UNIVERSITAT DE BARCELONA

Final degree project

MATHEMATICS DEGREE
BUSINESS MANAGEMENT DEGREE

Faculty of Mathematics and Computer Science
Faculty of Economics and Business
University of Barcelona

An introduction to the mathematical cornerstone of financial derivatives

Author:	Adrià Morera Morales
Directors:	Dr. Josep Vives i Santa-Eulàlia
	Dr. Oriol Roch Casellas
Realised in:	Department of Mathematics and Computer Science
	Department of Economic, Financial and Actuarial Mathematics
Barcelona,	January 18, 2019

Abstract

There are several paths that lead to the Black-Scholes formula. This project discusses two of them.

Chapters 2 and 3 depart from the discrete Cox-Ross-Rubinstein model of prices and reveal the Black-Scholes formula for European calls and puts.

Chapters 4 and 5 go one step further by considering since inception the continuous modelling of prices, in which a new concept of integral must be defined in order to formulate the Black-Scholes hypotheses from a stochastical point of view.

The project ends up debating the uses of derivatives and the appropriateness of the Black-Scholes model in the real world. Moreover, the annex contains Numerical Methods that implement the models covered in this project.

Acknowledgements

I owe special thanks to Dr. Josep Vives and Dr. Oriol Roch for having tutored me through this compelling topic and for having shared with me their time.

I must also thank Patrícia Crehuet, Óscar Pérez, Eugènia Serret, Laura Cadefau, Iván Marañón, Eugenio Tubio, Maria Camarasa, Martí Dalmases, Javier Carrasco, Maria Forés, Estefanía Alba, Marga García, Daniel Latre and José Miguel Barguilla for having introduced me to the world of financial derivatives and for all I learnt from them during my internship in the department of Product Development, DCM & FICC Sales, CaixaBank.

Last but not least, I thank my mother, my father and my brother for their continuous and unconditional love and care.

Contents

1 Introduction

1.1 A scheme of what is included in this project and how it is structured

My initial intention was to explain the mathematical results that conform the basis of discrete and continuous derivatives valuation models. My goal was to establish an analogy between the discrete theory rationale and the continuous theory rationale so that one could be better understood thanks to the other, and viceversa.

Thus, in Chapters 2 and 3 not only the CRR model - the discrete modelling of the evolution of the price of the underlying asset - is explained, but also the computation of the prices of European calls and puts - the valuation of derivatives.

By contrast, the reader will notice that in Chapters 4 and 5 only the continuous modelling of the evolution of the price of the underlying asset is discussed. I considered that the rigorous definition of the stochastic integral was of paramount importance in order to make a proper introduction of continuous modelling. Due to the established limit of number of pages, I had to leave continuous valuation models out of discussion. Because of the mentioned considerations, the stochastic integral became the main topic of this project.

1.2 Derivatives: usefulness and benefits

A derivative is a financial instrument that derives its performance from the performance of an underlying asset and that is created by means of a legal contract between two parties: the buyer and the seller of the contract.

Derivatives resemble insurance in the sense that both allow for the transfer of risk from the buyer to the seller of the contract, and/or viceversa. This transferred risk from one party to the other is usually the value of the underlying asset of the derivative, although it can also be some variable that is a function of the value of the underlying asset. Common derivatives underlyings are equities, fixed-income securities, currencies, and commodities, but may also be interest rates, credit, energy, weather, and even other derivatives.

There are two general classes of derivatives:

1. Forward commitments

They force the two parties to transact in the future at a previously agreed-on price. Forward contracts, futures contracts and swaps are the three types of forward commitment.

2. Contingent claims

They provide the right but not the obligation to buy (call) or sell (put) the underlying at a pre-determined price. The choice of buying or selling versus doing nothing depends on a particular random outcome. The primary contingent claim is called an option.

Derivatives allow market participants to practice more effective risk management.

Definition 1.1. *Risk management is the dynamic and ongoing process by which an organisation or an individual defines the level of risk it wishes to take, measures the level of risk it is taking and adjusts the latter to equal the former*[1]

In the pre-derivatives era, it was compulsory to engage in transactions in the underlyings if one wanted to set his actual level of risk to the desired level of risk. The problem was that such transactions typically had high transaction costs. Thanks to derivatives, one can trade the risk without trading the underlying itself.

Moreover, the advent of derivatives brought several benefits:

1. Information discovery

One of the characteristics of derivatives is their relatively high degree of leverage, which translates into less capital required to take a position in the derivatives market. Hence, information can flow into the derivatives market before it gets into the spot market.

Moreover, the price of an option, as will be shown in subsequent sections of this project, reflects two characteristics of the option (exercise price and time to expiration), three characteristics of the underlying (price, volatility and cash flows it might pay) and one general macroeconomic factor (risk-free rate). The only factor among those mentioned which is not easy to identify is volatility. Using the Black-Scholes model and some easily programmable non-linear solver, volatility that market participants are indeed using in their trade executions can be inferred. This volatility, called implied volatility, measures the expected risk of the underlying.

2. Operational Advantages

A derivatives market has lower transaction costs than the transaction costs of the underlyings. In addition, a derivatives market typically has greater liquidity than the underlying spot markets and short positions (simultaneous borrowing and selling) are easy to be taken.

3. Market efficiency

Competition, the relatively free flow of information and ease of trading tend to bring prices back in line with fundamental values. When prices deviate from fundamental values, derivative markets offer less costly ways to exploit the mispricing, which makes the financial markets function more effectively.

[1] Quoted from [8].

But how does a derivatives market work? How can the price of derivatives be ascertained? This is the pricing issue. Furthermore, how does the selling party cover itself from the potential losses that the contract entails? This is the hedging issue.

Forwards and futures are in essence the same financial instrument; the difference between the two only lies in how the contract is entered into by the parties: either through a futures exchange or through an over-the-counter (OTC) market participant. Swaps can be thought of as strings of futures or forward contracts, meaning that a swap can be decomposed into a series of forwards -if the swap contract is entered into through an OTC market participant- or into a series of futures -if the swap contract is entered into through a futures exchange-.

Likewise, a forward contract or a futures contract can be thought of as the simultaneously buying of a call option and selling of a put option. Therefore, the study of how futures, forwards and swaps can be priced and hedged is implicit in the study of how call and put options can be priced and hedged and that, indeed, is the focus of this project.

2 The Cox-Ross-Rubinstein model

This chapter is based on [14], the notes of the subject *Modelització* taught in the Mathematics degree of the University of Barcelona by Dr. Josep Vives i Santa-Eulàlia.

2.1 Definition of the model

Assume a market model with only two assets:

1. A riskless asset that evolves in a deterministic manner S^0.
2. A risky asset S, that evolves stochastically.

Assume time is discrete and finite, i.e. $\mathbb{T} := \{1, 2, ..., N\}, N \in \mathbb{N}$. Assume a constant interest rate $r \geq 0$ so that:

$$S^0 = \left\{S_t^0 = (1+r)^t, t \in \mathbb{T}\right\}, S_0^0 := 1$$

Assume the price of the risky asset can only go up or down at each step in time at a rate $u \geq 0$ for up moves or at a rate $d \leq r$[2] for down moves so that:

$$S_t \in \{(1+d) \cdot S_{t-1}, (1+u) \cdot S_{t-1}\}, \forall t \in \mathbb{T}, S_0 \in \mathbb{R}^+$$

Defining the random variable $T_t = \frac{S_t}{S_{t-1}}$, we can define the market model $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where:

- $\Omega = \left\{(1+d)^N, (1+d)^{N-1} \cdot (1+u), ..., (1+d) \cdot (1+u)^{N-1}, (1+u)^N\right\}$.
- $\mathbb{F} = \{\mathcal{F}_t, t \in \mathbb{T}\}$ with $\mathcal{F}_t := \sigma\{S_0, T_1, T_2, ..., T_N\}$.
- $\mathcal{F} = \mathcal{F}_N$.
- $\mathbb{P}$ is a probability measure over Ω such that:

$$\mathbb{P}(T_1 = x_1, T_2 = x_2, ..., T_N = x_N) > 0, \forall (x_1, ..., x_N) \in \Omega.$$

2.2 Viability and completeness of the model

We will begin this chapter with some definitions and some well-known results.

Definition 2.1. *A portfolio is a set of risky assets together with a riskless asset.*

[2] d does not have to be negative in order to cause a down move. Any value of d under r provokes a down move.

Definition 2.2. *An investment strategy is a series of random vectors*

$$\phi_n = \left\{ (\phi_n^0, \phi_n^1, ..., \phi_n^d) \in \mathbb{R}^{d+1}, n \in \mathbb{T} \backslash \{0\} \right\}$$

where, for every $i \in \{0, 1, ..., d\}$, ϕ_n^i is $\mathcal{F}_{n-1}$-measurable.

Definition 2.3. *The value of a portfolio at the end of day n is*

$$V_n(\phi) = \phi_n \times S_n$$

where $S_n = \left(S_n^0, S_n^1, ..., S_n^d\right)$ and $\times$ refers to the dot product in $\mathbb{R}^{d+1}$.

The discounted value of the portfolio is

$$\tilde{V}_n(\phi) = \phi_n \times \frac{S_n}{(1+r)^n} = \phi_n \times \tilde{S}_n,$$

and we refer to $\tilde{S}_n$ as the discounted price of the asset.

Definition 2.4. *An autofinanced strategy is an investment strategy such that*

$$\phi_n \times S_n = \phi_{n+1} \times S_n$$

Definition 2.5. *We say that ϕ is an admissible investment strategy if*

$$V_n(\phi) \geqslant 0, \forall n \in \mathbb{T}.$$

Definition 2.6. *We say that an autofinanced and admissible investment strategy ϕ is an arbitrage opportunity if $V_0(\phi) = 0$ and $V_N(\phi) > 0$ with strictly positive probability.*

Definition 2.7. *We say that a market is feasible if there exist no arbitrage opportunities in that market.*

Definition 2.8. *We say that a probability measure $\mathbb{P}^*$ is equivalent to $\mathbb{P}$ if*

$$\mathbb{P}(A) = \mathbb{P}^*(A), \forall A \in \mathcal{F}$$

We will refer to this relationship between $\mathbb{P}$ and $\mathbb{P}^*$ as $\mathbb{P} \sim^* \mathbb{P}^*$.

Theorem 2.9. *The first fundamental theorem of asset pricing asserts that:*

A finite market is feasible if and only if there exists a probability measure $\mathbb{P}^* \sim^* \mathbb{P}$ such that the discounted prices of the risky assets $\tilde{S}^i$, $i = 1, ..., d$, are martingales.

Theorem 2.10. *The second fundamental theorem of asset pricing asserts that:*

A feasible market is complete if and only if there exists a unique risk-neutral probability measure $\mathbb{P}^*$.

A proof of the first and second fundamental theorems of asset pricing can be found in [14].

Proposition 2.11. *The CRR model does not admit arbitrage opportunities if and only if $d < r < u$.*

Proof. We first prove that if the CRR model does not admit arbitrage opportunities then $d < r < u$. We will demonstrate this result by assuming that $d < r < u$ is false and concluding that the CRR model does admit arbitrage opportunities.

$$\neg (d < r < u) \Rightarrow (d \geqslant r) \vee (r \geqslant u)$$

1. Case $d \geqslant r$

Assume we borrowed S_0 and we bought the risky asset. Then, for every $t \in \mathbb{T}$, the value of our portfolio would be $V_t := S_t - S_0 \cdot (1+r)^t$. In the worst scenario, i.e. in the case that the price of the risky asset fell at each $t \in \mathbb{T}$, the value of our portfolio would be $S_0 \cdot (1+d)^t - S_0 \cdot (1+r)^t$. By hypothesis, $d \geqslant r \Rightarrow 1+d \geqslant 1+r \Rightarrow (1+d)^t \geqslant (1+r)^t \Rightarrow S_0 \cdot (1+d)^t - S_0 \cdot (1+r)^t =: V_t \geqslant 0$. In particular, $V_N \geqslant 0$, so there exist arbitrage opportunities in the CRR model.

2. Case $r \geqslant u$

Assume we short sold the risky asset, i.e. we borrowed the risky asset and sold it. Because in the market model only exist two assets, the risky asset and the riskless asset, we would invest the proceeds of this short selling in the riskless asset. The value of our portfolio at each $t \in \mathbb{T}$ would in this case be $V_t := S_0 \cdot (1+r)^t - S_t$. In the worst scenario, the price of the risky asset would go up at each step $t \in \mathbb{T}$ and we would have to buy it in order to honor our borrowing contract by returning the risky asset to the counterpart from whom we borrowed the risky asset. In this scenario, the value of our portfolio would be $S_0 \cdot (1+r)^t - S_0 \cdot (1+u)^t$. By hypothesis, $r \geqslant u \Rightarrow 1+r \geqslant 1+u \Rightarrow (1+r)^t \geqslant (1+u)^t \Rightarrow S_0 \cdot (1+r)^t - S_0 \cdot (1+u)^t =: V_t \geqslant 0$. In particular, $V_N \geqslant 0$, so there exist arbitrage opportunities in the CRR model.

We now turn to prove that if $d < r < u$ then arbitrage opportunities do not exist in the CRR model. By hypothesis, $r \in (d, u)$. Let us define the following application:

$$\begin{aligned} \mathbb{P}^* \colon \{1+d, 1+u\} &\to [0,1] \\ 1+u &\mapsto \mathbb{P}^*(T_t = 1+u) =: p^* = \frac{r-d}{u-d} \\ 1+d &\mapsto \mathbb{P}^*(T_t = 1+d) =: 1-p^* = \frac{u-r}{u-d} \end{aligned}$$

Let $T_1, T_2, ..., T_N$ be *iid* random variables, i.e. independent and identically distributed random variables, with Bernouilli law, with parameter p^*, over $\{1+d, 1+u\}$.

$$T_1, T_2, ..., T_N \Rightarrow E^*[T_t|\mathcal{F}_{t-1}] = E^*[T_t] := (1+d) \cdot \mathbb{P}^*(T_t = 1+d)$$
$$+ (1+u) \cdot \mathbb{P}^*(T_t = 1+u) = (1+d) \cdot \frac{r-d}{u-d} + (1+u) \cdot \frac{u-r}{u-d} = 1+r, \forall t \in \mathbb{T}$$

$$\Rightarrow 1+r = E^*[T_t|\mathcal{F}_{t-1}],\ \forall t \in \mathbb{T} \Rightarrow 1 = E^*\left[T_t \cdot \frac{1}{1+r}|\mathcal{F}_{t-1}\right] = E^*\left[\frac{S_t}{S_{t-1}} \cdot \frac{1}{1+r}|\mathcal{F}_{t-1}\right]$$

$$= E^*\left[\frac{S_t}{S_{t-1}} \cdot \frac{(1+r)^{t-1}}{(1+r)^t}|\mathcal{F}_{t-1}\right] = E^*\left[\frac{\frac{S_t}{(1+r)^t}}{\frac{S_{t-1}}{(1+r)^{t-1}}}|\mathcal{F}_{t-1}\right] = E^*\left[\frac{\tilde{S}_t}{\tilde{S}_{t-1}}|\mathcal{F}_{t-1}\right],$$

$$\forall t \in \mathbb{T} \Rightarrow E^*\left[\tilde{S}_t|\mathcal{F}_{t-1}\right] = \tilde{S}_{t-1},\ \forall t \in \mathbb{T}$$

$$\Rightarrow \tilde{S} \text{ is a } \mathbb{P}^*\text{-martingale}$$

Hence, $\mathbb{P}^*$ satisfies the hypotheses of the first fundamental theorem of asset pricing. Therefore, the CRR model does not admit arbitrage opportunities.

□

In the CRR model, viability implies completeness. We demonstrate this result in the following proposition.

Proposition 2.12. *If $d < r < u$ and $\tilde{S}$ is a $\mathbb{P}$-martingale, then $\mathbb{P} = \mathbb{P}^*$, $\mathbb{P}^*$ denoting the risk-neutral probability.*

Proof. By hypothesis, $\tilde{S}$ is a $\mathbb{P}$-martingale. This implies by definition that $E\left[\tilde{S}_t|\mathcal{F}_{t-1}\right] = \tilde{S}_{t-1}\ \forall t \in \mathbb{T} \Rightarrow E[T_t|\mathcal{F}_{t-1}] = 1+r\ \forall t \in \mathbb{T}$, as we have already shown in the proof of Proposition 2.11.

On the one hand,

$$E\left[T_t|\mathcal{F}_{t-1}\right] = (1+d) \cdot \mathbb{P}\left(T_t = 1+d|\mathcal{F}_{t-1}\right) + (1+u) \cdot \mathbb{P}\left(T_t = 1+u|\mathcal{F}_{t-1}\right)$$

by definition of $E\left[\ \cdot\ |\mathcal{F}_{t-1}\right]$. We already have ascertained that $E\left[T_t|\mathcal{F}_{t-1}\right] = 1+r$.

On the other hand, $\mathbb{P}\left(T_t = 1+d|\mathcal{F}_{t-1}\right) + \mathbb{P}\left(T_t = 1+u|\mathcal{F}_{t-1}\right) = 1$ because $\mathbb{P}(\ \cdot\ |\mathcal{F}_{t-1})$ is a probability measure over $\{1+d, 1+u\}$. Therefore, we can consider the following system of equations:

$$\begin{cases} (1+d) \cdot \mathbb{P}\left(T_t = 1+d|\mathcal{F}_{t-1}\right) + (1+u) \cdot \mathbb{P}\left(T_t = 1+u|\mathcal{F}_{t-1}\right) = 1+r \\ \mathbb{P}\left(T_t = 1+d|\mathcal{F}_{t-1}\right) + \mathbb{P}\left(T_t = 1+u|\mathcal{F}_{t-1}\right) = 1 \end{cases}$$

The solution of that system of equations is $\mathbb{P}^*\left(T_t = 1+u|\mathcal{F}_{t-1}\right) =: p^* = \frac{r-d}{u-d}$ and $\mathbb{P}^*\left(T_t = 1+d|\mathcal{F}_{t-1}\right) =: 1-p^* = \frac{u-r}{u-d}$ independently of $t \in \mathbb{T}$, i.e. $\forall t \in \mathbb{T}$. In order to prove the desired result of this Proposition, the only remaining fact that needs to be demonstrated is the independence of T_i and $T_j, j \neq i, \forall i, j \in \mathbb{T}$.

Let us remind the definition of conditional expectation: *Given a set $A \in \mathcal{F}$ such that* $\mathbb{P}(A) > 0$ and a random variable X, $E\left[X|A\right] := \frac{E[X \cdot \mathbb{1}_A]}{\mathbb{P}(A)}$. From this definition stems the following result: because $E\left[X|A\right] \in \mathbb{R}$,

$$E\left[X|A\right] \cdot \mathbb{P}(A) = E\left[E\left[X|A\right] \cdot \mathbb{1}_A\right] = E\left[X \cdot \mathbb{1}_A\right],$$

which we use fixing $X = \mathbb{1}_{\{T_t=1+u\}}$ and $X = \mathbb{1}_{\{T_t=1+d\}}$.

$$\begin{cases} E\left[E\left[\mathbb{1}_{\{T_t=1+d\}}|A\right]\cdot \mathbb{1}_A\right] = E\left[\mathbb{1}_{\{T_t=1+d\}}\cdot \mathbb{1}_A\right], \forall A \in \mathcal{F} \\ E\left[E\left[\mathbb{1}_{\{T_t=1+u\}}|A\right]\cdot \mathbb{1}_A\right] = E\left[\mathbb{1}_{\{T_t=1+u\}}\cdot \mathbb{1}_A\right], \forall A \in \mathcal{F} \end{cases}$$

Particularly, this result holds for $\mathcal{F} = \mathcal{F}_{t-1}$:

$$\begin{cases} E\left[E\left[\mathbb{1}_{\{T_t=1+d\}}|\mathcal{F}_{t-1}\right]\cdot \mathbb{1}_A\right] = E\left[\mathbb{1}_{\{T_t=1+d\}}\cdot \mathbb{1}_A\right], \forall A \in \mathcal{F}_{t-1} \\ E\left[E\left[\mathbb{1}_{\{T_t=1+u\}}|\mathcal{F}_{t-1}\right]\cdot \mathbb{1}_A\right] = E\left[\mathbb{1}_{\{T_t=1+u\}}\cdot \mathbb{1}_A\right], \forall A \in \mathcal{F}_{t-1} \end{cases}$$

Moreover, we have already proved that

$$E\left[\mathbb{1}_{\{T_t=1+u\}}|\mathcal{F}_{t-1}\right] = \mathbb{P}^*\left(T_t = 1+u|\mathcal{F}_{t-1}\right) =: p^* \text{ and}$$

$$E\left[\mathbb{1}_{\{T_t=1+d\}}|\mathcal{F}_{t-1}\right] = \mathbb{P}^*\left(T_t = 1+d|\mathcal{F}_{t-1}\right) =: 1-p^*,$$

so:

$$\begin{cases} E\left[(1-p^*)\cdot \mathbb{1}_A\right] = (1-p^*)\cdot E\left[\mathbb{1}_A\right] = E\left[\mathbb{1}_{\{T_t=1+d\}}\cdot \mathbb{1}_A\right], \forall A \in \mathcal{F}_{t-1} \\ E\left[p^*\cdot \mathbb{1}_A\right] = p^*\cdot E\left[\mathbb{1}_A\right] = E\left[\mathbb{1}_{\{T_t=1+u\}}\cdot \mathbb{1}_A\right], \forall A \in \mathcal{F}_{t-1} \end{cases}$$

$$\Rightarrow \begin{cases} (1-p^*)\cdot \mathbb{P}(A) = \mathbb{P}(\{T_t = 1+d\}\cap A), \forall A \in \mathcal{F}_{t-1} \\ p^*\cdot \mathbb{P}(A) = \mathbb{P}(\{T_t = 1+u\}\cap A), \forall A \in \mathcal{F}_{t-1} \end{cases}$$

We now proceed to prove by induction the independence of T_i and $T_j, j \neq i$, $\forall i, j \in \mathbb{T}$:

1. Case $N = 2$

On one hand, if $\{T_2 = 1+u\}$ and $A \in \{\{T_1 = 1+u\}, \{T_1 = 1+u\}\}$, then:

$$\mathbb{P}(\{T_2 = 1+u\}\cap A) = p^*\cdot \mathbb{P}(A) = \begin{cases} (p^*)^2, \text{ if A=\{T}_1 = 1+u\} \\ p^*\cdot(1-p^*), \text{ if } A = \{T_1 = 1+d\} \end{cases}$$

Hence, T_1 and T_2 are independent.

On the other hand, if $\{T_2 = 1+d\}$ and $A \in \{\{T_1 = 1+u\}, \{T_1 = 1+u\}\}$, then:

$$\mathbb{P}(\{T_2 = 1+d\}\cap A) = (1-p^*)\cdot \mathbb{P}(A) = \begin{cases} (1-p^*)^2, \text{ if A=\{T}_1 = 1+d\} \\ p^*\cdot(1-p^*), \text{ if A=\{T}_1 = 1+u\} \end{cases}$$

Hence, T_1 and T_2 are independent.

2. We assume that T_i and T_j are independent , $j \neq i, \forall i, j \leqslant N-1$. We will show that, under this hypothesis, T_i and T_j are independent , $j \neq i, \forall i, j \leqslant N$.

