



UPPSALA  
UNIVERSITET

U.U.D.M. Project Report 2013:21

# Kolmogorov Equations

Gustav Ludvigsson

Examensarbete i matematik, 15 hp  
Handledare och examinator: Johan Tysk  
Juni 2013

A large, faint watermark of the Uppsala University seal is visible in the bottom right corner of the page. The seal features a sun with rays and the Latin motto "VERITAS LIBERABIT VOS".

Department of Mathematics  
Uppsala University



## **Abstract**

Stochastic and partial differential equations are both useful tools when modeling problems in reality. In this thesis we are going to study two equations that can be seen as a link between these two theories. The equations we are going to study are the forward and backward Kolmogorov equation. We are going to start off with some theory needed to understand and prove these equations. This is followed up with proofs of both equations and some examples where the equations are used.

# Contents

Abstract .....	1
<b>1 Introduction .....</b>	<b>3</b>
1.1 History .....	4
1.2 Objective .....	4
<b>2 Theory .....</b>	<b>5</b>
2.1 Information .....	5
2.2 Martingale .....	6
2.3 Markov property .....	6
2.4 Strong Markov property .....	6
2.5 Brownian motion .....	7
2.6 Stochastic integration .....	8
2.7 Stochastic Differential Equations .....	8
2.8 Itô's Lemma .....	10
2.9 An important operator and Dynkin's formula .....	11
2.10 The Fourier transform .....	14
<b>3 The Kolmogorov Equations .....</b>	<b>15</b>
3.1 Kolmogorov Backward Equation .....	15
3.2 Kolmogorov Forward Equation .....	17
<b>4 Applications .....</b>	<b>20</b>
4.1 Scaled Brownian motion .....	20
4.2 Scaled Brownian motion with linear drift .....	21
4.3 Geometric Brownian motion .....	22
4.4 Ornstein-Uhlenbeck process .....	23
4.5 Black-Scholes equation and its connection to the backward Kolmogorov equation .....	25
4.6 Call Option .....	26
Referenses .....	29

# Chapter 1

## Introduction

Throughout this thesis the main references are Øksendal [5] and Björk [3].

The theory stochastic differential equations is a part of mathematics with many applications. One may know that ordinary and partial differential equations are heavily used in physics and real life applications. However often problems involve some kind of randomness. If we add this randomness to some of the coefficients we can make a more realistic model for the problem. When we add the randomness to the equation we changed the model from an ordinary or partial differential equation to a stochastic differential equation. An example of this is the following:

Consider the standard model for population growth

$$\frac{dN}{dt} = a(t)N(t), \quad (e1.1.1)$$

$$N(0) = N_0. \quad (e1.1.2)$$

In this model  $N(t)$  is the population size at time  $t$  and  $a(t)$  is the relative growth rate at time  $t$ . However, it might be the case that  $a(t)$  is affected by a lot of random effects. We can in this case improve the model by setting

$$a(t) = r(t) + \text{"random"},$$

where  $r(t)$  is the deterministic part of  $a(t)$  and "random" is the random part of  $a(t)$ . Putting this new  $a(t)$  into the ordinary differential equation (e1.1.1)-(e1.1.2) we get a stochastic differential equation. Part of this thesis is about learning how to solve these kinds of problems.

With the knowledge we have today, stochastic differential equations are often hard or impossible to solve. However mathematicians and physicists have worked with ordinary and partial differential equations for centuries and because of that we have many tools to solve them. This is where the Kolmogorov equations come in. The Kolmogorov equations can in some cases be used as a bridge from stochastic differential equations to partial differential equations. Because of this bridge the theory of stochastic differential equations can benefit from the tools developed in the theory of ordinary and partial differential equations. Not only can stochastic differential equations benefit from this bridge. There are also many cases where ordinary and partial differential equations are better looked at from a stochastic point of view.

We will start out with some history about the equations and some mathematical background needed. This is later followed up by stating and proving the Kolmogorov equations. We finish off with some examples and applications of the Kolmogorov equations. The section about the Kolmogorov equations are for higher dimensions and most of the theory is taken from

Øksendal [5]. The section about applications of the Kolmogorov equation uses the one dimensional Kolmogorov equation.

## **1.1 History**

In 1931 Andrey Nikolaewich Kolmogorov (1903-1987) started working with continuous time Markov chains; more precisely he studied their transition probability density. The same year he introduced some very important partial differential equations. These equations are known under the names the Kolmogorov backward equation and the Kolmogorov forward equation. Both equations are parabolic differential equations of the probability density function for some stochastic process. However the backward is mostly used in context with expected values. The names, forward and backward, come from the fact that the equations are diffusion equations that has to be solved in a certain direction, forward or backward.

The man behind the equations, Andrey Kolmogorov, is of many people considered the greatest mathematician in the history of Russia and also one of the most brilliant mathematicians the world has ever seen. He was a man of many interests and with his creativity and sharp intellect he was able to contribute to many different areas of mathematics. Fortunately, probability theory was one of these interests. Kolmogorov was the man who put probability theory in the category of rigorous mathematics. Before Kolmogorov there were not much difference between probability and stochastic modeling and one could doubt the results in probability the same way one would doubt the correctness of a model. Kolmogorov changed all of this by building a rigorous foundation for probability theory to stand on. Not only did he build the foundation he also contributed with advanced results to the field.

The following quote is taken from the book Kolmogorov's Heritage in Mathematics [4] and describes how brilliant Kolmogorov really was: "Most mathematicians prove what they can, Kolmogorov was of those who prove what they want".

Kolmogorov was however not the first person to come across the Kolmogorov forward equation. It was first introduced in the contents of physics by Adriaan Fokker and Max Planck and known as the Fokker-Planck equation as a way to describe the Brownian motion of particles.

## **1.2 Objective**

The objective of this thesis is to examine the Kolmogorov backward and forward equation. We will start with some basic theory that is needed to prove both equations moving on to proving both equations and after that using the Kolmogorov equations in some special cases.

# Chapter 2

## Theory

In this section we are going to present some theory that is going to be used later in the thesis when proving Kolmogorov equations. Some theory used in Chapter 4, where the equations are used in special cases, are left out and recommended literature will be given to the interested reader instead.

As mentioned earlier, the Kolmogorov equations can be used as a tool to solve stochastic differential equations. It is therefore important to have some knowledge about these stochastic differential equations. For this purpose we need to know some basic stochastic calculus. It is this chapters objective to give some basic understanding of stochastic calculus with the goal to use it in solving stochastic differential equations. We will start off with defining some important results and properties like for example the Brownian motion, from which much of the theory is built. We will then define the stochastic integral and a very important lemma, Itô's lemma. After this we have the tools to define what a stochastic differential equation is.

### 2.1 Information

In this subsection we will introduce the notation  $\mathcal{F}_t^X$  which intuitively can be thought of as "The information generated by  $X$  up to time  $t$ ". If we can say if an event  $A$  has occurred or not based on the observations of the trajectory of  $X$  up to time  $t$  we say that

$$A \in \mathcal{F}_t^X.$$

This is the same as to say that  $A$  is  $\mathcal{F}_t^X$ -measurable. The same can be said for processes. If a process  $Y(t)$  can be completely determined given only observations of the trajectory of  $X$  up to time  $t$ , one can write

$$Y(t) \in \mathcal{F}_t^X.$$

If this is true for every  $t \geq 0$  we say that  $Y(t)$  is adapted to the filtration  $\{\mathcal{F}_t^X\}_{t \geq 0}$ .

To understand this concept mathematically one need some background in measure theory and sigma algebras. However what is important to remember is the intuitive interpretation. A mathematical definition is given in 2.2.1 taken from [3].

