BACHELOR'S DEGREE THESIS

Degree in Mathematics

Title: Brownian motion: a random walk approximation

Author: Sergio Gutiérrez Moya

Advisor: Oriol Serra Albó

Department: Mathematics

Academic year: 2017-2018

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya
Facultat de Matemàtiques i Estadística

Degree in Mathematics

Bachelor's Degree Thesis

Brownian motion: a random walk approximation

Sergio Gutiérrez Moya

Supervised by Oriol Serra Albó

June, 2018

Thanks to my family and friends, who have supported me along this amazing venture. I wish to express my sincere thanks and indebtedness to Dr. Oriol Serra for supervising and leading me throughout this work. I also place on record my gratitude to all Faculty professors for providing me with valuable guidance and encouragement.

Abstract

Brownian motion is one of the most used stochastic models in applications to financial mathematics, communications, engineeering, physics and other areas. Many of the central results in the theory are obtained directly from its definition as a continuous process. As a mathematical object, Brownian motion also have some special and important properties that make it fundamental to understand related mathematical fields and state-of-the-art concepts.

The purpose of this work is to review a relatively recent approach which allows to reobtain these results via a random walks approximation. Brownian motion is the stochastic limit of suitably nested random walks, but some technical details are need to be checked in order to guarantee the convergence. The applications of this particular approach include the local time of Brownian motion and the Black-Scholes model in financial mathematics.

Keywords

Brownian motion, Wiener process, Random Walk, local time, Black-Scholes model, stochastic process, Markov chain, probability theory.

Contents

1. Introduction

Universe is sorrounded by a wide range of phenomena to which mathematical models provide a valuable tool of analysis and description. A mathematical model, in order to be useful, must satisfy two basic principles: *accuracy* and *simplicity*. It is clear that any model should be precise when describing a certain event, otherwise conclusions coming from it would be worthless. Simplicity is more subtile. Recalling *Occam's razor principle*, models should not make unnecessary assumptions. Provided a model is accurate, it comes useful to choose the less sophisticated, so that analyzing it in depth results into an easier procedure.

This work is devoted to the study of Brownian motion, the most globally spread stochastic model in the study of random phenomena. Its strength comes from the two features mentioned above, which Brownian motion widely satisfies. The simplicity of the Brownian motion arises from its background discrete model of random walks, based on simple random independent binary inputs. On the other hand, Brownian motion is based on the normal distribution, which through the Central Limit Theorem, enjoys a universal character describing the addition of a large number of independent random inputs.

This establishes the main motivation of this work. What is the essence that makes Brownian motion so special? Which are the components in its core that transform it into 'the model' for randomness? These questions will serve us as a starting point with the objective of cracking Brownian motion. After that, we will have a better understanding of the general ideas hidden behind it, which are fundamental to use in those models which are of our interest. However, Brownian motion itself would not be that relevant if it were not for its numerous applications in many fields. We also include a discussion on some of these applications.

From a global scale, this work intends to give a general view of Brownian motion. Firstly, a solid background is set: definitions, properties and results about regularity and its ordered random structure. Next, a construction of the Brownian motion through random walks will be given. Brownian motion is often described as a stochastic limit of random walks. Actually, it is so, but it is not that simple. Just a sequence of random walks is not enough to obtain a Brownian motion, but a suitably nested sequence of random walks that inherit its correlation along the sequence has to be defined in a delicate construction. This is a hard task that is completely detailed in this work. With little details left, an exhaustive proof of the convergence to Brownian motion is given. Our milestone is to produce a mathematically robust work, hence it is essential to add all details of the construction. We come up with the desired conclusion, which is the result used in the forthcoming applications.

The relevance of Brownian motion becomes aparent when looking at any text on random processes. Probably, when it was firstly introduced by the botanist Robert Brown in 1824, nobody was conscious about the impact it was going to have on the coming decades and centuries as a model in Statistical Physics, Mathematical Finances, Electrical Engineering and as a central object in Probability Theory, Statistics, Geometry among many other areas in mathematics and applications. It was one of the topics treated by Albert Einstein which eventually laid empirical foundations of the corpuscular nature of matter in Theoretical Physics, and as ealrly as 1900 it was already considered by Louis Bachelier as a model for the stock market evolution. Both directions have given substantial developments up to our days.

This work is organized as follows. We begin by giving a short general description and classification of stochastic processes, so that we can see fundamental properties of Brownian motion that also apply to other stochastic processes. This is the object of chapter 2. Following this chapter, special emphasis is given to Markov chains for two reasons: the particular vocabulary that appears recurrently throughout this thesis with which we need to be familiar with, and the crucial fact that Brownian motion and random walks are Markov processes, thus we need a good understanding of them.

After this introductory material, we are ready to thrive to the most classical view of Brownian motion: the probability approach. In chapter 4 the abstract probabilistic definition of Brownian motion is presented and some basic properties are proved. The existence of Brownian motion from its formal definition is backed up by the *Wiener's theorem*. The proof of this theorem uses an important technique which appears recurrently in this work. Surprisingly, some other important results are straightforward, such as the scaling invariance and the non-differentiability. As a Markov chain, Brownian motion has some special properties that makes it even more particular. They are studied in chapter 5. The most breathtaking is the reflection principle, not only by its own interest, but also because of its direct consequences.

Since its introduction at the beginning of the 19^{th} century, the feature which has given Brownian motion a greater entity is its several applications in science. From pure mathematics to natural sciences, Brownian motion has been a key piece in research throughout the time, allowing groundbreaking advances that make science progress. In this thesis two applications will be visited: Brownian local time and the Black-Scholes model.

We aim at replicating these results not only through a classical approach, but also using embedded random walks. To do so, the *twist and shrink* construction of Tamas Tsabados will be followed. Actually, chapter 6 is devoted to check that there is convergence to Brownian motion of random walks defined in this particular way. It contains the core of this work, the way Brownian motion can be approached through nested random walks. More precisely, some accurate bounds will be given as illustration. These inequalties will be instrumental to discuss later approximations of this approach. Many technical details are necessary to give an accurate proof. The lemmas stated lead to the a final theorem that asserts the desired convergence.

As an almost surely, nondifferentiable stochastic process, Brownian motion produces very irregular paths, in the sense that changes of value happen suddenly as the result of a random law. Thus, it could be of interest to analyze the time distribution that the process spends at a given level, from a probability point of view. This gives rise to *Brownian local time*. We approximate it via the classical results, the *Trotter's theorem*, and also by using the *twist and shrink* construction. The later is followed from the article [10], but some details are ommited due to its technical complexity and extension. The classical proof is based on the idea that underpins the proof of the *Ray-Knight* theorem, which is not included in this work.

Lastly, chapter 8 turns around the Black-Scholes model. It was a significant breakthrough in financial mathematics, as finally there was a unique formula to price some financial claims, and even for some of them, explicit formulas exist. First, a brief introduction on changes of measure is given, for reasons concerning risk-neutral pricing, in which much of financial options theory is based. Also, we state the fundamentals of the model in order to set a framework of reference. The approach used to retrieve the results is based on [11]. It is relatively straight to find these formulas using simple probability arguments and the *twist and shrink* embedding of random walks.

This work is based on the recent monography by Peters Morter and Yuval Peres entitled *Brownian motion* [8] and on a series of papers by Tamás Szabados and his collaborators [9, 10, 11], which were on the background motivation of this work. On one side, a thorough understanding of Brownian motion in the probabilsitic framework, and on the other side the detailed description of Brownian motion from the perspective of Random Walks for a better understanding of its nature and some of its applications.

2. Overview of Stochastic Processes

Consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and a random variable $X : \Omega \longrightarrow \mathbb{R}$. Ω, the set of possible successes, which a probability of occurrence is determined on $\mathbb{R}$ and $\mathcal{A}$ is a σ-algebra. Stochastic processes are the most general extension on $\mathcal{A}$ of random variables. Instead of a random distribution for the variable, an additional index set I is introduced, so that the distribution evolves over time. Another level of difficulty are random vectors. Now, events are not identified by numbers but a vector. Probability distributions (joint, marginals, conditionals) play a more important role when we try to describe randomness. Stochastic processes appear to give a infiniteness version of vectors extended to the whole real line, or a subset of it I depending on the case. They can be seen from two different perspectives, as they have two arguments: the state-space elements $\omega \in \Omega$ and a temporal parameter t. Hence two rather different interpretations can be done: fix t_0, observe how the state-space in configured at that time, or fixed $\omega \in \Omega$, a path is generated *path* along the time. The latter interpretation will be of essential importance for us in some sections, as we will be interested in the analysis of continuity and differentiation of *Brownian paths*. But, it is possible to find a connection of this two visions of a stochastic process in some cases, which is given by *Ergodic theorems*. They are a series of theorems, with applications in many mathematicals branches, in particular in probability theory that allow us to study the assymptotic properties of some random processes, such as Makov chains.

Also, it will be fundamental to decide or argue if a given process actually can exist and is well-defined. In the case of *Brownian motion* two different approaches will be detailed in this work. One of them is underpinned by the so-called *Wiener's theorem*, which definitely states that Brownian motion exists. The proof of this theorem was vital for the historical development of theory and applications that followed.

A first clear classification of processes is determined by the index parameter T. If $T = (0, 1 \dots)$ then the process X_t is said to be a *discrete-time* stochastic process. If $T = [0, \infty)$, then X_t is called a *continuous time* stochastic process.
Some general properties identify important classes of random processes. They not only describe natural features of random processes, but also allow mathemtical analysis in greater depth. These are the different stochastic processes:

a) *Processes with stationary independent increments*
If the random variables

$$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_n} - X_{t_{n-1}}$$

are independent for all $t_1 < t_2 < \dots < t_n$ then X_t is a process with independent increments (also allowing a first time t_0). Even if restrictive, this condition is satisfied by a wide class of random processes. It allows to obtain all joint distributions from the knowledge of the one for X_t and every t.
If the distribution of increments $X(t+h) - X(t)$ depends only on the length of h and not on the time t, then process is said to have *stationary increments*.
These properties hold for both *Brownian motion* and *Poisson processes*.

b) *Markov Processes*
These processes generalize the idea of Markov chains (see Section 3) for a wider range of state-space. The idea of these processes is that what happens in the future is only determined by the present of not the past, that cannot alter the future whatsoever, given the present. This leads to the *Markov property* of the process. In fact, conditional to an initial distribution of the state-space, the process is uniquely determined by a *transition probability matrix* between states that provides how states evolve in the Markov process. More formally, for each $t_1 < t_2 < ... t_k$

$$Pr(X_t = x | X_{t_1} = x_2, X_{t_2} = x_2 ... X_{t_k} = x_k) = Pr(X_t = x | X_{t_k} = x_k)$$

The two main stochastic process that are considered in this work, random walks and Brownian motion are Markov processes.

c) *Martingales*
The idea of martigales is similar to the one of Markov processes but in this case, instead of a probability idistribution and a transition matrix, the expectation of the process is the defining property for martingales. In other words, let $\{X_t\}$ a real-valued stochastic process. We say that $\{X_t\}$ is a *martingale* if $\mathbb{E}(|X_t|) < \infty$ for all t, and for any $t_1 < t_2 < \cdots < t_k < t_{k+1}$

$$\mathbb{E}(X_{t_{k+1}} | X_{t_1} = x_1, X_{t_2} = x_2, ... X_{t_k} = x_k) = x_k$$

This model is used in gambling games where fair games need to be defined, and after a game the player expects to have the same amount of money of the beggining of the game. They are also used in finance for events forecasting. Again, this property holds for Brownian motion and random walks.

d) *Stationary processes*
The concept of stationarity has very widely meanings, but in stochastic processes' context, we say that a process X_t is *stricly stationary* if the joint distributions

$$(X_{t_1+h}, ... X_{t_n+h}) \text{ and } (X_{t_1}, ... X_{t_n})$$

are equal for all $h > 0$ and for any choice of $t_1, ..., t_n$. A stochastic process X_t is said to be *widely sense stationary* if the process has finite second order moments and if $Cov(X_t, X_{t+h}) = E(X_t X_{t+h}) - E(X_t)E(X_{t+h})$ depends only on h for all t. A stationary process is *stationary in wide sense*, but the converse is not necessary so, it is a wider assumption (only at expectation level).

In addition to this processes, there are also *renewal processes* and *point processes*. They do concern the contents of this work, therefore no details about them are given. In interested, more information can be found in [6]. These are the main types of stochastic processes. As radom walks and Brownian motion are included in some types of the processes described, it is easier to extract information about them, or conversely, some properties of them can be generalised.

3. Markov chains

Markov chains are probably, together with Brownian motion, one of the most broadly studied stochastic processes. In this section, some of its most outstanding properties will be studied: from recurrent or transient states of discrete-time Markov chains to the connection of random walks to them through an assymptotic approxiamtion.
Despite this wide range of basic properties, the defining property that all Markov chains share is known as *Markov property*. Getting straight to it, it can be interpreted as a past independence given the present, in other words, for chains only matters the very last step. More formally, consider a sequence of random variables $X_0, X_1, ... X_n$, the property states that:

$$Pr(X_{n+1} = x_{n+1}|X_n = x_n, X_{n-1} = x_{n-1} ... X_0 = x_0) = Pr(X_{n+1} = x_{n+1}|X_n = x_n)$$

Markov chains and Markov processes are used in several applications, such as, weather forecast, simulation, game theory, birth-death processes (Galton-Watson)..., and, more importantly for us, in finance, in order to examine possible situations in which an arbitrage operation can be executed or to determine the volatitlity of prices of an underlying asset.

Although it is not proved in this chapter, Brownian motion is a continuous-time Markov chain, which in fact is also a martingale. Thus, in order to approximate it, we need a discrete process that taken to the limit converges to a Brownian motion. Here random walks appear to be exactly what fills the gap. It will be the tool used to approach Brownian motion through a discrete process. This joining process is detailed in chapter 6, where it will be seen that the embedding should be done suitably, so that we get a convergent process.

Markov chains are the vast majority of Markov processes. However, apart from Markov chains, there exist other Markov processes, which have more general state spaces, such as nonnumerable, which are not possible for Markov chains. The *Markov property* still holds for them. In this work, we will only work with discrete time Markov chains, as they will be our interest to see some properties of Brownian motion and random walks. Another type of Markov chains are continuous-time Markov chains. They have very similar properties to the discrete time chains, with some modifications. However, the general theory would be too extense to develop in deepth, and we will focus on the Markov property of Brownian motion in chapter 5, where important consequences will be deduced.

3.1 Discrete-time Markov Chains

Let I be a countable set. We call $i \in I$ the set of all possible states of the Markov chain. Then, I is known as the *state-space*. Consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a *distribution* of probability λ in this space. For a *random variable* $X : \Omega \to I$, set:

$$\lambda_i = \mathbb{P}(X = i) = \mathbb{P}(\omega : X(\omega) = i)$$

Definition 3.1. We say that a matrix $P = (p_{ij}, i, j \in I)$ is *stochastic* if every row $(p_{ij}, j \in I)$ is a probability distribution.

This condition on matrices representing the states transition is natural, as for a given state, a probability distribution will be established among all states, including itself. In this case, the transition to the next step is well-defined.

Definition 3.2. We say that $(X_n)_{n\geq 0}$ taking values in I is a *Markov chain* with initial distribution $\lambda = (\lambda_i)_{i\in I}$ and trasition matrix P if

(1) X_0 has a distribution λ;

(2) for $n \geq 0$, conditional on $X_n = i$, X_{n+1} has a distribution $(P_{ij} : j \in I)$ and is independent of $X_0, X_1, \dots, X_{n-1}$.

If the previous conditions are satisfied, for simplicity, we say that $(X_n)_{n\geq 0}$ is *Markov*(λ, P). Moreover, it is time-homogeneous, since P does not change with time. From the definition of *Markov chains* it is clear that:

Theorem 3.3. *A discrete-time random process* $(X_n)_{0\leq n\leq N}$ *is Markov*(λ, P) *if and only if*

$$\forall i_0, i_1, \dots i_N \in I \qquad \mathbb{P}(X_0 = i, X_1 = i_1, \dots, X_N = i_N) = \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{N-1} i_N} \tag{3.1}$$

A first-stage result on Markov chains is the *weak Markov property*. It is *weak* because later we will see a generalisation of it.

Theorem 3.4. (Weak Markov property). *Let* $(X_n)_{n\geq 0}$ *be Markov*(λ, P). *Then, conditional on* $X_m = i$, $(X_{m+n})_{n\geq 0}$ *is Markov*(δ_i, P) *and it is independent of the random variables* $X_0, X_1, \dots, X_m$.

A Markov chain can be very difficult to understand globally, but sometimes it is possible to split it in differents parts, so that its structure becomes easier to interpret and analize.

Definition 3.5. We say that i *leads to* j (and write $i \to j$) if

$$\mathbb{P}_i(X_n = j \text{ for some } n \geq 0) := \mathbb{P}_i(X_n = j \text{ for some } n \geq 0 | X_0 = i) > 0$$

If $i \to j$ and $j \to i$, then we say that i *communicates* j (write $i \leftrightarrow j$).

From definition and some basic properties, it is easy to check that $\leftrightarrow$ satisfies the conditions of an equivalence relation, whence it partitions I into *communicating classes*. We say that C is a *closed class* if

$$i \in C \quad i \to j, \text{ imply } j \in C$$

In other words, a closed class is a class from where it is not possible to scape. Particularly, if $\{i\}$ is a closed class, the state is said to be absorbing. A chain that has one unique communicating class is called *irreducible*.

3.2 Hitting times, recurrence and transience

Consider a Markov chain $(X_n)_{n\geq 0}$, with state-space I and a subset $A \subset I$.

Definition 3.6. The *hitting time* of a random variable: $H^A(\omega) : \Omega \to \{0, 1, \dots\} \cup \{\infty\}$ is defined as:

$$H^A(\omega) = inf\{n \geq 0 : X_n(\omega) \in A\}$$

The probability of hitting A starting from i is:

$$h_i^A = \mathbb{P}_i(H^A < \infty)$$

In the case A is closed, we say that H_i^A is the absorption probability.

Definition 3.7. A random variable $T : \Omega \to \{0, 1, \dots\} \cup \{\infty\}$ is called a *stopping time* if the event $\{T = n\}$ denpends only on $X_0, X_1, \dots, X_n$.

The most plausible interpretation of this definition is that just by observing the past of the process, it is known when a certain event is going to happen. There are many types of random variables that involve stopping times, but the following will be object of study in this project.

Example 3.8. The *first passage time* random variable:

$$T_j = inf\{n \geq 1 : X_n = j\}$$

is a stopping time. only has to be checked that:

$$\{T_j = n\} = \{X_1 \neq j, \dots, X_{n-1} \neq j, X_n = j\}$$

Now, we provide a stroger form of Theorem 3.4, using stopping times and its properties.

Theorem 3.9. (Strong Markov property) *Let $(X_n)_{n\geq 0}$ be a Markov(λ, P) and let be T a stopping time of $(X_n)_{n\geq 0}$. Then conditional on $T < \infty$ and $X_T = i$, $(X_{T+n})_{n\geq 0}$ is Markov(δ_i, P) and independent of $X_0, X_1, \dots X_T$.*

Actually, we are more interested in a generalisation of Example 3.8, considering the r-th passage time $T_i^{(r)}$ to state i

$$T_i^{(0)}(\omega) = 0, \quad T_i^{(r+1)}(\omega) = inf\{n \geq T_i^{(r)}(\omega) + 1 : X_n = i\}.$$

Definition 3.10. The length of the r-th excursion to i is:

$$S_i^{(r)} = \begin{cases} T_i^{(r)} - T_i^{(r-1)} & \text{if } T_i^{(r-1)} < \infty \\ 0 & \text{otherwise} \end{cases}$$

We denote the number of visits V_i to i by:

$$V_i = \sum_{n=0}^{\infty} 1_{\{X_n = i\}}$$

Then,

$$\mathbb{E}_i(V_i) := \mathbb{E}(V_i | X_0 = i) = \mathbb{E}_i\Big(\sum_{n=0}^{\infty} 1_{\{X_n=i\}}\Big) = \sum_{n=0}^{\infty} \mathbb{E}_i(1_{\{X_n=i\}}) = \sum_{n=0}^{\infty} \mathbb{P}_i(X_n = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)},$$

where the index i denotes the initial state i, and therefore, we are describing the probability (or the mean) of returning to the state $i \in I$, respectively.

