Utrecht University

BACHELOR THESIS

MATHEMATISCH INSTITUUT

Brownian motion and Option pricing

Author
Ben BALKENENDE
5966698

Supervisor
Dr. K. DAJANI

11th June 2018

Contents

Introduction

An assumption which is made often in economic theory is the no-arbitrage principle or the law of one price. This assumption states that when two goods are equivalent they should have the same price. This is the basis for option pricing.

An option gives the buyer the right to buy or sell a stock at a certain price K sometime in the future. If the buyer wants to exercise the option at a time T than the profit depends on the value of the underlying stock S_T. However, the future value of the stock is uncertain and therefore probability theory is needed. In this thesis we will give a strategy to apply the law of one price in order to determine the right price for an option. We will focus on European options.

This thesis is divided in three parts. The first part gives the mathematical background needed for the later chapters. In chapter 1 we will shortly discuss concepts from measure theory. In chapter 2 we will give a measure theoretic approach to conditional expectation and introduce stochastic processes and some of their properties.

The second part gives an introduction to stochastic calculus. This will be our toolkit for the option pricing problem. In chapter 3 we introduce the Brownian motion and discuss some properties, using reference [2] chapters 7 and 8 as our guideline. In chapter 4 we introduce the Itô-integral and discuss the martingale property. This chapter is is based on [2] chapter 10, and [9] chapter 3. In chapter 5 we introduce the Itô formula and give an outline for the proof. In chapter 6 we will give an introduction to stochastic differential equations. In which we will focus on the existence and uniqueness theorem. This chapter is based on [2] section 12.1.

The third part we will discuss the option pricing problem. In chapter 7 we will shortly introduce terms from finance and explain the law of one price in more detail. Then we will construct a portfolio such that becomes risk free and then use the law of one price to derive a general solution for the option pricing problem. This chapter is based on [2] section 16.2. In chapter 8 we will give an explicit solution for the option pricing formula for a European call option. This chapter is based on [2] section 16.3

Part I

Measure theory and probability

Chapter 1

Results from measure theory

In this chapter we will shortly discuss some definitions and theorems from measure theory which will be used during this paper. We will discuss the notion of a measure space, measurable functions and integrals. We will also translate these notions to the language of probability. For a more in depth analysis see for example [1]. If one is already familiar with these topics he/she can skip this chapter and move to chapter 2.

1.1 Measure spaces

In this section we will let X be an arbitrary set.

Definition 1.1.1: Let X be a set. A σ-algebra is a collection $\mathcal{F} \subset \mathcal{P}(X)$ of subsets of X which satisfies the following conditions

(i) $X \in \mathcal{F}$

(ii) If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$

(iii) If $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of sets in $\mathcal{F}$ then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$.

The pair $(X, \mathcal{F})$ is called a measurable space. The sets $A \in \mathcal{F}$ are often called measurable sets. We will 'measure' these sets in definition 1.1.3.

In probability theory the notion of a Borel measurable set is often used.

Definition 1.1.2: Let τ_{Eucl} be the topology induced by the Euclidean metric on $\mathbb{R}$. Let $\mathcal{O}$ be the collection of all open sets generated by τ_{Eucl}. Then the Borel-σ-algebra is the smallest σ-algebra on X which contains the collection $\mathcal{O}$. The Borel-σ-algebra is denoted by $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{O})$.

Example 1.1.1:
For every $x \in \mathbb{R}$ the set $(-\infty, x]$ belongs to $\mathcal{B}(\mathbb{R})$. To show this we begin by noting that for every $n \in \mathbb{N}_{\geq 1}$ we have $(-\infty, x - 1/n) \in \mathcal{O} \subset \mathcal{B}(\mathbb{R})$. Then by (ii) in definition 1.1.1 we must have $[x - 1/n, \infty) \in \mathcal{B}(\mathbb{R})$ for all $n \geq 1$. Now we can complete our argument

$$(-\infty, x] = \bigcap_{n=1}^{\infty} \left(-\infty, x - \frac{1}{n}\right) = \left(\bigcup_{n=1}^{\infty} \left[x - \frac{1}{n}, \infty\right)\right)^c \overset{(ii)/(iii)}{\in} \mathcal{B}(\mathbb{R})$$

♦

Definition 1.1.3: Let $(X,\mathcal{F})$ be a measurable space. A measure on $\mathcal{F}$ is a function $\mu : \mathcal{F} \to [0,\infty]$ which satisfies the following conditions

(i) $\mu(\varnothing) = 0$

(ii) Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of disjoint sets. Then

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$$

The triple $(X,\mathcal{F},\mu)$ is called a measure space.

Example 1.1.2:
Consider the measurable space $(\mathbb{N},\mathcal{P}(\mathbb{N}))$. On this space we can for example define the counting measure μ which is defined as

$$\mu(A) = |A|$$

Clearly $|\varnothing| = 0$ therefore $\mu(\varnothing) = 0$ and thus μ satisfies (i) from definition 1.1.3. Now let $\{A_j\}_{j\in\mathbb{N}}$ be a sequence of disjoint sets. Then

$$\mu\left(\bigcup_{j\in\mathbb{N}} A_j\right) = \left|\bigcup_{j\in\mathbb{N}} A_j\right| = \sum_{j\in\mathbb{N}} |A_j| = \sum_{j\in\mathbb{N}} \mu(A_j)$$

Thus μ also satisfies (ii) and therefore we conclude μ defines a measure.

♦

For this thesis we will mostly consider probability spaces. A probability space is a measure space $(\Omega,\mathcal{F},\mathbb{P})$ with $\mathbb{P}$ a probability measure i.e. $\mathbb{P}(\Omega) = 1$. The elements $A \in \mathcal{F}$ are called events.

1.2 Measurable functions and Integrals

In this section we will give a short introduction to measurable functions and integrals of these functions.

Definition 1.2.1: Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space. A function $X : \Omega \to \mathbb{R}$ is called measurable if for every $B \in \mathcal{B}(\mathbb{R})$ we have $X^{-1}(B) \in \mathcal{F}$. In the language of probability we call a measurable function a *random variable*.

Theorem 1.2.1: Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a random variable defined on this probability space. Then the function

$$\mathbb{P}_X(B) := \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\}) \quad \forall B \in \mathcal{B}(\mathbb{R})$$

defines a probability measure.

Proof: See for example [1] theorem 7.6. □

From definition 1.2.1 it follows that the value a random variable will take is uncertain. In many situations one may want to know what value to expect. This value is called the expected value of the random variable and is denoted by $\mathbb{E}[X]$. In elementary probability courses a distinction is made between continuous and discrete random variables. However, a more general definition exists.

Definition 1.2.2: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\{A_i\}_{i=1}^n$ be a sequence of disjoint subsets of Ω and let $\{a_i\}_{i=1}^n$ be a sequence of real valued non-negative numbers. Define $X := \sum_{i=1}^n a_i \mathbf{1}_{A_i}$. Then the integral of X with respect to the probability measure $\mathbb{P}$ equals

$$\mathbb{E}[X] = \int_\Omega X d\mathbb{P} := \sum_{i=1}^n a_i \mathbb{P}(A_i) = \sum_{i=1}^n a_i \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = a_i\}) = \sum_{i=1}^n a_i \mathbb{P}_X(\{a_i\})$$

Note that this definition equals the well known definition for the expected value of a discrete random variable. Now we will look at how this will generalize the expected value. A random variable which is a linear combination of indicator functions is called a *simple* function. It can easily been shown that every simple function is measurable. Now let $X : \Omega \to [0, \infty]$ be a non-negative function on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Now we will define the integral, hence the expected value.

Definition 1.2.3: The expected value of a non-negative function

$$\mathbb{E}[X] = \int_\Omega X d\mathbb{P} := \sup\left\{\int_\Omega \varphi d\mathbb{P} \mid \varphi \text{ a simple function }, \varphi \leq X\right\}$$

Theorem 1.2.2: Let $X \to [0, \infty]$ be a simple random variable. Then definition 1.2.2 and 1.2.3 are equivalent.

Proof: See for example [4] corollary 2.22

□

Now we can define the expected value for a random variable $X : \Omega \to \mathbb{R}$. Define $X^+ := \max(X, 0)$ and $X^- := \max(-X, 0)$. It can be shown that these are measurable functions. See for example [4] proposition 2.18

Definition 1.2.4: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a random variable. The expected value of X exists if and only if

$$\int_\Omega X^+ d\mathbb{P} < \infty \quad \text{or} \quad \int_\Omega X^- d\mathbb{P} < \infty$$

In which case we define the expected value of X as

$$\mathbb{E}[X] = \int_\Omega X d\mathbb{P} := \int_\Omega X^+ d\mathbb{P} - \int_\Omega X^- d\mathbb{P}$$

A random variable X is said to be integrable if $\mathbb{E}[|X|] < \infty$.

Definition 1.2.5: The set $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ with $p \in [1, \infty)$ is the set of all random variables $X : \Omega \to \mathbb{R}$ such that

$$\mathbb{E}[|X|^p] = \int_\Omega |X|^p d\mathbb{P} < \infty$$

On this set we define the semi-norm

$$||X||_p := \sqrt[p]{\mathbb{E}[|X|^p]}$$

To see that this is indeed a semi-norm see for example [1] remark 13.5(ii)

Remark 1.2.1: Suppose we have a sequence of random variables $\{X_j\}_{j \in \mathbb{N}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that $\{X_j\}_{j \in \mathbb{N}}$ converges to a random variable X in the $\mathcal{L}^p$ sense if $\lim_{j \to \infty} ||X_j - X||_p = 0$.

1.3 Theorems

In this section we will state some important theorems from measure theory which will be used during this thesis.

Definition 1.3.1: Let $(X,\mathcal{F})$ be a measurable space. Let μ,ν be measures defined on this space. If for every $A \in \mathcal{F}$ we have $\nu(A)=0$ if $\mu(A)=0$ then ν is called absolutely continuous with respect to μ. The notation for this is $\nu \ll \mu$

Definition 1.3.2: Let $(X,\mathcal{F},\mu)$ be a measure space. The measure μ is called σ-finite if there exists a sequence $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{F}$ such that $\mu(A_j)<\infty$ for every $j \in \mathbb{N}$ and $\bigcup_{j\in\mathbb{N}} A_j = X$.

Theorem 1.3.1 (Radon-Nikodym): Let $(X,\mathcal{F})$ be a measurable space and let μ,ν be σ-finite measures on this space such that $\nu \ll \mu$. Then there exist a $\mathcal{F}$-measurable function $f : X \to [0,\infty)$ such that

$$\nu(A) = \int_A f d\mu, \quad \forall A \in \mathcal{F}$$

Proof: See for example [4] theorem 5.11

□

The function f is called the Radon-Nikodym theorem of ν with respect to μ. The Radon-Nikodym derivative is often notated as $\frac{d\nu}{d\mu}$.

Theorem 1.3.2 (Tonelli): Let $(X,\mathcal{F},\mu)$ and $(Y,\mathcal{B},\nu)$ be σ-finite measure spaces and let $(X\times Y,\mathcal{F}\otimes\mathcal{B},(\mu\times\nu))$ be the product measure space. Now let $u : X\times Y \to [0,\infty]$ be $\mathcal{F}\otimes\mathcal{B}$ measurable. Then

$$\int_Y \int_X u(x,y)\mu(dx)\nu(dy) = \int_X \int_Y u(x,y)\nu(dy)\mu(dx)$$

Proof: See for example [1] theorem 13.8.

□

Theorem 1.3.3 Dominated convergence theorem:
Let $(X,\mathcal{F},\mu)$ be a measure space and let $g : X \to [0,\infty]$ be a function. Let f and (f_n) be $[-\infty,\infty]$-valued $\mathcal{F}$ measurable functions on X such that

$$f(x) = \lim_{n\to\infty} f_n(x) \text{ and } |f_n(x)| \le g(x)$$

for all $n \in \mathbb{N}$ and μ almost every $x \in X$ then

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu$$

Proof:
See for example [10] theorem 2.4.5.

□

Chapter 2

Results from probability theory

In this chapter we will discuss some results from probability theory which will be used during this paper. In the first section we will discuss the notion of conditional expectation in a more abstract way. Then we will discuss the notion of a stochastic process in which we will mainly focus on martingales.

2.1 Conditional expectation

In this section we will define the notion of conditional expectation. We will expand on the known definition. We will discuss the conditional expectation with respect to an event and with respect to a σ-algebra. Then we will discuss conditional expectation with respect to a random variable and the connection between the definitions.

Definition 2.1.1: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A \in \mathcal{F}$ be an event with positive measure. Let $X : \Omega \to \mathbb{R}$ be a random variable. Then the conditional expectation of X given A is defined by

$$\mathbb{E}[X|A] = \frac{\int_A X d\mathbb{P}}{\mathbb{P}(A)} = \frac{\mathbb{E}[\mathbf{1}_A X]}{\mathbb{P}(A)}$$

Definition 2.1.2: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a measurable function. Let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra. Then the conditional expectation of X given $\mathcal{G}$ is defined by a $\mathcal{G}$ measurable function Y such that

$$\int_B Y d\mathbb{P} = \int_B X d\mathbb{P} \quad \forall B \in \mathcal{G}$$

The conditional expectation is denoted by $\mathbb{E}[X|\mathcal{G}]$.

Theorem 2.1.1: The conditional expectation Y in definition 2.1.2 exists and is unique almost surely.