#### Definition 2.2.1

*Let  $\{X_t; t \geq 0\}$  be a random process, defined on the probability space  $(\Omega, \mathcal{F}, P)$ . We then define the sigma-algebra generated by  $X$  over the interval  $[0, t]$  by*

$$\mathcal{F}_t^X = \sigma\{X_s; s \leq t\}.$$

*Note:* One can sometimes write  $\mathcal{F}_t$  instead of  $\mathcal{F}_t^X$  since  $X$  is often implied.

## 2.2 Martingale

Martingale theory originates from betting strategies for making money on fair games. Here a fair game is defined as a game in which you have the same probability of winning as losing. The most common of these betting strategies is to go into a fair game with a starting bet, and for each consecutive loss you double the bet. This way you will regain all your lost money when you eventually win. Starting from this simple concept martingale theory evolved into an abstract mathematical theory which is one of the main tools in the theory of random processes. We will here present martingale theory in a simplified way since a rigorous presentation is out of the scope of this thesis. A way of thinking of this mathematically is

$$E[X_s | \mathcal{F}_t] = X_t,$$

which intuitively means that the expected future value of some stochastic process  $X_t$  (with the martingale property) based on the information at hand, is the present value.

## 2.3 Markov property

Some stochastic processes have the Markov property. This means that if you have a process and a complete description of the process at time  $t$ , then the future behavior of the process is independent of everything that happened before  $t$ . Theorem 2.3.1 is taken from [5] and tells us what the Markov property is.

### Theorem 2.3.1 (Markov property)

*Let  $f$  be a bounded Borel function from  $\mathbb{R}^n$  to  $\mathbb{R}$ . Then, for  $t, h \geq 0$*

$$E^x[f(X_{t+h}) | \mathcal{F}_t]_{(\omega)} = E^{X_t(\omega)}[f(X_h)]. \quad (e2.3.1)$$

## 2.4 Strong Markov property

The strong Markov property says that the Markov property still holds if the time  $t$  is replaced with a random time  $\tau$ , often called a stopping time or Markov time. The following theorem is taken from [5].

### Theorem 2.4.1 (The strong Markov property for Itô diffusions)

*Let  $f$  be a bounded Borel function on  $\mathbb{R}^n$ ,  $\tau$  is a stopping time with respect to  $\mathcal{F}_t^m$ ,  $\tau < \infty$  with probability 1. Then*

$$E^x[f(X_{\tau+h}) | \mathcal{F}_\tau^m] = E^{X_\tau}[f(X_h)] \quad \text{for all } h \geq 0. \quad (e2.4.1)$$



Here  $\mathcal{F}_t^m$  denotes the  $\sigma$ -algebra generated by  $\{W_{s\wedge t}; s > 0\}$ .

## 2.5 Brownian motion

Brownian motion, which is both Markov and Martingale, is arguably the most important stochastic process. It is commonly used in physical science and also several branches of social sciences. Its history speaks for the diversity of the process. Brownian motion was first introduced by the Scottish botanist Robert Brown after whom the process is named. Brown used the model to describe the motions of pollen grains suspended in water. After that Bachelier used the model in financial applications and Einstein found applications in physics for the model. The man who made the mathematical foundations for the model was Norbert Wiener after whom an alternative name of the standard Brownian motion is named. The alternative name is Wiener process.

An intuitive way of thinking of Brownian motion is imagining a game where you at each time step  $\Delta t$  flip a coin. If you get head you win  $\sqrt{\Delta t}$  dollar and if you get tail you lose  $\sqrt{\Delta t}$  dollar. The result of this is a random walk which has probability one half of going  $+\sqrt{\Delta t}$  dollar and probability one half to go  $-\sqrt{\Delta t}$  dollar. Brownian motion is now the case when the coin is tossed infinitely many times per second. The following function gives an intuitive description of a Brownian motion

$$f(t + \Delta t) = \begin{cases} f(t) + \sqrt{\Delta t}, & \text{with probability } \frac{1}{2} \\ f(t) - \sqrt{\Delta t}, & \text{with probability } \frac{1}{2} \end{cases}$$

The following definition is taken directly from [3] and gives a mathematical description of a standard Brownian motion.

### Definition 2.5.1 (Standard Brownian motion)

*The standard Brownian motion, or Wiener process,  $W$  is a stochastic process which satisfies the following conditions*

- 1)  $W(0) = 0$ .
- 2) *The process  $W$  has independent increments, i.e. if  $r < s \leq t < u$  then  $W(u) - W(t)$  and  $W(s) - W(r)$  are independent stochastic variables.*
- 3) *For  $s < t$  the stochastic variable  $W(t) - W(s)$  has the Gaussian distribution  $N[0, \sqrt{t - s}]$ .*
- 4)  *$W$  has continuous trajectories.*

*Note* From now on  $W$  will denote a standard Brownian motion.

## 2.6 Stochastic integration

To be able to work with stochastic differential equations we need to be able to integrate. Stochastic integration doesn't however work the same way as deterministic integration as we shall see later in one example. We are now going to define stochastic integration.

To define a stochastic integral we need a standard Brownian motion  $W$ , we also define  $g$  to be a stochastic process. Just like in the discrete case we need some conditions on  $g$  to make sure that it is integrable. It turns out that the class  $\mathcal{L}^2$  is natural. Definition 2.6.1 is taken from [3]

### Definition 2.6.1

i) We say that the process  $g$  belongs to the class  $\mathcal{L}^2[a, b]$  if the following conditions are satisfied.

$$*\int_a^b E[g^2(s)]ds < \infty.$$

\*The process  $g$  is adapted to the  $\mathcal{F}_t^W$ -filtration.

ii) We say that the process  $g$  belongs to the class  $\mathcal{L}^2$  if  $g \in \mathcal{L}^2[0, t]$  for all  $t > 0$ .

We are going to assume that all processes belong to the class  $\mathcal{L}^2$  from now on.

With  $g$  satisfying the conditions given in Definition 2.6.1 we can define a stochastic integral by

$$\int_0^t g(s)dW(s) = \lim_{n \rightarrow \infty} \sum_{i=1}^n g(t_{i-1})(W(t_i) - W(t_{i-1})). \quad (e2.6.1)$$

One thing that we can note when looking at this equation is that the integration is non-anticipatory. That is, we evaluate the integration at the left-hand point. From a financial point of view it makes sense since we don't want to use any information about the future in our calculations.

## 2.7 Stochastic Differential Equations

Stochastic differential equations have many applications. They are heavily used in finance but also in the natural sciences. For example they are used to model chemical reactions, systems in the human body, the dynamics of particles and systems in the human society like population growth and the dynamics of financial portfolios.

Stochastic differential equations are built up by a deterministic part and a random part. The random part is governed by Brownian motion, which was defined in Subsection 2.5. When one works with stochastic differential equations one often uses the following shorthand notation:

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad (e2.7.1)$$

$$X_0 = x_0 \quad (e2.7.2)$$

where  $W \in \mathbb{R}^d$  is a Wiener process,  $\mu: \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  is the so called drift function,  $x_0 \in \mathbb{R}^n$  is the initial condition and  $\sigma: \mathbb{R}_+ \times \mathbb{R}^n \rightarrow M(n, d)$  is the diffusion function.

What the equations (e2.7.1)-(e2.7.2) are shorthand for is

$$X_t = x_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s. \quad (e2.7.3)$$

$$X_0 = x_0 \quad (e2.7.4)$$

A way of thinking of a stochastic differential equation that might be helpful is

$$dX_t = \text{Deterministic} * dt + \text{Random} * dW_t.$$

To visualize this one can imagine a water molecule traveling down a river. It is driven by the flow of the water, the drift function. However this is not the whole truth, it is also affected by nearby molecules which our water molecule collide with. This later movement can be thought of as random with the diffusion function deciding how much the molecule is affected by this random movement.