Observe that if T_i and T_j are independent $, j \neq i, \forall i, j \leqslant N-1$ and

$$A = \bigcap_{k=1}^{N-1} \{T_k\} \in \mathcal{F}_{N-1}$$

then

$$\mathbb{P}(A) = \prod_{k=1}^{N-1} \mathbb{P}(\{T_k\}|\mathcal{F}_{k-1}),$$

where

$$\mathbb{P}(\{T_k\}|\mathcal{F}_{k-1}) = \begin{cases} p^*, \text{ if } \{T_k\} = \{T_k = 1+u\} \\ 1-p^*, \text{ if } \{T_k\} = \{T_k = 1+d\} \end{cases}$$

On one hand, if $\{T_N = 1+d\}$ and $A \in \mathcal{F}_{N-1}$, then:

$$\mathbb{P}(\{T_N = 1+d\} \cap A) = (1-p^*) \cdot \mathbb{P}(A) = (1-p^*) \cdot \prod_{k=1}^{N-1} \mathbb{P}(\{T_k\}|\mathcal{F}_{k-1})$$

Hence, $T_1, T_2, ..., T_N$ are independent.

On the other hand, if $\{T_N = 1+d\}$ and $A \in \mathcal{F}_{N-1}$, then:

$$\mathbb{P}(\{T_N = 1+u\} \cap A) = (p^*) \cdot \mathbb{P}(A) = (p^*) \cdot \prod_{k=1}^{N-1} \mathbb{P}(\{T_k\}|\mathcal{F}_{k-1})$$

Hence, $T_1, T_2, ..., T_N$ are independent.

$\Rightarrow \mathbb{P} = \mathbb{P}^*$, $\mathbb{P}^*$ being the risk-neutral probability. □

2.3 Valuation of European options according to the model

We have so far proved that it makes sense to consider financial derivatives in the CRR along with its hedging strategy as long as $d < r < u$. We now turn to give a closed formula for the price of a European call and a European put in the CRR model.

At any time $t \in \mathbb{T}$, the value of a European call with expiry date N and strike price K, which we will denote $C(t, S_t)$, is its payoff profile $G = (S_N - K)^+$ at the value of money at date t, i.e.:

$$C(t, S_t) = \frac{S_t^0}{S_N^0} \cdot E^* \left[(S_N - K)^+ | \mathcal{F}_t \right].$$

Hence:

$$C(t, S_t) = \frac{S_0^0 \cdot (1+r)^t}{S_0^0 \cdot (1+r)^N} \cdot E^* \left[(S_N - K)^+ | \mathcal{F}_t \right] = \frac{1}{(1+r)^{N-t}}$$

$$\cdot E^* \left[\left(S_t \cdot \prod_{i=t+1}^{N} T_i - K \right)^+ |\mathcal{F}_t \right] = \frac{1}{(1+r)^{N-t}} \cdot \sum_{j=0}^{N-t} \left(S_t \cdot (1+d)^{N-t-j} \cdot (1+u)^j - K \right)^+$$

$$\cdot \binom{N-t}{j} \cdot \left(\frac{u-r}{u-d} \right)^{N-t-j} \cdot \left(\frac{r-d}{u-d} \right)^j$$

Define $j^*(S_t)$ as the minimum of up moves that makes the payoff profile be positive at date N having fixed S_t. Then:

$$C(t, S_t) = \frac{1}{(1+r)^{N-t-j}} \cdot \frac{1}{(1+r)^j} \cdot \sum_{j=j^*(S_t)}^{N-t} S_t \cdot (1+d)^{N-t-j} \cdot (1+u)^j \cdot \binom{N-t}{j}$$

$$\cdot \left(\frac{u-r}{u-d} \right)^{N-t-j} \cdot \left(\frac{r-d}{u-d} \right)^j - \frac{1}{(1+r)^{N-t}} \cdot \sum_{j=j^*(S_t)}^{N-t} K \cdot \binom{N-t}{j} \cdot \left(\frac{u-r}{u-d} \right)^{N-t-j}$$

$$\cdot \left(\frac{r-d}{u-d} \right)^j = S_t \cdot \sum_{j=j^*(S_t)}^{N-t} \frac{(1+d)^{N-t-j}}{(1+r)^{N-t-j}} \cdot \frac{(1+u)^j}{(1+r)^j} \cdot \binom{N-t}{j} \cdot \left(\frac{u-r}{u-d} \right)^{N-t-j}$$

$$\cdot \left(\frac{r-d}{u-d} \right)^j - \frac{K}{(1+r)^{N-t}} \cdot \sum_{j=j^*(S_t)}^{N-t} \binom{N-t}{j} \cdot \left(\frac{u-r}{u-d} \right)^{N-t-j} \cdot \left(\frac{r-d}{u-d} \right)^j$$

Observe that if

$$\bar{p} = \frac{(1+u) \cdot (r-d)}{(1+r) \cdot (u-d)},$$

then:

$$C(t, S_t) = S_t \cdot \mathbb{P}\left(Bin\left(N-t, \bar{p} \right) \geqslant j^*\left(S_t \right) \right) - \frac{K \cdot \mathbb{P}\left(Bin\left(N-t, p^* \right) \geqslant j^*\left(S_t \right) \right)}{(1+r)^{N-t}}$$

Once this closed formula has been proved for the price at every $t \in \mathbb{T}$ of a European Call, it is easy to derive a similar closed formula for the price of a European Put at every $t \in \mathbb{T}$ by making use of the Call-Put Parity:

$$\begin{cases} \text{Call-Put Parity: } C(t, S_t) - P(t, S_t) = S_t - \frac{K}{(1+r)^{N-t}} \\ C(t, S_t) = S_t \cdot \mathbb{P}\left(Bin\left(N-t, \bar{p} \right) \geqslant j^*\left(S_t \right) \right) - \frac{K}{(1+r)^{N-t}} \cdot \mathbb{P}\left(Bin\left(N-t, p^* \right) \geqslant j^*\left(S_t \right) \right) \end{cases}$$

$$\Rightarrow P(t, S_t) = \frac{K \cdot \mathbb{P}\left(Bin\left(N-t, p^* \right) \leqslant j^*\left(S_t \right) - 1 \right)}{(1+r)^{N-t}}$$

$$- S_t \cdot \mathbb{P}\left(Bin\left(N-t, \bar{p} \right) \leqslant j^*\left(S_t \right) - 1 \right)$$

2.4 Building hedging portfolios for the selling of European options according to the model

Finally, we give a recursive function that can be implemented as a numerical method in a program with the objective of finding the hedge strategy at each step $t \in \mathbb{T}$.

Definition 2.13. *An investment strategy is a series of random vectors* $\phi_t := \{(\phi_t^0, \phi_t^1) \in \mathbb{R}^2, t \in \mathbb{T}\}$ *such that* ϕ_t^i *is* $\mathcal{F}_{t-1}$*-measurable for every* $i = 1, 2$.

ϕ_t^0 *denotes the quantity of riskless asset in our portfolio at step* t*, whereas* ϕ_t^1 *denotes the quantity of risky asset in our portfolio at step* t.

Forcing the hedging portfolio to have the same value as the European call at each step $t \in \mathbb{T}$:

$$\phi_t^0 \cdot (1+r)^t + \phi_t^1 \cdot S_t = C(t, S_t),$$

we can find ϕ_t^0 and ϕ_t^1 by solving the following system:

$$\begin{cases} \phi_t^0 \cdot (1+r)^t + \phi_t^1 \cdot (1+d) \cdot S_{t-1} = C(t, (1+d) \cdot S_{t-1}) \\ \phi_t^0 \cdot (1+r)^t + \phi_t^1 \cdot (1+u) \cdot S_{t-1} = C(t, (1+u) \cdot S_{t-1}) \end{cases}$$

$$\Rightarrow \phi_t^1 = \frac{C(t, (1+u) \cdot S_{t-1}) - C(t, (1+d) \cdot S_{t-1})}{(u-d) \cdot S_{t-1}}$$

Due to the Autofinancing property of ϕ_t:

$$V_{t-1}(\phi) = \phi_t^0 \cdot (1+r)^{t-1} + \phi_t^1 \cdot S_{t-1} \Rightarrow \phi_t^0 = \frac{V_{t-1}(\phi) - \phi_t^1 \cdot S_{t-1}}{(1+r)^{t-1}}$$

A program written in C++ that prices and builds the hedging strategy of different European derivatives can be found in 8.1.1.

3 From the Cox-Ross-Rubinstein model to the Black-Scholes formula

3.1 Adapting a time-continuous price process to the Cox-Ross-Rubinstein model

Let $S = \{S_t, t \in [0,T]\}$ be the price process in continuous time of the risky asset. Let $r \geqslant 0$ be the fixed interest rate. Fix $N \in \mathbb{N}$ as well.

We discretise the continuous time interval $[0,T]$ in $N+1$ time-points $t_i := \frac{i \cdot T}{N}$, for $i = 0, 1, ..., N$. Accordingly, the continuous price process S is also discretised defining $S_i = S(t_i)$ for $i = 0, 1, ..., N$.

We assume that at every time step of length $\frac{T}{N}$, only two changes are possible for a certain price: either increase or decrease in the same relative amount regulated by a fixed parameter σ. I.e:

$$1 + d_N = \frac{1 + r \cdot \frac{T}{N}}{e^{\frac{\sigma}{\sqrt{N}}}},$$

$$1 + u_N = \left(1 + r \cdot \frac{T}{N}\right) \cdot e^{\frac{\sigma}{\sqrt{N}}}$$

Therefore, we have got $S = \{S_i = S(t_i), 0 \leqslant i \leqslant N\}$ modelled with a CRR satisfying:

$$d_N = \frac{1 + r\frac{T}{N}}{e^{\frac{\sigma}{\sqrt{N}}}} - 1 \leqslant r\frac{T}{N} =: r_N < \left(1 + r\frac{T}{N}\right) \cdot e^{\frac{\sigma}{\sqrt{N}}} - 1 =: u_N$$

Consider de independent and identically distributed random variables T_0^N, T_1^N, ..., T_N^N with Bernouilli law on

$$\left\{ \frac{1 + r \cdot \frac{T}{N}}{e^{\frac{\sigma}{\sqrt{N}}}}, \left(1 + r \cdot \frac{T}{N}\right) \cdot e^{\frac{\sigma}{\sqrt{N}}} \right\}$$

and with parameter

$$p_N := \mathbb{P}\left(\left\{T_i^N = 1 + u_N\right\}\right) = \frac{r_N - d_N}{u_N - d_N} = \frac{1 - \frac{1}{e^{\frac{\sigma}{\sqrt{N}}}}}{e^{\frac{\sigma}{\sqrt{N}}} - \frac{1}{e^{\frac{\sigma}{\sqrt{N}}}}}, \forall i \in \{0, 1, ..., N\}$$

$$T_i^N \in \left\{ \frac{1 + r \cdot \frac{T}{N}}{e^{\frac{\sigma}{\sqrt{N}}}}, \left(1 + r \cdot \frac{T}{N}\right) \cdot e^{\frac{\sigma}{\sqrt{N}}} \right\}, \forall i \in \{0, 1, ..., N\} \Rightarrow \frac{T_i^N}{1 + r\frac{T}{N}} \in \left\{ \frac{1}{e^{\frac{\sigma}{\sqrt{N}}}}, e^{\frac{\sigma}{\sqrt{N}}} \right\}$$

$$\Rightarrow \log\left(\frac{T_i^N}{1 + r\frac{T}{N}}\right) \in \left\{ -\frac{\sigma}{\sqrt{N}}, \frac{\sigma}{\sqrt{N}} \right\} \Rightarrow X_i^N := \log\left(\frac{T_i^N}{1 + r\frac{T}{N}}\right) \in \left\{ -\frac{\sigma}{\sqrt{N}}, \frac{\sigma}{\sqrt{N}} \right\}$$

Hence, $X_0^N, X_1^N, ..., X_N^N$ is a series of independent and identically distributed random variables with Bernouilli law on $\left\{-\frac{\sigma}{\sqrt{N}}, \frac{\sigma}{\sqrt{N}}\right\}$ with parameter

$$p_N := \frac{1 - \frac{1}{e^{\frac{\sigma}{\sqrt{N}}}}}{e^{\frac{\sigma}{\sqrt{N}}} - \frac{1}{e^{\frac{\sigma}{\sqrt{N}}}}}$$

Observe that:

$$S_n = S_0 \cdot \prod_{i=0}^{n} T_i^N \Rightarrow \frac{S_n}{\left(1 + r\frac{T}{N}\right)^n} = S_0 \cdot \frac{\prod_{i=0}^{n} T_i^N}{\left(1 + r\frac{T}{N}\right)^n} = S_0 \cdot \prod_{i=0}^{n} \frac{T_i^N}{\left(1 + r\frac{T}{N}\right)}$$

$$\Rightarrow \log\left(\frac{S_n}{\left(1 + r\frac{T}{N}\right)^n}\right) = \log\left(S_0 \cdot \prod_{i=0}^{n} \frac{T_i^N}{\left(1 + r\frac{T}{N}\right)}\right) \Rightarrow \log(S_n) - \log\left(1 + r\frac{T}{N}\right)^n$$

$$= \log(S_0) + \sum_{i=0}^{n} \log\left(\frac{T_i^N}{\left(1 + r\frac{T}{N}\right)}\right) \Rightarrow \log(S_n) = \log(S_0) + \sum_{i=0}^{n} X_i^N + n \cdot \log\left(1 + r\frac{T}{N}\right),$$

$\forall n \in \{0, 1, ..., N\}$. Particularly, it holds for $n = N$, therefore:

$$\log(S_N) = \log(S_0) + \sum_{i=0}^{N} X_i^N + N \cdot \log\left(1 + r\frac{T}{N}\right),$$

Our intention is to discover the law of $\log(S_N)$ when $N \to \infty$. Because $\sum_{i=0}^{N} X_i^N$ is the only random variable in the above equation, we begin ascertaining the behaviour of its expectation and variance when $N \to \infty$.

3.2 Law of $\log S_N$ when $N \to \infty$

Lemma 3.1. *$E(X_j^N) = \frac{\sigma}{\sqrt{N}} \cdot (2p_N - 1)$, and $Var\left[X_j^N\right] = \frac{\sigma^2}{N} \cdot 4 \cdot p_N \cdot (1 - p_N)$, $\forall j \in \{0, 1, ..., N\}$*

Proof. For every $j \in \{0, 1, ..., N\}$,

$$E(X_j^N) = \frac{\sigma}{\sqrt{N}} \cdot p_N + \frac{-\sigma}{\sqrt{N}} \cdot (1 - p_N) = \frac{\sigma}{\sqrt{N}} \cdot (2p_N - 1)$$

For every $j \in \{0, 1, ..., N\}$,

$$Var\left[X_j^N\right] = E\left[\left(X_j^N - E(X_j^N)\right)^2\right] = E\left[\left(X_j^N\right)^2 - 2 \cdot X_j^N \cdot E(X_j^N) + \left(E(X_j^N)\right)^2\right]$$

$$= E\left[\left(X_j^N\right)^2 - 2 \cdot X_j^N \cdot \frac{\sigma}{\sqrt{N}} \cdot (2p_N - 1) + \left(\frac{\sigma}{\sqrt{N}} \cdot (2p_N - 1)\right)^2\right]$$

$$= E\left[\left(X_j^N\right)^2\right] - 2 \cdot \frac{\sigma}{\sqrt{N}} \cdot (2p_N - 1) \cdot E\left[X_j^N\right] + \left(\frac{\sigma}{\sqrt{N}} \cdot (2p_N - 1)\right)^2$$

$$= \left(\frac{\sigma}{\sqrt{N}}\right)^2 \cdot p_N + \left(\frac{-\sigma}{\sqrt{N}}\right)^2 \cdot (1 - p_N) - 2 \cdot \left(\frac{\sigma}{\sqrt{N}} \cdot (2p_N - 1)\right)^2 + \left(\frac{\sigma}{\sqrt{N}} \cdot (2p_N - 1)\right)^2$$

$$= \frac{\sigma^2}{N} \cdot p_N - \frac{\sigma^2}{N} \cdot p_N + \frac{\sigma^2}{N} - \frac{\sigma^2}{N} \cdot (2p_N - 1)^2 = \frac{\sigma^2}{N} \cdot \left[1 - (2p_N - 1)^2\right]$$

$$= \frac{\sigma^2}{N} \cdot \left[1 - (4{p_N}^2 - 4p_N + 1)\right] = \frac{\sigma^2}{N} \cdot 4 \cdot p_N \cdot (1 - p_N)$$

Therefore:

$$E(X_j^N) = \frac{\sigma}{\sqrt{N}} \cdot (2p_N - 1), \text{ and } Var\left[X_j^N\right] = \frac{\sigma^2}{N} \cdot 4 \cdot p_N \cdot (1 - p_N)$$

$\square$

Lemma 3.2.

$$E\left[\sum_{i=0}^{N} X_i^N\right] = \sigma \cdot \sqrt{N} \cdot (2p_N - 1)$$

and

$$Var\left(\sum_{j=0}^{N} X_j^N\right) = \sigma^2 \cdot 4 \cdot p_N \cdot (1 - p_N)$$

Proof. Because of the linearity of the expectation,

$$E\left[\sum_{i=0}^{N} X_i^N\right] = \sum_{i=0}^{N} E(X_j^N) = N \cdot E(X_j^N)$$

Similarly,

$$Var\left[\sum_{i=0}^{N} X_i^N\right] = \sum_{i=0}^{N} Var(X_j^N) = N \cdot Var(X_j^N),$$

because $X_0^N, X_1^N, ..., X_N^N$, are independent. $\square$

Lemma 3.3. $p_N \xrightarrow[N\to\infty]{} \frac{1}{2}$

Proof. By definition,

$$p_N = \frac{1 - \frac{1}{e^{\frac{\sigma}{\sqrt{N}}}}}{e^{\frac{\sigma}{\sqrt{N}}} - \frac{1}{e^{\frac{\sigma}{\sqrt{N}}}}}$$

Due to the fact that both $e^{\frac{\sigma}{\sqrt{N}}}$, $e^{\frac{-\sigma}{\sqrt{N}}} \xrightarrow[N\to\infty]{} e^0$, we are interested in studying the Taylor expansion of $p_N(x) := \frac{1-e^{-x}}{e^x - e^{-x}}$ around $x = 0$.

The Taylor expansions of e^x, e^{-x}, $1-e^{-x}$ and $e^x - e^{-x}$ around $x=0$ are, respectively,

$$\begin{aligned}
&1 + x + \frac{x^2}{2} + O(x^3),\\
&1 - x + \frac{x^2}{2} + O(x^3),\\
&x - \frac{x^2}{2} + O(x^3) \text{ and}\\
&2x + O(x^3)
\end{aligned}$$

Therefore:

$$p_N(x) := \frac{1-e^{-x}}{e^x - e^{-x}} = \frac{x - \frac{x^2}{2} + O(x^3)}{2x + O(x^3)}$$

$$p_N = p_N\left(\frac{\sigma}{\sqrt{N}}\right) = \frac{\frac{\sigma}{\sqrt{N}} - \frac{\left(\frac{\sigma}{\sqrt{N}}\right)^2}{2} + O\left(\left(\frac{\sigma}{\sqrt{N}}\right)^3\right)}{2\frac{\sigma}{\sqrt{N}} + O\left(\left(\frac{\sigma}{\sqrt{N}}\right)^3\right)} = \frac{\frac{\sigma}{\sqrt{N}} - \frac{1}{2}\frac{\sigma^2}{N} + O\left(\frac{1}{N\sqrt{N}}\right)}{2\frac{\sigma}{\sqrt{N}} + O\left(\frac{1}{N\sqrt{N}}\right)}$$

$$= \frac{\frac{1}{\sqrt{N}}}{\frac{1}{\sqrt{N}}} \cdot \frac{\sigma - \frac{1}{2}\frac{\sigma^2}{\sqrt{N}} + O\left(\frac{1}{N}\right)}{2\sigma + O\left(\frac{1}{N}\right)} = \frac{\sigma - \frac{1}{2}\frac{\sigma^2}{\sqrt{N}} + O\left(\frac{1}{N}\right)}{2\sigma + O\left(\frac{1}{N}\right)} \xrightarrow[N\to\infty]{} \frac{1}{2}$$

□

Proposition 3.4.

$$\lim_{N\to\infty} Var\left(\sum_{j=0}^{N} X_j^N\right) = \sigma^2$$

Proof. As we have already proved in Lemma 3.1, $Var\left(\sum_{j=0}^{N} X_j^N\right) = \sigma^2 \cdot 4 \cdot p_N \cdot (1-p_N)$. Applying Lemma 3.3:

$$Var\left(\sum_{j=0}^{N} X_j^N\right) = \sigma^2 \cdot 4 \cdot p_N \cdot (1-p_N) \xrightarrow[N\to\infty]{} \sigma^2 \cdot 4 \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right) = \sigma^2$$

□

Proposition 3.5.

$$\lim_{N\to\infty} E\left(\sum_{j=0}^{N} X_j^N\right) = \frac{-\sigma^2}{2}$$

Proof. We have already proved in Lemma 3.2 that

$$E\left[\sum_{i=0}^{N} X_i^N\right] = \sigma \cdot \sqrt{N} \cdot (2p_N - 1)$$

By definition of p_N,

$$E\left[\sum_{i=0}^{N} X_i^N\right] = \sigma \cdot \sqrt{N} \cdot (2p_N - 1) = \sigma \cdot \sqrt{N} \cdot \left(2 \cdot p_N \left(\frac{\sigma}{\sqrt{N}}\right) - 1\right)$$

$$= \sigma \cdot \sqrt{N} \cdot \left(2 \cdot p_N \left(\frac{\sigma}{\sqrt{N}}\right) - 1\right) = \sigma \cdot \sqrt{N} \cdot \left(2 \cdot \frac{\sigma - \frac{1}{2}\frac{\sigma^2}{\sqrt{N}} + O\left(\frac{1}{N}\right)}{2\sigma + O\left(\frac{1}{N}\right)} - 1\right)$$

$$= \sigma \cdot \sqrt{N} \cdot \left(\frac{2\sigma - \frac{\sigma^2}{\sqrt{N}} - 2\sigma + O\left(\frac{1}{N}\right)}{2\sigma + O\left(\frac{1}{N}\right)}\right) = \sigma \cdot \sqrt{N} \cdot \frac{\frac{-\sigma^2}{\sqrt{N}} + O\left(\frac{1}{N}\right)}{2\sigma + O\left(\frac{1}{N}\right)}$$

$$= \frac{\sqrt{N} \cdot \left(\frac{-\sigma^2}{\sqrt{N}} + O\left(\frac{1}{N}\right)\right)}{\frac{2\sigma + O\left(\frac{1}{N}\right)}{\sigma}} = \frac{-\sigma^2 + O\left(\frac{1}{N\sqrt{N}}\right)}{2 + O\left(\frac{1}{N}\right)} \xrightarrow[N\to\infty]{} \frac{-\sigma^2}{2}$$

$\square$

Corollary 3.6.

$$\log(S_N) \xrightarrow{\mathcal{L}} N\left(\log S_0 + rT - \frac{\sigma^2}{2}, \sigma^2\right)$$

Proof. Because Propositions 3.4 and 3.5 we can state that

$$\lim_{N\to\infty} E\left(\sum_{j=0}^{N} X_j^N\right) = \frac{-\sigma^2}{2}$$

and

$$\lim_{N\to\infty} Var\left(\sum_{j=0}^{N} X_j^N\right) = \sigma^2$$

The Central Limit Theorem ensures that

$$\left(\sum_{j=0}^{N} X_j^N\right) \xrightarrow{\mathcal{L}} N\left(\frac{-\sigma^2}{2}, \sigma^2\right)$$

