BACHELOR'S THESIS

# Degree in Mathematics

**Title:** Continuous Random Trees


**Author:** Ganjour Dmitri

**Advisor:** Prof. Juan José Rué Perna

**Department:** Facultat de Matemàtiques i Estadística.


**Academic year:** 2018-2019




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

# Continuous random trees

**Ganjour Dmitri**

Supervised by Prof. Juan José Rué Perna

June, 2019

I would like to express my gratitude to Professor Juan José Rué Perna for supervising my bachelor thesis throughout the semester. I want to thank Professor Guillem Perarnau for the help in the choice of the topic.
Special thanks to my parents Ganjour Olga and Serguei and my family for supporting me in my studies.

## Abstract

The Brownian motion has played an important role in the development of probability theory and stochastic processes. We are going to see that it appears in the limiting process of several discrete processes. In particular, we will define discrete processes on Galton-Watson trees to see 2 different types of limits, which are the local limits and the scaling limits. The first result, Kesten's theorem, is a result for the local limits. We are going to look at the trees up to an arbitrary fixed height and therefore only consider what happens at a finite distance from the root. The second result concerns the limit of the rescaled height processes of an infinite Galton-Watson forest. We are going to consider sequences of trees where the branches are scaled by some factor so that all the vertices remain at finite distance from the root. Due to the scaling, the branches have infinitesimal length. These scaling limits lead to the so-called continuous random trees.



### Keywords

Brownian motion, Galton-Watson process, discrete trees, random trees, Kesten's tree, local convergence of random trees, height process, continuous random trees, reflected Brownian motion

# Introduction

The Brownian motion is named after the botanist Robert Brown, who first described the phenomenon in 1827. Brown was studying pollen grains of the plant *Clarkia pulchella* suspended in water under a microscope when he observed minute particles, ejected by the pollen grains, executing a jittery motion ([6]). Later, mathematicians such as Thorvald N. Thiele and Norbert Wiener have brought mathematical explanations to the Brownian motion. Nowadays, the Brownian motion has played an important role in the development of probability theory and stochastic processes. In this thesis, we are going to see that the Brownian motion appears in the limiting process of the height process, obtained from a Galton-Watson forest. The Galton-Watson trees represent the development in generations of a population, so we might obtain trees with a large amount of individuals. There are many kinds of limits that can be considered in order to study large trees. We are going to see 2 different types of limits of Galton-Watson trees, which are the local limits and the scaling limits. The first main theorem of the thesis, Kesten's theorem, is a result for the local limits. We are going to look at the trees up to an arbitrary fixed height and therefore only consider what happens at a finite distance from the root. The second result concerns the limit of the rescaled height processes of an infinite Galton-Watson forest. We are going to consider sequences of trees where the branches are scaled by some factor so that all the vertices remain at finite distance from the root. Due to the scaling, the branches have infinitesimal length. These scaling limits lead to the so-called continuous random trees. In order to obtain these results, we have to start by studying the Brownian motion and developing two different constructions of it. We will mention its properties that will be needed later. Afterwards, we will introduce the discrete trees and a particular case of them, the Galton-Watson trees based on the the Galton-Watson process. From these notions, we can prove the result of Kesten's theorem about the local convergence. We will finish by explaining the important ideas behind the convergence of the rescaled height process of the so-called continuous random trees.

# 1. Reminders on probability theory

In this chapter we will remind some definitions and results that will be needed in the next parts of the thesis. In particular, we will mention the notion of stochastic process that is basic to define the Brownian motion by looking at [2]. Also, we need to recall the notion of convergence of a sequence of random variables with the help of [4]. Finally, we will remind some properties of the normal distribution. For the most basic concepts on probability theory, we can refer to [1].

## 1.1 Stochastic process

We start by providing a definition of a stochastic process. It will be necessary in the thesis as we are going to study several processes.

**Definition 1.1.** A stochastic process is a set of random variables $\{X(t)\colon\ t \in T\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$. $T$ is the index set, $\Omega$ the sample set, $\mathcal{F}$ a $\sigma$-algebra and $P$ a probability measure.

Depending on the index set $T$, we can distinguish several types of stochastic processes.

**Definition 1.2.** If $T$ is countable, we say that $\{X(t)\colon\ t \in T\}$ is a discrete time process. If $T$ is not countable, it's a continuous time process.

In the next chapters, we will see a continuous time stochastic process, the Brownian motion, and a discrete time stochastic process, the Galton-Watson process. We can define the notion of convergence for a stochastic process. Let us see different modes of convergence that we can define.

**Definition 1.3.** We say that the sequence $\{X_n\}_{n\geq 1}$ converges in distribution to the random variable $X$ as $n \to +\infty$ if $F_{X_n} \to F_X$. That is, if

$$F_{X_n}(x) \to F_X(x) \text{ for all } x \in C(F_X),$$

where

$$C(F_X) = \{x \in \mathbb{R} \text{ such that } F_X \text{ is continuous at } x\}.$$

We express that the sequence of random variables $\{X_n\}_{n\geq 1}$ converges in distribution to the random variable $X$ by

$$X_n \xrightarrow{(d)} X.$$

Also, we need the following mode of convergence, which will be very important in the definition of the Brownian motion.

**Definition 1.4.** We say that the sequence $\{X_n\}_{n\geq 1}$ converges almost surely to the random variable $X$ as $n \to +\infty$ if

$$\mathbb{P}(\{\omega \in \Omega \text{ such that } X_n(\omega) \to X(\omega) \text{ as } n \to +\infty\}) = 1.$$

We indicate that the sequence of random variables $\{X_n\}_{n\geq 1}$ converges almost surely to the random variable $X$ by

$$X_n \xrightarrow{a.s.} X.$$

Finally, the last type of convergence that will be needed is the convergence in probability.

**Definition 1.5.** We say that the sequence $\{X_n\}_{n\geq 1}$ converges in probability to the random variable $X$ as $n \to +\infty$ if, for all $\epsilon > 0$
$$\mathbb{P}(|X_n - X| \geq \epsilon) \to 0 \text{ as } n \to +\infty.$$
We show that the sequence of random variables $\{X_n\}_{n\geq 1}$ converges in probability to the random variable $X$ by
$$X_n \xrightarrow{(P)} X.$$

Later in the thesis, we will need some results about stochastic processes. They are listed in the next parts.

### 1.1.1 Borel-Cantelli lemmas

The Borel-Cantelli lemmas are used to determine whether an event can occur infinitely many times or not. This is important in order to prove convergence in the almost surely mode.

**Lemma 1.6.** *Let $\{E_i, i \in \mathbb{N}^*\}$ be a sequence of events in some probability space and let*

$$\limsup_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k\geq n} E_k$$

*be the event "$E_i$ occurs for infinitely many $i$". If*

$$\sum_{n=1}^{+\infty} \mathbb{P}(E_n) < \infty,$$

*then*
$$\mathbb{P}(\limsup_{n\to\infty} E_n) = 0.$$

*So only a finite number of events occur.*

*Proof.* Let $I_n = I\{E_n\}$ denote the indicator random variable of the event $E_n$, and let

$$N = \sum_{n=1}^{\infty} I_n,$$

be the total number of events that occur. Then

$$\mathrm{E}(N) = \mathrm{E}\left(\sum_{n=1}^{\infty} I_n\right) = \sum_{n=1}^{\infty} \mathrm{E}(I_n) = \sum_{n=1}^{\infty} \mathbb{P}(E_n),$$

and by the assumption,
$$\mathrm{E}(N) = \sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty.$$

This implies that
$$\mathbb{P}(N = \infty) = 0.$$

□

The next lemma is the second Borel-Cantelli lemma in which we consider the opposite case for the condition.

**Lemma 1.7.** *Let $\{E_i, i \in \mathbb{N}^*\}$ be a sequence of independent events in some probability space. If*

$$\sum_{n=1}^{+\infty} P(E_n) = \infty,$$

*then*

$$P(\limsup_{n\to\infty} E_n) = 1.$$

Note that in Lemma 1.7, the random variables have to be independent but it is not the case in Lemma 1.6.

## 1.2 Normal distribution and properties

Let us remind the definition of a normal distribution.

**Definition 1.8.** We say that a random variable $X$ is normally distributed if its density function is

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}},$$

where $\mu$ is the mean value of $X$ and $\sigma^2$ is the variance.

### 1.2.1 Central limit theorem

The central limit theorem provides a result on the convergence of the average random variable of independent and identically distributed (from now on, i.i.d.) random variables.

**Theorem 1.9.** *Let $\{X_n\}_{n\geq 1}$ be a sequence of random i.i.d. variables with finite mean $\mu$ and finite variance $\sigma^2$. Let*

$$S_n = X_1 + ... + X_n,$$

*then $\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{(d)} N(0,1)$.*

In the construction of the Brownian motion, this theorem will allow us to verify that the construction is valid. The proof of this theorem can be done using characteristic functions.

### 1.2.2 Random normal vector

We know that a random vector is a vector in which the components are random variables. We can restrict this notion to random vectors where each component has a standard normal distribution. Let us define it formally by using [5].

**Definition 1.10.** A real random vector $X = (X_1, ..., X_n)^T$ is called a standard normal random vector if all of its components $X_k$ are independent and each of them is normally distributed with mean 0 and variance 1, i.e. $X_k \sim N(0,1)$ for all $k$.

We can extend the previous definition to vectors with different mean and variance than those of the standard normal distribution.

**Definition 1.11.** A real random vector $X = (X_1, ..., X_n)^T$ is called a random normal vector if there exists a standard normal random vector $Y$ of size $l$, a vector $\mu$ of size $n$ and a $n \times l$ matrix $A$ such that $X = AY + \mu$. Formally:

$$X \sim N(\mu, \Sigma) \Leftrightarrow \text{there exist } \mu \in \mathbb{R}^n, A \in \mathbb{R}^{n\times l} \text{ such that } X = AY + \mu \text{ for } Y_k \sim N(0,1) \text{ i.i.d.},$$

where $\Sigma = AA^T$ is the covariance matrix.

### 1.2.3 Orthogonal transformation of a normal random vector

We can apply a linear transformation to a random normally distributed vector $X$ by multiplying it by a constant matrix $A$ to obtain the random vector $Y = AX$. Moreover, if this transformation is orthogonal, we can easily determine the distribution of the vector $Y$. We will need this result in the Lévy's construction of the Brownian motion in the following chapter. We will use a 2 dimensional vector as it is the one that we will need. The next proposition is a more general case of a similar result in [3] as we don't suppose that the determinant of $A$ is equal to 1.

**Proposition 1.12.** *Let* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ *be a random vector such that* $X_1$ *and* $X_2$ *are independant and normally distributed with mean 0 and finite variance* $\sigma^2$*. Let* $A$ *be an orthogonal matrix. Then* $Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = AX$ *consists of i.i.d. normally distributed variables.*

*Proof.* First, we can notice that, as $A$ is orthogonal,

$$A^T A = \frac{1}{k} Id \text{ where } k = (\mid det(A) \mid)^{-1} \text{ and } Id \text{ is the identity matrix.} \tag{1}$$

Thus, $A^{-1} = kA^T$. Now we can express the density function of the vector $X$ as

$$f_X(x) = f_X(x_1, x_2) = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^2} e^{-\frac{(x_1^2+x_2^2)}{2\sigma^2}} = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^2} e^{-\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} \frac{1}{2\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}. \tag{2}$$

Let us write

$$\Sigma = \begin{pmatrix} \frac{1}{2\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^2} \end{pmatrix} = \frac{1}{2\sigma^2} Id.$$

As $x = A^{-1}y$, we can replace $x$ by $A^{-1}y$ in Equation (2), obtaining that

$$f_X(x) = f_X(A^{-1}y) = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^2} e^{-(A^{-1}y)^T \Sigma (A^{-1}y)} = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^2} e^{-y^T((A^{-1})^T \Sigma A^{-1})y}. \tag{3}$$

This expression can be developed, obtaining that it is equal to

$$f_X(x) = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^2} e^{-\frac{(y_1^2+y_2^2)}{|det(A)|2\sigma^2}}. \tag{4}$$

The last equation holds because $y^T((A^{-1})^T\Sigma A^{-1})y$ can be expanded as:

$$y^T((A^{-1})^T\Sigma A^{-1})y = y^T\left((A^{-1})^T * \frac{1}{2\sigma^2}Id * A^{-1}\right)y = \frac{1}{2\sigma^2}y^T\left((A^{-1})^T * A^{-1}\right)y.$$

Using Equation (1), we obtain that the last expression is equal to

$$\frac{1}{2\sigma^2}y^T\left(k^2AA^T\right)y = \frac{1}{2\sigma^2}y^T\left(k * Id\right)y = \frac{k}{2\sigma^2}(y_1^2+y_2^2).$$

Let $\sigma'^2 = \mid det(A) \mid \sigma^2$. By replacing it in equation (3), we obtain

$$f_X(A^{-1}y) = \frac{\mid det(A) \mid}{\left(\sqrt{2\pi\sigma'^2}\right)^2}e^{-\frac{(y_1^2+y_2^2)}{2\sigma'^2}}dy. \tag{5}$$

Let $\Omega \in \mathbb{R}^2$, we can calculate $\mathbb{P}(Y \in \Omega)$ by applying the formula of integration by substitution:

$$\mathbb{P}(Y \in \Omega) = \mathbb{P}(AX \in \Omega) = \mathbb{P}(X \in A^{-1}\Omega) = \int_{A^{-1}\Omega} f_X(x)dx = \int_{\Omega}\frac{f_X(A^{-1}y)}{\mid det(A) \mid}dy. \tag{6}$$

Finally by replacing the expression from equation (5) in the equation (6) we get:

$$P(Y \in \Omega) = \int_{\Omega}\frac{1}{\left(\sqrt{2\pi\sigma'^2}\right)^2}e^{-\frac{(y_1^2+y_2^2)}{2\sigma'^2}}dy.$$

We can conclude that Y is a normally distributed i.i.d vector in which every component is normally distributed with mean 0 and variance $\mid det(A) \mid \sigma^2$.

□

The following proposition is taken from [7] and provides a tail bound for the normal distribution. In particular, we will see that the probability that a standard normal variable takes large values is exponentially small. It will be useful to prove that the Lévy's construction of the Brownian motion converges uniformely.

**Proposition 1.13.** *Let $Z$ be a random variable with standard normal distribution. Then, for $c > 1$ and large $n$, $\mathbb{P}(\mid Z \mid \geq c\sqrt{n}) \leq exp(\frac{-c^2n}{2})$*

*Proof.*

$$\begin{aligned}
\mathbb{P}(\mid Z \mid \geq c\sqrt{n}) &= \frac{1}{\sqrt{2\pi}}\int_{x \geq c\sqrt{n}} e^{\frac{-x^2}{2}}dx + \frac{1}{\sqrt{2\pi}}\int_{x \leq -c\sqrt{n}} e^{-\frac{x^2}{2}}dx \\
&= \frac{2}{\sqrt{2\pi}}\int_{x \geq c\sqrt{n}} e^{-\frac{x^2}{2}}dx \\
&\leq \frac{2}{\sqrt{2\pi}}\int_{x \geq c\sqrt{n}} e^{-\frac{x^2}{2}}\frac{x}{c\sqrt{n}}dx
\end{aligned}$$

The last inequality holds because $1 \leq \frac{x}{c\sqrt{n}}$. By calculating the last integral, we get the result. □

# 2. Brownian motion: existence and properties

In this chapter, we will introduce the standard Brownian motion, also known as the Wiener process. We will present its different constructions and some properties that will be useful later.