We won't prove the uniqueness and existence theorem for solutions to the stochastic differential equations since the proof is out of the scope of this thesis. The result is however presented here and is taken from [3].

### Proposition 2.7.1

Suppose that there exist a constant  $K$  such that the following conditions hold for all  $x, y$  and  $t$ :

$$\|b(t, x) - b(t, y)\| \leq K\|x - y\|, \quad (e2.7.3)$$

$$\|\sigma(t, x) - \sigma(t, y)\| \leq K\|x - y\|, \quad (e2.7.4)$$

$$\|b(t, x)\| + \|\sigma(t, x)\| \leq K(1 + \|x\|). \quad (e2.7.5)$$

Then there exists a unique solution to the SDE (e2.2.1)-(e2.22) which has the properties

- 1)  $X$  is  $\mathcal{F}_t^W$ -adapted.
- 2)  $X$  has continuous trajectories.
- 3)  $X$  is a Markov process.
- 4) There exists a constant  $C$  such that

$$E[\|X_t\|^2] \leq C e^{Ct} (1 + \|x_0\|^2).$$

In this thesis we are often going to consider Itô diffusions. An Itô diffusion is a solution to a stochastic differential equation on the form

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

where  $W \in \mathbb{R}^d$  is a Wiener process,  $\mu: \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\sigma: \mathbb{R}_+ \times \mathbb{R}^n \rightarrow M(n, d)$ . Here both  $\mu$  and  $\sigma$  are Lipschitz continuous. Itô diffusions have many properties which makes them easy to work with. For example they both have the Markov property and the strong Markov property. They also have some other nice properties like for example the existence of an infinitesimal generator and Dynkin's formula. The two later properties are going to be explained in Subsection 2.9.

## 2.8 Itô's Lemma

Itô's lemma is one of the building blocks of stochastic calculus. It's a very useful tool in financial mathematics. Some even describe financial mathematics with the abbreviation RAIL, Relentless Applications of Itô's Lemma. Itô's lemma for one dimension looks like this:

Let

$$dS = \mu(S, t)dt + \sigma(S, t)dW_t, \quad (e2.8.1)$$

and let  $f$  be a function of  $S$  and  $t$ . We then have

$$df = \left( \frac{\partial f}{\partial t} + \frac{\mu \partial f}{\partial S} + \frac{1}{2} \sigma^2 \left( \frac{\partial^2 f}{\partial S^2} \right) \right) dt + \sigma \frac{\partial f}{\partial S} dW_t. \quad (e2.8.2)$$

As we can see this has many similarities to the Taylor expansion for ordinary calculus. One can in fact think of Itô's lemma as Taylor expansion in stochastic calculus. The following heuristic motivation for the formula also promotes this connection between Itô's lemma and Taylor expansion.

Let  $S$  be defined like in equation (e.2.8.1). Also let  $f$  be a function of  $S$  and  $t$ . Then the Taylor expansion of  $f$  would look like this

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S} (\mu dt + \sigma dW) + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} (\mu^2 (dt)^2 + 2\mu\sigma dt dW + \sigma^2 (dW)^2) + \text{"higher order terms that can be shown to go to zero"}.$$

We now note that  $(dW)^2$  is the same as  $dt$ . This can seem a little weird in the beginning but you can motivate yourself if you think about the properties of Brownian motion, in particular calculations done with the variance. It is actually a fact that  $(dW)^2 = dt$  and it can be proven mathematically. So if we now have convinced ourselves that  $(dW)^2 = dt$  we can see that  $dt dW$  is smaller than  $dt$  and  $dx$ . We can also see that  $(dt)^2$  is smaller than  $dt$ . In fact it is true that both  $dt dW$  and  $(dt)^2$  goes to zero. If we know collect the terms we get

$$df = \left( \frac{\partial f}{\partial t} + \frac{\mu \partial f}{\partial S} + \frac{1}{2} \sigma^2 \left( \frac{\partial^2 f}{\partial S^2} \right) \right) dt + \sigma \frac{\partial f}{\partial S} dW_t.$$

Which is exactly equation (e2.8.2) and we have now given a heuristic motivation for Itô's lemma.

We will make one example to show the difference between ordinary integration and stochastic integration.

**Example:**

$$\int_0^t W(s)dW(s), \quad (e2.8.3)$$

where  $W(s)$  is a wiener process. To solve this we can do as follows:

Let  $X = W$ ,  $f(t, x) = x^2$  and  $dX = 0dt + 1dW$ . Using Itô's formula we get

$$d(f(t, W(t))) = dt + 2W(t)dW(t).$$

Integrating this result we get the following expression for (e2.8.3)

$$\int_0^t W(s)dW(s) = \frac{W^2(t)}{2} - \frac{t}{2}. \quad (e2.8.4)$$

Here we also used the fact that  $W(0) = 0$ . If we now replace  $W(t)$  with a deterministic function  $f(t)$  we would get the following result (given that  $f(0) = 0$ )

$$\int_0^t f(s)df(s) = \int_0^t f(s)f'(s)ds = \frac{f(t)^2}{2}.$$

This is a good example of the difference between stochastic calculus and ordinary calculus. It also tells us not to always trust our intuition when we look at stochastic problems.

## 2.9 An important operator and Dynkin's formula

In this subsection we are going to introduce an operator that is used to associate a partial differential operator to an Itô diffusion. This link between a partial differential operator and an Itô diffusion is important in many applications. Most of this subsection is inspired or taken directly from [5].

We are going to start off with a definition of the operator followed by a useful lemma. This lemma is then used to get the formula for the operator of an Itô diffusion. We follow this up with stating and proving the important Dynkin's formula.

### Definition 2.9.1

Let  $\{X_t\}$  be a time-homogenous Itô diffusion in  $\mathbb{R}^n$ . The generator  $A$  of  $X_t$  is defined by

$$Af(x) = \lim_{t \rightarrow 0} \frac{E^x[f(X_t)] - f(x)}{t}, \quad x \in \mathbb{R}^n.$$

The set of functions  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  such that the limit exist at  $x$  is denoted by  $D_A(x)$  and  $D_A$  denotes the set of functions for which the limit exists for all  $x \in \mathbb{R}^n$ .

With this definition in mind one can find the formula for the generator  $A$  of an Itô diffusion. The result is given in Lemma 2.9.2:

**Lemma 2.9.2**

Let  $Y_t = Y_t^x$  be an Itô process in  $\mathbb{R}^n$  of the form

$$Y_t^x(\omega) = x + \int_0^t u(s, \omega) ds + \int_0^t v(s, \omega) dW_s(\omega)$$

where  $W$  is a  $m$ -dimensional Wiener process. Let  $f \in C_0^2(\mathbb{R}^n)$ , i.e.  $f \in C^2(\mathbb{R}^n)$  and  $f$  has compact support, and let  $\tau$  be a stopping time w.r.t.  $\{\mathcal{F}_t^m\}$ , and assume that  $E^x[\tau] < \infty$ . Assume that  $u(t, \omega)$  and  $v(t, \omega)$  are bounded on the set of  $(t, \omega)$  such that  $Y(t, \omega)$  belongs to the support of  $f$ . Then

$$E^x[f(Y_\tau)] = f(x) + E^x \left[ \int_0^\tau \left( \sum_i u_i(s, \omega) \frac{\partial f}{\partial x_i}(Y_s) + \frac{1}{2} \sum_{i,j} (vv^T)_{i,j}(s, \omega) \frac{\partial^2 f}{\partial x_i \partial x_j}(Y_s) \right) ds \right],$$

where  $E^x$  is the expectation w.r.t. the natural probability law  $R^x$  for  $Y_t$  starting at  $x$ :

$$R^x[Y_t \in F_1, \dots, Y_{t_k} \in F_k] = P^0[Y_{t_1}^x \in F_1, \dots, Y_{t_k}^x \in F_k], F_i \text{ Borel sets.}$$

