \lambda\sigma\sqrt{v}\right)\frac{\partial V}{\partial v} - rV\end{aligned} \tag{6.54}$$

and the Feynman - Kac 2-dimensional PDE is

$$0 = \frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i,j}\sigma_{i,j}\frac{\partial^2 V}{\partial x_i \partial x_j} + \sum_i \mu_i^{\mathbb{Q}_\lambda}\frac{\partial V}{\partial x_i} - rV \tag{6.55}$$

where the $\{x_i\}$ are the underlying processes, in Heston we have two processes, the stock price and the variance SDE. $\sigma_{11} = vS^2$, the square root of the variance of the stock price SDE, $\sigma_{22} = v\sigma^2$ the square root of the variance in the variance SDE, $\sigma_{12} = \sigma_{21} = \rho\sigma vS$ is the correlation between the two Brownian motions. The drift of the stock price SDE, is $\mu_1 = rS$ and $\mu_2 = \kappa(\theta - v) - \lambda\sigma\sqrt{v}$ is the drift for the variance SDE.

6.1.3 European Call option Price, in Heston model

The derivation of Heston's pricing PDE, is similar to the Black-Scholes derivation, some would argue easier, if you are not afraid of complex numbers, but as we will see, in the solution, we will only use the real part, the imaginary part is just used in the derivation. I reproduce the Heston model. He assumed the following dynamics for the stock price, where the variance itself is a random process

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dZ_1 \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dZ_2 \end{aligned} \tag{6.56}$$

and in the risk - neutral setting, under the measure $\mathbb{Q}$ we get

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{v_t} S_t dZ_1^{\mathbb{Q}_\lambda} \\ dv_t &= \left(\kappa(\theta - v_t) - \lambda\sigma\sqrt{v_t}\right) dt + \sigma\sqrt{v_t} dZ_2^{\mathbb{Q}_\lambda} \end{aligned} \tag{6.57}$$

where the drift in the variance process gets adjusted by the market price of risk. Heston[3] made the variance process simpler by

$$dv_t = \left(\kappa(\theta - v_t) - \lambda v_t\right) dt + \sigma\sqrt{v_t} dZ_2^{\mathbb{Q}_\lambda} \tag{6.58}$$

Let us write the variance process a bit shorter to save space.

$$dv_t = \mu_v dt + \sigma\sqrt{v_t} dZ_2^{\mathbb{Q}_\lambda} \tag{6.59}$$

In the last chapter we showed that using Delta and Sigma hedging, we could solve this pricing PDE, which I reproduce here, it is the 2 dimensional version of the Black Scholes equation.

$$\begin{aligned} 0 = \frac{\partial V}{\partial t} + \frac{1}{2}\left(vS^2\frac{\partial^2 V}{\partial S^2} + \sigma^2 v\frac{\partial^2 V}{\partial v^2} + 2\rho\sigma v S\frac{\partial V^2}{\partial v \partial S}\right) + \\ rS\frac{\partial V}{\partial S} + \left(\kappa(\theta - v) - \lambda\sigma\sqrt{v}\right)\frac{\partial V}{\partial v} - rV \end{aligned} \tag{6.60}$$

The price of the derivative can be written as the expected value of the discounted terminal payoff

$$V_0 = \mathbb{E}^{\mathbb{Q}_\lambda}\left[e^{-rT}h(S_T)|S_0, v_0\right] \tag{6.61}$$

where $h(S_T)$ is the payoff function, so for an European call option it would $h(S_T) = \max(S_T - K, 0)$. Not that we are not using filtration, but conditional on the initial values, this is due to Markov properties. The Feynman Kac presents the link between the two representation (6.60) and (6.61) As in the Black Scholes model, it gets simpler if we make a transformation of the stock price to the log price of the stock price. $x = \log(S)$, apply Ito's lemma to the differential on both sides we get

$$dx = d\log(S) = \left(r - \frac{1}{2}v\right)dt + \sqrt{v_t} dZ_1^{\mathbb{Q}_\lambda} \tag{6.62}$$

Let us also transform the PDE (6.60) from being a function of S to being a function of x

$$0 = \frac{\partial V}{\partial t} + \frac{1}{2}v\frac{\partial^2 V}{\partial x^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2} + 2\rho\sigma v\frac{\partial V^2}{\partial v \partial x} + \left(r - \frac{1}{2}v\right)\frac{\partial V}{\partial x} + \mu_v)\frac{\partial V}{\partial v} - rV \tag{6.63}$$

[3](Breeden, 1979) Breeden states how to find a risk neutral measure for the market price of risk and the CAPM - model

For the BS case we could use the *Heat equation* to get a closed form solution, that can't be done in the Heston model. There are 3 variables in the Heston model, t, x, v, but the procedure to derive the pricing formula are step wise similar. The valuation formula givess us the following

$$V_0 = \mathbb{E}^{\mathbb{Q}_\lambda}\left[e^{-rT}h(S_T)|S_0, v_0\right] \tag{6.64}$$

The payoff for a European call option can be written as $h(S_T) = \max(S_T - K, 0) = S_T \mathbb{1}_{S_T>K} - K\mathbb{1}_{S_T>K}$, now we can split the payoff into two terms, due to the linearity of expectation. Let us go back to (6.61)

$$\begin{aligned} V_0 &= \mathbb{E}^{\mathbb{Q}_\lambda}\left[e^{-rT}\left(S_T\mathbb{1}_{S_T>K} - K\mathbb{1}_{S_T>K}\right)\right] \\ V_0 &= \mathbb{E}^{\mathbb{Q}_\lambda}\left[e^{-rT}S_T\mathbb{1}_{S_T>K}\right] - Ke^{-rT}\mathbb{E}^{\mathbb{Q}_\lambda}\left[\mathbb{1}_{S_T>K}\right] \end{aligned} \tag{6.65}$$

The second term is similar to Black- Scholes, but the first term is a bit more complicated, a change of numeraire is needed, the change of numeraire is the same as in the Black - Scholes example, where the first term were in the stock measure, and the second term were in the bank account numeraire.

$$\frac{V_0}{B_0} = \mathbb{E}^{\mathbb{Q}}\left[\frac{V_T}{B_T}\middle|\mathcal{F}_0\right] \quad \frac{V_0}{S_0} = \mathbb{E}^{\mathbb{S}}\left[\frac{V_T}{S_T}\middle|\mathcal{F}_0\right] \tag{6.66}$$

The value of an asset scaled by the value of the stock price, will be a martingale, under the measure induced by the stock price as the numeraire. Notice that S_0 is known at time, it is a constant

$$V_0 = \mathbb{E}^{\mathbb{Q}}\left[\frac{B_0}{B_T}V_T\middle|\mathcal{F}_0\right] \quad V_0 = \mathbb{E}^{\mathbb{S}}\left[\frac{S_0}{S_T}V_T\middle|\mathcal{F}_0\right] \tag{6.67}$$

as (6.67) and (6.67) is the price for the same asset, we come to

$$\frac{B_0}{B_T}d\mathbb{Q} = \frac{S_0}{S_T}dP^{\mathbb{S}} \quad S_T\frac{B_0}{B_T}d\mathbb{Q} = S_0 dP^{\mathbb{S}} \tag{6.68}$$

and since the bank account starts with 1 we get

$$S_T e^{-rT}d\mathbb{Q} = S_0 d\mathbb{P}^{\mathbb{S}} \tag{6.69}$$

inserting this in (6.65) we get

$$\begin{aligned} V_0 &= \mathbb{E}^{\mathbb{Q}}\left[S_T\mathbb{1}_{S_T>K}\right] - Ke^{-rT}\mathbb{E}^{\mathbb{Q}_\lambda}\left[\mathbb{1}_{S_T>K}\right] = \\ V_0 &= \mathbb{E}^{\mathbb{S}}\left[S_0\mathbb{1}_{S_T>K}\right] - Ke^{-rT}\mathbb{E}^{\mathbb{Q}_\lambda}\left[\mathbb{1}_{S_T>K}\right] \\ V_0 &= S_0P_1 - Ke^{-rT}P_2 \end{aligned} \tag{6.70}$$

P_1 is the probability that the stock price is greater than K, the strike price, under the stock measure. P_2 is the probability that the stock price is greater than K but under the risk-neutral measure. You can choose a time $\tau \in [0, T]$ to denote the price of the option V_τ with remaining time, $\tau = T - t$. Use the chain rule from regular calculus, notice that T is fixed, so there is the same derivative but with opposite signs, the chain rule produces an -1, so there will be a negative sign if working with τ instead of t. Our PDE is in x that is $x = \log(S)$ so our pricing function (6.65) will become

$$V_\tau = e^x P_1 - Ke^{-r\tau} P_2 \tag{6.71}$$

This is the price of the option with remaining maturity equal to τ

$$0 = \frac{\partial V}{\partial t} + \frac{1}{2} v \frac{\partial^2 V}{\partial x^2} + \frac{1}{2}\sigma^2 v \frac{\partial^2 V}{\partial v^2} + 2\rho\sigma v \frac{\partial V^2}{\partial v \partial x} + \left(r - \frac{1}{2}v\right) \frac{\partial V}{\partial x} + \mu_v \frac{\partial V}{\partial v} - rV \tag{6.72}$$

becomes with change of variables from t to τ

$$\frac{\partial V}{\partial \tau} = \frac{1}{2} v \frac{\partial^2 V}{\partial x^2} + \frac{1}{2}\sigma^2 v \frac{\partial^2 V}{\partial v^2} + 2\rho\sigma v \frac{\partial V^2}{\partial v \partial x} + \left(r - \frac{1}{2}v\right) \frac{\partial V}{\partial x} + \mu_v \frac{\partial V}{\partial v} - rV \tag{6.73}$$

Duffie showed that the stock can be written as

$$S_t = e^{a(t)+b(t).x} \tag{6.74}$$

so we need to replace in our price formula

$$V_t = S_0 P_1 - Ke^{-r\tau} P_2 \text{ replace } S_0 = e^x \rightarrow V_t = e^x P_1 - Ke^{-r\tau} P_2 \tag{6.75}$$

it relies on the fact that stock price can be written as

$$S_t = e^{a(t)+b(t).x} \tag{6.76}$$

where the x denotes a vector of two factors of $a(t)$ and $b(t)$ We now have the PDE, (6.73), and we have that is a solution (6.73), satisfies the PDE, but since V_τ is a linear combination, also note that the prices are linear combination of the two terms. of P_1 and P_2 both P_1 and P_2 must solve (6.73) by themselves. I carry out the calculations. Let $V_1 = e^x P_1$, and solve (6.73), first we must calculate the derivatives, put the results in the PDE and simplify.