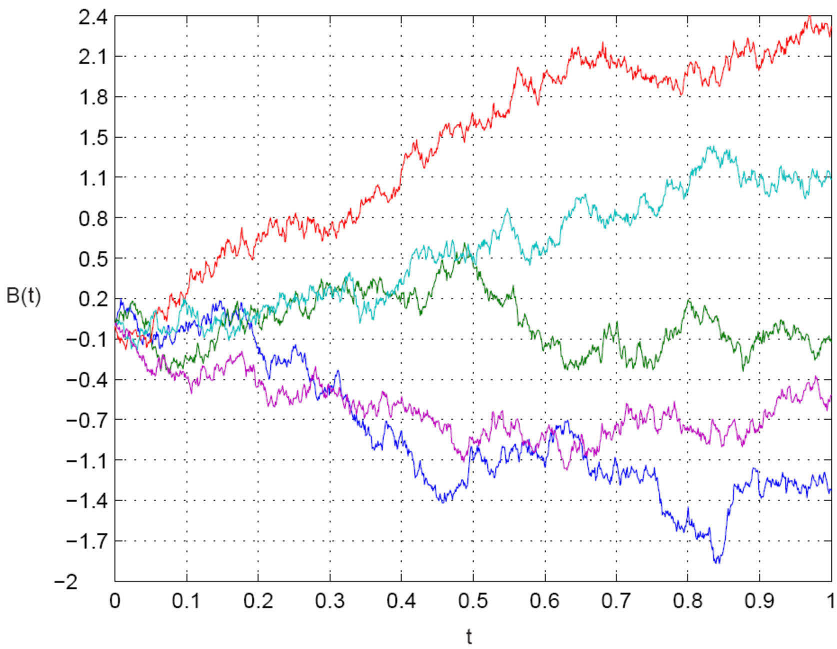
## 2.1 Definition

The Brownian motion is a stochastic process whose variations satisfy a normal distribution. We develop this idea more precisely in the following definition used in [7].

**Definition 2.1.** A standard Brownian motion is a stochastic process $\{B(t), t \geq 0\}$ such that:
1) $B(0) = 0$,
2) For all $t, s \geq 0$ and $s \neq 0$, $B(t+s) - B(t) \sim N(0, s)$,
3) If $0 \leq t_1 < t_2 < ... < t_n$, then the increments $B(t_2) - B(t_1), B(t_3) - B(t_2), ..., B(t_n) - B(t_{n-1})$ are independent.

In Figure 1 we can see different realizations of a standard Brownian motion.

Figure 1: Standard Brownian motion taken from [14]




We have given a definition of a standard Brownian motion but we do not know whether such a stochastic process exists or not. To prove its existence, we can consider 2 different constructions of it. The first construction was developped by the mathematician Paul Lévy and we will see it in details in the next part.

## 2.2 Lévy's construction

This part is essentially based on the material of articles [7] and [8]. Lévy's construction starts with a decomposition of the interval $[0,1]$ using a diadic argument. More precisely, we introduce the set $\mathcal{D}=\bigcup_{n\in\mathbb{N}}\mathcal{D}_n$ with $\mathcal{D}_n=\{\frac{k}{2^n},\ k=0,...,2^n\}$. We call $\mathcal{D}$ the set of dyadic numbers and the idea behind Lévy's construction is to build a function satisfying the properties of the Brownian motion on this set and to interpolate in between. We will use the density of $\mathcal{D}$ in $[0,1]$ to prove that the function obtained at the end satisfies the Definition 2.1. First, we need to prove that the set of dyadic numbers is dense in $[0,1]$.

**Proposition 2.2.** *$\mathcal{D}$ is dense in [0,1]*

*Remark* 2.3. We can notice that $\mathcal{D}_0\subset\mathcal{D}_1\subset...\subset\mathcal{D}_n\subset...$ because if $x=\frac{k}{2^n}\in\mathcal{D}_n$ for a $k\in\mathbb{Z}$, then $x=\frac{2k}{2^{n+1}}\in\mathcal{D}_{n+1}$.

*Proof.* Let $x\in[0,1]$ and $\delta>0$. Clearly, there exists $m$ large enough such that $\frac{k}{2^m}-\frac{k-1}{2^m}=\frac{1}{2^m}<\delta$ for all $k=1,...,2^m$. Thus, there exists $y\in\mathcal{D}_m$ such that $y\in B(x,\delta)$. □

Now, let us see how to proceed for the construction. Let $\{Z_t : t\in\mathcal{D}\}$ be a set of independant random variables with standard normal distribution. Observe that this family is indexed by $\mathcal{D}$. We then define a sequence $\{F_n\}_{n\geq 0}$ of continuous functions on [0,1] as follows:

$$F_0(t)=\begin{cases} Z_1 & \text{if t=1,}\\ 0 & \text{if t=0,}\\ linear & \text{between.}\end{cases}$$

$$F_n(t)=\begin{cases} 2^{-\frac{n+1}{2}}Z_t & \text{if t}\in\mathcal{D}_n\setminus\mathcal{D}_{n-1},\\ 0 & \text{if t}\in\mathcal{D}_{n-1},\\ linear & \text{between.}\end{cases}$$

In the definition of the sequence $\{F_n\}_{n\geq 0}$ above, we say that the function is linear between two consecutive points if the function takes the values of the linear interpolation. Thus, for all $n$, $F_n$ is a function linear by parts. Observe that in the definition of $F_n$, the expansion of $t$ in base 2 is important, as it determines the smallest subset $\mathcal{D}_n$ of $\mathcal{D}$ to which $t$ belongs to. Finally, for $t\in[0,1]$, we define:

$$B(t)=\sum_{i=0}^{\infty}F_i(t)\text{ and }B^{(m)}(t)=\sum_{i=0}^{m}F_i(t).$$

We can notice that for all $d\in\mathcal{D}$:

$$B(d)=\sum_{i=0}^{n}F_i(d)=\sum_{i=0}^{\infty}F_i(d),$$

where $n=min\{k\in\mathbb{N},d\in\mathcal{D}_k\}$.

The first step is to show that $B(t)$ converges for all $t \in [0,1]$. To do it, we can show the uniforme convergence of $B$ on $[0,1]$. From Lemma 1.13, we can remind that for all $d \in \mathcal{D}$, for $c > 1$ and $n$ large enough,

$$\mathbb{P}(\mid Z_d \mid \geq c\sqrt{n}) \leq exp\left(\frac{-c^2 n}{2}\right). \tag{7}$$

We will use Equation (7) to show that the values of $|Z_d|$ cannot be infinitely large as $n$ grows.

**Lemma 2.4.** *The event $\{$there exists $d \in \mathcal{D}_n$ with $|Z_d| \geq c\sqrt{n}\}$ happens only a finite number of times.*

*Proof.* We are going to use the Borel-Cantelli (Lemma 1.6) to prove the result. More precisely, we get the following bound:

$$\sum_{n=0}^{\infty} \mathbb{P}(\{\text{there exists } d \in \mathcal{D}_n \textit{with} \mid Z_d \mid \geq c\sqrt{n}\}) \leq \sum_{n=0}^{\infty} \sum_{d \in \mathcal{D}_n} \mathbb{P}(\mid Z_d \mid \geq c\sqrt{n}) \leq \sum_{n=0}^{\infty} (2^n + 1) exp\left(\frac{-c^2 n}{2}\right).$$

The last upper bound convergences if $c > \sqrt{2log2}$. By applying Borel-Cantelli lemma, there exists $N$ such that for all $n \geq N$ and for all $d \in \mathcal{D}_n$ we have $|Z_d| < c\sqrt{n}$. □

As a consequence, for $n \geq N$:

$$\sup_{t \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}} 2^{-\frac{n+1}{2}} \mid Z_t \mid < c\sqrt{n} 2^{-\frac{n}{2}}. \tag{8}$$

Now, to obtain the uniform convergence of $B(t)$, we introduce the infinite norm of a function $f$ defined on the set $I$ as

$$\parallel f \parallel_\infty = \sup_{x \in I} f(x).$$

Note that, according to Equation (8), we have

$$\parallel F_n \parallel_\infty < c\sqrt{n} 2^{-\frac{n}{2}}.$$

**Corollary 2.5.** *For all $t \in [0,1]$*

$$\parallel B \parallel_\infty < \infty.$$

*Proof.* We want to bound the infinite norm of $B$

$$\parallel B \parallel_\infty = \parallel \sum_{i=0}^{\infty} F_i \parallel_\infty.$$

By using the triangular inequality,

$$\parallel B \parallel_\infty \leq \sum_{i=0}^{\infty} \parallel F_i \parallel_\infty.$$

Finally, we use Equation (8) and we get that

$$\parallel B \parallel_\infty < \sum_{i=0}^{N} \parallel F_i \parallel_\infty + \sum_{i=N+1}^{\infty} c\sqrt{n} 2^{-\frac{n}{2}} < \infty.$$

□

This concludes the proof of the uniform convergence of $B(t)$ for all $t \in [0,1]$. We will now prove that, for $t \in [0,1]$, $B(t)$ is the standard brownian motion by verifying that it satisfies all properties of Definition 2.1.

**Theorem 2.6.** *The Lévy's construction gives a Brownian motion.*

*Proof.* The proof is made by induction on n. The induction hypothesis are:
$H_{1_n}$ : $\forall r < s < t \in \mathcal{D}_n$, $B(t) - B(s) \sim N(0, t-s)$ and it is independent of $B(s) - B(r)$
$H_{2_n}$ : the vectors $(B(d), d \in \mathcal{D}_n)$ and $(Z_t, t \in \mathcal{D} \setminus \mathcal{D}_n)$ are independant.

We start with n=0. In this case, $\mathcal{D}_0 = \{0,1\}$, $B(0) = 0$, $B(1) = Z_1$. Observe that $B(1) - B(0) = B(1) = Z_1$ so $H_{1_0}$ is satisfied. Let us argue for $H_{2_0}$: the values of $(Z_t, t \in \mathcal{D} \setminus \mathcal{D}_0)$ are not defined yet, so the vectors $(B(d), d \in \mathcal{D}_0)$ and $(Z_t, t \in \mathcal{D} \setminus \mathcal{D}_0)$ are independent so $H_{2_0}$ is satisfied. Then, we prove the induction hypothesis for all $n \geq 0$. We suppose that the hypothesis are true for $n$. We define $B(d)$ for $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$ as:

$$B(d) = \frac{B(d-2^{-n}) + B(d+2^{-n})}{2} + \frac{Z_d}{2^{(n+1)/2}}.$$

We can notice that the first summand is the linear interpolation of $B$ at $d$. We then add a perturbation $\frac{Z_t}{2^{-(n+1)/2}}$. This corresponds to the definition

$$B(d) = \sum_{i=0}^{\infty} F_i(d).$$

An iteration of the construction is presented in Figure 2.

As $d$ belongs to $\mathcal{D}_n \setminus \mathcal{D}_{n-1}$, we know that $d - 2^{-n}, d + 2^{-n} \in \mathcal{D}_{n-1}$ so $B(d-2^{-n})$ and $B(d+2^{-n})$ have already been calculated. Therefore $B(d)$ is independant of ($Z_t$ for $t \in \mathcal{D} \setminus \mathcal{D}_n$).

Let

$$C(d) = \frac{B(d-2^{-n}) - B(d+2^{-n})}{2} \text{ and } D(d) = \frac{Z_d}{2^{(n+1)/2}}.$$

$C$ and $D$ are independant and we notice that

$$C(d) + D(d) = \frac{B(d-2^{-n}) - B(d+2^{-n})}{2} + \frac{Z_d}{2^{(n+1)/2}} = B(d) - B(d-2^{-n}),$$

and

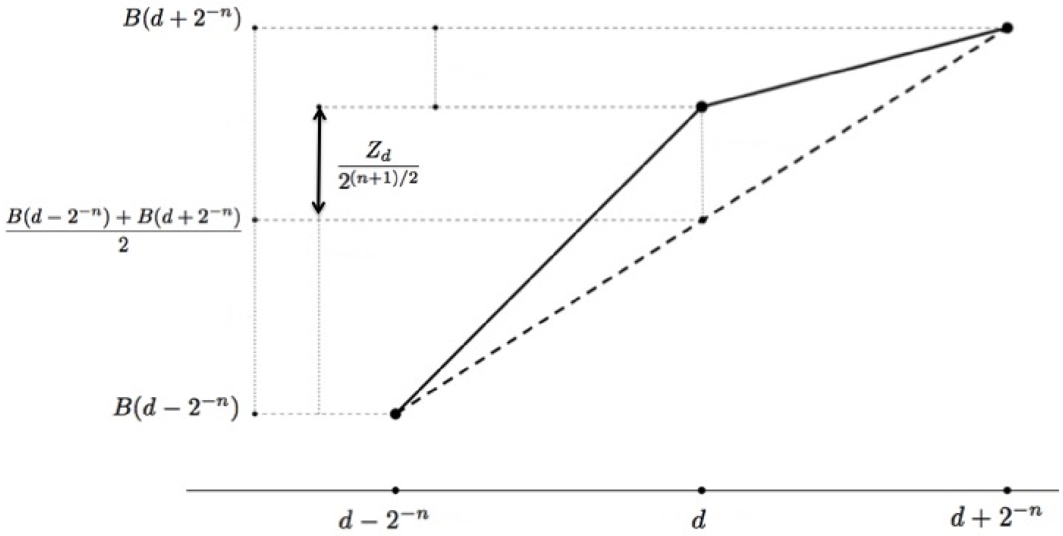
$$C(d) - D(d) = \frac{B(d-2^{-n}) - B(d+2^{-n})}{2} - \frac{Z_d}{2^{(n+1)/2}} = B(d+2^{-n}) - B(d).$$

Moreover, by induction hypothesis $C, D \sim N(0, 2^{-(n+1)})$. Let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

$A$ is orthogonal so we can use the Proposition 1.12 about an orthogonal transformation of a normal i.i.d. random vector. The vector $A\begin{pmatrix} C(d) \\ D(d) \end{pmatrix}$ is independent and normally distributed.

Figure 2: One iteration of Lévy's construction




Thus $B(d)-B(d-2^{-n})$ and $B(d+2^{-n})-B(d)$ are independant and their distribution is $N(0,2^{-n})$. The other case is when the increments are separated by some $d \in \mathcal{D}_{n-1}$. Let $d \in \mathcal{D}_j$ such that $d$ separates the increments and $j$ is minimal. The increments are contained in $[d-2^{-j}, d]$ and $[d, d+2^{-j}]$. By induction hypothesis the increments over these 2 intervals are independent. $H_{1_n}$ is verified. Finally, in order to obtain a Brownian motion for all $t \leq 0$, we can rescale the function obtained on $[0,1]$.

□

The Lévy's construction of the Brownian motion allows us to prove that it is a continuous function. The following result is important as it shows the continuity of the Brownian motion and it bounds the variations. We will use the previous construction to prove it.