Proof:
We have the following proof from [2] theorem 5.2. Let X be a random variable. Next we will define $X^+ := \max(X, 0)$ and $X^- := \max(-X, 0)$. Then $X = X^+ - X^-$. Then for a sub-σ-algebra $\mathcal{G}$ we define $\mathbb{Q} : \mathcal{G} \to [0, \infty)$ by

$$\mathbb{Q}(A) = \int_A X^+ d\mathbb{P} \quad \forall A \in \mathcal{G}$$

Then $\mathbb{Q}$ defines a measure on $\mathcal{G}$. It is obvious that $\mathbb{Q} \ll \mathbb{P}|_{\mathcal{G}}$ hence by the Radon-Nikodym theorem there exists a $\mathcal{G}$-measurable function Y^+ which satisfies

$$\int_A X^+ d\mathbb{P} = \mathbb{Q}(A) = \int_A Y^+ d\mathbb{P}, \quad \forall A \in \mathcal{G}$$

In a similar way we can find a $\mathcal{G}$-measurable function Y^- such that

$$\int_A X^- d\mathbb{P} = \int_A Y^- d\mathbb{P}, \quad \forall A \in \mathcal{G}$$

Next we define $Y := Y^+ - Y^-$. Then Y is a measurable function and satisfies

$$\int_A X d\mathbb{P} = \int_A Y d\mathbb{P}, \quad \forall A \in \mathcal{G}$$

Finally we must prove that Y is unique. Suppose that Y_1, Y_2 are random variables which satisfy

$$\int_A X d\mathbb{P} = \int_A Y_1 d\mathbb{P} = \int_A Y_2 d\mathbb{P}, \quad \forall A \in \mathcal{G}$$

Then

$$\int_A (Y_1 - Y_2) d\mathbb{P} = 0, \quad \forall A \in \mathcal{G}$$

Hence $Y_1 = Y_2$ almost surely.

□

Example 2.1.1:
Suppose we toss a coin three times. Then the sample space is given by $\Omega = \{(\omega_1, \omega_2, \omega_3) \mid \omega_i \in \{0,1\}\}$, where ω_i is the result of the i-th toss and 1 denotes heads and 0 tails. On this sample space we define the discrete σ-algebra $\mathcal{F} = \mathcal{P}(\Omega)$. On the measurable space $(\Omega, \mathcal{F})$ we define the following probability measure:

$$\mathbb{P}(A) := \frac{|A|}{8}$$

Let the random variable $X : \Omega \to \mathbb{R}$ model the number of times we toss heads. Consider the event $A \in \mathcal{F}$ where the first toss results in heads. Then we define the sub-σ-algebra $\mathcal{G}$ as $\mathcal{G} := \sigma(A) = \{\varnothing, A, A^c, \Omega\}$. Now we will determine the conditional expectation of X with respect to $\mathcal{G}$. We begin by noting that

$$X\mathbf{1}_A = \begin{cases} 3 & \text{if} \quad (\omega_1, \omega_2, \omega_3) \in \{(1,1,1)\} \\ 2 & \text{if} \quad (\omega_1, \omega_2, \omega_3) \in \{(1,1,0),(1,0,1)\} \\ 1 & \text{if} \quad (\omega_1, \omega_2, \omega_3) \in \{(1,0,0)\} \\ 0 & \text{else} \end{cases}$$

$$X\mathbf{1}_{A^c} = \begin{cases} 2 & \text{if} \quad (\omega_1, \omega_2, \omega_3) \in \{(0,1,1)\} \\ 1 & \text{if} \quad (\omega_1, \omega_2, \omega_3) \in \{(0,1,0),(0,0,1)\} \\ 0 & \text{else} \end{cases}$$

Hence $\mathbb{E}[X\mathbf{1}_\varnothing] = 0$, $\mathbb{E}[X\mathbf{1}_\Omega] = \frac{3}{2}$, $\mathbb{E}[X\mathbf{1}_A] = 1$ and $\mathbb{E}[X\mathbf{1}_{A^c}] = \frac{1}{2}$. Now we will define the random variable $\mathbb{E}[X|\mathcal{G}]$ as

$$\mathbb{E}[X|\mathcal{G}] = 2\mathbf{1}_A + \mathbf{1}_{A^c}$$

This is a simple function hence its integral is well defined. One can easily check that $\int_B \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_B X d\mathbb{P}$ for all $B \in \mathcal{G}$. Hence by theorem 2.1.1 $\mathbb{E}[X|\mathcal{G}]$ is the conditional expectation and is unique almost surely.

♦

Now we will consider the conditional expectation of a random variable X given another random variable Y.

Definition 2.1.3: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $X : \Omega \to \mathbb{R}$ be a random variable. Then we define the σ-algebra generated by X $\sigma(X)$ as the smallest σ-algebra such that $X : \Omega \to \mathbb{R}$ is measurable.

Definition 2.1.4: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $X, Y : \Omega \to \mathbb{R}$ be random variables defined on this probability space. The conditional expectation of X given Y is defined as

$$\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)]$$

We will now state and prove some important properties of conditional expectation which will be used in this thesis.

Theorem 2.1.2: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X, Y : \Omega \to \mathbb{R}$ be a random variables. Let $\mathcal{H}, \mathcal{G} \subset \mathcal{F}$ be sub-σ-algebras. Then the following properties hold.

P1 : *Law of total expectation*: $\mathbb{E}[\mathbb{E}[X|\mathcal{H}]] = \mathbb{E}[X]$

P2 : *Taking out what is known*: If X is $\mathcal{H}$-measurable then $\mathbb{E}[XY|\mathcal{H}] = X\mathbb{E}[Y|\mathcal{H}]$.

P3 : *Tower property*: If $\mathcal{H} \subset \mathcal{G}$ then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$

Proof:
For P1 note that $\Omega \in \mathcal{H}$ then the statement follows by definition. For P2 see for example [3] theorem 8.7(ii). For P3 note that $\int_H \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_H X d\mathbb{P}$ for all $H \in \mathcal{H}$ because $\mathcal{H} \subset \mathcal{G}$ the statement follows from theorem 2.1.1.

□

2.2 Stochastic processes

In this section we will shortly discuss the notion of a stochastic process. We will give the definition of a filtration and a stochastic process. Then we will focus on a special kind of stochastic process: a martingale

Definition 2.2.1: Let $(\Omega, \mathcal{F})$ be a measurable space. Let $\{\mathcal{F}_s\}_{s \in I}$ with $I \subset \mathbb{R}$ be a collection of sub-σ-algebras of $\mathcal{F}$. If $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for $s, t \in I$ such that $s \leq t$ then $\{\mathcal{F}_s\}_{s \in I}$ is called a filtration.

A filtration can be considered as the increment of information over time. Suppose we have a measurable space $(\Omega, \mathcal{F})$. We want to know which $\omega \in \Omega$ is the event that happened. We will illustrate the increment of information in the following example.

Example 2.2.1:
Suppose we toss a coin three times. Then the sample space is given by $\Omega = \{0,1\}^3$. The σ-algebra on this sample space is given by $\mathcal{F} = \mathcal{P}(\Omega)$. Before we toss the coin all we know is something in Ω will happen, hence we have

$$\mathcal{F}_0 = \{\varnothing, \Omega\}$$

Now suppose we have tossed the coin one time. Then there are two possibilities. If head comes up, the final outcome of the experiment will be in the set $A_H = \{(1,0,0),(1,0,1),(1,1,0),(1,1,1)\}$. If tale comes up than the final outcome will be in the set $A_T = \Omega / A_H$. Therefore we have

$$\mathcal{F}_1 = \{\varnothing, A_H, A_T, \Omega\}$$

After the second toss there are four possible outcomes: HH,HT,TH,TT. These outcomes imply that the final outcome of the experiment will be in one of the sets $A_{HH} = \{(1,1,0),(1,1,1)\}, A_{HT} = \{(1,0,0),(1,0,1)\}, A_{TH} = \{(0,1,0),(0,1,1)\}, A_{TT} = \{(0,0,0),(0,0,1)\}$. Therefore we have

$$\mathcal{F}_2 = \{\varnothing, A_H, A_T, A_{HH}, A_{HT}, A_{TH}, A_{TT}, \Omega\}$$

Finally after the third toss all information is available, hence $\mathcal{F}_3 = \mathcal{F}$. We note that $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \mathcal{F}$ hence this is a filtration.

♦

Definition 2.2.2: A stochastic process is a sequence of random variables $X_t : \Omega \to \mathbb{R}$ parameterized by $t \in I \subset \mathbb{R}$. The parameter t is often interpreted as time.

(i) If I is a discrete set, the process is called a discrete stochastic process. If I is an interval the process is called a continuous time stochastic process.

(ii) For each $\omega \in \Omega$ the mapping $t \mapsto X_t(\omega)$ is called a sample path.

(iii) The filtration $\{\mathcal{F}_t\}_{t\in I}$ generated by the process X_t, $\mathcal{F}_t := \sigma(\{X_s \mid 0 \le s \le t\})$ is called a natural filtration of X_t.

Definition 2.2.3: Let $\{\mathcal{F}_t\}_{t\in I}$ be a filtration and let $\{X_t\}_{t\in A}$ be a stochastic process. If for every $t \in I$ the random variable X_t is measurable with respect to $\mathcal{F}_t$. Then the stochastic process $\{X_t\}_{t\in I}$ is said to be adapted to the filtration $\{\mathcal{F}_t\}_{t\in I}$.

Definition 2.2.4: Let $\{X_t\}_{t\in I}$ be a stochastic process and let $\mathcal{F}_t = \sigma(X_u \mid 0 \le u \le t)$ be a sub-σ-algebra generated by the history of the process up to time t. The stochastic process $\{X_t\}_{t\in I}$ has the *Markov property* if for every $0 \le s \le t$ we have

$$\mathbb{P}(X_t \le y \mid \mathcal{F}_s) = \mathbb{P}(X_t \le y \mid X_s)$$

This means that if we know X_s the future movements of the process will be independent from what happened in the past of the process.

Definition 2.2.5: Let $\{X_t\}_{t\in I}$ be a stochastic process adapted to a filtration $\{\mathcal{F}_t\}_{t\in I}$ such that X_t is integrable for every $t \in I$. The process $\{X_t\}_{t\in I}$ is said to be a *martingale* with respect to $\{\mathcal{F}_t\}_{t\in I}$ if for every $0 \le s \le t$ we have

$$X_s = \mathbb{E}[X_t \mid \mathcal{F}_s]$$

Definition 2.2.6: Let $\{X_t\}_{0\le t\le T}$ be a stochastic process. Let $\mathcal{P} := \{t_0, t_1, ..., t_n\}$ be a partition of $[0,T]$ and let $\Delta(\mathcal{P})$ be the mesh of this partition. Define

$$Q(\mathcal{P}, X_t) = \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^2$$

Then the quadratic variation of the stochastic process $\{X_t\}_{0\le t\le T}$ is defined as

$$[X,X]_T := \lim_{\Delta(\mathcal{P})\to 0} Q(\mathcal{P}, X_t)$$

With convergence in the $\mathcal{L}^2$ sense.

Lemma 2.2.1:
The quadratic variation of the deterministic function $f(t) = t$ equals 0 with probability 1.

Proof:
We have by definition

$$[f(t), f(t)]_T := \lim_{\Delta(\mathcal{P})\to 0} \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2) \leq \lim_{\Delta(\mathcal{P})\to 0} \Delta(\mathcal{P}) \sum_{j=0}^{n-1} t_{j+1} - t_j = \lim_{\Delta(\mathcal{P})\to 0} \Delta(\mathcal{P}) T = 0$$

Note $[(f(t), f(t)]_T \geq 0$ hence the statement follows from the squeeze theorem.

□

Part II

Stochastic calculus

Chapter 3

Brownian motion

In this chapter we will introduce the Brownian motion. We will discuss the definition and elementary properties of the Brownian motion. We will also discuss the relation between the Brownian motion and martingales. The main goal of this chapter is to prove Girsanov's theorem.

3.1 Brownian motion: a stochastic process

In this section we will give the definition of the Brownian motion. Then we will derive the Brownian motion in such a way that will be useful for simulation. We will begin by giving the definition of the one-dimensional Brownian motion. In this section we follow the treatment as given by [2] section 7.2

Definition 3.1.1: A stochastic process $\{W_t\}_{t\geq 0}$ is called a Brownian motion if it has the following properties:

(i) $W_0 = 0$ and $t \mapsto W_t$ is continuous with probability 1

(ii) For $0 \leq s \leq t$ we have $W_t - W_s \sim \mathcal{N}(0, t-s)$

(iii) For $0 \leq t_1 < t_2 \leq t_3 < t_4 \leq \cdots \leq t_{2n-1} < t_{2n}$. The increments $W_{t_2} - W_{t_1}, W_{t_4} - W_{t_3}, ..., W_{t_{2n}} - W_{t_{2n-1}}$ are independent.

With this definition in mind we will now give a proper motivation for the existence of this stochastic process. Let Ω be the set of all continuous functions $\omega : [0,\infty) \to \mathbb{R}$ such that $\omega(0) = 0$. On this space we want to define a σ-algebra $\mathcal{F}$ such that we can define a probability measure on this space. In order to do this [2] considers cylinder subsets of Ω.