We compute the distribution of V_i under $\mathbb{P}_i$ in terms of the return probability:

$$f_i = \mathbb{P}_i(T_i < \infty)$$

Lemma 3.11. *For $r = 0, 1 \dots$, we have $\mathbb{P}(V_i > r) = f_i^r$*

Proof. We prove the result by induction over r. For $r = 0$ it is clearly true. For $r \geq 1$ it hold the $\{V_i > r\} = \{T_i^{(r)} < \infty\}$. If it is true for r, then:

$$\begin{aligned}\mathbb{P}_i(V_i > r+1) = \mathbb{P}_i(T_i^{(r+1)} < \infty) &= \mathbb{P}_i(T_i^{(r)} < \infty \text{ and } S_i^{(r+1)} < \infty) \\ &= \mathbb{P}_i(S_i^{(r+1)} < \infty | T_i^{(r)} < \infty)\mathbb{P}_i(T_i^{(r)} < \infty) \\ &= f_i f_i^r = f_i^{r+1}\end{aligned}$$

□

Definition 3.12. We say that a state i is **recurrent** if:

$$\mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 1$$

Alternatively, we say that a state is **transient** if:

$$\mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 0$$

Theorem 3.13. *The following dichotomy holds:*

(1) *If $\mathbb{P}_i(T_i < \infty) = 1$, then i is recurrent and $\sum_{i=0}^{\infty} p_{ii}^{(n)} = \infty$*

(2) *If $\mathbb{P}_i(T_i < \infty) < 1$, then i is transient and $\sum_{i=0}^{\infty} p_{ii}^{(n)} < \infty$*

Proof. If $\mathbb{P}_i(T_i < \infty) = 1$, then by Lemma 3.11,

$$\mathbb{P}_i(V_i = \infty) = \lim_{r\to\infty} \mathbb{P}_i(V_i > r) = 1.$$

Thus, i is recurrent and

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \mathbb{E}_i(V_i) = \infty.$$

Conversely, if $f_i = \mathbb{P}_i(T_i < \infty) < 1$,

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \mathbb{E}_i(V_i) = \sum_{r=0}^{\infty} \mathbb{P}_i(V_i > r) = \sum_{r=0}^{\infty} f_i^r = \frac{1}{1 - f_i} < \infty.$$

□

But for more generalisation in terms of the state-space, the preceding dichotomy can be extended to a class property. In other words:

Lemma 3.14. *Let C be a communicating class. Then either all states in C are recurrent or transient.*

Proof. Take any pair of states i, j such that i is transient. Then, there exist $m, n \geq 0$ such that $p_{ij}^{(n)}$ and $p_{ji}^{(m)}$, then, for all $r \geq 0$,

$$p_{ii}^{(n+r+m)} \geq p_{ij}^{(n)} p_{jj}^{(r)} p_{ji}^{(m)}.$$

Thus, if i is transient,

$$\sum_{r=0}^{\infty} p_{jj}^{(n)} \leq \frac{1}{p_{ij}^{(n)} p_{ji}^{(m)}} \sum_{r=0}^{\infty} p_{ii}^{(n+r+m)} < \infty.$$

Hence, j is transient, as required. □

3.3 Random Walks

Random walks are one type of Markov chains that require especial attention in this work due to their close connection to Brownian motion. They are also used in many applications in science and technology, embracing the model of a particle moving in a gas or to study the dynamic of a population, entre others. The first example for instance, is surprising as the reader could think of the movement of a particle as a continuous motion. Actually, it is, but in models and simulations continuous data and time intervals are impossible to get and there random walks appear as a good approximation to brownian motion. In principle, the motion could be modelled using other ideas and approximations, but as we will see later, brownian motion can be retrieved from a random walk taken to the limit of the discretization of $\mathbb{Z}^n$ grid with an appropiate scale.

A particular reason why they shoud be carefully studied is because random walks have a non-finite state space, what means that though its is irreducible and closed, it may be nonrecurrent. However, combinatorail arguments will be given to solve the problem of recurrence of random walks in low dimensions.

Remark 3.15. **(Stirling's formula)** For n sufficiently large, $n! \approx \sqrt{2\pi n}(\frac{n}{e})^n$, where $\approx$ denotes an assymptotic equivalence of both expressions.

For example, if a one-dimensional random walk is considered, the probability of return after an odd number of steps is zero. For even number of steps, the probability of return is:

$$p_{00}^{(2n)} = \binom{2n}{n} p^n q^n \approx \frac{(4pq)^n}{\sqrt{n\pi}}.$$

In the symetric case, $p = q = \frac{1}{2}$, $pq = \frac{1}{4}$ and thus, for some N and all $n \geq N$:

$$p_{00}^{(2n)} \geq \frac{1}{2\sqrt{2\pi n}} \approx \frac{1}{\sqrt{n\pi}}.$$

In consequence, we have:

$$\sum_{n=N}^{\infty} p_{00}^{(2n)} \geq \frac{1}{\sqrt{\pi}} \sum_{n=N}^{\infty} \frac{1}{\sqrt{n}} = \infty$$

So the symetric one-dimensional random walk is recurrent. In the non-symetric case, $4pq = r < 1$, and for some N:

$$\sum_{n=N}^{\infty} p_{00}^{(2n)} \leq \frac{1}{\sqrt{2\pi}} \sum_{n=N}^{\infty} r^n < \infty,$$

what shows that the one-dimensional random walk is transient, and by symmetry it also holds for $\sum_{n=N}^{\infty} p_{ii}^{(2n)}$ for any state i. In dimension two, this property still holds, due to the fact that the harmonic series is divergent, but from dimension three on, the series analysed are convergent and hence, the random walk comes back to the origin infinitely many times with zero probability.
In oder to connect random walks with discrte time Markov chains it is useful to introduce the so-called *Chapman-Kolmogorov* equations. These equations describe the relation between states in a Markov chain via the transition probabilities that uniquely determine the chain, up to an initial distribution. *Chapman-Kolmogorov equations are*:

$$p_{ij} = \sum_{k \in I} p_{ik} p_{kj}$$

These equations are not of interest in this work, but for example, they are useful to derive the partial differential equation that a Brownian satisfies.

4. Introduction to Brownian motion

Brownian motion has its origins in the $19^{th.}$ century, when the british scientist R. Brown discovered it by studying the erratic movement of particles on the surface of a liquid. The way he described the motion of particles was named after R. Brown, and was called *Brownian motion*. However, the concept today has been devolvoped and many generalizations and variations have been introduced. Actually, Brownian motion is widely used in many fields, and its construction and definition gives it some particular properties that make it appropiate for some applications. Usually it appears when in the process handled some uncertainty takes place. Its utility comes from the fact that it has no bias and, therefore, it faithfully simulates nature phenomena that happen in real life.

Suppose the motion of a particle is to be analysed. On the macroscopic level it is not clear at all which will be the path described by the particle at all. But, some inferences can be done when microscopic level is studied. At this level, we see that at any step of it, the particle undergoes a *discrete* path which is modeled by a random walk. Say S_n is the random walk started at X_0. Every step is a random variable that takes the value $X_k = \pm 1$ with equal probability, and then $S_n = S_0 + \sum_{i=1}^{n} X_i$.

However, some results are surprising, because not all features of the microscopic view will have an effect on the macroscopic view. In other words, it is possible to obtain similar path for random walks that share mean and covariance matrices. That makes Brownian motion a very general object, which will link different processes.

Physical interpretation aside, the objective of this section will be making a first approach to its most famous construction and visiting basic properties that will be the seed for the following developments, where more specific topics will be discussed. This will be a fist overlook of the macroscopic figure generated by the process.

4.1 A joint probabilities approach. The Wiener process

Brownian motion is usually defined by a process that must satisfy some conditions. This definition is very versatile, as it allows to study it using arguments underpinned by the normal distribution. Since this distribution is well-known, it requires a softer introduction to acquire a first idea of the nature of this particular process. However, it must be ensured that the definition gives raise to the process desired. This is given by the *Wiener's theorem*. After this is done, the random walk approach is more sensitive, provided the local behaviour of the process.

Definition 4.1. A stochastic process $\{X(t) : t \geq 0\}$ is a *Brownian motion* if:

i) Paths are almost surely continuous for all $t \geq 0$.

ii) Every increment $X(t+s) - X(s)$ is normally distributed with mean 0 and variance t.

iii) for any $t_1 < t_2 < t_3 < t_4$, the increments: $X(t_2) - X(t_1)$ and $X(t_4) - X(t_3)$ are independent with the distributions specified in *ii*).

In addition, if $X(0) = 0$ then we say that the process is a *standard* Brownian motion.

This definition has many important consequences that are followed almost inmediately from it. The first remarkable fact is that Brownian motion satisfies the Markov property, as the distribution of the increments only depends on the time elapsed independently of the values taken by the random variable. More formally,

$$Pr(X(t) \leq x | X(t_k) \leq x_k, ... X(t_0) \leq x_0) = Pr(X(t) \leq x | X(t_k) \leq x_k).$$

It is particularly interesting to compute easily the previous probability and the joint probability of the increments from the definition of Brownian motion.

$$p(x - x_0, t - t_0) := Pr(X(t) \leq x | X(t_0) \leq x_0) = Pr(X(t) - X(0) \leq x - x_0) \\ = \frac{1}{\sqrt{2\pi(t - t_0)}} \int_{-\infty}^{x - x_0} e^{-\frac{y^2}{2(t-t_0)}} dy$$

Now recall that, from a very basic convolution product, the sum of normal random variables is also normal, so the last part of the definition as, for any $t_1 < t_2 < t_3$:

$$X(t_3) - X(t_1) = (X(t_3) - X(t_2)) + (X(t_2) - X(t_1))$$

Applying recurrently the Markov property and the independence assumed by hypothesis, using the same notation it holds that the joint probability distribution of increments in terms of the normal density function is:

$$f(x_1, ... x_n) = p(x_1, t_1) p(x_2 - x_1, t_2 - t_1) \cdots p(x_n - x_{n-1}, t_n - t_{n-1})$$

4.2 Wiener's theorem

The next step is to show that a process defined as above exists, that is, the conditions given are compatible and define a unique process. Previously it is convinient to proof this result:

Lemma 4.2. *Let $X \geq 0$ be a non-negative random variable and $p > 0$, then:*

$$\mathbb{E}(X^p) = \int_0^\infty p x^{p-1} \mathbb{P}(X > x) dx$$

Proof. This is a generalization of the result already known for $p = 1$ and the proof is mainly the same, with some modifications using the Fubini's theorem.

$$\mathbb{E}(X^p) = \int_0^\infty x^p f(x) dx = \int_0^\infty \int_0^x p t^{p-1} f(x) dt dx = \\ \int_0^\infty \int_t^\infty p t^{p-1} f(x) dx dt = \int_0^\infty p t^{p-1} \mathbb{P}(X > t) dt$$

□

Reviewing again the definiton of the process, it seems that the conditions imposed to marginals densities of the process can lead to a contradiction or may generate discontinuous paths. This is the main concern of the following theorem. In order to prove it, we will construct Brownian motion as a uniform limit of continuous functions, producing a continuous limit process. Mainly, dyadic intervals will be used to generate the distributions imposed, and then this properties will be extended to any interval.

Theorem 4.3. (Wiener's theorem) *Brownian motion exists.*

Proof. Consider the set of $D = \cup D_N$, with D_N defined as the set of integer multiples of 2^{-N} in $[0, \infty)$ for $N = 0, 1, 2, \ldots$. The steps to follow to complete this proof are delicate, from point of view of mathematical analysis. First we will construct a Brownian motion for some set of indeces and then the difficult task will be how to extend it to the semi-real line continuously and see that, the extension is a Brownian motion.
For any $t \in D$, consider an independent Gaussian random variable Y_t with mean 0 and varaince 1. For $t \in D_0 = \mathbb{Z}^+$ set:

$$B_t = \sum_{i=0}^{t} Y_i$$

then, B_t is a Brownian motion indexed by D_0. Now, by induction we should see that this process can be extended for all other indexes D_N as a Brownian motion. Suppose that $(B_t : t \in D_0)$ is brownian motion for D_{N-1}. For $t \in D_N \backslash D_{N-1}$ set $r = t - 2^{-N}$ and $s = t + 2^{-N} \Rightarrow r, s \in D_{N-1}$
Consider

$$Z_t = 2^{-\frac{N+1}{2}} Y_t$$

$$B_t = \frac{1}{2}(B_r + B_s) + Z_t$$

The new increments obtained are:

$$B_t - B_s = \frac{1}{2}(B_s - B_r) + Z_t$$

$$B_t - B_s = \frac{1}{2}(B_s - B_r) - Z_t$$

Then with a little compututation we get:

$$\mathbb{E}[(B_t - B_r)^2] = \mathbb{E}[(B_s - B_t)^2] = \frac{1}{4}2^{-(N-1)} + 2^{-(N+1)} = 2^{-N}$$

$$\mathbb{E}[(B_t - B_r)(B_s - B_t)] = \frac{1}{4}2^{-(N-1)} - 2^{-(N+1)} = 0$$

Thus, the increments are independent and they have the variance required for a Brownian motion. Hence, $(B_t : t \in D_N)$ is a Brownian motion for all N, and in consequence, $(B_t : t \in D)$ is also a Brownian motion.
Now, for each N, denote by $(B_t^{(N)})_{t \geq 0}$ the continuous process obtained by linear interpolation of $(B_t : t \in D_N)$. Also, $Z_t^{(N)} = B_t^{(N)} - B_t^{(N-1)}$. It is clear, by construction that $Z_t^{(N)} = 0$ for $t \in D_N \backslash D_{N-1}$ and

$$Z_t^{(N)} = B_t - \frac{1}{2}(B_{t-2^{-N}} + B_{t+2^{-N}}) = Z_t = 2^{-\frac{(N+1)}{2}} Y_t$$

Set $M_t = \sup\limits_{t \in [0,1]} |Z_t^{(N)}| = \sup\limits_{t \in D_N \backslash D_{N-1} \cap [0,1]} 2^{-\frac{(N+1)}{2}} |Y_t|$
Note that there are 2^{N-1} points in $(D_N \backslash D_{N-1} \cap [0,1])$, so for $\lambda > 0$ we have:

$$\mathbb{P}(M_N > \lambda 2^{-\frac{(N+1)}{2}} \leq 2^{N-1} \mathbb{P}(|Y_1| > \lambda)$$

Now, using Lemma 4.2 for the random variable $2^{p\frac{N+1}{2}} \mathbb{E}(M_N^p)$, we have:

$$2^{p\frac{N+1}{2}} \mathbb{E}(M_N^p) = \int_0^\infty px^{p-1} \mathbb{P}(2^{\frac{(N+1)}{2}} M_N > x) dx \leq 2^{N-1} \int_0^\infty px^{p-1} \mathbb{P}(|Y_1| > x) = 2^{N-1} \mathbb{E}(|Y_1|^p).$$

Thus, for any $p > 2$, recalling *Hölder's inequalty*:

$$\mathbb{E}\Big(\sum_{N=0}^{\infty} M_N\Big) = \sum_{N=0}^{\infty} \mathbb{E}(M_N) \leq \sum_{N=0}^{\infty} \mathbb{E}(M_N^p)^{\frac{1}{p}} \leq \mathbb{E}(|Y_1|^p)^{\frac{1}{p}} \sum_{N=0}^{\infty} (2^{\frac{p-2}{2p}})^{-N} < \infty.$$

Using this result, it follows from *Borel-Cantelli*'s lemmas that with probability 1, as $N \to \infty$,

$$B_t^{(N)} = B_t^{(0)} + Z_t^{(1)} + \cdots + Z_t^{(N)}$$

converges uniformly in $t \in [0, 1]$, and by extension, for any bounded closed interval. Therefore, $(B_t : t \in D)$ has a continuous extension $(B_t)_{t \geq 0}$.
The last thing that remains to be proved is that the increments of the random variables built for the process are also independent. As we are in a compact set, for any $0 < t_1 < \cdots < t_n$, we can define sequences $(t_k^m)_{m \in \mathbb{N}}$ such that $t_k^m \to t_k$ for all k. Then, by continuity of the covariance function of these random variables (it is zero since they are Gaussian), it holds that

$$B_{t_1} - B_{t_0}, \dots B_{t_n} - B_{t_{n-1}}$$

are independent, and normally distributed, as required. □

This Theorem shows that Gaussian processes with continuous paths do exist, although nothing else can be said about regularity of the paths, so far. The next step is to study differentiability of Brownian motion. Before, we need some auxiliar results and properties. This will be a direct approach using up and down derivatives. However, in chapter 5 a more straight proof of non-differentiability will be given using the reflection principle.

4.3 Time scaling and time inversion

Before getting deeper in regularity issues of Brownian motion, from the definition it is straightforward to deduce two key properties that will be recurrently used. They can be summerized in the following results:

Lemma 4.4. (Scaling invariance) *Suppose $\{B(t) : t \geq 0\}$ is a standard Brownian motion and let $a > 0$. Then the (scaled) process $\{X(t) : t \geq 0\}$ defined by $X(t) = \frac{1}{a}B(a^2 t)$ is also a standard Brownian motion.*

Proof. The only property that is not obvious to hold is that the distribution of increments is normally distributed and with the right parameters. Observe that:

$$X(t) - X(s) = \frac{1}{a}(B(a^2 t) - B(a^2 s)),$$

which is a normal distribution with the parameters required for a standard Brownian motion, since it has expectation 0 and variance $\frac{1}{a^2}(a^2 t - a^2 s) = t - s$. □

This principle will be retrieved in the next chapter. To anticipate it, just consider $T(a, b) = inf\{t \geq 0 : B(t) = a \text{ or } B(t) = b\}$, that is, $T(a, b)$ is the first exit time of the interval $[a, b]$ for a one-dimensional Brownian motion . Then, with the scaled Brownian motion $X(t) = \frac{1}{a^2}B(a^2 t)$:

$$\mathbb{E}(T(a, b)) = a^2 \mathbb{E}\Big(inf\Big\{t \geq 0 : X(t) = 1 \text{ or } X(t) = \frac{b}{a}\Big\}\Big) = a^2 \mathbb{E}\Big(T(1, \frac{b}{a})\Big).$$

The conclusion is that the exit time for the motion only depends on the ratio of a and b.