Therefore,

$$\log(S_N) = \log(S_0) + \sum_{i=0}^{N} X_i + N \cdot \log\left(1 + r\frac{T}{N}\right) \xrightarrow{\mathcal{L}} N\left(\log S_0 + rT - \frac{\sigma^2}{2}, \sigma^2\right)$$

$\square$

3.3 The Black-Scholes formula

Consider the price $P_0^{(N)}$ of a put option with expiry date T and strike price K:

$$P_0^{(N)} = E\left[\left(\frac{K}{\left(1+r\frac{T}{N}\right)^N} - S_0 e^{\sum_{j=0}^N X_j^N}\right)^+\right]$$

Let us define $Y_N := \sum_{j=0}^N X_j^N$ for a better reading. Recall that in the previous section we proved that $Y_N \to Y \sim N\left(\frac{-\sigma^2}{2}, \sigma^2\right)$

Definition 3.7.

$$\begin{aligned}\varphi\colon \mathbb{R} &\to \mathbb{R}\\ y \mapsto \varphi(y) &:= \left(\frac{K}{e^{rT}} - S_0 e^y\right)^+\end{aligned}$$

Observation: φ is continuous and $\varphi(y) \leqslant \frac{K}{e^{rT}}\ \forall y \in \mathbb{R}$

Lemma 3.8.

$$\lim_{N\to\infty} P_0^{(N)} = \lim_{N\to\infty} E\left[\varphi\left(Y_N\right)\right]$$

Proof.

$$P_0^{(N)} = E\left[\varphi\left(Y_N\right)\right] + P_0^{(N)} - E\left[\varphi\left(Y_N\right)\right]$$

Notice that:

$$\left|P_0^{(N)} - E\left[\varphi\left(Y_N\right)\right]\right| = \left|E\left[\left(\frac{K}{\left(1+r\frac{T}{N}\right)^N} - S_0 e^{Y_N}\right)^+\right] - E\left[\varphi\left(Y_N\right)\right]\right|$$

$$= \left|E\left[\left(\frac{K}{\left(1+r\frac{T}{N}\right)^N} - S_0 e^{Y_N}\right)^+\right] - E\left[\left(\frac{K}{e^{rT}} - S_0 e^{Y_N}\right)^+\right]\right|$$

$$= \left|E\left[\left(\frac{K}{\left(1+r\frac{T}{N}\right)^N} - S_0 e^{Y_N}\right)^+ - \left(\frac{K}{e^{rT}} - S_0 e^{Y_N}\right)^+\right]\right| \leqslant \left|E\left[\frac{K}{1+r\frac{T}{N}^N} - \frac{K}{e^{rT}}\right]\right|$$

$$= K\cdot\left|E\left[\frac{1}{1+r\frac{T}{N}^N} - \frac{1}{e^{rT}}\right]\right| \xrightarrow[N\to\infty]{} K\cdot\left|E\left[\frac{1}{e^{rT}} - \frac{1}{e^{rT}}\right]\right| = 0$$

$$\Rightarrow \left|P_0^{(N)} - E\left[\varphi\left(Y_N\right)\right]\right| \xrightarrow[N\to\infty]{} 0 \Rightarrow \lim_{N\to\infty} P_0^{(N)} = \lim_{N\to\infty} E\left[\varphi\left(Y_N\right)\right]$$

$\square$

Proposition 3.9.

$$E\left[\varphi\left(Y_N\right)\right] \xrightarrow[N\to\infty]{} E\left[\varphi\left(Y\right)\right]$$

Proof. Because $\lim_{N\to\infty} P_0^{(N)} = \lim_{N\to\infty} E\left[\varphi\left(Y_N\right)\right]$, φ is bounded, φ is continuous and $Y_N \to Y \sim N\left(\frac{-\sigma^2}{2}, \sigma^2\right)$, under the hypotheses of Portmanteau Theorem we have that

$$E\left[\varphi\left(Y_N\right)\right] \xrightarrow[N\to\infty]{} E\left[\varphi\left(Y\right)\right]$$

A proof of Portmanteau Theorem can be found in [2]. □

Corollary 3.10. *The Black-Scholes formula for a European put with expiry date T and strike price K is:*

$$P_0 := \lim_{N\to\infty} P_0^{(N)} = \frac{K}{e^{rT}} \cdot \Phi\left(d_+\right) - S_0 \cdot \Phi\left(d_-\right)$$

where Φ denotes the cumulative distribution function of a standard normal distribution, and d_+ and d_- are:

$$d_+ := \frac{1}{\sigma} \cdot \log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}$$

$$d_- := d_+ - \sigma$$

Proof. Thanks to Proposition 3.9, we know that

$$P_0 := \lim_{N\to\infty} P_0^{(N)} = \lim_{N\to\infty} E\left[\varphi\left(Y_N\right)\right] = E\left[\varphi\left(Y\right)\right]$$

We focus on the computation of $E\left[\varphi\left(Y\right)\right]$:

$$E\left[\varphi\left(Y\right)\right] = \int_{-\infty}^{\infty} \varphi(y) \cdot f_Y(y) dy = \int_{-\infty}^{\infty} \left(\frac{K}{e^{rT}} - S_0 e^y\right)^+ \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-\left(y - \left(\frac{-\sigma^2}{2}\right)\right)^2}{2\sigma^2}} dy$$

$$= \int_{-\infty}^{\infty} \left(\frac{K}{e^{rT}} - S_0 e^y\right)^+ \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-\left(y + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}} dy$$

Standardising variable Y, we come up with the change of variables $u = \frac{y + \frac{\sigma^2}{2}}{\sigma}$ $\left(\Rightarrow y = \sigma u - \frac{\sigma^2}{2}\right)$ and $dy = \sigma du$. Notice that $-\infty < y < \infty \Rightarrow -\infty < u < \infty$. Therefore:

$$E\left[\varphi\left(Y\right)\right] = \int_{-\infty}^{\infty} \left(\frac{K}{e^{rT}} - S_0 e^{\sigma u - \frac{\sigma^2}{2}}\right)^+ \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-\left(\sigma u - \frac{\sigma^2}{2} + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}} \sigma du$$

$$= \int_{-\infty}^{\infty} \left(\frac{K}{e^{rT}} - S_0 e^{\sigma u - \frac{\sigma^2}{2}}\right)^+ \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-(\sigma u)^2}{2\sigma^2}} du$$

$$= \int_{-\infty}^{\infty} \left(\frac{K}{e^{rT}} - S_0 e^{\sigma u - \frac{\sigma^2}{2}}\right)^+ \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-u^2}{2}} du$$

In order to remove the $(\ \cdot\ )^+$ function from the integral, we discuss for which $u \in \mathbb{R}$ the inequality

$$\left(\frac{K}{e^{rT}} - S_0 e^{\sigma u - \frac{\sigma^2}{2}}\right)^+ > 0$$

is satisfied:

$$\left(\frac{K}{e^{rT}} - S_0 e^{\sigma u - \frac{\sigma^2}{2}}\right)^+ > 0 \iff \frac{K}{e^{rT}} > S_0 e^{\sigma u - \frac{\sigma^2}{2}} \iff \frac{K}{S_0 e^{rT}} > e^{\sigma u - \frac{\sigma^2}{2}}$$

$$\iff \log\left(\frac{K}{S_0 e^{rT}}\right) > \log\left(e^{\sigma u - \frac{\sigma^2}{2}}\right) \iff \log\left(\frac{K}{S_0}\right) - \log\left(e^{rT}\right) > \sigma u - \frac{\sigma^2}{2}$$

$$\iff \frac{\log\left(\frac{K}{S_0}\right) - \log\left(e^{rT}\right) + \frac{\sigma^2}{2}}{\sigma} > u \iff \frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2} > u$$

Hence:

$$E\left[\varphi\left(Y\right)\right] = \int_{-\infty}^{\frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}} \left(\frac{K}{e^{rT}} - S_0 e^{\sigma u - \frac{\sigma^2}{2}}\right) \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-u^2}{2}} du$$

$$= \int_{-\infty}^{\frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}} \frac{K}{e^{rT}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-u^2}{2}} du - \int_{-\infty}^{\frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}} S_0 e^{\sigma u - \frac{\sigma^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-u^2}{2}} du$$

$$= \frac{K}{e^{rT}} \cdot \int_{-\infty}^{\frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}} \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-u^2}{2}} du - \int_{-\infty}^{\frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}} S_0 e^{\sigma u - \frac{\sigma^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-u^2}{2}} du$$

Observe that the left hand side integral satisfies

$$\int_{-\infty}^{\frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}} \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-u^2}{2}} du = \int_{-\infty}^{\frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}} \phi(u) du,$$

ϕ being the distribution function of a standardised normal distribution. Therefore,

$$\int_{-\infty}^{\frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}} \phi(u) du = \Phi(d_+),$$

Φ being the density function of a standardised normal distribution and $d_+ := \frac{1}{\sigma} \cdot \log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}$.

On the other hand, in order to solve the right hand integral, we must realize that

$$e^{\sigma u - \frac{\sigma^2}{2}} \cdot e^{\frac{-u^2}{2}} = e^{-\frac{u^2}{2} + u\sigma - \frac{\sigma^2}{2}} = e^{\frac{-(u-\sigma)^2}{2}}$$

Standardising variable u using the change of variables $u = v + \sigma$, $dv = du$ forces us to adapt the upper limit of integration because

$$-\infty < u < \frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2} \Rightarrow -\infty < v < \frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2} - \sigma$$

$$= \frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} - \frac{\sigma}{2} =: d_-$$

Hence:

$$E\left[\varphi\left(Y\right)\right] = \frac{K}{e^{rT}}\Phi(d_+) - \int_{-\infty}^{\frac{1}{\sigma}\log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} - \frac{\sigma}{2}} \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-v^2}{2}} dv = \frac{K}{e^{rT}}\Phi(d_+) - S_0\Phi(d_-)$$

$$\Rightarrow P_0 := \lim_{N\to\infty} P_0^{(N)} = \frac{K}{e^{rT}} \cdot \Phi\left(d_+\right) - S_0 \cdot \Phi\left(d_-\right)$$

$\square$

Corollary 3.11. *The Black-Scholes formula for the price of a European call with expiry date T and strike price K is*

$$C_0 = S_0 \cdot \Phi\left(-d_-\right) - \frac{K}{e^{rT}} \cdot \Phi\left(-d_+\right)$$

Proof. Because the Black-Scholes formula for a European put has to be consistent with Call-Put Parity, we know that:

$$C_0 = P_0 + S_0 - \frac{K}{e^{rT}} = \frac{K}{e^{rT}} \cdot \Phi\left(d_+\right) - S_0 \cdot \Phi\left(d_-\right) + S_0 - \frac{K}{e^{rT}}$$

$$\Rightarrow C_0 = \frac{K}{e^{rT}} \cdot \left(\Phi(d_+) - 1\right) - S_0 \cdot \left(\Phi(d_-) - 1\right) = S_0 \cdot \left(1 - \Phi(d_-)\right) - \frac{K}{e^{rT}} \cdot \left(1 - \Phi(d_+)\right)$$

Finally, thanks to the $1 - \Phi(x) = \Phi(-x)$, $\forall x \in \mathbb{R}$, property of the cumulative distribution function of a standardised normal distribution, we have:

$$C_0 = S_0 \cdot \Phi\left(-d_-\right) - \frac{K}{e^{rT}} \cdot \Phi\left(-d_+\right).$$

$\square$

A program written in C++ that uses the Black-Scholes formula to compute the implied volatility σ of the market can be found in 8.1.2.

4 Stochastic calculus

4.1 The Wiener process and the Brownian motion process

Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geqslant 0}, \mathbb{P})$ be a filtered probability space, i.e.:

- Let $t \in [0, \infty)$.
- Let Ω be the set of all possible continuous paths of the evolution of prices of a particular risky asset.
- $\mathbb{F} = \{\mathcal{F}_t, t \in [0, \infty], \mathcal{F}_t\ \sigma\text{-algebra of } \Omega \text{ and } \mathcal{F} = \mathcal{F}_\infty\}$.
- Let $\mathbb{P}$ be a probability measure over Ω.

The motivation of considering continuous-time models comes from the fact that, in practice, the price changes in the market are so frequent that a discrete-time model can barely follow the moves. On the other hand, the most widely used model is the continuous-time Black-Scholes model, which leads to more explicit computations than the discrete-time models.

We shall begin defining the Brownian motion process since it is the core concept of the Black-Scholes model and appears in most financial asset models.

Definition 4.1. *A Brownian motion process is a real-valued, continuous stochastic process $(W_t)_{t\geqslant 0}$, with independent and stationary increments. In other words:*

$$\begin{aligned} W\colon [0, \infty) \times \Omega &\to \mathbb{R} \\ (t, \omega) &\mapsto W(t, \omega) \end{aligned}$$

is a Brownian motion process if it satisfies:

- *it is continuous: $\mathbb{P}$ a.s. the map $s \mapsto W(s, \omega)$ is continuous.*
- *it has independent increments: if $s \leqslant t$, $W(t, \omega) - W(s, \omega)$ is independent of $\mathcal{F}_s = \sigma(W(u, \omega), u \leqslant s)$.*
- *it has stationary increments: if $s \leqslant t$, then $W(t, \omega) - W(s, \omega)$ and $W(t - s, \omega) - W(0, \omega)$ have the same probability law.*

The previous definition contains properties of the behaviour of the evolution of prices in the markets that a trader can observe in his screen at work. This definition induces the distribution of the process $W(t, \omega)$, but the result is difficult to prove and the reader ought to consult the book by Gikhman and Skorokhod (1969) for a proof of the following theorem.

Theorem 4.2. *If W is a Brownian motion process, then $W(t, \omega) - W(0, \omega)$ is a normal random variable with mean rt and variance $\sigma^2 t$, where r and σ are constant real numbers.*

This theorem establishes the grounds of another definition of the Brownian motion process, more friendly to handle in the technical details of the upcoming sections.

Definition 4.3.

$$\begin{aligned} W\colon [0,\infty)\times\Omega &\to \mathbb{R} \\ (t,\omega) &\mapsto W(t,\omega) \end{aligned}$$

is a Brownian motion process relative to $\mathbb{F}$ *if and only if for every* $\omega\in\Omega$*:*

- $W(0,\omega)\overset{\text{a.s.}}{=}0$ *and* $W(t,\omega)$ *is* $\mathcal{F}_t$*-measurable.*
- $W(t,\omega)-W(s,\omega)$ *is independent of* $\mathcal{F}_s$*,* $s\leqslant t\ \forall s,t\in[0,\infty)$*.*
- $W(t,\omega)-W(s,\omega)\sim N(0,t-s)$*,* $\forall s,t\in[0,\infty)$*.*

Moreover, the following definition of Wiener process is closely related to Definition 4.3..

Definition 4.4.

$$\begin{aligned} W\colon [0,\infty)\times\Omega &\to \mathbb{R} \\ (t,\omega) &\mapsto W(t,\omega) \end{aligned}$$

is a Wiener process relative to $\mathbb{F}$ *if and only if for every* $\omega\in\Omega$*:*

- $W(0,\omega)\overset{\text{a.s.}}{=}0$ *and* $W(t,\omega)$ *is* $\mathcal{F}_t$*-measurable.*
- $E\left[W(t,\omega)|\mathcal{F}_s\right]=W(s,\omega)\ \forall s<t,\ s,t\in[0,\infty)$*. I.e.,* W *is a martingale with respect to* $\mathbb{P}$*.*
- $E\left[(W(t,\omega))^2\right]<\infty\ \forall t\in[0,\infty)$*.*
- $E\left[(W(t,\omega)-W(s,\omega))^2\right]=t-s,\ \forall s\leqslant t,\ s,t\in[0,\infty)$*.*
- *For every fixed* $\omega\in\Omega$*,* $t\mapsto W(t,\omega)$ *is continuous.*

In fact, the following theorem states that we can use Definition 4.3. and Definition 4.4. interchangeably:

Theorem 4.5. W *is a Wiener process relative to* $(\mathcal{F}_t)_t$*,* $t\in[0,\infty)$ *if and only if* W *is a Brownian motion process relative to* $(\mathcal{F}_t)_t$*,* $t\in[0,\infty)$*.*

Proof. The details of this proof can be found in a series of theorems, propositions and remarks in [12]. □

4.2 Properties of the Wiener process

In order to reach a better understanding of the concept of Wiener process, we will show some of its properties. New concepts have to be introduced so that the Wiener process' properties can be better explained:

Definition 4.6. *For any $[a, b] \subset [0, \infty)$, $(\pi_n)_n$ is a series of partitions of $[a, b]$ such that, for every $n \in \mathbb{N}$, $\pi_n = \{t_0, t_1, ..., t_n\}$ satisfies $a = t_0 < t_1 < ... < t_n = b$ and*

$$\Pi := \max_{j=0,\ \ldots,\ n-1} \{t_{j+1} - t_j\} \xrightarrow[n \to \infty]{} 0.$$

In some parts of this chapter we will require every π_n to be an equi-spaced partition, i.e., if $\pi_n = \{t_0, t_1, ..., t_n\}$ then $t_{j+1} - t_j = \frac{b-a}{n}$ $\forall j \in \{0, 1, ..., n-1\}$. If such assumption is needed, we will explicitly mention that π_n are equi-spaced.

Definition 4.7. *Consider, for any $t \in [0, \infty)$, a series of partitions $(\pi_n)_n$ of the interval $[0, t]$ and*

$$V_{[0,t]}\left[(W(t_k, \omega))_{k \leqslant n}\right] := \sum_{j=0}^{n-1} |W(t_{j+1}, \omega) - W(t_j, \omega)|$$

If the limit $\lim_{\pi \to 0}\left(V_{[0,t]}\left[(W(t_k, \omega))_{k \leqslant n}\right]\right)$ exists, we define it as the $Variation$ of process W along the interval $[0, t]$ and we call it $Variation_t(W)$:

$$Variation_t(W) := \lim_{\Pi \to 0}\left(V_{[0,t]}\left[(W(t_k, \omega))_{k \leqslant n}\right]\right)$$

Notice that $Variation_t(W)$, if it exists, is a stochastic process too. We will show that a Wiener process has infinite variation.

Definition 4.8. *Consider, for any $t \in [0, \infty)$, a series of partitions $(\pi_n)_n$ of the interval $[0, t]$ and*

$$V^2_{[0,t]}\left[(W(t_k, \omega))_{k \leqslant n}\right] := \sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2$$

If the limit $\lim_{\pi \to 0}\left(V^2_{[0,t]}\left[(W(t_k, \omega))_{k \leqslant n}\right]\right)$ exists, we define it as the $Quadratic$ $Variation$ of process W along the interval $[0, t]$ and we call it $[W, W]_t$:

$$[W, W]_t := \lim_{\Pi \to 0}\left(V^2_{[0,t]}\left[(W(t_k, \omega))_{k \leqslant n}\right]\right)$$

Lemma 4.9. *Let $t \in [0, \infty)$ and let π_n be a partition of the interval $[0, t]$ for every $n \in \mathbb{N}$. For every $n \in \mathbb{N}$, if W is a Wiener process, then:*

$$E\left[\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 - t\right] = 0$$

Proof. For every $n \in \mathbb{N}$,

$$\sum_{j=0}^{n-1} (t_{j+1} - t_j)$$

is a telescoping sum. Therefore,

$$\sum_{j=0}^{n-1} (t_{j+1} - t_j) = t_n - t_0,$$

π_n is a partition of the interval $[0, t]$ for every $n \in \mathbb{N}$, then $0 = t_0$ and $t_n = t$ implying

$$\sum_{j=0}^{n-1} (t_{j+1} - t_j) = t_n - t_0 = t$$

$$E\left[\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 - t\right] = E\left[\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 - (t_{j+1} - t_j)\right]$$

and because the expectation is a linear function this is equal to:

$$= \sum_{j=0}^{n-1} E\left[(W(t_{j+1}, \omega) - W(t_j, \omega))^2\right] - E\left(t_{j+1} - t_j\right)$$

W Wiener process $\Rightarrow E\left[(W(t_{j+1}, \omega) - W(t_j, \omega))^2\right] = t_{j+1} - t_j, \forall j \in \{0, 1, ..., n-1\}$

Because $t_{j+1}, t_j \in \mathbb{R}$ and are not random variables, $E\left(t_{j+1} - t_j\right) = t_{j+1} - t_j$. Therefore:

$$E\left[\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 - t\right] = \sum_{j=0}^{n-1} [t_{j+1} - t_j - (t_{j+1} - t_j)] = 0$$

□

Lemma 4.10. *Let $t \in [0, \infty)$ and let π_n be a partition of the interval $[0, t]$ for every $n \in \mathbb{N}$. For every $n \in \mathbb{N}$, if W is a Wiener process, then:*

$$Var\left[\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 - t\right] = E\left[\left(\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 - t\right)^2\right]$$

Proof. This lemma is a direct implication of Lemma 4.9.

□

Lemma 4.11. *Let $t \in [0, \infty)$ and let π_n be a partition of the interval $[0, t]$ for every $n \in \mathbb{N}$. For every $n \in \mathbb{N}$, if W is a Wiener process then:*

$$Var\left[\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 - t\right] = \sum_{j=0}^{n-1} Var\left[(W(t_{j+1}, \omega) - W(t_j, \omega))^2\right]$$

Proof.