Derivatives in the solution of Heston

V_1	P_1	V_2	P_2
$\frac{\partial V_1}{\partial \tau}$	$e^x \frac{\partial P_1}{\partial \tau}$	$\frac{\partial V_2}{\partial \tau}$	$-re^{-r\tau} P_2 + e^{-r\tau} \frac{\partial P_2}{\partial \tau}$
$\frac{\partial V_1}{\partial x}$	$e^x P_1 + e^x \frac{\partial P_1}{\partial x}$	$\frac{\partial V_2}{\partial x}$	$e^{-r\tau} \frac{\partial P_2}{\partial x}$
$\frac{\partial^2 V_1}{\partial x^2}$	$e^x P_1 + 2e^x \frac{\partial P_1}{\partial x} + e^x \frac{\partial^2}{\partial x^2}$	$\frac{\partial V_2}{\partial x^2}$	$e^{-r\tau} \frac{\partial^2 P_2}{\partial x^2}$
$\frac{\partial V_1}{\partial v}$	$e^x \frac{\partial P_1}{\partial v}$	$\frac{\partial V_2}{\partial v}$	$e^{-r\tau} \frac{\partial P_2}{\partial v}$
$\frac{\partial^2 V_1}{\partial v^2}$	$e^x \frac{\partial^2 P_1}{\partial x^2}$	$\frac{\partial^2 V_2}{\partial v^2}$	$e^{-r\tau} \frac{\partial^2 P_2}{\partial v^2}$
$\frac{\partial^2 V_1}{\partial v \partial x}$	$e^x \frac{\partial P_1}{\partial v} + e^x \frac{\partial^2 P_1}{\partial x \partial v}$	$\frac{\partial^2 V_2}{\partial v \partial x}$	$e^{-r\tau} \frac{\partial P_2}{\partial v} + e^x \frac{\partial^2 P_2}{\partial x \partial v}$

Then we insert the substitution into (6.73), and we get, notice that e^x appears in all terms for $V_1 = e^x P_1$ so it cancels. The second term, $V_2 = e^{-r\tau} P_2$, and notice that $e^{-r\tau}$ appears in all terms in the second line, so we can cancel it.

$$\frac{\partial P_1}{\partial \tau} = \frac{1}{2}v\frac{\partial^2 P_1}{\partial x^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_1}{\partial v^2} + \rho\sigma v\frac{\partial P_1^2}{\partial v \partial x} + \left(r + \frac{1}{2}v\right)\frac{\partial P_1}{\partial x} + (\mu_v + \rho\sigma v)\frac{\partial P_1}{\partial v}$$
$$\frac{\partial P_2}{\partial \tau} = \frac{1}{2}v\frac{\partial^2 P_2}{\partial x^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_2}{\partial v^2} + \rho\sigma v\frac{\partial P_2^2}{\partial v \partial x} + \left(r - \frac{1}{2}v\right)\frac{\partial P_2}{\partial x} + \mu_v\frac{\partial P_2}{\partial v} \tag{6.77}$$

These two lines look very similar, the only difference is in coefficients of $\frac{\partial P_1}{\partial x}$ and $\frac{\partial P_1}{\partial v}$ $\frac{\partial P_1}{\partial x}$ and $\frac{\partial P_2}{\partial x}$, looks like that geometric Brownian motion under the risk - neutral measure and under the stock measure. The drift will be r under the risk neutral measure, which in the logarithm becomes $r - \frac{1}{2}\sigma^2$. In the Heston model the stock price diffusion term is $\sqrt{v}$ which in the log form will give us a drift of $r - \frac{1}{2}v$ which the coefficient in P_2 equation for the stock price under the risk-neutral measure. $\frac{\partial P_2}{\partial x}$ The drift term for gBm under the stock measure will be $dS = (r + \sigma^2)Sdt + \sigma S dW^{\mathbb{S}}$, in Heston, it means that we add v for the drift under P_1, which is the coefficient for $\frac{\partial P_1}{\partial x}$ It also explains the extra term in the coefficient for $\frac{\partial P_1}{\partial v}$, because ρ is the correlation of stock price SDE and dZ_2 $\rho\sigma v$ is as the covariance.

As the two PDE are similar, we can write them as one PDE with a level j, let $u_1 = 0.5, u_2 = -0.5$, and $b_1 = \kappa + \lambda - \rho\sigma$ and $b_2 = \kappa + \lambda$ then (6.77) can be written as a generic PDE

$$\frac{\partial P_j}{\partial \tau} = \frac{1}{2}v\frac{\partial^2 P_j}{\partial x^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_j}{\partial v^2} + \rho\sigma v\frac{\partial P_j^2}{\partial v \partial x} + \left(r + \frac{1}{2}u_j v\right)\frac{\partial P_j}{\partial x} + (a - b_j v)\frac{\partial P_j}{\partial v} \tag{6.78}$$

To solve numerically (6.78) subject to its terminal condition, which has become the initial condition, because $\tau = T - t$, and remember that the value $V_t = e^x P_1 - Ke^{-r\tau}P_2$, and that P_1 is equal to the probability that the stock price is greater or equal to strike K at maturity, under the different measures. So the initial condition is $P_0 = \mathbb{1}_{S>K}$

Characteristic function to solve Heston PDE

It should not be hard to solve (6.78), using numerical methods, but let us try to find an analytic solution, knowing that separation of variables is not a possible solution for this model, and we turn to the characteristic function method. The Feynman-Kac's gives us the expectation form us this problem.

$$P_j = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{X_T > \log K} \middle| S_0, v_0\right] \tag{6.79}$$

where the function value at maturity is in term of the indicator function, The characteristic function is

$$f_j = \mathbb{E}^{\mathbb{Q}}\left[e^{i\phi X_T} \middle| S_0, v_0\right] \tag{6.80}$$

and it must satisfy the same PDE (6.78), only the terminal condition is changing, from indicator function, to exponential, the remember that the indicator function, is the same as the probability to the indicated event, under some measure.

$$\frac{\partial f_j}{\partial \tau} = \frac{1}{2}v\frac{\partial^2 f_j}{\partial x^2}+\frac{1}{2}\sigma^2 v\frac{\partial^2 f_j}{\partial v^2}+\rho\sigma v\frac{\partial f_j^2}{\partial v \partial x}+\left(r+\frac{1}{2}u_j v\right)\frac{\partial f_j}{\partial x}+(a-b_j v)\frac{\partial f_j}{\partial v} \quad (6.81)$$

If X is a normal distributed r.v. $X \sim N(m, \sigma^2)$ then its characteristic function will be

$$\begin{aligned} \mathbb{E}[e^{i\phi x}] &= \int_{-\infty}^{\infty} e^{i\phi x} p(x)\, dx \\ \mathbb{E}[e^{i\phi x}] &= e^{i\phi m + \frac{1}{2}(i\phi)^2\sigma^2} \\ \mathbb{E}[e^{i\phi x}] &= e^{um+\frac{1}{2}(u)^2\sigma^2} \end{aligned} \quad (6.82)$$

where $p(x)$ is the probability function, in this case for the normal density, and let $u = i\phi$ be a complex number. For a geometric Brownian motion, where the stock price has the following dynamics, under the risk neutral measure

$$dS = rS\, dt + \sigma S\, dW_t \quad (6.83)$$

and if we take the logarithm of the stock price dynamics, it will be normally distributed

$$\log(S_T) \sim N\left[x + \left(r - \frac{1}{2}\sigma^2\right)\tau, \sigma^2\tau\right] \quad (6.84)$$

which can be viewed as a marginal distribution, and its characteristic function will be note that $x = \log(S_0)$, τ is the time to maturity, we are using filtration now, and the subscript x in X_t^x represents the value of a Markov process X, at time t, where it started at small x at time 0. . ψ and ϕ are some generic functions of τ, u. An affine function is function of this form

$$f(x) = a + bx \quad (6.85)$$

so it is a linear transformation plus translation.

$$\mathbb{E}\left[e^{uX_T^x}\middle|\mathcal{F}_t\right] = e^{ux+u(r-\frac{1}{2}\sigma^2)\tau+\frac{1}{2}u^2\sigma^2\tau} \quad (6.86)$$

in a multi-dimensional setting it would be

$$\mathbb{E}\left[e^{u.X_T^x}\middle|\mathcal{F}_t\right] = e^{\psi(\tau,u).x+\phi(\tau,u)} \quad (6.87)$$

this is also called the affine, i.e. linear transformation plus translation exponential. If $\{X_t\}$ is a stochastic process with a dynamics given by