Definition 3.1.2: Consider the space Ω. For arbitrary time points $0 = t_0 < t_1 < \cdots < t_n = t$ and arbitrary intervals $I_1, ..., I_n \subset \mathbb{R}$ we define the cylinder subset as

$$\mathcal{C}(t_1, ..., t_n, I_1, ..., I_n) = \{\omega \in \Omega \mid \omega(t_1) \in I_1, ..., \omega(t_n) \in I_n\}$$

Next we define the σ-algebra $\mathcal{F}_t$ as the σ-algebra generated by the cylinder subsets $\mathcal{C}(t_1, ..., t_n, I_1, ..., I_n)$ where $n \geq 1$ and $0 = t_0 < t_1 < \cdots < t_n \leq t$. Because we are considering more cylinder subsets as t becomes larger we can conclude that $\{\mathcal{F}_t\}_{t\geq 0}$ is a filtration. We can now define the σ-algebra $\mathcal{F}$ on Ω by

$$\mathcal{F} := \sigma\left(\bigcup_{t\geq 0} \mathcal{F}_t\right)$$

The filtered measurable space $(\Omega, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{F})$ is called the sample space of the Brownian motion. Finally we will construct a probability measure $\mathbb{P}$.

Definition 3.1.3: Let $\Delta t_j = t_t - t_{j-1}$ and define $\tilde{\mathbb{P}}(\mathcal{C}(t_1,...,t_n,I_1,...,I_n)$ by

$$\int_{I_1} \cdots \int_{I_n} p(\Delta_{t_1};0;x_1)p(\Delta_{t_2};x_1;x_2)...p(\Delta_{t_n}:x_{n-1};x_n)dx_n...dx_1$$

where $p(\Delta t_j;x;y)$ denotes the density function of a $\mathcal{N}(x,\Delta t_j)$ random variable. Then by the Kolmogorov extension theorem $\tilde{\mathbb{P}}$ can be extended to a probability measure $\mathbb{P}$ on $(\Omega,\mathcal{F})$.

We have now constructed a probability space $(\Omega,\mathcal{F},\mathbb{P})$. On this probability space we will construct a stochastic process $\{W_t\}_{t\geq 0}$ for which we will show that this satisfies the conditions in definition 3.1.1 and therefore is a Brownian motion.

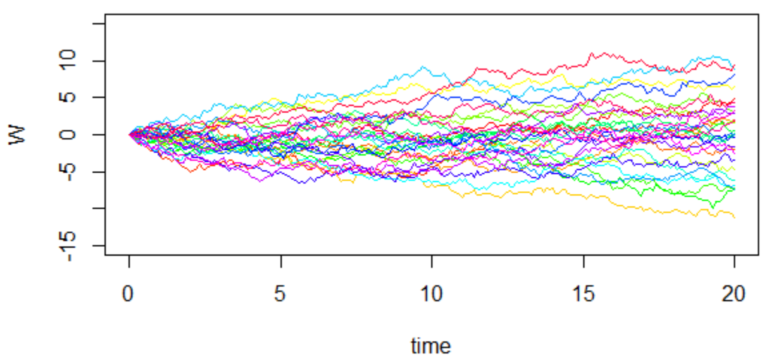

Figure 3.1: Simulation: 30 sample paths of a Brownian motion

Definition 3.1.4: Consider the probability space $(\Omega,\mathcal{F},\mathbb{P})$. Define the stochastic process $\{W_t\}_{t\geq 0}$ on this probability space by $W_t(\omega) = \omega(t)$. Note that $W_t^{-1}(I) = \mathcal{C}(t,I)$ for a Borel set $I \subset \mathbb{R}$. Therefore $\sigma(W_t) \subset \mathcal{F}_t$ for all t and thus $\{W_t\}_{t\geq 0}$ is adapted to $\{\mathcal{F}_t\}_{t\geq 0}$.

Theorem 3.1.1: The process $\{W_t\}_{\geq 0}$ from definition 3.1.4 is a Brownian motion.

Proof: See for example [2] theorem 7.8.

□

Theorem 3.1.2: Let $\{W_t\}_{t\geq 0}$ be a Brownian motion then we have $[W_t,W_t]_T = T$.

Proof:
Let $Z \sim \mathcal{N}(0,1)$ then it can be shown that $\mathbb{E}[Z^4] = 3$ for example by using the moment generating function.
Let $tT > 0$ and consider the interval $[0,T]$. Let $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$ be a partition $\mathcal{P}$ of $[0,T]$. Define $\Delta(\mathcal{P}) := \max_{1\leq i\leq n} |t_i - t_{i-1}|$. Now let $\mathcal{I}_i = (W_{t_i} - W_{t_{i-1}})^2 - (t_i - t_{i-1})$. By the independent increment property of a Brownian motion and using the normal distribution property

we have $\mathbb{E}[\mathcal{I}_i] = 0$ for $i = 1, ..., n$. Next we note that

$$\sum_{i=1}^{n} \mathcal{I}_i = Q(\mathcal{P}) - T$$

Next we determine the second moment of this random variable

$$\begin{aligned}\mathbb{E}\left[(Q(\mathcal{P}) - T)^2\right] &= \sum_{i=1}^{n} \mathbb{E}[\mathcal{I}_i^2] = \sum_{i=1}^{n} \mathbb{E}[(W_{t_i} - W_{t_{i-1}})^4 - 2(W_{t_i} - W_{t_{i-1}})^2(t_i - t_{i-1}) + (t_i - t_{i-1})^2] \\ &= \sum_{i=1}^{n} \mathbb{E}[(W_{t_i} - W_{t_{i-1}})^4] - 2\sum_{i=1}^{n}(t_i - t_{i-1})^2 + \sum_{i=1}^{n}(t_i - t_{i-1})^2\end{aligned}$$

Then by using the $\mathbb{E}[Z^4] = 3$ property we obtain the following estimation.

$$\begin{aligned}\mathbb{E}\left[(Q(\mathcal{P}) - T)^2\right] &= 3\sum_{i=1}^{n} \mathbb{E}(t_i - t_{i-1})^2 - 2\sum_{i=1}^{n}(t_i - t_{i-1})^2 + \sum_{i=1}^{n}(t_i - t_{i-1})^2 = 2\sum_{i=1}^{n}(t_i - t_{i-1})^2 \\ &\leq 2\Delta(\mathcal{P})\sum_{i=1}^{n}(t_i - t_{i-1}) = 2\Delta(\mathcal{P})T\end{aligned}$$

Now suppose we have a sequence $\{\mathcal{P}_j\}_{j\in\mathbb{N}}$ of partitions of $[0,T]$ such that $\lim_{j\to\infty} \Delta(\mathcal{P}_j) = 0$. For every $j \in \mathbb{N}$ we have

$$0 \leq \mathbb{E}[|Q(\mathcal{P}_j) - T|^2] \leq 2\Delta(\mathcal{P}_j)T$$

Then applying the squeeze theorem gives us

$$\lim_{j\to\infty} \mathbb{E}[|Q(\mathcal{P}_j) - T|^2] = 0$$

From which we conclude that $Q(\mathcal{P}_j)$ converges to T in the $\mathcal{L}^2$-sense.

□

3.2 Brownian motion and martingales

In this section we will discuss the relation between the Brownian motion and martingales. We will give examples of martingales which come from the Brownian motion.

Lemma 3.2.1: Let $\{W_t\}_{t\geq 0}$ be a Brownian motion and let θ be a real constant. Then the stochastic process $\{L_t\}_{t\geq 0}$ defined by

$$L_t = e^{-\frac{1}{2}\theta^2 t - \theta W_t}$$

is a martingale.

Proof: Let $0 \leq s \leq t$.

$$\begin{aligned}
\mathbb{E}[L_t|\mathcal{F}_s] &= \mathbb{E}[e^{-\frac{1}{2}\theta^2 t-\theta W_t}|\mathcal{F}_s] \\
&= e^{-\frac{1}{2}\theta^2 t}\mathbb{E}[e^{-\theta(W_t-W_s)}e^{-\theta W_s}|\mathcal{F}_s] \\
&= e^{-\frac{1}{2}\theta^2 t}\mathbb{E}[e^{-\theta(W_t-W_s)}|\mathcal{F}_s]\mathbb{E}[e^{-\theta W_s}|\mathcal{F}_s] \quad W_t - W_s \text{ are } W_s = W_s - W_0 \text{ are independent} \\
&= e^{-\frac{1}{2}\theta^2 t}\mathbb{E}[e^{-\theta(W_t-W_s)}]\mathbb{E}[e^{-\theta W_s}|\mathcal{F}_s] \\
&= e^{-\frac{1}{2}\theta^2 t}e^{\frac{1}{2}\theta^2(t-s)}e^{-\theta W_s} \\
&= e^{-\frac{1}{2}\theta^2 s-\theta W_s} \\
&= L_s
\end{aligned}$$

□

We will now give some examples of a martingale which come from the Brownian motion.

Theorem 3.2.1: Let $\{W_t\}_{t\geq 0}$ be a Brownian motion. Then the following stochastic processes are martingales:

(i) $\{W_t\}_{t\geq 0}$

(ii) $\left\{e^{W_t-\frac{1}{2}t}\right\}_{t\geq 0}$.

Proof:
We will first prove (i). Let $0 \leq s \leq t$ then by (iii) from definition 3.1.1 we have $W_t - W_s$ independent from $\mathcal{F}_s$. Hence

$$\mathbb{E}[W_t|\mathcal{F}_s] = \mathbb{E}[W_t - W_s + W_s|\mathcal{F}_s] = \mathbb{E}[W_t - W_s|\mathcal{F}_s] + \mathbb{E}[W_s|\mathcal{F}_s] = \mathbb{E}[W_t - W_s] + W_s = W_s$$

this proves (i). For (ii) note that we can apply lemma 3.2.1 with $\theta = -1$.

□

We will finish this section with a theorem which will be used in proving Girsanov's theorem in the next section.

Theorem 3.2.2: Let $\{M_t\}_{t\geq 0}$ be a martingale with respect to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and a probability measure $\mathbb{Q}$. If M_t is continuous, $M_0 = 0$ and $[M,M]_t = t$ then $\{M_t\}_{t\geq 0}$ is a Brownian motion.

Proof: This is theorem 7.13 from [2]

□

3.3 Girsanov's Theorem

Suppose we have a Brownian motion $\{W_t\}_{\geq 0}$ on a measurable space $(\Omega, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{F})$ with respect to a probability measure $\mathbb{P}$. Let θ be a real constant and consider the stochastic process $X_t = W_t + t\theta$ which is called a Brownian motion with drift. In this section we want to find a probability measure $\mathbb{Q}$ such that the process $\{X_t\}_{0\leq t\leq T}$ for some $T > 0$ is a Brownian motion with respect to the probability measure $\mathbb{Q}$. An additional requirement is that the probability measures $\mathbb{P}$ and $\mathbb{Q}$ must be equivalent.

Remark 3.3.1: If X is a random variable and $\mathbb{P}$ is a probability measure we will denote the expectation of X with respect to this probability measure as $\mathbb{E}^{\mathbb{P}}[X]$.

Definition 3.3.1: Let $\{L_t\}_{t\geq 0}$ be the stochastic process as defined in lemma 3.2.1. We know that L_t is a martingale and we note that $\mathbb{E}[L_t] = \mathbb{E}[L_0] = 1$. Let $T > 0$ then we can define the probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}_T)$ by $d\mathbb{Q} = L_T d\mathbb{P}$. which means that for every $A \in \mathcal{F}_T$ we have

$$\mathbb{Q}(A) = \int_A L_T d\mathbb{P}$$

Theorem 3.3.1: Let $0 \leq s \leq t$ and let Y_t be an $\mathcal{F}_t$ measurable random variable. Then we have

$$\mathbb{E}^{\mathbb{Q}}[Y_t|\mathcal{F}_s] = \mathbb{E}^{\mathbb{P}}\left[Y_t \frac{L_t}{L_s}|\mathcal{F}_s\right]$$

Proof: See for example [2] lemma 8.1

□

Lemma 3.3.1: Let $\{W_t\}_{t\geq 0}$ be a Brownian motion and let θ be a real constant, then

$$\mathbb{E}[W_t e^{\theta W_t}] = \theta t e^{\frac{1}{2}\theta^2 t} \text{ and } \mathbb{E}[W_t^2 e^{\theta W_t}] = (t + \theta^2 t^2) e^{\frac{1}{2}\theta^2 t}$$

Proof:
We note that $W_t \sim \mathcal{N}(0, t)$. Let $k \in \{1, 2\}$ then

$$\mathbb{E}[W_t^k e^{\theta W_t}] = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} w^k e^{\theta w} e^{-\frac{1}{2}\frac{w^2}{t}} dw = e^{\frac{1}{2}\theta^2 t} \int_{\mathbb{R}} \frac{w^k}{\sqrt{2\pi t}} e^{-\frac{1}{2}\left(\frac{w-\theta t}{\sqrt{t}}\right)^2} dw$$

The last integral equals the k-th moment of a $\mathcal{N}(\theta t, t)$ distribution. For $k = 1$ this equals θt and for $k = 2$ this equals $t + \theta^2 t^2$ proving the lemma.

□

Lemma 3.3.2: Let $\{W_t\}_{t\geq 0}$ be a Brownian motion and let $0 \leq s < t$ and θ a real constant. Then

$$\mathbb{E}[W_t e^{\theta W_t}|\mathcal{F}_s] = [W_s + \theta(t-s)] e^{\frac{1}{2}\theta^2(t-s) + \theta W_s}$$

Proof:
By using linearity of expectation we can write

$$\mathbb{E}[W_t e^{\theta W_t} \mid \mathcal{F}_s] = \mathbb{E}[(W_t - W_s) e^{\theta(W_t - W_s)} e^{\theta W_s}|\mathcal{F}_s] + \mathbb{E}[W_s e^{\theta(W_t - W_s)} e^{\theta W_s}|\mathcal{F}_s]$$

Taking out what is known then gives us

$$\mathbb{E}[(W_t - W_s) e^{\theta(W_t - W_s)}|\mathcal{F}_s] e^{\theta W_s} + W_s e^{\theta W_s} \mathbb{E}[e^{\theta(W_t - W_s)}|\mathcal{F}_s]$$

Now we can use that $W_t - W_s$ is independent from $\mathcal{F}_s$,$W_t - W_s \sim \mathcal{N}(0, t-s)$ and apply lemma 3.3.1 to obtain the following expression

$$\theta(t-s) e^{\frac{1}{2}\theta^2(t-s)} e^{\theta W_s} + W_s e^{\theta W_s} e^{\frac{1}{2}\theta^2(t-s)}$$

Rewriting this expression gives us the desired result.