W Wiener process $\Rightarrow$ Brownian motion process $\Rightarrow \forall s, s', t, t' \in [0, \infty], s \leqslant t$, $s' \leqslant t', s \neq s'$ and $t \neq t', W(t,\omega) - W(s,\omega)$ and $W(t',\omega) - W(s',\omega)$ are independent $\Rightarrow \forall s, s', t, t' \in [0,\infty], s \leqslant t, s' \leqslant t', s \neq s'$ and $t \neq t', (W(t,\omega) - W(s,\omega))^2$ and $(W(t',\omega) - W(s',\omega))^2$ are independent $\Rightarrow \forall s, s', t, t' \in [0,\infty], s \leqslant t, s' \leqslant t', s \neq s'$ and $t \neq t'$, and $\forall a, b \in \mathbb{R}, (W(t,\omega) - W(s,\omega))^2 - a$ and $(W(t',\omega) - W(s',\omega))^2 - b$ are independent $\Rightarrow \forall s, s', t, t' \in [0,\infty], s \leqslant t, s' \leqslant t', s \neq s'$ and $t \neq t'$,$\forall a, b \in \mathbb{R}$:

$$Var\left[(W(t',\omega) - W(s',\omega))^2 - a + (W(t',\omega) - W(s',\omega))^2 - b\right]$$
$$= Var\left[(W(t',\omega) - W(s',\omega))^2 - a\right] + Var\left[(W(t',\omega) - W(s',\omega))^2 - b\right]$$

Applying this result recursively:

$$Var\left[\sum_{j=0}^{n-1} (W(t_{j+1},\omega) - W(t_j,\omega))^2 - (t_{j+1} - t_j)\right]$$
$$= \sum_{j=0}^{n-1} Var\left[(W(t_{j+1},\omega) - W(t_j,\omega))^2 - (t_{j+1} - t_j)\right]$$

Finally, because adding a constant does not change the variance:

$$Var\left[\sum_{j=0}^{n-1} (W(t_{j+1},\omega) - W(t_j,\omega))^2 - t\right] = \sum_{j=0}^{n-1} Var\left[(W(t_{j+1},\omega) - W(t_j,\omega))^2\right]$$

□

Proposition 4.12. *For every $t \in [0,\infty)$,*

$$W \text{ Wiener process } \Rightarrow [W,W]_t = t \text{ in the } L^2 \text{ sense}$$

Proof. Let π_n be a partition of the interval $[0,t]$ for every $n \in \mathbb{N}$.

$$[W,W]_t = t \text{ in the } L^2 \text{ sense} \iff E\left[\left(\sum_{j=0}^{n-1} (W(t_{j+1},\omega) - W(t_j,\omega))^2 - t\right)^2\right] \xrightarrow[n\to\infty]{} 0$$

$$E\left[\left(\sum_{j=0}^{n-1} (W(t_{j+1},\omega) - W(t_j,\omega))^2 - t\right)^2\right]$$
$$\overset{\text{Lemma 4.10}}{=} Var\left[\sum_{j=0}^{n-1} (W(t_{j+1},\omega) - W(t_j,\omega))^2 - t\right]$$

$$\stackrel{\text{Lemma 4.11}}{=} \sum_{j=0}^{n-1} Var\left[(W(t_{j+1},\omega) - W(t_j,\omega))^2\right]$$

Therefore, if we define $Y_j := W(t_{j+1},\omega) - W(t_j,\omega)\ \forall j = 0, 1, ..., n-1$ we know that:

$$E\left[\left(\sum_{j=0}^{n-1}(W(t_{j+1},\omega) - W(t_j,\omega))^2 - t\right)^2\right] \xrightarrow[n\to\infty]{} 0 \iff \sum_{j=0}^{n-1} Var({Y_j}^2) \xrightarrow[n\to\infty]{} 0$$

Applying the definition of variance in $Var({Y_j}^2)$, we know that:

$$Var({Y_j}^2) = E\left[\left({Y_j}^2 - E\left({Y_j}^2\right)\right)^2\right] = E\left[{Y_j}^4 - 2{Y_j}^2 E\left({Y_j}^2\right) + \left(E\left({Y_j}^2\right)\right)^2\right]$$

$$= E({Y_j}^4) - \left(E\left({Y_j}^2\right)\right)^2$$

W Wiener process $\Rightarrow Y_j := W(t_{j+1},\omega) - W(t_j,\omega) \sim N(0, t_{j+1} - t_j) \Rightarrow E(Y_j) = 0$

$$\Rightarrow Var(Y_j) = E({Y_j}^2) = t_{j+1} - t_j$$

Therefore:

$$Var({Y_j}^2) = E({Y_j}^4) - (t_{j+1} - t_j)^2$$

$$Y_j \sim N(0, t_{j+1} - t_j) \Rightarrow E({Y_j}^4) = 3(t_{j+1} - t_j)^2$$

$$\Rightarrow Var({Y_j}^2) = 3(t_{j+1} - t_j)^2 - (t_{j+1} - t_j)^2 = 2(t_{j+1} - t_j)^2$$

Hence:

$$E\left[\left(\sum_{j=0}^{n-1}(W(t_{j+1},\omega) - W(t_j,\omega))^2 - t\right)^2\right] \xrightarrow[n\to\infty]{} 0 \iff \sum_{j=0}^{n-1} 2(t_{j+1} - t_j)^2 \xrightarrow[n\to\infty]{} 0$$

$$\sum_{j=0}^{n-1} 2(t_{j+1} - t_j)^2 \leqslant 2 \max_{j=0,\ ...,\ j=n-1}\{t_{j+1} - t_j\} \sum_{j=0}^{n-1}(t_{j+1} - t_j) = 2\Pi t \xrightarrow[n\to\infty]{} 0.$$

□

Proposition 4.13. *For every* $t \in [0,\infty)$,

$$W \text{ Wiener process } \Rightarrow Variation_t(W) = \infty$$

Proof. Consider a Wiener process W. For every $t \in [0,\infty)$, let $(\pi_n)_n$ be a series of partitions of $[0,t]$, $n \in \mathbb{N}$. Suppose $Variation_t(W) < \infty$. We will reach a contradiction.

$$Variation_t(W) := \lim_{\Pi\to 0} \sum_{j=0}^{n-1} |W(t_{j+1},\omega) - W(t_j,\omega)| < \infty$$

$$\Rightarrow \sup_{\pi_n}\left(\sum_{j=0}^{n-1} |W(t_{j+1},\omega) - W(t_j,\omega)|\right) < \infty$$

W Wiener process $\Rightarrow$ for every $\omega \in \Omega$, $\tilde{t} \mapsto W(\tilde{t}, \omega)$ is continuous on $\tilde{t} \in [0, t]$ for every $t \in [0, \infty]$.

$\tilde{t} \mapsto W(\tilde{t}, \omega)$ is continuous on $\tilde{t} \in [0, t]$ for every $t \in [0, \infty]$ and $[0, t]$ is a bounded interval $\Rightarrow W(\tilde{t}, \omega)$ is uniformly continuous on $[0, t]$

$$\Rightarrow \max_{j=0,\ldots,n-1} \{W(t_{j+1}, \omega) - W(t_j, \omega)\} \xrightarrow[\Pi \to 0]{} 0$$

$$\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 \text{ is less than}$$

$$\max_{j=0,\ldots,n-1} \{W(t_{j+1}, \omega) - W(t_j, \omega)\} \sum_{j=0}^{n-1} |W(t_{j+1}, \omega) - W(t_j, \omega)|$$

$$\sup_{\pi_n} \left(\sum_{j=0}^{n-1} |W(t_{j+1}, \omega) - W(t_j, \omega)| \right) < \infty \Rightarrow \sum_{j=0}^{n-1} |W(t_{j+1}, \omega) - W(t_j, \omega)| < \infty$$

for every $n \in \mathbb{N}$. Together with the already proved fact that

$$\max_{j=0,\ldots,n-1} \{W(t_{j+1}, \omega) - W(t_j, \omega)\} \xrightarrow[\Pi \to 0]{} 0$$

we get:

$$\max_{j=0,\ldots,n-1} \{W(t_{j+1}, \omega) - W(t_j, \omega)\} \sum_{j=0}^{n-1} |W(t_{j+1}, \omega) - W(t_j, \omega)| \xrightarrow[\Pi \to 0]{} 0$$

Therefore:

$$\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 \xrightarrow[\Pi \to 0]{} 0$$

But we have already proved in Proposition 4.12 that

$$\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 \xrightarrow[\Pi \to 0]{} t \text{ in the } L^2 \text{ sense}$$

$$\Rightarrow \exists \text{ some partition } \pi_{n_k} \text{ such that}$$

$$\sum_{j=0}^{n-1} (W(t_{j+1}, \omega) - W(t_j, \omega))^2 \xrightarrow[\max_{j=0,\ldots,n-1} \left\{ t^k_{j+1} - t^k_j \right\} \to 0]{} t$$

Having reached a contradiction, we know that it cannot be true that $Variation_t(W) < \infty$. Hence,

$$Variation_t(W) = \infty$$

$\square$

Proposition 4.14. *For every $t \in [0,\infty)$, almost all paths of W are not differentiable at t.*

Proof. Fix $t \in [0,\infty)$. Consider the set A defined as:

$$A := \{\omega \in \Omega : s \mapsto W(s,\omega) \text{ has derivative at } t\}$$

We will show that $\mathbb{P}(A) = 0$.

Let us define A_n^k, for $k \in \mathbb{N}$:

$$A_n^k := \left\{\omega \in \Omega : \text{for } \epsilon \text{ s.t. } 0 < \epsilon < \frac{1}{n}, n \in \mathbb{N}, -k\sqrt{\epsilon} \leqslant \frac{W(t+\epsilon,\sigma) - W(t,\sigma)}{\sqrt{\epsilon}} \leqslant k\sqrt{\epsilon}\right\}$$

Observe that, W Wiener process $\Rightarrow W(t+\epsilon,\sigma)-W(t,\sigma) \sim N(0,t+\epsilon-t) = N(0,\epsilon)$. Therefore, $\frac{W(t+\epsilon,\sigma)-W(t,\sigma)}{\sqrt{\epsilon}} \sim N(0,1)$. Hence, $\mathbb{P}(A_n^k) = \Phi(k\sqrt{\epsilon}) - \Phi(-k\epsilon)$, where Φ denotes the cumulative distribution function of a standardised normal distribution. Recall that $\Phi(-x) = 1 - \Phi(x)\ \forall x \in \mathbb{R}$, then:

$$\mathbb{P}(A_n^k) = 2\Phi(k\sqrt{\epsilon}) - 1 = \Phi(k\sqrt{\epsilon}) - \Phi\left(-k\epsilon\right) \leqslant 2\Phi\left(\frac{k}{\sqrt{n}}\right) - 1$$

$$\Rightarrow \mathbb{P}(A_n^k) \leqslant 2\Phi\left(\frac{k}{\sqrt{n}}\right) - 1 \xrightarrow[n\to\infty]{} 2\Phi\left(0\right) - 1 = 2\cdot\frac{1}{2} - 1 = 0 \Rightarrow \mathbb{P}(A_n^k) \xrightarrow[n\to\infty]{} 0$$

But it is $\mathbb{P}(A)$ that we want to compute:

$$\sigma \in A \Rightarrow s \mapsto W(s,\sigma) \text{ has derivative at } t$$

$$\Rightarrow \exists \lim_{\epsilon\to 0} \frac{W(t+\epsilon,\sigma) - W(t,\sigma)}{\epsilon}$$

$$\Rightarrow \text{for } \epsilon \text{ s.t. } 0 < \epsilon < \frac{1}{n}, n \in \mathbb{N}, \exists k \in \mathbb{R} \text{ such that } -k \leqslant \frac{W(t+\epsilon,\sigma) - W(t,\sigma)}{\epsilon} \leqslant k$$

$$\Rightarrow \text{for } \epsilon \text{ s.t. } 0 < \epsilon < \frac{1}{n}, n \in \mathbb{N}, \exists k \in \mathbb{R} \text{ such that}$$

$$-k\sqrt{\epsilon} \leqslant \frac{W(t+\epsilon,\sigma) - W(t,\sigma)}{\sqrt{\epsilon}} \leqslant k\sqrt{\epsilon}$$

$$\Rightarrow \exists k \in \mathbb{N} \text{ such that } \sigma \in A_n^k \Rightarrow A \subset \bigcup_{k=1}^{\infty}\bigcup_{n=1}^{\infty} A_n^k$$

$$A \subset \bigcup_{k=1}^{\infty}\bigcup_{n=1}^{\infty} A_n^k \Rightarrow \mathbb{P}(A) \leqslant \mathbb{P}\left(\bigcup_{k=1}^{\infty}\bigcup_{n=1}^{\infty} A_n^k\right) = \sum_{k=1}^{\infty}\left[\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^k\right)\right]$$

Observe that, for any $k, n \in \mathbb{N}$ $A_{n+1}^k \subset A_n^k$ because:

$$\omega \in A_{n+1}^k \Rightarrow \frac{-k}{\sqrt{n+1}} < -k\sqrt{\epsilon} \leqslant \frac{W(t+\epsilon,\omega) - W(t,\omega)}{\sqrt{\epsilon}} \leqslant k\sqrt{\epsilon} < \frac{k}{\sqrt{n+1}}$$

$$\Rightarrow \frac{-k}{\sqrt{n}} < -k\sqrt{\epsilon} \leqslant \frac{W(t+\epsilon,\omega)-W(t,\omega)}{\sqrt{\epsilon}} \leqslant k\sqrt{\epsilon} < \frac{k}{\sqrt{n}} \Rightarrow \omega \in A_n^k \Rightarrow A_{n+1}^k \subset A_n^k$$

$$A_{n+1}^k \subset A_n^k \Rightarrow \left(A_n^k\right)^c \subset \left(A_{n+1}^k\right)^c \Rightarrow \bigcup_{n=1}^{\infty} A_n^k = \left(\bigcap_{n=1}^{\infty} \left(A_n^k\right)^c\right)^c$$

$$\Rightarrow \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^k\right) = \mathbb{P}\left(\left(\bigcap_{n=1}^{\infty} \left(A_n^k\right)^c\right)^c\right) = 1 - \lim_{n\to\infty}\left(1-\mathbb{P}\left(A_n^k\right)\right) = \lim_{n\to\infty}\mathbb{P}\left(A_n^k\right) = 0$$

$$\mathbb{P}(A) \leqslant \sum_{k=1}^{\infty}\left[\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^k\right)\right] = 0 \Rightarrow \mathbb{P}(A) = 0$$

□

Theorem 4.15. *If W is a Wiener process, $W(t,\omega)$ and $e^{\sigma\cdot W(t,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}$ are martingales with respect to $\mathcal{F}_t^W$, the natural filtration of W.*

Proof. We will deal first with $W(t,\omega)$. Let $0 < s < t$. Because $\mathcal{F}_t^W$ is the natural filtration of W, we know that $W(s,\omega)$ is $\mathcal{F}_s^W$-measurable. Therefore,

$$E\left(W(t,\omega)|\mathcal{F}_s^W\right) - W(s,\omega) = E\left(W(t,\omega)-W(s,\omega)|\mathcal{F}_s^W\right)$$

W is a Wiener process $\Rightarrow$ W is a Brownian motion process $\Rightarrow$ $W(t,\omega)-W(s,\omega)$ and $\mathcal{F}_s^W$ are independent.Therefore, $E\left(W(t,\omega)-W(s,\omega)|\mathcal{F}_s^W\right) = E\left(W(t,\omega)-W(s,\omega)\right)$

$$W \text{ Brownian motion process } \Rightarrow W(t,\omega)-W(s,\omega) \sim N(0,t-s)$$

$$\Rightarrow E\left(W(t,\omega)-W(s,\omega)\right) = 0$$
$$E\left(W(t,\omega)-W(s,\omega)\right) = E\left(W(t,\omega)-W(s,\omega)|\mathcal{F}_s^W\right) = 0$$
$$\Rightarrow E\left(W(t,\omega)|\mathcal{F}_s^W\right) = W(s,\omega)$$

Hence, $W(t,\omega)$ is a martingale with respect to its natural filtration.

We turn to deal with $e^{\sigma\cdot W(t,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}$:

$$E\left[e^{\sigma\cdot W(t,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}|\mathcal{F}_s^W\right] = E\left[e^{\sigma\cdot W(t,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t+\sigma W(s,\omega)-\sigma W(s,\omega)}|\mathcal{F}_s^W\right]$$

$$= E\left[e^{\sigma\cdot W(t,\omega)-\sigma W(s,\omega)}e^{\sigma W(s,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}|\mathcal{F}_s^W\right]$$

Because $W(s,\omega)$ is $\mathcal{F}_s^W$-measurable, $e^{\sigma W(s,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}$ is also $\mathcal{F}_s^W$-measurable. Therefore:

$$E\left[e^{\sigma\cdot W(t,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}|\mathcal{F}_s^W\right] = e^{\sigma W(s,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}\cdot E\left[e^{\sigma\cdot W(t,\omega)-\sigma W(s,\omega)}|\mathcal{F}_s^W\right]$$

W is a Wiener process $\Rightarrow$ W is a Brownian motion process $\Rightarrow$ $W(t,\omega)-W(s,\omega)$ and $\mathcal{F}_s^W$ are independent. Therefore:

$$e^{\sigma W(s,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}\cdot E\left[e^{\sigma\cdot W(t,\omega)-\sigma W(s,\omega)}|\mathcal{F}_s^W\right] = e^{\sigma W(s,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}\cdot E\left[e^{\sigma[W(t,\omega)-W(s,\omega)]}\right]$$

W is a Wiener process $\Rightarrow W$ is a Brownian motion process $\Rightarrow W(t,\omega) - W(s,\omega)$

$$\sim N(0, t-s) \Rightarrow \sigma\left(W(t,\omega) - W(s,\omega)\right) \sim N(0, \sigma^2\,(t-s))$$

$$\sigma\left(W(t,\omega) - W(s,\omega)\right) \sim N(0, \sigma^2\,(t-s)) \Rightarrow E\left(e^{\sigma(W(t,\omega)-W(s,\omega))}\right) = e^{\frac{1}{2}\cdot\sigma^2\cdot(t-s)}$$

Finally:

$$E\left[e^{\sigma\cdot W(t,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}|\mathcal{F}_s^W\right] = e^{\sigma W(s,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}\cdot E\left[e^{\sigma[W(t,\omega)-W(s,\omega)]}\right]$$

$$= e^{\sigma W(s,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}\cdot e^{\frac{1}{2}\cdot\sigma^2\cdot(t-s)} = e^{\sigma W(s,\omega)-\frac{1}{2}\cdot\sigma^2\cdot s}$$

$\Rightarrow e^{\sigma\cdot W(t,\omega)-\frac{1}{2}\cdot\sigma^2\cdot t}$ is a martingale with respect to the natural filtration of W.

□

4.3 The stochastic integral

Let $W = (\Omega, \mathcal{F}, (\mathcal{F}_t)_t, (W_t)_t, \mathbb{P})$ be a Wiener process. In this subsection we want to give meaning to

$$\int_0^T X(s,w)dW(s,w),$$

where $(X(s,w))_{0\leqslant s\leqslant T}$ is a stochastic process enjoying certain properties.

4.3.1 Elementary processes

Let $(\pi_n)_n$ be a series of partitions of the time interval $[0,T]$ such that for every $n\in\mathbb{N}$ $\pi_n = \{0 = t_0, t_1, ..., t_n = T\}$, $0 = t_0 < t_1 < ... < t_n = T$.

Let $(X_i)_{i<n}$ be a series of random variables such that for every $n\in\mathbb{N}$ and $\forall i = 0, ..., n-1$ the random variable

$$\begin{aligned} X_i\colon \Omega &\to \mathbb{R} \\ \omega &\mapsto X_i(\omega) \end{aligned}$$

is $\mathcal{F}_{t_i}$-measurable. This fact deserves a special highlight: once the time t_i has occurred, the value $X_i(\omega)$ is known and is a real value.

Definition 4.16. *A stochastic process*

$$\begin{aligned} X\colon [0,T]\times\Omega &\to \mathbb{R} \\ (t,\omega) &\mapsto X(t,\omega) \end{aligned}$$

is an elementary process if and only if

$$X(t,\omega) = \sum_{i=0}^{n-1} X_i(\omega)\mathbb{1}_{[t_i,t_{i+1})}(t)$$

where X_i is a $\mathcal{F}_{t_i}$-measurable random variable for every $i = 0, ..., n-1$

Notice that an elementary process is right-continuous.

Definition 4.17. *Let X and X' be stochastic processes.*

$$X \sim^{**} X' \iff \int_0^T |X(s,\omega) - X'(s,\omega)| ds \overset{\text{a.s.}}{=} 0$$

Proposition 4.18. *$\sim^{**}$ is an equivalence relation*

Proof. For any X, Y and Z stochastic processes,

$$\int_0^T |X(s,\omega) - X(s,\omega)| ds = 0 \overset{\text{a.s.}}{=} 0 \Rightarrow X \sim^{**} X$$

If we assume that $X \sim^{**} Y$, then:

$$0 \overset{\text{a.s.}}{=} \int_0^T |X(s,\omega) - Y(s,\omega)| ds = \int_0^T |Y(s,\omega) - X(s,\omega)| ds \Rightarrow Y \sim^{**} X$$

If we assume that $X \sim^{**} Y$ and $Y \sim^{**} Z$, then:

$$\begin{aligned} 0 \leqslant \int_0^T |X(s,\omega) - Z(s,\omega)| ds &= \int_0^T |X(s,\omega) - Y(s,\omega) - Z(s,\omega) + Y(s,\omega)| ds \\ &\leqslant \int_0^T |X(s,\omega) - Y(s,\omega)| ds + \int_0^T |Y(s,\omega) - Z(s,\omega)| ds \overset{\text{a.s.}}{=} 0 + 0 \\ &\Rightarrow \int_0^T |X(s,\omega) - Z(s,\omega)| ds \overset{\text{a.s.}}{=} 0 \Rightarrow X \sim^{**} Z \end{aligned}$$

□

Definition 4.19.

$$\mathbb{X}([0,T]) := {\left\{ X \textit{ progressively measurable stochastic process : } X = \left(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, (X_t)_{0 \leqslant t \leqslant T}, \mathbb{P}\right) \right\}}\big/_{\sim^{**}}$$

The objective of the definition of $\mathbb{X}([0,T])$ is to avoid considering two equivalent stochastic processes in the sense of $\sim^{**}$ from now on. Even though an element of $\mathbb{X}([0,T])$ is a class of equivalence, we will not use any special notation to refer to it. I.e., we will refer to elements in $\mathbb{X}([0,T])$ with the notation of a stochastic process and we will just keep in mind that each element is "unique" in $\mathbb{X}([0,T])$ in the sense of $\sim^{**}$.

Because every right-continuous process is progressively measurable, we know that $\mathbb{X}([0,T])$ contains all elementary processes. A proof of such a result can be found in [1].

Definition 4.20.

$$M^p([0,T]) := \left\{ X \in \mathbb{X}([0,T]) \textit{ such that } E\left(\int_0^T |X(s,\omega)|^p ds \right) < \infty \right\}$$

Proposition 4.21. *Let X be an elementary process, i.e.:*

$$X(t,\omega) = \sum_{i=0}^{n-1} X_i(\omega)\mathbb{1}_{[t_i,t_{i+1})}(t)$$

Then:

$$X \in M^p\left([0,T]\right) \iff E\left(|X_i|^p\right) < +\infty,\ \forall i \in \{0,1,...,n-1\}$$

In particular,

$$X \in M^2\left([0,T]\right) \iff X_i \text{ is square-integrable } \forall i \in \{0,1,...,n-1\}$$

Proof.

$$X \in M^p\left([0,T]\right) \iff E\left(\int_0^T |X(s,\omega)|^p ds\right) < \infty$$

$$E\left(\int_0^T |X(s,\omega)|^p ds\right) = E\left(\int_0^T \left|\sum_{i=0}^{n-1} X_i(\omega)\mathbb{1}_{[t_i,t_{i+1})}(s)\right|^p ds\right)$$

$$= E\left(\sum_{i=0}^{n-1} |X_i(\omega)|^p \cdot (t_{i+1}-t_i)\right) = \sum_{i=0}^{n-1} E\left(|X_i(\omega)|\right) \cdot (t_{i+1}-t_i)$$

Therefore:

$$X \in M^p\left([0,T]\right) \iff \sum_{i=0}^{n-1} E\left(|X_i(\omega)|\right) \cdot (t_{i+1}-t_i) < +\infty$$

$$\iff E\left(|X_i|^p\right) < +\infty,\ \forall i \in \{0,1,...,n-1\}$$

□

Definition 4.22.

$$E^p\left([0,T]\right) := \{X \in M^p\left([0,T]\right) \text{ such that } X \text{ is an elementary process}\}$$

4.3.2 Stochastic integral of an elementary process

Definition 4.23. *Let $X \in E^2\left([0,T]\right)$ be an elementary process and W be a Wiener process. Then:*

$$\int_0^T X(t,\omega)dW(t,\omega) := \sum_{i=0}^{n-1} X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right]$$

is the stochastic integral of X with respect to W.

Notice that $\int_0^T X(t,\omega)dW(t,\omega)$ is itself a random variable, and not a stochastic process.