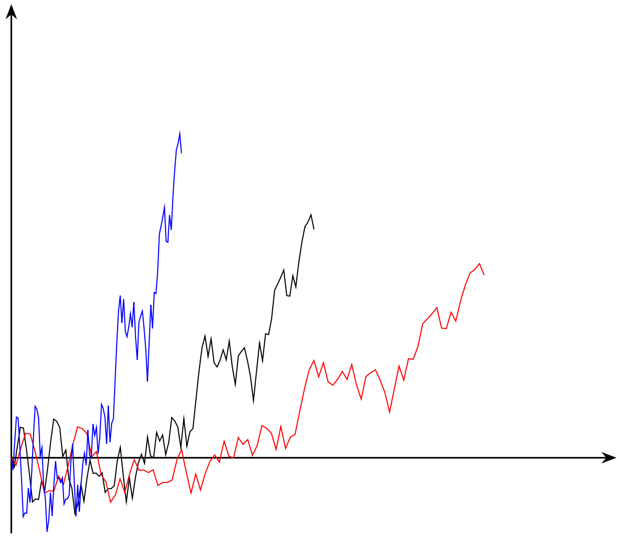


Figure 4.1: Brownian motion and scaled processes with parameters a=1.25 and a=0.75

Lemma 4.5. *Let $\{B_t : t \geq 0\}$ a standard Brownian motion. Then for all $s < t$:*

$$Cov(B_s, B_t) = s$$

Proof. Write $B_s B_t = B_s^2 + B_s(B_t - B_s)$, then, using that their mean is 0:

$$Cov(B_s, B_t) = \mathbb{E}(B_s B_t) - \mathbb{E}(B_s)\mathbb{E}(B_t) = \mathbb{E}(B_s^2) - \mathbb{E}(B_s)\mathbb{E}(B_t)$$
$$= \mathbb{E}(B_s^2) = Var(B_s) + \mathbb{E}(B_s)^2 = s$$

since $B_s \sim \mathcal{N}(0, s)$ for any $s \leq t$ and increments of non-overlapping intervals are independent. □

Theorem 4.6. (Time inversion) *Suppose $\{B(t) : t \geq 0\}$ is a standard Brownian motion, then the process $\{X(t) : t \geq 0\}$ defined by:*

$$X(t) = \begin{cases} 0 & \text{for } t = 0 \\ tB(\frac{1}{t}) & \text{for } t \neq 0 \end{cases}$$

is also a standard Brownian motion.

Proof. $\{X(t) : t \geq 0\}$ is also a Gaussian process and the random vectors $(X(t_1), \ldots, X(t_n))$ have expectation 0. Using the previous lemma, covariance for $h > 0$ is:

$$Cov(X(t+h), X(t)) = (t+h)Cov\Big(B\Big(\frac{1}{t+h}\Big), B\Big(\frac{1}{t}\Big)\Big) = t$$

Then, the covariance is also the one of a Brownian motion. Now, it only remains to see that paths are continuous. Away from the origin, they are clearly continuous. At $t = 0$, on the rationals X is almost surely continuous on $(0, \infty)$, so:

$$\lim_{t \to 0} X(t) = 0 \quad \text{almost surely}$$

Hence, $\{X(t) : t \geq 0\}$ has almost surely continuous paths and so it is a standard Brownian motion. □

The inversion will be a very useful in the analysis of assymptotic properties of Brownian motion as it will suffice to study it in a neighbourhood of $t = 0$.

Corollary 4.7. *Almost surely,*

$$\lim_{t\to\infty} \frac{B(t)}{t} = 0.$$

Proof. Let $\{X(t) : t \geq 0\}$ be defined as in Theorem 4.6, then by continuity of $X(t)$:

$$\lim_{t\to\infty} \frac{B(t)}{t} = \lim_{t\to\infty} X\Big(\frac{1}{t}\Big) = X(0) = 0.$$

□

4.4 Nondifferentiability

The natural questions that arise at this point are mainly two:

i) How strong is the continuity of Brownian motion?

ii) Is Brownian motion more regular than continuous, i.e., it can be diferenciated at some points or almost surely every point is nondifferentiable?

This two question will be answered in this section. The second one will be answered in brief, whereas the first one will be quantified by the notion of α-*Hölder continuity*.

Definition 4.8. A function $f : [0, \infty) \to \mathbb{R}$ is said to be **locally α-Hölder continuous** at $x \geq 0$, if there exists $\epsilon > 0$ and $c > 0$ such that

$$|f(x) - f(y)| \leq c|x - y|^{\alpha}, \qquad \text{for all } y \geq 0 \text{ with } |y - x| < \epsilon$$

We refer to $\alpha > 0$ as the **Hölder exponent** and $c > 0$ as the **Hölder constant**.

This is a measure on the strength of continuity of a function: as α increases the continuity gets stronger. For the particular case of Brownian motion, $\alpha = \frac{1}{2}$ is the critical value, as the following result states. We do not give a formal proof of it, since it requires a previous theorem, but it can be found in detail in [8], when Brownian motion continuity is extensively discussed.

Proposition 4.9. *If $\alpha < \frac{1}{2}$, then, almost surely, Brownian motion is everywhere locally α-Hölder continuous.*

It is important to recall that $\alpha < \frac{1}{2}$ is the optimal result meaning that for larger values of $\alpha > \frac{1}{2}$ the property does not holds, and there exist points where Brownian motion is not α-Hölder continuous, almost surely.
Next point to discuss is the differentiability of Brownian motion. Firstly, some local properties will be given, and then the nondifferentiability property will be extended to any compact set. Here, it will be fundamental to retrieve the *inversion property* of Brownian motion together with two new notions of derivatives.

Proposition 4.10. *Almost surely, for all $0 < a < b < \infty$, Brownian motion is not monotone on the interval $[a, b]$.*

Proof. Suppose $[a, b]$ is an interval of monotonicity, i.e., $B(s) < B(t)$ for all $a < s < t < b$. Then we can split the interval in $a = a_1 < a_2 < \cdots < a_n = b$ which are n subintervals of $[a, b]$. Then, every increment $B(a_{i+1}) - B(a_i)$ has to have the same sign, as Brownian motion is monotone on the interval. Also, increments are independent, hence the probability for all them to have the same sign is $2 \cdot 2^{-n}$. Taking $n \to \infty$, the probability of $[a, b]$ of being an interval of monotonicity is 0. Taking countable unions gives us that there are no such intervals with rational endpoints, but any interval must contain rational numbers, so we have the result required. □

Now, to continue studying the differentiability of Brownian motion, we are interested in its assymptotic behaviour, and then applying time inversion property, we will be able to deduce new properties for times near $t = 0$.

Proposition 4.11. *Almost surely,*

$$\limsup_{n\to\infty} \frac{B(n)}{\sqrt{n}} = +\infty \qquad \liminf_{n\to\infty} \frac{B(n)}{\sqrt{n}} = -\infty$$

In order to prove this proposition, we need an auxiliary result: the Hewitt-Savage $0 - 1$ law for exchangeable events.

Definition 4.12. Let $X_1, \dots, X_n$ be a sequence of random variables in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and consider a set A such that:

$$\{X_1, \dots X_n \in A\} \in \mathcal{F}$$

The event $\{X_1, \dots X_n \in A\}$ is called **exchangeable** if

$$\{X_1, \dots X_n \in A\} \subset \{X_{\sigma_1}, \dots X_{\sigma_n} \in A\}$$

for all finite permutations of $\mathbb{N}$, i.e., permutations with a finite number of nonfixed elements (Note that the inclusion can eventually become an equalty, in some cases).

Plainly, an event is exchangeable if it is closed under permutations of random variables. For instance, in the exmaple below, if Brownian motion is condidered in the discrete version, i.e., as a random walk, it is clear that many paths lead to the same event at the n-th step, the only condition that they must share is that the number of up/down moves is equal. All events that share this condition are exhangeable. The following law was firstly introduced in [5]:

Lemma 4.13. (Hewitt-Savage $0 - 1$ law)
Let be (X_n) a sequence of independent identically distributed random variables. Then the σ-field of exchangable events $\mathcal{E}$ is trivial.

Proof. We need to prove that for any $A \in \mathcal{E}$, $P(A)$ is either 0 or 1. Take $A \in \mathcal{E}$. Given a partition of the σ-field $\mathcal{E} = \cup F_n$, approximate A by $A_n \in F_n$ such that $P(A \triangle A_n) \to 0$. We write $A_n = \{(X_1, \dots X_n) \in B_n\}$ and $\tilde{A}_n = \{(X_{n+1}, \dots X_{2n}) \in B_n\}$ and condider the permutation the sends A_n to $\tilde{A}_n$.
By exchangeability, $P(\tilde{A}_n \triangle A) = P(A_n \triangle A) \Rightarrow P(A_n \cap \tilde{A}_n) \to P(A)$. Since the sequence (X_n) is i.i.d., we have: $P(A_n \cap \tilde{A}_n) = P(A_n)P(\tilde{A}_n) \to P(A)^2$. Hence, $P(A) \in \{0, 1\}$. □

Now, we have everything necessary to proof the non-differentiability, applying the Hewitt-Savage $0-1$ law:

Proof. (Proposition 4.11)
By Fatou's lemma, we have:

$$\mathbb{P}(B(n) > c\sqrt{n} \text{ infinitely often}) \geq \limsup_{n\to\infty} \quad \mathbb{P}(B(n) > c\sqrt{n}) > 0$$

Taking $X_n = B(n+1) - B(n)$, note that:

$$\{B(n) > c\sqrt{n} \text{ infinitely often}\} = \{\sum_{j=1}^{n} X_j > c\sqrt{n} \text{ infinitely often}\}$$

is an exhangeable event. Hence, by lemma 4.13 with probability 1, $B(n) > c\sqrt{n}$ infinitely often. As the constant c can be taken as large as desired, we deduce the first equalty of the proposition. The other part is proved analogously. □

In Corollary 4.7 it has been proved that the local growth of Brownian motion is slower than a linear function. The Proposition 4.11 shows that the growth is larger than the square root, so it is natural to try to find an intermediate function that expresses the assymptotic growth of Brownian motion more accurately. The answer to this question will be given by the *law of iterated logartihm*, which will not be stated in this work, although more details can be found in [8].

Definition 4.14. Given function f, the **upper derivative** is defined by:

$$D^* f(t) = \limsup_{h\downarrow 0} \frac{f(t+h) - f(t)}{h}$$

Analogously, the **lower derivative**:

$$D_* f(t) = \liminf_{h\downarrow 0} \frac{f(t+h) - f(t)}{h}$$

Theorem 4.15. *Fix $t \geq 0$. Then, almost surely, Brownian motion $B(t)$ is not differentiable at t. Furthermore,*

$$D^* B(t) = +\infty \text{ and } D_* B(t) = -\infty$$

Proof. Given a standard Brownian motion $B(t)$, we construct another by time inversion that satisfies:

$$D^* X(0) \geq \limsup_{n\to\infty} \frac{X(\frac{1}{n}) - X(0)}{\frac{1}{n}} \geq \limsup_{n\to\infty} \sqrt{n} X(\frac{1}{n}) = \limsup_{n\to\infty} \frac{B(n)}{\sqrt{n}} = +\infty.$$

Thus, $X(t)$ is not differentiable at $t = 0$. Now, consider $B(t) = X(t+s) - X(s)$, and given the non-differentiability of $X(t)$ at $t = 0$, now it is equivalent to the nondifferentiability of $B(t)$ at t. □

From the theorem above we cannot still deduce the desired property, it is a weaker form of it. We know that for a given t, Brownian motion will be nondifferentiable, almost surely, but it is not equivalent that almost surely every point is a nondiferentiability point. There exists a subtile but important difference between both statements, due to quatificator orders. However, the second statement is proved in the next theorem.

Theorem 4.16. *Almost surely, Brownian motion is nowhere differentiable. In addition, almost surely, for all t, either $D^*B(t) = +\infty$ or $D_*B(t) = -\infty$ or both.*

Proof. Suppose there is $t_0 \in [0, 1]$ such that $-\infty < D_*B(t) < D^*B(t) < \infty$, then

$$\limsup_{h\downarrow 0} \frac{|B(t+h) - B(t)|}{h} < \infty.$$

By continuity of Brownian motion in the interval $[0, T]$, we would have that:

$$\underset{h\in[0,1]}{sup} \frac{|B(t+h) - B(t)|}{h} < M.$$

Whence it suffices to prove that this happens with zero probability. Take $t_0 \in [\frac{k-1}{2^n}, \frac{k}{2^n}]$, for $n > 2$ and $1 \le j \le 2^n - k$, by the triangle inequalty:

$$\Big|B\Big(\frac{k+j}{2^n}\Big) - B\Big(\frac{k+j-1}{2^n}\Big)\Big| \le \\ \Big|B\Big(\frac{k+j}{2^n}\Big) - B(t_0)\Big| + \Big|B(t_0) - B\Big(\frac{k+j-1}{2^n}\Big)\Big| \le M\frac{j}{2^n} + M\frac{j-1}{2^n} = M\frac{2j+1}{2^n}.$$

Define the events

$$\Omega_{n,k} = \Big\{\Big|B\Big(\frac{k+j}{2^n}\Big) - B\Big(\frac{k+j-1}{2^n}\Big)\Big| \le M\frac{2j+1}{2^n} \text{ for } j = 1, 2, 3, \dots l\Big\}.$$

Then by independence of marginal distributions of increments in a Brownian motion and properly scaling of the process:

$$\mathbb{P}(\Omega_{n,k}) = \prod_{j=1}^{l} \mathbb{P}\Big\{\Big|B\Big(\frac{k+j}{2^n}\Big) - B\Big(\frac{k+j-1}{2^n}\Big)\Big| \le M\frac{2j+1}{2^n}\Big\} \le \mathbb{P}\{|B(1)| \le \frac{7M}{\sqrt{2^n}}\}^l.$$

Now, we have to think accurately of the distribution of $B(1)$. Note that we are considering a discrete approach of Brownian motion, i.e., a random walk. Then, $B(1)$ con only take a finite number of values, following a binomial distribution. Thus, when $n \to \infty$ by the central limit theorem A.12 the distribution tends to a $\mathcal{N}(0, \frac{1}{4})$, thus:

$$\mathbb{P}(|B(1)| \le \frac{7M}{\sqrt{2^n}}) = \int_{-\frac{7M}{\sqrt{2^n}}}^{\frac{7M}{\sqrt{2^n}}} \frac{2}{\sqrt{2\pi}} e^{-x^2} dx \le \frac{7M}{\sqrt{2^n}} = 7M2^{-\frac{n}{2}}.$$

Hence, given n, we have:

$$\mathbb{P}(\bigcup_{k=1}^{2^n-l} \Omega_{n,k}) \le 2^n(7M2^{-\frac{n}{2}})^l < (7M)^3 2^{-\frac{n}{2}},$$

given, $l \ge 3$, which is necessary for this events to form a convergent sum for all n. Now, using the Borel-Cantelli lemma:

$$\mathbb{P}\Big\{ \text{ there is } t_0 \in [0, 1] : \underset{h\in[0,1]}{sup} \frac{|B(t_0+h) - B(t_0)|}{h} \le M\Big\} \le \mathbb{P}\Big(\bigcup_{k=1}^{2^n-l} \Omega_{n,k} \text{ for infinitely many } n\Big) = 0,$$

where the inequalty holds as a consequence of the more restrictive condition, as for the latter case the condition must be satisfied infinitely many cases. ☐

5. The Markov property of Brownian motion

The introduction made in chapter 3 will now be contextualized with Brownian motion. Some results seen there will be revisited, may be with some modifications, adapted to this particular process. It aims at stablishing the *Markov property* (weak and strong versions) using the notion of *stopping times* that holds for Brownian motion, and also some consequences will be included, such as the reflection principle.

5.1 Markov property and Blumenthal's $0-1$ law

In this chapter our mission is to state in which sense Markov property holds for the case of Brownian motion, even in multidimensional cases, and from here deduce some interesting properties and laws. In a more general property, we will see that some processes that proceed from Brownian motion are also Markov processes. This will lead to the reflection principle, which will be dicussed in the next section, whose consequences are surprinsing and interesting from a mathematical point of view.

Definition 5.1. If $B_1, \dots B_d$ are independent Brownian motions started in $(x_1, \dots, x_d)^T$, the stochastic process $\{B(t) : t \geq 0\}$ given by

$$B(t) = (B_1(t), \dots, B_d(t))$$

is a d-**dimensional Brownian motion** started in $(x_1, \dots, x_d)^T$.

If we recall the Markov property for a stochastic process, we have that a process $\{B(t) : t \geq 0\}$ with such property is only determined by the distribution at some s, and information in the interval $[0, s]$ can be disregarded, and it will have no influence in the future of the process.

Also note that two stochastic processes $\{X(t) : t \geq 0\}$ and $\{Y(t) : t \geq 0\}$ are said to be **independent** is for any set of $t_1, t_2, \cdots, t_n$ and $s_1, s_2, \cdots, s_n$ the random vectors $(X(t_1), X(t_2), \dots, X(t_n))$ and $(Y(s_1), Y(s_2), \dots, Y(s_n))$ are independent vectors.

Theorem 5.2. (Markov property) *Suppose that $\{B(t) : t \geq 0\}$ is a d-dimensional Brownian motion started in $x \in \mathbb{R}^d$. Let $s > 0$, then the process $\{B(t+s) - B(s) : t \geq 0\}$ is a standard Brownian motion independent of the process $\{B(t) : 0 \leq t \leq s\}$.*

Proof. It is clear that $\{B(t+s) - B(s) : t \geq 0\}$ satisfies the definition of a d-dimensional Brownian motion, and the independence of $\{B(t) : 0 \leq t \leq s\}$ is a consequence of the independence of increments that characterizes the Brownian motion, by definition. □

In this context, information that a random process gives us can be gathered in a *filtration*, which abstractly will come to express exactly this. More formally:

Definition 5.3. A **filtration** on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a family $(\mathcal{F}(t) : t \geq 0)$ of σ-algebras such that $\mathcal{F}(s) \subset \mathcal{F}(t) \subset \mathcal{F} \quad \forall s < t$. A prrbability space with a filtration defined, is called a **filtered probability space**. A stochastic process $\{X(t) : t \geq 0\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is called **adapted** if $X(t)$ is $\mathcal{F}(t)$-measurable for any $t \geq 0$.

Remark 5.4. A filtration $\mathcal{F}(t)$ can also be denoted by $\mathcal{F}_t$, which is generally used, but we will denote it in the first way because the latter notation can result ambiguous for the next steps that we are going to consider in this chapter.

Given a random process, such as Brownian motion $\{B(t) : t \geq 0\}$, it is easy to build an adpated filtration. We just need to take the filtration $(\mathcal{F}^0(t))$, where $\mathcal{F}^0(t)$ is the σ-algebra generated by the random variables $\{B(s) : 0 \leq s \leq t\}$. Obviously, Brownian motion is adpated to this filtration, that contains the information of the process observed up to t. From theorem 5.2 we can easily deduce that the process $\{B(t+s) - B(s) : t \geq 0\}$ is independent of $\mathcal{F}^0(s)$, but can this be extended? What if we know the information up to an infinitessimal time instant after s?
Consider the filtration

$$\mathcal{F}^+(s) = \cap_{t>s}\mathcal{F}^0(t)$$

Then, clearly $\mathcal{F}^0(s) \subset \mathcal{F}^+(s)$

Theorem 5.5. *For $s \geq 0$, the process $\{B(t+s) - B(s) : t \geq 0\}$ is independent of $\mathcal{F}^+(s)$.*

This theorem is just a consequence of the continuity of Brownian paths. It is necessary to take a decreasing sequence $\{s_n\}_{n\geq 0}$ convergent to s, and by continuity, the theorem holds.
From all σ-algebras, there is one, the so-called **germ σ-algebra**, defined as $\mathcal{F}^+(0)$, which contains information known at an infinitessimal time after the start of the process. So, the following result was found and named after its discoverer in [3]:

Theorem 5.6. (Blumenthal's $0-1$ law) *Let $A \in \mathcal{F}^+(0)$, then $\mathbb{P}(A) \in \{0,1\}$*

Proof. Apply Theorem 5.5 for $s = 0$, which implies that for any $A \in \sigma\{B(t) : t \geq 0\}$, A is independent of $\mathcal{F}^+(0)$. In particular applies if $A \in \mathcal{F}^+(0)$, whence A is independent of itself, thus $\mathbb{P}(A) \in \{0,1\}$ ☐

Remark 5.7. The previous proof is interesting from the point of view of the idea used. $0-1$ laws have appeared in this work, and all of them are proved using the same idea. If an event A has probability either 0 or 1, then it suffices to prove that this event is independent of itself.