□

Theorem 3.3.2: Let $\mathbb{Q}$ be the probability measure from definition 3.3.1 and let $X_t = W_t + t\theta$ with W_t a Brownian motion with respect to $\mathbb{P}$. Then X_t is a martingale with respect to $\mathbb{Q}$.

Proof:
Let $0 \leq s \leq t$ then from lemma 3.3.2 we can conclude

$$\mathbb{E}^{\mathbb{P}}[W_t L_t|\mathcal{F}_s] = e^{-\frac{1}{2}\theta^2 t}\mathbb{E}^{\mathbb{P}}[W_t e^{-\theta W_t}|\mathcal{F}_s] = [W_s - \theta(t-s)]L_s$$

Then by using theorem 3.3.1

$$\mathbb{E}^{\mathbb{Q}}[X_t|\mathcal{F}_s] = \mathbb{E}^{\mathbb{P}}[X_t L_t L_s^{-1}|\mathcal{F}_s] = \mathbb{E}^{\mathbb{P}}[(W_t + t\theta)L_t L_s^{-1}|\mathcal{F}_s]$$

Taking out what is known and applying the definition of L_t gives us

$$L_s^{-1}\mathbb{E}^{\mathbb{P}}[W_t L_t|\mathcal{F}_s] + L_s^{-1} t\theta \mathbb{E}^{\mathbb{P}}[L_t|\mathcal{F}_s]$$

Using that L_t is a martingale with respect to $\mathbb{P}$ and using the first comment in this proof we obtain the desired result.

$$L_s^{-1}([W_s - \theta(t-s)]L_s) + L_s^{-1} t\theta L_s = W_s + \theta s = X_s$$

□

Theorem 3.3.3: Girsanov's theorem:
Let $\{W_t\}_{0\leq t\leq T}$ be a Brownian motion with respect to the probability measure $\mathbb{P}$ and the filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let θ be a real constant and let $X_t = W_t + t\theta$. Let $\mathbb{Q}$ be the probability measure from definition 3.3.1. Then X_t is a Brownian motion with respect to $\mathbb{Q}$.

Proof:
Because L_t is non-negative it is obvious that $\mathbb{P}$ and $\mathbb{Q}$ are equivalent probability measures. From theorem 3.3.2 we know that X_t is a martingale with respect to $\mathbb{Q}$. Because W_t is almost surely continuous we have $X_t = W_t + t\theta$ continuous almost surely. It is obvious that $X_0 = 0$ and $[X, X]_t = t$ from lemma 2.2.1. We see that X_t satisfies the conditions from theorem 3.2.2 and thus we conclude that X_t is a Brownian motion with respect to $\mathbb{Q}$

□

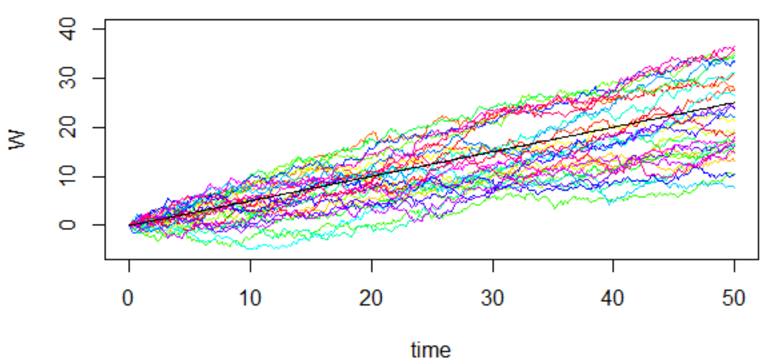

Figure 3.2: Simulation: 30 sample paths of a Brownian motion with drift $\theta = \frac{1}{2}$

Chapter 4

The Itô integral

In this chapter we will construct the Itô integral and discuss important properties such as Itô isometry and the martingale property.

4.1 Definition of the Itô integral

In this section we will construct the Itô integral. We begin by introducing a class of functions as defined in [9] definition 3.4

Definition 4.1.1: Let $\mathcal{V} = \mathcal{V}(S,T)$ be the class of functions $f : [0,\infty) \times \Omega \to \mathbb{R}$ such that

(i) $(t,\omega) \mapsto f(t,\omega)$ is $\mathcal{B}([0,\infty)) \otimes \mathcal{F}$ measurable

(ii) $f(t,\omega)$ is $\mathcal{F}_t$ adapted

(iii)
$$\mathbb{E}\left[\int_S^T |f_t(\omega)|^2 dt\right] < \infty$$

Definition 4.1.2: Let $T > 0$ and consider the interval $[0,T]$. Let $= 0 = t_0 < t_1 < t_2 < \cdots < t_n = T$ be a partition of this interval. Suppose that for every $i \in \{0,...,n-1\}$ there exist a $\mathcal{F}_{t_i}$ measurable random variable $\xi_{t_i} \in \mathcal{L}^2(\Omega,\mathcal{F},\mathbb{P})$. Then a stochastic process of the form

$$f_t(\omega) = \sum_{i=0}^{n-1} \xi_{t_i} \mathbf{1}_{[t_1,t_{i-1})}(t)$$

is called a simple process.

Corollary 4.1.1: The set $\mathcal{H}_0^2$ of all simple processes is a subset of $\mathcal{V}$.

Proof:
Because $\xi_{t_i} \in \mathcal{L}^2(\Omega,\mathcal{F},\mathbb{P})$ we have by definition $\mathbb{E}[|\xi_{t_i}|^2] < \infty$. Hence we have

$$\sum_{i=0}^{n-1} \mathbb{E}[|\xi_{t_i}|^2](t_{i+1} - t_i) < \infty$$

□

Definition 4.1.3: Itô integral of a simple process:
Let $\{f_t(\omega)\}$ be a simple process and let $\{W_t\}_{t\geq 0}$ be a Brownian motion. Then the Itô integral is defined as

$$I(f) = \sum_{i=0}^{n-1} \xi_{t_i}(\omega)(W_{t_{i+1}}(\omega) - W_{t_i}(\omega))$$

We can now begin by deriving some important properties of the Itô integral.

Theorem 4.1.1: Itô isometry:
Let $f \in \mathcal{H}_0^2$. Then $I(f) \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ and the map $I : \mathcal{H}_0^2 \to \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ is a norm preserving map.

Proof:
In this proof we will use the notation $\Delta W_{t_i} := W_{t_{i+1}} - W_{t_i}$. Let $f \in \mathcal{H}_0^2$.

$$\mathbb{E}|I(f)|^2] = \mathbb{E}\left[\left(\sum_{i=0}^{n-1} \xi_{t_i}(W_{t_{i+1}} - W_{t_i})\right)^2\right] = \mathbb{E}\left[\left(\sum_{i=0}^{n-1} \xi_{t_i} \Delta W_{t_i}\right)^2\right]$$

Rewriting this expression gives us

$$\mathbb{E}\left[\sum_{i=0}^{n-1} \xi_{t_i}^2 (\Delta W_{t_i})^2\right] + 2\mathbb{E}\left[\sum_{j<i} \xi_{t_i}\xi_{t_j} \Delta W_{t_i} \Delta W_{t_j}\right]$$

Note that because ξ_t and ΔW_t are independent we can take the product of the expectations. Furthermore we have by the independent increment property of the Brownian motion that $\mathbb{E}[\Delta W_{t_i} \Delta W_{t_j}] = \mathbb{E}[\Delta W_{t_i}]\mathbb{E}[\Delta W_{t_j}] = 0$. Also we have $\mathbb{E}[(\Delta W_{t_i})^2] = t_{i+1} - t_i$. Using these arguments we can finish our expression

$$\mathbb{E}[I(f)^2] = \sum_{i=0}^{n-1} \mathbb{E}[|\xi_{t_i}|^2](t_{i+1} - t_i) < \infty$$

Thus we have $I(f) \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. We will now consider the norm of $f \in \mathcal{H}_0^2$. We have

$$\begin{aligned} |f_t(\omega)|^2 &= \left(\sum_{i=0}^{n-1} \xi_{t_i}(\omega)\mathbf{1}_{[t_i,t_{i+1})}(t)\right)^2 = \sum_{i=0}^{n-1} |\xi_{t_i}|^2 \mathbf{1}_{[t_i,t_{i+1})} + 2\sum_{j<i} \xi_{t_i}\xi_{t_j}\mathbf{1}_{[t_i,t_{i+1})}(t)\mathbf{1}_{[t_j,t_{j+1})} \\ &= \sum_{i=0}^{n-1} |\xi_{t_i}(\omega)|^2 \mathbf{1}_{[t_i,t_{i+1})}(t) \end{aligned}$$

Hence

$$\int_0^\infty |f_t(\omega)| dt = \sum_{i=0}^{n-1} |\xi_{t_i}(\omega)|^2 (t_{i+1} - t_i)$$

From which we can see

$$\mathbb{E}\left[\int_0^\infty |f_t(\omega)|^2 dt\right] = \sum_{i=0}^{n-1} \mathbb{E}[|\xi_{t_i}|^2](t_{i+1} - t_i)$$

Which proves the theorem.

□

With this theorem we can expand the notion of the Itô integral to a bigger set of stochastic processes. We present the arguments of [9]. Let $f \in \mathcal{V}$ then there exists a sequence of simple processes (f_n) that converges to f see appendix B theorem B.2. Because this is a convergent sequence this is a Cauchy-sequence. Because the Itô integral is norm preserving the sequence $I(f_n) \in \mathcal{L}^2(\Omega,\mathcal{F},\mathbb{P})$ is a Cauchy-sequence as well. Because the space $\mathcal{L}^2(\Omega,\mathcal{F},\mathbb{P})$ is complete there exists an almost surely unique limit $I(f) \in \mathcal{L}^2(\Omega,\mathcal{F},\mathbb{P})$ i.e.

$$\lim_{n\to\infty} \mathbb{E}[|I(f) - I(f_n)|^2] = 0$$

Therefore we have the following definition.

Definition 4.1.3: Let $f \in \mathcal{V}$ and let (f_n) be a sequence of simple processes converging to f. Then we define the Itô integral of f to be

$$I(f) = \lim_{n\to\infty} I(f_n)$$

Before we will look at an example we will first give a new definition.

Definition 4.1.4: Let $\{f_t\}_{t\geq 0}$ be a stochastic process adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Then the process $\left\{\mathbf{1}_{[0,T]}(t) f_t\right\}_{t\geq 0}$ is adapted to $\{\mathcal{F}_t\}_{t\geq 0}$ as well. Then for some $T > 0$ we define the Itô integral of $\{f_t\}_{t\geq 0}$ on $[0,T]$ by

$$\int_0^T f_t dW_t := \int_0^\infty \mathbf{1}_{[0,T]}(t) f_t dW_t$$

Example 4.1.1:
We will show that

$$\int_0^T W_t dW_t = \frac{1}{2} W_T^2 - \frac{1}{2} T$$

We will begin by introducing the sequence of simple processes (f_n) by

$$f_n(t,\omega) := \sum_{i=0}^{n-1} W_{t_i} \mathbf{1}_{[t_i,t_{i+1})}(t)$$

Note that the sequence (f_n) converges to $W_t \mathbf{1}_{[0,T]}(t)$ because

$$\begin{aligned}
\mathbb{E}\left[\int_0^\infty |W_t \mathbf{1}_{[0,T]}(t) - f_n(t)|^2 dt\right] &= \sum_{i=0}^{n-1} \left(\int_{t_i}^{t_{i+1}} \mathbb{E}[(W_t - W_{t_i})^2] dt\right) \\
&= \sum_{i=0}^{n-1} \left(\int_{t_i}^{t_{i+1}} (t - t_i) dt\right) = \frac{1}{2} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 \\
&\leq \frac{1}{2} \max_{0\leq i\leq n-1} |t_{i+1} - t_i| \sum_{i=0}^{n-1} (t_{i+1} - t_i) = \frac{1}{2} T \max_{0\leq i\leq n-1} |t_{i+1} - t_i|
\end{aligned}$$

Note that we may interchange the expectation and the integral because of Tonelli's theorem. Let $\max_{0\leq i\leq n-1} |t_{i+1} - t_i| \to 0$ if $n \to 0$. Therefore we have convergence in $\mathcal{V}$. Note that the Itô integral of this process is given by

$$I(f_n) = \sum_{i=0}^{n-1} W_{t_i}(W_{t_{i+1}} - W_{t_i}) = \frac{1}{2} \sum_{i=0}^{n-1} (W_{t_{i+1}}^2 - W_{t_i}^2) - \frac{1}{2} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2$$

Note that the first part of the last expression equals $\frac{1}{2} W(T)^2$. In the last part of the expression we recognize the formula for the quadratic variation. Hence by theorem 3.1.2 if we take the limit

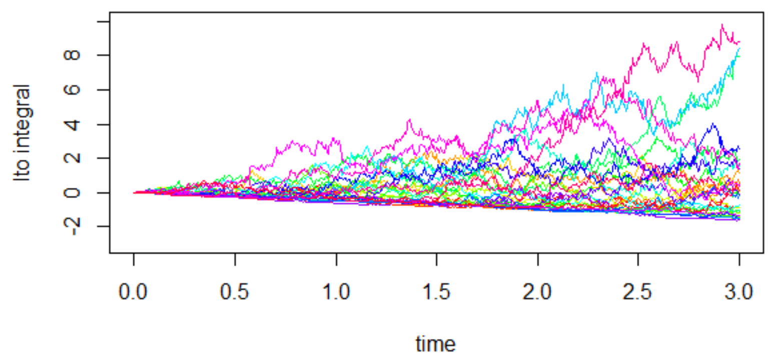

Figure 4.1: Simulation: 30 sample paths of the stochastic process $\left(\int_0^t W_s dW_s\right)_{0\leq t\leq 3}$

this will converge to $-\frac{1}{2}T$ in the $\mathcal{L}^2$-sense. Hence we conclude

$$\int_0^T W_t dW_t = \frac{1}{2}W_T^2 - \frac{1}{2}T$$

4.2 The martingale property of the Itô integral

In this section we will prove the martingale property of the Itô integral.