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \quad (6.88)$$

then X_t is called an affine process if the drift and diffusion are affine functions and the drift and diffusion are function that can be written as follows, note that it is drift and variance and not drift and volatility

$$\mu(x) = c_0 + c_1 x, \quad \sigma^2(x) = k_0 + k_1 x \tag{6.89}$$

Then we characteristic function will be in the *affine*, the market price of risk, and the square root in the diffusion model of Heston's model are examples of affine exponential form. In Heston's model we have that the log of the stock price is given by the following SDE

$$dx = (r - 0.5v)dt + \sqrt{v}dZ_1 \tag{6.90}$$

and the variance SDE is given by

$$dv = (\kappa\theta - \kappa v - \lambda v)dt + \sigma\sqrt{v}dZ_2 \tag{6.91}$$

which is affine. Heston wrote the characteristic function as

$$f(x, v, \tau) = e^{C(\tau)+D(\tau)v+i\phi x} \tag{6.92}$$

we need to calculate the derivative of (6.78) in terms of (6.92)

$$\begin{aligned} &\frac{\partial f}{\partial \tau} = \left(\frac{\partial C}{\partial \tau} + v\frac{\partial D}{\partial \tau}\right) f, \quad \frac{\partial f}{\partial x} = i\phi f \\ &\frac{\partial^2 f}{\partial x^2} = -\phi^2 f, \quad \frac{\partial f}{\partial v} = Df \\ &\frac{\partial^2 f}{\partial v^2} = D^2 f, \quad \frac{\partial^2 f}{\partial v \partial x} = i\phi D f \end{aligned} \tag{6.93}$$

and we put them into (6.78) we get

$$\left(\frac{\partial C}{\partial \tau} + v\frac{\partial D}{\partial \tau}\right) f = -\frac{1}{2}v\phi^2 f + \frac{1}{2}\sigma^2 v D^2 f + \rho\sigma i\phi D f + (r + u_j v)i\phi f + (a - b_j v)Df \tag{6.94}$$

and collect terms we get and notice that f is in all terms and we can cancel f

$$\left(-\frac{\partial D}{\partial \tau} - \frac{1}{2}\phi^2 + \frac{1}{2}\sigma^2 D^2 + \rho\sigma i\phi D + u_j i\phi - b_j D\right) v - \frac{\partial C}{\partial \tau} + ri\phi + aD = 0 \tag{6.95}$$

for an affine function to be zero, the coefficients before v must be zero, and the sum of the other terms must also be zero.

$$-\frac{\partial D}{\partial \tau} - \frac{1}{2}\phi^2 + \frac{1}{2}\sigma^2 D^2 + \rho\sigma i\phi D + u_j i\phi - b_j D = 0 \tag{6.96}$$

and that

$$-\frac{\partial C}{\partial \tau} + ri\phi + aD = 0 \tag{6.97}$$

This is the Ricatti equation system, begin with (6.81), put the time-derivative on the l.h.s.

$$\frac{\partial D}{\partial \tau} = \left(i\mu_j\phi - \frac{1}{2}\phi^2\right) + (i\rho\sigma i\phi - b_j)D + +\frac{1}{2}\sigma^2 D^2 \tag{6.98}$$

and moving the time derivative in (6.98) gives

$$\frac{\partial C}{\partial \tau} = +ri\phi + aD \tag{6.99}$$

if $\tau = 0$, then our function $f(x, v, 0) = e^{i\phi x}$, then it follows, that we have the initial conditions

$$D(0, \phi) = 0, \quad C(0, \phi) = 0 \tag{6.100}$$

The general form for a Riccati equation is

$$\frac{dy}{d\tau} = a + by + cy^2 \tag{6.101}$$

Start with solving (6.97), we can identify the coefficient in the Riccati equation

$$a = i\mu_j\phi - \frac{1}{2}\phi^2, \quad b = i\rho\sigma i\phi - b_j, \quad c = \frac{1}{2}\sigma^2 \tag{6.102}$$

The solution can be written as, given that we have a initial equation $y(0) = 0$ and $d = \pm\sqrt{b^2 - 4ac}$, I will only use the solution with a plus[4]. Use the transformation to the second order linear equation and using the characteristic equation to solve it, given the condition $y(0) = 0$

$$\begin{aligned} y &= -\frac{1}{2c}\frac{-(b-d)e^{d\tau} + (b-d)}{-\frac{b-d}{b+d}e^{d\tau} + 1} \\ D(\tau) &= \frac{d-b}{2c}\frac{1 - e^{d\tau}}{1 - \frac{b-d}{b+d}e^{d\tau}} \\ d_j &= \sqrt{(i\rho\sigma\phi)^2 - (2i\mu_j\phi - \phi^2)\sigma^2} \\ g_j &= \frac{b_j - i\rho\sigma\phi + d_j}{b_j - i\rho\sigma\phi - d_j} \\ D(\tau) &= \frac{d_j + b_j - i\rho\sigma\phi}{\sigma^2}\frac{1 - e^{d_j\tau}}{1 - g_j e^{d_j\tau}} \end{aligned} \tag{6.103}$$

Solving for $C(0, \phi) = 0$ in the (6.99)

$$C(\tau) = ir\phi\tau + \frac{a}{\sigma^2}\left(-2\log\left(\frac{1 - g_j e^{d_j\tau}}{1 - g_j}\right) + (d_j + b_j - i\rho\sigma\phi)\tau\right) \tag{6.104}$$

[4] The solution with a minus sign is called the Heston trap

So the characteristic function $C(\tau)$ and $D(\tau)$ are the two that correspond to P_1 and P_2, that appear in the European option price in the Heston model. We have deduced the solution of the characteristic function

$$f(x, v, \tau) = e^{C(\tau)+D(\tau)t+i\phi x} \tag{6.105}$$

Remember that we need the option pricing formula to be able to price

$$V_\tau = e^x P_1 - Ke^{r\tau} P_2 \tag{6.106}$$

We also need the Lévy inversion formula, to transform the characteristic function into probabilities. One version of it is here

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathcal{R} \left[\frac{e^{-i\phi \log K} f_j}{i\phi} \right] d\phi \tag{6.107}$$

To calculate the price of the option, we first need to find the characteristic functions, we can thereafter determine the probability by some numerical integration methods and when we have the probabilities, we have the price of the option.

6.1.4 Bibliographical notes

The starting point in this chapter is (Heston, 1993). There are many preinstalled package to solve numerical this problem, in R there is the package NMOF, (Gilli, Maringer, & Schumann, 2019). The reason for making the the asset prices in the log scale, is due to (Duffie et al., 1999), There is also (Breeden, 1979), that is about the market price of risk.

Chapter 7

Forward start options

7.1 Forward start option

7.1.1 Rubinstein

A starting point for Forward start options is Rubinstein's article from 1990[1] Rubinstein's setup. You have an underlying asset, how much would you be will to pay today, time 0 for an asset that becomes valid at a time t, the grant date, this date is determined and has no randomness, and has a maturity at T. The strike is set to be ATM. Let it behave as an European call option. Rubinstein's make four assumptions, that are quite general, and is applicable to all types of options that I am looking at in this thesis.

- homogeneity, the call option value, when it is granted, will be homogeneous of degree one in the underlying asset price and the strike price
- state variable, all uncertainty in valuing the option after time t is resolved once the underlying asset price after time t is known
- data-invariance, the variables determining the value of the option are not date-dependent
- payout, the underlying asset through the grant date has a known constant payout rate d

Furthermore, let

- $S \equiv$ current value of the underlying
- $S_t \equiv$ the (random) value of the underlying after time t, the grant date.

[1] "Pay Now, Choose Later", RISK 4 (Februry 1991), p.13 Rubinstein, Mark, found on the internet as "Forward-Start Option", https://ramurapt.files.wordpress.com/2009/10/forwardstartoptions.doc, retr. 2021-03-15

- $C(X, Y, T-t)$ value of a call option with X as the underlying, Y as the strike price and $T-t$ remaining time to maturity.

The value of a forward starting at the money call option is, on the grant date, t, by using the homogeneity assumption

$$C(S_t, S_t, T-t) = S_t C(1, 1, T-t) \tag{7.1}$$

Since all randomness comes from S_t, the second factor is non-random, $C(1, 1, T-t)$.

By using the replicating portfolio assumption, if we can make an investment now, that will for sure produce the same outcome at time t, $S_t C(1, 1, T-t)$, then the current cost of the investment must equal the value of the forward start option. Let $C(1, 1, T-t)$ be the number of shares, to replicate the value of the option after time t, we need to hold $C(1, 1, T-t)$, correcting for the dividends until time t

$$S d^{-1} C(1, 1, T-t) \tag{7.2}$$

is the current value of the forward-start option. Using homogeneity it can be written as

$$d^{-1} C(S, S, T-t) \tag{7.3}$$

Thus *"the value a forward-start option is simply the current value of d^{-1} calls which are currently at-the-money, with time to expiration $T-t$*

We can split the time until maturity, into two parts, one part is the current time until the grant time, $0-t$, during that period we need to hold $C(1, 1, T-t)$ shares of the stock.

7.2 Background

Here I follow (Musiela & Rutkowski, 2005), to write the theoretical background of the problem. The value of a forward start option changes with volatility. In Black - Scholes setting, with constant and deterministic σ, volatility, the Forward- start option becomes very simple. In their notation, the payoff will be

$$FS_T = (S_T - KS_{T_0})^+ \tag{7.4}$$

At the grant date, T_0 it becomes

$$FS_{T_0} = C_{T_0}(S_{T_0}, T-t, KS_{T_0}) \tag{7.5}$$

that is a European Call option, with starting point at time T_0. Its price is

$$C_{T_0}(S_{T_0}, T-T_0, KS_{T_0}) = S_{T_0} c(1, T-T_0, K, r, \sigma) \tag{7.6}$$

where everything in the parentheses in the right hand side of the above equation are deterministic.