**Theorem 2.7.** *There exists a constant $C > 0$ such that, almost surely, for $h > 0$ small enough and $0 \leq t \leq 1-h$ we have*

$$\mid B(t+h) - B(t) \mid \leq C\sqrt{hlog(1/h)}.$$

*Proof.* From the Lévy's construction, we know that the Brownian motion can be written as $B(t) = \sum_{n=0}^{\infty} F_n(t)$. For n fixed, the derivative of $F_n$ exists almost everywhere and it is piecewise constant, so $\parallel F_n' \parallel_\infty \leq \frac{2\|F_n\|_\infty}{2^{-n}}$ because the numerator is the maximum absolute difference and the denominator is the distance between 2 consecutive points. We recall from Equation (8) that for $c > \sqrt{2log(2)}$ and $n$ large enough, there exists $N \in \mathbb{N}$ such that $\forall n > N$:

$$\parallel F_n \parallel_\infty < c\sqrt{n}2^{-\frac{n}{2}}. \tag{9}$$

Therefore

$$\| F_n' \|_\infty \leq \frac{2 \| F_n \|_\infty}{2^{-n}} \leq 2c\sqrt{n}2^{n/2}. \tag{10}$$

Let $t, t+h \in [0,1]$ and $l \in \mathbb{N}$:

$$\begin{aligned} | B(t+h) - B(t) | &\leq \sum_{n=0}^{\infty} | F_n(t+h) - F_n(t) | \leq \sum_{n=0}^{\infty} h \| F_n' \|_\infty \\ &\leq \sum_{n=0}^{l} h \| F_n' \|_\infty + \sum_{n=l+1}^{\infty} 2 \| F_n \|_\infty . \end{aligned}$$

We have used the mean value theorem in the second equality and in the last equality we have used the fact that $| F_n(t+h) - F_n(t) | \leq 2 \| F_n \|_\infty$

By using Equations (9) and (10) and assuming that $l > N$ we get:

$$\begin{aligned} | B(t+h) - B(t) | &\leq \sum_{n=0}^{l} h \| F_n' \|_\infty + \sum_{n=l+1}^{\infty} 2 \| F_n \|_\infty \\ &\leq h \sum_{n=0}^{N} \| F_n' \|_\infty + 2ch \sum_{n=N}^{l} \sqrt{n}2^{n/2} + 2c \sum_{n=l+1}^{\infty} \sqrt{n}2^{-n/2}. \end{aligned}$$

We have to choose $h$ such that $h < \sqrt{hlog(1/h)}$ and $l$ such that $2^{-l} < h \leq 2^{-l+1}$ and larger than $N$. Therefore the final sum that bounds $| B(t+h) - B(t) |$ is bounded by $C\sqrt{hlog(1/h)}$ where $C$ is a constant.

□

Finally, as

$$\lim_{h \to 0} \sqrt{hlog(1/h)} = 0$$

this theorem confirms that the Brownian motion is a continuous function.

## 2.3 Construction using random walks

In this section we will present an other construction of the standard Brownian motion. This construction is based on the usage of random walks and allows us the build a Brownian motion for all $t \geq 0$. We can start with some definitions and properties of random walks.

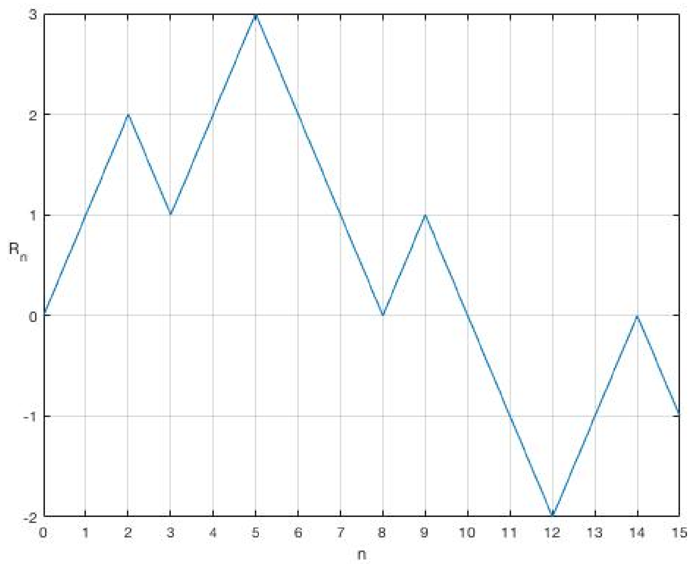
**Definition 2.8.** A simple symmetric random walk $R_n$ is a walk such that:
1) $R_0 = 0$
2) $R_n = \sum_{i=1}^{n} \Delta_i$ where the $\Delta_i$ are i.i.d. such that $P(\Delta_i = -1) = P(\Delta_i = 1) = \frac{1}{2}$

We consider $n$ as the index of time and $R_n$ the position of the particle at time $n$. On Figure 3, we can see a symmetric random walk with $R_0 = 0, R_1 = 1, R_2 = 2, R_3 = 1, \ldots$

We know that, for all $i$, $E(\Delta_i) = 0$ and $Var(\Delta_i) = E(\Delta_i^2) - E(\Delta_i)^2 = 1$. Thus $E(R_n) = 0$ and $Var(R_n) = n$. For large k, we define

$$B_k(t) = \frac{1}{\sqrt{k}} \sum_{i=1}^{tk} \Delta_i.$$

Figure 3: A realization of a symmetric random walk




We can consider it as a random walk in which the particle changes position with time step $\frac{1}{k}$ and with step size $\frac{\Delta_i}{\sqrt{k}}$. If $tk$ is not an integer, we replace $tk$ by $\lfloor tk \rfloor = max\{j \in \mathbb{N}, j \leq tk\}$. The next theorem gives the result that a rescaled random walk with mean 0 and finite variance converges to the Brownian motion.

**Theorem 2.9.** *(Donsker's): as $k \to \infty$, the process $\{B_k(t), t \geq 0\}$ converges to the Brownian motion $\{B(t), t \geq 0\}$.*

*Proof.* For fixed $k$ and $s < t$, we can calculate

$$B_k(t) - B_k(s) = \frac{1}{\sqrt{k}} \sum_{i=sk}^{tk} \Delta_i.$$

We can notice that the distribution of $B_k(t) - B_k(s)$ depends only on the $k(t-s)$ $\Delta_i$ in the interval $(s,t]$, so this difference depends only on $t-s$. Thus, for the limiting process $\{B(t), t \geq 0\}$ as $k \to \infty$, the distribution of any increment $B(t) - B(s)$ depends only on the value $t - s$. For $0 \leq t_1 < t_2 < t_3 < t_4$, the differences

$$B_k(t_4) - B_k(t_3) = \frac{1}{\sqrt{k}} \sum_{i=t_3 k}^{t_4 k} \Delta_i,$$

and

$$B_k(t_2) - B_k(t_1) = \frac{1}{\sqrt{k}} \sum_{i=t_1 k}^{t_2 k} \Delta_i,$$

are independant because the $\Delta_i$ used in the construction of the walk on the intervals $[t_1, t_2]$ and $[t_3, t_4]$ are different and i.i.d. Thus the limiting process $\{B(t), t \geq 0\}$ as $k \to \infty$ will also contain independant increments for non-overlapping intervals. We observe that $E(B_k(t)) = 0$ and $Var(B_k(t)) = \frac{\lfloor tk \rfloor}{k} \to t$ as $k \to \infty$. Thus the limiting process is such that $E(B(t)) = 0$ and $Var(B(t)) = t$. We want to see whether $B_k(t)$ converges to a known probability distributions. We notice that the sequence $(\Delta_i, i \leq 1)$ is made of i.i.d random variables with mean 0 and variance 1. Thus, we can apply the central limit theorem (Theorem 1.9) for $tk$ $\Delta_i$s, as all hypothesis are satisfied. In particular, we get

$$\frac{1}{\sqrt{tk}} \sum_{i=1}^{tk} \Delta_i \xrightarrow{(d)} N(0,1) \text{ as } k \to \infty.$$

We can multiply this result by $\sqrt{t}$ to obtain, using variance properties, that

$$B_k(t) = \sqrt{t} \left( \frac{1}{\sqrt{tk}} \sum_{i=1}^{tk} \Delta_i \right) \xrightarrow{(d)} N(0,t) \text{ as } k \to \infty.$$

We have obtained the limiting probability distribution for $B_k(t)$. Now, let us see if we can get a limiting process for $B_k(t) - B_k(s)$ for $s < t$, in order to verify that the variations follow a normal distribution. By using the same arguments as above, we get:

$$\frac{1}{\sqrt{k(t-s)}} \sum_{i=sk}^{tk} \Delta_i \xrightarrow{(d)} N(0,1) \text{ as } k \to \infty.$$

By multiplying this result by $\sqrt{t-s}$, we obtain

$$B_k(t) - B_k(s) = \sqrt{t-s} \left( \frac{1}{\sqrt{k(t-s)}} \sum_{i=sk}^{tk} \Delta_i \right) \xrightarrow{(d)} N(0, t-s) \text{ as } k \to \infty.$$

For the limiting we process we have the property

$$B(t) - B(s) \sim N(0, t-s),$$

and for every non overlaping intervals, the differences are independent. We can conclude that using this construction, the process obtained satisfies the conditions of a Brownian motion. □

We have seen 2 different constructions of a Brownian motion, in the next part we can discuss some properties.

## 2.4 Properties of the Brownian motion

In this section, we will see that the Brownian motion is a good example of a function continuous everywhere but differentiable nowhere. The content of the proof is based on [7]. First, we need a property that will be useful in the proof of the non-differentiability.

**Lemma 2.10.** *(Scaling property): Let $\{B(t), t \geq 0\}$ be a standard Brownian motion and $a > 0$. Then the process $\{X(t), t \geq 0\}$ such that $X(t) = \frac{1}{a} B(a^2 t)$ is also a standard Brownian motion.*

*Proof.* 1) It is obvious that $X(0) = 0$ because $B$ is a standard Brownian motion.
2) To show that the difference $X(t) - X(s)$ is normally distributed with mean 0 and variance $t - s$ for $t > s$, we write $X(t) - X(s) = \frac{1}{a}(B(a^2t) - B(a^2s))$. We know that $B(a^2t) - B(a^2s)$ has probability law $N(0, a^2(t - s))$, so $X(t) - X(s)$ has probability law $N(0, t - s)$ by applying a basic variance property.
3) Non overlapping increments are independent because we have supposed that $\{B(t), t \geq 0\}$ is a standard Brownian motion. □

### 2.4.1 Non differentiability of the Brownian motion

We have the results to prove that the Brownian motion is nowhere differentiable. First, we introduce some notations: for $h > 0$, we denote by

$$D_+f(t) = \limsup_{h\to 0} \frac{f(t+h) - f(t)}{h} \text{ the upper right derivative of } f \text{ at } t,$$

and by

$$D_-f(t) = \liminf_{h\to 0} \frac{f(t+h) - f(t)}{h} \text{ the lower right derivative of } f \text{ at } t.$$

**Theorem 2.11.** *Almost surely, the Brownian motion is nowhere differentiable. Moreover at least one of these statements is satistified:*
*1)$D_+B(t) = +\infty$ or*
*2)$D_-B(t) = -\infty$.*

*Proof.* Suppose that there exists $t_0 \in [0, 1]$ such that $-\infty < D_-B(t_0) \leq D_+B(t_0) < +\infty$ then

$$\limsup_{h\to 0} \frac{\mid B(t_0 + h) - B(t_0) \mid}{h} < \infty.$$

So there exists a constant M such that:

$$\sup_{h\in[0,1]} \frac{\mid B(t_0 + h) - B(t_0) \mid}{h} \leq M.$$

Therefore, for all $t_0 \in [0, 1]$ and $h$ such that $t_0 + h \in [0, 1]$:

$$\frac{\mid B(t_0 + h) - B(t_0) \mid}{h} \leq M. \tag{11}$$

Let us show that this event has probability 0 to happen. We fix $M$, $t_0$ and $k$ such that $t_0 \in [(k - 1)/2^n, k/2^n]$ where $n > 2$. For $1 \leq j \leq 2^n - k$, we have that

$$\begin{aligned}
\mid B((k + j)/2^n) - B((k + j - 1)/2^n) \mid &\leq \mid B((k + j)/2^n) - B(t_0) \mid + \mid B(t_0) - B((k + j - 1)/2^n) \mid \\
&\leq M\left(\frac{k}{2^n} - t_0 + \frac{k-1}{2^n} - t_0 + \frac{2j}{2^n}\right) \\
&\leq M\left(\frac{1}{2^n} + \frac{2j}{2^n}\right) \leq M(2j + 1)/2^n.
\end{aligned}$$

The second inequality is obtained using equation (11). The third inequality is true because $t_0 \in [(k - 1)/2^n, k/2^n]$. We define the set of events:

$$A_{n,k} = \{\mid B((k + j)/2^n) - B((k + j - 1)/2^n) \mid \leq M(2j + 1)/2^n \text{ for } j = 1, 2, 3\}.$$

By applying the scalling property and knowing that the increments are independent, we get:

$$
\begin{aligned}
\mathbb{P}\left(A_{n,k}\right) &= \prod_{j=1}^{3} \mathbb{P}\left(\mid B((k+j)/2^n) - B((k+j-1)/2^n) \mid \leq M(2j+1)/2^n\right) \\
&\leq \prod_{j=1}^{3} \mathbb{P}\left(\mid B((k+j)/2^n) - B((k+j-1)/2^n) \mid \leq 7M/2^n\right) \\
&= \prod_{j=1}^{3} \mathbb{P}\left(\frac{1}{\sqrt{2^n}} \mid B(k+j) - B(k+j-1) \mid \leq 7M/2^n\right) \\
&\leq \mathbb{P}\left(\mid B(1) \mid \leq 7M/\sqrt{2^n}\right)^3 \\
&\leq (7M2^{-n/2})^3.
\end{aligned}
$$

The first step is due to the independence of the events, the second step is true because $j \leq 3$, in the third step we apply the scaling property, the fourth step is true because $B(k+j)-B(k+j-1)$ has probability law $N(0,1)$ and the last step is due to the fact that the density distribution function of $B(1)$ is bounded by $1/2$.

We can now bound the probability of $\bigcup_{k=1}^{2^n-3} A_{n,k}$:

$$
\mathbb{P}\left(\bigcup_{k=1}^{2^n-3} A_{n,k}\right) \leq 2^n(7M2^{-n/2})^3 = (7M)^3 2^{-n/2}.
$$

The upper bound is sumable over $\mathbb{N}$. By applying Borel-Cantelli lemma:

$$
\mathbb{P}\left(\exists t_0 \text{ with } \sup_{h\in[0,1]} \frac{\mid B(t_0+h) - f(t_0) \mid}{h} \leq M\right) \leq \mathbb{P}\left(\bigcup_{k=1}^{2^n-3} A_{n,k} \text{ for infinitely many n}\right) = 0.
$$

We conclude that we obtain a contradiction with our hypothesis so the Brownian motion is, almost surely, nowhere differentiable. □

In this chapter we have built the Brownian motion using 2 different constructions and he have shown that it is nowhere differentiable. In the next chapters, the goal is to show that the Brownian motion is the limiting process of some processes for discrete trees.

# 3. Discrete and Galton-Watson trees

In this chapter, we will study 2 types of trees: discrete trees and Galton-Watson trees. The material of this chapter is taken from [10], [11], [12]. Actually, we will see that Galton-Watson trees are a particular case of discrete trees, obtained with the help of the Galton-Watson process. That is why we first start by introducing discrete trees.

## 3.1 Discrete trees

Let us state the formal definition of a tree. Afterwards, we are going to use an other definition that will be more convenient for the next parts.