*Proof.*

Put  $Z = f(Y)$  and apply Itô's formula ( $Y_1, \dots, Y_n$  and  $W_1, \dots, W_m$  denotes the coordinates of  $Y$  and  $B$ )

$$\begin{aligned} dZ &= \sum_i \frac{\partial f}{\partial x_i}(Y) dY_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(Y) dY_i dY_j = \sum_i u_i \frac{\partial f}{\partial x_i} dt \\ &\quad + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} (vdW)_i (vdB)_j + \sum_i \frac{\partial f}{\partial x_i} (vdW)_i. \end{aligned}$$

Since

$$(vdB)_i \cdot (vdB)_j = (vv^T)_{i,j} dt,$$

this gives

$$f(Y_t) = f(Y_0) + \int_0^t \left( \sum_i u_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j} (vv^T)_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} \right) ds + \sum_{i,k} \int_0^t v_{i,k} \frac{\partial f}{\partial x_i} dW_k. \quad (e2.9.1)$$

Hence

$$E^x[f(Y_\tau)] = f(x) + E^x \left[ \int_0^\tau \left( \sum_i u_i \frac{\partial f}{\partial x_i}(Y) + \frac{1}{2} \sum_{i,j} (vv^T)_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(Y) \right) \right] \\ + \sum_{i,k} E^x \left[ \int_0^\tau v_{i,k} \frac{\partial f}{\partial x_i}(Y) dW_k \right]. \quad (e2.9.2)$$

If  $g$  is a bounded Borel function,  $|g| \leq M$  say, then for all integers  $k$  we have

$$E^x \left[ \int_0^{\tau \wedge k} g(Y_s) dW_s \right] = E^x \left[ \int_0^k \chi_{\{s < \tau\}} g(Y_s) dW_s \right] = 0,$$

since  $g(Y_s)$  and  $\chi_{\{s < \tau\}}$  are both  $\mathcal{F}_s^m$ -measurable. Moreover

$$E^x \left[ \left( \int_0^\tau g(Y_s) dW_s - \int_0^{\tau \wedge k} g(Y_s) dW_s \right)^2 \right] = E^x \left[ \int_{\tau \wedge k}^\tau g^2(Y_s) ds \right] \leq M^2 E^x[\tau - \tau \wedge k] \rightarrow 0.$$

Therefore

$$0 = \lim_{k \rightarrow \infty} E^x \left[ \int_0^{\tau \wedge k} g(Y_s) dW_s \right] = E^x \left[ \int_0^\tau g(Y_s) dW_s \right].$$

Combining this with equation (e2.9.2) we get the proof of this lemma. ■

From Lemma 2.9.2 we also get the formula for the generator  $A$ . All we need to do is to replace  $\tau$  with  $t$  and use the definition of  $A$  given in Definition 2.9.1. This result is stated in Theorem 2.9.3 below.

### Theorem 2.9.3

Let  $X_t$  be an Itô diffusion

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t.$$

If  $f \in C_0^2(\mathbb{R}^n)$  then  $f \in D_A$  and

$$Af(x) = \sum_i b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j} (\sigma\sigma^T)_{i,j}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}. \quad (e2.9.3)$$

We are now going to look at Dynkin's formula. The formula is used frequently in stochastic analysis. It gives us the expected value of an Itô diffusion at a stopping time. To get an idea of how the formula works one may draw analogues to the second fundamental theorem of calculus. Dynkin's formula is given in theorem 2.9.4 and it is taken from [5].

### Theorem 2.9.4 (Dynkin's formula).

Let  $f \in C_0^2(\mathbb{R}^n)$ . Suppose that  $\tau$  is a stopping time and that  $E^x[\tau] < \infty$ . Then

$$E^x[f(X_\tau)] = f(x) + E^x\left[\int_0^\tau Af(X_s)ds\right]. \quad (e2.9.4)$$

*Proof.*

We look at equation (e2.9.2) in the proof of lemma 2.9.2 and combine this with equation (e2.9.3) in theorem 2.9.3. Noting that the last term in (e2.9.2) is zero we get Dynkin's formula. ■

## 2.10 The Fourier transform

In Chapter 4 we want to use the Kolmogorov equations to get information about some common stochastic processes. To solve the Kolmogorov equations we need some mathematical tools. One useful tool is the Fourier transform which will be explained briefly here. For a more detailed explanation see [1] and [2].

The Fourier transform of an integrable function  $f$  is denoted  $\hat{f}$  and given by

$$\hat{f}(\lambda) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i x \lambda} dx, \quad (e2.10.1)$$

where  $\lambda \in \mathbb{R}$ . If we instead want to have our original function  $f$  given that we have  $\hat{f}$  we can use the inverse Fourier transform

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\lambda)e^{2\pi i \lambda x} d\lambda, \quad (e2.10.2)$$

where  $x \in \mathbb{R}$ .

When we use the Fourier transform in Chapter 4 we are going to see that the Fourier transform for a normal distribution will show up. Here is the probability density function  $f$  for a normal distribution with mean  $\mu$  and variance  $\sigma^2$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (e2.10.3)$$

We now get that the Fourier transform of this normal distribution  $f$  is given by

$$\hat{f}(\lambda) = \int_{-\infty}^{\infty} f(x) e^{i\lambda x} dx = e^{i\mu\lambda - \frac{1}{2}\sigma^2\lambda^2}. \quad (e2.10.4)$$

As mentioned, this formula is used later in Chapter 4 where we use it to solve the forward Kolmogorov equation for some common processes.



## Chapter 3

### The Kolmogorov Equations

In this section we are going to state and prove the backward and forward Kolmogorov equation. The reason to why we want these equations in the first place is because they are a bridge between stochastic differential equations and partial differential equations. This bridge gives us a larger toolbox for solving both stochastic and partial differential equations. The names come from the fact that the equations need to be solved in different directions. The backward equation need to be solved backwards in time and the forward needs to be solved forward in time. The forward Kolmogorov equation is an equation of the probability density of some stochastic process and the backward equation is an equation for the expected value of some stochastic process.

#### 3.1 Kolmogorov Backward Equation

The following theorem taken from [5].

##### **Theorem 3.1.1 (The Kolmogorov backward equation)**

Let  $X_t$  be an Itô diffusion in  $\mathbb{R}^n$  satisfying the SDE

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \quad (e3.1.1)$$

with generator  $A$  defined in Theorem 2.9.3. We also let  $f \in C_0^2(\mathbb{R}^n)$ .

i) Define

$$u(t, x) = E^x[f(X_t)]. \quad (e3.1.4)$$

Then  $u(t, \cdot) \in D_A$  for each  $t$  and

$$\frac{\partial u}{\partial t} = Au, \quad t > 0, x \in \mathbb{R}^n \quad (e3.1.5)$$

$$u(0, x) = f(x); \quad x \in \mathbb{R}^n \quad (e3.1.6)$$

Here the right hand side is to be interpreted as  $A$  applied to the function  $x \rightarrow u(t, x)$ .

ii) Moreover if  $w(t, x) \in C^{1,2}(\mathbb{R} \times \mathbb{R}^n)$  is a bounded function satisfying equation (e3.1.5) and (e3.1.6) then  $w(t, x) = u(t, x)$ , given by equation (e3.1.4).