7.2.1 Deterministic volatility

Were we to expand the classic Black Scholes model with deterministic volatility $\sigma(t)$ (Musiela & Rutkowski, 2005). The volatility will be different at different times. It will give a flat smile in the implied volatility surface. Let the maturity T be known, as is the case for a forward - start option, then the mapping $K \mapsto \widehat{\sigma}_0(T,K)$, is the implied volatility curve for the maturity date T. The market-based Black - Scholes implied volatility surface $\widehat{\sigma}_0(T,K)$ is thus implicitly defined:

$$C_0^m(T,K) = c(S_0, T, K, r, \widehat{\sigma}_0(T,K) \tag{7.7}$$

where $c(S_0, T, K, r, \sigma)$ is the Black - Scholes price of a call option. Let $C_0^M(T,K)$ be a family of market prices of European call options with all strikes $K > 0$, and all maturities $0 < T < T^*$ for some $T^* > 0$. I will treat the parameter r as constant in this thesis. Assume that the implied volatility $\widehat{\sigma}_0(T,K)$ inferred from the call option prices is flat in K, for each the maturity date T, the implied volatility does not depend on strike K, in this case $\widehat{\sigma} : (0, T^*) \mapsto \mathbb{R}^+$. To match market data, we only need an extension to Black - Scholes model, assume time dependent volatility function $\widehat{\sigma} : \mathbb{R}^+ \mapsto \mathbb{R}^+$. The extension to Black - Scholes would be driven by the following SDE, under the risk - neutral measure $\mathbb{Q}$

$$dS(t) = rS(t)dt + \widehat{\sigma}(t)dW^{\mathbb{Q}}(t) \tag{7.8}$$

and the volatility function satisfies

$$\widehat{\sigma}_0^2(T,K) = \frac{1}{T}\int_0^T \widehat{\sigma}^2(u)\, du \tag{7.9}$$

We will have a flat smile in the implied volatility surface. Going back to the forward - start option, if we assume a flat implied volatility surface we have

$$C_{T_0} = (S_{T_0}, T - T_0, KS_{T_0}) = S_{T_0} c(1, T - T_0, K, r, \sigma(T_0, T)) \tag{7.10}$$

where the average future volatility is

$$\widehat{\sigma}^2(T_0, T) = \frac{1}{T - T_0}\int_{T_0}^T \sigma^2(t)\, dt \tag{7.11}$$

and thus

$$FS_0 = S_0\, c(1, T - T_0, K, r, \sigma(T_0, T)) = c(S_0, T - T_0, KS_0, r, \sigma(T_0, T)) \tag{7.12}$$

Since the forward start option start its life at T_0, before that time, that time it behaves as process growing at a risk free rate, when the option becomes active it will have a volatility, that is difference from the frozen asset at S_{T_0} and the call option future variance between T_0 and T, for the underlying asset S_t. The forward implied volatility is

$$\widehat{\sigma}^2(T_0, T) = \frac{T\widehat{\sigma}_0^2(T,K) - T_0\widehat{\sigma}_0^2(T_0,K)}{T - T_0} \tag{7.13}$$

note that the implied volatility is independent of K

7.2.2 Case for random volatility, Musiela Rudkowski

If there is a volatility smile, we can no longer derive uniquely the forward volatility from the implied volatility surface. To deal with this case, we make the assumption that S satisfies

$$dS_t = S_t(r\,dt + \sigma_t)dW_t^* \tag{7.14}$$

for some stochastic volatility process σ. Suppose that the volatility process σ is given, and we want to find a closed - form expression solution for the price $C(S_{T_0}, T - T_0, KS_{T_0})$, to find a forward start option for the time $t \in [0, T_0]$, we need to compute

$$\boldsymbol{FS}_t = e^{-r(T_0-t)}\mathbb{E}^{\mathbb{Q}}\left[C(S_{T_0}, T - T_0, KS_{T_0})|\mathcal{F}_t\right] \tag{7.15}$$

and it is difficult. A help can be the terminal condition, as it is an European option, the terminal payoff is

$$\boldsymbol{FS}_T = (S_T - KS_{T_0})^+ = S_{T_0}\,(Y - K)^+ = \hat{S}_T\,(Y - K)^+ \tag{7.16}$$

where $\hat{S}$ is given by $\hat{S}_T = S_{t\wedge T_0}$ for every $t \in [0, T]$ and where $Y = \frac{S_T}{S_{T_0}} = \frac{S_T}{\hat{S}_T}$, define an measure $\hat{\mathbb{Q}}$ equivalent to $\mathbb{Q}$, by

$$\eta_T = \frac{d\hat{\mathbb{Q}}}{d\mathbb{Q}} = a\frac{\hat{S}_T B_0}{\hat{S}_0 B_{T_0}} = \frac{\hat{S}_{T_0} B_0}{S_0 B_{T_0}} \text{ where } a = e^{-r(T-T_0)} \tag{7.17}$$

and for $t \in [0, T_0]$

$$\eta_t = \frac{d\hat{\mathbb{Q}}}{d\mathbb{Q}|\mathcal{F}_t} = \mathbb{E}^{\hat{\mathbb{Q}}}\left[\frac{S_{T_0}B_0}{S_0 B_{T_0}}\middle|\mathcal{F}_t\right] = \frac{S_t B_0}{S_0 B_t} \tag{7.18}$$

by writing $b_t = aS_t$, and changing the probability measure from $\mathbb{Q} \sim \hat{\mathbb{Q}}$ we get

$$\boldsymbol{FS}_t = b_t\mathbb{E}^{\hat{\mathbb{Q}}}\left[(Y - K)^+|\mathcal{F}_t\right] \tag{7.19}$$

The process Y is constant before the delivery date T_0, the random variable Y value at maturity T is a process with vanishing volatility for every $t \in [0, T_0]$, that leads us to the following The arbitrage free price at time $t \in [0, T_0]$ of a forward start option is

$$\boldsymbol{FS}_t = S_t\mathbb{E}^{\hat{\mathbb{Q}}}\left[e^{\int_{T_0}^T \sigma_t d\hat{W}t - \frac{1}{2}\int_{T_0}^T \sigma_t^2\,dt} - Ke^{-r(T-T_0)}\middle|\mathcal{F}_t\right] \tag{7.20}$$

where the process

$$\hat{W}_t = W_t^* - \int_0^t \sigma_u \mathbb{1}_{[0,T_0]}(u)\,du \tag{7.21}$$

is a standard BM under $\hat{\mathbb{Q}}$. It follows from the Girsanov theorem, it is obviously a BM for $t \in [0, T_0]$ and for the $t \in [T_0, T]$, notice that $Y = \frac{S_t}{S_{T_0}}$ and has the following representation

$$Y = \exp\left(r(T - T_0) + \int_{T_0}^T \sigma_t\,d\hat{W}_t - \frac{1}{2}\int_{T_0}^T \sigma_t^2\,dt\right) \tag{7.22}$$

In order to find the option price, we need to find the volatility process σ under $\hat{\mathbb{Q}}$, it can be done if we know the SDE, governing volatility process σ under $\mathbb{Q}$. Let σ have the following dynamics under $\mathbb{Q}$

$$d\sigma_t = \tilde{a}(\sigma_t, t)dt + b(\sigma_t, t)d\tilde{W}_t \tag{7.23}$$

where $\tilde{W}$ is a one-dimensional standard BM, possible correlated with $\hat{W}$, s.t. $d < \hat{W}, W^* >= \rho dt$, taking values in $[-1, 1]$, while under the measure $\hat{\mathbb{Q}}$, the volatility process σ_t is

$$d\sigma_t = \overline{a}(\sigma_t, t)dt + b(\sigma_t, t)d\overline{W}_t \tag{7.24}$$

where

$$\overline{W}_= \tilde{W}_t - \int_0^t \rho_u \sigma_u \mathbb{1}_{[0,T_0]}(u)\, du \tag{7.25}$$

and the adjusted drift coefficient $\overline{a}(t, \sigma_t)$ is given by

$$\overline{a}(\sigma_t, t) = \tilde{a}(\sigma_t, t) + \sigma_t \rho_t b(\sigma_t, t) \mathbb{1}_{[0,T_0]}(t) \tag{7.26}$$

This is the model that Lucic and Kruse-Nögel used to incorporate the Heston model for stochastic volatility.

7.3 Lucic' solution

His (Lucic, 2003) article describes how you can price a forward - start option via the change of numeraire. The terminal payoff can be written in two ways. The first way as

$$(S_T - K S_{T_0})^+ \tag{7.27}$$

where T is the maturity, $T_0 < T$, is the strike set date, and K, is the (percentage) strike. A forward - start contract can also be seen as the building block of cliquet options, we write the payoff as

$$\left(\frac{S_T}{S_{T_0}} - K \right)^+ \tag{7.28}$$

The problem is to value the options. Assume that you two independent Brownian motions, one is driving the asset process, and the other driving the variance process, under some Martingale measure $\mathbb{Q}$ by

$$\begin{aligned} dS_t &= r_t S_t\, dt + \sigma_t(v_t, S_t) S_t\, dW_t^{(1)} \\ dv_t &= \alpha_t(v_t)\, dt + \beta_t(v_t) \left(\rho\, dW_t^{(1)} + \sqrt{1-\rho^2} dW_t^{(2)} \right) \end{aligned} \tag{7.29}$$

Assume some regularity assumptions, and that the discounted asset price should be a Martingale. In this general framework we can study many different models for stochastic volatility (Hull and White, Stein - Stein, Heston) and the local volatility model from Dupire. Let

$$P(s,t) = \exp\left(-\int_s^t r_u \mathbb{1}_{s\le u}\, du\right) \tag{7.30}$$

Let $S_t^{T_0}$ be the asset price process stopped at T_0

$$S_t^{T_0} = S_{t\wedge T_0} \tag{7.31}$$

then the payoff in equation (7.27) can be written as

$$\left(S_T - KS_T^{T_0}\right)^+ \tag{7.32}$$

and the value of the option will be

$$V^{(1)} = P(t,T)\mathbb{E}^{\mathbb{Q}}\left[\left(S_T - KS_T^{T_0}\right)^+\middle|\mathcal{F}_t\right] \tag{7.33}$$

You can split the forward - start option into two parts, one part before the grant date, $t\in[0,T_0]$ and the other part $t\in[T_0,T]$. Fix a t in the former part, and study the asset price process.

$$S_u = S_0\exp\left(\int_0^u\left(r_s - \frac{1}{2}\sigma_s^2\right)ds + \int_0^u \sigma_s\, dW_s^{(1)}\right) \quad u\in[0,T_0] \tag{7.34}$$

Do a change of numeraire

$$N_u = \frac{S_u^{T_0}}{P(T_0,u)} \tag{7.35}$$

Than we can re-write equation (7.33) as

$$\begin{aligned} V^{(1)} &= N_t\mathbb{E}^{\mathbb{N}}\left[\left(\frac{S_T P(T_0,T)}{S_T^{T_0}} - KP(T_0,T)\right)^+\middle|\mathcal{F}_t\right] \\ V^{(1)} &= S_t P(T_0,T)\mathbb{E}^{\mathbb{N}}\left[\left(\exp\left(\int_{T_0}^T\left(r_s - \frac{1}{2}\sigma_s^2\right)ds + \int_{T_0}^T\sigma_s\, dW_s^{(1)}\right) - K\right)^+\middle|\mathcal{F}_t\right] \end{aligned} \tag{7.36}$$

where

$$\frac{d\mathbb{N}}{d\mathbb{Q}} = \frac{N_T P(0,T)}{N_0} = \exp\left(-\frac{1}{2}\int_0^T \sigma_s^2\mathbb{1}_{s\le T_0}\, ds + \int_0^T \sigma_s \mathbb{1}_{s\le T_0}\, dW_s^{(1)}\right) \tag{7.37}$$