**Definition 3.1.** A tree is an undirected connected graph $G = (V, E)$ with no cycles. We have denoted as $V$ the set of vertices and $E$ is the set of edges.

*Remark* 3.2. The set of vertices $V$ can be infinite and a vertex can have infinite degree.

Let $\mathcal{U} = \bigcup_{n \geq 0} (\mathbb{N}^*)^n$ be the set of finite sequences of positive integers. For each vertex of a tree, we can associate a sequence $u = (u_1, ..., u_n) \in \mathcal{U}$ as shown in Figure 4.

Figure 4: Illustration of a tree with labelled vertices using the set $\mathcal{U}$.

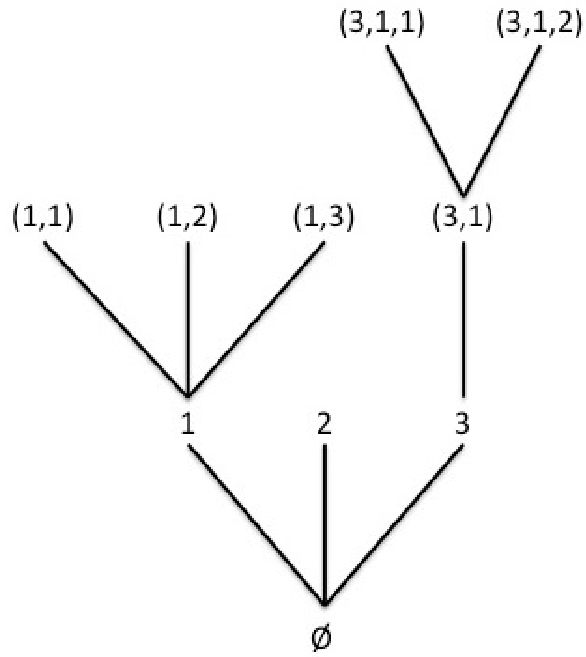



Figure 4 represents a finite tree. We can explain in details the bijection of the set of vertices with the set $\mathcal{U}$. For each vertex $u$ of the tree $\mathbf{t}$, we label its offsprings (if any) from 1 to $k_u(\mathbf{t})$. Then, for each vertex of the tree, we create a sequence of labels of its ancestors ordered from the oldest one to itself. We denote by $n$ the length of $u$ and by convention $|\emptyset| = 0$. Let $u, v \in \mathcal{U}$, we denote by $uv$ the concatenation of the sequences such that $uv = u$ if $v = \emptyset$ and $uv = v$ if $u = \emptyset$. We say that $v$ is an ancestor of u (denoted as $v \prec u$) if $\exists w \in \mathcal{U}$, $w \neq \emptyset$ such that $u = vw$. The set of ancestors of $u$ is $A_u$. Now, we can state a different definition of a tree. We will rather use this one afterwards.

**Definition 3.3.** A tree $\mathbf{t}$ is a subset of $\mathcal{U}$ such that:
1) $\emptyset \in \mathbf{t}$,
2) if $u \in \mathbf{t}$, then $A_u \in \mathbf{t}$,
3) for all $u \in \mathbf{t}$, there exists $k_u(\mathbf{t}) \in \mathbb{N} \cup \{+\infty\}$ such that, for all $i \in \mathbb{N}^*$, $ui \in \mathbf{t}$ if and only if $1 \leq i \leq k_u(\mathbf{t})$.

We call $k_u(\mathbf{t})$ the number of offsprings of the vertex $u$ in $\mathbf{t}$. If $k_u(\mathbf{t}) = 0$, $u$ is a leaf. If $k_u(\mathbf{t}) = +\infty$, $u$ is infinite. $\emptyset$ is the root of $\mathbf{t}$. We denote by $\mathbb{T}_\infty$ the set of trees and by $\mathbb{T}$ the subset of trees where all vertices have finite degree:

$$\mathbb{T} = \{\mathbf{t} \in \mathbb{T}_\infty, k_u(\mathbf{t}) < \infty \ \forall u \in \mathbf{t}\}.$$

For $\mathbf{t} \in \mathbb{T}$, $|\mathbf{t}|$ is the size of the tree (number of vertices). We can notice that

$$\sum_{u \in \mathbf{t}} k_u(\mathbf{t}) = |\mathbf{t}| - 1,$$

because the root does not have ancestors. We denote by $\mathbb{T}_0$ the set of finite trees:

$$\mathbb{T}_0 = \{\mathbf{t} \in \mathbb{T}, |\mathbf{t}| < +\infty\}.$$

*Remark* 3.4. We can notice that we have the following inclusions:

$$\mathbb{T}_0 \subsetneq \mathbb{T} \subsetneq \mathbb{T}_\infty.$$

In fact, if a tree has a finite number of vertices, then all vertices have finite degree.

For $\mathbf{t} \in \mathbb{T}_\infty$, we denote by $L_0(\mathbf{t})$ the set of leaves of $\mathbf{t}$:

$$L_0(\mathbf{t}) = \{u \in \mathbf{t} \text{ such that } k_u(\mathbf{t}) = 0\}.$$

We can define functions characterizing $\mathbf{t}$ such as the global height function $H$ and the width at level $h$, $z_h$.

$$H(\mathbf{t}) = sup\{|u|, u \in \mathbf{t}\} \text{ and } z_h(\mathbf{t}) = \mid \{u \in \mathbf{t}, \mid u \mid = h\} \mid .$$

In other words, $z_h(\mathbf{t})$ is the number of vertices at level $h$.

**Example 3.5.** In the tree of Figure 4, $H(\mathbf{t}) = 3$ and $z_0(\mathbf{t}) = 1$, $z_1(\mathbf{t}) = 3$, $z_2(\mathbf{t}) = 4$ and $z_3(\mathbf{t}) = 2$.

We can define $\mathbb{T}^{(h)}$ as the subset of trees with global height less than $h$:

$$\mathbb{T}^{(h)} = \{\mathbf{t} \in \mathbb{T}, H(\mathbf{t}) \leq h\}.$$

Finally, for $\mathbf{t} \in \mathbb{T}$, we introduce the restriction function from $\mathbb{T}$ to $\mathbb{T}^{(h)}$ defined as

$$r_h(\mathbf{t}) = \{u \in \mathbf{t}, |u| \leq h\}.$$

From now on, we denote as $r_h(\mathbf{t})$ the subtree of $\mathbf{t}$ obtained by cutting the tree at height $h$.

## 3.2 Galton-Watson process

We are going to introduce the Galton-Watson process. It is a stochastic process that corresponds to a population organised in generations. We define this process in the following definition.

**Definition 3.6.** Let $X$ be a random variable taking values in $\mathbb{N}$ with distribution $p = (\mathbb{P}(X = n), n \in \mathbb{N})$. We denote by $p(k)$ the value of $\mathbb{P}(X = k)$ and by $m$ the expectation $E(X)$, and we let $g(s) = \sum_{k \in N} \mathbb{P}(X = k)r^k$ be the probability generating function of $p$.

Now, suppose that we have a population such that:
1) the population evolves in generations. Let $Z_k$ be the number of members in the $k-th$ generation. We assume that $Z_0 = 1$
2) each member of the $k-th$ generation gives birth independently to $X$ children following the distribution $p$. The set of children of the $k-th$ generation is the $(k+1)-th$ generation.

Then, we say that $Z = (Z_k, k \in \mathbb{N})$ is a Galton-Watson process with offspring distribution $p$.

We can distinguish Galton-Watson processes according to the value of the mean $m$.

**Definition 3.7.** The offspring distribution is sub-critical if $m < 1$, critical if $m = 1$ and super-critical if $m > 1$.

### 3.2.1 Extinction probability

We can imagine that the population gets extinct at a certain generation, in other words, this generation doesn't produce offsprings. We will explain how to calculate the probability that this event happens.

**Definition 3.8.** We say that the population is extinct at generation n if $Z_n = 0$.

It is obvious that if there exists $N$ such that $Z_N = 0$, then $Z_n = 0$ for all $n \geq N$.

We denote as $E$ the event of extinction of the population:

$$E = \{\exists N \in \mathbb{N} \text{ such that } Z_N = 0\}.$$

First we can consider some basic cases and calculate the probability of extinction for them:
1) If $p(0) = 0$ then $\mathbb{P}(E) = 0$ because every individual produces at least one offspring.
2) If $p(0) = 1$ then $\mathbb{P}(E) = 1$ and we only have the first element in generation 0.
3) If $p(1) = 1$ then we have one member at every generation so the process is infinite and $\mathbb{P}(E) = 0$.

Now we can calculate the probability of extinction omitting these basic cases.
We need to introduce $E_m$ the event $\{Z_m = 0\}$ and let $e_m = \mathbb{P}(Z_m = 0)$.

$$e_m = \mathbb{P}(E_m) = \sum_{k \geq 0} \mathbb{P}(E_m \mid Z_1 = k)\mathbb{P}(Z_1 = k) = \sum_{k \geq 0} (e_{m-1})^k \mathbb{P}(X = k) = g(e_{m-1}).$$

The third equality is true because we have $k$ independant Galton-Watson processes with one generation less. Therefore we obtain the relation

$$e_m = g(e_{m-1}). \tag{12}$$

We can notice that $e_0 \leq e_1 \leq ... \leq e_m \leq ... \leq 1$, so

$$e = \lim_{m \to \infty} e_m$$

exists. We can take the limit as $m \to \infty$ of Equation (12) so we get $e = g(e)$. The probability of extinction is the smallest solution of $s = g(s)$.

## 3.3 Galton-Watson trees

Let $p$ be a probability distribution over $\mathbb{N}$ as defined in the previous section, $X$ a random variable with distribution $p$ and $g$ the probabiliy generating function of $p$. We will provide a representation of a Galton-Watson process using discrete trees. Let us show how to proceed. First, we define a branching process, as it will be necessary to introduce Galton-Watson trees.

**Definition 3.9.** A $\mathbb{T}$-valued random variable $\tau$ is said to have a branching property if for $n \in \mathbb{N}^*$, conditionnaly on $k_\emptyset(\tau) = n$, the sub-trees $(S_1, ..., S_n)$ are independent and distributed like the original tree $\tau$.

Now, by putting more conditions on the distribution of $\tau$, we can give the definition of the Galton-Watson trees.

**Definition 3.10.** A $\mathbb{T}$-valued random variable $\tau$ is a Galton-Watson tree if it has the branching property and the distribution of $k_\emptyset(\tau)$ is $p$.

Let us give the first result concerning Galton-Watson trees. For $\mathbf{t} \in \mathbb{T}^{(h)}$, it gives the probability that the restriction to height $h$ of a Galton-Watson tree is equal to a fixed tree $\mathbf{t}$.

**Proposition 3.11.** *If $\tau$ is a Galton-Watson tree with offspring distribution $p$ then for every $h \in \mathbb{N}^*$ and a tree $\mathbf{t} \in \mathbb{T}^{(h)}$*

$$\mathbb{P}(r_h(\tau) = \mathbf{t}) = \prod_{u \in \mathbf{t}, |u| < h} p(k_u(\mathbf{t})). \tag{13}$$

*Proof.* $r_h(\tau)$ is the restriction of $\tau$ to level $h$. As every $k_u(\mathbf{t})$ is distibuted independently following the distribution $p$, Equation (13) is verified. □

In particular, the restriction of $\tau$ on $\mathbb{T}_0$ satisfies:

$$\forall \mathbf{t} \in \mathbb{T}_0, \; \mathbb{P}(\tau = \mathbf{t}) = \prod_{u \in \mathbf{t}} p(k_u(\mathbf{t})).$$

We can apply the section 3.2.1 about extinction probability to Galton-Watson trees by identifying a tree with a Galton-Watson process.

In the next chapter, we will introduce another type of discrete trees, Kesten's trees and then prove some results about them.

# 4. Kesten's trees and Kesten's convergence theorem

In this chapter, we will discuss Kesten's trees. First, we will present some topology results about the set of trees $\mathbb{T}$ that we have seen in the previous chapter. It will allow us to introduce Kesten's trees and the main theorem of this chapter: Kesten's theorem. The reference [12] was used for this chapter.

## 4.1 Topology

We recall some notations from the previous chapter. Recall that $\mathbb{T}_\infty$ is the set of trees. We denote by $\mathbb{T}$ the subset of trees with all vertices to be finite:

$$\mathbb{T} = \{\mathbf{t} \in \mathbb{T}_\infty, k_u(\mathbf{t}) < \infty \ \forall u \in \mathbf{t}\}.$$

We denote by $\mathbb{T}_0$ the set of finite trees, namely $\mathbb{T}_0 = \{\mathbf{t} \in \mathbb{T}, \mid \mathbf{t} \mid < +\infty\}$. Finally, we write $\mathbb{T}^{(h)}$ for the subset of trees with global height less than $h$:

$$\mathbb{T}^{(h)} = \{\mathbf{t} \in \mathbb{T}, H(\mathbf{t}) \leq h\}.$$

We can remind the restriction function $r_h$ from $\mathbb{T}$ to $\mathbb{T}^{(h)}$ defined as

$$r_h(\mathbf{t}) = \{u \in \mathbf{t}, |u| \leq h\}.$$

For this chapter we need to introduce a set of trees which is the the subset of trees with only one infinite spine. First, for $v = (v_k, k \in \mathbb{N}^*) \in (\mathbb{N}^*)^{\mathbb{N}}$, we set $\bar{v_n} = (v_1, ..., v_n)$ for $n \in \mathbb{N}$, with the convention that $\bar{v_0} = \emptyset$ and $\bar{v} = \{\bar{v_n}, n \in \mathbb{N}\}$ defines an infinite spine or branch. We denote by $\mathbb{T}_1$ the subset of trees with only one infinite spine:

$$\mathbb{T}_1 = \{\mathbf{t} \in \mathbb{T}, \text{ there exists a unique } v \in (\mathbb{N}^*)^{\mathbb{N}} \text{ such that } \bar{v} \subset \mathbf{t}\}.$$

Now, let us endow the set $\mathbb{T}$ with the distance $d$ defined as

$$d(\mathbf{t}, \mathbf{t}') = 2^{-sup\{h \in \mathbb{N}, r_h(\mathbf{t}) = r_h(\mathbf{t}')\}}. \tag{14}$$

We can notice that $sup\{h \in \mathbb{N}, r_h(\mathbf{t}) = r_h(\mathbf{t}')\}$ is the last generation when the tree $\mathbf{t}$ coincides with the tree $\mathbf{t}'$.

*Remark* 4.1. We call the distance $d$ the local topology because it compares the trees progressively at every level. We only know the minimal height when both trees differ for the first time. We don't know how the trees are after this level. Nonetheless, we know that until a certain height they are equal.