Theorem 4.2.1: Let $(\Omega, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a probability space with a filtration. Let $\{W_t\}_{t\geq 0}$ be a $(\mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$-Brownian motion and let $\{g_t(\omega)\}_{t\geq 0}$ be a $\mathcal{F}_t$ predictable process such that $\mathbb{E}[|g(t,\omega)|^2] < \infty$. Then the stochastic process

$$M_t := \int_0^t g(s,\omega) dW_s$$

is a martingale.

Lemma 4.2.1: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra. Then the map $\mathbb{E}[\cdot|\mathcal{G}] : \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}) \to \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ maps convergent sequences to convergent sequences.

Proof of the lemma:
This proof is lemma 6.11 from [7]. Note that $x \mapsto x^2$ is convex hence we can use Jensen's inequality to obtain the following inequality

$$\mathbb{E}[|X_n - X||\mathcal{G}]^2 \leq \mathbb{E}[|X_n - X|^2|\mathcal{G}]$$

Then the result is obtained by taking the expectation on both sides and raising to the power $\frac{1}{2}$.

□

Proof of theorem 4.2.1:
Let us first assume $g(s,\omega) \in \mathcal{H}_0^2$. Let $0 \leq r \leq t$. Let $0 = t_0 < t_1 < \cdots < t_n = t$ be the partition of $[0,t]$ for which g is defined. We then rewrite the partition $0 = t_0 < t_1 < \cdots < t_j \leq r < t_{j+1} < \cdots < t_n = t$

such that

$$g(s,\omega) = \sum_{i=0}^{j-1} \xi_{t_i} \mathbf{1}_{[t_i,t_{i+1})}(s) + \xi_{t_j} \left(\mathbf{1}_{[t_j,r)}(s) + \mathbf{1}_{[r,t_{j+1})}(s) \right) + \sum_{i=j+1}^{n-1} \xi_{t_i} \mathbf{1}_{[t_i,t_{i+1})}(s)$$

From this we see

$$\begin{aligned} \mathbb{E}[M_t|\mathcal{F}_r] &= \mathbb{E}\left[\int_0^t g(s,\omega)dW_s|\mathcal{F}_r\right] = \mathbb{E}\left[\int_r^t g(s,\omega)dW_s + \int_0^r g(s,\omega)dW_s|\mathcal{F}_r\right] \\ &= \mathbb{E}\left[\int_r^t g(s,\omega)dW_s|\mathcal{F}_r\right] + \mathbb{E}\left[\int_0^r g(s,\omega)dW_s|\mathcal{F}_r\right] \\ &= \mathbb{E}\left[\int_r^t g(s,\omega)dW_s\right] + \mathbb{E}\left[\int_0^r g(s,\omega)dW_s|\mathcal{F}_r\right] \end{aligned}$$

The last step comes from the fact that $W_{t_{i+1}} - W_{t_i}$ is independent of $\mathcal{F}_r$ if $t_i \geq r$. We will continue with the expression

$$= \mathbb{E}[\xi_{t_j}(\omega)]\mathbb{E}[W_{t_{j+1}} - W_r] + \sum_{i=j+1}^{n-1} \mathbb{E}[\xi_{t_i}(\omega)]\mathbb{E}[\Delta W_{t_j}] + \int_0^r g(s,\omega)dW_s$$

Note that $\mathbb{E}[\Delta W_{t_j}] = 0$ for all j hence the sum will be cancelled. This leaves us with

$$\mathbb{E}[M_t|\mathcal{F}_r] = \int_0^r g(s,\omega)dW_s = M_r$$

Thus if $g(s,\omega) \in \mathcal{H}_0^2$ the process $\{M_t\}$ is a martingale. Now assume $g(s,\omega) \in \mathcal{V}$. Then there exists a sequence $(g_n) \subset \mathcal{H}_0^2$ converging to g. The conditional expectation map maps convergent sequences to convergent sequences by lemma 4.2.1 hence

$$\mathbb{E}\left[\int_0^t g(s,\omega)dW_s|\mathcal{F}_r\right] = \lim_{n\to\infty} \mathbb{E}\left[\int_0^t g_n(s,\omega)dW_s|\mathcal{F}_r\right] = \lim_{n\to\infty} \int_0^r g_n(s,\omega)dW_s = \int_0^r g(s,\omega)dW_s$$

Where we used the definition of the Itô integral.

$$\lim_{n\to\infty} \int_0^r g_n(s,\omega)dW_s = \int_0^r g(s,\omega)dW_s$$

□

Corollary 4.2.1:

$$\mathbb{E}\left[\int_0^t f(s,\omega)dW_s\right] = 0$$

Proof:
By theorem 4.2.1 we know that the Itô integral is a martingale. Next we note that

$$\mathbb{E}[M_0] := \mathbb{E}\left[\int_0^0 f(s,\omega)dW_s\right] = \mathbb{E}[0] = 0$$

Then using the martingale property

$$\mathbb{E}\left[\int_0^t f(s,\omega)dW_s\right] = \mathbb{E}[M_t] = \mathbb{E}[M_0] = 0$$

□

Chapter 5

The Itô formula

In this chapter we will discuss the Itô formula. We will begin by stating the formula and discussing some examples. In the second part of this chapter we will give an intuitive derivation of the Itô formula.

5.1 The formula and examples

Theorem 5.1.1: Let $f \in C^2([0,\infty) \times \mathbb{R})$ and let $T > 0$ then the following equality holds a.s.

$$f(T,W_T) - f(0,W_0) = \int_0^T f_t(s,W_s)ds + \int_0^T f_x(s,W_s)dW_s + \frac{1}{2}\int_0^T f_{xx}(s,W_s)ds \tag{5.1}$$

We will next consider some examples.

Example 5.1.1:
Suppose we have $f(t,x) = x$. Then we obviously have $f \in C^2([0,\infty) \times \mathbb{R})$. Therefore the partial derivatives exist and are continuous. We can now apply the Itô formula

$$W_T = W_T - W_0 = \int_0^T dW_s$$

♦

Example 5.1.2:
Suppose we want to evaluate the stochastic integral

$$\int_0^T e^{W_s} dW_s$$

To do this we will consider the function $f(t,x) = e^x$. We have $f \in C^2([0,\infty) \times \mathbb{R})$ hence we can apply the Itô formula. This gives us

$$e^{W_T} - 1 = e^{W_t} - e^0 = \int_0^T 0ds + \int_0^T e^{W_s} dW_s + \frac{1}{2}\int_0^T e^{W_s} ds$$

Then rewriting this expression gives us the following result.

$$\int_0^T e^{W_s} dW_s = e^{W_T} - 1 - \frac{1}{2}\int_0^T e^{W_s} ds$$

5.2 Proof of the Itô formula

In this section we will give an outline of the proof. For an in depth treatment of the proof see for example [2] theorem 11.1. We will begin with a lemma which will be useful in the proof for the Itô formula.

Lemma 5.2.1: Cross-variation
Let $T > 0$ and consider the interval $[0,T]$. Let $t_0 < t_1 < \cdots < t_n$ be a partition of this interval and let $\Delta(\mathcal{P})$ denote the mesh of this interval. We then have for the cross-variation of W_t with t

$$\lim_{\Delta(\mathcal{P})\to 0}\sum_{j=0}^{n-1}(W_{t_{j+1}}-W_{t_j})(t_{j+1}-t_j)=0$$

Proof
We begin by noting that $[t_j,t_{j+1}]$ are compact sets for all j and that $W_{t_{j+1}}-W_{t_j}$ are continuous paths hence a maximum exists. Therefore we can estimate the limit by

$$\begin{aligned}\lim_{\Delta(\mathcal{P})\to 0}\sum_{j=0}^{n-1}(W_{t_{j+1}}-W_{t_j})(t_{j+1}-t_j) &\leq \lim_{\Delta(\mathcal{P})\to 0}\max_{1\leq u\leq n-1}|W_{t_{u+1}}-W_{t_u}|\sum_{j=0}^{n-1}(t_{j+1}-t_j)\\ &= \lim_{\Delta(\mathcal{P})\to 0}\max_{1\leq u\leq n-1}|W_{t_{u+1}}-W_{t_u}|T=0\end{aligned}$$

Where the last step comes from the fact that $\{W_t\}$ is continuous on $[0,T]$.

□

Proof of the Itô formula:
Let $f:[0,\infty)\times\mathbb{R}\to\mathbb{R}$ satisfy the conditions from theorem 5.1.1. We will also assume that all partial derivatives are bounded. This will not change the final result, see for example [2] theorem 11.1. Consider a time interval $[0,T]$ and let $0=t_0<t_1<\cdots<t_n=T$ be a partition of this interval. The mesh of this partition will be denoted by $\Delta(\mathcal{P})$. In this proof we write $\delta t_j = t_{j+1}-t_j$ and $\delta W_{t_j} = W_{t_{j+1}}-W_{t_j}$. We will write f_u for the partial derivative with respect to u evaluated in the point (t_j,W_{t_j}). From this we note that for a point (t_j,W_{t_j}) the Taylor expansion of f without the higher order terms is given by:

$$f(t_{j+1},W_{t_{j+1}})=f(t_j,W_{t_j})+f_t\delta t_j+f_x\delta W_{t_j}+\frac{1}{2}\left(f_{tt}(\delta t_j)^2+f_{xt}\delta t_j\delta W_{t_j}+f_{xx}(\delta W_{t_j})^2\right)$$

$$\begin{aligned}f(T-W_t)-f(0,W_0)&=\sum_{i=0}^{n-1}\left\{f(t_{j+1},W_{t_{j+1}})-f(t_j,W_{t_j})\right\}\\ &=\sum_{j=0}^{n-1}f_t\delta t_j+f_x\delta W_{t_j}+\frac{1}{2}\left(f_{tt}(\delta t_j)^2+f_{xt}\delta t_i\delta W_{t_j}+f_{xx}(\delta W_{t_j})^2\right)\end{aligned}$$

We will now evaluate each sum as we take the limit for $\Delta(\mathcal{P})\to 0$. We first consider the partial derivative f_t

$$\lim_{\Delta(\mathcal{P})\to 0}\sum_{j=0}^{n-1}f_t(t_{j+1}-t_j)=\int_0^T f_t(s,W_s)ds$$

which is a Riemann integral. Next we consider the partial derivative f_x.

$$\lim_{\Delta(\mathcal{P})\to 0} \sum_{j=0}^{n-1} f_x(W_{t_{j+1}} - W_{t_j}) = \int_0^T f_x(s, W_s) dW_s$$

which follows from the definition of the Itô integral. We now consider the partial derivative f_{xx}. We note that this sum resembles theorem 3.1.2. Therefore we conclude

$$\lim_{\Delta(\mathcal{P})\to 0} \sum_{j=0}^{n-1} \frac{1}{2} f_{xx}(W_{t_{j+1}} - W_{t_j})^2 = \frac{1}{2}\int_0^T f_{xx}(s, W_s) ds$$

for a more rigorous argument of this statement see for example [2] theorem 11.1. Finally we will show the remaining sums converge to 0. We then have

$$\lim_{\Delta(\mathcal{P})\to 0} \sum_{j=0}^{n-1} f_{tt}\delta t_j^2 \leq \lim_{\Delta(\mathcal{P})\to 0} \Delta(\mathcal{P}) \sum_{j=0}^{n-1} f_t \delta t_j = 0$$

and

$$\lim_{\Delta(\mathcal{P})\to 0} \sum_{j=0}^{n-1} f_{xt}\delta t_j \delta W_{t_j} \leq \lim_{\Delta(\mathcal{P})\to 0} \max|W_{t_{j+1}} - W_{t_j}| \sum_{j=0}^{n-1} f_{xt}\delta t_j = 0$$

We have that the sums converge to finite integrals. We have $\Delta(\mathcal{P})$ converging to 0 by definition and $\max|W_{t_{j+1}} - W_{t_j}|$ converging to 0 because of the continuity of $\{W_t\}$. In [9] theorem 4.2 it is stated that the higher terms are given by $O(|\delta t_j|^2 + |\delta W_{t_j}|^2)$. Therefore the total remainder equals

$$O\left(\sum_i \delta t_j^2 + \delta W_{t_j}^2\right)$$

We will now show that the remainder converges to 0 in the $\mathcal{L}^2$ sense. We have

$$\mathbb{E}\left[\left(O\left(\sum_i \delta t_j^2 + \delta W_{t_j}^2\right)\right)^2\right] = \sum_j \mathbb{E}[O(\delta t_j^4 + 2\delta t_j^2 \delta W_{t_j}^2 + \delta W_{t_j}^4)] + \sum_{i,j} \mathbb{E}[O(\delta t_j^2 + \delta W_{t_j}^2)(\delta t_i^2 + \delta W_{t_i}^2)]$$

We note that these terms converge to 0 as $\delta t_j \to 0$ because of lemma 2.2.1 and 5.2.1 and because $\mathbb{E}[\delta W_{t_j}^4] = 3\delta t_j^2$.