Theorem 5.8. *Suppose $\{B(t) : t \geq 0\}$ a standard Brownian motion.*
Define $\tau = inf\{t > 0 : B(\tau) > 0\}$ and $\sigma = inf\{t > 0 : B(\tau) = 0\}$. Then,

$$\mathbb{P}\{\tau = 0\} = \mathbb{P}\{\sigma = 0\} = 1$$

Proof.

$$\{\tau = 0\} = \bigcap_{n=1}^{\infty} \left\{ \text{there is } 0 < \epsilon < \frac{1}{n} : B(\epsilon) > 0 \right\}$$

Obviously, the event is in $\mathcal{F}^+(0)$, and in addition, $\mathbb{P}(\tau < t) > \mathbb{P}(B(t) > 0) = \frac{1}{2}$, which proves the first part, as a consequence of theorem 5.6. The same argument is valid for the case $B(t) < 0$, thus, by continuity of Brownian motion, using the mean-value theorem, we have that the property for σ holds. ☐

5.2 The strong Markov property

As in chapter 3, to introduce the idea of the strong Markov property, it is necessary to use stopping times, which intuitively are related with the Markov property of only near-present dependence. Roughly, a stopping time is a random variable whose value at time t can be deduced from the information of the past values of itself. More formally,

Definition 5.9. A random variable T with values in $[0,\infty]$, defined in a probability space with an associated filtration $(\mathcal{F}(t) : t \geq 0)$ is called a **stopping time** if $\{T < t\} \in \mathcal{F}(t)$, for every $t \geq 0$. Additionally, if $\{T \leq t\} \in \mathcal{F}(t)$ for every $t \geq 0$, then T is a **strict stopping time**.

From the definition is is clear that every strict stopping time is also a stopping time, just observing the following:

$$\{T < t\} = \bigcup_{n=1}^{\infty} \{T \leq t + \frac{1}{n}\} \in \mathcal{F}(t).$$

The main problem is to indentify in which cases the converse in also true. Fortunately, there are nice enough filtrations where this happens. To overcome this problem in case of Brownian motion it is appropiate to work with the filtration $(\mathcal{F}^+(t) : t \geq 0)$. Recall that this filtration satisfies $\mathcal{F}^0(s) \subset \mathcal{F}^+(s)$, what results in a larger filtration, and hence more stopping times accepted by the process. The key point why we consider this filtration instead of the natural one is due to the *right-continuity* of $\mathcal{F}^+(t)$.

Theorem 5.10. *Every stopping time T with respect to the filtration $(\mathcal{F}^+(t) : t \geq 0)$ is also a strict stopping time.*

For every stopping time T, we define:

$$\mathcal{F}^+(T) = \{A \in \mathcal{A} : A \cup \{T < t\} \in \mathcal{F}^+(t) \text{ for all } t \geq 0\}$$

This introduces a slightly different notion, as now we are only considering the events that are completely known or determined up to a stopping time T.

Theorem 5.11. (Strong Markov property) *For every almost surely finite stopping time T, the process $\{B(t+T) - B(T) : t \geq 0\}$ is a standard Brownian motion independent of $\mathcal{F}^+(T)$.*

Proof. We start by an upper discrete approximation of the stopping time T by $T_n = \frac{(m+1)}{2^n}$ if $\frac{m}{2^n} \leq T < \frac{(m+1)}{2^n}$. Also, consider the following processes:

$$B_k(t) = B\Big(t + \frac{k}{2^n}\Big) - B\Big(\frac{k}{2^n}\Big)$$

and, fixing n, consider:

$$B_n(t) = B(t + T_n) - B(T_n)$$

Suppose we have an event $E \in \mathcal{F}^+(T_n)$. Then, for every event $\{B_n \in A\}$, we have:

$$\begin{aligned}\mathbb{P}(\{B_n \in A\} \cup E) = \sum_{k=0}^{\infty} \mathbb{P}(\{B_k \in A\} \cap E \cap \{T_n = k2^{-n}\}) \overset{(*)}{=} \\ = \sum_{k=0}^{\infty} \mathbb{P}(\{B_k \in A\}) \cap \mathbb{P}(E \cap \{T_n = k2^{-n}\}) \overset{(**)}{=} \\ = \mathbb{P}(\{B \in A\} \sum_{k=0}^{\infty} \mathbb{P}(E \cap \{T_n = k2^{-n}\}) = \mathbb{P}\{B \in A\}\mathbb{P}(E)\end{aligned}$$

which shows that B_n is a Brownian motion independent of E, hence of $\mathcal{F}^+(T_n)$. In the proof it has been used that:

(*) $\{B_k \in A\}$ is independent of $\{E \cap T_n = k2^{-n}\} \in \mathcal{F}^+(k2^{-n})$.

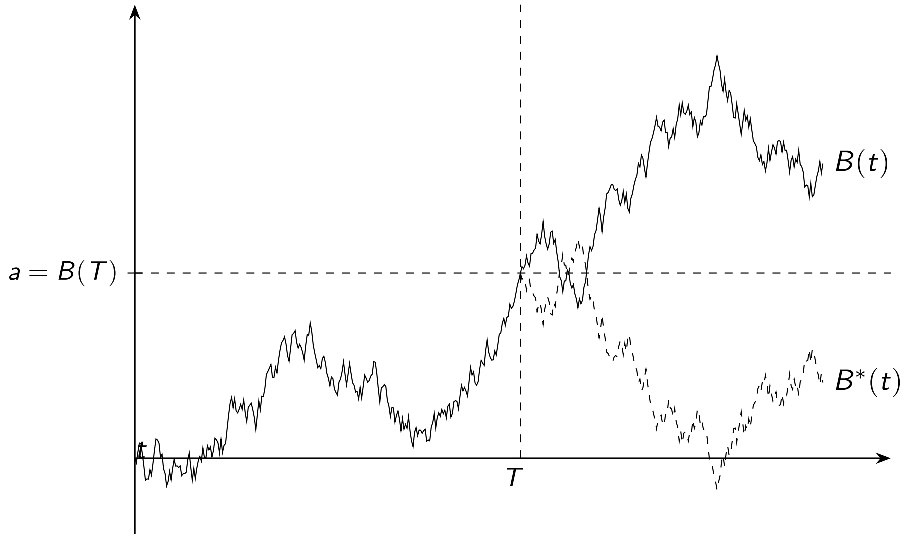
(*) $\mathbb{P}(B_k \in A) = \mathbb{P}(B \in A)$ is independent of k.

The generalisation to stopping times T follows from the continuity of Brownian motion for any sequence T_n such that $T_n \downarrow T$. ☐

5.3 The reflection principle

The first and most important consequence of the Markov property of Brownian motion is the *reflection principle*. It is almost a raw application of it. We consider a stopping time T, which can be arbitrary, and the process:

$$B^*(t) = B(t)\mathbb{1}_{\{t \le T\}} + (2B(T) - B(t))\mathbb{1}_{\{t > T\}} \tag{5.1}$$



Theorem 5.12. (Reflection principle) *If T is a stopping time and $\{B^*(t) : t \ge 0\}$ is a standard Brownian motion, then the process defined in (5.1), called* **Brownian motion reflected at** T *is also a standard Brownian motion.*

Proof. If T is finite, then using the strong Markov property,

$$\{B(t+T) - B(T) : t \ge 0\} \qquad \text{and} \qquad \{-(B(t+T) - B(T)) : t \ge 0\}$$

are Brownian motion independent of the original $\{B(t) : t \ge 0\}$. Hence the process $B^*(t)$ is a standard Brownian motion, as the first part and the second one have the same distirbution and both are standard Brownian motions. ☐

Let $a = B(T)$, that is, the level at the stopping time. This level clearly exhibits a symmetry for those Brownian motions that at $t \in [T, T']$ are above or below a, as for any path that is above there is a relfected path that is below and conversely.
Let $M(t) = \max_{0 \le s \le t} B(s)$. This random variable has unknown distribution, a priori, but in application of the reflection principle, we can determine it:

Corollary 5.13. *If $a > 0$, then $P\{M(t) > a\} = 2P\{B(t) > a\}$*

Proof. Let $T = inf\{t \ge 0 : B(t) = a\}$ and $\{B^*(t) : t \ge 0\}$ is the refelected Brownian motion, then:

$$P(M(t) > a) = P(M(t) > a, B(t) \ge a) + P(M(t) > a, B(t) < a) = \\ P(B(t) > a) + P(B^*(t) > a) = 2P(B(t) > a)$$

since both are Brownian motions. ☐

In chapter 4 the regularity of Brownian motion have been studied with a direct approach. Now, also as a consequence of the reflection principle, the non-differentiation results is retrieved.

Corollary 5.14. *If $\{B(t) : 0 \leq t \leq T\}$ is a standard Brownian motion, then, almost surely, $\forall t \geq 0$ $B(t)$ is nondifferentiable.*

Proof. We argue by contradiction. Suppose that at some t, the derivative exists. Then:

$$|B(t+\epsilon') - B(t)| \leq \epsilon A$$

Let $M(\epsilon) = \max_{\epsilon'} |B_{t+\epsilon'} - B_t| \leq \epsilon A$. By Corollary 5.13, we have

$$P(M(t) > \epsilon A) = 2P(B(\epsilon) > \epsilon A) = 2P(\mathcal{N}(0,\epsilon) > \epsilon A) = 2P(\mathcal{N}(0,1) > \sqrt{\epsilon} A) \underset{\epsilon \to 0}{\longrightarrow} 1$$

where $\mathcal{N}$ denotes a normal distribution.
Thus, there is a contradiction, since the maximum is not bounded for any ϵ with probability 1, and therefore, $B(t)$ is nondifferentiable at any point. □

6. From random walks to Brownian motion

6.1 Preliminaries

Throughout this chapter, there will be a permanent idea overflying theorems and arguments: how to approximate a Brownian motion through a random walk. There are many constructions possible, but the approach given in the proof of the Ray-Knight theorem has the advantadge that it is natural and elementary. Sometimes, the embedding process is called *twist and shrink* for obvious reasons. The key point is to construct a sequence of random walks uniformly convergent to Brownian motion.

Firstly, imagine that we observe the motion of a particle when it hits to coordinates $\{j, j \in \mathbb{Z}\}$. Then, we observe the particle exclusively when coordinates $\{\frac{j}{2}, j \in \mathbb{Z}\}$ are hit. And so on. This generates a sequence of random walks, which suitably managed, converge appropiately.
We consider a sequence of independent identically distributed random variables $X_m(k)$ in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}(X_m(k) = \pm 1) = \frac{1}{2}$, for $m \geq 0$ and $k \geq 1$. Then the random walk considered is: $S_m(0) = 0$ and $S_m(n) = \sum_{k=1}^{n} X_m$. Thus, $\mathbb{E}(X_m(k)) = 0$ and $Var(X_m(k)) = 1$. Hence:

$$\mathbb{E}(S_m(n)) = 0 \qquad Var(S_m(n)) = n. \tag{6.1}$$

This expression suggests that time and space scaling can not be done arbitrarely. The question is: if the space is halved, what should be time reduction in order to converge to a Brownian motion? From (6.1), we know that the square root of the average squared distance of a random walks from the origin at time n is $\sqrt{n}$, so an appropiate shrink after n would require a lenght of $\frac{1}{\sqrt{n}}$.
So far, the construction of the random walks is independent of m, but this does not give a convengence to a limit as desired. We need to find a link, which will be a successive refinement of them, that joins them correctly. With this construction in mind, we define the *stopping times* $T_m(0) = 0$, and:

$$T_m(k+1) = inf\{n > T_m(k) : |S_m(n) - S_m(T_m(k))| = 2\} \tag{6.2}$$

These are random (stopping) times when the random walk visits even integers different from the provious one. A key point is that the distribution of $T_m(k)$ is known, and it is given by the following Lemma:

Lemma 6.1. *Let $T_m(k)$ be random variables defined as in (6.2). Then $T_m(k)$ follows the distribution of the double negative binomial random variable with parameters k and $p = \frac{1}{2}$.*

Proof. Consider the random variables $\tau_j = inf\{n > \tau_{j-1} : |S_m(n) - S_m(\tau_{j-1})| = 2\} - \tau_{j-1}$, with $\tau_0 = 0$. These random variables have a known distribution. Imagine the random walks as a sequence of independent pairs of steps: *returns* or *change of magnitud* ± 2. Then, $P\{\tau_{k+1} = 2j\} = \frac{1}{2^j}$, what implies that $\tau_{k+1} = 2Y_{k+1}$, where Y_{k+1} is a geometric random variable with paraemter $p = \frac{1}{2}$. This results follow from the fact that $T_m(k)$ can be seen as the sum of k independent geometric random variables τ_j:

$$T_m(k) = \sum_{j=0}^{k} \tau_j = 2\sum_{j=0}^{k} Y_j,$$

that leads to a negative binomial distribution, and also, $\mathbb{E}(T_m(k)) = 4k$ and $Var(T_m(k)) = 4\frac{p}{(1-p)^2}k = 8k$. □

As a consequence the Central Limit Theorem 6.1 and A.12, we have that for any real fixed x and $k \to \infty$:

$$P\Big\{\frac{T_k - 4k}{\sqrt{8k}}\Big\} \to \Phi(x) \tag{6.3}$$

where $\Phi(x)$ denote the cummulative distribution of a standard normal random variable. Now, we can define recursively *twisted* random variables for $T_m(k) < n < T_m(k+1)$:

$$\tilde{X}_m(n) = \begin{cases} X_m(n) & \text{if } S_m(T_m(k+1)) - S_m(T_m(k)) = 2\tilde{X}_{m-1}(k+1) \\ -X_m(n) & \text{otherwise} \end{cases} \tag{6.4}$$

and then $\tilde{S}_m(n) = \sum_{j=1}^{n} \tilde{X}_m(j)$.

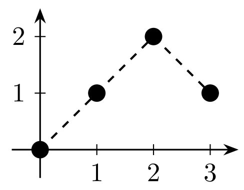


Figure 6.1: $S_0(t, \omega)$

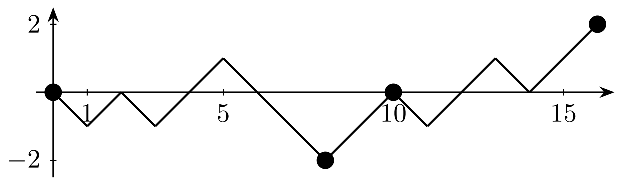


Figure 6.2: $S_1(t; \omega)$

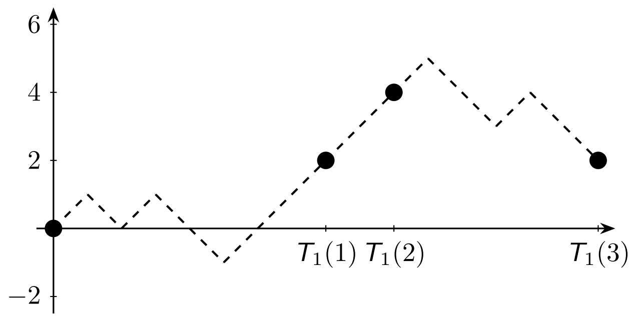


Figure 6.3: $\tilde{S}_1(t; \omega)$

The figures below illustrate the *twist* introduced to the random walks that correlate two consecutive terms of the sequence. However, this correlation must satisfy some restrictions, commented above, and next Lemma is going to give us certainty that the construction respects the random walk structure.

Lemma 6.2. *For each $m \geq 0$, $\tilde{S}(t), (t \geq 0)$ is a random walk, that is, $\tilde{X}_m(1), \tilde{X}_m(2), \ldots$ is a sequence of independent random variables indentically distributed such that:*

$$P\{\tilde{X}_m(n) = 1\} = P\{\tilde{X}_m(n) = -1\} = \frac{1}{2} \tag{6.5}$$

The proof of this lemma is ommited, since its very extense. The difficult point of it is to prove the independence of the random walks defined. For more detials, see [9, Lemma 1].

6.2 Twist and shrink embedding

The main consequence of Lemma 6.2 is the *twist* property that implies the correlation between two consecutive random walks in the following way:

$$\tilde{S}_m(T_m(k)) = \sum_{j=1}^{k}(\tilde{S}_m(T_m(j)) - \tilde{S}_m(T_m(j-1))) = \sum_{j=1}^{k}\tilde{X}_{m-1}(j) = \tilde{S}_{m-1}(k) \tag{6.6}$$

The next step is the *shrink* of the random walk. We know that time steps must be the square of spatial steps. Therefore, we define the m-th approximation to the Wiener process by:

$$\tilde{B}_m\Big(\frac{t}{2^{2m}}\Big) = \frac{1}{2^m}\tilde{S}_m(t) \quad (t \geq 0, m \geq 0) \tag{6.7}$$

Rewriting it with a change of variable:

$$\tilde{B}_m(t) = 2^{-m}\tilde{S}_m(2^{2m}t)$$

These expressions for $(m \geq 0)$ give an approximation to the Wiener process on the coordinates $x = \frac{j}{2^m} \quad j \in \mathbb{Z}$. Thus, for the case of Brownian motion, (6.6) becomes the following *refinement property*:

$$\tilde{B}_m\Big(\frac{T_m(k)}{2^{2m}}\Big) = \frac{1}{2^m}\tilde{S}_m(T_m(k)) = \frac{1}{2^m}\tilde{S}_{m-1}(k) = \tilde{B}_{m-1}\Big(\frac{k}{2^{2(m-1)}}\Big) \tag{6.8}$$

for any $m \geq 1$ and $k \geq 0$.

The main objective of this chapter is to see how, from the *refinement property* proved above, the times $\frac{T_m(k)}{2^{2m}}$ and $\frac{k}{2^{2(m-1)}}$ get arbitrarely close, so that, $B_m(t)$ converges to the Wiener process, donoted by $W(t)$, hereinafter. Firstly, we establish a technical lemma, which will be useful for other auxiliar results.

Lemma 6.3. *Suppose that for $j \geq 0$, we have $\mathbb{E}(Z_j) = 0$ and $Var(Z_j) = j$ and with some $a > 0$ and $b > 0$,*

$$P(\{|Z_j| \geq t\}) \leq 2a^j e^{-bt} \quad (t \geq 0)$$

(exponential-Chebyshev inequalty).
Assume as well that there exists $j_0 > 0$ such that for any $j > j_0$,

$$P\Big\{\frac{|Z_j|}{\sqrt{j}} > x_j\Big\} \leq e^{-\frac{x_j}{2}}$$

whenever $X_j \to \infty$ and $x_j = o(j^{1/6})$ (large deviation type inequalty).
Then, for any $C > 1$,

$$P\Big\{\max_{0\leq j\leq N}|Z_j| \geq (2CNlog(N))^{\frac{1}{2}}\Big\} \leq \frac{2}{N^{1-C}} \tag{6.9}$$

if N is large enough, $N \geq N_0(C)$.

Proof. Firstly we introduce a bound for the maximum, which, although is not accurate, can be very useful in this lemma. For any random variable Z_j, it holds that:

$$P(\max_{1\leq j\leq N} Z_j > t) = P(\cup_{j=1}^{N}\{Z_j > t\}) \leq \sum_{j=1}^{N} P\{Z_j > t\} \tag{6.10}$$

Then, the condition in (6.9) can be directly found with the previous bound. We split the sum in two parts, which will be bounded with the Chebishev-type inequalty and large deviation type inequalty respectively. In the second case, x_j will be taken equal to $\sqrt{2Clog(N)}$ and since $j \leq N$ then $x_j \to \infty$. For $j \geq log^4(N)$ the conditions for x_j hold, and both inequalties can be applied in the following way:

$$P\Big\{\max_{0\leq j\leq N}|Z_j| \geq \sqrt{2CNlog(N)}\Big\} \leq$$

$$\sum_{j=0}^{\lfloor log^4(N)\rfloor} 2a^j e^{-b\sqrt{2CNlog(N)}} + \sum_{j=\lfloor log^4(N)\rfloor}^{N} P\{|Z_j|/\sqrt{j} \geq \sqrt{2Clog(N)}\} \leq$$

$$\frac{2a}{a-1} e^{log(a)log^4(N)-b\sqrt{2CNlog(N)}} + Ne^{-Clog(N)} \leq \frac{2}{N^{C-1}}$$

□

The crucial fact that happens in this case is that the assumption made in the previous lemma holds for any sequence Z_j of partial sums of independent identically distributed random variables with mean 0. The justification of this properties can be found in [10]. Next result is a consequence of Lemma 6.3 applied to the random variable $\frac{T_m(k)-4k}{8k}$, with $N = K2^{2m}$. So $log(N) = log(K) + 2log(N)m$ and consequently, for m large enough, $logN \leq \frac{3}{2}m$ and $\sqrt{2CNlog(N)} \leq \sqrt{3CKm}2^m$.