Lemma 4.24. *Let $X \in E^2([0,T])$ and W a Wiener process. Then, for every $a, b \in [0,T]$, $a < b$, consider the stochastic integral of X with respect to W in the interval $[a,b]$. Then:*

1.
$$E\left(\int_a^b X(t,\omega)dW(t,\omega)|\mathcal{F}_a\right) = 0$$

2.
$$E\left(\left(\int_a^b X(t,\omega)dW(t,\omega)\right)^2 |\mathcal{F}_a\right) = E\left(\int_a^b (X(t,\omega))^2 dt|\mathcal{F}_a\right)$$

In particular,
$$E\left(\left(\int_a^b X(t,\omega)dW(t,\omega)\right)^2\right) = E\left(\int_a^b (X(t,\omega))^2 dt\right)$$

Proof. We will first deal with 1. Let X be an elementary process, i.e.
$$X(t,\omega) = \sum_{i=o}^{n-1} X_i(\omega)\mathbb{1}_{[t_i,t_{i+1})}(t)$$
where each random variable X_i is $\mathcal{F}_{t_i}$-measurable.

By definition of stochastic integral of an elementary process:
$$E\left(\int_a^b X(t,\omega)dW(t,\omega)|\mathcal{F}_a\right)$$
$$= E\left(\sum_{i=0}^{n-1} X(t_i,\omega)\cdot[W(t_{i+1},\omega) - W(t_i,\omega)]|\mathcal{F}_a\right)$$
$$\xRightarrow{E[\cdot|\mathcal{F}_a] \text{ is linear}} E\left(\int_a^b X(t,\omega)dW(t,\omega)|\mathcal{F}_a\right)$$
$$= \sum_{i=0}^{n-1} E\left[X(t_i,\omega)\cdot[W(t_{i+1},\omega) - W(t_i,\omega)]|\mathcal{F}_a\right]$$

Due to one of the properties of conditional expectation, because
$$a \leqslant t_i \Rightarrow \mathcal{F}_a \subset \mathcal{F}_{t_i} \forall i \in \{0,1,...,n-1\}$$
$$\Rightarrow E\left[X(t_i,\omega)\cdot[W(t_{i+1},\omega) - W(t_i,\omega)]|\mathcal{F}_a\right]$$
$$= E\left[E\left[X(t_i,\omega)\cdot[W(t_{i+1},\omega) - W(t_i,\omega)]|\mathcal{F}_{t_i}\right]|\mathcal{F}_a\right]$$

Because $X(t_i,\omega)$ is $\mathcal{F}_{t_i}$-measurable, we know that
$$E\left[X(t_i,\omega)\cdot[W(t_{i+1},\omega) - W(t_i,\omega)]|\mathcal{F}_{t_i}\right] = X(t_i,\omega)\cdot E\left[W(t_{i+1},\omega) - W(t_i,\omega)|\mathcal{F}_{t_i}\right]$$

Because W is a Wiener process, it is also a Brownian motion process and therefore $W(t_{i+1},\omega) - W(t_i,\omega)$ is independent of $\mathcal{F}_{t_i}$. Then:

$$E\left[W(t_{i+1},\omega) - W(t_i,\omega)|\mathcal{F}_{t_i}\right] = E\left[W(t_{i+1},\omega) - W(t_i,\omega)\right]$$

Because W is a Wiener process, $W(t_{i+1},\omega) - W(t_i,\omega) \sim N(0, t_{i+1} - t_i)$. Then:

$$E\left[X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right]|\mathcal{F}_{t_i}\right] = E\left[W(t_{i+1},\omega) - W(t_i,\omega)\right] = 0$$

Hence:

$$E\left[X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right]|\mathcal{F}_a\right]$$
$$= E\left[E\left[X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right]|\mathcal{F}_{t_i}\right]|\mathcal{F}_a\right] = E\left[0|\mathcal{F}_a\right] = 0$$
$$E\left(\int_a^b X(t,\omega)dW(t,\omega)|\mathcal{F}_a\right)$$
$$= \sum_{i=0}^{n-1} E\left(X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right]|\mathcal{F}_a\right) = \sum_{i=0}^{n-1} 0 = 0$$

We turn to prove 2. By definition of stochastic integral of X with respect to W:

$$E\left(\left(\int_a^b X(t,\omega)dW(t,\omega)\right)^2|\mathcal{F}_a\right)$$

$$= E\left(\left(\sum_{i=0}^{n-1} X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right]\right)^2|\mathcal{F}_a\right)$$

$$= E\left(\sum_{i=0}^{n-1} X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right]\sum_{i=0}^{n-1} X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right]|\mathcal{F}_a\right)$$

$$= E\left(\sum_{i,j=0}^{n-1} (X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\left[W(t_{j+1},\omega) - W(t_j,\omega)\right])|\mathcal{F}_a\right)$$

and because $E\left[\ \cdot\ |\mathcal{F}_a\right]$ is linear:

$$= \sum_{i,j=0}^{n-1} E\left((X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\cdot\left[W(t_{j+1},\omega) - W(t_j,\omega)\right])|\mathcal{F}_a\right)$$

We break the summation with index j in three summations in the following manner:

$$\sum_{i,j=0}^{n-1} E\left((X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\cdot\left[W(t_{j+1},\omega) - W(t_j,\omega)\right])|\mathcal{F}_a\right)$$

$$= \sum_{i=0}^{n-1}\sum_{j=0}^{i-1} E\left((X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\left[W(t_{j+1},\omega) - W(t_j,\omega)\right])|\mathcal{F}_a\right)$$

$$+\sum_{i=0}^{n-1} E\left((X(t_i,\omega))^2 \cdot [W(t_{i+1},\omega) - W(t_i,\omega)]^2 | \mathcal{F}_a\right)$$

$$+\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1} E\left((X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\left[W(t_{j+1},\omega) - W(t_j,\omega)\right]) | \mathcal{F}_a\right)$$

Notice that $a \leqslant t_i, t_j \Rightarrow \mathcal{F}_a \subset \mathcal{F}_{t_i}$ and $\mathcal{F}_a \subset \mathcal{F}_{t_j}$. Further: if $t_i < t_j \Rightarrow \mathcal{F}_{t_i} \subset \mathcal{F}_{t_j}$ and if $t_j < t_i \Rightarrow \mathcal{F}_{t_j} \subset \mathcal{F}_{t_i}$. Therefore, in the summation

$$\sum_{i=0}^{n-1}\sum_{j=0}^{i-1} E\left((X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\left[W(t_{j+1},\omega) - W(t_j,\omega)\right]) | \mathcal{F}_a\right),$$

$\mathcal{F}_{t_j} \subset \mathcal{F}_{t_i}$ holds. Analogously, in the summation

$$\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1} E\left((X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\left[W(t_{j+1},\omega) - W(t_j,\omega)\right]) | \mathcal{F}_a\right),$$

$\mathcal{F}_{t_i} \subset \mathcal{F}_{t_j}$ holds. Hence:

If $i < j$ and X_i is $\mathcal{F}_i$-measurable $\Rightarrow X_i$ is $\mathcal{F}_j$-measurable andX_j is $\mathcal{F}_j$-measurable. Moreover, at time t_j the values of $X(t_i,\omega)$, $X(t_j,\omega)$,$W(t_{i+1},\omega)$ and $W(t_i,\omega)$ are already known, so:

$$E\left((X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\cdot\left[W(t_{j+1},\omega) - W(t_j,\omega)\right]) | \mathcal{F}_a\right)$$

$$= E\left[E\left((X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\left[W(t_{j+1},\omega) - W(t_j,\omega)\right]) | \mathcal{F}_{t_j}\right) | \mathcal{F}_a\right]$$

$$= E\left[X(t_i,\omega)\cdot X(t_j,\omega)\cdot [W(t_{i+1},\omega) - W(t_i,\omega)] E\left(W(t_{j+1},\omega) - W(t_j,\omega) | \mathcal{F}_{t_j}\right) | \mathcal{F}_a\right]$$

W Wiener process $\Rightarrow W$ is a Brownian motion process $\Rightarrow W(t_{j+1},\omega) - W(t_j,\omega)$ and $\mathcal{F}_{t_j}$ are independent $\Rightarrow E\left(W(t_{j+1},\omega) - W(t_j,\omega) | \mathcal{F}_{t_j}\right) = E\left(W(t_{j+1},\omega) - W(t_j,\omega)\right)$

W Brownian motion process $\Rightarrow W(t_{j+1},\omega) - W(t_j,\omega) \sim N(0, t_{j+1} - t_j)$, so

$$E\left(W(t_{j+1},\omega) - W(t_j,\omega)\right) = 0$$

This is the reason why

$$E\left((X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\cdot\left[W(t_{j+1},\omega) - W(t_j,\omega)\right]) | \mathcal{F}_a\right) = 0$$

Analogously, if $j < i$

$$E\left((X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\cdot\left[W(t_{j+1},\omega) - W(t_j,\omega)\right]) | \mathcal{F}_a\right) = 0$$

holds. Finally:

$$\sum_{i,j=0}^{n-1} E\left((X(t_i,\omega)\cdot\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\cdot\left[W(t_{j+1},\omega) - W(t_j,\omega)\right]) | \mathcal{F}_a\right)$$

$$= \sum_{i=0}^{n-1} \sum_{j=0}^{i-1} E\left((X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\left[W(t_{j+1},\omega) - W(t_j,\omega)\right])|\mathcal{F}_a\right)$$

$$+ \sum_{i=0}^{n-1} E\left((X(t_i,\omega))^2 \cdot \left[W(t_{i+1},\omega) - W(t_i,\omega)\right]^2 |\mathcal{F}_a\right)$$

$$+ \sum_{i=0}^{n-1} \sum_{j=i+1}^{n-1} E\left((X(t_i,\omega)\left[W(t_{i+1},\omega) - W(t_i,\omega)\right])(X(t_j,\omega)\left[W(t_{j+1},\omega) - W(t_j,\omega)\right])|\mathcal{F}_a\right)$$

$$= 0 + \sum_{i=0}^{n-1} E\left((X(t_i,\omega))^2 \cdot \left[W(t_{i+1},\omega) - W(t_i,\omega)\right]^2 |\mathcal{F}_a\right) + 0$$

$$= \sum_{i=0}^{n-1} E\left((X(t_i,\omega))^2 \cdot \left[W(t_{i+1},\omega) - W(t_i,\omega)\right]^2 |\mathcal{F}_a\right) =: E\left(\int_a^b (X(t,w))^2 dt |\mathcal{F}_a\right)$$

by the definition of stochastic integral of an elementary process and the linearity of conditional expectation.

□

Equip $E^2([0,T])$ with the inner product $< \cdot, \cdot >$ defined as:

$$< \cdot, \cdot >: E^2([0,T]) \times E^2([0,T]) \to \mathbb{R}$$
$$(X,Y) \mapsto < X, Y > := E\left(\int_a^b X(t,\omega) \cdot Y(t,\omega) dt\right)$$

So that the norm

$$\|X\|_{E^2([0,T])} := \sqrt{< X, X >}, \forall X \in E^2([0,T])$$

can be defined. Consequently, a distance in $E^2([0,T])$ can be defined too:

$$d_{E^2([0,T])}(X,Y) := \|X - Y\|_{E^2([0,T])}, \forall X, Y \in E^2([0,T])$$

Equip $L^2(\Omega)$ with the inner product $<< \cdot, \cdot >>$ defined as:

$$<< \cdot, \cdot >>: L^2(\Omega) \times L^2(\Omega) \to \mathbb{R}$$
$$(X,Y) \mapsto << X, Y >> := E(X \cdot Y)$$

So that the norm

$$\|X\|_{L^2(\Omega)} := \sqrt{<< X, X >>}, \forall X \in L^2(\Omega)$$

can be defined. Consequently, a distance in $L^2(\Omega)$ can be defined too:

$$d_{L^2(\Omega)}(X,Y) := \|X - Y\|_{L^2(\Omega)}, \forall X, Y \in L^2(\Omega)$$

Recall that, if $X \in E^2([0,T])$, then $\int_0^T X(t,\omega) dW(t,\omega) \in L^2(\Omega)$.

The previous Lemma is of paramount importance because it establishes an isometry between the metric spaces $\left(E^2([0,T]), d_{E^2([0,T])}\right)$ and $\left(L^2(\Omega), d_{L^2(\Omega)}\right)$. Indeed:

Corollary 4.25. *The map*

$$\int_0^T \cdot\, dW(t,\omega)\colon \left(E^2\left([0,T]\right), d_{E^2([0,T])}\right) \to \left(L^2\left(\Omega\right), d_{L^2(\Omega)}\right)$$
$$X \mapsto \int_0^T X(t,\omega) dW(t,\omega)$$

is an isometry.

Proof.

$$\left\| \int_0^T X(t,\omega) dW(t,\omega) \right\|_{L^2(\Omega)} = E\left(\left(\int_0^T X(t,w) dW(t,w) \right)^2 \right)$$
$$= E\left(\int_a^b \left(X(t,w)\right)^2 dt \right) = \|X\|_{E^2([0,T])}$$

□

Later on we will show that this isometry can be extended to the whole $M^2\left([0,T]\right)$.

4.3.3 Stochastic integral of a stochastic process

We need an approximation result that states that an M^p-process can be suitably approximated by elementary processes.

Let $\left(\pi_n\right)_n$ be a series of equi-spaced partitions of $[0,T]$.

Definition 4.26. *Let $f \in M^p([0,T])$. We define:*

$$G_n f(t,\omega) := \sum_{i=0}^{n-1} f_i(\omega) \cdot \mathbb{1}_{[t_i,t_{i+1})}(t)$$

where:

$$f_i(\omega) := \begin{cases} 0, \text{ if } i = 0 \\ \frac{1}{t_{i+1}-t_i} \cdot \int_{t_{i-1}}^{t_i} f(s,\omega) ds = \frac{1}{\frac{T-0}{n}} \cdot \int_{t_{i-1}}^{t_i} f(s,\omega) ds = \frac{n}{T} \cdot \int_{t_{i-1}}^{t_i} f(s,\omega) ds, \text{ if } i \neq 0 \end{cases}$$

Observe that f_i is a $\mathcal{F}_{t_i}$-measurable random variable. I.e., once t_i has occurred, we do know the real value $f_i(\omega)$. Therefore, $G_n f(t,\omega)$ is $\mathcal{F}_{t_i}$-measurable too.

Lemma 4.27. *If $f \in M^p\left([0,T]\right)$, then:*

1. $$\int_0^T |G_n f(s,\omega)|^p ds \leqslant \int_0^T |f(s,\omega)|^p ds.$$

2. $G_n f \in M^p\left([0,T]\right)$.

In particular, $G_n f \in E^p([0,T])$.

Proof. Let $f \in M^p([0,T])$ and consider $G_n f(t,\omega) := \sum_{i=0}^{n-1} f_i(\omega) \cdot \mathbb{1}_{[t_i,t_{i+1})}(t)$.

$$\int_0^T |G_n f(s,\omega)|^p ds = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} |G_n f(s,\omega)|^p ds$$

$$= \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \left| \sum_{i=0}^{n-1} f_i(\omega) \cdot \mathbb{1}_{[t_i,t_{i+1})}(t) \right|^p ds = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} |f_j(\omega)|^p ds$$

For every $j = 0, 1, ..., n-1$, $f_j(\omega)$ does not depend on s. Therefore:

$$\sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} |f_j(\omega)|^p ds = \sum_{j=0}^{n-1} |f_j(\omega)|^p \int_{t_j}^{t_{j+1}} ds = \sum_{j=0}^{n-1} |f_j(\omega)|^p \cdot (t_{j+1} - t_j)$$

$$= \frac{T}{n} \cdot \sum_{j=0}^{n-1} |f_j(\omega)|^p = \frac{T}{n} \cdot \sum_{j=0}^{n-1} \left| \frac{n}{T} \int_{t_{i-1}}^{t_i} f(s,\omega) ds \right|^p$$

Because $x \mapsto |x|^p$ is a convex function, Jensen's inequality applies and:

$$\frac{T}{n} \cdot \sum_{j=0}^{n-1} \left| \frac{n}{T} \int_{t_{i-1}}^{t_i} f(s,\omega) ds \right|^p = \frac{T}{n} \cdot \sum_{j=0}^{n-1} \frac{n}{T} \int_{t_{i-1}}^{t_i} |f(s,\omega)|^p ds$$

$$= \sum_{j=0}^{n-1} \int_{t_{i-1}}^{t_i} |f(s,\omega)|^p ds = \int_0^T |f(s,\omega)|^p ds$$

Hence:

$$\int_0^T |G_n f(s,\omega)|^p ds \leqslant \int_0^T |f(s,\omega)|^p ds$$

$$\Rightarrow E\left(\int_0^T |G_n f(s,\omega)|^p ds \right) \leqslant E\left(\int_0^T |f(s,\omega)|^p ds \right)$$

$$f \in M^p([0,T]) \Rightarrow E\left(\int_0^T |f(s,\omega)|^p ds \right) < +\infty$$

$$\Rightarrow E\left(\int_0^T |G_n f(s,\omega)|^p ds \right) < +\infty \Rightarrow G_n f \in M^p([0,T])$$

$G_n f$ satisfies the definition of elementary process, therefore $G_n f \in E^p([0,T])$.

□

Lemma 4.28.

$$f \textit{ continuous on } t \Rightarrow \lim_{n\to\infty} \int_0^T |G_n f(s,\omega) - f(s,\omega)|^p ds = 0$$

Proof. f continuous on t and $[0,T]$ compact $\Rightarrow$ f is uniformly continuous on $[0,T]$. Fix $\epsilon > 0$, f being uniformly continuous on $[0,T]$ means that $\exists n \in \mathbb{N}$ big enough such that for $s, s' \in [0,T]$, $|s - s'| \leqslant \frac{2T}{n} \Rightarrow |f(s) - f(s')| \leqslant \epsilon$.

Consider the function

$$\left(\int_0^{\cdot} |f(s,\omega)|^p ds\right)^{\frac{1}{p}} : [0,T] \to [0,\infty)$$
$$t \mapsto \left(\int_0^{t} |f(s,\omega)|^p ds\right)^{\frac{1}{p}}$$

$\left(\int_0^t |f(s,\omega)|^p ds\right)^{\frac{1}{p}}$ is continuous because the integral is a continuous function and f is continuous by hypothesis, moreover $[0,T]$ is compact $\Rightarrow$ $\left(\int_0^t |f(s,\omega)|^p ds\right)^{\frac{1}{p}}$ is uniformly continuous on $[0,T]$ $\Rightarrow \exists t_1 \geqslant 0$ close enough to 0 so that

$$\left(\int_0^{t_1} |f(s,\omega)|^p ds\right)^{\frac{1}{p}} \leqslant \epsilon$$
$$\Rightarrow \int_0^{t_1} |f(s,\omega)|^p ds \leqslant \epsilon^p$$

For any $i \in \{0, 1, ..., n-1\}$, $|[t_{i-1}, t_i) \bigcup [t_i, t_{i+1})| = 2\frac{T}{n}$ because π_n is an equi-spaced partition for evey $n \in \mathbb{N}$.

Let $u \in [t_{i-1}, t_i)$ and $s \in [t_i, t_{i+1})$. Then:

$$|u - s| \leqslant \frac{2T}{n} \xRightarrow{f \text{ uniformly continuous on } [0,T]} |f(s,\omega) - f(u,\omega)| \leqslant \epsilon(*)$$

$$G_n f(s,\omega) = f_i(\omega) = \frac{\int_0^T f(s,\omega) ds}{t_i - t_{i-1}} = \frac{\lim_{N\to\infty} \frac{t_i - t_{i-1}}{N} \sum_{k=1}^N f(t_{i-1} + k\frac{t_i - t_{i-1}}{N}, \omega)}{t_i - t_{i-1}}$$
$$= \lim_{N\to\infty} \frac{\sum_{k=1}^N f(t_{i-1} + k\frac{t_i - t_{i-1}}{N}, \omega)}{N}, \text{ which is the average of } f \text{ on } [t_{i-1}, t_i)$$
$$\Rightarrow f(t_{i-1},\omega) \leqslant = \lim_{N\to\infty} \frac{\sum_{k=1}^N f(t_{i-1} + k\frac{t_i - t_{i-1}}{N}, \omega)}{N} = G_n f(s,\omega) \leqslant f(t_i,\omega)$$
$$G_n f(s,\omega) \leqslant f(t_i,\omega) \Rightarrow G_n f(s,\omega) - f(s,\omega) \leqslant f(t_i,\omega) - f(s,\omega)$$
$$\Rightarrow |G_n f(s,\omega) - f(s,\omega)| \leqslant |f(t_i,\omega) - f(s,\omega)|$$
$$\xRightarrow{(*)} |G_n f(s,\omega) - f(s,\omega)| \leqslant |f(t_i,\omega) - f(s,\omega)| < \epsilon$$
$$\Rightarrow |G_n f(s,\omega) - f(s,\omega)|^p \leqslant |f(t_i,\omega) - f(s,\omega)|^p < \epsilon^p$$

Observe that if $t \in [0, t_1) \Rightarrow G_n f(t, \omega) := \sum_{i=0}^{n-1} f_i(\omega) \cdot \mathbb{1}_{[t_i, t_{i+1})}(t) = f_0 := 0$. Therefore:

$$\int_0^T |G_n f(s,\omega) - f(s,\omega)|^p ds = \int_0^{t_1} |G_n f(s,\omega) - f(s,\omega)|^p ds$$

$$+ \int_{t_1}^T |G_n f(s,\omega) - f(s,\omega)|^p ds = \int_0^{t_1} |-f(s,\omega)|^p ds + \int_{t_1}^T |G_n f(s,\omega) - f(s,\omega)|^p ds$$

We have already shown that $\int_0^{t_1} |f(s,\omega)|^p ds \leqslant \epsilon^p$ and $|G_n f(s,\omega) - f(s,\omega)|^p \leqslant \epsilon^p$. Hence:

$$\int_0^{t_1} |f(s,\omega)|^p ds + \int_{t_1}^T |G_n f(s,\omega) - f(s,\omega)|^p ds \leqslant \epsilon^p + \int_{t_1}^T \epsilon^p ds = \epsilon^p \cdot (1 + T - t_1)$$

$$\Rightarrow \int_0^T |G_n f(s,\omega) - f(s,\omega)|^p ds \leqslant \epsilon^p \cdot (1 + T - 0) \xrightarrow[\epsilon \to 0]{} 0$$

Finally, because $\epsilon \to 0 \iff n \to \infty$:

$$\int_0^T |G_n f(s,\omega) - f(s,\omega)|^p ds \xrightarrow[n \to \infty]{} 0$$

$\square$

This lemma can be proved to hold for a general stochastic process $f \in M^p([0,T])$. Such a result stems from the fact that continuous stochastic processes are dense in $M^p([0,T])$. A full prove is not provided in this text, but it can be found in [13].

Lemma 4.29.

$$\forall p \in \mathbb{N}, p \geqslant 1, \forall a, b \in \mathbb{R}, |a| \cdot |b|^p + |b| \cdot |a|^p \leqslant |a|^{p+1} + |b|^{p+1}$$

Proof. Induction on p.