*Remark 1:* The partial differential equation (e.3.1.5)-(e3.1.6) is here written with an initial condition. Most often see the backward Kolmogorov equation with an end condition and a negative sign in front of the time derivative. We get both of these changes if we make the variable change from  $t$  to  $T - t$ . The equation would in this case look like

$$\begin{aligned} -\frac{\partial u}{\partial t} &= Au, & t > 0, x \in \mathbb{R}^n \\ u(T, x) &= f(x), & x \in \mathbb{R}^n \end{aligned}$$

which is how it is mostly seen.

*Remark 2:* If we use the equation we got in Remark 1 and make a slight change in the end time condition we get the following equation

$$\begin{aligned} -\frac{\partial u}{\partial t} &= Au, & t > 0, x \in \mathbb{R}^n \\ u(T, y) &= I_B(y), & x \in \mathbb{R}^n \end{aligned}$$

where  $I_B$  is the indicator function of the set  $B$ . We can know see that

$$u(r, y) = E_{r,y}[I_B(X_T)] = P(X_T \in B | X_r = y).$$

This shows us that we can in fact write the backward Kolmogorov equation for transition probabilities

$$\begin{aligned} -\frac{\partial P}{\partial r}(y, r; B, t) &= AP(y, r; B, t), & 0 < s < t, & \quad y \in \mathbb{R}^n \\ P(y, r; B, t) &= I_B(y), \end{aligned}$$

From here one can see that the equation can also be written in terms of transition probability densities. In this case we get the equation

$$\begin{aligned} -\frac{\partial p}{\partial r}(y, r; B, t) &= Ap(y, r; x, t), & 0 < s < t, & \quad y \in \mathbb{R}^n \\ p(y, r; B, t) &\rightarrow \delta_x \text{ as } s \rightarrow t. \end{aligned}$$

The reason why we want to look at the equation this way is because it is of the same form as the forward Kolmogorov equation. When you compare the two equations in this form it gets obvious that the backward equation uses the backward variables meanwhile the forward equation uses the forward variables.

Now we are moving on to the proof of Theorem 3.1.1 which also is inspired from [5].

*Proof:*

i) We introduce  $g(x) = u(t, x)$ . Since  $t \rightarrow u(t, x)$  is differentiable we have

$$\begin{aligned} \frac{E^x[g(X_r)] - g(x)}{r} &= \frac{1}{r} E^x[E^{X_r}[f(X_t)] - E^x[f(X_t)]] \\ &= \frac{1}{r} E^x[E^x[f(X_{t+r}) | \mathcal{F}_r] - E^x[f(X_t) | \mathcal{F}_r]] \\ &= \frac{1}{r} E^x[f(X_{t+r}) - f(X_t)] \end{aligned}$$

$$= \frac{1}{r} (u(t+r, x) - u(t, x)) \rightarrow \frac{\partial u}{\partial t}, \text{ as } r \rightarrow 0.$$

Hence we have that

$$Au = \lim_{r \rightarrow 0} \frac{E^x[g(X_r)] - g(x)}{r} \text{ exists and } \frac{\partial u}{\partial t} = Au, \text{ as asserted.}$$

In this line of thought we first used that  $g(x) = u(t, x) = E^x[f(X_t)]$ . We then used the Markov property followed by a result for conditional expectations.

- ii) To prove statement ii) we assume that a function  $w(t, x) \in C^{1,2}(\mathbb{R} \times \mathbb{R}^n)$  satisfies (e3.1.6) and (e3.1.7). Then

$$\tilde{A}w := -\frac{\partial w}{\partial t} + Aw = 0, \text{ for } t > 0, x \in \mathbb{R}^n, \quad (\text{e3.1.7})$$

and

$$w(0, x) = f(x), \quad x \in \mathbb{R}^n. \quad (\text{e3.1.8})$$

We fix  $(s, x) \in \mathbb{R} \times \mathbb{R}^n$ . Define the process  $Y_t$  in  $\mathbb{R}^{n+1}$  by  $Y_t = (s-t, X_t^{0,x})$ ,  $t \geq 0$ . Then  $Y_t$  has generator  $\tilde{A}$  and by (e3.1.8) and by Dynkin's formula, theorem (2.10.1), we have, for all  $t \geq 0$ ,

$$E^{s,x}[w(Y_{t \wedge \tau_R})] = w(s, x) + E^{s,x} \left[ \int_0^{t \wedge \tau_R} \tilde{A}w(Y_r) dr \right] = w(s, x),$$

where  $\tau_R = \inf\{t > 0; |X_t| \geq R\}$ . Letting  $R \rightarrow \infty$  we get

$$w(s, x) = E^{s,x}[w(Y_t)]; \quad \text{for all } t \geq 0.$$

In particular, choosing  $t = s$  we get

$$w(s, x) = E^{s,x}[w(Y_s)] = E[w(0, X_s^{0,x})] = E[f(X_s^{0,x})] = E^x[f(X_s)].$$

And the proof is completed. ■

### 3.2 Kolmogorov Forward Equation

We are going to prove to the forward Kolmogorov equation using the backward Kolmogorov equation. Again we let  $X_t$  be an Itô diffusion in  $\mathbb{R}^n$  so that it satisfies the stochastic differential equation (e3.1.1) with generator  $A$  defined in Theorem 2.9.2. We also assume that the transition measure of  $X_t$  has a density  $p(x, r; y, t)$ . Where the notation  $p(x, r; y, t)$  means the probability to be in  $y$  at time  $t$  given that we are at  $x$  at time  $r$ . i.e. that

$$E^{x,r}[f(X_t)] = \int_{\mathbb{R}^n} f(y)p(x,r;y,t)dy; \quad f \in C_0^2. \quad (e3.2.1)$$

We assume that  $y \rightarrow p(x,r;y,t)$  is smooth for each  $t, x$ .

The following theorem and proof is inspired from a problem in [5].

**Theorem 3.2.1 ( The forward Kolmogorov equation)**

Let  $p(x,r;y,t)$  be defined as above. Then  $p$  satisfies The forward Kolmogorov equation:

$$\frac{\partial}{\partial t}p(x,r;y,t) = A^*p(x,r;t,y), \quad (e3.2.2)$$

$$\lim_{t \rightarrow s^+} p(x,r;\cdot,t) = \delta_x, \quad (e3.2.3)$$

where the operator  $A^*$  is defined by

$$A^*f = - \sum_i \left( b_i \frac{\partial}{\partial x_i} f \right) + \frac{1}{2} \sum_{i,j} a_{i,j} \left( \frac{\partial^2}{\partial x_i \partial x_j} \right) f \quad (e3.2.4)$$

where

$$a_{i,j} = \sum_k \sigma_{i,k}(t,x)\sigma_{j,k}(x). \quad (e3.2.5)$$

We can here note that with respect to the  $L^2$  inner product,  $A^*$  is the dual of  $A$ .