Lemma 6.4. *a) For any $C > 1$, $K > 0$, and for any $m \geq m_0(C, K)$ we have*

$$P\{\max_{0\leq k/2^{2m}\leq K}|T_{m+1}(k) - 4k| \geq \sqrt{24CKm}2^m\} < 2(K2^{2m})^{1-C} \tag{6.11}$$

b) For any $K > 0$,

$$\max_{0\leq k/2^{2m}\leq K}\left|\frac{T_{m+1}(k)}{2^{2(m+1)}} - \frac{k}{2^{2m}}\right| < \sqrt{2Km}2^{-m} \tag{6.12}$$

with probability 1 for all but finitely many m.

Proof. a) (6.11) is a direct consequence of Lemma 6.3

b) Take $C = \frac{4}{3} > 1$ in *a*) and difine the events:

$$A_m = \Big\{\max_{0\leq k/2^{2m}\leq K}|T_{m+1}(k) - 4k| \geq \sqrt{32Km}2^m\Big\} \tag{6.13}$$

By 6.11, we know that $P\{A_m\} < 2(K2^{2m})^{-\frac{1}{3}}$. This implies that $\sum_{m=0}^{\infty} P(A_m) < \infty$ and by the first Borel-Cantelli lemma, with probability 1 only finitely many of the events A_m occur. That is, almost surely for all but finitely many m:

$$\max_{0\leq k/2^{2m}\leq K}|T_{m+1}(k) - 4k| < \sqrt{32Km}2^m$$

which is equivalent to (6.12)

□

Lemma 6.5. *a) For any $C \geq \frac{3}{2}$, $K > 0$ and for any $n \geq n_0(C)$ we have*

$$P\Big\{\max_{0\leq k/2^{2n}\leq K}|B_{n+1}(T_{n+1}(k)/2^{2(n+1)}) - B_{n+1}(k/2^{2n})| \geq (1/8)n2^{-\frac{n}{2}}\Big\} \leq 3(K2^{2n})^{1-C} \tag{6.14}$$

and

$$P\Big\{\max_{0\leq t\leq K}|B_{n+j}(t) - B_n(k/2^{2n})| \geq n2^{-\frac{n}{2}} \quad \text{for some } j \geq 1\Big\} < 6(K2^{2n})^{1-C}. \tag{6.15}$$

b) For any $K > 0$,
$$\max_{0\le t\le K}|B_{n+j}(t) - B_n(t)| < n2^{-\frac{n}{2}} \tag{6.16}$$
with probability 1 *for all* $j \geq 1$ *and for all but finitely many n.*

Proof. First of all we consider the difference between two consecutive approximations $B_{m+1}(t) - B_m(t)$, whose maximum can be approximated by the maximum over dyadic intervals of the form $\frac{k}{2^{2m}}$ by construction. Moreover the increment between two points $\frac{k}{2^{2m}}$ and $\frac{k+1}{2^{2m}}$ is always equal to 2^{-m}. We take, as before, $t_m = \lfloor t2^{2m} \rfloor$ for each $t \in [0, K]$, thus one has $\frac{t_m}{2^{2m}} \leq t < \frac{t_{m+1}}{2^{2m}}$. Consequently:

$$|B_m(t) - B_m(t_m/2^{2m})| < 2^{-m},$$

and
$$|B_{m+1}(t) - B_{m+1}(4t_m/2^{2(m+1)})| < 2^{-m}.$$

Using these bounds directly and the triangular inequalty with maximum, we obtain:

$$\max_{0\le t\le K}|B_{m+1}(t) - B_m(t)| \leq \max_{0\le k/2^{2m}\le K}|B_{m+1}(4k/2^{2(m+1)}) - B_m(k/2^{2m})| + 3\cdot 2^{-m}$$

Moreover, by (6.8) and (6.7)

$$B_{m+1}(4k/2^{2(m+1)} - B_m(k/2^{2m}) = 2^{-(m+1)}\tilde{S}_{m+1}(4k) - 2^{-(m+1)}\tilde{S}_{m+1}(T_{m+1}(k)) \tag{6.17}$$

Thus,

$$\begin{aligned} &P\Big\{\max_{0\le t\le K}|B_{n+1}(t) - B_{n+1}(t)| \geq (1/4)m2^{-\frac{m}{2}}\Big\} \\ \leq\ &P\Big\{\max_{0\le k/2^{2m}\le K}|B_{m+1}(4k/2^{2(m+1)}) - B_m(k/2^{2m})| \geq (1/8)m2^{-\frac{m}{2}}\Big\} \\ =\ &P\Big\{\max_{0\le k/2^{2m}\le K}|\tilde{S}_{m+1}(4k) - \tilde{S}_{m+1}(T_{m+1}(k))| \geq (1/4)m2^{\frac{m}{2}}\Big\} \end{aligned}$$

for m large enough. By Lemma 6.3, for m large enough again, the probability of:

$$A_m = \Big\{\max_{0\le k\le K2^{2m}}|T_{m+1}(k) - 4k| \geq \sqrt{24CKm}2^m\Big\}$$

is very small.
The next part of the proof corresponds to the split of the previous expression in two parts accordidng to A_m and A_m^c. This includes a lot of technical details in order to be able to apply the previous lemmas for suitable constants N' and C' which will not be included here. The complete proof of the statement a) of this lemma can be found in [10]. From here, we can conclude that:

$$P\Big\{\max_{0\le k/2^{2n}\le K}|B_{n+1}(T_{n+1}(k)/2^{2(n+1)}) - B_{n+1}(k/2^{2n})| \geq (1/8)n2^{-\frac{n}{2}}\Big\} \leq 3(K2^{2n})^{1-C}$$

for m large enough.
By (6.18), $\max_{0\le t\le K}|B_{m+1}(t) - B_m(t)| > (1/4)m2^{-\frac{m}{2}}$ for all $m \geq n \geq 15$, which implies, for $j \geq 1$:

$$\begin{aligned} &\max_{0\le t\le K}\Big|B_{n+j}(t) - B_n(t)\Big| = \max_{0\le t\le K}|\sum_{m=n}^{n+j-1} B_{m+1}(t) - B_m(t)| \\ \leq &\sum_{m=n}^{n+j-1}\max_{0\le t\le K}|B_{m+1}(t) - B_m(t)| < \sum_{m=n}^{\infty}(1/4)m2^{-\frac{m}{2}} < n2^{-\frac{n}{2}} \end{aligned}$$

With all those bounds, we can conclude the following:

$$P\Big\{\max_{0\le t\le K}|B_{n+j}(t)-B_n(k/2^{2n})|\ge n2^{-\frac{n}{2}}\quad\text{for some } j\ge 1\Big\}$$
$$\le\sum_{m=n}^{\infty}P\Big\{\max_{0\le t\le K}|B_{m+1}(t)-B_m(t)|\ge(1/4)m2^{-\frac{m}{2}}\Big\}$$
$$\sum_{m=n}^{\infty}3(K2^{2m})^{1-C}=3(K2^{2n})^{1-C}\frac{1}{1-2^{2(1-C)}}<6(K2^{2n})^{1-C}$$

whenever $C>\frac{3}{2}$, for any n large enough. The part remaining of the statement in $b)$ is proved analogously to Lemma 6.3 using the Borel-Cantelli lemma. □

Remark 6.6. In the proof we have used the elementary bound:

$$\sum_{m=n}^{\infty}m2^{-\frac{m}{2}}\le 4n2^{-\frac{n}{2}}\tag{6.18}$$

for $n\ge 15$, using that the previous sum as a derivative of $\sum_{m=n}^{\infty}x^m$ at the point $x=\frac{1}{\sqrt{2}}$.

Theorem 6.7. *As $n\to\infty$, almost surely, for all $t\in[0,\infty]$, $B_n(t,\omega)$ converge to $W(t,\omega)$ such that:*

i) $W(0,\omega)=0$, and $W(t,\omega)$ is a continuous function of t on the interval $[0,\infty]$;

ii) for any $0\le s<t$, $W(t)-W(s)$ is a normally distributed random variable with expectation 0 and variance $t-s$;

iii) for any $0\le s<t\le u<v$, the increments $W(t)-W(s)$ and $W(v)-W(u)$ are independent random variables.

By definition, $W(t)$ $(t\ge 0)$ is called a **Wiener process**, *as defined in chapter 4.*
Further, we have the following estimates for the difference of the Wiener process and its approximations.

a) For any $C\ge\frac{3}{2}$, $K>0$ and for any $n\ge n_0(C)$ we have

$$P\Big\{\max_{0\le t\le K}|W(t)-B_n(t)|\ge n2^{-\frac{n}{2}}\Big\}\le 6(K2^{2n})^{1-C}\tag{6.19}$$

b) For any $K>0$,

$$\max_{0\le t\le K}|W(t)-B_n(t)|<n2^{-\frac{n}{2}}\tag{6.20}$$

with probability 1 for all $j\ge 1$ and for all but finitely many n.

Proof. Recall that $B_n(0,\omega)$ for any n, and by (6.5) $B_n(t,\omega)$ is convergent, thus it must hold that $W(0,\omega)$ for any $\omega \in \Omega$.
Taking $j \to \infty$ in (6.15), (6.19) follows. By (6.16), the convergence of $B_n(t)$ is uniform on any bounded interval $[0,K]$ and, since $B_n(t)$ is continuous for all n, we have that the limit $W(t)$ is also continuous, since we are considering a compact set. This proves i).
Take arbitrary $t > s \geq 0$. With $K > t$ fixed, (6.19) shows that for any $\delta > 0$ there exists an $n \geq n_0(C,K)$ such that:

$$P\Big\{\max_{0\leq u\leq K}|W(u) - B_n(u)| \geq \delta\Big\} \leq \delta \tag{6.21}$$

Observe that:

$$P\{W(t) - W(s) \leq x\} = P\{B_n(t) - B_n(s) \leq x - (W(t) - B_n(t)) + (W(s) - B_n(s))\},$$

(6.21) implies that

$$P\{B_n(t)-B_n(s) \leq x-2\delta\}-2\delta \leq P\{W(t)-W(s) \leq x\} \leq P\{B_n(t)-B_n(s) \leq x+2\delta\}+2\delta \tag{6.22}$$

These bounds of the probability distribution of $W(t) - W(s)$ indicate that can be deduced from $B_n(t) - B_n(s)$, whose expression is already known. More precisely,

$$B_n(t) - B_n(s) = 2^{-n}\tilde{S}_n(2^{2n}t) - 2^{-n}\tilde{S}_n(2^{2n}s) \tag{6.23}$$

Consider the integers $j_n = \lfloor 2^{2n}t \rfloor$ and $i_n = \lfloor 2^{2n}s \rfloor$, and $j_n \geq i_n$. Then 6.23 differs from

$$2^{-n}(\tilde{S}_n(j_n) - \tilde{S}_n(i_n)) = 2^{-n}\sum_{i_n+1}^{j_n}\tilde{X}_k \tag{6.24}$$

by an error not greater than $2 \cdot 2^{-n} \leq \delta$. Also, $j_n - i_n$ differs from $2^{2n}(t-s)$ by at most 1. In particular $j_n - i_n \longrightarrow \infty$ as $n \longrightarrow \infty$. Using the Central Limit Theorem A.12, for n large enough so that al the previous holds, we have that for any real x'

$$\Phi(x') - \delta \leq P\{\frac{1}{\sqrt{j_n - i_n}}\sum_{i_n+1}^{j_n}\tilde{X}_k \leq x'\} \leq \Phi(x') + \delta \tag{6.25}$$

Taking an approximation of $\sqrt{j_n - i_n}$ by $2^n\sqrt{t-s}$, for n large enough we have:

$$1 - \delta < \frac{2^n\sqrt{t-s}}{\sqrt{j_n - i_n}} < 1 + \delta \tag{6.26}$$

Conbining the expressions 6.23- 6.26, we obtain that

$$\Phi\Big((1-\delta)\frac{x}{\sqrt{t-s}} - \delta\Big) - \delta \leq P\Big\{B_n(t) - B_n(s) \leq x\Big\} \leq \Phi\Big((1+\delta)\frac{x}{\sqrt{t-s}} + \delta\Big) + \delta.$$

This shows that when $\delta \to 0$, the distribution of $B_n(t) - B_n(s)$ is assymptotically normal with mean 0 and variance $t-s$ when $n \to \infty$. Hence, by (6.2), the distribution of $W(t) - W(s)$ is normal with the parameters desired, what proves ii).
In order to prove iii), the same construction done in (6.23)-(6.26) is valid for arbitrary $v > u \geq t > s \geq 0$ applied to any real numbers x, y such that:

$$P\Big\{W(t) - W(s) \leq x, W(v) - W(u) \leq y\Big\} = P\Big\{W(t) - W(s) \leq x\Big\} \cdot P\Big\{W(v) - W(u) \leq y\Big\}.$$

□

If suitably manipulated, Theorem 6.7 can be transformed into the following result:

Theorem 6.8. *On any bounded interval, the sequence* (B_m) *almost surely uniformly converges as* $m \to \infty$ *and the limit process is Brownian motion* W. *For any* $C > 1$, *and for any* $K > 0$ *and* $m \geq 1$ *such that* $K2^{2m} \geq N_2(C)$, *we have*

$$P\Big\{ \sup_{0 \leq t \leq K} |W(t) - B_m(t)| \geq 27CK_*^{1/4}(\log_* K)^{3/4} m^{3/4} 2^{-m/2} \Big\} \leq \frac{6}{1 - 4^{1-C}} (K2^{2m})^{1-C},$$

where $K_* = \max\{1, K\}$ *and* $\log_*(x) = \max\{1, \log(x)\}$

Now, using the Borel-Cantelli lemma we get that for any fixed $K > 0$ there is a constant $c_K = 28K_*^{1/4}(\log_* K)^{3/4}$ (taking $C = 1 + \frac{1}{27}$, say) such that almost surely,

$$\limsup_{m\to\infty} \quad m^{-3/4} 2^{m/2} \sup_{0 \leq t \leq K} |W(t) - B_m(t)| < c_K \tag{6.27}$$

Similarly, for any fixed $m \geq 1$, there is a constant $c_m = 55m^{3/4}2^{-m/2}$ such that, almost surely,

$$\limsup_{K\to\infty} \quad K^{-1/4}(\log(K))^{-3/4} \sup_{0 \leq t \leq K} |W(t) - B_m(t)| < c_m \tag{6.28}$$

Since $K^{1/4}(\log(K))^{3/4}$ and for any $\omega \in \Omega$, $\sup_{0 \leq t \leq K} |W(t) - B_m(t)|$ are non-decreasing, it is enough to see that:

$$\limsup_{K\to\infty} \quad n^{-1/4}(\log(n))^{-3/4} \sup_{0 \leq t \leq n+1} |W(t) - B_m(t)| < c_m \tag{6.29}$$

when n only takes integer values. To prove it, it suffices to apply the Borel-Cantelli Lemma with $C = 2 + \frac{1}{27}$ in the previous theorem.

As a conclusion of this part of the section, we have developed an elementary construction of random walks that converge to Brownian motion. Although the construction might seem simple apparently, there are many technical details needed to prove the convergence. It might be possible to approach Brownian motion through other simple, symmetric random walks sequences, but this particular construction will allow us in the following sections to tackle other classical problems using these approximations.

7. Brownian local time

How can we measure the time that a Brownian motion or a more general stochastic process spends at a certain level or state? It is possible to show that this occupation times are absolutely continuous measures, so that their densities are a viable measure for the time spend at a level a up to time t or equivalently, in the time interval $[0, t]$. This section will lead us to an important result, the so-called *Trotter's theorem*, which will not only be valuable by its central role in local time theory of Brownian motion, but also because a non-typical argument, so far, will be provided as the main tool to proceed with the proof. We will use random walk approximations to a Brownian motion. The approximation will be done via the *twist and shrink* construction seen in section 6. In this section the results will be basically the ones given by [8], and for the last part of it, we will follow results already obtained by [9].

7.1 Brownian local time at zero

We start considering a Brownian motion $\{B(t) : t \geq 0\}$. Our mission is to measure the amount of time spend at 0. It is possible to prove that the zero set has $\frac{1}{2}$ *Haussdorf dimension*, e.g. see in [8], but the proof is out of the scope of this project. Anyway, the Haussdorf dimension of this set does not provide nontrivial information.
One way to tackle local time problem is through the study of downcrossings of a sequence of nested subintervals. More precisely, given $a < b$, and a Brownian motion $\{B(t) : t \geq 0\}$, we define $\tau_0 = 0$ and for $j \geq 1$:

$$\sigma_j = inf\{t > \tau_{j-1} : B(\sigma_j) = b\} \quad \text{and} \quad \tau_j = inf\{t > \sigma_j : B(\sigma_j) = a\}$$

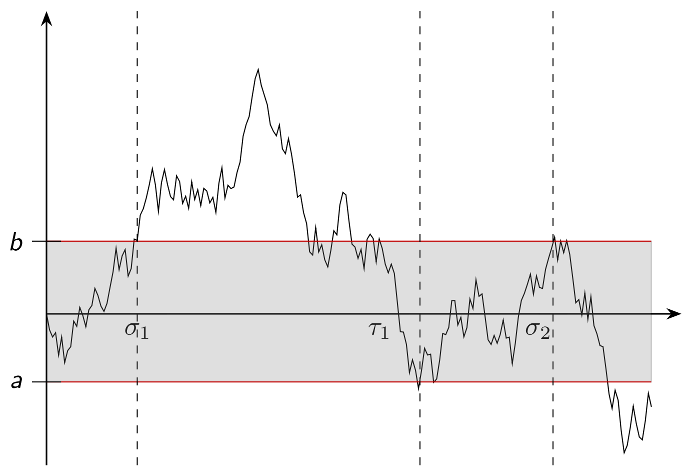
Then, we to define the downcrossings:

$$B^{(j)} : [0, \tau_j - \sigma_j] \to \mathbb{R}, \qquad B^{(j)}(s) = B(\tau_j + s)$$

for the j-th downcrossing of the interval $[a, b]$, and for the number of downcrossings, we donote:

$$D(a, b, t) = max\{j \in \mathbb{N} : \tau_j \leq t\}$$

From the definition, $D(a, b, t)$ is almost surely finite, as Brownian motion is absolutely continuous on any compact interval $[0, t]$. The final result we want to prove is the following one.



Theorem 7.1. (Downcrossing representation of the local time at zero) *There exists a nontrivial stochastic process* $\{L(t) : t \geq 0\}$ *called* **the local time at zero** *such that for any sequences* $a_n \uparrow 0$ *and* $b_n \downarrow 0$*, almost surely*

$$L(t) := \lim_{n\to\infty} 2(b_n - a_n)D(a_n, b_n, t)$$

for every $t \geq 0$*. Moreover this process is* γ*-Hölder continuous for any* $\gamma < \frac{1}{2}$

This theorem is the objective of the chapter. Some technical details must be introduced and developed.