□

Remark 5.2.1: It can be shown that is is sufficient to state that only f_x, f_{xx} and f_{xt} exist and are continuous. Because this propery will not be used in this thesis and because we only give an outline for the proof we only consider the case for $f \in C^2([0,\infty) \times \mathbb{R})$.

5.3 The general Itô formula

Definition 5.3.1: Itô process:
A stochastic process $\{X_t\}_{t\geq 0}$ of the form

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t b_s dW_s \tag{5.2}$$

is called an Itô process if $\{a_t\}, \{b_t\}$ are adapted to $\{\mathcal{F}_t\}_{t\geq 0}$ such that

$$\mathbb{P}\left(\int_0^T |a_s| ds < \infty\right) = 1$$

and $\left\{b_t \mathbf{1}_{[0,T]}(t)\right\}_{t\geq 0} \in \mathcal{V}$ for all $T > 0$.

Remark 5.3.1: We will often write (5.2) as $dX_t = a_t dt + b_t dW_t$

Theorem 5.3.1: General Itô formula
Consider an Itô process $\{X_t\}_{t\geq 0}$ such that $dX_t = a_t dt + b_t dW_t$ and let $f : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ be a function whose partial derivatives f_t, f_x and f_{xx} exist and are continuous. We further assume $f_x(t, X_t) b_t \mathbf{1}_{[0,T]}(t) \in \mathcal{V}$ for all $T > 0$. Then $f(t, X_t)$ is an Itô process and we have

$$f(t, X_t) - f(0, X_0) = \int_0^t \left(f_t(s, X_s) + f_x(s, X_s) a_s + \frac{1}{2} f_{xx}(s, X_s) b_s^2\right) ds + \int_0^t f_x(s, X_s) dW_s \quad (5.3)$$

almost surely.

Proof
See for example [9] theorem 4.2.

□

Remark 5.3.2: We will keep the notation as introduced in remark 5.3.1. Then we can rewrite (5.3) as

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} a_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} b_t^2\right) dt + \frac{\partial f}{\partial x} b_t dW_t$$

Remark 5.3.3: In theorem 4.2 from [9] it is noted that (5.3) is equivalent to the expression

$$df(t, X_t) = \frac{\partial f}{\partial t}(t, X_t) dt + \frac{\partial f}{\partial x}(t, X_t) dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X_t)(dX_t)^2$$

Chapter 6

Stochastic differential equations

In this chapter we will give an introduction to the theory of stochastic differential equations (SDE). We will give the definition of a SDE and give some examples on how to solve them. Then we will give the conditions for which there exists a unique solution to the SDE.

6.1 Definition and example

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t\}_{t \geq 0}$ be a $\mathbb{P}$-Brownian motion defined on this probability space. Let $\mu : [0, \infty) \times \mathbb{R} \to \mathbb{R}$, $\sigma : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be continuous functions. Suppose we are given the following SDE

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x_0 \tag{6.1}$$

Definition 6.1.1: Let $\{X_t\}$ be a stochastic process which satisfies (6.1). Then $\{X_t\}$ is called a solution to the SDE.

We will now give a theorem for the uniqueness of the solution. This theorem will be proved in the next section.

Theorem 6.1.1: Existence and uniqueness
We are given (6.1) with $\mu, \sigma : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ continuous functions such that $|\mu(t,x) - \mu(t,y)|, |\sigma(t,x) - \sigma(t,y)| \leq K|x - y|$ for some constant $K \geq 0$. Then there exists a unique solution $\{X_t\}$ that is continuous and adapted.

Example 6.1.1: The geometric Brownian motion
The geometric Brownian $\{S_t\}$ motion is used to model the behaviour of a stock. It's SDE is given by $dS_t = \mu S_t dt + \sigma S_t dW_t$. We also have the initial value condition $S_0 = s_0$. We will now solve this equation. Assume $S_t = f(t, W_t)$ then by the Itô-formula we have

$$dS_t = \left(\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \frac{\partial f}{\partial x} dW_t$$

From this we obtain the following set of partial differential equations

$$\begin{cases} \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} & = & \mu f(t,x) \\ \frac{\partial f}{\partial x} & = & \sigma f(t,x) \end{cases}$$

From the second equation we note that $f(t,x) = \phi(t)e^{\sigma x}$ with $\phi(t)$ a function depending only on t. Then by the first equation we can see

$$\phi'(t) + \frac{1}{2}\sigma^2\phi(t) = \mu\phi(t)$$

From which we deduce $\phi(t) = Ce^{(\mu-\frac{1}{2}\sigma^2)t}$ with C a constant. We want $f(0,0) = s_0$ hence the solution for the SDE is given by $f(t,x) = s_0e^{\sigma x+(\mu-\frac{1}{2}\sigma^2)t}$.

To show that this solution is unique we rewrite the SDE as

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \quad \text{if} \quad S_t \neq 0, \quad \forall t$$

which is equivalent to the original SDE according to [9] example 5.1 Now note that $\mu(t,x) = \mu$ and $\sigma(t,x) = \sigma$ are constant hence Lipschitz-continuous functions hence by theorem 6.1.1 the solution is unique a.s.

♦

6.2 Proof of the existence and uniqueness theorem

In this section we will prove theorem 6.1.1. We present the arguments as given by [2] theorem 12.1. We differ in naming the solution from [2]. In [2] the solution is called a strong solution, we call it a solution.

Proof:
We will first prove the existence of a solution. We will begin with $X_t^0 := x_0$ and then define X_t^n inductively for $n \geq 0$ by

$$X_t^{n+1} = x_0 + \int_0^t \mu(s, X_s^n)ds + \int_0^t \sigma(s, X_s^n)dW_s$$

For $n \geq 1$ define

$$\mu^n(t) := \mathbb{E}\left[\sup_{0\leq s\leq t} |X_s^n - X_s^{n-1}|^2\right]$$

Now let $T \geq 1$ then for $0 \leq t \leq T$ we have

$$\begin{aligned}
\mu^1(t) := \mathbb{E}&\left[\sup_{0\leq s\leq t}\left|\int_0^s \mu(0,x_0)dr + \int_0^s \sigma(0,x_0)dW_r\right|^2\right] \\
&\leq 2\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\int_0^s \mu(0,x_0)dr\right|^2\right] + 2\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\int_0^s \sigma(0,x_0)dW_r\right|^2\right] \\
&\leq 2t\mathbb{E}\left[\int_0^t |\mu(0,x_0)|^2 dr\right] + 8\mathbb{E}\left[\int_0^t |\sigma(0,x_0)|^2 dr\right] \\
&\leq 2t^2|\mu(0,x_0)|^2 + 8t|\sigma(0,x_0)|^2 \\
&\leq 10Tt(|\mu(0,x_0)|^2 + |\sigma(0,x_0)|^2)
\end{aligned}$$

Where the first inequality comes from $(a+b)^2 \leq 2(a^2+b^2)$ and the second inequality comes from Doob's $\mathcal{L}^2$-inequality (theorem B.4) because we have a martingale by theorem 4.2.1. We also apply the Cauchy-Schwartz inequality (theorem B.5). Using the same arguments for $n \geq 1$ we have

$$\begin{aligned} \mu^{n+1}(t) &\leq 2\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\int_0^s (\mu(r,X_r^n)-\mu(r,X_r^{n-1}))dr\right|^2\right] + 2\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\int_0^s (\sigma(r,X_r^n)-\sigma(r,X_r^{n-1}))dr\right|^2\right] \\ &\leq 10K^2T\int_0^t \mu^n(r)dr \end{aligned}$$

Then we can show by using an induction argument that for $0 \leq t \leq T$ we have

$$\mu^n(t) \leq C\frac{(10TKt)^n}{n!}$$

For some constant C. Therefore we have

$$\mathbb{E}\left[\sum_{n=1}^{\infty}\sup_{0\leq s\leq T}|X_s^n - X_s^{n-1}|^2\right]^{\frac{1}{2}} \leq \sum_{n=1}^{\infty}\mathbb{E}\left[\sup_{0\leq s\leq T}|X_s^n - X_s^{n-1}|^2\right]^{\frac{1}{2}} < \infty$$

Thus we can see that the sequence $(X_t^n)_{n\geq 0}$ is almost surely uniformly convergent on $[0,T]$. From this we conclude there exists a continuous adapted process $\{X_t\}$ such that

$$\mathbb{P}\left(\lim_{n\to\infty}\sup_{0\leq s\leq t}|X_s^n - X_s| = 0\right) = 1$$

Therefore we have

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\int_0^s \mu(r,X_r^n)dr - \int_0^s \mu(r,X_r)dr\right|^2\right] = 0$$

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\int_0^s \sigma(r,X_r^n)dr - \int_0^s \sigma(r,X_r)dr\right|^2\right] = 0$$

Therefore we conclude that the process $\{X_t\}$ must satisfy the SDE.

Finally we will show the uniqueness. Assume X_t and Y_t are solutions to the SDE. Now we define the function

$$g(t) := \mathbb{E}\left[\sup_{0\leq s\leq t}|X_s - Y_s|^2\right]$$

Using the same arguments as before we obtain:

$$g(t) \leq 10K^2T\int_0^t g(s)ds, \quad 0 \leq t \leq T$$

Then by Gronwall's inequality (theorem B.3) we conclude $g(t) = 0$ hence $X_t = Y_t$ almost surely.

□

Part III

Option pricing

Chapter 7

Basics of finance

In this chapter we will give a short introduction to financial terms. We will cover the time value of money, the law of one price and we will give the definition of an option.

7.1 The time value of money

In this section we will give a short introduction to the notion of time value of money. For a more in depth analysis one can for example see [5] chapter 4 or [6] chapter 2.

Suppose you are given the option to get 10 euro today or 11 euro next year, what should you do? Now assume the interest rate equals $r = 0.2$. If you take the 10 euro today next year you will have 12 euros. Therefore you have made a profit. This gives us an idea that money has different values at different points in time.

More generally assume we currently have a value V_0 and that the interest rate for a period equals r. Then in T periods we will have $V_T = (1+r)^T V_0$. This implies that the *present value* of V_T equals

$$PV(V_T) = \frac{V_T}{(1+r)^T}$$

Now suppose the interest rate in a period equals r and the interest is compounded in m sub-periods. Then the amount we will have at the end of T periods

$$V_T = \left(1+\frac{r}{m}\right)^{mT} V_0$$

In this case we call r the *nominal interest* and $r' := (1+r/m)^m - 1$ the *effective interest*. If we now let m go to infinity we will have continuous compounding. This gives us a familiar limit from calculus

$$\lim_{m\to\infty}\left(1+\frac{r}{m}\right)^{m} = e^r$$

Thus with continuous compounding the present value of V_T equals

$$PV(V_T) = e^{-rT} V_T$$

In this thesis we will look at payments in continuous time and therefore we will use e^{-rT} as the discount factor.

7.2 Options and arbitrage

In this section we will shortly discuss what options are and how we can use the principle of no arbitrage to determine the price of an option. This will be done with a very simple model: the one-period binomial model.

In a market we say there is arbitrage if it is possible to take advantage of a price difference between equivalent goods. If there is an arbitrage opportunity one can expect investors to buy the cheaper product and sell it at the higher price. This is assumed to happen very fast hence the price difference between the goods will disappear. As a result the arbitrage opportunity will disappear as well. Because it is assumed that any price difference between equivalent goods will disappear very fast we will assume the price of equivalent goods will be the same. This is called the *law of one price* and this will be an assumption when pricing options.

An *option* gives the buyer the right to buy or sell a stock at a certain price K which is called the strike price of the option at a predetermined time T called the expiry date. An option that gives the right to buy is called a *call* option and an option that gives the right to sell is called a *put* option.

Let S_T denote the price of the stock at time T. Then the payoff of a call option equals $(S_t - K)^+ := \max(S_T - K, 0)$ and the payoff of a put option equals $(K - S)^+ := \max(K - S_T, 0)$. In general we will denote C_T for the payoff of an option at time T.

Option pricing with the one-period binomial tree

We will now illustrate the principle of option pricing with the law of one price. Assume a stock currently has price S_0. The next period the stock can be either $S_u = uS_0$ or $S_d = dS_0$ with $u > d > 0$. Let $r > 0$ be the risk free interest rate. Note that we must have $u > 1 + r > d$ or else there would be arbitrage. To see why consider the following situations:

Case 1: $1 + r < d$
If you borrow S_0 from the bank at time $t = 0$ and buy the asset. At time $t = 1$ you will have made dS_0 or uS_0 and the amount you owe is $(1 + r)S_0 < uS_0, dS_0$ hence you will have made a profit.

Case 2: $1 + r > u$
If you short sell the asset for S_0 and deposit this amount in the bank. At time $t = 1$ you will have $(1 + r)S_0 > uS_0$ hence you will have made a profit.