*Proof:*

We fix  $T$  and let  $0 \leq r \leq t \leq T$  and  $f \in C_0^2(\mathbb{R}^n)$ . We also define

$$u(x,r) = E^{x,r}[f(X_T)] = \int_{\mathbb{R}^n} f(y)p(x,r;y,t)dy,$$

And hence  $u$  satisfies Kolmogorov's backward equation, namely equation (e3.1.5) and (e3.1.6). By the Markov property we get

$$u(x,r) = E^{x,r}[f(X_t)] = \int p(x,r;y,t)u(y,t)dy$$

for all  $t \in [s, T]$ . By differentiating in  $t$  and using the fact that  $u$  satisfies Kolmogorov's backward equation we get

$$\begin{aligned} 0 &= \frac{\partial u(x,r)}{\partial t} = \int_{\mathbb{R}^n} \left[ \frac{\partial}{\partial t} p(x,r;y,t)u(y,t) + p(x,r;y,t) \frac{\partial}{\partial t} u(y,t) \right] dy \\ &= \int_{\mathbb{R}^n} \left[ \frac{\partial}{\partial t} p(x,r;y,t)u(y,t) - p(x,r;y,t)A_y u(y,t) \right] dy \\ &= \int_{\mathbb{R}^n} \left[ \frac{\partial}{\partial t} p(x,r;y,t) - A_y^* p(x,r;y,t) \right] u(y,t) dy \end{aligned}$$

Here we used integration by parts together with the fact that all the boundary terms involved are 0 since  $p$  vanishes at infinity. We also used the fact that  $A$  and  $A^*$  are dual to each other with respect to the  $L^2$  inner product i.e.

$$\int_{\mathbb{R}^n} Af(x)g(x)dx = \int_{\mathbb{R}^n} f(x)A^*g(x)dx.$$

Now if we choose  $t = T$ , we get that

$$\int_{\mathbb{R}^n} \left[ \frac{\partial}{\partial t} p(x, r; y, T) - A_y^* p(x, r; y, T) \right] f(y) dy = 0$$

and since  $f \in C_0^2$  is arbitrary

$$\frac{\partial}{\partial t} p(x, r; y, T) - A_y^* p(x, r; y, T) = 0$$

and thus the proof is completed. ■

# Chapter 4

## Applications

In this subsection we are going to give examples on cases where both the Kolmogorov backward and Kolmogorov forward equations can be used. As mentioned earlier the equations are used in many different areas ranging from the natural sciences to the social sciences. One reason why they are so successful is because problems in the real world are often macroscopic systems that can depend on a huge amount of microscopic variables. To solve this macroscopic system we often have to solve all the microscopic variables in the system. This is generally an overwhelming task. Instead of doing this huge amount of calculations we can use a stochastic description of the problem. This is done by describing the system like a system of macroscopic variables that fluctuate. We will here look closer on some specific processes and problems, namely

- Scaled Brownian motion
- Scaled Brownian motion with drift
- Geometric Brownian motion
- The Ornstein-Uhlenbeck process
- Black-Scholes equation and its connection to the Kolmogorov equations
- Pricing a call option

### 4.1 Scaled Brownian motion

A solution to the following stochastic differential equation

$$\begin{aligned}dX_t &= \sigma dW_t, \\ X_0 &= 0,\end{aligned}$$

is called a scaled Brownian motion. The name gives us a good description of the process. A scaled Brownian motion is in fact just a Brownian motion with a coefficient in front of the random part. This means that the increments are scaled.

Scaled Brownian motion is somewhat special when it comes to the Kolmogorov equations. The forward and backward Kolmogorov equations are most often two completely different equations. However for a scaled Brownian motion the two equations look very much alike. The backward equation for a stochastic Brownian motion with end time  $t = T$  is

$$u_t + \frac{1}{2}\sigma^2 \Delta u = 0, \quad \text{for } t < T,$$

and the forward equation with starting time  $t = 0$  is

$$p_t - \frac{1}{2}\sigma^2\Delta p = 0, \quad \text{for } t > 0.$$

We can see that the equations are the same only with different directions in time.

Another interesting thing about a scaled Brownian motion is that the Kolmogorov equations are on the exact same form as the, in physics, well known heat equation. If we for example look at the forward equation we know that, because of the fact that it is the heat equation, the solution can be given by the Gaussian density

$$p(t, x) = \frac{1}{\sigma\sqrt{2\pi t}} e^{-\frac{1}{2}\frac{x^2}{\sigma^2 t}}.$$

## 4.2 Scaled Brownian motion with linear drift

The standard Brownian motion was explained earlier in Subsection 2.1 and as mentioned it can be viewed as a random walk centered on zero. Scaled Brownian motion with drift can be viewed as a scaled Brownian motion with a trend. This means that the center of the scaled Brownian motion varies with time. The stochastic differential equation for a scaled Brownian motion with drift,  $X_t$ , is given by

$$dX_t = \mu dt + \sigma dW_t \quad (e4.2.1)$$

$$X_0 = x_0 \quad (e4.2.2)$$

If we integrate this equation we get the following solution to (e4.2.1)-(e4.2.2)

$$X_t = x_0 + \mu t + \sigma(W_t - W_0).$$

This model is a standard example of a Brownian motion. One could imagine that it has many uses in finance with the volatility term and a drift term that could for example represent the market trend or inflation rate. However the process has one problem. It can take on negative numbers. This is a big problem since it isn't really a good realization if the model shows negative stock prices. A model that doesn't have this problem is the geometric Brownian motion. This is the most commonly used process in finance and we are looking more in to it in Subsection 4.3.

If we are interested in the probability density function of our Brownian motion with drift we can get it using the Kolmogorov equations. The forward Kolmogorov equation for (e4.1.1)-(e4.2.2) is the following

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

$$p(0, x) = \delta(x - x_0).$$

Now using the Fourier transform on this we get the following equation

$$\begin{aligned}\frac{\partial \hat{p}}{\partial t}(t, \lambda) &= -\mu i \lambda \hat{p}(t, \lambda) - \frac{1}{2} \sigma^2 \lambda^2 \hat{p}(t, \lambda), \\ \hat{p}(0, \lambda) &= e^{-i \lambda x_0}.\end{aligned}$$

From this we get

$$\begin{aligned}\hat{p}(t, \lambda) &= \hat{p}(0, \lambda) e^{-\mu i \lambda t - \frac{1}{2} \sigma^2 \lambda^2 t} = \\ \hat{p}(t, \lambda) &= e^{-i \lambda \mu t - i \lambda x_0 - \frac{1}{2} \sigma^2 \lambda^2 t}\end{aligned}$$

We recognize this as the Fourier transform for a normal distribution function (defined in (e2.10.4)) with mean  $\mu t + x_0$  and variance  $\sigma^2 t$  we get the probability distribution function

$$p(t, x) = \frac{1}{\sigma \sqrt{2\pi t}} e^{-\frac{(x - (\mu t + x_0))^2}{2\sigma^2 t}}.$$

### 4.3 Geometric Brownian motion

As mentioned earlier the geometric Brownian motion is the most commonly used process for modeling in financial mathematics. There are several reasons why this is the case. The process can't take on negative numbers, which agrees with the reality of stock prices. The movement of a geometric Brownian motion also agrees well with the movement you see in stock prices. The expected return of stock prices is independent of the price the process currently has. This is also the case for geometric Brownian motion. Last and not least the model is easy to work with mathematically. Geometric Brownian motion is however not perfect for modeling stock prices. It has some down sides. Geometric Brownian motion has normally distributed returns; this is not the case in reality. In reality there is also a higher chance for large changes than it is for geometric Brownian motion. Another simplification is that a geometric Brownian motion has constant volatility meanwhile this is not the case in reality.

If a stochastic process  $X_t$  that satisfies the following stochastic differential equation

$$dX_t = \mu X_t dt + \sigma X_t dW_t, \quad (e4.3.1)$$

$$X_0 = x_0, \quad (e4.3.2)$$

where  $W_t$  is a Wiener process, then  $X_t$  is called a geometric Brownian motion. If we take a closer look at the geometric Brownian motion we can see that it has an exponential behavior. The logarithm of the process actually follows a Brownian motion with drift. This motivates an alternative name for the process, the exponential Brownian motion.