Lemma 7.2. *Let* $a < m < b$ *and suppose* $\{B(t) : 0 \leq t \leq T\}$ *is a Brownian motion stopped at time* T *when it first hits a given level above* b*. Let :*

- D *be the number of downcrossings of the interval* $[a, b]$
- D_l *be the number of downcrossings of the interval* $[a, m]$
- D_u *be the number of downcrossings of the interval* $[m, b]$

Then, there exist two sequences of random variables $X_0, X_1, \dots$ *and* $Y_0, Y_1, \dots$ *of independent non negative random variables and independent of* D*, such that for* $j \geq 1$*:*

- $X_j \sim Geom(\frac{(b-a)}{(m-a)})$
- $Y_j \sim Geom(\frac{(b-a)}{(b-m)})$

and the following equalties hold:

$$D_l = X_0 + \sum_{j=1}^{D} X_j \qquad \text{and} \qquad D_u = Y_0 + \sum_{j=1}^{D} Y_j$$

For a proof, see e.g. [8, Lemma 6.3] based on [8, Theorem 2.45].

Definition 7.3. A stochastic process $\{X_n : n \geq 0\}$ is a **martingale** with respect to a filtration $(\mathcal{F}_n : n \geq 0)$ if:

- X_n is measurable with respect to $\mathcal{F}_n$
- $\mathbb{E}|X_n| < \infty$
- $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$

Additionally, if $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geq X_n$, then $\{X_n : n \geq 0\}$ is a **submartingale** and if $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n$, then $\{X_n : n \geq 0\}$ is a **supermatingale**.

Lemma 7.4. *For any two sequences $a_n \uparrow 0$ and $b_n \downarrow 0$ with $a_n < b_n$, the discrete stochastic process*

$$\{2(b_n - a_n)D(a_n, b_n, T), n \in \mathbb{N}\}$$

is a submartingale with respect to the filtration $(\mathcal{F}_n : n \in \mathbb{N})$.

Proof. First of all, we suppose without loss of generality that in each step:

$$\text{either} \quad a_{n+1} = a_n \quad \text{or} \quad b_{n+1} = b_n$$

This can be assumed, because in case neither of them were true in the n-th. step, this step could be split in two different steps, one step for each sequence, whilst the other remains constant.
Suppose $a_{n+1} = a_n$.Using Lemma 7.2 for D_l, the number of downcrossings $D(a_n, b_{n+1}, T)$ given $\mathcal{F}_n$, is the sum of $D(a_n, b_n, T)$ independent geometric random variables with parameter $\frac{b_{n+1}-a_n}{b_n-a_n}$ plus a nonnegetive factor. Thus:

$$\begin{aligned} \mathbb{E}(2(b_{n+1} - a_{n+1})D(a_{n+1}, b_{n+1}, T)) &\overset{a_{n+1}=a_n}{=} \\ &= 2(b_{n+1} - a_n)\mathbb{E}(D(a_n, b_{n+1}, T)) \\ \geq 2(b_{n+1} - a_n)\frac{b_n - a_n}{(b_{n+1} - a_n)}D(a_n, b_n, T) &= 2(b_n - a_n)D(a_n, b_n, T) \end{aligned} \tag{7.1}$$

Using again Lemma 7.2 for D_u, we have that the same relation holds for the case $b_{n+1} = b_n$, hence the result follows. ☐

Lemma 7.5. *Fix $T_b \in \mathbb{R}$ large enough. For any two sequences $a_n \uparrow 0$ and $b_n \downarrow 0$ with $a_n < b_n$ the limit*

$$L(T_b) = \lim_{n\to\infty} 2(b_n - a_n)D(a_n, b_n, T_b)$$

exists almost surely and does not depend on the choice of sequences.

Proof. $D(a_n, b_n, T_b)$ is a geometric random variable with parameter $\frac{b_n-a_n}{b-a_n}$, hence, its variance can be bounded by $2(\frac{b-a_n}{b_n-a_n})^2$. Thus,

$$\mathbb{E}((2(b_n - a_n)D(a_n, b_n, T_b))^2) \leq \mathbb{E}(2(b_n - a_n)D(a_n, b_n, T_b))^2 \leq 4(b - a_n)^2$$

Thus the submartingale in lemma 7.4 is L^2-bounded and by the submartingale convergence theorem (see [8]), the limit

$$\lim_{n\to\infty} 2(b_n - a_n)D(a_n, b_n, T_b)$$

exists almost surely, and it is nontrivial, where T_b is a stopping time with $b > b_1$.

It just remains to prove that $L(T_b)$ does not depend on the choice of a_n and b_n, what is also true, as if we were given sequences with different limits, we could construct a sequence of alternating intervals that would lead to a limiteless sequence, what can not happen, by this theorem. ☐

Lemma 7.6. *For any fixed time $t \geq 0$, the limit*

$$L(t) = \lim_{n\to\infty} 2(b_n - a_n)D(a_n, b_n, t) \qquad \textit{exists}$$

Proof. We define an auxiliary Brownian motion $\{B_t(s) : s \geq 0\}$ such that $B_t(s) = B(t+s)$. Also, for any integer $b > b_1$, we denote $D_t(a_n, b_n, T_b)$ the number of downcrossings of the interval $[a_n, b_n]$ by the auxiliary Brownian motion before it hits b. Then, by previous lemma,

$$L(T_b) = \lim_{n\to\infty} 2(b_n - a_n)D_t(a_n, b_n, T_b)$$

exists almost surely. For $t \geq 0$, we fix a Brownian motion such that this limits exists for all $b > b_1$. Then, picking b so large that $T_b > t$, define:

$$L(t) := L(T_b) - L_t(T_b)$$

Now, by considering the process suitably to avoid t coinciding with a downcrossing time:

$$D(a_n, b_n, T_b) - D_t(a_n, b_n, T_b) - 1 \leq D(a_n, b_t, t) \leq D(a_n, b_n, T_b) - D_t(a_n, b_n, T_b)$$

Then, multiplying by $2(b_n - a_n)$ and taking limits, we get $L(t) = L(T_b) - L_t(T_b)$ for both inequalties, which proves the convergence to $L(t)$. □

The next Lemma, which is stated without proof, will give an estimate on the infinitessimal growth of local times. For a proof, see e.g. [8, Lemma 6.7].

Lemma 7.7. *Let $\gamma < \frac{1}{2}$ and $0 < \epsilon < \frac{1-2\gamma}{3}$. Then for all $t \geq 0$ and $0 < h < 1$*

$$\mathbb{P}\{L(t+h) - L(t) > h^\gamma\} \leq 2e^{-\frac{1}{2}h^{-\epsilon}}$$

Lemma 7.8. *Almost surely,*

$$L(t) = \lim_{n\to\infty} 2(b_n - a_n)D(a_n, b_n, t)$$

exists for every $t \geq 0$.

Proof. We proof it for all $0 \leq t \leq 1$. Consider the grid:

$$\mathcal{G} = \bigcup_{m\in\mathbb{N}} \mathcal{G}_m \cup \{1\}, \quad \text{where } \mathcal{G}_m = \Big\{\frac{k}{m}, k \in \{0, 1, \dots, m-1\}\Big\}.$$

We show that convergence holds in the set:

$$\bigcup_{t\in\mathcal{G}} \{L(t) \text{ exists }\} \cup \bigcup_{M<m} \bigcup_{t\in\mathcal{G}_m} \Big\{L(t + \frac{1}{m}) - L(t) > \Big(\frac{1}{m}\Big)^\gamma\Big\},$$

which has probability close to one by the previous Lemmas, for M large enough. Given $t \in [0, 1)$, we find $t_1, t_2 \in \mathcal{G}_m$ with $t_2 - t_1 = \frac{1}{m}$. We have

$$2(b_n - a_n)D(a_n, b_n, t_1) \leq 2(b_n - a_n)D(a_n, b_n, t) \leq 2(b_n - a_n)D(a_n, b_n, t_2).$$

Both of sides converge to $L(t_1)$ and $L(t_2)$ respectively, and consequently, $L(t)$ exists, since $L(t_2) - L(t_1) \leq m^{-\gamma}$. Thus, the proof is complete. □

Lemma 7.9. *For $\gamma < \frac{1}{2}$, almost surely, the process $\{L(t) : t \geq 0\}$ is γ-Hölder continuous.*

Proof. Again, we only proof it for $0 \leq t \leq 1$, without loss of generality. We reconsider the set of the previous lemma and see that γ-Hölder continuity still holds in this set, that is, almost surely. Take $0 \leq s < t < 1$ and $t - s < \frac{1}{M}$, then pick $m > M$ such that

$$\frac{1}{m+1} \leq t - s < \frac{1}{m}.$$

Now, just taking $t_1 < s$ such that $t_1 \in \mathcal{G}_m$ and $s - t_1 < \frac{1}{m}$ and $t_2 > t$ with $t_2 \in \mathcal{G}_m$ and $t_2 - t < \frac{1}{m}$, note that:

$$L(t) - L(s) \leq L(t_2) - L(t_1) \leq 2\Big(\frac{1}{m}\Big)^{\gamma} \leq 2\Big(\frac{m+1}{m}\Big)^{\gamma}(t-s)^{\gamma} \leq (t-s)^{\gamma}$$

□

These lemmas jointly, complete the proof of the downcrossing representation Theorem.

7.2 A random walk approach to the local time process

The previous section will be the basis over which we contruct a more general idea of Brownian local time. For a given linear Brownian motion $\{B(t) : t \geq 0\}$, let $\{L^a(t) : a \in \mathbb{R}, t \geq 0\}$ be the local time at zero of the auxiliary Brownian motion $\{B^a(t) : t \geq 0\}$ such that $B^a(t) = B(t) - a$. The aim of this section is to study $\{L^a(t) : a \in \mathbb{R}, t \geq 0\}$, i.e., the local time as a function of the level a. We give the following result without proof:

Theorem 7.10. *For a Brownian motion $\{B(t) : t \geq 0\}$ and for any measurable function $g : \mathbb{R} \to \mathbb{R}$ and $t \geq 0$:*

$$\int g(a) d\mu_t(a) = \int_0^t g(B(s)) ds = \int_{-\infty}^{\infty} g(a) L^a(t) da$$

This theorem, in particular, implies that $\{L^a(t) : a \in \mathbb{R}\}$ is the density of the occupation measure μ_t, what can be defined as:

$$\mu_t(A) := \int_0^t \mathbb{1}_A(B(s)) ds$$

This measure has the property that is absolutely continuous with respect to the Lebesgue measure. The result that must be shown is the continuity of the density $\{L^a(t) : a \in \mathbb{R}\}$. In order to study it, we need to extend the idea of the downcrossing representation theorem, and in this case, we will use a random walk embedded in a Brownian motion. But, this ones must be done carefully, this is why we pay especial attention, as this argument or similars will be used more than once to replicate other Brownian motion results through random walks embeddings, which is one the main goals of this work. The embedding we will use is the following: we define the stopping times

$$\tau_k := \tau_k^{(n)} := inf\{t > t_{k-1} : |B(t) - B(\tau_{k-1})| = 2^{-n}\}$$

and define the n-th. embedded random walk by:

$$X_k := X_k^{(n)} := 2^n B(\tau_k^{(n)})$$

For any time $t > 0$, the length of the n-th. random walk is defined by:

$$N := N^{(n)}(t) := max\{k \in \mathbb{N} : t \leq \tau_k\}$$

Given $a \in \mathbb{R}$, we can choose $j(a) \in \{0, 1, \dots\}$ such that $j(a)2^{-n} \leq a < (j(a)+1)2^{n}$. Then, we denote the number of downcrossings of $2^n a$ by:

$$D^{(n)}(a,t) := \#\{k \in \{0,1,\dots,N^{(n)}(t)\} : X_k^{(n)} = j(a)+1, X_{k+1}^{(n)} = j(a)\}$$

Theorem 7.11. (Trotter's theorem) *Let $\{B(t) : t \geq 0\}$ be a Brownian motion and let $D^{(n)}$ be the number of downcrossings of $2^n a$ by the n-th. embedded random walk stopped at time $N^{(n)}(t)$. Then, almost surely,*

$$L^a(t) := \lim_{n\to\infty} 2^{-n+1} D^{(n)}(a,t) \qquad \textit{exists for all } a \in \mathbb{R} \textit{ and } t \geq 0$$

Moreover, for $\gamma < \frac{1}{2}$ the process

$$\{L^a(t) : a \in \mathbb{R}, t \geq 0\}$$

is almost surely locally γ-Hölder continuous.

Remark 7.12. $\{L^a(t) : a \in \mathbb{R}, t \geq 0\}$ is usally considered a random field, instead of a process, as it depends on more than one parameter, but the random nature of the process is conserved.

The proof of this theorem is especially important and relevant for this work, since it aims at proving a classical result of Brownian motion using random walk approximation, but in this case, without using the approach developed in the previous chapter. Actually, there is a more general result, the *Ray-Knight theorem*, which will not be included, but can be found in [8] in which this idea is also used.

Fix $\gamma \leq \frac{1}{2}$ and a large enough integer N, that determines the stopping time of the Brownian motion at T_N when it hits the level N. Consider the abbreviation $D^{(n)}(a) = D^{(n)}(a, T_N)$, then:

Lemma 7.13. *Denote by $\Omega(m)$ the event that for every $n \geq m$:*

a) $|D^{(n)}(a) - \frac{1}{2}D^{(n+1)}(a)| \leq 2^{n(1-\gamma)}$ *for all $a \in [-N, N)$*

b) $|D^{(n)}(a) - D^{(n)}(b)| \leq 2 \cdot 2^{n(1-\gamma)}$ *for all $a, b \in [-N, N)$ such that $|a-b| \leq 2^{-n}$*

Then,

$$\lim_{m\to\infty} P(\Omega(m)) = 1$$

Thus, if we are capable of proving the results for the elements satisfying the conditions of $\Omega(m)$, then the results will be valid almost surely, as a consequence of the previous Lemma.

Lemma 7.14. *On the set $\Omega(m)$ we have that:*

$$L^a(T_N) := \lim_{n\to\infty} 2^{-n-1} D^{(n)}(a)$$

for every $a \in [-N, N)$

Proof. The easiest way to proof this lemma is by showing that the sequence $2^{-n-1}D^{(n)}(a)$ is a Cauchy sequence and hence, convergent. By definition, on $\Omega(m)$, for any $a \in [-N, N)$, and $n \geq m$,

$$|2^{-n-1}D^{(n)}(a) - 2^{-n-2}D^{(n+1)}(a)| \leq 2^{-n\gamma}$$

Thus, we have:

$$\sup_{k\geq n} |2^{-n-1}D^{(n)}(a) - 2^{-k-1}D^{(k)}(a)| \leq \sum_{k=n}^{\infty} |2^{-k-1}D^{(k)}(a) - 2^{-k-2}D^{(k+1)}(a)| \leq \sum_{k=n}^{\infty} |2^{-n\gamma}| \overset{n\to\infty}{\longrightarrow} 0$$

□

Lemma 7.15. *On the set $\Omega(m)$, the process $\{L^a(T_N), a \in [N, N)\}$ is γ-Hölder continuous.*

Proof. Fix $a, b \in [-N, N)$ with $2^{-n-1} < |a-b| < 2^{-n}$ for some $n \geq m$. Then, for all $k \geq n$,

$$|2^{-k-1}D^{(k)}(a) - 2^{-k-1}D^{(k)}(b)| \leq |2^{-n-1}D^{(n)}(a) - 2^{-n-1}D^{(n)}(b)|$$
$$+ \sum_{j=n}^{k-1} |2^{-j-2}D^{(j+1)}(a) - 2^{-j-1}D^{(j)}(a)| + \sum_{j=n}^{k-1} |2^{-j-2}D^{(j+1)}(b) - 2^{-j-1}D^{(j)}(b)|$$
$$\leq 2^{-n\gamma+1} + 2\sum_{j=n}^{\infty} 2^{-j\gamma}$$

Now, letting $k \uparrow \infty$,

$$|L^a(T_N) - L^b(T_N)| \leq (2 + \frac{2}{1-2^{-\gamma}})2^{-n\gamma} \leq (2^{1+\gamma} + \frac{2^{1+\gamma}}{1-2^{-\gamma}})|a-b|^\gamma$$

thus, $\{L^a(T_N), a \in [N, N)\}$ is γ-Hölder continuous, as we wanted.

□

Lemma 7.16. *For any fixed $t > 0$, almost surely, the limit*

$$L^a(t) := \lim_{n\to\infty} 2^{-n-1}D^{(n)}(a)$$

exists for any $a \in \mathbb{R}$, and $\{L^a(T_N), a \in \mathbb{R}\}$ is γ-Hölder continuous.

Proof. In this lemma, we define an auxiliary Brownian motion $\{B_t(s) : s \geq 0\}$ by $B_t(s) = B(t+s)$ and the number of downcrossing of the auxiliary process $D_t^{(n)}(a)$. Thus, by the previous Lemma, almost surely, the limit $L_t^a(T_N) := \lim_{n\to\infty} 2^{-n-1}D_t^{(n)}(a)$ exists for all $a \in \mathbb{R}$ and for all N. Taking N large enough so that $T_N > t$, define: $L^a(t) := L^a(T_N) - L_t^a(T_N)$, which is γ-Hölder continuous and:

$$D^{(n)}(a, T_N) - D_t^{(n)}(a, T_N) - 1 \leq D^{(n)}(a, t) \leq D^{(n)}(a, T_N) - D_t^{(n)}(a, T_N)$$

Then, multiplying by 2^{-n-1} and taking limits, we have the convergence desired. □

Theorem 7.17. *Almost surely,*

$$L^a(t) := \lim_{n\to\infty} 2^{-n-1}D^{(n)}(a)$$

exists for all $t > 0$ and for every $a \in \mathbb{R}$, and $\{L^a(T_N), a \in \mathbb{R}\}$ is γ-Hölder continuous.

The complete proof of this Theorem can be found in [8, Lemma 6.26]. This ends the proof of Theorem 7.11.The most important fact of it is that it ensures the well-definition and non-vanishing property of the process. This can be surprising as the level sets have zero Lebesgue measure, but taken to limit it results in such process.

7.3 Local time approach through *twist and shrink* random walks

Retriving the *twist and shrink* construction of random walks that converge to the Wiener process, we aim to give the approximation of random walks to the local time process. So far, Brownian local time has been looked from a downcrossing and upcrossing point of view, but this must be refined in order to define the successive random walks.
We define the local time of the random walks $\tilde{S}_m(k)$ at a point $x \in \mathbb{Z}$ at time $k \in \mathbb{B}$ as $l_m(0,x) = 0$ and

$$l_m(k,x) = \#\{j : 0 \leq j < k, \tilde{S}_m(j) = x\}$$

The local time of the m-th approximation B_m at a point $x \in 2^{-m}\mathbb{Z}$ at time $t \in 2^{-2m}\mathbb{N}$ is defined as: $L(t,x) = 2^{-m} l_m(t2^{2m}, x2^m)$. The extension of $L(t,x)$ to arbitrary $t \in \mathbb{R}_+$ and $t \in \mathbb{R}$ is done by linear interpolation, resulting in a continuous process. One can define the *one-sided up and down local times* $l^{\pm}(k,x)$ in the following way:

$$l^{\pm}_m(t,x) = \#\{j : 0 \leq j \leq k : \tilde{S}_m(j) = x, \tilde{S}_m(j+1) = x \pm 1\} \quad (k \geq 1)$$

As we know that the distribution of the local time $\tilde{l}$ of the simple, symmetric random walks, it follows that:

$$P\{\tilde{l}(2k) = j\} = P\{\tilde{l}(2k+1) = j\} = \frac{1}{2^{2k-j}}\binom{2k-j}{k}$$

Hence, the Moivre-Laplace theorem gives the assymptotic approximation of its distribuition:

$$P\{\tilde{l}(2k) = j\} \sim \sqrt{\frac{2}{\pi k}} exp\Big(-\frac{j^2}{2k}\Big)$$

whenever $k \to \infty$ and for any $0 \leq j \leq K_k = o(k^{2/3})$. Then, for any sequence $u_k \to \infty$, such that $u_k = o(j^{1/6})$. We obtain the following inequalty:

$$P\Big\{\frac{l^{\pm}(k,x)}{\sqrt{k}} \geq u_k\Big\} \leq P\Big\{\frac{l(k,x)}{\sqrt{k}} \geq u_k\Big\} \leq exp\Big(-\frac{u_k^2}{2}\Big), \tag{7.2}$$

if k is large enough, i.e.,$k \geq k_0$, for some k_0.