Using equations (7.29), and (7.36) we have the following dynamics, under the measure $\mathbb{N}$, and use the fact that the Girsanov theorem

$$\begin{aligned} W_u^{N(1)} &= W_u^{(1)} - \int_0^u \sigma_s\mathbb{1}_{s\le T_0}\, ds \\ W_u^{N(2)} &= W_u^{(2)} \end{aligned} \tag{7.38}$$

gives us two independent Brownian Motion

$$\begin{aligned}
dS_u &= \left(r_u + \sigma_u^2 \mathbb{1}_{u \le T_0} S_u\right) dt + \sigma_u S_u \, dW_u^{N(1)} \\
dv_u &= \left(\alpha_u + \rho\beta_u\sigma_u \mathbb{1}_{u \le T_0}\right) du + \beta_u \left(\rho dW_u^{N(1)} + \sqrt{1-\rho^2} dW_t^{N(2)}\right) \\
V^{(1)} &= S_t P(T_0,T) \mathbb{E}^{\mathbb{N}} \left[\left(\exp\left(\int_{T_0}^{T} \left(r_s - \frac{1}{2}\sigma_s^2\right) ds + \int_{T_0}^{T} \sigma_s \, dW_s^{(1)} \right) - K \right)^+ \middle| \mathcal{F}_t \right]
\end{aligned} \tag{7.39}$$

By the risk - neutral valuation theorem, the value process scaled by the numeraire of the asset will be a martingale, under the measure induced by the numeraire. So the $\frac{V_t^{(1)}}{S_t}$ is the value of the European Call options, and the asset dynamics is under the risk - neutral measure $\mathbb{Q}$ is

$$\begin{aligned}
\hat{S}_u &= \hat{r}_u \hat{S}_u \, du + \hat{\sigma}_u \hat{S}_u \, dW_u^{N(1)} \quad \hat{S}_0 = 1 \\
\hat{r}_u &= r_u \mathbb{1}_{T_0 \le 0} \\
\hat{\sigma}_u &= \sigma_u(v_v, S_u) \mathbb{1}_{T_0 \le 0}
\end{aligned} \tag{7.40}$$

With this change of numeraire, the asset $\hat{S}_t$ is frozen until T_0, the grant date, or the time when the strike is set. If we use the payout in (7.28)

$$V^{(2)} = P(t,T) \mathbb{E}^{\mathbb{Q}} \left[\left(\exp\left(\int_{T_0}^{T} \left(r_s - \frac{1}{2}\sigma_s^2\right) ds + \int_{T_0}^{T} \sigma_s \, dW_s^{(1)} \right) - K \right)^+ \middle| \mathcal{F}_t \right] \tag{7.41}$$

The pricing of a forward start call option is thus reduced to pricing vanilla call options.

Lucic, Forward start option in the Heston model

In the Heston model we have the following dynamics.

$$\begin{aligned}
dS_t &= r_t S_t \, dt + \sqrt{v_t} S_t \, dW^{(1)} \\
dv_t &= \lambda(\overline{v} - v_t \, dt + \eta\sqrt{v_t} \left(\rho dW^{(1)} + \sqrt{1-\rho^2} dW^{(2)}\right)
\end{aligned} \tag{7.42}$$

This is a case of (7.29). Under the risk - neutral measure $\mathbb{Q}$ we can write for the two types of payouts, (7.27), 7.28) as $V^{(m)}$ for $m = 1, 2$, we have the following dynamics.

$$\begin{aligned}
d\hat{S}_t^{(m)} &= r_t \hat{S}_t^{(m)} \mathbb{1}_{T_0 \le t} \, dt + \sqrt{v_t^{(m)}} S_t^{(m)} \mathbb{1}_{T_0 \le t} \, dW^{(1)} \\
dv_t^{(m)} &= \left(\lambda\overline{v} - (\lambda - \rho\eta(2-m)\mathbb{1}_{t \le T_0}) v_t^{(m)}\right) dt + \eta\sqrt{v_t^{(m)}} \left(\rho dW^{(1)} + \sqrt{1-\rho^2} dW^{(2)}\right)
\end{aligned} \tag{7.43}$$

The difference from (7.42) to the above, is that the coefficients are (time) piece-wise constant coefficients. So the original Heston procedure can be used, over the discrete intervals where the coefficients are constants. Lucic follow the steps as outlined in Gatheral's notes. Denote $\tau = T - t$, and use, as in Heston, the log scale of the asset process, $x = \log(S)$.

$$\begin{aligned} \frac{\partial C}{\partial \tau} &= \lambda D, \quad C(0) = 0 \\ \frac{\partial D}{\partial \tau} &= \alpha_t - \beta_t^{(m)} D + \frac{\eta^2 D^2}{2}, \quad D(0) = 0 \end{aligned} \tag{7.44}$$

Now we need to integrate (7.44) over $[0, \tau]$ which is done in two separate cases, if $\tau \in [0, T - T_0]$, then it is a vanilla call option, we know the asset value at S_{T_0} since the filtration is after T_0, so we get constants for all parameters

$$\begin{aligned} D(m,k,\tau) &= r_-^{(m)} \frac{1 - e^{-d^{(m)}\tau}}{1 - g^{(m)} \exp(-d^m \tau)} \\ C(m,k,\tau) &= \lambda \left(r_-^{(m)} \tau - \frac{2}{\eta^2} \log \left(\frac{1 - g^{(m)} e^{-d^{(m)}\tau}}{1 - g^{(m)}} \right) \right) \\ d^{(m)} &= \sqrt{\left(\beta_0^{(m)}\right)^2 - 2\alpha_0 \eta^2} \\ r_\pm^{(m)} &= \frac{\beta_0^{(m)} \pm d^{(m)}}{\eta^2} \\ g^{(m)} &= \frac{r_-^{(m)}}{r_+^{(m)}} \end{aligned} \tag{7.45}$$

for the genuine forward - start option, where $\tau > T - T_0$, we are integrating over $[T - T_0, \tau]$ and using C and D from (7.45) as the initial conditions.

$$\begin{aligned} D(m,k,\tau) &= \frac{2\beta_T^{(m)}}{\eta^2 \left(1 + c \exp(\beta_T^{(m)}(\tau - T + T_0))\right)} \\ C(m,k,\tau) &= C(m,k,T - T_0) + \frac{2\beta_T^{(m)} \lambda(\tau - T + T_0)}{\eta^2} \\ &\quad - \frac{2\lambda}{\eta^2} \log \left(\frac{1 + \exp\left(\beta_T^{(m)}(\tau - T + T_0)\right)}{1 + c} \right) \\ c &= \frac{2\beta_T^{(m)}}{\eta^2 D(m,k,T - T_0)} - 1 \end{aligned} \tag{7.46}$$

This solves the calculations for the Fourier transform in the option price.

7.4 Kruse: FSO under Heston model

Here I am following the following article (?, ?) A forward start option starts somewhere in the future, the determination time of the strike, when the strike is set equal to a proportion of the current price. In the BS setting, one can easily transform the pricing problem of a FSO, into a valuation problem of a vanilla option at the determination time. The option price at the determination time, has only one stochastic component at the determination time, the asset stock price. In a stochastic volatility model we add the randomness of the volatility of the underlying. It makes the today's price to rely on today's volatility, and the assumption of the SDE of the volatility process.

The payoff structure is

$$P_{FWS}(S(T, S(t^*)) = (S(T) - kS(t^*))^+ \tag{7.47}$$

where k is the percentage of the strike price. In Heston's model we have following structure for the asset is under measure $\mathbb{Q}$ and for the volatility under measure $\mathbb{Q}^\lambda$

$$dS(t) = rS(t) + \sqrt{\nu(t)}S(t)dW_1(t) \tag{7.48}$$

$$d\nu(t) = \kappa(\theta - \nu(t))dt + \sigma\sqrt{\nu(t)}d\left(\rho W_1 + \sqrt{1-\rho^2}W_2(t)\right) \tag{7.49}$$

assuming some regularity conditions, and note that $d(W_1, W_2) = 0dt$ they are uncorrelated. Heston showed that for a European vanilla option, with payoff

$$P(S(T)) = \left(S(T) - K)^+\right) \tag{7.50}$$

where K, the strike is known, by using a Delta - Sigma hedging, we end up with P_1 and P_2, and using Fourier transformation, and Riccati equation for solving parabolic PDE. The option price at time $t \in [0, T]$ is

$$C(t, S(t), \nu(t)) = S(t)P_1(t, S(t), \nu(t), \overline{K}) - \overline{K}e^{-r(T-t)}P_2(t, S(t), \nu(t), \overline{K}) \tag{7.51}$$

where P_j for $j = 1, 2$ are given by

$$P_j(t, S(t), \nu(t), \overline{K}) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\frac{e^{-i\phi\log(\overline{K})}f_j(S(t), \nu(t), T-t, \phi}{i\phi}\,d\phi \tag{7.52}$$

and f_j have the characteristic function

$$f_j(S(t), \nu(t), T-t, \phi) = \exp\left(i\phi\log(S(t)) + C_j(\phi, T-t) + D_j(\phi, T-t)\nu(t)\right) \tag{7.53}$$

note that C_j and D_j is an affine function. And it uses Riccati equation for solving it.