We now want to find a solution to the geometric Brownian motion given by equation (4.3.1) and (4.3.2). To do this we are going to use Itô's lemma and we let the exponential behavior of the geometric Brownian motion motivate us to investigate the following process



$$Z_t = \ln(X_t). \quad (e4.3.3)$$

Here  $X_t$  is assumed to be strictly positive. If we apply Itô's lemma to the process  $Z_t$  given in (e4.3.3) we get

$$\begin{aligned} dZ &= \left( 0 + \frac{\mu X}{X} - \frac{1}{2} \frac{\sigma^2 X^2}{X^2} \right) dt + \frac{\sigma X}{X} dW_t \\ dZ &= \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t. \end{aligned} \quad (e4.3.4)$$

This equation is shorthand for an integral formula. If we write out the equation we get

$$\ln(X_t) = \ln(X_0) + \left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t.$$

Taking the exponential of this equation we get

$$X_t = x_0 \exp \left( \left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right). \quad (e4.3.5)$$

This is however not enough to prove that  $X_t$  is the solution of (e4.3.1) and (e4.3.2). We must first assume that there exists a positive solution  $X_t$ . However if we view this derivation as heuristic and then show that this solution in fact solves the stochastic differential equation (e4.3.1)-(e4.3.2) we can avoid this problem.

If we want to know the probability distribution function for our geometric Brownian motion we can use the same approach as we did in Subsection 4.2. However the Kolmogorov equation for the geometric Brownian motion is a bit messier to solve this way. One can here instead use the fact that the logarithm of the geometric Brownian motion follows a scaled Brownian motion with drift. We can see that if  $X_t$  follows the geometric Brownian motion defined in (e4.3.1)-(e4.3.2). Then  $\ln(X_t)$  follows the scaled Brownian motion with drift given in (e4.3.4). We now use what we found out in Section 4.2 about the probability distribution function for scaled Brownian motion with drift and that  $P(X_t < x) = P(Y < \ln(x))$ . What we get is that the probability density function looks like

$$p(t, x) = \frac{1}{x\sigma\sqrt{2\pi t}} e^{-\frac{(\ln(\frac{x}{x_0}) - (\mu - \frac{1}{2}\sigma^2)t)^2}{2\sigma^2 t}}$$

#### 4.4 Ornstein-Uhlenbeck process

The process  $X_t$  generated by the following stochastic differential equation

$$dX_t = -\mu X_t dt + \sigma dW_t \quad (e4.4.1)$$

$$X_0 = x_0 \quad (e4.4.2)$$

is called an Ornstein-Uhlenbeck process. To find an analytic solution for this equation we introduce the factor  $e^{\mu t}$  and investigate the process  $e^{\mu t} X_t$ . We get

$$\begin{aligned} d(e^{\mu t} X_t) &= e^{\mu t} dX_t + X_t d(e^{\mu t}) = \\ e^{\mu t} (-\mu X_t dt + \sigma dW_t) + \mu X_t e^{\mu t} dt &= \sigma e^{\mu t} dW_t. \end{aligned}$$

Now if we integrate from 0 to  $t$  and integrate by parts we get

$$\begin{aligned} e^{\mu t} X_t &= x_0 + \int_0^t \sigma e^{\mu s} dW_s \\ X_t &= x_0 e^{-\mu t} + \sigma \int_0^t e^{\mu(s-t)} dW_s \\ X_t &= x_0 e^{-\mu t} + \sigma \left( W_t - \mu \int_0^t e^{\mu(s-t)} W_s ds \right). \end{aligned}$$

Once again we can get probability density function for the solution to the stochastic differential equation using the forward Kolmogorov equation. The forward Kolmogorov equation for Ornstein-Uhlenbeck process looks like this

$$\begin{aligned} \frac{\partial p}{\partial t} &= \mu \frac{\partial}{\partial x} (xp) + \frac{1}{2} \sigma^2 \frac{\partial^2 p}{\partial x^2} \\ p(0, x) &= \delta(x - x_0). \end{aligned}$$

We now use the Fourier transform and get

$$\frac{\partial \hat{p}}{\partial t} + \mu \lambda \frac{\partial \hat{p}}{\partial \lambda} = -\frac{1}{2} \sigma^2 \lambda^2 \hat{p}$$

We now use the method of characteristics. Information about the method of characteristics can be found in [1] and [2]. Using this method it leads us to suppose that

$$\frac{\partial \hat{p}}{\partial t} = \frac{\partial \hat{p}}{\partial t} + \frac{\partial \hat{p}}{\partial \lambda} \frac{d\lambda}{dt} = \frac{\partial \hat{p}}{\partial t} + \mu \lambda \frac{\partial \hat{p}}{\partial \lambda},$$

from this we get that

$$\frac{\partial \lambda}{\partial t} = \mu \lambda,$$

which gives us that

$$\lambda = \lambda(0) e^{\mu t}.$$

We get

$$\frac{\partial \hat{p}}{\partial t} = -\frac{1}{2} \sigma^2 \lambda^2 \hat{p} = -\frac{1}{2} \sigma^2 \lambda(0)^2 e^{2\mu t} \hat{p}.$$

Since we know that

$$\hat{p}(0, \lambda) = e^{-i\lambda x_0}$$

we get

$$\begin{aligned} \hat{p}(t, \lambda) &= \hat{p}(0, \lambda(0)) e^{-\int_0^t \frac{1}{2} \sigma^2 \lambda(0)^2 e^{2\mu s} ds} = \\ &= e^{-i\lambda x_0 e^{-\mu t} - \frac{1}{2} \frac{\sigma^2}{2\mu} e^{-2\mu t} (e^{2\mu t} - 1)}. \end{aligned}$$

Now we recognize this to be the same form as the Fourier transform of the probability density function with mean  $x_0 e^{-\mu t}$  and variance  $\sigma^2 e^{-2\mu t} \frac{1}{2\mu} (e^{2\mu t} - 1)$ . What we get when using the inverse Fourier transform is then

$$p(t, x) = \frac{1}{\sqrt{2\pi \frac{\sigma^2}{2\mu} e^{-2\mu t} (e^{2\mu t} - 1)}} e^{-\frac{(x - x_0 e^{-\mu t})^2}{2 \frac{\sigma^2}{2\mu} e^{-2\mu t} (e^{2\mu t} - 1)}}.$$

#### 4.5 Black-Scholes equation and its connection to the backward Kolmogorov equation

The equation

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \left( \frac{\partial^2 V}{\partial S^2} \right) + rS \left( \frac{\partial V}{\partial S} \right) - rV = 0, \quad (e4.5.1)$$

where  $V$  is the price of some derivative,  $S$  is the price of some stock,  $r$  is the risk free interest rate and  $\sigma$  the volatility of the stock returns, is known as the Black-Scholes equation. This is an equation that models the classical market. The model comes from the following assumptions:

- i) *There are no arbitrage opportunities, which mean that you can't make a riskless profit.*
- ii) *The asset prices follow a geometric Brownian motion.*
- iii) *There exist a constant risk-free interest rate at which you can borrow and lend cash.*
- iv) *You can buy and sell any amount of stock, fractions of stocks included.*
- v) *The market is frictionless, which means that you can borrow and lend cash as well as buy or sell stocks without fees.*
- vi) *Now dividends are paid during the time period of which we are calculating.*

This model is heavily used in finance to describe the prices of options over time. One main purpose of the model is also to hedge the option, i.e. lowering risk. There is a connection between this very important model and the backward Kolmogorov equation.