Lemma 7.18. *For any $C > 1$ and for any $K > 0$ and $m \geq 1$such that $K2^{2m} \geq N(C)$, we have:*

$$P\Big\{\sup_{j\in\mathbb{Z}}\sup_{0\leq t_k\leq K}|L_{m+1}(t_k,x_j) - L_m(t_k,x_j)| \geq 6CK_*^{1/4}(log_*(K))^{3/4}m^{3/4}2^{-m/2}\Big\} \leq 12(K2^{2m})^{1-C}$$

where $t_k = k2^{-2m}$ and $x_j = 2^{-m}$.

Lemma 7.19. *For any $C > 1$ and for any $K > 0$ and $m \geq 1$ such that $K2^{2m} \geq N(C)$, we have:*

$$P\Big\{\sup_{r\geq 1}\sup_{(t,x)\in[0,K]x\mathbb{R}}|L_{m+r}(t,x) - L_m(t,x)| \geq 79CK_*^{1/4}(log_*(K))^{3/4}m^{3/4}2^{-m/2}\Big\} \leq \frac{15}{1-4^{1-C}}(K2^{2m})^{1-C}$$

where $K_* = max\{1,K\}$.

Theorem 7.20. *On any strip $[0,K]x\mathbb{R}$ the sequence $(L_m(t,x))$ almost surely converges as $m \to \infty$ and the limit process $L(t,x)$ is jointly continuous in (t,x), the local time of the Brownian motion $W(t)$.*

Proof. It is a direct consequence of the previous Lemma, we just need to apply the Borel-Cantelli Lemma.
Thus, there is uniform convergence to $L(t,x)$ and, since the functions are continuous, the limit is also jointly continuous in (t,x) □

8. Approximation to the Black-Scholes model

In financial mathematics, if there has been a breakthrough in recent times, this is the model used in option pricing introduced by Black, Scholes and Merton in 1973, which is outlined in [2], was awarded with an Economics Nobel prize. It definitely gave an expression which both counterparts should agree with to buy and sell options on an underlying, what was a revolution at that point. Naturally, further studies were done to to exploit the via ignited, but we will not go deeper in details in this work. In this chapter we aim at giving a brief description of the model, its assumptions and a strong discrete approximation of it using suitable nested sequence of simple, symmetric random walks. This approximation can be extended to stock prices, replicating protfolios and the greeks. However, we will only focus on the model, although some comments will be necessarily done to reference other applications.

We aim at finding the formula for the price of basic financial instruments such as European calls and puts. The pricing must be objective, meaning that both counterparts would agree to buy and sell, respectively, knowing that a fair price has been set for the option. To this extent, *fair* must be contextualized: it would be absolutely nonsense to give a ubiased price without the risk-neutral probabilities, that is, neglecting any market view subject to uncertain considerations. Under this measure, we know the discounted prices and the value of a portfolio with risky and riskless assets behave as a martingale with respect to a filtration.

We will construct a random walk approximation that leads us to the same results obtained using the Black-Scholes model. In other words, using random walks will allow us replicate some results, which will only be a discrete version of the model, which taken to limit will retieve it most known formulas for european options.

8.1 Remark on changes of measure and the Cameron-Martin-Girsanov theorem

Consider a space $(\Omega, \mathcal{A})$, where $\mathcal{A}$ is a σ-algebra and two equivalent probability measures on $\mathcal{A}$, $\mathbb{P}, \mathbb{Q}$. Let $W(t)$ be a Brownian motion under $\mathbb{P}$. How does $W(t)$ looks for $\mathbb{Q}$? Will the change of measure affect $W(t)$? The relation between the measures is given by the Radon-Nikodym derivative:

Definition 8.1. Suppose $\mathbb{P}$ and $\mathbb{Q}$ two equivalent measures, that is, they produce the same null-measure sets. Given a random process $X : \Omega x[0, \infty) \longrightarrow \mathbb{R}$ adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Let ω be a path and an ordered mesh $\{t_1, t_2, \dots t_n\}$, we define x_i to be $W(t_i, \omega)$ and the *Radon-Nikodym derivative* $\frac{d\mathbb{Q}}{d\mathbb{P}}$ up to time $t_n = T$ is defined to be the limit of the liklihood ratios:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \lim_{n\to\infty} \frac{f^n_{\mathbb{Q}}(x_1, \dots, x_n)}{f^n_{\mathbb{P}}(x_1, \dots, x_n)}$$

as the mesh becomes dense in $[0, T]$. This continuous time derivative satisfies:

(a) $\mathbb{E}_{\mathbb{Q}}(X_T) = \mathbb{E}_{\mathbb{P}}(\frac{d\mathbb{Q}}{d\mathbb{P}} X_T)$

(b) $\mathbb{E}_{\mathbb{Q}}(X_t|\mathcal{F}_s) = \zeta_s^{-1}\mathbb{E}_{\mathbb{P}}(\zeta_t X_t|\mathcal{F}_s)$

where $\zeta_t = \mathbb{E}_{\mathbb{P}}(\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t)$, which is called the *Radon-Nikodym process*.

The example we are interested in is given by the following definition of the derivative:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = exp\{-\gamma W_T - \frac{1}{2}\gamma^2 T\}$$

By using the moment generating function of normal random variables or an identification criteria it is possible to prove that, under the probability $\mathbb{Q}$, $W(t)$ is a Brownian motion with constant drift $-\gamma$. Hence, defining

$$\tilde{W}(t) = W(t) + \gamma t$$

we have that the process $\tilde{W}(t)$ is a Brownian motion under $\mathbb{Q}$.
The generalization of the generation of new Brownian motions under different measures using a drift is given in [4], and essentially the theorem that underpins the previous results is:

Theorem 8.2. ***(Cameron-Martin-Girsanov)*** *If $W(t)$ is a $\mathbb{P}$-Brownian motion and γ_t is a $\mathcal{F}_t$ measurable function satisfying the condition $\mathbb{E}_{\mathbb{P}}\Big(exp(\frac{1}{2}\int_0^T \gamma_t^2 dt)\Big) < \infty$, then there exists a measure $\mathbb{Q}$ such that:*

a) $\mathbb{Q}$ *is equivalent to* $\mathbb{P}$

b) $\frac{d\mathbb{Q}}{d\mathbb{P}} = exp\Big(-\int_0^T \gamma_t dW_t - \frac{1}{2}\int_0^T \gamma_t^2 dt\Big)$

c) $\tilde{W}(t) = W(t) + \int_0^t \gamma_s ds$ *is a* $\mathbb{Q}$*-Brownian motion.*

This theorem will be very useful in our case. We are interested in finding a measure that ensures that a certain process, which will be detailed later, is a martingale with respect to a filtration. Therefore, using this theorem we will have this certainty, and in addtion, Brownian motion in the model will be drifted suitably, exactly with the drift given by the Cameron-Martin-Girsanov theorem, with a known parameter γ, so the right Brownian motion is obtained.

8.2 Introduction to the Black-Scholes Model

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, and let $(B(t))_{t\geq 0}$, $B(0) = 0$ be a Brownian motion, and denote the generated filtration by the stochastic process by $(\mathcal{F}_t)_{t\geq 0}$. We consider a portfolio with two type of assets: risky and risk-free assets (bonds), that sum up the values of the portfolio at time t given by:

$$V(t) := a(t)S(t) + b(t)\beta(t) \tag{8.1}$$

where $a(t)$ and $b(t)$ are deterministic known processes, with no intrinsic random component. $S(t)$ denotes the price of the risky asset, which the model assumes to be a geometric random walk:

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t), \qquad S(0) = S_0 > 0 \tag{8.2}$$

$(\mu \in \mathbb{R}, \sigma > 0)$ and denote $\beta(t)$ the price of the riskless asset:

$$d\beta(t) = r\beta(t)dt, \qquad \beta(0) = \beta_0 > 0$$

In general, r could be time-dependent $r = r(t)$ but in this case, for simplicity, we consider it constatnt, so that the previous equation has the following solution, if $\beta_0 = 1$: $\beta(t) = e^{rt}$. Moreover $S(t)$ is the unique solution of the Stochastic Differential Equation 8.2:

$$S(t) = S_0 e^{((\mu - \frac{\sigma^2}{2})t + \sigma B(t))}. \tag{8.3}$$

It is also assumed that the portfolio is *self-financing*, that is, the change of its value throughout time is only the consequence of the change of the value of the assets, and any quantity of money can not be deposited or withdrawn:

$$dV(t) = a(t)dS(t) + b(t)d\beta(t). \tag{8.4}$$

One of the most important assumptions underpinning the model is the absence of arbitrage or *free-lunch* operations. This does not happen in real world, but it is key to assume non-arbitrage markets and at the same time, not far from reality, as there are not many arbitrage opportunities and they are necessary for market corrections. This assumption leads to the so-called *risk-neutral probability* or pricing, which is a probability measure that ensures absence of arbitrage and, what is more important, that the discounted prices are martingales with respect to the filtration $\mathcal{F}_t$.
This mearsure on a fixed interval $[0, T]$, expressed in terms of the Radon-Nikodym derivative is:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := exp\Big(\frac{r-\mu}{\sigma}B(T) - \frac{1}{2}\Big(\frac{r-\mu}{\sigma}\Big)^2 T\Big). \tag{8.5}$$

Under the probability $\mathbb{Q}$, we have that $S(t)$ satisfies the SDE

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \qquad S(0) = S_0 > 0 \tag{8.6}$$

where

$$W(t) := B(t) + \frac{\mu - r}{\sigma}t, \tag{8.7}$$

is a $\mathbb{Q}$-Brownian motion, since the condition for $\gamma_t = \frac{\mu - r}{\sigma}$ in Theorem 8.2 holds automatically, given it is a constant function, and the drifted random walk $W(t)$ solves the SDE:

$$S(t) = S_0 e^{((\mu - \frac{\sigma^2}{2})t + \sigma W(t))}, \tag{8.8}$$

We will focus on pricing options whose payoff is of the form $g(S(T))$, where $T > 0$ is the maturity of the option and g must satisfy regularity conditions that will be especified later in this chapter. These are the simplest type of options, since they are neither path-dependent nor barrier options, which are not in general analytically solvable. For example, for European put options, $g(S(T)) = max\{K - S(T), 0\}$, where K is the strike price (right to buy/sell at this price at maturity). This gives a unique answer for the price at maturity, but what about the price at $t \in [0, T]$?
From the martingale property of the discounted prices of a portfolio under risk-neutral probability,the answer to the question is the $\mathbb{Q}$-expectation of the discounted claim of the option:

$$f(t, x) := \mathbb{E}_{\mathbb{Q}}\Big(e^{-(T-t)}g(S(T))|S(t) = x\Big). \tag{8.9}$$

Then, after long computations and observations, one obtains the expression for the price of an European call option:

$$C(t, x) = x\Phi(d_+(T - t, x)) - e^{-r(T-t)}K\Phi(d_-(T - t, x)), \tag{8.10}$$

where:

$$d_{\pm}(t, x) := \frac{1}{\sigma\sqrt{T - t}}\Big(log(\frac{x}{K}) + (r \pm \frac{\sigma^2}{2}(T - t))\Big) \quad . \tag{8.11}$$

Using the general *Itô rule* for functions of two or more variables and the self-financing condition on the portfolio, it is possible to deduce that $f(x, t)$ solves the Black-Scholes partial differential equation:

$$\partial_t f(t, x) + rx\partial_x f(t, x) + \frac{1}{2}\sigma^2 x^2 \partial_{xx} f(t, x) - rf(t, x) = 0, \tag{8.12}$$

with the final condition $f(T,x) = g(x)$. Here, $f(x,t)$ is required regulartiy conditions, at least two times differentiable in space, for instance, which actually, $C(t,x)$ will satisfy. Hence it will solve the PDE. This PDE can be transformed into the heat equation with some changes of variables, and thus, as a parabolic equation, initial (or final) conditions must be applied so that we get the price for the option. Note that in this case, the condition is set at maturity $t = T$, hence the equation has to be solved backward in time, as opposite to general methods for parabolic PDEs. The result that aims at being proved is:

Theorem 8.3. *Suppose that $g \in \mathcal{C}_c(\mathbb{R}_+)$ and $T > 0$. As $m \to \infty$, the price $f_m(t^{(m)},x)$ of the option whose payoff is $g(S_m(T))$ obtained by the above discrete approximation converges to its value $f(t,x)$ obtained by the Black-Scholes model, uniformly for $t \in [0,T]$ and $x > 0$:*

$$\lim_{m\to\infty} f_m(t^{(m)},x) = \lim_{m\to\infty} \mathbb{E}_{\mathbb{Q}_m}\Big(r_m^{\lfloor t2^{2m}\rfloor - \lfloor T2^{2m}\rfloor} g(S_m(T^{(m)}))|S_m(t^{(m)}) = x\Big) =$$
$$\mathbb{E}_{\mathbb{Q}}\Big(e^{-r(T-t)}g(S(T))|S(t) = x\Big) = f(t,x),$$

where $t^{(m)} = \lfloor t2^{2m}\rfloor 2^{-2m}$

Plainly, this theorem states that the classical formula for pricing the claim $f(t,x)$ used in BSM is still valid for approximations, in the limit. Thus the approximations taken will converge to the model.

8.3 Discrete random walk approximation of BSM

This is the main part of the section. We want to construct the approximation of Brownian motion using tools already seen in chapter 6, but towards the obtention of the change of measure that is required for an arbitrage-free derivative pricing, using the risk-neutral probability. From here, we will be able to obtain a general formula valid for any claim that will be applied, in particular, for the case of plain vanilla options, such as calls and puts.
In order to be clear in the approximations we need some notation. For every $m = 0,1,...$ we fix $\Delta t = 2^{-2m}$ and $\Delta x = 2^{-m}$ and $t_k = k\Delta t$. Random walks will be denoted by $B_m(t_k)$ and the time filtration used is denoted by $(\mathcal{F}^m_{t_k})_{t\geq 0}$.
Recall that from [9], we know that *twist and shrink* construction through random walks converge to a Brownian motion, and more precisely:

Theorem 8.4. *The sequence of twist and shrink random walks B_m uniformly converge to Brownian motion B on bounded intervals, almost surely. For all $T > 0$, as $m \to \infty$:*

$$\sup_{0\leq t\leq T} |B(t) - B_m(t)| = \mathcal{O}(m^{\frac{3}{4}}2^{-\frac{m}{2}}). \tag{8.13}$$

Let $a(t_k)$ and $b(t_k)$ be the processes defined in the beginning of the section, which should be measurable over the period $[t_k, t_{k+1})$ with respect to the filtration $\mathcal{F}^m_{t_k}$ for each $k \geq 1$. Thus the market value of a portfolio like in (8.1) is given by:

$$V_m(t_k) := a_m(t_k)S_m(t_k) + b_m(t_k)\beta_m(t_k). \tag{8.14}$$

Recall again that one of the assumptions of the model, that actually simplifies the solution of the SDE seen in 8.6, is that both μ_t (drift) and σ_t (volatility) are linear with respect the risky asset price S_t, that is in discrete notation: $\mu(t_k) = \mu S(t_k)$ and $\sigma(t_k) = \sigma S(t_k)$. Thus, increments of price of risky asset can be written as:

$$\Delta S_m(t_{k+1}) = \mu S_m(t_k)\Delta t + \sigma S_m(t_k)\Delta B_m(t_{k+1}), \quad S_m(t_0) = S_0 > 0. \tag{8.15}$$

Hence, we deduce the recursion of S_m in terms of time indexes:

$$S_m(t_{k+1}) = S_m(t_k)(1 + \mu\Delta t + \sigma\Delta B_m(t_{k+1})) \quad (k \geq 0), \tag{8.16}$$

where the Brownian motion should be interpreted as infinitessimal steps of a fair coin toss, i.e.,

$$X_m(t_{k+1}) := 2^m \Delta B_m(t_{k+1}) = \pm 1, \tag{8.17}$$

Lemma 8.5. *Denoting*

$$\tilde{S}_m(t) := S_0 exp\Big(\Big(\mu - \frac{\sigma^2}{2}\Big)t + \sigma B_m(t)\Big) \quad (t \geq 0),$$

we obtain for any $m \geq 0$ that:

$$\sup_{0 \leq t_k \leq T} |S_m(t_k) - \tilde{S}_m(t_k)| \leq c_1(\mu, \sigma, T)2^{-m}, \quad c_1 \geq 0,$$

where S_m is the solution of (8.16).

Proof. Using the Taylor expansion of the exponential function:

$$\begin{aligned}\Delta\tilde{S}_m(t_{k+1}) &:= \tilde{S}_m(t_{k+1}) - \tilde{S}_m(t_k) = \tilde{S}_m(t_k)\{exp((\mu - \frac{\sigma^2}{2})\Delta t + \sigma\Delta B_m(t_{k+1})) - 1\} \\ &= \tilde{S}_m(t_k)\Big\{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma\Delta B_m(t_{k+1}) + \frac{1}{2}((\mu - \frac{\sigma^2}{2})\Delta t + \sigma\Delta B_m(t_{k+1}))^2\Big\},\end{aligned}$$

where $0 < |t| < (\mu - \frac{\sigma^2}{2})2^{-2m} + \sigma 2^{-m}$. Thus,

$$|\Delta\tilde{S}_m(t_{k+1}) - \tilde{S}_m(t_k)(\mu\Delta t + \sigma\Delta B_m(t_{k+1}))| \leq C_1' 2^{-3m}$$

and recursively, it implies the statement of the lemma. □

For the riskless asset we will use the most natural approximation, given $r(t) = r > 0$ and $\beta_0 = 1$:

$$\beta_m(t_k) = (1 + r\Delta t)^k \quad \text{and} \quad \beta_m(t) = (1 + r\Delta t)^{\lfloor \frac{t}{\Delta t} \rfloor}, \quad (t \geq 0). \tag{8.18}$$

Then, it holds that for any $m \geq 0$:

$$\sup_{0 \leq t \leq T} |\beta(t) - \beta_m(t)| = \sup_{0 \leq t \leq T} |e^{rt} - (1 + r2^{-2m})^{\lfloor t2^{2m} \rfloor}| \leq c_2(r)2^{-2m}. \tag{8.19}$$

We have seen the approximation to the Brownian motion using the same *twist and shrink* construction than in the section 6. However, one of the key points that underpines the results is the use of the so-called *risk-neutral measure or probability*. We have already introduced it for the continous case for the filtration $\mathcal{F}_t$, but we need a definition for the discrete case of the filtration $\mathcal{F}^m_{t_k}$. Based on (8.16) and (8.18) we set:

$$r_m := 1 + r\Delta t = 1 + r2^{-2m}$$
$$u_m := 1 + \mu 2^{-2m} + \sigma 2^{-m}$$
$$d_m := 1 + \mu 2^{-2m} - \sigma 2^{-m},$$

and the risk neutral probabilities $q^{\pm}_m$ corresponding to the up or down move of the price of an underlying asset (or commonly explained through a coin toss whose outcomes are head or tail):

$$q^+_m = \frac{r_m - d_m}{u_m - d_m} = \frac{1}{2} + \frac{1}{2}\frac{r-\mu}{\sigma}2^{-m}$$

$$q^-_m = 1 - q^+_m = \frac{1}{2} - \frac{1}{2}\frac{r-\mu}{\sigma}2^{-m}.$$

Recall that the returns given in the discrete model are u_m with probability q^+_m and d_m with probability q^-_m. Then, we define the probability measure $\mathbb{Q}_m$ with its Radon-Nikodym derivative on the filtration $(\mathcal{F}^m_{t_k})_{k\geq 0}$:

$$\frac{d\mathbb{Q}_m}{d\mathbb{P}} = \Big(\frac{q^+_m}{\frac{1}{2}}\Big)^{\#\text{Heads(T)}}\Big(\frac{q^-_m}{\frac{1}{2}}\Big)^{\#\text{Tails(T)}}$$