7.4.1 Kruse's solution

Choose a time t before the determination time t^*, an European Call option and some regularity conditions, its price is

$$C(t, \nu(t), S(t)) = S(t)\hat{P}_1(t, \nu(t)) - ke^{-r(T-t^*)}S(t)\hat{P}_2(t, \nu(t)) \tag{7.54}$$

with P_j the probabilities as in Heston model

$$\hat{P}_j(t, \nu(t)) := \int_0^\infty P_j(t^*, 1, \nu(t), k) f(\nu(t^*)|\nu(t)\, d\nu(t^*) \tag{7.55}$$

$$f(\nu(t^*)|\nu(t)) = \frac{B}{2} e^{-(B\nu(t^*)+\Lambda)/2} \left(\frac{B(\nu(t^*)}{\Lambda}\right)^{(R/2-1)/2} \mathbb{1}_{R/2-1}\left(\sqrt{\Lambda B \nu(t^*)}\right) \tag{7.56}$$

Proof: The price of a call option is a numeraire under $\mathbb{Q}$, we change the numeraire to the stock measure, it will also be a Martingale $\mathbb{P}^{\mathbb{S}}$, the BM, that is driving the asset process and the volatility process, will with Girsanov, have a new dynamics.

$$W_1^S(t) = W_1(t) - \int_t^T \sqrt{\nu(s)}\, ds \tag{7.57}$$

$$W_2^S(t) = W_2(t) \tag{7.58}$$

note that $d(W_1^S, W_2^S) = 0dt$ Under the new stock measure, the dynamics can be written as

$$dS(t) = rS(t) + \sqrt{\nu(t)}S(t)dW_1^S(t) \tag{7.59}$$

$$d\nu(t) = \tilde{\kappa}(\hat{\theta} - \nu(t))dt + \sigma\sqrt{\nu(t)}d\left(\rho W_1^S(t) + \sqrt{1-\rho^2}W_2^S(t)\right) \tag{7.60}$$

where

$$\hat{\kappa} = \kappa - \rho\sigma, \quad \hat{\theta} = \frac{\kappa\theta}{\kappa - \rho\sigma} \tag{7.61}$$

The option price, by use of the Tower property can be re-written as

$$C_{FWS}(t, \nu(t), S(t)) = \mathbb{E}^{\mathbb{S}}\left[S(t)\left(1 - k\frac{S(t^*)}{S(T)}\right)^+ |\mathcal{F}_t\right] \tag{7.62}$$

remember that t^* is the delivery time, and if the valuation time is $t < t^*$ we get, again the tower property of expectation

$$C_{FWS}(t, \nu(t), S(t)) = S(t)\mathbb{E}^{\mathbb{S}}\left[\mathbb{E}^{\mathbb{S}}\left[\left(1 - k\frac{S(t^*)}{S(T)}\right)^+ |\mathcal{F}_{t^*}\right] |\mathcal{F}_t\right] \tag{7.63}$$

Since $t < t^*$ and is measurable we can re-write the last equation as

$$\mathbb{E}^{\mathbb{S}}\left[\mathbb{E}^{\mathbb{S}}\left[\left(\frac{S(t^*)\,(S(T) - kS(t^*))}{S(T)}\right)^+ |\mathcal{F}_{t^*}\right] |\frac{1}{S(t^*)}|\mathcal{F}_t\right] \tag{7.64}$$

where the inside expectation is the value of the call option at determination point.

$$\mathbb{E}^{\mathbb{S}}\left[\left(\frac{S(t^*)\left(S(T)-kS(t^*)\right)}{S(T)}\right)^+|\mathcal{F}_{t^*}\right]| = C_{FWS}(t^*,\nu(t^*),S(t^*)) \tag{7.65}$$

that can we insert in Heston option pricing formula

$$S(t^*)\left(P_1(t^*,S(t^*),\nu(t^*),kS(t^*)-ke^{-r(T-t^*)}P_2(t^*,S(t^*),\nu(t^*),kS(t^*)\right) \tag{7.66}$$

By the definition of P_j we know that the probabilities don't depend on $S(t^*)$, so to obtain the option price at a time $t < t^*$, prior to the determination of the strike is

$$C_{FWS}(t,\nu(t),S(t)) = S(t)\mathbb{E}^{\mathbb{S}}\left[\left(\frac{C_{FWS}(t^*,\nu(t^*),S(t^*)}{S(t^*)}\right)\Big|\mathcal{F}_t\right] \tag{7.67}$$

so the pricing formula becomes

$$C_{FWS}(t,\nu(t),S(t)) = S(t)\mathbb{E}^{\mathbb{S}}\left[P_1(t^*,1,v(t^*),k|\mathcal{F}_t\right]-kS(t)e^{-r(T-t^*)}\mathbb{E}^{\mathbb{S}}\left[P_2(t^*,1,v(t^*),k|\mathcal{F}_t\right] \tag{7.68}$$

for $j = 1, 2$ Heston's model gives us

$$\hat{P}_j(t,\nu(t)) = \mathbb{E}^{\mathbb{S}}\left[P_j(t^*,1,\nu(t^*),k)|\mathcal{F}_t\right] \tag{7.69}$$

47 To calculate the conditional expectations involves Bessel function.

7.5 Forward start option using the CF

Here I will follow (Oosterlee & Grzelak, 2019). A forward start option can be viewed as a performance option. Let there be two maturity days, T_1 and T_2 with $t_0 < T_1 < T_2$. A forward start option payoff is defined as

$$V^{fwd}(T_2,S(T_2)) := \max\left(\frac{S(T_2)-S(T_1)}{S(T_1)}-K,0\right) \tag{7.70}$$

with a strike price K, that is fixed, usually a percentage of the stock value at time $S(T_1)$. The case when $t_0 = T_1$, the option will just be normal vanilla option, and its payoff will be

$$V^{fwd}(T_2,S(T_2)) = \frac{1}{S_0}\max\left(S(T_2)-S_0K^*,0\right) \quad K^* = K+1 \tag{7.71}$$

The value of the contract depends on the performance, (percentage) of the asset under two time points, T_1, T_2. That is the building blocks for cliquets. We can re-write (7.70) as

$$V^{fwd}(T_2,S(T_2)) := \max\left(\frac{S(T_2)}{S(T_1)}-K^*,0\right) \tag{7.72}$$

Assume that there is an equivalent martingale measure $\mathbb{Q}$ and a bank account, with a constant deterministic interest rate. Than the RNVF will give us the today's value from the payoff at time T_2

$$V^{fwd}(t_0, S_0) = B(t_0)\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{B(T)}\max\left(\frac{S(T_2)}{S(T_1)} - K^*, 0\right)\bigg| \mathcal{F}(t_0)\right] \tag{7.73}$$

Using (Duffie et al., 1999), to have an affine process, we need to take the logarithm of the asset values, using (7.72) and the fact that the logarithm of the quotient is the difference of the logarithms we get

$$V^{fwd}(t_0, S_0) = \frac{B(t_0)}{B(T_2}\mathbb{E}^{\mathbb{Q}}\left[\max\left(e^{x(T_1,T_2)} - K^*, 0\right)\Big| \mathcal{F}(t_0)\right] \tag{7.74}$$

where $x(T_1, T_2) = \log(S(T_2)) - \log(S(T_1))$. Now we can derive the characteristic function

$$\phi_x(u) \equiv \phi_x(u, t_0, T_2) = \mathbb{E}^{\mathbb{Q}}\left[e^{iu(\log(S(T_2)) - \log(S(T_1))}\Big| \mathcal{F}(t_0)\right] \tag{7.75}$$

Using the property of iterated expectation, the Tower property, we can condition (7.75) on the time T_1 and write the characteristic function as

$$\phi_x(u) = \mathbb{E}^{\mathbb{Q}}\left\{\mathbb{E}^{\mathbb{Q}}\left[e^{iu(\log(S(T_2)) - \log(S(T_1))}\Big| \mathcal{F}(T_1)\right]\Big| \mathcal{F}(t_0)\right\} \tag{7.76}$$

in the inner expectation in (7.76), which is conditioned at time T_1, that is, we know the log asset price at that time, and can take it outside the inner expectation, also make use of the fact that $e^{-r(T_2-T_1)} \cdot e^{r(T_2-T_1)} = 1$, in my models r is not stochastic we get

$$\phi_x(u) = \mathbb{E}^{\mathbb{Q}}\left\{e^{-iu(\log(S(T_1))}e^{r(T_2-T_1)}\mathbb{E}^{\mathbb{Q}}\left[e^{iu(\log(S(T_2))}e^{-r(T_2-T_1)}\Big| \mathcal{F}(T_1)\right]\Big| \mathcal{F}(t_0)\right\} \tag{7.77}$$

The inner expectation is the *discounted* characteristic function of $X(T_2) = \log(S(T_2))$, so we have

$$\phi_x(u) = \mathbb{E}^{\mathbb{Q}}\left[e^{-iu(\log(S(T_1))}e^{r(T_2-T_1)}\psi_X(u, T_1, T_2)\Big| \mathcal{F}(t_0)\right] \tag{7.78}$$

where we will derive the $\psi_X(u, T_1, T_2)$ for two different asset classes, the Black - Scholes and the Heston model.