With this in mind we will continue with deriving the right price for an option. Now assume there is an option with strike price K. Let C_d be the payoff of the option if the stock goes down and C_u be the payoff if the stock goes up. We want to replicate the payoff of this option by investing some money B_0 in a risk free bond and buying Δ shares of the stock. Then we want

$$B_0(1 + r) + \Delta u S_0 = C_u$$
$$B_0(1 + r) + \Delta d S_0 = C_d$$

Because the portfolio (B_0, Δ) has the same payoff as the option it is equivalent, hence by the law of one price it must have the same value as the option. Therefore the value of the option is given by

$$C_0 = B_0 + \Delta S_0$$

Solving the equations gives us

$$B_0 = \frac{1}{1+r}\left(\frac{-dC_u + uC_d}{u-d}\right) \quad \Delta = \frac{C_u - C_d}{S_0(u-d)}$$

Therefore the price of the option equals

$$C_0 = \frac{1}{1+r}\left(\frac{1+r-d}{u-d}C_u + \frac{u-(1+r)}{u-d}C_d\right) = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}[C_1]$$

where the probability measure $\mathbb{Q}$ defined by $q := \frac{1+r-d}{u-d}$, the probability that the price goes up, is called the risk neutral probability. Note that the discounted expectation of the option price with respect to $\mathbb{Q}$ is a martingale. This will be the central idea in the next section.

7.3 Option pricing by hedging

In this section we will use the the principle of hedging to get an expression for the value of a European option. A European option is an option for which the owner can exercise the right to buy or sell the stock at time $T > 0$ which is called the expiration time of the option. The option can be bought at any time $t \in [0,T]$.

The discounted price of an asset

Let $\{W_t\}_{t\geq 0}$ be a $\mathbb{P}$-Brownian motion. We assume that the price of a stock $\{S_t\}_{t\geq 0}$ follows a geometric Browian motion, i.e.

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{7.1}$$

with $\mu \in \mathbb{R}, \sigma > 0$ constants. The unique solution to (7.1) is given by

$$S_t = S_0 e^{\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t}$$

by example 6.1.1.

Now let $r > 0$ be the risk free interest rate and let $\theta = \frac{\mu - r}{\sigma}$. Define $X_t := W_t + \theta t$ and let $\mathbb{Q}$ be the probability measure from definition 3.3.1. Then by theorem 3.3.3 (Girsanov's theorem) $\mathbb{P}$ and $\mathbb{Q}$ are equivalent and X_t is a $\mathbb{Q}$-Brownian motion (hence a $\mathbb{Q}$-martingale as well). Let $\left\{\tilde{S}_t\right\}_{t\geq 0}$ with $\tilde{S}_t := e^{-rt}S_t$ be a discounted asset price then

$$\begin{aligned}\tilde{S}_t &= e^{-rt}S_t = e^{-rt}S_0 e^{\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t} \\ &= e^{-rt}S_0 e^{\sigma(X_t - \theta t + (\mu - \frac{1}{2}\sigma^2)t} \\ &= S_0 e^{\sigma X_t - \frac{1}{2}\sigma^2 t}\end{aligned}$$

Then by using theorem 5.1.1 we conclude

$$d\tilde{S}_t = \sigma \tilde{S}_t dX_t$$

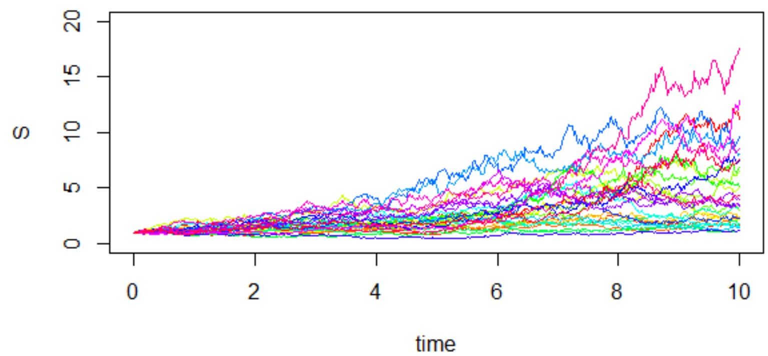

Figure 7.1: Simulation: 30 sample paths of a geometric Brownian motion with $\mu = 0.15$ and $\sigma = 0.25$

Option pricing by hedging

We will construct a portfolio consisting of selling an option, buying a risk free bond B_0 and Δ_t shares of the risky stock. Then the value of the portfolio at time t equals

$$\Pi_t = -V_t + B_0 e^{rt} + \Delta_t S_t$$

If Y_t is a stochastic process we define the discounted process by $\tilde{Y}_t := e^{-rt} Y_t$. Therefore the discounted portfolio process is given by

$$\tilde{\Pi}_t = -\tilde{V}_t + B_0 + \Delta_t \tilde{S}_t$$

Next we choose Δ_t such that Π_t becomes risk free, i.e. $\Pi_t = \Pi_0 e^{rt}$, which we will denote by:

$$\Delta_t = \frac{\partial V}{\partial S}$$

A more rigorous explanation will be given in section 7.4. Note that Δ_t depends on S_t and therefore $\{\Delta_t\}$ is a stochastic process adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$. Note that we have $\tilde{\Pi}_t = \Pi_0$ for all $0 \leq t \leq T$ hence we have $d\tilde{\Pi}_t = 0$. Now let $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$ be a partition of the interval $[0, T]$. Because the processes $\tilde{\Pi}_t, \tilde{V}_t . \tilde{S}_t$ are continuous with probability one we have

$$\tilde{\Pi}_{t_{i+1}} - \tilde{\Pi}_{t_i} = -(\tilde{V}_{t_{i+1}} - \tilde{V}_{t_i}) + \Delta_{t_i}(\tilde{S}_{t_{i+1}} - \tilde{S_{t_i}}) = 0$$

Hence

$$\tilde{V}_{t_{i+1}} - \tilde{V}_{t_i} = \Delta_{t_i}(\tilde{S}_{t_{i+1}} - \tilde{S_{t_i}})$$

From which we can see

$$\tilde{V}_T - \tilde{V}_0 = \sum_{i=0}^{n-1} \Delta_{t_i}((\tilde{S}_{t_{i+1}} - \tilde{S_{t_i}})$$

Then taking the limit for $n \to \infty$

$$\tilde{V}_T = \tilde{V}_0 + \int_0^T \Delta_t d\tilde{S}_t = \tilde{V}_0 + \int_0^T \Delta_t \sigma \tilde{S}_t dX_t \tag{7.2}$$

And because X_t is a $\mathbb{Q}$-Brownian motion we conclude by theorem 4.2.1 that $\tilde{V}_t$ is a martingale. Now consider the interval $[t,T]$ we then have

$$\tilde{V}_T = \tilde{V}_t + \int_t^T \Delta_s \sigma d\tilde{S}_s$$

From corollary 4.2.1 and by using that $\{\tilde{V}_t\}_{0\leq t\leq T}$ is a martingale we conclude that

$$\tilde{V}_t = \mathbb{E}^{\mathbb{Q}}[\tilde{V}_t|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[\tilde{V}_T|\mathcal{F}_t]$$

Then using the definition of the discounted process.

$$e^{-rt}V_t = \mathbb{E}^{\mathbb{Q}}[e^{-rT}V_T|\mathcal{F}_t]$$

Rewriting this expression then gives us

$$V_t = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[V_T|\mathcal{F}_t]$$

In particular we have

$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}[V_T]$$

We can now state a general solution for the option pricing problem.

Theorem 7.3.1: For a European option with expiration date T and payoff C_T the value at $t=0$ is given by

$$V_0 = \mathbb{E}^{\mathbb{P}}[e^{-rT}C_T(S_0 e^{\sigma W_T + (r-\frac{1}{2}\sigma^2)T})] \tag{7.3}$$

Proof:

$$\begin{aligned} V_0 &= \mathbb{E}^{\mathbb{Q}}[e^{-rT}C_T(S_T)] \\ &= \mathbb{E}^{\mathbb{Q}}[e^{-rT}C_T(e^{rT}\tilde{S}_T)] \\ &= \mathbb{E}^{\mathbb{Q}}[e^{-rT}C_T(e^{rT}S_0 e^{\sigma X_T - \frac{1}{2}\sigma^2 T})] \\ &= \mathbb{E}^{\mathbb{Q}}[e^{-rT}C_T(S_0 e^{\sigma X_T + (r-\frac{1}{2}\sigma^2)T})] \\ &= \mathbb{E}^{\mathbb{P}}[e^{-rT}C_T(S_0 e^{\sigma W_T + (r-\frac{1}{2}\sigma^2)T})] \end{aligned}$$

□

7.4 The Black-Scholes-Merton equation

In the previous section we stated that we chose Δ_t to be $\frac{\partial V}{\partial S}$ without exactly stating what this meant. A more rigorius explanation will be given in this section which is based on [8]. Let $V(t,S_t)$ denote the value of the option at time t. Note that this is a function of S_t which is an Itô process hence we can apply theorem 5.3.1 to see

$$dV = \left(\frac{\partial V}{\partial t} + \frac{\partial V}{\partial S}\mu S_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2\right)dt + \frac{\partial V}{\partial S}\sigma S_t dW_t$$

Next like in the previous section we want to construct a portfolio Π_t by longing the option and shorting the risky stock. Let Δ_t denote the number of stocks we have at time t. Then if we trade at the time interval dt the change in value of the portfolio will equal $d\Pi_t = dV - \Delta_t dS_t$. If we assume Δ_t is fixed during the time-step we find

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \frac{\partial V}{\partial S}\mu S_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2 - \mu S_t \Delta_t\right) dt + \left(\frac{\partial V}{\partial S} - \Delta_t\right)\sigma S_t dW_t$$

Then by choosing $\Delta_t = \frac{\partial V}{\partial S}$ the term with the Brownian motion is cancelled and we are left with a deterministic expression. This leaves us with

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

Now note that because Π_t has become risk free the growth rate of this portfolio must equal $d\Pi_t = r\Pi_t dt$ because of the law of one price. This leaves us with the following equation:

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

No we can replace Π_t by $V_t - \Delta_t S_t$ and $\Delta_t = \frac{\partial V}{\partial S}$ this gives us the following partial differential equation.

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

This is called the Black-Scholes-Merton equation.

Chapter 8

Solution to the option pricing problem

In this chapter we will give the solution for the option pricing problem for a European call option.

8.1 The probability distribution of the asset price

In the previous chapter is was shown that the option price follows the geometric Brownian motion i.e. $S_T = S_0 e^{\sigma W_T + (\mu - \frac{1}{2}\sigma^2)T}$. This expression can be rewritten as

$$W_T = \frac{\log(S_T/S_0) - (\mu - \frac{1}{2}\sigma^2)T}{\sigma}$$

Let $y = y(x)$ be the realisation of W_T then the pdf of W_T can be written as

$$f_{W_T}(y(x)) = \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{1}{2T}\left(\frac{\log(x/S_0) - (\mu - \frac{1}{2}\sigma^2)T}{\sigma}\right)^2\right)$$

Therfore the price of the option is given by

$$\begin{aligned}
e^{-rT}\mathbb{E}[C_T(S_T)] &= e^{-rT}\int_0^\infty C_T(x) f_{S_T}(x)dx \\
&= e^{-rT}\int_0^\infty C_T(x) f_{W_T}(y(x))\frac{dy}{dx}dx \\
&= e^{-rT}\int_0^\infty \frac{C_T(x)}{\sigma x\sqrt{2\pi T}} \exp\left(-\frac{1}{2T}\left(\frac{\log(x/S_0) - (\mu - \frac{1}{2}\sigma^2)T}{\sigma}\right)^2\right)dx
\end{aligned}$$

To apply the risk-neutral method we substitute $\mu = r$ giving us the following expression.

$$V_0 = e^{-rT}\int_0^\infty \frac{C_T(x)}{\sigma x\sqrt{2\pi T}} \exp\left(-\frac{1}{2T}\left(\frac{\log(x/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma}\right)^2\right)dx$$

8.2 Solution to the option pricing problem

In this section we will determine the value of a European call option with strike price K. The payoff function in this case equals $C(S_T) = (S_T - K)^+$. By using the proof of theorem 7.3.1 we have

$$\begin{aligned}
V_0 &= e^{-rT}\mathbb{E}^{\mathbb{Q}}[(S_T - K)^+] \\
&= e^{-rT}\mathbb{E}^{\mathbb{Q}}[(e^{rT}\tilde{S}_T - K)^+] \\
&= e^{-rT}\mathbb{E}^{\mathbb{Q}}[(S_0 e^{\sigma X_T + (r-\frac{1}{2}\sigma^2)T} - K)^+] \\
&= e^{-rT}\mathbb{E}^{\mathbb{Q}}[(S_0 e^{\sigma\sqrt{T}Z + (r-\frac{1}{2}\sigma^2)T} - K)^+] \\
&= e^{-rT}\int_{\mathbb{R}} (S_0 e^{\sigma\sqrt{T}x + (r-\frac{1}{2}\sigma^2)T} - K)^+ \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx
\end{aligned}$$

Note that $(S_T - K)^+ \geq 0$ if $S_T \geq K$ therefore the integration region is given by

$$\left\{ x \in \mathbb{R} \mid S_0 e^{\sigma\sqrt{T}x + (r-\frac{1}{2}\sigma^2)T} \geq K \right\}$$

We can see that this is equal to

$$\left\{ x \in \mathbb{R} \mid x \geq x_0 := \frac{\log(K/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \right\}$$

Therefore the price of the option equals

$$e^{-rT}\int_{x_0}^{\infty} (S_0 e^{\sigma\sqrt{T}x + (r-\frac{1}{2}\sigma^2)T} - K) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \tag{8.1}$$

Now we define

$$d_1, d_2 = \frac{\log(S_0/K) + (r \pm \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

Let $y = -x$ then $dy = -dx$ hence we can see (8.1) equals

$$e^{-rT}\int_{-\infty}^{d_2} \frac{S_0}{\sqrt{2\pi}} e^{-\sigma\sqrt{T}y + (r-\frac{1}{2}\sigma^2)T} e^{-\frac{1}{2}y^2} dy - e^{-rT}K\int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy \tag{8.2}$$

This equation can then be rewritten as

$$\begin{aligned}
V_0 &= S_0\int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\sigma\sqrt{T})^2} dy - e^{-rT}K\int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy \\
&= S_0\Phi(d_2 + \sigma\sqrt{T}) - e^{-rT}K\Phi(d_2) \\
&= S_0\Phi(d_1) - e^{-rT}K\Phi(d_2)
\end{aligned}$$

Where $\Phi(.)$ denotes the CDF of a $\mathcal{N}(0,1)$ distribution.