If $p = 1$:

$$|a| \cdot |b|^1 + |b| \cdot |a|^1 = 2|a| \cdot |b| \leqslant |a|^{1+1} + |b|^{1+1}$$
$$\iff 0 \leqslant |a|^2 + |b|^2 - 2|a| \cdot |b| = (|a| - |b|)^2$$

Let us make the following hypothesis of induction: assume $|a| \cdot |b|^n + |b| \cdot |a|^n \leqslant |a|^{n+1} + |b|^{n+1}$ holds $\forall n \leqslant p - 1$. We will prove:

$$|a| \cdot |b|^p + |b| \cdot |a|^p \leqslant |a|^{p+1} + |b|^{p+1}$$

Because the above inequality is symmetric, we can assume that $|a| \leqslant |b|$ without loss of generality.

$$\text{I.H.: } |a| \cdot |b|^{p-1} + |b| \cdot |a|^{p-1} \leqslant |a|^p + |b|^p$$

$$\Rightarrow |a| \cdot |b| \cdot \left(|a| \cdot |b|^{p-1} + |b| \cdot |a|^{p-1}\right) \leqslant |a| \cdot |b| \cdot (|a|^p + |b|^p)$$

$$
\begin{aligned}
&\Rightarrow |a|^2 \cdot |b|^p + |b|^2 \cdot |a|^p \leqslant |a| \cdot |b|^{p+1} + |b| \cdot |a|^{p+1} \\
&\Rightarrow |a|^2 \cdot |b|^p - |a| \cdot |b|^{p+1} \leqslant |b| \cdot |a|^{p+1} - |b|^2 \cdot |a|^p \\
&\Rightarrow |a| \cdot \left(|a||b|^p - |b|^{p+1}\right) \leqslant |b| \left(|a|^{p+1} - |b||a|^p\right) \\
&\xRightarrow{|a|\leqslant|b|} |a||b|^p - |b|^{p+1} \leqslant |a|^{p+1} - |b||a|^p \\
&\Rightarrow |a| \cdot |b|^p + |b| \cdot |a|^p \leqslant |a|^{p+1} + |b|^{p+1}
\end{aligned}
$$

□

Lemma 4.30.

$$\forall p \in \mathbb{N}, p \geqslant 1, \forall a, b \in \mathbb{R}, |a-b|^p \leqslant 2^{p-1}(|a|^p + |b|^p)$$

Proof. Induction on p.

If $p = 1$:

$$|a-b|^1 \leqslant 2^{1-1}\left(|a|^1 + |b|^1\right) = |a| + |b|, \text{which is a triangle inequality}$$

Let us make the following hypothesis of induction: assume $|a-b|^n \leqslant 2^{n-1}(|a|^n + |b|^n)$ holds $\forall n \leqslant p$. We will prove:

$$|a-b|^{p+1} \leqslant 2^p\left(|a|^{p+1} + |b|^{p+1}\right)$$

Consider $|a-b|^{p+1}$: $|a-b|^{p+1} = |a-b| \cdot |a-b|^p$. By hypothesis of induction:

$$
\begin{aligned}
|a-b| \cdot |a-b|^p &\leqslant 2^{1-1}\left(|a|^1 + |b|^1\right) \cdot 2^{p-1}(|a|^p + |b|^p) = \\
&= 2^{p-1} \cdot \left(|a|^{p+1} + |b|^{p+1} + |a| \cdot |b|^p + |b| \cdot |a|^p\right) \\
&\leqslant 2^{p-1} \cdot \left(|a|^{p+1} + |b|^{p+1} + |a|^{p+1} + |b|^{p+1}\right) \text{ because of Lemma 4.29}
\end{aligned}
$$

Therefore:

$$|a-b|^{p+1} \leqslant 2^p\left(|a|^{p+1} + |b|^{p+1}\right)$$

□

Proposition 4.31.

$$X \in M^p([0,T]) \Rightarrow \exists (G_n X)_n \subset E^p([0,T]) \text{ such that:}$$

$$\lim_{n\to\infty} E\left(\int_0^T |X(s,\omega) - G_n X(s,\omega)|^p ds\right) = 0$$

Moreover, the elementary processes $G_n X$ can be chosen in such a way that

$$n \mapsto \int_0^T |G_n X(s,\omega)|^p ds$$

is increasing.

Proof. Let $X \in M^p([0,T])$. Consider $(GnX)_n \subset E^p([0,T])$.

Consider $\int_0^T |X(s,\omega) - G_nX(s,\omega)|^p ds$ and apply Lemma 4.30 on it:

$$\int_0^T |X(s,\omega) - G_nX(s,\omega)|^p ds \leqslant \int_0^T 2^{p-1} \left(|X(s,\omega)|^p + |G_nX(s,\omega)|^p\right) ds$$

We already know that $\int_0^T |G_nX(s,\omega)|^p ds \leqslant \int_0^T |X(s,\omega)|^p ds$, therefore:

$$2^{p-1}\left(\int_0^T |X(s,\omega)|^p ds + \int_0^T |G_nX(s,\omega)|^p ds\right)$$

$$\leqslant 2^{p-1}\left(\int_0^T |X(s,\omega)|^p ds + \int_0^T |X(s,\omega)|^p ds\right) = 2^p \int_0^T |X(s,\omega)|^p ds$$

$$X \in M^p([0,T]) \Rightarrow 2^p E\left(\int_0^T |X(s,\omega)|^p ds\right) < +\infty$$

$$\Rightarrow E\left(\int_0^T |X(s,\omega) - G_nX(s,\omega)|^p ds\right) < +\infty$$

In Lemma 4.28 we have already shown that $\lim_{n\to\infty} \int_0^T |G_nf(s,\omega) - f(s,\omega)|^p ds = 0$. Hence, the hypotheses of the Dominated Convergence Theorem are satisfied and we can assert that

$$\lim_{n\to\infty} E\left(\int_0^T |X(s,\omega) - G_nX(s,\omega)|^p ds\right) = 0$$

Realise that $G_{2^n}f = G_{2^n}(G_{2^{n+1}}f)$ since $G_{2^{n+1}}f$ has an implicit partition of $[0,T]$ that is finer than that of $G_{2^n}f$'s. Therefore, the elementary process G_nX can be chosen in such a way that

$$n \mapsto \int_0^T |G_nX(s,\omega)|^p ds$$

is increasing. □

In Corollary 4.25 we saw that the stochastic integral of elementary processes is an isometry. Now the previous lemma has showed that elementary processes are dense in $M^p([0,T])$, so that the isometry can be extended to the whole $M^p([0,T])$, thus defining the stochastic integral for every $X \in M^p([0,T])$.

Definition 4.32. *Let $X \in M^2([0,T])$ and $(G_nX)_n \subset E^2([0,T])$ approximating X in the sense of Proposition 4.31. We define:*

$$\int_0^T X(t,\omega)dW(t,\omega) := \lim_{n\to\infty} \int_0^T G_nX(t,\omega)dW(t,\omega) \text{ in the } L^2(\Omega) \text{ sense}$$

That is, $\int_0^T G_nX(t,\omega)dW(t,\omega)$ is defined as the random variable that satisfies

$$\lim_{n\to\infty} E\left[\left(\int_0^T X(t,\omega)dW(t,\omega) - \int_0^T G_nX(t,\omega)dW(t,\omega)\right)^2\right] = 0$$

Lemma 4.33. *$\int_0^T X(t,\omega)dW(t,\omega)$ for an arbitrary process $X \in M^2([0,T])$ is well defined.*

Proof. Two results have to be proven in order to prove the lemma:

1. The existence of $\lim_{n\to\infty} \int_0^T G_nX(t,\omega)dW(t,\omega)$ in $L^2(\Omega)$.
2. If two elementary processes approximate X in the sense of Proposition 4.31, then these two elementary processes approximate each other in the sense of Proposition 4.31.

We will first deal with 1.

Let $X \in M^2([0,T])$ and let $(G_nX)_n \subset E^2([0,T])$ be such that

$$\lim_{n\to\infty} E\left(\int_0^T |X(s,\omega) - G_nX(s,\omega)|^p ds\right) = 0$$

That is, $G_nX \longrightarrow X$ in $M^2([0,T])$. Let us define $I(G_nX := \int_0^T G_nX(t,\omega)dW(t,\omega)$. We want to prove that the sequence $(I(G_nX))_n$ converges in $L^2(\Omega)$.

$$G_nX \longrightarrow X \text{ in } M^2([0,T]) \Rightarrow (G_nX)_n \text{ is a Cauchy series in } M^2([0,T])$$

$$\Rightarrow \exists N \in \mathbb{N} \text{ sucht that for } n,m \geqslant N, d_{M^2([0,T])}(G_nX, G_mX) \xrightarrow[n,m\to\infty]{} 0$$

$$d_{M^2([0,T])}(G_nX, G_mX) = \|G_nX - G_mX\|_{M^2([0,T])}$$

$$= E\left(\int_0^T (G_nX(t,w) - G_mX(t,w))^2 dt\right)$$

$$= E\left(\left(\int_0^T G_nX(t,w) - G_mX(t,w)dW(t,w)\right)^2\right) = \|G_nX - G_mX\|_{L^2(\Omega)}$$

Therefore:

$$d_{L^2(\Omega)}(G_nX, G_mX) \xrightarrow[n,m\to\infty]{} 0$$

$$\Rightarrow (I(G_nX))_n \text{ is a Cauchy series in } L^2(\Omega)$$

$$\Rightarrow \exists \lim_{n\to\infty} \int_0^T G_nX(t,\omega)dW(t,\omega) \text{ in } L^2(\Omega)$$

We turn now to prove 2..

Let $(X_n)_n, (Y_n)_n \subset E^2([0,T])$ such that

$$\lim_{n\to\infty} E\left(\int_0^T |X(s,\omega) - X_n(s,\omega)|^p ds\right) = 0$$

and

$$\lim_{n\to\infty} E\left(\int_0^T |X(s,\omega) - Y_n(s,\omega)|^p ds\right) = 0$$

Then:

$$\lim_{n\to\infty} E\left(\int_0^T |X(s,\omega)-X_n(s,\omega)|^p ds\right) + \lim_{n\to\infty} E\left(\int_0^T |X(s,\omega)-Y_n(s,\omega)|^p ds\right)$$

$$= \lim_{n\to\infty} E\left(\int_0^T |X(s,\omega)-X_n(s,\omega)|^p + |X(s,\omega)-Y_n(s,\omega)|^p ds\right) = 0$$

$$\Rightarrow \lim_{n\to\infty} E\left(\int_0^T 2^{p-1}\left(|X(s,\omega)-X_n(s,\omega)|^p + |X(s,\omega)-Y_n(s,\omega)|^p\right) ds\right) = 0$$

By Lemma 4.30 we know that:

$$0 \leqslant \lim_{n\to\infty} E\left(\int_0^T |X(s,\omega)-X_n(s,\omega)-(X(s,\omega)-Y_n(s,\omega))|^p ds\right)$$

$$= \lim_{n\to\infty} E\left(\int_0^T |Y_n(s,\omega)-X_n(s,\omega)|^p ds\right)$$

$$\leqslant \lim_{n\to\infty} E\left(\int_0^T 2^{p-1}\left(|X(s,\omega)-X_n(s,\omega)|^p + |X(s,\omega)-Y_n(s,\omega)|^p\right) ds\right) = 0$$

$$\Rightarrow \lim_{n\to\infty} E\left(\int_0^T |Y_n(s,\omega)-X_n(s,\omega)|^p ds\right) = 0$$

□

Lemma 4.34. *Let $X \in M^2([0,T])$ and W a Wiener process. Then, for every $a,b \in [0,T]$, $a<b$, consider the stochastic integral of X with respect to W in the interval $[a,b]$. Then:*

1.

$$E\left(\int_a^b X(t,\omega)dW(t,\omega)|\mathcal{F}_a\right) = 0$$

2.

$$E\left(\left(\int_a^b X(t,\omega)dW(t,\omega)\right)^2 |\mathcal{F}_a\right) = E\left(\int_a^b (X(t,\omega))^2 dt|\mathcal{F}_a\right)$$

In particular,

$$E\left(\left(\int_a^b X(t,\omega)dW(t,\omega)\right)^2\right) = E\left(\int_a^b (X(t,\omega))^2 dt\right)$$

Proof. We will first deal with 1.

Let $X \in M^2([0,T])$ and $(G_nX)_n \subset E^2([0,T])$ approximating X in the sense of Proposition 4.31. Because $(G_nX)_n \subset E^2([0,T])$, we know from Lemma 4.24. that

$$E\left(\int_a^b G_nX(t,\omega)dW(t,\omega)|\mathcal{F}_a\right) = 0$$

holds. Moreover, in Lemma 4.33. we have demonstrated the existence in $L^2(\Omega)$ of

$$\lim_{n\to\infty}\int_0^T G_nX(t,\omega)dW(t,\omega) =: \int_a^b X(t,\omega)dW(t,\omega)$$

Therefore:

$$E\left(\int_a^b G_nX(t,\omega)dW(t,\omega)|\mathcal{F}_a\right) = 0$$

$$\Rightarrow \lim_{n\to\infty} E\left(\int_a^b G_nX(t,\omega)dW(t,\omega)|\mathcal{F}_a\right) = 0$$

Because $E[\ \cdot\ |\mathcal{F}_a]$ is continuous, we know that:

$$\lim_{n\to\infty} E\left(\int_a^b G_nX(t,\omega)dW(t,\omega)|\mathcal{F}_a\right) = E\left(\lim_{n\to\infty}\int_a^b G_nX(t,\omega)dW(t,\omega)|\mathcal{F}_a\right)$$

Hence:

$$E\left(\int_a^b X(t,\omega)dW(t,\omega)|\mathcal{F}_a\right) = 0$$

We now turn to focus on 2.

Let $X \in M^2([0,T])$ and $(G_nX)_n \subset E^2([0,T])$ approximating X in the sense of Proposition 4.31. Because $(G_nX)_n \subset E^2([0,T])$, we know from Lemma 4.24. that

$$E\left(\left(\int_a^b G_nX(t,\omega)dW(t,\omega)\right)^2|\mathcal{F}_a\right) = E\left(\int_a^b (G_nX(t,\omega))^2dt|\mathcal{F}_a\right)$$

$$\Rightarrow \lim_{n\to\infty} E\left(\left(\int_a^b G_nX(t,\omega)dW(t,\omega)\right)^2|\mathcal{F}_a\right) = \lim_{n\to\infty} E\left(\int_a^b (G_nX(t,\omega))^2dt|\mathcal{F}_a\right)$$

Concentrate on the left hand side term of the last equality. Because $E[\ \cdot\ |\mathcal{F}_a]$ is linear, we know that:

$$\lim_{n\to\infty} E\left(\left(\int_a^b G_nX(t,\omega)dW(t,\omega)\right)^2|\mathcal{F}_a\right)$$

$$= E\left(\lim_{n\to\infty}\left(\int_a^b G_nX(t,\omega)dW(t,\omega)\right)^2|\mathcal{F}_a\right)$$

$$\int_a^b G_nX(t,\omega)dW(t,\omega) \xrightarrow[n\to\infty]{L^2(\Omega)} \int_a^b X(t,\omega)dW(t,\omega)$$

$$\Rightarrow \left(\int_a^b G_nX(t,\omega)dW(t,\omega)\right)^2 \xrightarrow[n\to\infty]{L^1(\Omega)} \left(\int_a^b X(t,\omega)dW(t,\omega)\right)^2$$

$$\Rightarrow \lim_{n\to\infty} E\left(\left(\int_a^b G_nX(t,\omega)dW(t,\omega)\right)^2|\mathcal{F}_a\right) = E\left(\left(\int_a^b X(t,\omega)dW(t,\omega)\right)^2|\mathcal{F}_a\right)$$

Concentrate now on the right hand side of the equality. Thanks to Proposition 4.31. we can assume that

$$\left(\int_a^b (G_n X(t,\omega))^2 dt\right)_n$$

is an increasing sequence. Therefore, Beppo Levi's Theorem[3] is applicable and

$$\lim_{n\to\infty} E\left(\int_a^b (G_n X(t,\omega))^2 dt|\mathcal{F}_a\right) = E\left(\int_a^b (X(t,\omega))^2 dt|\mathcal{F}_a\right)$$

Hence, the statement of the Lemma follows. □

4.3.4 The stochastic integral as a stochastic process and its martingale property

Let $X \in M^2([0,T])$, then the restriction of X to $[0,t]$, $t \leqslant T$, also belongs to $M^2([0,t])$ and we can consider its integral $\int_0^t X(s,\omega)dW(s,\omega)$. Let the real-valued process I_X be defined as:

$$\begin{aligned} I_X \colon [0,T]\times\Omega &\to \mathbb{R} \\ (t,\omega) \mapsto I_X(t,\omega) &:= \int_0^t X(s,\omega)dW(s,\omega) \end{aligned}$$

Theorem 4.35. *If $X \in M^2([0,T])$, then I_X is an $\mathcal{F}_t$-square integrable martingale.*

Proof. If $s < t$, thanks to Lemma 4.34:

$$E\left[I_X(t,\omega) - I_X(s,\omega)|\mathcal{F}_s\right] = E\left[\int_s^t X(u,\omega)dW(u,\omega)|\mathcal{F}_s\right] = 0$$

Because $I_X(s,\omega)$ is $\mathcal{F}_s$-measurable and together with the above equality we know that

$$\begin{aligned} E\left[I_X(t,\omega) - I_X(s,\omega)|\mathcal{F}_s\right] &= E\left[I_X(t,\omega)|\mathcal{F}_s\right] - I_X(s,\omega) = 0 \\ &\Rightarrow E\left[I_X(t,\omega)|\mathcal{F}_s\right] = I_X(s,\omega) \end{aligned}$$

□

4.4 Stochastic differential equations, Itô processes and the Itô formula

In this subsection we introduce the Itô formula, a key tool that makes computations in stochastic calculus easier and which we will use in the following chapter when dealing with the Black-Scholes model.

[3] A proof of Beppo Levi's Theorem can be found in [1].

Definition 4.36. *Let W be a Wiener process. An $\mathbb{R}$-valued stochastic process X is an Itô process if it can be written as*

$$X(t,\omega) = X(0,\omega) + \int_0^t K(s,\omega)ds + \int_0^t H(s,\omega)dW(s,\omega) \ , \mathbb{P} \ a.s. \ \forall t \leqslant T,$$

where:

- $X(0,\omega)$ *is $\mathcal{F}_0$-measurable.*
- $(K(t,\omega))_{0\leqslant t\leqslant T}$ *and* $(H(t,\omega))_{0\leqslant t\leqslant T}$ *are $\mathcal{F}_t$-adapted processes.*
- $\int_0^t |K(s,\omega)|\, ds < +\infty$ $\mathbb{P}$ *a.s.*
- $\int_0^t (H(s,\omega))^2 ds < +\infty$ $\mathbb{P}$ *a.s.*

Because in the previous sections we have given meaning to the integral

$$\int_0^t H(s,\omega)dW(s,\omega),$$

the Itô process concept is well defined.

Definition 4.37. *If X is an Itô process which can be written as*

$$X(t,\omega) = X(0,\omega) + \int_0^t K(s,\omega)ds + \int_0^t H(s,\omega)dW(s,\omega) \ , \mathbb{P} \ a.s. \ \forall t \leqslant T,$$

we say that X admits the stochastic differential equation

$$dX(t,\omega) = K(t,\omega)dt + H(t,\omega)dW(t,\omega)$$

The equation $dX(t,\omega) = K(t,\omega)dt + H(t,\omega)dW(t,\omega)$ lacks of mathematical meaning because a Wiener process is almost nowhere differentiable. The equation $dX(t,\omega) = K(t,\omega)dt + H(t,\omega)dW(t,\omega)$ is just an abbreviation of the equation in Definition 4.36.

Theorem 4.38. *Let W be a Wiener process.*

Let

$$\begin{aligned} f\colon \mathbb{R}\times\mathbb{R} &\to \mathbb{R} \\ (t,x) &\mapsto f(t,x) \end{aligned}$$

be such that $f \in \mathcal{C}^{1,2}(\mathbb{R}\times\mathbb{R})$. Let f_t, f_x and f_{xx} denote, respectively, de derivative of f with respect to t, the derivative of f with respect to x and the second derivative of f with respect to x. Then:

$$\begin{aligned} f(t,W(t,\omega)) = f(0,W(0,\omega)) &+ \int_0^t f_s(s,W(s,\omega))ds + \int_0^t f_x(s,W(s,\omega))dW(s,\omega) \\ &+ \frac{1}{2}\int_0^t f_{xx}(s,W(s,\omega))d[W,W]_s, \end{aligned}$$

where $[W,W]$ is the variation of the Wiener process S as defined in Definition 4.8.

Proof. A proof of this Theorem and other more generalised Itô Formula Theorems can be found in [1]. $\square$

5 The role of the Stochastic Integral in the Black-Scholes model

The model suggested by Black and Scholes to describe the behaviour of prices is a continuous-time model with:

1. A riskless asset S_t^0 whose evolution through time is encapsulated in the differential equation $dS_t^0 = rS_t^0 dt$, where $r \in \mathbb{R}^+$ is an instantaneous interest rate.
2. A risky asset $S(t,\omega)$ that evolves according to the stochasctic differential equation
$$dS(t,\omega) = \mu S(t,\omega)dt + \sigma dW(t,\omega)$$
where $\mu, \sigma \in \mathbb{R}$ and W is a Wiener process.

The idea behind the equation $dS(t,\omega) = \mu S(t,\omega)dt + \sigma dW(t,\omega)$ is that the price of the risky asset moves randomly around some linear tendency, so that σ stands for the volatility of the risky asset. Note that such an assumption could not be made without a formal definition of the stochastic integral.

Observe that
$$\int_0^t |\mu S(s,\omega)|\, ds < +\infty \ \mathbb{P} \text{ a.s.}$$
and
$$\int_0^t (\sigma S(s,\omega))^2 ds < +\infty \ \mathbb{P} \text{ a.s.}$$
Therefore, it will make sense solving $dS(t,\omega) = \mu S(t,\omega)dt + \sigma dW(t,\omega)$.

These assumptions lead to a concrete formula for the evolution of the price of the risky asset.

Theorem 5.1. *The process*
$$\begin{aligned} S\colon \mathbb{R}\times\Omega &\to \mathbb{R} \\ (t,\omega) &\mapsto S(t,\omega) = S_0 e^{\mu t - \frac{\sigma^2}{2}t + \sigma W(t,\omega)} \end{aligned}$$
where $S_0 := S(0,\omega)$, solves the stochastic differential equation
$$dS(t,\omega) = \mu S(t,\omega)dt + \sigma dW(t,\omega)$$

Proof. Consider function f defined as
$$\begin{aligned} f\colon \mathbb{R}\times\mathbb{R} &\to \mathbb{R} \\ (t,x) &\mapsto f(t,x) = S_0 e^{\mu t - \frac{\sigma^2}{2}t + \sigma x} \end{aligned}$$

f is infinitely differentiable with respect both variables x and t. Therefore, Itô formula from Theorem 4.38 is applicable to $f(t, W(t,\omega)) = S(t,\omega)$. We compute first f_t, f_x and f_{xx}:

$$f_t(t,x) = \left(\mu - \frac{\sigma^2}{2}\right) f(t,x)$$
$$f_x(t,x) = \sigma f(t,x)$$
$$f_{xx}(t,x) = \sigma^2 f(t,x)$$

Applying Itô's formula:

$$f(t, W(t,\omega)) = f(0, W(0,\omega)) + \int_0^t f_s(s, W(s,\omega))ds + \int_0^t f_x(s, W(s,\omega))dW(s,\omega)$$
$$+\frac{1}{2}\int_0^t f_{xx}(s, W(s,\omega))d[W,W]_s.$$

In Proposition 4.12. we demonstrated that

$$[W,W] = t \text{ in the } L^2 \text{ sense}$$
$$\Rightarrow S(t,\omega) = S_0 + \int_0^t \left(\mu - \frac{\sigma^2}{2}\right) S(s,\omega)ds + \int_0^t \sigma S(s,\omega)dW(s,\omega)$$
$$+\frac{1}{2}\int_0^t \sigma^2 S(s,\omega)ds$$
$$\Rightarrow S(t,\omega) = S_0 + \int_0^t \left(\mu - \frac{\sigma^2}{2} + \frac{\sigma^2}{2}\right) S(s,\omega)ds + \int_0^t \sigma S(s,\omega)dW(s,\omega)$$
$$\Rightarrow S(t,\omega) = S_0 + \int_0^t \mu S(s,\omega)ds + \int_0^t \sigma S(s,\omega)dW(s,\omega)$$
$$\Rightarrow S(t,\omega) = S_0 e^{\mu t - \frac{\sigma^2}{2}t + \sigma W(t,\omega)} \text{ solves } dS(t,\omega) = \mu S(t,\omega)dt + \sigma dW(t,\omega)$$

□

A proof of the uniqueness of that solution can be found in [5].