7.5.1 Pricing under the Black - Scholes model

Under the Black - Scholes we have discounted characteristic function, on the log stock asset price $X(t)$, conditioned on the information until the time T_1

$$\psi_x(u, T_1, T_2) = \exp\left[\left(r - \frac{1}{2}\sigma^2\right)iu\Delta T - \frac{1}{2}\sigma^2 u^2 \Delta T - r\Delta T + iuX(T_1)\right] \tag{7.79}$$

where $T_2 - T_1 = \Delta T$. Insert the above in (7.78) we get

$$\begin{aligned}\phi_x(u) &= \mathbb{E}^{\mathbb{Q}}\left[e^{\left(r-\frac{1}{2}\sigma^2\right)iu\Delta T - \frac{1}{2}\sigma^2 u^2 \Delta T}\middle| \mathcal{F}(t_0)\right] \\ \phi_x(u) &= \left(r - \frac{1}{2}\sigma^2\right) iu\Delta T - \frac{1}{2}\sigma^2 u^2 \Delta T\end{aligned} \tag{7.80}$$

and the last line is the characteristic function for normally distributed random variable with mean equal to $\left(r - \frac{1}{2}\sigma^2\right)\Delta T$ and variance equal to $\sigma^2\Delta T$. The above equation does not depend on S_t, it is a consequence that in the Black - Scholes model, the ratio of two, assets, only depend on the r, the interest rate, and σ the volatility. So we can get a pricing formula at time $t_0 = 0$ for the forward start option under the Black - Scholes model.

$$\begin{aligned}V^{fwd}(t_0, S_0) &= e^{-rT_2}\mathbb{E}^{\mathbb{Q}}\left[\max\left(\frac{S(T_2)}{S(T_1)} - K^*\right)\middle| \mathcal{F}(t_0)\right] \\ V^{fwd}(t_0, S_0) &= e^{-rT_1}\Phi(d_1) - K^* e^{-rT_2}\Phi(d_2) \\ &\text{where} \\ d_1 = \frac{\log\left(\frac{1}{K^*}\right) + (r + \frac{1}{2}\sigma^2)\Delta T}{\sigma\sqrt{\Delta T}} &\quad d_2 = \frac{\log\left(\frac{1}{K^*}\right) + (r - \frac{1}{2}\sigma^2)\Delta T}{\sigma\sqrt{\Delta T}}\end{aligned} \tag{7.81}$$

Proof of (7.81), first note that

$$\frac{S(T_2)}{S(T_1)} = e^{r - \frac{1}{2}\sigma^2\Delta T + \sigma(W(T_2) - W(T_1))} \tag{7.82}$$

insert it in the expectation of (7.81)

$$\begin{aligned}V^{fwd}(t_0, S_0) &= e^{-rT_2}\mathbb{E}^{\mathbb{Q}}\left[\max\left(\frac{S(T_2)}{S(T_1)} - K^*\right)\middle| \mathcal{F}(t_0)\right] \\ V^{fwd}(t_0, S_0) &= \frac{e^{-rT_2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\max\left(e^{(r-\frac{1}{2}\sigma^2)\Delta T + \sigma(T_2 - T_1)x} - K^*, 0\right)e^{-\frac{1}{2}x^2}\end{aligned} \tag{7.83}$$

The integral can be split into two integrals, and note that we are only interest positive payout

$$\begin{aligned}V^{fwd}(t_0, S_0) = \frac{e^{-rT_2}}{\sqrt{2\pi}}&\int_a^{\infty} e^{(r-\frac{1}{2}\sigma^2)\Delta T + \sigma(T_2 - T_1)x)}e^{-\frac{1}{2}x^2} \\ &- K^* e^{-rT_2}\left(1 - \Phi(a)\right) \quad \text{with} \\ a &= \frac{1}{\sigma\sqrt{\Delta T}}\left(\log K^* - \left(r - \frac{1}{2}\sigma^2\right)\Delta T\right)\end{aligned} \tag{7.84}$$

The integral in the last expression can be simplified, by taking the constants out of the integrals.

$$\begin{aligned}&= \frac{e^{r\Delta T}}{\sqrt{2\pi}}\int_a^{\infty} e^{(r-\frac{1}{2}\sigma^2)\Delta T + \sigma(T_2 - T_1)x)}e^{-\frac{1}{2}x^2} \\ &= \frac{e^{(r-\frac{1}{2}\sigma^2)\Delta T + \frac{1}{2}\sigma^2\Delta T}}{\sqrt{2\pi}}\int_a^{\infty} e^{-\frac{1}{2}\left(x - \sigma\sqrt{\Delta T}\right)}\end{aligned} \tag{7.85}$$

make the normal random variable standard, by extracting its mean and divide with its deviation and we get

$$\begin{aligned} &\frac{e^{(r-\frac{1}{2}\sigma^2)\Delta T+\frac{1}{2}\sigma^2\Delta T}}{\sqrt{2\pi}}\int_a^\infty e^{-\frac{1}{2}\left(x-\sigma\sqrt{\Delta T}\right)} \\ &= e^{r\Delta T}\left[1-\Phi\left(a-\sigma\sqrt{\Delta T}\right)\right] \end{aligned} \tag{7.86}$$

use the fact the standard normal is symmetric around zero, we get

$$\begin{aligned} V^{fwd}(t_0,S_0) &= e^{rT_1}\Phi\left(\sigma\sqrt{\Delta T}-a\right)-K^*e^{-rT_2}\Phi(-a) \\ &\text{with} \\ d_1 &= \frac{\log\left(\frac{1}{K^*}\right)+\left(r+\frac{1}{2}\sigma^2\right)\Delta T}{\sigma\sqrt{\Delta T}} \\ d_2 &= \frac{\log\left(\frac{1}{K^*}\right)+\left(r-\frac{1}{2}\sigma^2\right)\Delta T}{\sigma\sqrt{\Delta T}} \end{aligned} \tag{7.87}$$

7.5.2 Pricing under the Heston model

Before we looked at the Black Scholes model, we had found that pricing the forward start function, is the same as calculating this characteristic function

$$\phi_x(u) = \mathbb{E}^{\mathbb{Q}}\left[e^{-iu(X(T_1))}e^{r(T_2-T_1)}\psi_X(u,T_1,T_2)\middle|\,\mathcal{F}(t_0)\right] \tag{7.88}$$

under the Heston model this is to find $\psi_X(u,T_1,T_2)$, now the state- space u is a two-dimensional vector, $\mathbf{u}^T=[u,0]^T$, with the second parameter is set to zero. We want to find the asset price at maturity, not the variance. Of course the variance influences the asset price, but it is already captured in the asset price. Using the fact that Heston model belongs to the class of AJD, affine jump diffusions, we know that its characteristic function can be written as

$$\psi_X(u,T_1,T_2) = e^{\overline{A}(u,\tau)+\overline{B}(u,\tau)X(T_1)+\overline{C}(u,\tau)v(T_1)} \tag{7.89}$$

where $\overline{A},\overline{B},\overline{C}$ are complex valued function. In the Heston model, the variance follows a CIR - model, a squared root diffusion, with mean-reversion, and no jumps. In the Heston model $\overline{B}(u,\tau)=iu$ and in (7.89) the constant $\overline{A}(u,\tau)$, is not stochastic. I let $\tau=T-t$, time to maturity. So this simplifies (7.88) to the following

$$\phi_x(u) = e^{\overline{A}(u,\tau)+r(T_2-T_1)}\mathbb{E}^{\mathbb{Q}}\left[e^{\overline{C}(u,\tau)v(T_1)}\psi_X(u,T_1,T_2)\middle|\,\mathcal{F}(t_0)\right] \tag{7.90}$$

This formula does not depend on the asset price $S(t)$ or the log asset price $X(t)=\log(S(t))$. In order to get an affine system we are using the log of the asset price. The idea behind solving (7.90) is to use moment - generating function for the CIR - model. In the Heston model we have the in (Oosterlee & Grzelak, 2019) writes the representation for the dynamics of the variance thus

$$dv(t) = \kappa(\overline{v} - v(t))\, dt + \gamma\sqrt{v(t)}\, dW_v(t) \tag{7.91}$$

That is the same as in Heston's representation.

$$dv(t) = \kappa(\theta - v(t))\, dt + \sigma\sqrt{v(t)}\, dW_v(t) \tag{7.92}$$

the moment generating function has the following form.

$$\mathbb{E}^{\mathbb{Q}}\left[e^{uv(t)}\middle|\,\mathcal{F}(t_0)\right] = \left(\frac{1}{1-2u\overline{c}(t,t_0)}\right)^{\frac{1}{2}\delta}\exp\left(\frac{u\overline{c}(t,t_0)\overline{\kappa}(t,t_0)}{1-2u\overline{c}(t,t_0)}\right) \tag{7.93}$$

with the following parameters

$$\begin{aligned}
\overline{c}(t,t_0) &= \frac{\gamma^2}{4\kappa}\left(1-e^{-\kappa(t-t_0)}\right)\\
\delta &= \frac{4\kappa\overline{v}}{\gamma^2}\\
\overline{\kappa}(t,t_0) &= \frac{4\kappa v_0 e^{-\kappa(t-t_0)}}{\gamma^2\left(1-e^{-\kappa(t-t_0)}\right)}
\end{aligned} \tag{7.94}$$

The density of a noncentral chi - square distribution is $\chi^2(\delta, \overline{\kappa}(t,t_0))$

$$f_{\chi^2(\delta,\overline{\kappa}(t,t_0))}(x) = \sum_{k=0}^{\infty}\frac{1}{k!}e^{-\frac{\overline{\kappa}(t,t_0)}{2}}\left(\frac{\overline{\kappa}(t,t_0)}{2}\right)^k f_{\chi^2(\delta+2k)}(x) \tag{7.95}$$

which is the chi-squared distribution with $\delta + 2k$ degrees of freedom. Than the moment - generating function becomes.

$$\begin{aligned}
\mathcal{M}_{v(t)}(u) &:= \mathbb{E}^{\mathbb{Q}}\left[e^{uv(t)}\middle|\,\mathcal{F}(t_0)\right]\\
&= \frac{1}{\overline{c}(t,t_0)}\sum_{k=0}^{\infty}\frac{1}{k!}e^{-\frac{\overline{\kappa}(t,t_0)}{2}}\left(\frac{\overline{\kappa}(t,t_0)}{2}\right)^k\int_0^{\infty}e^{uy}f_{\chi^2(\delta+2k)}\left(\frac{y}{\overline{c}(t,t_0)}\right)dy
\end{aligned} \tag{7.96}$$

Change of variables $y = \overline{c}(t,t_0)x$ gives us

$$\mathcal{M}_{v(t)}(u) = \sum_{k=0}^{\infty}\frac{1}{k!}e^{-\frac{\overline{\kappa}(t,t_0)}{2}}\left(\frac{\overline{\kappa}(t,t_0)}{2}\right)^k\int_0^{\infty}e^{u\overline{c}(t,t_0)}f_{\chi^2(\delta+2k)}(x)\, dx \tag{7.97}$$