Conclusion

Solution to the option pricing problem

Suppose we have an option regarding some risky stock S_t. We assumed that the risky stock follows a geometric Brownian motion i.e.

$$dS_t = \mu S_t st + \sigma S_t dW_t$$

For which we know there exists an almost surely unique solution by example 6.1.1. We then constructed a portfolio by selling the option, investing B_0 in a risk free bond and buying Δ_t shares of the risky stock.

$$\Pi_t = -V_t + B_0 e^{rt} + \Delta_t S_t$$

where $r > 0$ denotes the risk free interest rate. By chosing Δ_t in such a way that the portfolio becomes risk free so that $d\tilde{\Pi}_t = 0$ hence we found,

$$\tilde{V}_T = V_0 + \int_0^T \sigma \Delta_t \tilde{S}_t dX_t$$

We defined $X_t := W_t + \frac{\mu - r}{\sigma} t$ which was a Brownian motion with respect to the probability measure $\mathbb{Q}$ by theorem 3.3.3. Now we can see that the integrand of the integral is a $\mathbb{Q}$-Brownian motion hence the integral is an Itô integral. Thus by theorem 4.2.1 we know that $\tilde{V}_t$ is a martingale. In particular we have

$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}[V_T] = \mathbb{E}^{\mathbb{P}}[C_T(S_0 e^{\sigma W_T + (r - \frac{1}{2}\sigma^2)T})]$$

In particular, for a European call option we have $V_T = (S_T - K)^+$. Then using that S_T follows a geometric Brownian motion we obtain the particular solution:

$$V_0 = S_0 \Phi(d_1) - e^{-rT} K \Phi(d_2)$$

Further research

In this thesis we followed the assumption that the price of a risky stock kan be described by a Geometric Brownian motion. However as mentioned by my supervisor Dr. K. Dajani one can also use a Lévy process to describe the randomness of the stock instead of a Brownian motion. An advantage of a Lévy process is that it allows jumps in its sample paths where the Brownian motion does not. For more information once can see for example [11].

Furthermore we mainly focussed on the pricing of European options. There are however other types of options such as American and Asian options. This is also an interesting topic to explore further.

Appendix A

Code for simulation

Here we will present the code used for the simulation in this thesis. The programme used for simulation is R. For more information about this programe see for example

https://www.r-project.org

The codes are based on the codes presented by [2] which are written in Matlab.

The code for figure 3.1: [2] Simulation 7.1

```
# In this program we will simulate sample paths of the Brownian motion

N <- 200            # number of time steps
Y <- 20
dt <- Y/N
M <- 30             # number of sample paths

time <- seq(0,Y,by=dt)

W <- matrix(0,nrow=M,ncol=N+1)
dW <- sqrt(dt)*matrix(rnorm(N*M),M)

for(i in 1:N){
  W[,i+1] <- W[,i]+dW[,i]
}

cl <- rainbow(M)
plot(time,W[1,],col=cl[1],type="l",ylab="W",
     ylim=c(-15,15))
for(j in 2:M) lines(time,W[j,],col=cl[j])
```

The code for figure 3.2: [2] Simulation 8.1

```
# In this program we will simulate sample paths of a Brownian motion with drift

N <- 300                              # Number of time steps
Y <- 50
dt <- Y/N
```

```
theta <- 0.5

time <- seq(0,Y,by=dt)
M <- 30                              # Number of sample paths
X <- matrix(0,ncol=N+1,nrow=M)
dW <- sqrt(dt)*matrix(rnorm(M*N),M)

for(j in 1:M){
  for(i in 1:N){
    X[j,i+1] <- X[j,i]+dW[j,i]+theta*dt
  }
}

cl <- rainbow(M)
plot(time,X[1,],col=cl[1],type="l",ylab="W",ylim=c(-5,40),xlim=c(0,50))
for(u in 2:M) lines(time,X[u,],col=cl[u])
lines(time,theta*time)
```

Code for figure 4.1: [2] Simulation 10.2

```
# In this program we will simulate sample paths of a stochastic process
# defined by the Ito integral

Y <- 3
M <- 30          # number of sample paths
N <- 500          # number of time steps
dt <- Y/N
dW <- sqrt(dt)*matrix(rnorm(M*N),M)
time <- seq(0,Y,by=dt)



# We will first simulate the sample paths of the Brownian motion again.
W <- matrix(0,ncol=N+1,nrow=M)
for(i in 1:N){
  W[,i+1] <- W[,i]+dW[,i]
}

# No we will use the saple paths to simulate the stochastic process
# defined by the Ito integral
Integral <- matrix(0,nrow=M,ncol=N+1)
for(j in 1:N){
  Integral[,j+1] <- Integral[,j]+W[,j]*dW[,j]
}

# Now we will plot the sample paths of the integral process
cl <- rainbow(M)
plot(time,Integral[1,],col=cl[1],type="l",ylab="Ito integral",
     ylim=c(-3,10))
for(u in 2:M) lines(time,Integral[u,],col=cl[u])
```

Code for figure 6.1: [2] Simulation 11.3

```
# In this program we will simualte 30 sample paths for the geometric
# Brownian motion

N <- 500            # Number of time steps
M <- 30             # Number of sample paths
mu <- 0.15          # Drift coefficent
sigma <- 0.25       # Standard deviantion
Y <- 10
dt <- Y/N
time <- seq(0,Y,by=dt)

# We will now simualte the process
dW <- sqrt(dt)*matrix(rnorm(N*M),M)
S_0 <- 1                              # Starting value of the Geo. Brow. Motion
S <- matrix(S_0,ncol=N+1,nrow=M)
for(i in 1:N){
  S[,i+1]<-S[,i]+mu*S[,i]*dt+sigma*S[,i]*dW[,i]
}

# We will now plot the sample paths
cl <- rainbow(M)
plot(time,S[1,],col=cl[1],type="l",ylab="S",
     ylim=c(0,20))
for(j in 2:M) lines(time,S[j,],col=cl[j])
```

Appendix B

Additional results

In this appendix we will shortly mention and prove some results from analysis and measure theoretic probability which will be used in this thesis.

Theorem B.1: Let $(X, ||\cdot||)$ be a normed space then every convergent sequence $(x_n) \subset X$ is a Cauchy-sequence.

Proof:
Let $\varepsilon > 0$. Because (x_n) is a convergent sequence, say it converges to some $x \in X$, there exists $N \in \mathbb{N}$ such that for all $m, n > N$ we have $||x_n - x||, ||x_m - x|| < \frac{\varepsilon}{2}$. Then by the triangle inequality we have

$$||x_n - x_m|| = ||x_n - x + x - x_m|| \leq ||x_n - x|| + ||x - x_m|| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Thereby proving the lemma.

□

Theorem B.2: Let $\mathcal{V}$ denote the set of stochastic processes as defined in section 4.1 and let $f \in \mathcal{V}$. Then there exists a sequence of simple processes (f_n) converging to f.

Proof:
We will present the proof as given by [9](p.24-25)

Lemma B.2.1: Let $g \in \mathcal{V}$ be bounded and continuous for every $\omega \in \Omega$. Then there exists a sequence (f_n) of simple processes such that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T (g - f_n)^2 dt\right] = 0$$

Proof:
Define $f_n(t,\omega) = \sum_{j=0}^{n-1} g(t_j,\omega)\mathbf{1}_{[t_j,t_{j+1}]}$ then f_n is a simple process because $g_n(t_j,\cdot)$ is $\mathcal{F}_{t_j}$ measurable. Note that $\lim_{n\to\infty} \int_0^T (g - f_n)^2 dt = 0$ because $g(\cdot,\omega)$ is continuous for every ω. Therefore we have

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T (g - f_n)^2 dt\right] = 0$$

by the bounded convergence theorem.

□

Lemma B.2.2: Let $h \in \mathcal{V}$ be bounded. Then there exist a sequence of bounded functions $(g_n) \subset \mathcal{V}$ that are continuous for every ω and n and

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T (h-g_n)^2 dt\right] = 0$$

Proof:
Assume $|h(t,\omega)| \leq M$ for all (t,ω). For every n let ϕ_n be a non-negative function continuous function on $\mathbb{R}$ defined such that

(i) $\phi_n(x) = 0$ if $x \leq -\frac{1}{n}$ and $x \geq 0$

(ii) $\int_{\mathbb{R}} \phi_n(x)dx = 1$

Define

$$g_n(t,\omega) = \int_0^t \phi_n(s-t)h(s,\omega)ds$$

Then $g(\cdot,\omega)$ is continuous for each ω and $|g_n(t,\omega)| \leq M$. Because $h \in \mathcal{V}$ we note that $g_n(t,\cdot)$ is $\mathcal{F}_t$ measurable for all t. From this $\lim_{n\to\infty} \int_0^T (h-g_n)^2 dt = 0$ for every ω. Hence by the bounded convergence theorem we have

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T (h-g_n)^2 dt\right] = 0$$

□

Lemma B.2.3: Let $f \in \mathcal{V}$. Then there exists a sequence $(h_n) \subset \mathcal{V}$ of bounded functions such that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T (f-h_n)^2 dt\right] = 0$$

Proof:
Define

$$h_n(t,\omega) = \begin{cases} -n & \text{if} \quad f(t,\omega) < -n \\ f(t,\omega) & \text{if} \quad -n \leq f(t,\omega) \leq n \\ n & \text{if} \quad f(t,\omega) > n \end{cases}$$

Then the statement follows from the dominated convergence theorem.

□

We will now conclude the proof of theorem B.2. Let $f \in \mathcal{V}$. Then by lemma B.2.1, B.2.2 and B.2.3 we can find a sequence of simple functions (f_n) such that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T (f(t,\omega)-f_n(t,\omega))^2 dt\right] = 0$$

thereby proving the theorem.

□

Theorem B.3: Gronwall's Inequality:
Let $g : [0,a] \to [0,\infty)$ be a continuous function. Assume there are constants $C, K \geq 0$ such that

$$g(t) \leq C + K\int_0^t g(d)ds$$

for every $t \in [0,a]$ then $g(t) \leq Ce^{Kt}$ for every $t \in [0,a]$

Proof:
Assume $C > 0$ and define the function

$$G(t) = C + K\int_0^t g(s)ds$$

We then have $G(t) \geq g(t)$ and $G(t) \geq C > 0$. We have $G'(t) = Kg(t) \leq KG(t)$ hence we see $\frac{d}{dt}\log G(t) \leq K$. Also note that $G(0) = C$ hence $\log G(t) \leq \log G(0) + tK$. From this we conclude

$$g(t) \leq G(t) \leq G(0)e^{Kt} = Ce^{Kt}$$

□

The following two theorems are stated in [2] theorem 12.1

Theorem B.4: Doob's $\mathcal{L}^2$ inequality
Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a probability space with a filtration. Let $\{M_t\}_{t\geq 0}$ be a martingale defined on this probability space. Let $p \in (1, \infty)$ then for $T > 0$ we have

$$\left|\left|\sup_{0\leq t\leq T} M_t\right|\right|_2 \leq 2||M_T||_2$$

Theorem B.5: Cauchy-Schwarz:
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{Y_s\}_{0\leq s\leq t}$ be a predictable process such that $Y_s(\omega)$ is continuous for almost every $\omega \in \Omega$. Then

$$\sup_{0\leq s\leq t}\left|\int_0^s Y_r dr\right|^2 \leq t\int_0^t Y_r^2 dr$$

Bibliography

[1] Rene L. Schilling *Measures, Integrals and Martingales*. Cambridge University Press 2017, ISBN: 978-1-316-62024-3

[2] Geon Ho Choe *Stochastic Analysis for Finance with Simulations* Springer 2016, ISBN: 978-3319-255897

[3] P.J.C. Spreij *Measure Theoretic Probability* Universiteit van Amsterdam, Augustus 2016

[4] F. Ziltener *Maat en Integratie, WISB312* Universiteit Utrecht, 3 November 2017

[5] Jonathan Berk, Peter DeMarzo *Corporate Finance* Pearson Education Limited 2011, ISBN: 978-0-27-375603-3

[6] S. David Promislow *Fundamentals of Actuarial Mathematics* John Wiley & Sons Ltd 2011, ISBN: 978-0-470-68411-5

[7] http://www.math.ucla.edu/ biskup/275b.1.14w/PDFs/conditional-expectation.pdf

[8] R. Chan *http://www.math.cuhk.edu.hk/ rchan/teaching/math4210/chap08.pdf*

[9] B. Øksendal *Stochastic Differential Equations, An introduction with applications* Springer, 1995, ISBN 3-540-60243-7

[10] Donald L. Cohn *Measure theory* Birkhäuser, 2013, ISBN 978-1-4614-6955-1

[11] A. Papapantoleon *An introduction to Levy processes with applications in finance* http://www.math.ntua.gr/ papapan/papers/introduction.pdf