6 Theory put into practice

6.1 Do the theoretical models fit the real world?

In their own abstract world, the theoretical models produce a perfect hedge of derivatives. But the theoretical world of derivatives models differs the real world in many aspects. We discus some of these differences in this section.

1. Jumps

The stochastic integral that has been defined in this project is based on the Wiener process as its building block. As it has already been pointed out, the Wiener process is characterised by having *continuous paths*. Therefore, the stochastic integral is only defined over continuous processes.

The real world price changes exhibit systematic discontinuous jumps though. Thus, hedging errors in using the models presented in this project will stem from this difference.

Jumps could be captured by the models if *discontinuous paths* were blended into the Wiener process. Such a merged process exists; the Poisson process. It is said that the Wiener process is fit for modelling "normal events" whereas the Poisson process enables the possibility of modelling "rare events". The interested reader may find in [9] a further discussion on the topic.

2. Discrete Trading

In the theoretical models, the hedging portfolio is assumed to be adjusted continuously. In the real world, hedging portfolios are rebalanced in a discrete fashion, thus existing some hedge error due to this difference.

On the one hand, even if a very powerful computer executed the trades in unnoticeable short time lapses, the adjusting of the hedging portfolio would still be discrete. On the other hand, in the real world, adjusting the hedging portfolio comes with transaction costs, therefore continuous adjusting becomes an unfeasible practice.

In [10], the author carries out a study based on real life numbers concluding that the discrete trading difference entails no practical problem for a big derivatives firm. More concretely, the conclusion is that, on-average (i.e., taking into account all the hedging strategies that the firm simultaneously enters into instead of focusing on each individual hedging strategy), discrete rebalancing of the hedging portfolios corrects the hedge error made by implicitly assuming continuous adjusting in the models.

3. Transaction Costs

In the theoretical models there exist no transaction costs. But in the real world there are indeed transaction costs, namely [4]:

[4] The four transaction costs and their explanation is quoted from [11].

- *Commissions: the fee that must be paid to a broker to execute an order.*
- *Market impact: the displacement of prices resulting from order arrival.*
- *Opportunity costs: the costs (which may be negative) of market price movements due to forces other than the particular trade during the time elapsed from the moment the decision to trade was made until the actual completion of the trade.*
- *Miscellaneous other costs: this includes items such as transfer taxes and execution errors.*

These transaction costs have not a negligible effect on the prices of derivatives. Furthermore, *they contribute to the uncertainty with respect to the ultimate outcome of the hedging process, as sometimes the hedge will require relatively little adjustment and therefore generate high transaction costs*[5].

One of the mentioned transaction costs, the taxes levied on trade executions, brings back the philosophical discussion on whether the Tobin tax should be present in hedging-related trade executions as they provoke more expensive derivatives prices and, consequently, harm the efficiency of the derivatives market and, by extension, society as a whole. Ultimately, the main problem is that it is difficult to differenciate hedging versus speculative trade executions.

4. Volatility Misprediction

One of the inputs of the Black-Scholes model is volatility, which is an *a priori* unknown variable. Even though the value of volatility is uncertain, some value must be put into the model in order to function. There are two possible measures of volatility that can be put into the model:

- Historical volatility: computed by means of linear regression or other statistical methodology that uses the past data of the underlying.
- Implied volatility: computed by means of non-linear solvers applied to the Black-Scholes model that take into account the actual in-the-market quoted prices of derivatives sharing the same underlying and contract details.

On the one hand, because past behaviour of the markets is not a good forecast of future behaviour of the markets, historical volatility may not be a good input in the model.

On the other hand, implied volatility inherits al the Black-Scholes model assumptions that might be wrong, thus conveying a potentially wrong insight of future volatility.

Difficulties in ascertaining the volatility input of the model produce hedging errors as well as further uncertainty on the outcome of the hedge.

[5]Quoted from [11].

6.2 Alternative hedging strategies

A way of avoiding the hedging errors mentioned in the previous section is, simply, not building the hedging strategy that the theoretical model proposes. In this section we briefly mention some possible hedging strategies that a derivatives firm may adopt.

6.2.1 Common sense hedging: back-to-back dealing

This approach is straight-forward: the derivatives firm has a client asking for a concrete derivative and the derivatives firm, before entering into the contract with the client, looks for another counterparty willing to enter into that exact contract but at a better price. If the details of the two derivatives contracts are exactly the same except for their prices, the hedge is perfect and no complex mathematics have had to be used.

6.2.2 Common sense hedging: parity relationships

This approach shares the same basic principle of back-to-back dealing, in the sense that the derivatives firm has a client asking for a concrete derivative and the derivative firm, before entering into the contract with the client, looks for other counterparties from whom will get a perfect hedge at a better price. The difference between back-to-back dealing and this strategy lies in the fact that the derivatives firm looks for other counterparties (in plural) each willing to enter into a part of the contract.

This possibility of breaking the original contract into parts stems from parity relations: sometimes the payoff of a contract can be decomposed into the payoffs of two or more other contracts. The classical parity relation example is that of the equivalence of the buying of a forward contract and a simultaneous buying of a call option and selling of a put option.

Again, if the details of the different contracts involved match in all terms except for the prices, the hedge is perfect and no complex mathematics have had to be used.

6.2.3 Taylor series hedging

Let V be the price of an arbitrary option. Assume the price of that arbitrary option is determined by two variables: one being I (the reference index of the option), and the other being σ (the implied volatility of the option).

Such a hypothetical option is therefore a function of these two variables, i.e. $V = f(I, \sigma)$, for some $f : \mathbb{R}^2 \to \mathbb{R}$ that we will consider, for simplicity, infinitely derivable on both variables I and σ.

Consider the Taylor series expansion of $V = f(I, \sigma)$ around the actual (I_0, σ_0)

at which the option is considered to be trading[6]:

$$\Delta V = \frac{\partial f(I,\sigma)}{\partial I}(I - I_0) + \frac{1}{2}\frac{\partial^2 f(I,\sigma)}{\partial I^2}(I - I_0)^2 + \frac{1}{6}\frac{\partial^3 f(I,\sigma)}{\partial I^3}(I - I_0)^3 + ...$$
$$+\frac{\partial f(I,\sigma)}{\partial \sigma}(\sigma - \sigma_0) + \frac{1}{2}\frac{\partial^2 f(I,\sigma)}{\partial \sigma^2}(\sigma - \sigma_0)^2 + \frac{1}{6}\frac{\partial^3 f(I,\sigma)}{\partial \sigma^3}(\sigma - \sigma_0)^3 + ...$$

Because the objective of a hedging strategy is being able to deliver the value of the option at the expiry date, then a good hedge is one that is able to keep track of every change ΔV.

The above formula conveys that one can keep track of ΔV by possessing a basket of other options dependent on the same variables I and σ. We explain this in detail:

Suppose that we take a position in a basket of options and that this basket as a whole has a price dependant on I and σ. Suppose that this dependency is captured by function $g : \mathbb{R}^2 \to \mathbb{R}$ so that $g(I,\sigma)$ is the price of that basket of options. We will again consider, for simplicity, that g is infinitely derivable on both variables I and σ. Suppose the following equalities hold:

$$\frac{\partial g(I,\sigma)}{\partial I} = \frac{\partial f(I,\sigma)}{\partial I}, \frac{\partial^2 g(I,\sigma)}{\partial I^2} = \frac{\partial^2 f(I,\sigma)}{\partial I^2}, \frac{\partial^3 g(I,\sigma)}{\partial I^3} = \frac{\partial^3 f(I,\sigma)}{\partial I^3}, ...$$
$$\frac{\partial g(I,\sigma)}{\partial \sigma} = \frac{\partial f(I,\sigma)}{\partial \sigma}, \frac{\partial^2 g(I,\sigma)}{\partial \sigma^2} = \frac{\partial^2 f(I,\sigma)}{\partial \sigma^2}, \frac{\partial^3 g(I,\sigma)}{\partial \sigma^3} = \frac{\partial^3 f(I,\sigma)}{\partial \sigma^3}, ...$$

Then the basket of options keeps track of ΔV as is, therefore, a good hedge strategy.

The non-mathematical explanation of the rationale behind this hedging strategy is the following: if we are able to find a set of options that have the same sensibility to the changes in I and σ, then we have found a good hedge.

In the previous two alternative strategies, a the derivatives firm looked for a new counterparty who entered into a contract with the same exact details as the former contract. This Taylor series strategy is easier to achieve as that contract must not be the same, but just have the same sensibility to the changes in the variables I and σ.

Because of their importance, the sensitivities

$$\frac{\partial f(I,\sigma)}{\partial I}, \frac{\partial^2 f(I,\sigma)}{\partial I^2}, \frac{\partial^3 f(I,\sigma)}{\partial I^3}$$

are called *delta*, *gamma* and *omega*, respectively.

Likewise, the sensitivities

$$\frac{\partial f(I,\sigma)}{\partial \sigma}, \frac{\partial^2 f(I,\sigma)}{\partial \sigma^2}$$

are called *vega* and *omega*, respectively.

[6] I_0 is observable in the market and σ_0 is implicit in the quote of similar options.

6.3 An example of target client of a derivatives firm

6.3.1 Dealing with the exposure to foreign exchange risk

Consider a business that periodically buys a commodity, processes it and then sells a finished product to the market. The managers of the company know a priori the N dates that the purchases of commodity will have to be made because the company has closed contracts with several clients for this year. The company has limited storing space for supplies, that is the reason why purchases have to be made periodically.

In order to show how can foreign exchange risk arise, let us make three assumptions about this company: the price of the commodity it buys never changes, it buys that commodity in the same currency as it sells its finished products and this company finances its operations without the need of asking for loans nor any other type of debt. Such a company faces several risks, namely[7]:

1. *Non-financial risks:*
   - (a) *Legal risk: the risk of being sued over a transaction and the risk that the terms of a contract will not be upheld by the legal system.*
   - (b) *Compliance risk: the risk derived from the matter of conforming to policies, laws, rules and regulations, as set by governments and authoritative bodies.*
   - (c) *Model risk: the risk of a valuation error from improperly using a model, by using the wrong model or using the right one incorrectly.*
   - (d) *Operational risk: the risk that arises from the people and processes that the company combines to produce the output of the organisation.*
   - (e) *Solvency risk: the risk that the entity does not survive or succeed because it runs out of cash, even though it might otherwise be solvent.*
2. *Financial risks:*
   - (a) *Credit risk: the risk of loss if one of its clients fails to pay an amount owed on an obligation.*
   - (b) *Liquidity risk: the risk that stems from the lack of marketability of an investment that cannot be bought or sold quickly enough to prevent or minimise a loss.*

Let us change one of the assumptions: assume that this company buys the commodity in a foreign currency. Now this company faces a new financial risk:

Definition 6.1. *Market risk is the risk derived from the uncertainties arising from the movements in interest rates, stock prices, exchange rates and commodities*[8].

[7]Quoted from [7].
[8]Quoted from [7].

In particular, if the foreign currency suddenly becomes expensive in terms of the company's local currency, the company's costs rise. Of course, if that foreign currency becomes cheap in terms of the company's local currency, the company's costs decrease. This rising or decreasing of the costs and the management of that company are completely unconnected. Still, a rise of costs can damage severely the P&L of that company, so this problem must be tackled.

Derivatives offer an effective solution for this kind of business hazards. Many derivatives would be fit for this company to avoid future foreign exchange uncertainties. We mention two:

1. Buying a European call option on the foreign currency exchange rate with expiry date one year from now and with N exercise dates.

The company buys the right but not the obligation of being able to exercise a predetermined foreign exchange rate. Therefore, the danger of the foreign exchange to rise is limited to this predetermined rate.

2. Buying a European call option on the foreign currency exchange rate with a knock-out barrier with expiry date one year from now and with N exercise dates.

The company has the same advantages as in the previous case but together with one benefit and a disadvantage. The benefit is that the price of buying this option is cheaper. The disadvantage is that if the level of the knock-out barrier is reached, this company looses the protection that the derivative offers.

7 Conclusions

On one hand, we have began studying the Cox-Ross-Rubinstein model, the discrete model that assumes that prices of risky assets either go a little bit up or a little bit down at each step. Once a certain number of steps have occurred, the model assigns a probability to each discrete possible price, therefore allowing an expectation to be computed with which the fair (risk-neutral) price of a derivative can be found.

Moreover, we have seen that if we shorten infinitely the time lapse between each step, the valuation of European calls and puts (assuming that prices of risky assets move according to the CRR model) is given by the Black-Scholes formula.

On the other hand, the Wiener process serves as a good model for random continuous evolutions of prices of risky assets, but, as we have demonstrated, its drawback is that it is not differentiable. This characteristic of the Wiener process unveils the need of a different concept of integral. The stochastic integral allows the definition of some *"derivative of the Wiener process"* with which stochastic differential equations can be considered.

Finally, we have shown that the stochastic differential equation in which the price of a risky asset moves with some random volatility around some linear tendency (the stochastic differential equation that stems from the assumptions of the Black-Scholes model) has a unique solution which is the model of the continuous evolution of the prices of risky assets.

8 Annex

8.1 Numerical methods

In this section, we study three of the many programs that [6] contains.

8.1.1 Pricing and hedging European options using the CRR model

This program prices European options using the CRR model and builds their associated theoretical hedging strategy.

The principal idea that the author of [6] highlights when dealing with option numerical method tasks is that programs should be built in a way such that adding new components does not force the programmer to rearrange all the already written code.

In order to reach that objective, the author of [6] argues that the code should be casted in the style of object-oriented programming. The classes defined in the program should reflect relationships between real entities; the binomial model and European options in this case. If that principle is followed, not only the program gives the desired output but it can also be expanded with new options without interfering with existing files. We will explain how this is done in comments to different parts of the code.

Listing 1: Main11.cpp

```
#include "BinModel02.h"
#include "Options06.h"
#include "DoubDigitOpt.h"
#include "Strangle.h"
#include "Butterfly.h"
#include <iostream>
#include <cmath>
using namespace std;

int main()
{

        BinModel Model;

        if (Model.GetInputData() == 1) {
                return 1;
        }

        Call Option1;
        Option1.GetInputData();
        BinLattice<double> RiskyDeltaHedgeCall;
        BinLattice<double> RisklessDeltaHedgeCall;
        cout << "European call option price = " <<
        Option1.PriceByCRR(Model, RiskyDeltaHedgeCall,
        RisklessDeltaHedgeCall) << endl << endl;
        cout << "RiskyDeltaHedgeCall tree: " << endl;
        RiskyDeltaHedgeCall.Display();
```

```
cout << "RisklessDeltaHedgeCall tree: " << endl << endl;
RisklessDeltaHedgeCall.Display();

Put Option2;
Option2.GetInputData();
BinLattice<double> RiskyDeltaHedgePut;
BinLattice<double> RisklessDeltaHedgePut;
cout << "European put option price = " <<
Option2.PriceByCRR(Model, RiskyDeltaHedgePut,
RisklessDeltaHedgePut) << endl << endl;
cout << "RiskyDeltaHedgePut tree: " << endl;
RiskyDeltaHedgePut.Display();
cout << "RisklessDeltaHedgePut tree: " << endl << endl;
RisklessDeltaHedgePut.Display();

DoubDigitOpt Option3;
Option3.GetInputData();
BinLattice<double> RiskyDeltaHedgeDoubDigitOpt;
BinLattice<double> RisklessDeltaHedgeDoubDigitOpt;
cout << "European double-digital option price = " <<
Option3.PriceByCRR(Model, RiskyDeltaHedgeDoubDigitOpt,
RisklessDeltaHedgeDoubDigitOpt) << endl << endl;
cout << "RiskyDeltaHedgeDoubDigitOpt tree: " << endl;
RiskyDeltaHedgeDoubDigitOpt.Display();
cout << "RisklessDeltaHedgeDoubDigitOpt tree: " << endl
<< endl;
RisklessDeltaHedgeDoubDigitOpt.Display();

Strangle Option4;
Option4.GetInputData();
BinLattice<double> RiskyDeltaHedgeStrangle;
BinLattice<double> RisklessDeltaHedgeStrangle;
cout << "Strangle option price = " <<
Option4.PriceByCRR(Model, RiskyDeltaHedgeStrangle,
RisklessDeltaHedgeStrangle) << endl << endl;
cout << "RiskyDeltaHedgeStrangle tree: " << endl;
RiskyDeltaHedgeStrangle.Display();
cout << "RisklessDeltaHedgeStrangle tree: " << endl
<< endl;
RisklessDeltaHedgeStrangle.Display();

Butterfly Option5;
Option5.GetInputData();
BinLattice<double> RiskyDeltaHedgeButterfly;
BinLattice<double> RisklessDeltaHedgeButterfly;
cout << "Butterfly option price = " <<
Option5.PriceByCRR(Model, RiskyDeltaHedgeButterfly,
RisklessDeltaHedgeButterfly)
<< endl << endl;
cout << "RiskyDeltaHedgeButterfly tree: " << endl;
RiskyDeltaHedgeButterfly.Display();
cout << "RisklessDeltaHedgeButterfly tree: "
<< endl << endl;
RisklessDeltaHedgeButterfly.Display();
```

```
        return 0;
}
```

- Comment on Listing 1:

The $main()$ function simply creates $Option1$, $Option2$, $Option3$, $Option4$ and $Option5$ and treats these objects by calling other functions of the program.

Listing 2: BinModel02.h

```
#ifndef BinModel02_h
#define BinModel02_h

class BinModel {
private:
        double S0;
        double U;
        double D;
        double R;

public:
        //computing risk-neutral probability
        double RiskNeutProb();

        //computing the stock price at node n,i
        double S(int n, int i);

        //Inputting, displaying and checking model data
        int GetInputData();

        double GetR();
};

#endif
```

- Comment on Listing 2:

In this header file a new class of type BinModel is created. The objective of defining this class is leaving out anything related to options in the binomial model and making it to only consist of a market account and a stock.

Members of the BinModel class can either be private or public. The private members of the class BinModel are only accessible in $main()$ and other parts of the program via this class. In contrast, public members of the class BinModel will be accessible outside the class.

Observe that there are no variables mentioned in the parentheses that follow the name of the functions in the class. The reason why these lines can be compiled is that every member function of the class has access to variables specified as members of the class. Note also that in this case all functions are public members, but this need not be the general rule.

Listing 3: BinModel02.cpp

```
#include "BinModel02.h"
#include <iostream>
#include <cmath>
using namespace std;

double BinModel::RiskNeutProb() {
        return (R - D) / (U - D);
}

double BinModel::S(int n, int i) {
        return S0 * pow(1 + U, i)*pow(1 + D, n - i);
}

int BinModel::GetInputData() {
        //entering data
        cout << "Enter S0: ";
        cin >> S0;
        cout << "Enter U: ";
        cin >> U;
        cout << "Enter D: ";
        cin >> D;
        cout << "Enter R: ";
        cin >> R;
        cout << endl;

        //making sure that 0<S0, -1<D<U, -1<R
        if (S0 <= 0.0 || U <= -1.0 || D <= -1.0 || U <= D || R <=
                -1.0) {
                cout << "Illegal data ranges" << endl;
                cout << "Terminating program" << endl;
                return 1;
        }

        //checking for arbitrage
        if (R >= U || R <= D) {
                cout << "Arbitrage exists" << endl;
                cout << "Terminating program" << endl;
                return 1;
        }

        cout << "Input data checked" << endl;
        cout << "There is no arbitrage" << endl << endl;

        return 0;
}

double BinModel::GetR() {
        return R;
}
```

- Comment on Listing 3:

Just like in Listing 2, the header of Listing 3, there are no declarations of parameters passed to any member function whenever the parameters are members of the class.

Listing 4: BinLattice02.h

```
#ifndef BinLattice02_h
#define BinLattice02_h

#include "pch.h"
#include <iostream>
#include <iomanip>
#include <vector>
using namespace std;

template<typename Type> class BinLattice {
private:
        int N;
        vector<vector<Type>> Lattice;

public:
        void SetN(int N_) {
                N = N_;
                Lattice.resize(N+1);
                for (int n = 0; n <= N;n++) {
                        Lattice[n].resize(n+1);
                }
        }

        void SetNode(int n, int i, Type x) {
                Lattice[n][i] = x;
        }

        Type GetNode(int n, int i) {
                return Lattice[n][i];
        }

        void Display() {
                cout << setiosflags(ios::fixed) << setprecision(3);
                for (int n = 0; n <= N;n++) {
                        for (int i = 0; i <= n;i++) {
                                cout << setw(7) << GetNode(n,i);
                                cout << endl;
                        }
                }
        }
};

#endif
```

- Comment on Listing 4:

BinLattice is a class template with type parameter *Type*. Because a class template can only be compiled after an object has been declared using the template with a specific data type, there is no *.cpp* file corresponding to *BinLattice02.h*.

The BinLattice class consists of $N+1$ vectors whose components are vectors too. This construction allows the program to build nodes. In these nodes, the strategy of how many shares and how much money in the account have to be kept at every node will be stored.

Listing 5: Options06.h

```
#pragma once
#ifndef Options06_h
#define Options06_h
#include "pch.h"

#include "BinModel02.h"
#include "BinLattice02.h"

class EurOption {
private:
        //steps to expiry
        int N;

public:
        void SetN(int N_) { N = N_; }
        virtual double Payoff(double z) = 0;
        //pricing European option
        double PriceByCRR(BinModel Model,
        BinLattice<double>& x, BinLattice<double>& y);
};

//computing call payoff
double CallPayoff(double z, double K);

class Call : public EurOption {
private:
        double K; //strike price
public:
        void SetK(double K_) { K = K_; }
        int GetInputData();
        double Payoff(double z);
};



class Put : public EurOption {
private:
        double K; //strike price

public:
        void SetK(double K_) { K = K_; }
        int GetInputData();
        double Payoff(double z);
};

#endif
```

- Comment on Listing 5:

In Listing 5, the *EurOption* class is defined, referring to the entity of European Options. Again, private members of the class can only be used by the rest of the program via the EurOption class. In contrast, public members can be used without problem in the rest of the program.

All European Options have N steps to expiry and a payoff function. In this program, European Options are priced using the CRR model, and their hedging strategy is built under this model too. x and y will contain this hedging strategy.

In Listing 5 there are two subclasses of *EurOption* declared too: *Call* and *Put*. These two subclasses have their own strike price K and their specific $Payoff()$ function attached to their class.