Imagine for a while that we have bought a stock following a geometric Brownian motion with known  $\mu$  and  $\sigma$  and want to know how much money we are going to make. What we need is the probability density function  $p(x, t; x', t')$ . We assume that the stock is following a geometric Brownian motion, so we know that

$$dS = \mu S dt + \sigma S dW_t \quad (e4.5.2)$$

To get this probability density function we are going to use the backward Kolmogorov equation

$$\frac{\partial p}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 p}{\partial x^2} + \mu \frac{\partial p}{\partial x} = 0. \quad (e4.5.3)$$

As we can see this already looks a lot like Black-Scholes equation. If we know want to price an option using this equation we introduce the option value  $V$ . We let  $p(S, T)$  be the payoff at T. Now, discounting from the payoff at T we get

$$V(S, t) = e^{-r(T-t)} p(S, t). \quad (e4.5.4)$$

Using the backward Kolmogorov equation again we get that V satisfies

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \mu S \frac{\partial V}{\partial S} - rV = 0. \quad (e4.5.5)$$

This looks even more like Black-Scholes equation. The only difference is that  $\mu$  should equal  $r$ . So the process we should look at to get the Black-Scholes equation from the backward Kolmogorov equation is the risk-neutral random walk given by the stochastic differential equation

$$dS_t = r S_t dt + \sigma S_t dW_t. \quad (e4.5.5)$$

The risk-neutral random walk is given in equation (e4.5.5) is the random walk that is used to satisfy the Black-Scholes model. If we had another value on  $\mu$  than  $r$  we would allow arbitrage opportunities.

## 4.6 Call Option

A call option is one of the most basic forms of options there is. It gives the buyer of the call option the right to buy a particular asset for a certain price at a predefined time in the future. In our case we say that we have the right to buy the stock at time T for K. The payoff of the option is given by

$$V(S, T) = \max(S - K, 0), \quad (e4.6.1)$$

where  $S$  is the value of the underlying asset following the risk-neutral model

$$dS_t = rS_t + \sigma S_t dW_t. \quad (e4.6.2)$$

The probability transition density,  $p(S, t; S', t')$  (reads probability to be at  $S'$  at time  $t'$  given that we are at  $S$  at time  $t$ ) have the value of the payoff at time  $T$ ,

$$V(S, T) = p(S, T). \quad (e4.6.3)$$

If we look at the times before  $T$  we can discount from  $p$  and get the expected value of the option.

$$V(S, t) = e^{-r(T-t)}p(S, t). \quad (e4.6.4)$$

We can write this as

$$p(S, t) = e^{r(T-t)}V(S, t). \quad (e4.6.5)$$

We also know that  $p$  satisfies the backward Kolmogorov equation. If we substitute (e4.6.5) in to the backward Kolmogorov equation we get

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0. \quad (e4.6.6)$$

This we recognize as the Black-Scholes equation. We have now reduced the problem of pricing the call option to solving the partial differential equation (e4.6.6). One way of solving this partial differential equation is to transform it to the heat equation. We do this by making the following transformations:

$$S = Ke^x, \quad (e4.6.7)$$

$$t = T - \frac{\tau}{\frac{1}{2}\sigma^2}, \quad (e4.6.8)$$

$$k = \frac{r}{\frac{1}{2}\sigma^2}, \quad (e4.6.9)$$

$$V = Ke^{-\frac{1}{2}(k-1)x - \frac{1}{4}(k+1)^2\tau} u(x, \tau). \quad (e4.6.10)$$

Now substituting (e4.6.10) into (e4.6.6) we get

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < \infty, \tau > 0. \quad (e4.6.11)$$

This is actually a famous equation known as the heat equation. The heat equation is used to model diffusions, for example heat conduction. Because of this there exists much theory and many tools to solve this equation. The solution to this equation is therefore already known, we get

$$u(x, \tau) = \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} u_0(y) e^{-\frac{(x-y)^2}{4\tau}} dy, \quad (e4.6.12)$$

where  $u_0(x)$  is the initial condition  $u(x, 0)$ . We can get the initial condition  $u_0(x)$  from equation (e4.6.10). We get

$$u_0(x) = \frac{1}{K} e^{\frac{1}{2}(k-1)x} V(S, T). \quad (e4.6.13)$$

Since we know  $V(S, T)$  for our option is given by equation (e4.6.1). Equation (e4.6.13) together with (e4.6.1) gives us

$$u_0(x) = \max\left(e^{\frac{1}{2}(k+1)x} - e^{\frac{1}{2}(k-1)x}, 0\right). \quad (e4.6.14)$$

We now make the following change of variable

$$z = \frac{y - x}{\sqrt{2\tau}}, \quad (e4.6.15)$$

put into equation (e4.6.12) gives us

$$u(x, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\pi}}}^{\infty} e^{\frac{1}{2}(k+1)(z\sqrt{2\tau}+x) - \frac{1}{2}z^2} dz - \frac{1}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\pi}}}^{\infty} e^{\frac{1}{2}(k-1)(z\sqrt{2\tau}+x) - \frac{1}{2}z^2} dz. \quad (e4.6.16)$$

This reminds us of the cumulative distribution function for a normal distribution, which motivates us to rewrite equation (e4.6.16) to

$$\begin{aligned} u(x, \tau) = & \frac{e^{\frac{1}{2}(k+1)x + \frac{1}{4}(k+1)^2\tau}}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\pi}}}^{\infty} e^{-\frac{1}{2}\left(z - \frac{1}{2}(k+1)\sqrt{2\tau}\right)^2} dz \\ & - \frac{e^{\frac{1}{2}(k-1)x + \frac{1}{4}(k-1)^2\tau}}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\pi}}}^{\infty} e^{-\frac{1}{2}\left(z - \frac{1}{2}(k-1)\sqrt{2\tau}\right)^2} dz. \quad (e.4.6.17) \end{aligned}$$

This is now exactly of the form of the cumulative distribution function for a normal distribution. Since we know a lot of this distribution we can show that it equals

$$u(x, \tau) = e^{\frac{1}{2}(k+1)x + \frac{1}{4}(k+1)^2\tau} N(d_1) - e^{\frac{1}{2}(k-1)x + \frac{1}{4}(k-1)^2\tau} N(d_2), \quad (e4.6.18)$$

where  $d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$  and  $d_2 = \frac{\ln\left(\frac{S}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$ .

Why we got these values of  $d_1$  and  $d_2$  requires an inequality. If you are interested in the details they can be found in [3].

## Referenses

- [1] Yehuda Pinchover and Jacob Rubinstein, *An introduction to Partial Differential Equations*. Cambridge University Press, (2005).
- [2] Richard Haberman, *Applied Partial Differential Equations: with Fourier Series and Boundary Value Problems*. Pearson Prentice Hall, Upper Saddle River, New Jersey, 4<sup>th</sup> Edition, (2004)
- [3] Tomas Björk, *Arbitrage Theory in Continuous Time*. Oxford University Press Inc, New York, 3<sup>rd</sup> Edition, (2009).
- [4] Erik Charpentier, Annick Lesne, Nikolaï Nikolski, *Kolmogorov's Heritage in Mathematics*. Springer-Verlag, Berlin, (2007).
- [5] Bernt Øksendal, *Stochastic Differential Equations: An introduction with Applications*. Springer-Verlag, Berlin, 6<sup>th</sup> Edition, (2007).
- [6] Hannes Risken, *The Fokker-Planck Equation*. Springer-Verlag, Berlin Heidelberg, (1989).
- [7] Paul Wilmott, Sam Howison and Jeff Dewynne, *The MATHEMATICS of FINANCIAL DERIVATIVES: A Student Introduction*, The Press Syndicate of the University of Cambridge, New York, (1996).