**Proposition 4.2.** *The distance $d$ is ultra-metric, which means that for all $\mathbf{t}, \mathbf{t}', \mathbf{t}'' \in \mathbb{T}$*

$$d(\mathbf{t}, \mathbf{t}') \leq max(d(\mathbf{t}, \mathbf{t}''), d(\mathbf{t}', \mathbf{t}'')).$$

*Proof.* We can prove this proposition by contradiction. In fact, if

$$d(\mathbf{t}, \mathbf{t}') > max(d(\mathbf{t}, \mathbf{t}''), d(\mathbf{t}', \mathbf{t}'')),$$

then $d(\mathbf{t},\mathbf{t}') > d(\mathbf{t},\mathbf{t}'')$ and $d(\mathbf{t},\mathbf{t}') > d(\mathbf{t}',\mathbf{t}'')$. Let

$$H = sup\{h \in \mathbb{N}, r_h(\mathbf{t}) = r_h(\mathbf{t}')\},$$

so there exists $H'$ and $H''$ greater than $H$ such that

$$H' = sup\{h \in \mathbb{N}, r_h(\mathbf{t}) = r_h(\mathbf{t}'')\},$$

and

$$H'' = sup\{h \in \mathbb{N}, r_h(\mathbf{t}') = r_h(\mathbf{t}'')\}.$$

Therefore, we obtain that $r_{H'}(\mathbf{t}) = r_{H'}(\mathbf{t}'')$ and $r_{H''}(\mathbf{t}') = r_{H''}(\mathbf{t}'')$. Let $H_m = min(H', H'')$. Obviously $H_m > H$ but $r_{H_m}(\mathbf{t}') = r_{H_m}(\mathbf{t}'') = r_{H_m}(\mathbf{t})$. $\square$

Furthermore, in an ultra-metric space, we have the property that all open balls are also closed.

We can now define an open (and closed) ball in $\mathbb{T}$ centered in $\mathbf{t}$ and of radius $h$ as

$$r_h^{-1}(\{r_h(\mathbf{t})\}) = \{\mathbf{t}' \in \mathbb{T}, d(\mathbf{t},\mathbf{t}') \leq 2^{-h}\}.$$

We have to introduce the notion of convergence of a sequence of trees that we will use in the next lemma to show that the space $(\mathbb{T}, d)$ is complete.

**Definition 4.3.** Let $(\mathbf{t}_n, n \in \mathbb{N})$ be a sequence of trees in $\mathbb{T}$. This sequence converges to $\mathbf{t} \in \mathbb{T}$ with respect to $d$ if and only if for every $h$, $r_h(\mathbf{t}_n) = r_h(\mathbf{t})$ for n large enough.

**Lemma 4.4.** *The metric space $(\mathbb{T}, d)$ is a Polish metric space (complete space and there exists a separable subset).*

*Proof.* Let's first prove that $\mathbb{T}_0$ is a separable subset of $\mathbb{T}$. It is countable because $\mathbb{T}_0 = \bigcup_{k \geq 0} \mathbb{T}^k$ where $\mathbb{T}^k$ is the set of rooted trees of size $k$. To prove that $\mathbb{T}_0$ is dense in $\mathbb{T}$, we consider $\mathbf{t} \in \mathbb{T}$ and the sequence $(r_h(\mathbf{t}), h \in \mathbb{N})$. This sequence converges to $\mathbf{t} \in \mathbb{T}$, which proves that $\mathbb{T}_0$ is dense in $\mathbb{T}$.
Now we need to prove that $\mathbb{T}$ is complete, so any Cauchy sequence in $\mathbb{T}$ converges. Let $(\mathbf{t}_n, n \in \mathbb{N})$ be a Cauchy sequence in $\mathbb{T}$. For all $h \in \mathbb{N}$, the sequence $(r_h(\mathbf{t}_n), n \in \mathbb{N})$ is a Cauchy sequence in $\mathbb{T}^{(h)}$, In fact, for $\mathbf{t}', \mathbf{t}'' \in \mathbb{T}^{(h)}$, $d(\mathbf{t}',\mathbf{t}'') \leq 2^{-(h+1)}$ implies that $\mathbf{t}' = \mathbf{t}''$ because $d(\mathbf{t}',\mathbf{t}'') \leq 2^{-(h+1)}$ implies $r_{h+1}(\mathbf{t}') = r_{h+1}(\mathbf{t}'')$.
We deduce that for $n$ large enough, the sequence $(r_h(\mathbf{t}_n), n \in \mathbb{N})$ is constant, let's say equal to a $\mathbf{t}^h$. This implies that for any $h$, the sequence $(r_h(\mathbf{t}_n), n \in \mathbb{N})$ converges to $\mathbf{t}^h$. The restriction function $r_h$ is continuous so $r_h(\mathbf{t}^{h'}) = \mathbf{t}^h$ for $h' \geq h$. This implies that $\mathbf{t} = \bigcup_{h \in \mathbb{N}} \mathbf{t}^h$ is a tree. Moreover

$$r_h(\mathbf{t}) = r_h\left(\bigcup_{h \in \mathbb{N}} \mathbf{t}^h\right) = \mathbf{t}^h.$$

We can conclude that $(\mathbf{t}_n, n \in \mathbb{N})$ converges to $\mathbf{t}$. $\square$

### 4.1.1 Convergence in distribution of a sequence of trees

We begin by statting the definition of the convergence in distribution for the local topology.

**Definition 4.5.** A sequence of random trees $(T_n, n \in \mathbb{N})$ converges in distribution with respect to the distance $d$, as defined in Equation (14), towards a random tree $T$, if and only if for all $h \in \mathbb{N}$ and $\mathbf{t} \in \mathbb{T}^{(h)}$:

$$\lim_{n\to\infty} \mathbb{P}(r_h(T_n) = \mathbf{t}) = \mathbb{P}(r_h(T) = \mathbf{t}).$$

We write

$$T_n \xrightarrow{(d)} T.$$

We want to prove the result stated in proposition 4.11 about the convergence in distribution for the local topology. First, we need to state some preliminary results and notations.

If $\mathbf{t} \in \mathbb{T}$ is a tree, $x \in L_0(\mathbf{t})$ a leaf of $\mathbf{t}$ and $\mathbf{t}' \in \mathbb{T}$ another tree, we denote by $\mathbf{t} \circledast_x \mathbf{t}'$ the tree obtained by grafting the tree $\mathbf{t}'$ on the leaf $x$:

$$\mathbf{t} \circledast_x \mathbf{t}' = \mathbf{t} \cup \{xu, u \in \mathbf{t}'\}$$

For $\mathbf{t} \in \mathbb{T}$ and $x \in L_0(\mathbf{t})$, the set of trees obtained by grafting a tree $\mathbf{t}'$ on $x$ is

$$\mathbb{T}(\mathbf{t}, x) = \{\mathbf{t} \circledast_x \mathbf{t}', \mathbf{t}' \in \mathbb{T}\}$$

We denote by $\mathcal{F}$ the subclass of Borel sets of $\mathbb{T}_0 \cup \mathbb{T}_1$:

$$\mathcal{F} = \{\{\mathbf{t}\}, \mathbf{t} \in \mathbb{T}_0\} \cup \{\mathbb{T}(\mathbf{t}, x), \mathbf{t} \in \mathbb{T}_0, x \in L_0(\mathbf{t})\}.$$

For later results, we need to show that the set $\mathcal{F}$ is stable under intersection. If it is the case, we say that the set is a $\pi$-system. Let us define it formally.

**Definition 4.6.** A $\pi$-system on a set $\Omega$ is a collection $P$ of some subsets of $\Omega$ such that:
1) $P$ is non-empty.
2) if $A$,$B$ are subsets in $P$ then $A \cap B \in P$.

We can now state the result that $\mathcal{F}$ is a $\pi$-system.

**Lemma 4.7.** *The family $\mathcal{F}$ is a $\pi$-system.*

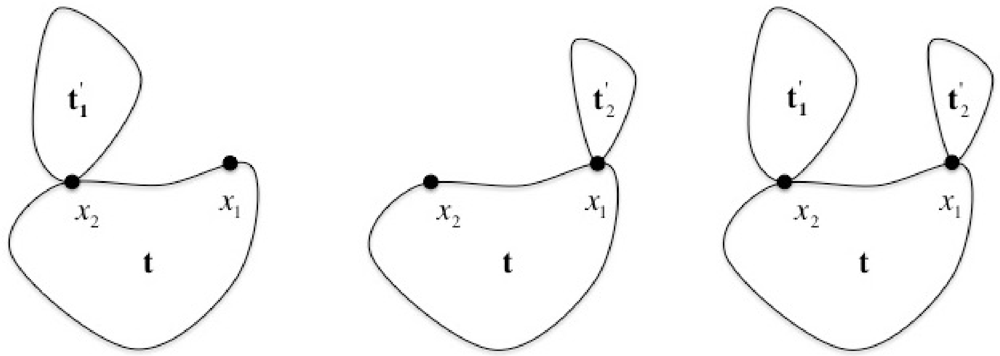
*Proof.* Let $\mathbf{t}_1, \mathbf{t}_2 \in \mathbb{T}_0$ and $x_1 \in L_0(\mathbf{t}_1)$ and $x_2 \in L_0(\mathbf{t}_2)$. We suppose that $\mathbb{T}(\mathbf{t}_1, x_1) \neq \mathbb{T}(\mathbf{t}_2, x_2)$. Then:

$$\mathbb{T}(\mathbf{t}_1, x_1) \cap \mathbb{T}(\mathbf{t}_2, x_2) = \begin{cases} \mathbb{T}(\mathbf{t}_1, x_1) & \text{if } \mathbf{t}_1 \in \mathbb{T}(\mathbf{t}_2, x_2) \text{ and } \mathbf{t}_1 \neq \mathbf{t}_2 \\ \mathbb{T}(\mathbf{t}_2, x_2) & \text{if } \mathbf{t}_2 \in \mathbb{T}(\mathbf{t}_1, x_1) \text{ and } \mathbf{t}_1 \neq \mathbf{t}_2 \\ \{\mathbf{t}_1 \cup \mathbf{t}_2\} & \text{if } \mathbf{t}_1 = \mathbf{t} \circledast_{x_2} \mathbf{t}_1',\ \mathbf{t}_2 = \mathbf{t} \circledast_{x_1} \mathbf{t}_2' \text{ and } x_1 \neq x_2 \\ \emptyset & \text{in any other case} \end{cases} \tag{15}$$

The first 2 cases are obvious. The third case is detailled in Figure 5 which represents the trees $\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_1 \cup \mathbf{t}_2$.

Finally the intersection $\mathbb{T}(\mathbf{t}_1, x_1) \cap \{\mathbf{t}_2\}$ is either equal to the empty set, $\mathbf{t}_2$ or a subtree of $\mathbf{t}_2$. We can conclude that $\mathcal{F}$ is stable under intersection and thus, it is a $\pi$-system. □

Figure 5: Representation of the third cases above. The figures reprensent $\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_1 \cup \mathbf{t}_2$ respectively




We have shown that $\mathcal{F}$ is stable under intersection. We will now show that any subset of $\mathcal{F}$ is open.

**Definition 4.8.** Let $X$ be a metric space. We define an open neighborhood system for the point $x$ as the collection of all open neighborhoods of $x$.

We say that a set $Y$ is an open neighborhood system in $X$ if for all $x \in X$ , the neighborhood system for the point $x$ is open. In other words, every neighborhood of $x$ is open.

**Lemma 4.9.** *The family $\mathcal{F}$ is an open neighborhood system in $\mathbb{T}_0 \cup \mathbb{T}_1$.*

*Proof.* We first verify that all elements of $\mathcal{F}$ are open. For $\mathbf{t} \in \mathbb{T}$ and $\epsilon > 0$, let $B(\mathbf{t}, \epsilon)$ be the open (and closed) ball centered at $\mathbf{t}$ with radius $\epsilon$. If $\mathbf{t}$ belongs to $\mathbb{T}_0$, we have $\{\mathbf{t}\} = B(\mathbf{t}, 2^h)$ for every $h > H(\mathbf{t})$ and thus, $\{\mathbf{t}\}$ is open. Moreover, for every $\mathbf{s} \in \mathbb{T}(\mathbf{t}, x)$ and for a leaf $x \in L_0(\mathbf{t})$, we have

$$B(\mathbf{s}, 2^{-H(\mathbf{t})-1}) \subset \mathbb{T}(\mathbf{t}, x),$$

so $\mathbb{T}(\mathbf{t}, x)$ is open. Finally, we need to check that $\mathcal{F}$ is a neighborhood system. In particular, we need to verify that, since all elements of $\mathcal{F}$ are open, for all tree $\mathbf{t} \in \mathbb{T}_0 \cup \mathbb{T}_1$ and $\epsilon > 0$, there exists an element $A'$ of $\mathcal{F}$ which is a subset of $B(\mathbf{t}, \epsilon)$ and contains $\mathbf{t}$.
If $\mathbf{t} \in \mathbb{T}_0$, we only need to consider $A' = \{\mathbf{t}\}$. We suppose that $\mathbf{t} \in \mathbb{T}_1$. Let $(u_n, n \in \mathbb{N}^*)$ be the infinite spine of $\mathbf{t}$ such that $\bar{u_n} = u_1 ... u_n \in \mathbf{t}$ for all $n \in \mathbb{N}^*$. Let $n \in \mathbb{N}^*$ such that $2^{-n} < \epsilon$ and set a tree $\mathbf{t}'$ such that $\mathbf{t}' = \{v \in \mathbf{t}, \bar{u_n} \notin A_v\}$. We notice that $\bar{u_n} \in \mathbf{t}'$ and we set $A' = \mathbb{T}(\mathbf{t}', \bar{u_n})$ such that $A'$ belongs to the $\pi$-system $\mathcal{F}$. As a result, we get $\mathbf{t} \in A' \subset B(\mathbf{t}, \epsilon)$. □

*Remark* 4.10. $\mathbf{t}'$ is a finite tree.

**Proposition 4.11.** *Let $(T_n, n \in \mathbb{N})$ and $T$ be random trees taking values in $\mathbb{T}_0 \cup \mathbb{T}_1$. The sequence $(T_n, n \in \mathbb{N})$ converges in distribution for the local topology (according to the distance $d$) to $T$ if and only if the following statements hold:*
*1) for every finite tree $\mathbf{t} \in \mathbb{T}_0$, $\lim_{n\to+\infty} \mathbb{P}(T_n = \mathbf{t}) = \mathbb{P}(T = \mathbf{t})$*
*2) for every $\mathbf{t} \in \mathbb{T}_0$ and leaf $x \in L_0(\mathbf{t})$, $\liminf_{n\to+\infty} \mathbb{P}(T_n \in T(\mathbf{t}, x)) \geq \mathbb{P}(T \in T(\mathbf{t}, x))$*

*Proof.* First, we note that the set $\mathbb{T}_0 \cup \mathbb{T}_1$ is a separable metric space because $\mathbb{T}_0$ is dense and countable in $\mathbb{T}_0 \cup \mathbb{T}_1$. In fact, for every $\mathbf{t} \in \mathbb{T}_1$, $\mathbf{t} = \lim_{h\to\infty} r_h(\mathbf{t})$. Now let $G$ be an open subset of $\mathbb{T}_0 \cup \mathbb{T}_1$. By using

Lemma 4.9, we can write $G$ as

$$G = \bigcup_{i \in \mathbb{N}} A_i,$$

where $(A_i, i \in \mathbb{N})$ is a family of elements of $\mathcal{F}$. For any $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$\mathbb{P}(T \in G) \leq \epsilon + \mathbb{P}\left(T \in \bigcup_{i \leq n_0} A_i\right).$$

We can use the proof of Lemma 4.7 which implies that, for any $I \subset \mathbb{N}$, the intersection $\bigcap_{i \in I} A_i$ is either empty, equal to a $A_{i_I}$ for a $i_I \in I$ or equal to $\{t_I\}$ for some $t_I \in \mathbb{T}_0$. By applying the inclusion-exclusion formula, we obtain that there exists $n_1 \leq n_0$, $\mathbf{t}_j \in \mathbb{T}_0$, $x_j \in L_0(\mathbf{t}_j)$ for $j \leq n_1$ and $n_2 < \infty$, $\mathbf{t}_l \in \mathbb{T}_0$, $\alpha_l \in \mathbb{Z}$ for $l \leq n_2$ such that for any random variable $T'$ taking values in $\mathbb{T}_0 \cup \mathbb{T}_1$:

$$\mathbb{P}\left(T' \in \bigcup_{i \leq n_0} A_i\right) = \sum_{j \leq n_1} \mathbb{P}(T' \in \mathbb{T}(\mathbf{t}_j, x_j)) + \sum_{l \leq n_2} \alpha_l \mathbb{P}(T' = \mathbf{t}_l).$$

Now, from the assumptions 1) and 2), we deduce that

$$\liminf_{n \to +\infty} \mathbb{P}(T_n \in G) \geq \liminf_{n \to +\infty} \mathbb{P}\left(T_n \in \bigcup_{i \leq n_0} A_i\right) \geq \mathbb{P}\left(T \in \bigcup_{i \leq n_0} A_i\right) \geq \mathbb{P}(T \in G) - \epsilon.$$

As $\epsilon$ is arbitrary, we deduce that

$$\liminf_{n \to +\infty} \mathbb{P}(T_n \in G) \geq \mathbb{P}(T \in G),$$

and thus, $T_n \xrightarrow{(d)} T$. □

## 4.2 Kesten's tree

We are going to introduce the Kesten's trees which will be important in the study of local convergence of a sequence of random Galton-Watson trees. In the next part, we are going to state the main result of local convergence where Kesten's tree are actually the limiting trees.