The crucial code line in this Listing is *virtual double* $Payoff($*double* $z) = 0$;. Because every different European Option has a different payoff function, the class *EurOption* is built in such a way that the program will be able to differenciate between different payoff functions of options inside the class. This job is possible because of this *virtual* function. Being declared as a *virtual* function in the *EurOption* class makes it possible for $Payoff()$ to recognise, at run time, when an object of a subclass is passed to it via a pointer to the parent class: the version of $Payoff()$ belonging to that subclass is then executed.

Listing 6: Options06.cpp

```
#include "Options06.h"
#include "BinModel02.h"
#include "BinLattice02.h"
#include <math.h>
#include <iostream>
#include <cmath>
#include <vector>
using namespace std;

double EurOption::PriceByCRR(BinModel Model,
        BinLattice<double>& x, BinLattice<double>& y) {
        double q = Model.RiskNeutProb();
        vector<double> Price(N + 1, 0);
        x.SetN(N);
        y.SetN(N);
        for (int i = 0; i <= N; i++) {
                Price[i] = Payoff(Model.S(N, i));
        }
        for (int n = N-1; n >= 0; n--) {
                for (int i = 0; i <= n; i++) {
                        x.SetNode(n,i,(Price[i+1]-Price[i])/
                        (Model.S(n+1,i+1)-Model.S(n+1,i)));
                        y.SetNode(n, i, (1 / (pow(1 + Model.GetR(), 1)))*
                        (Price[i] - x.GetNode(n, i)*Model.S(n+1, i)));
                        Price[i] = (q*Price[i + 1] + (1 - q)*Price[i]) /
                        (1 + Model.GetR());
                }
        }
        return Price[0];
}

int Call::GetInputData() {
        cout << "Enter call option data: " << endl;
```

```
        int N = 0;
        cout << "Enter steps to expiry N: ";
        cin >> N;
        SetN(N);
        cout << "Enter strike price K: ";
        cin >> K;
        cout << endl;
        if (K < 0 || N < 0) {
                cout << "Wrong data: neither the strike price nor"<<
                "the number of steps to expiry can be negative" << endl;
                return 1;
        }
        return 0;
}

double Call::Payoff(double z) {
        if (z > K) {
                return z - K;
        }
        return 0.0;
}

int Put::GetInputData() {
        cout << "Enter put option data: " << endl;
        int N = 0;
        cout << "Enter steps to expiry N: ";
        cin >> N;
        SetN(N);
        cout << "Enter strike price K: ";
        cin >> K;
        cout << endl;
        if (K < 0 || N < 0) {
                cout << "Wrong data: neither the strike price"<<
                "nor the number of steps to expiry can be negative"
               << endl;
                return 1;
        }
        return 0;
}

double Put::Payoff(double z) {
        if (z < K) {
                return K - z;
        }
        return 0.0;
}
```

- Comment on Listing 6:

This Listing contains the code that executes the CRR model price for an option. Instead of using the closed formulas for calls and puts demonstrated in 2.3.,this program uses an iterative process to compute prices such that can be used in a *EurOption* general subclass. As the program computes that price, the hedging strategy is being computed too and is saved in vectors x and y.

Furthermore, the payoffs of *Call* and *Put* are coded in this Listing.

Listing 7: Strangle.h

```
#ifndef Strangle_h
#define Strangle_h
#include "pch.h"
#include "Options06.h"
#include "BinModel02.h"


class Strangle : public EurOption {
private:
        double K1; //parameter 1
        double K2; //parameter 2

public:
        int GetInputData();
        double Payoff(double z);
};

#endif
```

Listing 8: Strangle.cpp

```
#include "BinModel02.h"
#include "Options06.h"
#include "DoubDigitOpt.h"
#include "Strangle.h"
#include <iostream>
using namespace std;

int Strangle::GetInputData() {
        cout << "Enter strangle option data; " << endl;
        int N;
        cout << "Enter steps to expiry N: ";
        cin >> N;
        SetN(N);
        cout << "Enter parameter K1:      ";
        cin >> K1;
        cout << "Enter parameter K2:      ";
        cin >> K2;
        cout << endl;
        return 0;
}

double Strangle::Payoff(double z) {
        if (z<=K1) {
                return K1-z;
        }
        if (K2<z) {
                return z-K2;
        }
        return 0;
}
```

Listing 9: DoubDigitOpt.h

```
#ifndef DoubDigitOpt_h
#define DoubDigitOpt_h
#include "pch.h"
#include "Options06.h"
#include "BinModel02.h"


class DoubDigitOpt : public EurOption {
private:
        double K1; //parameter 1
        double K2; //parameter 2

public:
        int GetInputData();
        double Payoff(double z);
};

#endif
```

Listing 10: DoubDigitOpt.cpp

```
#include "BinModel02.h"
#include "Options06.h"
#include "DoubDigitOpt.h"
#include <iostream>
using namespace std;

int DoubDigitOpt::GetInputData() {
        cout << "Enter double-digital option data; " << endl;
        int N;
        cout << "Enter steps to expiry N: ";
        cin >> N;
        SetN(N);
        cout << "Enter parameter K1:      ";
        cin >> K1;
        cout << "Enter parameter K2:      ";
        cin >> K2;
        cout << endl;
        return 0;
}

double DoubDigitOpt::Payoff(double z) {
        if (K1<z && z<K2) {
                return 1.0;
        }
        return 0;
}
```

Listing 11: Butterfly.h

```
#ifndef Butterfly_h
#define Butterfly_h
#include "pch.h"
#include "Options06.h"
#include "BinModel02.h"
```

```
class Butterfly : public EurOption {
private:
        double K1; //parameter 1
        double K2; //parameter 2

public:
        int GetInputData();
        double Payoff(double z);
};

#endif
```

Listing 12: Butterfly.cpp

```
#include "BinModel02.h"
#include "Options06.h"
#include "DoubDigitOpt.h"
#include "Butterfly.h"
#include <iostream>
using namespace std;

int Butterfly::GetInputData() {
        cout << "Enter butterfly option data; " << endl;
        int N;
        cout << "Enter steps to expiry N: ";
        cin >> N;
        SetN(N);
        cout << "Enter parameter K1:      ";
        cin >> K1;
        cout << "Enter parameter K2:      ";
        cin >> K2;
        cout << endl;
        return 0;
}

double Butterfly::Payoff(double z) {
        if (K1<z && z<=(K1+K2)/2) {
                return z - K1;
        }
        if (((K1+K2)/2)<z && z<=K2) {
                return K2 - z;
        }
        return 0;
}
```

- Comment on Listings 7, 8, 9, 10, 11 and 12:

These Listings show how the program can be expanded in order to add different European options to be priced and hedged using the CRR model. Because the program is built in an object-oriented manner, all that is needed is a *.h* file that specifies that a new subclass of *EurOption* is built together with a *.cpp* in which the specific $Payoff()$ function is coded.

8.1.2 Computing implied volatility using the Black-Scholes model

In this section we show a program that computes the implied volatility of a European call. This can be done by assuming the Black-Scholes model holds, then looking at the quoted prices of similar calls in the market and taking their price in the market as an input to the Black-Scholes formula. I.e., instead of using the formula

$$C_0 = S_0 \cdot \Phi(-d_-) - \frac{K}{e^{rT}} \cdot \Phi(-d_+)$$

with

$$d_+ := \frac{1}{\sigma} \cdot \log\left(\frac{K}{S_0}\right) - \frac{rT}{\sigma} + \frac{\sigma}{2}$$
$$d_- := d_+ - \sigma$$

to compute C_0, we use it to compute σ. In order to do so, the use of non-linear solvers is required: the Bisection Method and the Newton-Raphson method.

Listing 13: EurCall.h

```
#ifndef EurCall_h
#define EurCall_h

class EurCall {
public:
        double T, K;
        EurCall(double T_, double K_) { T = T_; K = K_; }
        double d_plus(double S0, double sigma, double r);
        double d_minus(double S0, double sigma, double r);
        double PriceByBSFormula(double S0, double sigma, double r);
        double VegaByBSFormula(double S0, double sigma, double r);
};

#endif
```

- Comment on Listing 13:

This header file simply defines the *EurCall* class and its members.

Listing 14: EurCall.cpp

```
#include "EurCall.h"
#include <cmath>
#include "math.h"

double N(double x) {
        double gamma = 0.2316419;
        double a1 = 0.319381530;
        double a2 = -0.356563782;
        double a3 = 1.781477937;
        double a4 = -1.821255978;
        double a5 = 1.330274429;
        double pi = 4.0*atan(1.0);
        double k = 1.0 / (1.0 + gamma*x);
```

```
        if (x>=0.0) {
                return 1.0-((((a5*k+a4)*k+a3)*k+a2)*k+a1)*k*
                exp(-x*x/2.0)/sqrt(1.0*pi);
        }
        else {
                return 1.0- N(-x);
        }
}

double EurCall::d_plus(double S0, double sigma, double r) {
        return (log(S0/K)+(r+0.5*pow(sigma, 2.0))*T)/
        (sigma*sqrt(T));
}

double EurCall::d_minus(double S0, double sigma, double r) {
        return d_plus(S0, sigma, r) - sigma * sqrt(T);
}

double EurCall::PriceByBSFormula(double S0, double sigma,
        double r) {
        return S0 * N(d_plus(S0,sigma,r)) - K * exp(-r*T)*
        N(d_minus(S0, sigma, r));
}

double EurCall::VegaByBSFormula(double S0, double sigma,
        double r) {
        double pi = 4.0*atan(1.0);
        return S0 * exp(-d_plus(S0, sigma, r)*d_plus(S0,
        sigma, r)/2)*sqrt(T) / sqrt(2.0*pi);
}
```

- Comment on Listing 14:

This Listing contains the code of the public member functions defined in class *EurCall*. The Black-Scholes formula for a Eurpean call demonstrated in 3.11. is used. Notice that $VegaByBSFormula$ is the derivative with respect to σ of the formula in 3.11.. Another remarkable fact of these lines of code is that the accumulated normal distribution function Φ is approximated by means of a rational function and an exponential function so that $\Phi(x)$ for any $x \in \mathbb{R}$ can be executed quickly.

Listing 15: Solver03.h

```
#ifndef Solver03_h
#define Solver03_h

template<typename Function>double SolveByBisect(Function* Fct,
        double Tgt, double LEnd, double REnd, double Acc) {
        double left = LEnd, right = REnd, mid = (left + right) / 2;
        double y_left = Fct->Value(left) - Tgt,
        y_mid = Fct->Value(mid) - Tgt;

        while (mid - left > Acc) {
```

```
            if ((y_left > 0 && y_mid > 0) || (y_left < 0
                    && y_mid < 0)) {
                    left = mid;
                    y_left = y_mid;
            }
            else {
                    right = mid;
            }
            mid = (left + right) / 2;
            y_mid = Fct->Value(mid) - Tgt;
        }
        return mid;
}

template<typename Function> double SolveByNR(Function* Fct,
        double Tgt, double Guess, double Acc) {
        double x_prev = Guess;
        double x_next = x_prev - (Fct->Value(x_prev) - Tgt) /
        Fct->Deriv(x_prev);

        while (x_next - x_prev > Acc || x_prev - x_next > Acc) {
                x_prev = x_next;
                x_next = x_prev - (Fct->Value(x_prev) - Tgt) /
                Fct->Deriv(x_prev);
        }
        return x_next;
}

#endif
```

- Comment on Listing 15:

This Listing implements the Bisection method and the Newton-Raphson method. The attentive reader will have noticed that these methods have been implemented in the form of templates. Because each option has a different Black-Scholes formula (at least those options that do have a Black-Scholes formula), therefore each different option has a different derivative with respect to σ too. The fact that these two functions are different for each option establishes the grounds for constructing Bisection and Newton-Raphson lines of code that do not depend on those functions. Thus, the use of templates solves that problem.

Further, *the compiler can decide on how to compile the template functions by looking at the first parameter passed to SolveByBisect() and SolveByNR. If* $MyF1$ *is passed, it substitutes class* $F1$ *for the parameter Function when compiling the code. When* $MyF2$ *is passed, then it substitutes* $F2$*, and compiles another version of the code. There will be two different versions of SolveByBisect(), one to work with class* $F1$ *and one with class* $F2$*, as well as two versions of* $SolveByNR()$ *in the compiled code. If more functions were involved, each with its own class, there would be even more versions of SolveByBisect() and* $SolveByNR()$ *in the compiled code. This can result in long compile times and large .exe files. It may or may not* *be a price worth paying for gains in the speed of computation at run time*[9].

Listing 16: Main18.cpp

```
#include "EurCall.h"
#include "Solver03.h"
#include <iostream>
using namespace std;

class Intermediary: public EurCall{
private:
        double S0, r;
public:
        Intermediary(double S0_, double r_, double T_,
        double K_) :EurCall(T_, K_) { S0 = S0_; r = r_; }
        double Value(double sigma){
                return PriceByBSFormula(S0, sigma, r);
        }
        double Deriv(double sigma) {
                return VegaByBSFormula(S0, sigma, r);
        }
};

int main()
{
        double S0 = 100.0;
        double r = 0.1;
        double T = 1.0;
        double K = 100.0;
        Intermediary Call(S0,r, T, K);

        double Acc = 0.001;
        double LEnd = 0.01, REnd = 1.0;
        double Tgt = 12.56;
        cout << "Implied volatility by bisect:" <<
        SolveByBisect(&Call, Tgt, LEnd, REnd, Acc)
        << endl;
        double Guess = 0.23;
        cout << "Implied volatility by Newton-Raphson: "
                << SolveByNR(&Call, Tgt, Guess, Acc) << endl;

        return 0;
}
```

- Comment on Listing 16:

The function *main*() simply declares the inputs of the Black-Scholes formula and calls the other functions in the program with the objective of computing the implied volatility.

A remark must be made in these lines of code: a class called *Intermediary* is defined. *Such class serves as an intermediary between the EurCall class and*

[9] Quoted from [5].

the solvers, translating $PriceByBSFormula()$ and $VegaByBSFormula()$ into the $Value()$ and $Deriv()$ functions, which the solvers can understand[10].

8.1.3 Pricing path dependent options by means of the Black-Scholes model and Monte Carlo simulation

In this section we show how Path Dependent Options are priced using the Black-Scholes model. We will explain the procedure that makes it possible: Monte Carlo Simulation.

Listing 17: BSModel01.h

```
#include <ctime>
#pragma once

#ifndef BSModel01_h
#define BSModel01_h

using namespace std;
#include <vector>

typedef vector<double> SamplePath;

class BSModel{
public:
        double S0, r, sigma;
        BSModel(double S0_, double r_, double sigma_) {
                /*construction function*/
                S0 = S0_;
                r = r_;
                sigma = sigma_;
                srand((unsigned int)time(NULL));
        }
        void GenerateSamplePath(double T, int m, SamplePath& S);
};

#endif
```

Listing 18: BSModel01.cpp

```
#include "BSModel01.h"
#include <cmath>

const double pi = 4.0*atan(1.0);

double Gauss() {
        double U1 = (rand()+1.0) / (RAND_MAX + 1.0);
        double U2 = (rand() + 1.0) / (RAND_MAX + 1.0);

        return sqrt(-2.0*log(U1))*cos(2.0*pi*U2);
}

void BSModel::GenerateSamplePath(double T, int m,
```

[10]Quoted from [5]

```
        SamplePath& S) {
        double St = S0;
        for (int k = 0; k < m;k++) {
                S[k] = St * exp((r-sigma*sigma*0.5)*
                (T/m)+sigma*sqrt(T/m)*Gauss());
                St = S[k];
        }
}
```

- Comment on Listings 17 and 18:

Consider a money market account that continuously compounds interest rate $r \in \mathbb{R}$:

$$A(t) = e^{rt}$$

where $t \geqslant 0$ denotes time.

Consider a risky asset whose price evolves under the Black-Scholes hypotheses:

$$S(t, \omega) = S_0 e^{rt - \frac{\sigma^2}{2} t + \sigma W(t, \omega)}$$

where $\sigma \in \mathbb{R}$ stands for the volatility and $W(t, \omega)$ is a Wiener process under a measure of probability that is risk-neutral[11].

Because $W(t, \omega) - W(s, \omega) \sim N(0, t-s)$ for any $t \geqslant s$, we can write

$$S(t_k, \omega) = S(t_{k-1}, \omega) e^{\left(r - \frac{\sigma^2}{2}\right)(t_k - t_{k-1}) + \sigma \sqrt{t_k - t_{k-1}} Z_k}$$

for $t_k = \frac{k}{m}$ and $k = 1, ..., m$ and where $Z_1, ..., Z_m$ are copies of $Z \sim N(0,1)$.

Definition 8.1. *Let $\hat{Z}_1, ..., \hat{Z}_m$ be a sequence of independent samples of $Z_1, ..., Z_m$ respectively.*

An ordered sequence $\left(\hat{S}_{t_1}, ..., \hat{S}_{t_m}\right)$ is called a sample path if

$$\hat{S}_{t_1} = S_0 e^{rt_1 - \frac{\sigma^2}{2} t_1 + \sigma \sqrt{t_1} \hat{Z}_1}$$

and

$$\hat{S}_{t_k} = \hat{S}_{t_{k-1}} e^{r(t_k - t_{k-1}) - \frac{\sigma^2}{2}(t_k - t_{k-1}) + \sigma \sqrt{t_k - t_{k-1}} \hat{Z}_k}, \text{ for } k = 2, ..., m$$

The following theorem allows to generate a random number with distribution $N(0,1)$:

Theorem 8.2. *If U_1, U_2 are independent random variables with uniform distribution on an interval $[0,1]$, then the random variable*

$$Z = \sqrt{-2 ln \left(U_1\right)} cos \left(2 \pi U_2\right)$$

has distribution $N(0,1)$.

[11] All the theory of discounted prices being martingales in the CRR model discussed in Chapter 2 works analogously for continuous price processes such as $S(t, \omega)$. The basic idea is that a risk-neutral probability measure can be built in continuous time so that μ in Chapter 5 satisfies $\mu = r$. For further discussion on that topic the interested reader may refer to [3].

Proof. A proof of this theorem can be found in [1]. □

The former theorem receives the name of *Box-Muller method*. Listing 18 uses this theorem: builds random integers k_1 and k_2 with function $rand()$, resizes them so that they lie in the interval $[0, 1]$ and then the Box-Muller method is computed. That method is repeated m-times in order to obtain $\hat{Z}_1, ..., \hat{Z}_m$, with which the sequence $\hat{S}_{t_1}, ..., \hat{S}_{t_m}$ is computed

Listing 19: PathDepOption01.h

```
#ifndef PathDepOption01_h
#define PathDepOption01_h

#include "BSModel01.h"

 class PathDepOption {
 public:
         double T;
         int m;
         double PriceByMC(BSModel Model, long N);
         virtual double Payoff(SamplePath& S) = 0;
 };

 class ArthmAsianCall : public PathDepOption {
 public:
         double K;
         ArthmAsianCall(double T_, double K_, int m_) {
                 T = T_;
                 K = K_;
                 m = m_;
         }
         double Payoff(SamplePath& S);
 };

#endif
```

Listing 20: PathDepOption01.cpp

```
#include "PathDepOption01.h"
#include <cmath>

double PathDepOption::PriceByMC(BSModel Model,
         long N) {
         double H = 0.0;
         SamplePath S(m);
         for (long i = 0; i < N; i++) {
                 Model.GenerateSamplePath(T,m, S);
                 H = (i*H+Payoff(S)) / (i+1.0);
         }

         return exp(-Model.r*T)*H;
}

double ArthmAsianCall::Payoff(SamplePath& S) {
         double Ave = 0.0;
         for (int k = 0; k < m;k++) {
```

```
            Ave = (k*Ave+S[k]) / (k+1.0);
        }
        if (Ave<K) {
            return 0.0;
        }

        return Ave - K;
}
```

- Comment on Listings 19 and 20:

Definition 8.3. *A path dependent option is a financial derivative with payoff H at expiry date T such that:*

$$\begin{aligned} h\colon \mathbb{R}^m &\to \mathbb{R} \\ (S_{t_1}, ..., S_{t_m}) &\mapsto h\left(S_{t_1}, ..., S_{t_m}\right) = H(T) \end{aligned}$$

A typical example of path dependent option is the arithmetic Asian call, whose payoff function is:

$$h\left(S_{t_1}, ..., S_{t_m}\right) = \left(\frac{1}{m}\sum_{i=1}^{m} S_{ti} - K\right)^+$$

This is the kind of option that Listing 19 treats as *ArthAsianCall*, a subclass of a class *PathDepOption*.

Listing 21: Main19.cpp

```
#include "PathDepOption01.h"
#include <iostream>
using namespace std;

int main()
{
        double S0 = 100.0, r = 0.03, sigma = 0.2;
        BSModel Model(S0, r, sigma);

        double T = 1.0 / 12.0, K = 100.0;
        int m = 30;
        ArthmAsianCall Option(T, K, m);

        long N = 30000;
        cout << "Asian Call Price  = " <<
        Option.PriceByMC(Model, N) << endl;

        return 0;
}
```

- Comment on Listing 21:

We say that

$$H(0) = e^{-rT}\mathbb{E}\left(H(T)\right)$$

is the price of a path dependent option, where $\mathbb{E}$ stands for the expectancy under the risk-neutral probability.

Let $\left(\hat{S}^i_{t_1}, ..., \hat{S}^i_{t_m}\right)$, for $i \in \mathbb{N}$, be a sequence of independent sample paths. By the law of large numbers:

$$\mathbb{E}\left(h\left(S_{t_1}, ..., S_{t_m}\right)\right) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} h\left(\hat{S}^i_{t_1}, ..., \hat{S}^i_{t_m}\right)$$

Listing 21 gives a very large number for N, 30000 indeed, with which the price of the Arithmetic Asian Call is computed.

References

[1] Baldi, P.: *Stochastic Calculus*, Springer, Italy, 2017.

[2] Bilingsley, P.: *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.

[3] Capiński, M.; Kopp, E.: *Discrete Models of Financial Markets*, Cambridge University Press, United Kingdom, 2012.

[4] Capiński, M.; Kopp, E.: *The Black-Scholes Model*, Cambridge University Press, United Kingdom, 2012.

[5] Capiński, M.; Kopp, E.; Traple, J.: *Stochastic Calculus for Finance*, Cambridge University Press, United Kingdom, 2012.

[6] Capiński, M. J.; Zastawniak, T.: *Numerical Methods in Finance with C++*, Cambridge University Press, United Kingdom, 2012.

[7] CFA Institute: *Corporate Finance and Portfolio Management*, Wiley, Charlottesville, 2018.

[8] CFA Institute: *Derivatives and Alternative Investments*, Wiley, Charlottesville, 2018.

[9] Hirsa, A.; Neftci, S.N.: *An Introduction to the Mathematics of Financial Derivatives*, Academic Press, India, 2014.

[10] Hull, J.C.: *Options, Futures and other Derivatives*, Prentice Hall, Toronto, 2012.

[11] Kat, H. M.: *Structured Equity Derivatives*, Wiley Finance, United Kingdom, 2001.

[12] Lamberton, D., Lapeyre, B.: *Introduction to Stochastic Calculus Applied to Finance*, Chapman Hall, United Kingdom, 2008.

[13] Rudin, W.: *Análisis Real y Complejo*, Editorial Alhambra, Madrid, 1979.

[14] Vives, J.: *Modelització*, Universitat de Barcelona, Barcelona, 2018.