The integral is the moment generating function for a chi-squared distribution with $\delta + 2k$ degrees of freedom. So we have

$$\mathcal{M}_{\chi^2(\delta+2k)}(u\overline{c}(t,t_0)) = \int_0^{\infty}e^{u\overline{c}(t,t_0)}f_{\chi^2(\delta+2k)}(x)\, dx = \left(\frac{1}{1-2u\overline{c}(t,t_0)}\right)^{\frac{1}{2}\delta+k} \tag{7.98}$$

adding and subtracting an exponential term

$$\mathcal{M}_{v(t)}(u) = \left(\frac{1}{1-2u\overline{c}(t,t_0)}\right)^{\frac{1}{2}\delta} \exp\left(\frac{\overline{\kappa}(t,t_0)}{2(1-2u\overline{c}(t,t_0)} - \frac{\overline{\kappa}(t,t_0)}{2}\right) \\ \cdot \sum_{k=0}^{\infty} \frac{1}{k!} e^{-\frac{\overline{\kappa}(t,t_0)}{2(1-2u\overline{c}(t,t_0))}} \left(\frac{\overline{\kappa}(t,t_0)}{2(1-2u\overline{c}(t,t_0))}\right)^k \tag{7.99}$$

The expression under the sum, is the probability that $\mathbb{P}(Y = k)$ for a Poisson distributed random variable, the probability mass function, for a Poisson distributed random variable (with parameter $\hat{\alpha}$ is

$$\mathbb{P}[Y = k] = \frac{1}{k!} e^{\hat{\alpha}} \hat{\alpha}^k \tag{7.100}$$

in our example $\hat{\alpha} = \frac{\overline{\kappa}(t,t_0)}{2(1-2u\overline{c}(t,t_0))}$ we get

$$\mathcal{M}_{v(t)}(u) = \left(\frac{1}{1-2u\overline{c}(t,t_0))}\right)^{\frac{1}{2}\delta} \exp\left(\frac{\overline{\kappa}(t,t_0)}{2(1-2u\overline{c}(t,t_0)} - \frac{\overline{\kappa}(t,t_0)}{2}\right) \sum_{k=0}^{\infty} \mathbb{P}[Y = k] \tag{7.101}$$

but as the sum of all probabilities is equal to one, the last sum vanish, and we get

$$\mathcal{M}_{v(t)}(u) = \left(\frac{1}{1-2u\overline{c}(t,t_0)}\right)^{\frac{1}{2}\delta} \exp\left(\frac{\overline{\kappa}(t,t_0)}{2(1-2u\overline{c}(t,t_0))} - \frac{\overline{\kappa}(t,t_0)}{2}\right) \tag{7.102}$$

So we can insert this in (7.90) we get the $\phi_X(u, T_1, T_2)$ when solving the coupled Riccait equations

$$\phi_x(u) = \exp\left(\overline{A}(u,\tau) + r\tau + \frac{\overline{C}(u,\tau)\overline{c}(T_1,t_0)\overline{\kappa}(T_1,t_0)}{1-2\overline{C}(u,\tau)\overline{c}(T_1,t_0)}\right) \\ \cdot \left(\frac{1}{1-2\overline{C}(u,\tau),\overline{c}(T_1,t_0)}\right)^{\frac{1}{2}\delta} \tag{7.103}$$

7.6 Appendix

I reproduce the variance dynamics in the Heston model (CIR- model)

$$dv(t) = \kappa(\overline{v} - v(t))\,dt + \gamma\sqrt{v(t)}\,dW_v(t) \tag{7.104}$$

The process $v(t)|v(s)$ with $0 < s < t$ under the CIR dynamics is distributed as $\overline{c}(t,s)$ times a non-central χ^2 random variable $\chi^2(\delta, \overline{\kappa}(t,s))$ where δ is the degree of freedom, and $\overline{\kappa}(t,s)$ is the non-centrality parameter. This gives us

$$v(t)|v(s) \sim \overline{c}(t,s)\chi^2(\delta, \overline{\kappa}(t,s)) \quad t > s > 0$$

$$\begin{aligned} &\text{with} \\ &\overline{c}(t,s) = \frac{1}{4\kappa}\gamma^2\left(1 - e^{-\kappa(t-s)}\right) \\ &\delta = \frac{4\gamma\overline{v}}{\gamma^2} \\ &\overline{\kappa}(t,s) = \frac{4\kappa v(s) e^{-\kappa(t-s)}}{\gamma^2\left(1 - e^{-\kappa(t-s)}\right)} \end{aligned} \tag{7.105}$$

The cumulative distribution function, CDF, will look like

$$F_{v(t)}(x) = \mathbb{Q}[v(t) \le x] = \mathbb{Q}\left[\chi^2(\delta, \overline{\kappa}(t,s)) \le \frac{x}{\overline{c}(t,s)}\right] = F_{\chi^2(\delta,\overline{\kappa}(t,s))}\left(\frac{x}{\overline{c}(t,s)}\right) \tag{7.106}$$

where

$$F_{\chi^2(\delta,\overline{\kappa}(t,s))}(y) = \sum_{k=0}^{\infty} \exp\left(-\frac{\overline{\kappa}(t,s)}{2}\right) \frac{\left(\frac{(\overline{\kappa}(t,s)}{2}\right)^2}{k!} \frac{\gamma(k + \frac{\delta}{2}, \frac{y}{2})}{\Gamma\left(k + \frac{\delta}{2}\right)} \tag{7.107}$$

and the lower incomplete Gamma function $\gamma(a,z)$, and the Gamma function $G(z)$ are

$$\gamma(a,z) \int_0^z t^{a-1} e^{-t}\, dt, \quad \Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt \tag{7.108}$$

and the probability density function, pdf, is

$$f_{\chi^2(\delta,\overline{\kappa}(t,s))}(y) = \frac{1}{2} e^{-\frac{1}{2}(y + \overline{\kappa}(t,s))} \left(\frac{y}{\overline{\kappa}(t,s)}\right)^{\frac{1}{2}(\frac{\delta}{2}-1)} \mathcal{B}_{\frac{\delta}{2}-1}\left(\sqrt{\overline{\kappa}(t,s)y}\right) \tag{7.109}$$

where the $\mathcal{B}$ is the modified Bessel function of the first kind.

$$\mathcal{B}_a(z) = \left(\frac{z}{2}\right)^a \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^k}{k!\Gamma(a+k+1)} \tag{7.110}$$

Now the density function for $v(t)$ is thus

$$\begin{aligned} f_{v(t)}(x) &:= \frac{d}{dx} F_{v(t)}(x) = \frac{d}{dx} F_{\chi^2(\delta,\overline{\kappa}(t,s))}\left(\frac{x}{\overline{c}(t,s)}\right) \\ &= \frac{1}{\overline{c}(t,s)} f_{\chi^2(\delta,\overline{\kappa}(t,s))}\left(\frac{x}{\overline{c}(t,s)}\right) \end{aligned} \tag{7.111}$$

and the mean and variances are

$$\begin{aligned} \mathbb{E}[v(t)|\mathcal{F}_0] &= \overline{c}(t,0)(\delta + \overline{\kappa}(t,0)) \\ \text{Var}[v(t)|\mathcal{F}_0] &= \overline{c}^2(t,0)(2\delta + 4\overline{\kappa}(t,0)) \end{aligned} \tag{7.112}$$

7.7 Bibliographical notes

The starting point is Rubinstein's article from 1990. A good book that describes the Forward- start problem is (Musiela & Rutkowski, 2005), thereafter I used two articles, (Lucic, 2003) and (?, ?), that independently produced a closed form expression for the Heston model. Another article is (Ahlip & Rutkowski, 2009), that also set up the framework for stochastic interest rate. I only reproduced their result with deterministic interest rate. A recent book (Oosterlee & Grzelak, 2019) makes use of the moment generating function to derive the Black Scholes and the Heston model.

References

Ahlip, R., & Rutkowski, M. (2009). Forward start options under stochastic volatility and stochastic interest rates. *International Journal of Theoretical and Applied Finance*, *12*, 209-225.

Björk, T. (2009). *Arbitrage theory in continuous time*. Oxford university press.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics (JFE)*, *7*(3).

Carr, P., & Madan, D. (1999). Option valuation using the fast fourier transform. *Journal of Computational Finance*, *2*, 61-73.

Carr, P., & Wu, L. (2004). Time-changed lévy processes and option pricing. *Journal of Financial economics*, *71*(1), 113–141.

Černỳ, A. (2004). Introduction to fast fourier transform in finance. *The Journal of Derivatives*, *12*(1), 73–88.

DERMAN, E., & Kani, I. (1994). Riding on a smile. *Risk*, *7*, 32–39.

Duffie, D., Pan, J., & Singleton, K. (1999). Transform analysis and asset pricing for affine jump-diffusions. *Capital Markets eJournal*.

Dupire, B., et al. (1994). Pricing with a smile. *Risk*, *7*(1), 18–20.

Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, *322*(8), 549-560. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19053220806` doi: https://doi.org/10.1002/andp.19053220806

Evans, L. C. (2012). *An introduction to stochastic differential equations* (Vol. 82). American Mathematical Soc.

Gatheral, J. (2011). *The volatility surface: a practitioner's guide* (Vol. 357). John Wiley & Sons.

Gilli, M., Maringer, D., & Schumann, E. (2019). *Numerical methods and optimization in finance* (Second ed.). Waltham, MA, USA: Elsevier/Academic Press. Retrieved from `http://www.enricoschumann.net/NMOF/` (ISBN 978-0128150658)

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, *6*, 327-343.

Iacus, S. M. (2011). *Option pricing and estimation of financial models with r*. John Wiley & Sons.

Kendall, M. G., et al. (1946). *The advanced theory of statistics.* (No. 2nd Ed). Charles Griffin and Co., Ltd., London.

Lindström, E., Madsen, H., & Nielsen, J. N. (2015). *Statistics for finance.* CRC Press.

Lucic, V. (2003). Forward-start options in stochastic volatility model. *Wilmott*, *2003*, 72-75.

Matsuda, K. (2004). Introduction to option pricing with fourier transform: Option pricing with exponential lévy models. white paper. *City University of New York*.

Musiela, M., & Rutkowski, M. (2005). Martingale methods in financial modelling springer-verlag. *Springer-Verlag, Berlin, Heidelberg, New York*.

Oosterlee, C. W., & Grzelak, L. A. (2019). *Mathematical modeling and computation in finance: With exercises and python and matlab computer codes.* World Scientific.

Pascucci, A. (2011). *Pde and martingale methods in option pricing.* Springer Science & Business Media.

Singleton, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, *102*(1), 111–141.