Now, using that the number of heads/tails up to time T (sum of up/down moves) for a random walk in terms is given by the realtion of the Brownian motion that equals the difference of up/down moves. So:

$$\#\text{Heads(T)} = \frac{1}{2}(T2^{2m} + B_m(T)2^m) \quad \text{and} \quad \#\text{Tails(T)} = \frac{1}{2}(T2^{2m} - B_m(T)2^m)$$

Also, using the Taylor expansion of the logarithm, we obtain:

$$log(2q^{\pm}_m) = log(1 \pm \frac{r-\mu}{\sigma}2^{-m}) = \pm\frac{r-\mu}{\sigma}2^{-m} - \frac{1}{2}\Big(\frac{r-\mu}{\sigma}\Big)^2 2^{-2m} + \mathcal{O}(2^{-3m})$$

Finally, we obtain:

$$\frac{d\mathbb{Q}_m}{d\mathbb{P}} = \exp\Big\{\frac{r-\mu}{\sigma}B_m(T) - \frac{1}{2}\Big(\frac{r-\mu}{\sigma}\Big)^2 T + \mathcal{O}(2^{-m})\Big\} \tag{8.20}$$

Lemma 8.6. *a) The process*

$$\Lambda(t_k) := (2q^+_m)^{\frac{1}{2}(t_k 2^{2m} + B_m(t_k)2^m)}(2q^-_m)^{\frac{1}{2}(t_k 2^{2m} - B_m(t_k)2^m)} \tag{8.21}$$

is a positive $\mathbb{P}$-martingale with respect to $(\mathcal{F}^m_{t_k})_{k\geq 0}$ with expectation 1.

b) For the total variation distance between the probabilities $\mathbb{Q}_m$ and $\mathbb{Q}$, we have:

$$\lim_{m\to\infty} \sup_{A\in\mathcal{F}} |\mathbb{Q}_m(A) - \mathbb{Q}(A)| = 0 \tag{8.22}$$

c) If we consider a random walk B_m constructed with the twist and shrink process, plus a suitable drift:

$$W_m(t_k) := B_m(t_k) + \frac{\mu - r}{\sigma}t_k \tag{8.23}$$

then $W_m(t_k)$ is a $\mathbb{Q}_m$-martingale with respect to $(\mathcal{F}^m_{t_k})_{k\geq 0}$

d) Extending W_m by linear interpolation to arbitrary $t \in \mathbb{R}_+$, for any $T > 0$, we have:

$$\sup_{0\le t\le T} |W(t) - W_m(t)| = \mathcal{O}(m^{\frac{3}{4}} 2^{-\frac{m}{2}}) \tag{8.24}$$

Proof. a) First we need a little computation that will be used after:

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\Big((2q_m^+)^{\frac{1+X_m(t_{k+1})}{2}}(2q_m^-)^{\frac{1-X_m(t_{k+1})}{2}}\Big) &= (2q_m^+)\frac{1}{2} + ((2q_m^+)(2q_m^-))^{\frac{1}{2}}\frac{1}{2}\\ &= q_m^+ + (q_m^+ q_m^-)^{\frac{1}{2}} = \frac{1}{2} + \sqrt{\frac{1}{4}} = \frac{1}{2} + \frac{1}{2} + \mathcal{O}(2^{-m}) = 1,\end{aligned} \tag{8.25}$$

Now, using the expression (8.25) and $\Delta t = 2^{-m}$ and $\Delta B_m(t_k) = 2^{-2m}$:

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}(\Lambda_m(t_{k+1})|\mathcal{F}^m_{t_k}) &= \mathbb{E}_{\mathbb{P}}\Big((2q_m^+)^{\frac{(t_k+\Delta t_k)2^{-2m}+(B_m(t_k)+X_m(t_{k+1}))2^{-m}}{2}}(2q_m^-)^{\frac{(t_k+\Delta t_k)2^{-2m}+(B_m(t_k)+X_m(t_{k+1}))2^{-m}}{2}}\Big)\\ &= \Lambda_m(t_k)\mathbb{E}_{\mathbb{P}}\Big((2q_m^+)^{\frac{1+X_m(t_{k+1})}{2}}(2q_m^-)^{\frac{1-X_m(t_{k+1})}{2}}\Big) = \Lambda_m(t_k)\end{aligned} \tag{8.26}$$

since $\Lambda(t_k)$ is $\mathcal{F}^m_{t_k}$-measurable, and where $X_m(t_{k+1})$ is defined as in (8.17).

b) By Scheffé's theorem, e.g., see [1], it is enough to prove that $\frac{d\mathbb{Q}_m}{d\mathbb{P}}$ converge to $\frac{d\mathbb{Q}}{d\mathbb{P}}$, which is clear from (8.20).

c) It is enough to show that $\Lambda_m(t_k)W_m(t_k)$ is a $\mathbb{P}$-martingale with respect to $\mathcal{F}^m_{t_k}$:

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}(\Lambda_m(t_{k+1})W_m(t_{k+1})|\mathcal{F}^m_{t_k}) &= \mathbb{E}_{\mathbb{P}}\Big\{\Big(W_m(t_k) + X_m(t_{k+1})2^{-m} + \frac{\mu - r}{\sigma}2^{-2m}\Big)\\ &\quad \Lambda_m(t_k)(2q_m^+)^{\frac{1+X_m(t_{k+1})}{2}}(2q_m^-)^{\frac{1-X_m(t_{k+1})}{2}}|\mathcal{F}^m_{t_k}\Big\} = \Lambda_m(t_k)W_m(t_k)\end{aligned} \tag{8.27}$$

since $W_m(t_k)$ and $\Lambda(t_k)$ are $\mathcal{F}^m_{t_k}$-measurable random variables.

d) It follows from (8.7),(8.13) and (8.23).

□

We return to the model describing the evolution of risky assets, that under the risk-neutral probabilities $\mathbb{Q}_m$ results into the following expression:

$$\Delta S_m(t_{k+1}) = rS_m(t_k)\Delta t + \sigma S_m(t_k)\Delta W_m(t_{k+1}). \tag{8.28}$$

Until now, we have been introducing the random walk approximation to the Wiener process and seeing how to establish the relations between them and all the properties from Brownian motion that can be extrapolated to these approximations. The next step is to prove that this property still hold we approximate the process through random walks.

Lemma 8.7. *The disocunted price process $r^{-m}S_m(t_k)$ and the discounted value of the portfolio $r_m^{-k}V_m(t_k)$ are $\mathbb{Q}_m$-martingales with respect to the $\mathcal{F}_{t_k}^m$.*

Proof. We will mainly use the fact that $W_m(t_k)$ is a $\mathbb{Q}$-Brownian motion and 8.28.

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}_m}(r_m^{-(k+1)}S_m(t_{k+1})|\mathcal{F}_{t_k}^m) &= \mathbb{E}_{\mathbb{Q}_m}(r_m^{-(k+1)}(S_m(t_k)+\Delta S_m(t_{k+1}))|\mathcal{F}_{t_k}^m)\\ &= r_m^{-(k+1)}S_m(t_k)+\mathbb{E}_{Q_m}(r_m^{-(k+1)}(rS_m(t_k)\Delta t+\sigma S_m(t_k)\Delta W_m(t_{k+1}))|\mathcal{F}_{t_k}^m)\\ &= r_m^{-(k+1)}S_m(t_k)+r_m^{-(k+1)}S_m(t_k)\mathbb{E}_{Q_m}(r\Delta t+\sigma S_m(t_k)\Delta W_m(t_{k+1}))|\mathcal{F}_{t_k}^m)\\ &= r_m^{-(k+1)}S_m(t_k)(1+r\Delta t) = r_m^{-k}S_m(t_k).\end{aligned}$$

Then, since $\beta_m(t_k) = r_m^k$, we have:

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}_m}(r_m^{-(k+1)}V_m(t_{k+1})|\mathcal{F}_{t_k}^m) &= r^{-(k+1)}\mathbb{E}_{\mathbb{Q}_m}(a_m(t_{k+1})S_m(t_{k+1})+b_m(t_{k+1})r_m^{k+1})|\mathcal{F}_{t_k}^m)\\ &= r^{-(k+1)}\mathbb{E}_{\mathbb{Q}_m}(a_m(t_k)S_m(t_{k+1})+b_m(t_k)r_m^{k+1})|\mathcal{F}_{t_k}^m)\\ &= r_m^{-k}a_m(t_k)S_m(t_k)+b_m(t_k) = r_m^{-k}V_m(t_k),\end{aligned}$$

using the self-financing condition and (8.28). □

In the next step we will give an arbitrage-free price of the claim $g(S_m(T))$ at maturity $T > 0$. By the completeness of our models, we know that it is always possible to hedge the claim $g(S_m(T))$ of value $f_m(t_k)$ at any moment $t_k \in [0.T]$ with a replicating self-financing portfolio $V_m(t_k)$ of the claim. Thus, using the previous lemma:

Corollary 8.8. *The arbitrage-free price of an option at time $t_k \in [0, T]$ whose claim is $g(S_m(T))$ with maturity $T = N\Delta t$ is:*

$$f_m(t_k) = V_m(t_k) = \mathbb{E}_{\mathbb{Q}_m}(r_m^{k-N}g(S_m(T))|\mathcal{F}_{t_k}^m),$$

where $V_m(t_k)$ is the value of the replicating portfolio of the claim at t_k.

This corollary gives an explicit way of computing the claim $g(S_m(T))$ using the discrete model. Note that $f_m(t_k)$ includes implicitly the depedence on the price x up to t_k in the filtration condition of probability. Following with the notation u_m and d_m for up/down moves of the stock prices with probabilities $q_m^{\pm}$. Every step is independent of the other ones, which gives S_m the structure of a discrete-time Markov chain, and thus, the valuation of the claim is only influenced by the value of the stock price at maturity. More formally:

$$\begin{aligned}f_m(t_k, x) &= r_m^{k-N}\mathbb{E}_{\mathbb{Q}_m}(g(S_m(T))|S_m(t_k) = x)\\ &= r_m^{n-N}\sum_{i=0}^{N-k}\binom{N-k}{i}(q_m^+)^i(q_m^-)^{N-k-i}g(xu_m^i d_m^{N-k-i}). \quad (8.29)\end{aligned}$$

For the case of an European call, $g(S_m(T)) = (S_m(T) - K)_+$ and consequently:

$$C_m(t_k, x) = r_m^{n-N}\sum_{i=0}^{N-k}\binom{N-k}{i}(q_m^+)^i(q_m^-)^{N-k-i}(xu_m^i d_m^{N-k-i} - K)_+.$$

So, the first term that adds something to the sum must satisfy:

$$xu_m^i d_m^{N-k-i} - K > 0 \Longleftrightarrow i > \frac{log(\frac{K}{x}) - (N-k)log(d_m)}{log(\frac{u_m}{d_m})}$$

Thus, taking $i_{m,k} = \lceil \frac{log(\frac{K}{x})-(N-k)log(d_m)}{log(\frac{u_m}{d_m})} \rceil$, we have:

$$C_m(t_k, x) = xBin(j_{m,k}, N-k, \tilde{q}_m^+) - r_m^{k-N} Bin(j_{m,k}, N-k, q_m^+), \tag{8.30}$$

where $\tilde{q}_m^+ = \frac{u_m}{r_m} q_m^+$ and $Bin(j, n, p) = \sum_{i=j}^{n} \binom{n}{i} p^i (1-p)^{n-i}$
The same argument for put options would lead to a similar formula and, if wanted, using the put-call parity, the price for a foward claim would also be obtained.

Theorem 8.9. *Suppose that $g \in \mathcal{C}_c(\mathbb{R}_+)$ and $T > 0$. As $m \to \infty$, the price $f_m(t^{(m)}, x)$ of the option $g(S_m(T))$ obtained by the above discrete approximation converges to its value $f(t, x)$ obtained by the Black-Scholes model, uniformly for $t \in [0, T]$ and $x > 0$:*

$$\lim_{m\to\infty} f_m(t^{(m)}, x) = \lim_{m\to\infty} \mathbb{E}_{\mathbb{Q}_m}\Big(r_m^{\lfloor t2^{2m} \rfloor - \lfloor T2^{2m} \rfloor} g(S_m(T^{(m)}))|S_m(t^{(m)}) = x\Big) =$$
$$\mathbb{E}_{\mathbb{Q}}\Big(e^{-r(T-t)} g(S(T))|S(t) = x\Big) = f(t, x),$$

where $t^{(m)} = \lfloor t2^{2m} \rfloor 2^{-2m}$.

Proof. By (8.13) applied to the case of $S_m(t)$ it is clear that there is uniform convergence of $S_m(t)$ to $S(t)$ on $[0, T]$. Using Lemma $2(b)$ for the time-homogeneous Markov chains $S(t)$ and $S_m(t)$, we have:

$$\lim_{m\to\infty} \sup_{t\in[0,T], x>0} |\mathbb{E}^x_{\mathbb{Q}_m}\Big(g(S_m(T^{(m)} - t^m))\Big) - \mathbb{E}^x_{\mathbb{Q}}\Big(g(S(T-t))\Big)| = 0.$$

This and (8.19) prove the theorem. □

Hence we have convergence for the case of $C_m(t, x)$ and $P_m(t, x)$ to $C(t, x)$ and $P(t, x)$, respectively. The convergence of these expressions to the explicit formulas of the Black-Scholes model is a consequence of the approximation to higher order terms of the normalization of the binomial distributions of the discrete model.

9. Bibliography

References

[1] Billingsley, P. (1968) *Convergence of probability measures*. Wiley, New York.

[2] Black, F. and Scholes, M (1973) *The pricing of options and corporate liabilities*. J. Polit. Econ. **81**, 637–654

[3] Blumenthal, Robert M. (1957), *An extended Markov property*. Trans. Amer. Math. Soc., n. **85**.

[4] Cameron, R. H.; Martin, W. T. (1944). *Transformations of Wiener Integrals under Translations*. *Annals of Mathematics*, n. **45**.

[5] Hewitt, E., Savage, L. J. (1955). *Symmetric measures on Cartesian products*, Trans. Amer. Math. Soc. n. **80**.

[6] Karlin, Samuel, Taylor, Howard M. *A first course in stochastic processes*. Academic Press, 2nd. edition.

[7] Norris, J.R. *Markov Chains*. Cambridge series on statistical and probabilistic mathematics. N. **02**, USA.

[8] Peres, Yuval, Morters, Peter. (2008) *Brownian motion*. University of California, Berkeley.

[9] Szabados, T. (1996) *An elementary introduction to the Wiener process and stochastic integrals*. Studia Sci. Math. Hung. **31**, 31-60.

[10] Szabados, T., Székely, B. (1996) *An elementary approach to Brownian local time based on simple, symmetric random walks*. Budapest University of Technology and Economics. *arxiv.org*.

[11] Szabados, T., Nika, Z. (2018) *Strong approximation of Black-Scholes theory based on simple random walks*. Budapest University of Technology and Economics. *arxiv.org*.

A. Remark on normal distributions and convergence of random variables

In this appendix, the definitions and results will be given for continuous random variables, as the normal distribution will be of main importance, but the same results could be extended to discrete random variables.

Definition A.1. Let X be a continuous random variable with density function $f(x)$, then the **mean** of X, also the first moment, denoted by $\mathbb{E}$, is defined by:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)dx$$

whenever this integral is absolutely convergent.

Definition A.2. In the same conditions, the **variance** of X is defined by:

$$Var(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

whenever both moments exist.

Moments of higher order can be also defined, but in general we will only be interested in the two defined above. However, it will be more explicitly especified, if necessary.

Definition A.3. Two random variables X,Y defined in the same probability space have a continuous **joint distribution** if their joint distribution function $F(x,y) = Pr(X \leq x, Y \leq y)$ can be written as:

$$F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u,v)dudv$$

for all $(x,y) \in \mathbb{R}^2$. In this case, we willl say that $f(x,y)$ is the **joint density function** of (X,Y)

Proposition A.4. *Let (X,Y) a random bivariant continuous variable with joint distribution $F(x,y)$ and density $f(x,y)$. Then, X and Y are continuous random variables with* **marginal density function**

$$f_X(x) = \int_{\mathbb{R}} f(x,y)dy \qquad f_Y(y) = \int_{\mathbb{R}} f(x,y)dx$$

Proposition A.5. *For any pair of random variables X and Y defined in the same probability space, X and Y are independent if and only if*

$$f(x,y) = f_X(x)f_Y(y)$$

Definition A.6. The random vector $(X_1, \dots X_n)$ has a non-degenerate **normal multivariant distribution**, denoted by $X \sim \mathcal{N}(\mu, \Sigma)$ if its density function is

$$f(x_1, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

where $\mu \in \mathbb{R}^n$ and Σ is a symetric postive definite matrix.

Theorem A.7. *If $X \sim \mathcal{N}(\mu, \Sigma)$, then:*

i) $\mathbb{E}(x) = \mu$

ii) $\Sigma = (\sigma_{ij})$ *is the variance and covariance matrix of X. In particular, if $U_1, \dots U_n$ are independent normal random variables $\mathcal{N}(0, 1)$, then:*

$$U = (U_1, \dots U_n) \sim \mathcal{N}(0, Id_n)$$

Theorem A.8. *A random vector X has a non-degenerate multidimensional normal distribution if and only if there exists a non-singular matrix A and a vector b such that:*

$$X = AU + b$$

where $U = (U_1, \dots U_n)$ with $U_i \sim \mathcal{N}(0, 1)$ independent.

Corollary A.9. *A marginal distribution of any dimension of a multimensional normal distribution is also a multidimensional normal.*

Corollary A.10. *Let $X \sim N(\mu, \Sigma)$, and $a_1, \dots, a_n$ real numbers such that $\sum_{i=0}^{n} a_i^2 > 0$, then:*

$$\sum_{i=0}^{n} a_i X_i \sim N(\sum_{i=0}^{n} a_i \mu_i, \sum_{i=0}^{n} a_i^2 + 2 \sum_{i<j} a_i a_j \sigma_{ij})$$

As a consequence of this corollary, the linear combination of normal distributions is also normal and the marginal distribution of any joint normal distribution is normal. This behaviour of normal random varaibles and vector makes it more straight when dealing with normal distributions. The importance of the normal random variables remains on the fact that they are *central* variables, that is, under some conditions, the sum random variables, generally taken independent identically distributed, tends to distribute as a normal random variable.

Definition A.11. Let $X_1, \dots X_n$ be a sequence of random variables defined in some probability space. Then, we say that $X_n \to X$ when $n \to \infty$ in **distribution** if:

$$Pr(X_n \leq x) \overset{n \to \infty}{\longrightarrow} Pr(X \leq x)$$

for all x where $F_X(x) = Pr(X \leq x)$ is continuous and it will be denoted by:

$$X_n \overset{D}{\longrightarrow} X$$

Theorem A.12. (Central Limit Theorem)
Let $X_1, X_2, \dots, X_n$ independent identically distributed random variables with $0 < V(X_1) = \sigma^2 < \infty$ and $\mathbb{E}(X_1) = \mu$. Sigui $S_n = \sum_{i=1}^{n} X_i$, $n \geq 1$, then

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \overset{D}{\longrightarrow} \mathcal{N}(0, 1), \text{ when } n \to \infty$$

Theorem A.13. (Multidimensional Central Limit Theorem)
Let $X_i = (X_{i1}, X_{i2} \dots, X_{in})$ with $i \in \mathbb{N}$ a sequence of random vectors with independent indentically distributed random variables, and $\mu = (\mathbb{E}(X_1), \mathbb{E}(X_2), \dots \mathbb{E}(X_n))^T$ and covariance matrix Σ. Then, if $S_n = \sum_{i=1}^{n} X_i$

$$\frac{S_n - n\mu}{\sqrt{n}} \overset{D}{\longrightarrow} \mathcal{N}(0, \Sigma)$$