**Definition 4.12.** Let $p = (p(n), n \in \mathbb{N})$ be a probability distribution on $\mathbb{N}$ with finite mean $m$. We define the corresponding size-biased distribution $p^* = (p(n), n \in \mathbb{N})$ as:

$$p^*(n) = \frac{np(n)}{m}.$$

*Remark* 4.13. We can notice that $p^*(0) = 0$.

If $p$ is a probability distribution defined as above, from now on, we consider distributions $p$ satisfying

$$0 < p(0) < 1 \text{ and } p(0) + p(1) < 1. \tag{16}$$

**Definition 4.14.** Let $p$ be an offspring distribution with finite mean $m$ satisfying (16) and $p^*$ the corresponding size-biased distribution. A Kesten's tree associated to the probability distribution $p$ is a Galton-Watson tree $\tau^*$ such that:

1) individuals are normal or special
2) a normal individual gives birth only to normal individuals according to $p$.
3) a special individual gives birth to individuals according to $p^*$. One of them, chosen uniformely at random, is special, the others (if any) are normal.
4) the root of $\tau^*$ is special.

*Remark* 4.15. As $p^*(0) = 0$ and $m < \infty$, a Kesten's belongs to the set $\mathbb{T}$. Moreover it has at least one infinite spine and at each level, there is only one special individual.

The next lemma gives a link between the distribution of $\tau$ and $\tau^*$.

**Lemma 4.16.** *Let $p$ be an offspring distribution with finite mean $m$ satisfying (16), $\tau$ a Galton-Watson tree associated to $p$ and $\tau^*$ a Kesten's tree associated to $p$. For all $n \in \mathbb{N}$, $\mathbf{t} \in \mathbb{T}_0$ and $v \in \mathbf{t}$ such that $H(\mathbf{t}) = |v| = n$:*

$$\mathbb{P}(r_n(\tau^*) = \mathbf{t}, v \textit{ is special }) = \frac{1}{m^n}\mathbb{P}(r_n(\tau) = \mathbf{t}), \tag{17}$$

$$\mathbb{P}(r_n(\tau^*) = \mathbf{t}) = \frac{z_n(\mathbf{t})}{m^n}\mathbb{P}(r_n(\tau) = \mathbf{t}). \tag{18}$$

*where $z_n(\mathbf{t})$ is the number of individuals at generation $n$.*

*Proof.* We can notice that if $u$ is special, the probability that it has $k_u$ children and $ui$ is special with $i$ given is $p^*(k_u)\frac{1}{k_u} = \frac{p(k_u)}{m}$. Using Equation (13) we get:

$$\mathbb{P}(r_n(\tau^*)) = \mathbf{t},\ \text{v is special}) = \prod_{u\in\mathbf{t}, u\notin A_v} p(k_u(\mathbf{t})) \prod_{u\in A_v} \frac{p(k_u(\mathbf{t}))}{m} = \frac{1}{m^n}\mathbb{P}(r_n(\tau) = \mathbf{t}) \tag{19}$$

As there is only one special individual among $z_n(\mathbf{t})$ individuals at level $n$ of $\mathbf{t}$, we get the Equation (18) by multiplying the Equation (17) by $z_n(\mathbf{t})$. $\square$

## 4.3 Kesten's theorem

We can now state the Kesten's theorem which is the first main result of the thesis. It provides the result for the local convergence of a sequence of random Galton-Watson trees in the sub-critical and critical cases.

**Theorem 4.17.** *Let $p$ be a critical or sub-critical offspring distribution satisfying (16). Let $\tau$ be a Galton-Watson tree with offspring distribution $p$ and let $\tau^*$ a Kesten's tree associated to $p$. For every $n \in \mathbb{N}$, let $\tau_n$ be a random tree distributed as $\tau$ conditionally on $\{H(\tau_n) \geq n\}$. Then*

$$\tau_n \xrightarrow{(d)} \tau^*,$$

*where $\xrightarrow{(d)}$ is the convergence for the local topology as defined in Definition 4.5.*

The proof of this theorem will be made step by step as we need to introduce other notions for it.

We are now interested in computing the probability that a Galton-Watson or Kesten's tree belongs to the set $\mathbb{T}(\mathbf{t}, x)$ for a finite tree $\mathbf{t} \in \mathbb{T}_0$ and $x \in L_0(\mathbf{t})$. For example, for $\tau$ a Galton-Watson tree with offspring distribution $p$, $\mathbf{t} \in \mathbb{T}_0$ and $x \in L_0(\mathbf{t})$ we have

$$\mathbb{P}(\tau = \mathbf{t}) = \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x), k_x(\tau) = 0) = \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x))p(0). \tag{20}$$

The last equility is obtained due to the independence of $k_x(\tau)$ for all vertices $x$.

The next lemma provides the computation of the probability for a Kesten's tree to belong to $\mathbb{T}(\mathbf{t}, x)$.

**Lemma 4.18.** *Let $p$ an offspring distribution satisfying the condition (16) and with finite mean $m$. Let $\tau$ a Galton-Watson tree with offspring distribution $p$ and $\tau^*$ a Kesten's tree associated to $p$. Then for all finite tree $\mathbf{t} \in \mathbb{T}_0$ and leaf $x \in L_0(\mathbf{t})$:*

$$\mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) = \frac{1}{m^{|x|}} \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x)). \tag{21}$$

*In the case of critical offspring distrbution $p$ ($m = 1$) we get*

$$\mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) = \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x)).$$

*Remark* 4.19. In the case of critical offspring distribution $p$, $\mathbb{P}(\tau \in \mathbb{T}_0) = 1$ and $\mathbb{P}(\tau^* \in \mathbb{T}_1) = 1$.

*Proof.* Let $\mathbf{t} \in \mathbb{T}_0$ and a leaf $x \in L_0(\mathbf{t})$. If $\tau^* \in \mathbb{T}(\mathbf{t}, x)$, then the vertex $x$ has to be special because $\mathbf{t}$ is finite and $\tau^*$ is almost surely infinite.
The condition $\tau^* \in \mathbb{T}(\mathbf{t}, x)$ is equivalent to $r_{|x|}(\tau^*) = \mathbf{t}$ and $x$ is special. The same statement applies for $\tau$: $\tau \in \mathbb{T}(\mathbf{t}, x)$ is equivalent to $r_{|x|}(\tau) = \mathbf{t}$.
By using lemma 4.16 we get:

$$\mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) = \prod_{u \in \mathbf{t}, u \notin A_x} p(k_u(\mathbf{t})) \prod_{u \in A_x} \frac{p(k_u(\mathbf{t}))}{m} = \frac{1}{m^{|x|}} \mathbb{P}(r_{|x|}(\tau) = \mathbf{t}) = \frac{1}{m^{|x|}} \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x)).$$

□

## 4.4 Functional for Kesten's theorem

In this section we will present the last required results needed to prove the Kesten's theorem (Theorem 4.17). We need to introduce the concept of a functional which is a function $A$ such that $A : \mathbb{T}_0 \to \mathbb{N}$ satisfies the condition that the set $\{\mathbf{t}, A(\mathbf{t}) \geq n\}$ is non empty for all $n \in \mathbb{N}^*$.
In the results of this part, we need more properties on $A$. We will specify for each result which properties are needed.
For all $\mathbf{t} \in \mathbb{T}_0$ and all leaf $x \in L_0(\mathbf{t})$, there exists $n_0 \in \mathbb{N}^*$ and $D(\mathbf{t}, x) \geq 0$ (only for the (additivity) property) such that for all $\mathbf{t}' \in \mathbb{T}_0$ satisfying $A(\mathbf{t} \circledast_x \mathbf{t}') \geq n_0$:

$$\begin{aligned}
A(\mathbf{t} \circledast_x \mathbf{t}') &\geq A(\mathbf{t}') && \text{(monotonicity)},\\
A(\mathbf{t} \circledast_x \mathbf{t}') &= A(\mathbf{t}') + D(\mathbf{t}, x) && \text{(additivity)},\\
A(\mathbf{t} \circledast_x \mathbf{t}') &= A(\mathbf{t}') && \text{(identity)}.
\end{aligned}$$

*Remark* 4.20. We can notice that the 3 properties are listed from the weakest to the strongest one. In fact, the property (identity) is a particular case of the property (additivity) with $D(\mathbf{t}, x) = 0$ and the property (additivity) is a particular case of the property (monotonicity).

Here are some examples of functionals that we can use:
1) Maximal degree $M(\mathbf{t}) = max\{k_u(\mathbf{t}), u \in \mathbf{t}\}$ with the property (identity) and $n_0 = M(\mathbf{t}) + 1$. In fact, if $M(\mathbf{t} \circledast_x \mathbf{t}') \geq M(\mathbf{t}) + 1$, $M(\mathbf{t}') > M(\mathbf{t})$.
2) The cardinal $|\mathbf{t}|$ =Card$(\mathbf{t})$. This functional satisfies the property (monotonicity) if $n_0 = 0$ or the property

(additivity) if $n_0 = 0$ and $D(\mathbf{t}, x) = |\mathbf{t}| - 1 \geq 0$
3) The global height $H(\mathbf{t}) = max\{|u|, u \in \mathbf{t}\}$. This functional satisfies the property (additivity) if $n_0 = H(\mathbf{t})$ and $D(\mathbf{t}, x) = |x|$.

We need 2 theorems, stated below, in which we separate the critical and sub-critical cases. After proving them, it will be easy to deduce the Kesten's theorem.

**Theorem 4.21.** *Let $\tau$ be a Galton-Watson tree with a critical offspring distribution $p$ and $\tau^*$ a Kesten's tree associated to $p$. Let $\tau_n$ be a random tree distributed according to $\tau$ conditionally on the event $\mathbb{A}_n = \{A(\tau) \in [n, n + n_1[\}$ and we assume that $\mathbb{P}(\mathbb{A}_n) > 0$ for $n$ large enough. If one of the following conditions is satisfied:*
*1) $n_1 \in \mathbb{N}^* \cup \{+\infty\}$ and $A$ satisfies the condition (identity)*
*2) $n_1 = +\infty$ and $A$ satisfies the condition (monotonicity)*
*3) $n_1 \in \mathbb{N}^* \cup \{+\infty\}$, $A$ satisfies the condition (additivity) and*

$$\limsup_{n \to +\infty} \frac{\mathbb{P}(\mathbb{A}_{n+1})}{\mathbb{P}(\mathbb{A}_n)} \leq 1. \tag{22}$$

*Then:*

$$\tau_n \xrightarrow[n \to +\infty]{(d)} \tau^*.$$

**Theorem 4.22.** *Let $\tau$ be a Galton-Watson tree with a sub-critical offspring distribution $p$ and $\tau^*$ a Kesten's tree associated to $p$. Let $\tau_n$ be a random tree distributed according to $\tau$ conditionally on the event $\mathbb{A}_n = \{A(\tau) \in [n, n + n_1[\}$ and we assume that $\mathbb{P}(\mathbb{A}_n) > 0$ for $n$ large enough. If $A$ satisfies the property (Additivity) with $D(\mathbf{t}, x) = |x|$ and*

$$\limsup_{n \to +\infty} \frac{\mathbb{P}(\mathbb{A}_{n+1})}{\mathbb{P}(\mathbb{A}_n)} \leq m. \tag{23}$$

*Then:*

$$\tau_n \xrightarrow[n \to +\infty]{(d)} \tau^*.$$

*Remark* 4.23. We can notice that if we have

$$\limsup_{n \to +\infty} \frac{\mathbb{P}(\mathbb{A}_{n+1})}{\mathbb{P}(\mathbb{A}_n)} \leq m,$$

then, we have

$$\liminf_{n \to +\infty} \frac{\mathbb{P}(\mathbb{A}_n)}{\mathbb{P}(\mathbb{A}_{n+1})} \geq \frac{1}{m}.$$

Moreover, for $k \leq n$, as we can write

$$\frac{\mathbb{P}(\mathbb{A}_{n-k})}{\mathbb{P}(\mathbb{A}_n)} = \frac{\mathbb{P}(\mathbb{A}_{n-k})}{\mathbb{P}(\mathbb{A}_{n-k+1})} \cdot \frac{\mathbb{P}(\mathbb{A}_{n-k+1})}{\mathbb{P}(\mathbb{A}_{n-k+2})} \cdots \frac{\mathbb{P}(\mathbb{A}_{n-2})}{\mathbb{P}(\mathbb{A}_{n-1})} \cdot \frac{\mathbb{P}(\mathbb{A}_{n-1})}{\mathbb{P}(\mathbb{A}_n)},$$

we can deduce that

$$\liminf_{n \to +\infty} \frac{\mathbb{P}(\mathbb{A}_{n-k})}{\mathbb{P}(\mathbb{A}_n)} \geq \frac{1}{m^k}.$$

We can now provide a proof of theorems 4.21 and 4.22.

*Proof.* We first assume that the functional $A$ satisfies the property (identity) or (additivity). In both theorems, we consider only the critical and sub-critical cases so, almost surely, the trees $\tau_n$ belong to the set $\mathbb{T}_0$ and $\tau^*$ belong to the set $\mathbb{T}_1$. We are going to use Proposition 4.11 to prove the local convergence of the theorems. To cover all cases of the theorems, we set $n_1 \in \{1, +\infty\}$ and $\mathbb{A}_n = \{A(\tau) \in [n, n+n_1[\}$. Let $\mathbf{t} \in \mathbb{T}_0$. We have

$$\mathbb{P}(\tau_n = \mathbf{t}) = \frac{\mathbb{P}(\tau_n = \mathbf{t}, \mathbb{A}_n)}{\mathbb{P}(\mathbb{A}_n)} \leq \frac{1}{\mathbb{P}(\mathbb{A}_n)} \mathbb{1}_{\{A(\mathbf{t}) \in [n, n+n_1[\}}.$$

As $\mathbf{t} \in \mathbb{T}_0$, $A(\mathbf{t})$ is finite, so $\mathbb{1}_{\{A(\mathbf{t}) \in [n, n+n_1[\}} = 0$ for $n > A(\mathbf{t})$. We deduce that

$$\lim_{n \to \infty} \mathbb{P}(\tau_n = \mathbf{t}) = 0 = \mathbb{P}(\tau^* = \mathbf{t}), \tag{24}$$

because $\tau^*$ is almost surely infinite. The condition 1) of Proposition 4.11 is verified. Now, we need to verify the condition 2) of Proposition 4.11. Let $\mathbf{t} \in \mathbb{T}_0$ and $x \in L_0(\mathbf{t})$ a leaf. As $\tau^*$ is almost surely infinite

$$\mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x), \mathbb{A}_n) = \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}(\tau = \mathbf{t} \circledast_x \mathbf{t}') \mathbb{1}_{\{A(\mathbf{t} \circledast_x \mathbf{t}') \in [n, n+n_1[\}}. \tag{25}$$

By using the definition of a Galton-Watson tree which states that the offsprings of each vertex are independent and Equation (20), we get

$$\mathbb{P}(\tau = \mathbf{t} \circledast_x \mathbf{t}') = \frac{1}{p(0)} \mathbb{P}(\tau = \mathbf{t}) \mathbb{P}(\tau = \mathbf{t}'),$$

and by applying Equation (21) we obtain

$$\mathbb{P}(\tau = \mathbf{t} \circledast_x \mathbf{t}') = m^{|x|} \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) \mathbb{P}(\tau = \mathbf{t}'). \tag{26}$$

We deduce from Equation (25) and (26) that

$$\mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x), \mathbb{A}_n) = m^{|x|} \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}(\tau = \mathbf{t}') \mathbb{1}_{\{A(\mathbf{t} \circledast_x \mathbf{t}') \in [n, n+n_1[\}}. \tag{27}$$

Now assume that the offspring distribution $p$ is critical ($m = 1$) and that the property (identity) holds for $A$. In this case, for $n \geq n_0$, we have

$$\sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}(\tau = \mathbf{t}') \mathbb{1}_{\{A(\mathbf{t} \circledast_x \mathbf{t}') \in [n, n+n_1[\}} = \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}(\tau = \mathbf{t}') \mathbb{1}_{\{A(\mathbf{t}') \in [n, n+n_1[\}} = \mathbb{P}(\mathbb{A}_n). \tag{28}$$

We can apply this result in the critical case in Equation (27) to obtain that for $n \geq n_0$ we have

$$\mathbb{P}(\tau_n \in \mathbb{T}(\mathbf{t}, x)) = \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)).$$

This equality holds because the tree $\tau_n$ is such that $A(\tau) \in [n, n+n_1[$. The condition 2) of Proposition 4.11 holds. This proves that the sequence $(\tau_n, n \in \mathbb{N}^*)$ converges in distribution to $\tau^*$. Therefore the Theorem 4.21 is proved if the condition 1) is satisfied.
Now, let us assume that the property (additivity) holds for $A$. Then, by using Equation (27), for $n \geq n_0$, we have

$$\begin{aligned}
\mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x), \mathbb{A}_n) &= m^{|x|} \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}(\tau = \mathbf{t}') \mathbb{1}_{\{A(\mathbf{t}') + D(\mathbf{t}, x) \in [n, n+n_1[\}} \\
&= m^{|x|} \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) \mathbb{P}(n - D(\mathbf{t}, x) \leq A(\tau) \leq n - D(\mathbf{t}, x) + n_1) \\
&= m^{|x|} \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) \mathbb{P}(\mathbb{A}_{n - D(\mathbf{t}, x)}).
\end{aligned}$$

Thus, we obtain that

$$\mathbb{P}(\tau_n \in \mathbb{T}(\mathbf{t},x)) = \frac{\mathbb{P}(\tau \in \mathbb{T}(\mathbf{t},x), \mathbb{A}_n)}{\mathbb{P}(\mathbb{A}_n)} = m^{|x|}\mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t},x))\frac{\mathbb{P}(\mathbb{A}_{n-D(\mathbf{t},x)})}{\mathbb{P}(\mathbb{A}_n)}.$$

Finally, if the conditions (22) for the critical case and (23) for the sub-critical case are satisfied, then by using the Remark 4.23 with $D(\mathbf{t},x) = |x|$, we get

$$\liminf_{n\to\infty} \mathbb{P}(\tau_n \in \mathbb{T}(\mathbf{t},x)) = \liminf_{n\to\infty} m^{|x|}\mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t},x))\frac{\mathbb{P}(\mathbb{A}_{n-D(\mathbf{t},x)})}{\mathbb{P}(\mathbb{A}_n)} \geq \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t},x)). \tag{29}$$

So in both cases, the condition 2) of Proposition 4.11 is satisfied. Therefore, the Theorem 4.22 and the Theorem 4.21 if the condition 3) is satisfied are proved.
Finally, let us prove the last case: Theorem 4.21 if the condition 2) occurs. We suppose that $p$ is a critical offspring distribution and the property (monotonicity) holds for $A$. Let $\mathbf{t} \in \mathbb{T}_0$ and $x \in L_0(\mathbf{t})$ a leaf of $\mathbf{t}$. As $A(\mathbf{t})$ is finite, the Equation (24) holds. Moreover, $\tau$ is almost surely finite, so for $n \geq n_0$:

$$\mathbb{P}(\tau \in \mathbb{T}(\mathbf{t},x), \mathbb{A}_n) = \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}(\tau = \mathbf{t} \circledast_x \mathbf{t}')\mathbb{1}_{\{n \leq A(\mathbf{t}\circledast_x \mathbf{t}')\}} \geq \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}(\tau = \mathbf{t} \circledast_x \mathbf{t}')\mathbb{1}_{\{n \leq A(\mathbf{t}')\}},$$

where we have used the property (monotonicity) in the inequality. Then, from Equation (27) and (28), we deduce that

$$\mathbb{P}(\tau \in \mathbb{T}(\mathbf{t},x), \mathbb{A}_n) \geq \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t},x))\mathbb{P}(\mathbb{A}_n).$$

Therefore, we can say that Equation (29) is also satisfied in this case because

$$\mathbb{P}(\tau_n \in \mathbb{T}(\mathbf{t},x)) = \frac{\mathbb{P}(\tau \in \mathbb{T}(\mathbf{t},x), \mathbb{A}_n)}{\mathbb{P}(\mathbb{A}_n)}.$$

By using the condition 2) of Proposition 4.11, we get the desired result. □

## 4.5 Proof of Kesten's theorem

Now, let us consider the functional that is the global height of the tree: $A(\mathbf{t}) = H(\mathbf{t})$. It satisfies the property (additivity) because for every $\mathbf{t}, \mathbf{t}' \in \mathbf{T}_0$ and a leaf $x \in L_0(\mathbf{t})$ such that $H(\mathbf{t} \circledast_x \mathbf{t}') \geq H(\mathbf{t}') + 1$, we have

$$H(\mathbf{t} \circledast_x \mathbf{t}') = H(\mathbf{t}') + |x|.$$

We first state the last preliminary result that we need to prove Theorem 4.17.

**Lemma 4.24.** *Let $p$ be a critical or sub-critical offspring distribution satisfying (16) with finite mean $m$ and let $\tau$ be a Galton-Watson tree with offspring distribution $p$. We set $n_1 \in \{1, +\infty\}$ and $\mathbb{A}_n = \{A(\tau) \in [n, n+n_1[\}$ for $n \in \mathbb{N}^*$. Then*

$$\lim_{n\to+\infty} \frac{\mathbb{P}(\mathbb{A}_{n+1})}{\mathbb{P}(\mathbb{A}_n)} = m.$$

*Proof.* We consider the case when $n_1 = 1$, so $\mathbb{A}_n = \{H(\tau) = n\}$. We remind the notations introduced in the previous parts, for any tree $\mathbf{t} \in \mathbb{T}$, $z_n(\mathbf{t})$ is the size of the n-th generation of the tree $\mathbf{t}$. We set $Z_n = z_n(\tau)$ so $(Z_n, n \in \mathbb{N})$ is a Galton-Watson process. We can notice that

$$\mathbb{A}_n = \{Z_{n+1} = 0\} \cap \{Z_n = 0\}^C.$$

We remind that $g$ is the probability generating function of the offspring distribution $p$ and $g_n$, the one of $Z_n$. We can remind, from Equation (12), that

$$g_{n+1}(s) = g(g_n(s)).$$

We have $\mathbb{P}(\mathbb{A}_n) = \mathbb{P}(Z_{n+1} = 0) - \mathbb{P}(Z_n = 0) = g(g_n(0)) - g_n(0)$. Because $\tau$ is critical or sub-critical, it is almost surely finite so we can deduce that

$$\lim_{n\to+\infty} g_n(0) = \lim_{n\to+\infty} \mathbb{P}(z_n(\tau) = 0) = 1.$$

As $m = g'(1)$, we get

$$\mathbb{P}(\mathbb{A}_n) = (1 - g_n(0))(1 - m + o(1)) = (g(1) - g(g_{n-1}(0)))(1 - m + o(1))).$$

We deduce that

$$\frac{\mathbb{P}(\mathbb{A}_{n+1})}{\mathbb{P}(\mathbb{A}_n)} = \frac{g(1) - g(g_n(0))}{1 - g_n(0)} \cdot \frac{1 - m + o(1)}{1 - m + o(1)} = m + o(1).$$

□

Finally, by applying the theorems 4.22 and 4.21 in the case 3) with the functional to be the global height of the tree and Lemma 4.24, we complete the proof of Kesten's theorem.

# 5. Convergence of the rescaled height process

In this last chapter we are going to see the second type of convergence concerning random trees using the material of [13]. More precisely, we will introduce a new stochastic process, the height process, based on a sequence of random Galton-Watson trees and we will show that this process, if rescaled, converges to an almost known for us stochastic process. In fact, the limit is called the reflected Brownian motion. It has similar properties as the Brownian motion but it is non-negative. Let us see more in details this result. We start by introducing the height function associated to a discrete tree $\mathbf{t}$. It will be needed to define the height process.

We know from the Section 3.1 that in any discrete tree $\mathbf{t}$, we can represent each vertex of the tree with a sequence of integers. We need to define a special ordering of vertex in order to define later the height function of the tree.

**Definition 5.1.** Let $s_1 = a_1 a_2 ... a_k$ and $s_2 = b_1 b_2 ... b_l$ be 2 finite sequences of elements of $\mathbb{N}$. We say that they are ranked in the lexicographical order and we denote it as $s_1 \prec s_2$ if and only if $a_i < b_i$ for the first $i$ when $s_1$ differ from $s_2$. Obviously, if one sequence is shorter, we complete it with zeros.

Now, we can define the height function that will be esential to define the height process.

**Definition 5.2.** Let $\mathbf{t} \in \mathbb{T}_0$ be a finite discrete tree. Denote by $u_0, u_1, ..., u_{\#(\mathbf{t})-1}$ the sequence of vertices of $\mathbf{t}$ ordered in the lexiographical way. By convention, $u_0 = \emptyset$. The height function $(h_{\mathbf{t}}(n), 0 \leq n \leq \ |\mathbf{t}| - 1)$ is defined as

$$h_{\mathbf{t}}(n) = |u_n| \text{ for } 0 \leq n \leq \ |\mathbf{t}| - 1.$$

Note that the height function defined in the last Definition differ from the global height function presented in Section 3.1.

Let us give an example. We use the tree from Figure 4. The lexicographical ordering of the vertices is $(\emptyset, 1, (1,1), (1,2), (1,3), 2, 3, (3,1), (3,1,1), (3,1,2))$. The height function obtained is presented in Figure 6.

We need to define the concepts that will be needed in Theorem 5.6. First, let us consider a Galton-Watson tree with offspring distribution $p$. From now on, we denote it as a $p$-Galton-Watson tree. Let $(R_n, n \geq 0)$ be a random walk on $\mathbb{Z}$ with initial value $R_0 = 0$ and jump distribution $\nu(k) = p(k+1)$ for every $k \geq -1$. Let us fix a critical offspring distribution $p$ with mean $m = 1$ and finite variance $\sigma^2$. Let $\tau_1, \tau_2, ...$ be a sequence of independent $p$-Galton Watson trees. For each $\tau_i$, we can associate its height function $h_{\tau_i}$ defined for $0 \leq n \leq |\tau_i| - 1$. Then, we define the height process $(H_n, n \geq 0)$ of the forest by concatenating the height functions of every tree:

$$H_n = h_{\tau_i}(n - (|\tau_1| + ... + |\tau_{i-1}|)),$$

if $|\tau_1| + ... + |\tau_{i-1}| \leq n < |\tau_1| + ... + |\tau_i|$.

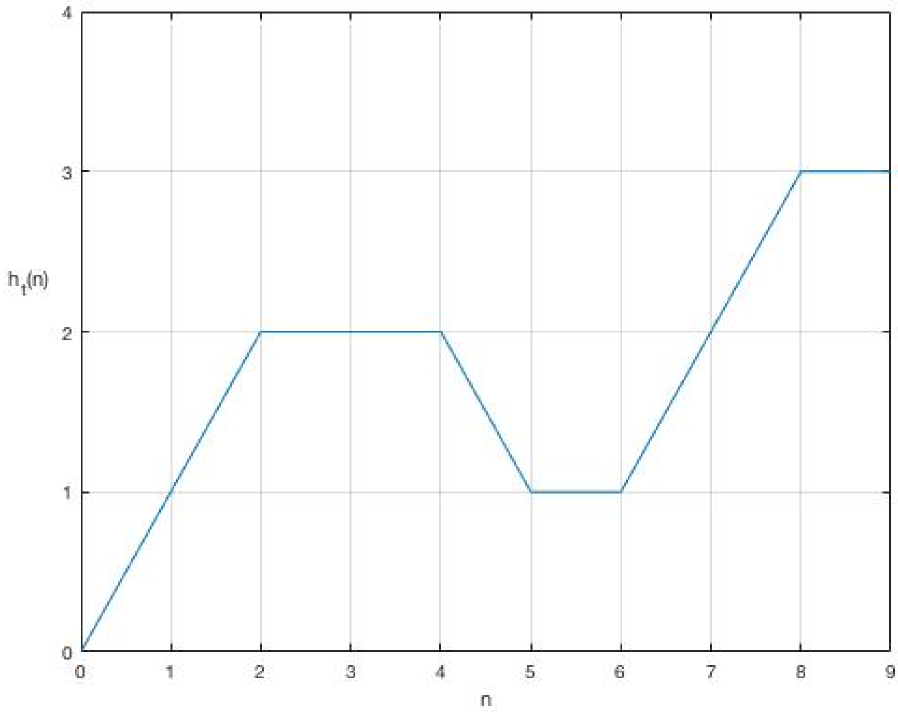
*Remark* 5.3. The height process $(H_n, n \geq 0)$ determines the sequence of trees. In fact, the $k$-th excursion of $H$ from 0 (values of $H$ between the $k$-th and the $k+1$-th zero) is the height function of the tree $k$-th tree in the sequence.

**Proposition 5.4.** *For every $n \geq 0$, we have*

$$H_n = \#\{k \in \{0, 1, ..., n-1\}, R_k = \inf_{k \leq j \leq n} R_j\}.$$

*where $(R_n, n \geq 0)$ is a random walk distributed as above.*

Figure 6: Height function of the tree presented in Figure 4




In the next theorem, we will see that the limit of a rescaled height process contains the reflected Brownian motion. Let us define it formally.

**Definition 5.5.** Let $(B(t), t \geq 0)$ be a standard Brownian motion. For all $t \geq 0$, we define a reflected Brownian motion as

$$\gamma(t) = B(t) - \inf_{0 \leq s \leq t} B(s).$$

We can now state the theorem which gives the convergence of the rescaled height process.

**Theorem 5.6.** *Let $\tau_1, \tau_2, \ldots$ be a sequence of independent p-Galton-Watson trees, and let $(H_n, n \geq 0)$ be the associated height process. Then*

$$\left( \frac{1}{\sqrt{p}} H_{[pt]}, t \geq 0 \right) \xrightarrow[p \to \infty]{(d)} \left( \frac{2}{\sigma} \gamma(t), t \geq 0 \right),$$

*where $\gamma$ is a reflected Brownian motion.*

*Remark* 5.7. The convergence in Theorem 5.6 holds in the sense of weak convergence on the Skorokhod space $\mathbb{D}(\mathbb{R}^+, \mathbb{R}^+)$. We remind that the Skorokhod space $\mathbb{D}(\mathbb{R}^+, \mathbb{R}^+)$ is the space of functions from $\mathbb{R}^+$ to $\mathbb{R}^+$ that are everywhere right-continuous and have left limits everywhere.

In Theorem 5.6, we notice that the left term $\frac{1}{\sqrt{p}} H_{[pt]}$ is the rescaled height process of the Galton-Watson forest. As $p \to \infty$, we determine that the height tends to 0 for a fixed vertex and vertices get closer to each other. The tree discribed by the limit of the rescaled height process is called a continuous random tree. The goal is to prove Theorem 5.6. We start by introducing new tools needed for the proof. Let $R = (R_n, n \geq 0)$ be a random walk as in Proposition 5.4. We notice that the jump distribution $\nu$ has mean 0 and variance $\sigma^2$. Thus, by looking at the random walk as a Markov chain where the states are the ordinates, we know that the random walk is recurrent. By Theorem 2.9, we get

$$\left(\frac{1}{\sqrt{p}} R_{[pt]}, t \geq 0\right) \xrightarrow[p\to\infty]{(d)} (\sigma B(t), t \geq 0), \tag{30}$$

where $(B(t), t \geq 0)$ is a standard Brownian motion. In fact, the proof of Theorem 2.9 doesn't require that the random walk is symmetric and it is valid with the random walk $R = (R_n, n \geq 0)$. Finally, let us define

$$M_n = \sup_{0\leq k\leq n} R_k \text{ and } I_n = \inf_{0\leq k\leq n} R_k,$$

and we will need

$$K_n = \#\{k \in \{1, 2, ..., n\}, R_k = M_k\}.$$

**Lemma 5.8.** *Let $\{T_j\}_{j\geq 0}$ be a sequence of stopping times defined inductively by $T_0 = 0$ and for $j \geq 1$*

$$T_j = \inf\{n > T_{j-1}, S_n = M_n\}.$$

*Then the random variables $R_{T_j} - R_{T_{j-1}}$ for $j = 1, 2, ...$ are independent and identically distributed, with distribution*

$$\mathbb{P}(R_{T_j} - R_{T_{j-1}} = k) = \nu([k, +\infty[).$$

*Proof.* The fact that the random variables $R_{T_j} - R_{T_{j-1}}$ for $j = 1, 2, ...$ are independent and identically distributed is a consequence of the strong Markov property. Now we have to compute the distribution of $R_{T_j} - R_{T_{j-1}}$ which is the same the one of $R_{T_1}$ because $R_{T_0} = 0$. If $S_0 = \inf\{n \geq 1, R_n = 0\}$, then, for every $i \in \mathbb{Z}$

$$\mathrm{E}\left(\sum_{n=0}^{S_0-1} \mathbb{1}_{\{R_n=i\}}\right) = 1.$$

Notice that $T_1 \leq S_0$ and the random walk takes positive values on $[T_1, S_0]$. It follows that for $i \leq 0$

$$\mathrm{E}\left(\sum_{n=0}^{T_1-1} \mathbb{1}_{\{R_n=i\}}\right) = 1. \tag{31}$$

Let $g : \mathbb{Z} \to \mathbb{Z}^+$ be a function. Then, we have for any function $g$ as defined

$$\begin{aligned}\mathrm{E}\left(\sum_{n=0}^{T_1-1} g(R_n)\right) &= \mathrm{E}\left(\sum_{n=0}^{T_1-1} g(R_n) \sum_{i\leq 0} \mathbb{1}_{\{R_n=i\}}\right) = \sum_{i\leq 0} \mathrm{E}\left(\sum_{n=0}^{T_1-1} g(R_n)\mathbb{1}_{\{R_n=i\}}\right)\\ &= \sum_{i\leq 0} g(i)\, \mathrm{E}\left(\sum_{n=0}^{T_1-1} \mathbb{1}_{\{R_n=i\}}\right) = \sum_{i=0}^{-\infty} g(i).\end{aligned}$$

The first equality is due to the fact that for every $n$,

$$\sum_{i\leq 0} \mathbb{1}_{\{R_n=i\}} = 1,$$

because $R_n$ takes exactly one negative value. The last equality is obtained using Equation (31). Then, for any function $f : \mathbb{Z} \to \mathbb{Z}^+$, we have

$$\begin{aligned}
\mathrm{E}\left(f(R_{T_1})\right) &= \mathrm{E}\left(\sum_{k=0}^{\infty} \mathbb{1}_{\{k<T_1\}} f(R_{k+1}) \mathbb{1}_{\{R_{k+1}\geq 0\}}\right) = \sum_{k=0}^{\infty} \mathrm{E}\left(\mathbb{1}_{\{k<T_1\}} f(R_{k+1}) \mathbb{1}_{\{R_{k+1}\geq 0\}}\right) \\
&= \sum_{k=0}^{\infty} \mathrm{E}\left(\mathbb{1}_{\{k<T_1\}} \sum_{j=0}^{\infty} \nu(j) f(R_k + j) \mathbb{1}_{\{R_k + j\geq 0\}}\right) = \sum_{i=0}^{-\infty} \sum_{j=0}^{\infty} \nu(j) f(i+j) \mathbb{1}_{\{i+j\geq 0\}} \\
&= \sum_{m=0}^{\infty} f(m) \sum_{j=m}^{\infty} \nu(j)
\end{aligned}$$

By choosing $f$ as the identity function, we get the desired result. $\square$

We have to define the time-reversed random walk that is obtained from an existing random walk. This Definition will be useful to prove the next Proposition.

**Definition 5.9.** Let $R = (R_n, n \geq 0)$ be a random walk as in Proposition 5.4. For $n \leq 0$, we define the time-reversed random walk $\hat{R}^n = (\hat{R}^n_k, 0 \leq k \leq n)$ as

$$\hat{R}^n_k = R_n - R_{n-k}.$$

We need the following result to prove the Proposition 5.11.

**Proposition 5.10.** *The time-reversed random walk* $(\hat{R}^n_k, 0 \leq k \leq n)$ *has the same distribution as* $(R_k, 0 \leq k \leq n)$.

*Proof.* Denote by $\Delta_i$ the steps of the random walk $(R_k, 0 \leq k \leq n)$. Then

$$R_k = \sum_{i=1}^{k} \Delta_i.$$

Then

$$\hat{R}^n_k = R_n - R_{n-k} = \sum_{i=n-k+1}^{n} \Delta_i.$$

Both processes are the sum of $k$ steps of a random walk with the same jump distribution. This completes the proof of the Proposition. $\square$

We will use the time-reversed random walk and the previous result to prove the next Proposition.

**Proposition 5.11.** *The pair* $(M_n, K_n)$ *has the same distribution as* $(R_n - I_n, H_n)$.

*Proof.* We need to replacing the random walk $R$ by the time-reversed random walk $\hat{R}^n$. We start by proving that $M_n$ has the same distribution as $R_n - I_n$. Let $(\hat{R}^n_k, 0 \leq k \leq n)$ be the time-reversed random walk. Let us define the processes $(\hat{M}^n, 0 \leq k \leq n)$ and $(\hat{I}^n, 0 \leq k \leq n)$ as

$$\hat{M}^n = \sup_{0\leq k\leq n} \hat{R}^n_k \text{ and } \hat{I}^n = \sup_{0\leq k\leq n} \hat{R}^n_k.$$

Let us calculate

$$\hat{M}^n = \sup_{0\le k\le n} \hat{R}^n_k = \sup_{0\le k\le n} R_n - R_{n-k} = R_n + \sup_{0\le k\le n} -R_{n-k} = R_n - \inf_{0\le k\le n} R_k = R_n - I_n,$$

and we have,

$$\hat{R}^n_n - \hat{I}^n = R_n - \inf_{0\le k\le n} \hat{R}^n_k = R_n - \inf_{0\le k\le n} \hat{R}^n_k = R_n - \inf_{0\le k\le n} R_n - R_{n-k} = - \inf_{0\le k\le n} -R_{n-k} = \sup_{0\le k\le n} R_{n-k} = M_n.$$

By replacing the random walk $R$ by the time reversed random walk, we have that $M_n$ has the same distribution as $R_n - I_n$. Finally, to prove that $K_n$ has the same distribution $H_n$, we use Proposition 5.4, the definition of $K_n$ and the same reasoning as for the first part of the proof. □

The next lemma is the last result that is needed to prove Theorem 5.6.

**Lemma 5.12.** *We have*

$$\frac{H_n}{R_n - I_n} \xrightarrow[n\to\infty]{(P)} \frac{2}{\sigma^2},$$

*where the notation* $\xrightarrow{(P)}$ *means the convergence in probability as defined in Definition 1.5.*

*Proof.* From the definitions, we have that

$$M_n = \sum_{T_k\le n} (R_{T_k} - R_{T_{k-1}}) = \sum_{k=1}^{K_n} (R_{T_k} - R_{T_{k-1}}).$$

Now we can apply the Lemma 5.8 and the law of large numbers (as $K_n \to \infty$) to get that

$$\frac{M_n}{K_n} \xrightarrow[n\to\infty]{a.s.} \mathrm{E}(R_{T_1}) = \frac{\sigma^2}{2}.$$

Finally, we complete the proof by using Proposition 5.11. We obtain that

$$\frac{R_n - I_n}{H_n} \xrightarrow[n\to\infty]{(P)} \mathrm{E}(R_{T_1}) = \frac{\sigma^2}{2}.$$

□

The final point to prove Theorem 5.6 is to use Equation (30) to obtain that, for $0 \le t_1 < \dots < t_m$, we have

$$\frac{1}{\sqrt{p}} \left(R_{[pt_1]} - I_{[pt_1]}, \dots, R_{[pt_m]} - I_{[pt_m]}\right) \xrightarrow[p\to\infty]{(d)} \sigma \left(B_{t_1} - \inf_{0\le s\le t_1} B_s, \dots, B_{t_m} - \inf_{0\le s\le t_m} B_s\right),$$

and with Lemma 5.12, we get

$$\frac{1}{\sqrt{p}} \left(H_{[pt_1]}, \dots, H_{[pt_m]}\right) \xrightarrow[p\to\infty]{(d)} \frac{2}{\sigma} \left(B_{t_1} - \inf_{0\le s\le t_1} B_s, \dots, B_{t_m} - \inf_{0\le s\le t_m} B_s\right).$$

We can conclude that the height process of a Galton-Watson random forest converges to a reflected Brownian motion.

# Conclusion

We can conclude the thesis by recapping the results that we have seen. In the first time, we have studied the standard Brownian motion. We have concluded that it is a continuous function in which the variations have a normal distribution. We have seen two different constructions of it. The Lévy's construction consists of building a function satisfying the properties of the Brownian motion on the dense set of dyadic numbers. The density of the set allowed us to conclude that we have build the Brownian motion. In the second construction, using the random walks, we have proved the important result of the Donsker's theorem which mentions that a rescaled random walk converges to the Brownian motion. Finally, the main property of the Brownian motion is the non-differentiability at any point, although it is continuous everywhere. In the second time, we have introduced the set of discrete trees which we have endowed with an ultra-metric distance. It allowed us the consider the set of discrete trees as a metric space. As a particular type of discrete trees, we have seen the Galton-Watson trees based on the Galton-Watson process. It consists of representing a population in generations where the number of offsprings is random and independent for every individual. We were interested in determining whether there exists a limit for such a tree if we condition it on the size. In order to talk about a limit for a sequence of random trees, we have introduced the notion of convergence in distribution. The Kesten's theorem showed us that the limit is the so-called Kesten's tree which is almost surely a tree with exactly one infinite spine. Finally, we have proved that the limit of the rescaled height process of an infinite random Galton-Watson forest is a reflected Brownian motion which is a function with random variations but it is always non-negative. For future works, we can study more in details the processes behind the continuous random trees, for example the properties of a reflected Brownian motion.

# References


[1] Lukacs, Eugene. *Probability and Mathematical Statistics*. Academic Pr., 1972.

[2] Stochastic Process. *Wikipedia*, Wikimedia Foundation, 24 June 2019, en.wikipedia.org/wiki/Stochastic_process.

[3] Panchenko, Dmitry. Multivariate Normal Distribution and Multivariate CLT. 2006, Massachusetts Institute of Technology, Massachusetts Institute of Technology.

[4] Fabrega Canudas, Josep. Convergence of Sequence of Random Variables. 2019, Universitat Politècnica de Catalunya, Universitat Politècnica De Catalunya.

[5] Multivariate Normal Distribution. *Wikipedia*, Wikimedia Foundation, 12 June 2019, en.wikipedia.org/wiki/Multivariate_normal_distribution.

[6] Brownian Motion. *Wikipedia*, Wikimedia Foundation, 24 June 2019, en.wikipedia.org/wiki/Brownian_motion.

[7] Mörters, Peter. Brownian Motion. 2011, University of Bath, University of Bath.

[8] Diez, Antoine. Levy's Construction of Brownian Motion. ENS Rennes, ENS Rennes.

[9] Sigman, Karl. Notes on Brownian Motion. IEOR 4700. 2006, Columbia University, Columbia University.

[10] Goldschmidt, Christina. An Introduction to Random Trees. CIMPA-DAAD Research School on Stochastic Processes and their Applications. CIMPA-DAAD Research School on Stochastic Processes and their Applications, 28 July 2015, Ulaanbaatar, National University of Mongolia.

[11] Fabrega Canudas, Josep. Population growth: Galton-Watson process. 2019, Universitat Politècnica de Catalunya, Universitat Politècnica De Catalunya.

[12] Romain Abraham, Jean-François Delmas. "An introduction to Galton-Watson trees and their local limits." 2015. hal-01164661

[13] Le Gall, Jean-François. Random Trees and Applications. *Probability Surveys*, The Institute of Mathematical Statistics and the Bernoulli Society, projecteuclid.org/euclid.ps/1132583290.

[14] Vincent, Granville. A Simple Introduction to Complex Stochastic Processes. *Data Science Central*, 27 Dec. 2017, www.datasciencecentral.com/profiles/blogs/a-simple-introduction-to-complex-stochastic-processes.