CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL
Unidad Zacatenco
Departamento de Matemáticas

Modelos discretos con ambiente aleatorio

Tesis que presenta

M. C. María Isabel Castro Martínez

para obtener el Grado de

Doctor en Ciencias

en la Especialidad de

Matemáticas

Director de Tesis:

Dr. Carlos G. Pacheco

Ciudad de México

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL
Zacatenco Unit
Department of Mathematics

Discrete Models with Random Environment

Thesis that presents

MSc. María Isabel Castro Martínez

to obtain the Degree of

Doctor of Philosophy

in the Specialty of

Mathematics

Doctoral Advisor:

Ph.D Carlos G. Pacheco

Mexico City

Discrete Models with Random Environment

Agradecimientos

Le agradezco al Dr. Carlos Pacheco por su paciencia y apoyo en todo momento.

Doy gracias a mi familia por darme su amor y comprensión.

A David y Renata por ser mi motivación.

Gracias también a mis amigos que siempre han estado dando ánimo.

Agradezco al Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV) y al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico proporcionado que me permitió realizar este trabajo.

Resumen

Estudiamos de manera general la caminata aleatoria de Sinai, es decir, teoremas clásicos, su comportamiento asintótico y resultados sobre como recuperar el ambiente dada una trayectoria de la caminata, resultado del cual propusimos una mejora. El análisis de la caminata de Sinai, nos dió la pauta a proponer un modelo parsimonioso para fijar el precio de un derivado financiero, incorporando la existencia de ambiente aleatorio. Tal construcción puede pensarse como una extensión del modelo Cox-Ross-Rubinstein (CRR). Mencionamos las dificultades de aplicar el modelo de Sinai al intentar usarlo en el procedimiento CRR. Además, presentamos algunas simulaciones y un experimento numérico. Mostramos algunos resultados sobre el comportamiento de el modelo propuesto y así como ocurre con la caminata de Sinai, proporcionamos un método para conectar los sitios más visitados de el modelo con los puntos mínimos de una cierta función del ambiente. Por último, a manera de conjetura propusimos como obtener información en el problema de recuperación de el ambiente para el caso de el nuevo modelo.

Summary

We study in a general way the Sinai's random walk, i.e., classical theorems, its asymptotic behaviour and results about recovering the environment given a trajectory of the walk, where we propose some refinements of an algorithm to recover the environment. The analysis about the Sinai's walk gave us the guideline to propose a parsimonious model for financial pricing that incorporates the existence of a random environment, and such construction can be thought as an extension of the Cox-Ross-Rubinstein (CRR) model. We mention the difficulty of applying Sinai's model if we try to use it for the CRR procedure. In addition, we present some simulations and a numerical experiment. We prove some results about the behaviour of the proposed model, and as it was done with the Sinai's walk, we also provide a method to connect the most visited sites of the model with the minimum points of a certain function of the environment. Finally, in a conjecture we provide how we can obtain information of the problem about recovering environment for the case of the new model.

Contents

Introduction

A random walk in a random environment (it is abbreviated by RWRE) is a stochastic process with double randomness, that is, the randomness induced by the environment and the randomness by the walk itself, which produces that its behaviour differs substantially from the classical random walk. This kind of models emerged from problems in mathematical physics, and eventually they were applied in problems of DNA replication. The models in random environment present many interesting phenomena. In general, we realise that each environment affects the position of the particle. Roughly speaking, we can think in Sinai's random walk as a walk where each time that the particle reaches a state x, it goes up a unit with probability α_x or goes down a unit with probability $1-\alpha_x$, when the environment (in this case, such probabilities α_x) is fixed.

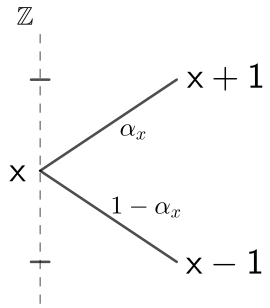

Figure 1: Dynamic of Sinai's walk.

Sometimes we have a trajectory of the walk, but we do not know the environment that implicitly produced such trajectory. Hence to obtain a procedure to recover the environment is a relevant task, which was studied in (Andreoletti, 2011) for the case of the Sinai walk. It is our motivation to study in a general way the Sinai walk, its asymptotic behaviour and Andreoletti's algorithm to recover the environment given a trajectory. Such survey gave us the guideline to propose other model in random

environment which is applied in Chapter 2 and studied with more detail in Chapter 3.

In the field of financial mathematics, when we study valuation of options we think that it is possible to suggest the existence of a random medium. For that we propose a model different of the Sinai walk, but that it saves its spirit. We are not claiming that behind the price of an underlying asset there is a fixed environment throughout the life of the asset. But it could happen that the environment behind the price of the asset is changing more slowly than the fluctuation of prices. In this thesis, it is not our intention to compare and measure which model is better, we only want to present our model as a first approximation maybe closer to the reality for modelling financial series and asset pricing.

In Chapter 1, we introduce the Sinai walk described in (Sinai, 1982) and summarise some important results. For instance, the Theorem 1.2.10 about recurrence we explain the proof with detail and clearly using the Chung-Fuchs theorem. We explain about the asymptotic behaviour in a more intuitive way and the relation between the most visited sites by the walk and the potential associated to the environment. Moreover, we study the problem of recovering the environment given that we know a trajectory of the Sinai walk in a finite time window. Mainly we review the first ideas of (Andreoletti, 2011) using only the local time. Also, we propose some refinements of the algorithm of Andreoletti to improve the results under certain situations. My contributions in this chapter are mainly two, on hand I give a proof clearly and complete of the recurrence theorem and on the other hand I propose some refinements of the algorithm to recover the environment.

In Chapter 2, we make a new proposal in the context of financial pricing that incorporates the existence of a random environment. Such construction can be though as an extension of the Cox-Ross-Rubinstein (CRR) model proposed in (Cox et al., 1979). For that, we consider the essence of the Sinai walk to present a new model in a random environment to apply it for option pricing; in fact, we mention the difficulty of applying exactly the Sinai model if we try to use it for the CRR procedure. For such reason, the proposed model is of different nature. Furthermore, we provide a way to connect the most visited sites of the model with the minimum points of a function of the environment, in a similar way as the connection between the Sinai walk and the potential. We present some simulations and a numerical experiment to bring a new perspective. In conclusion, my contribution in this chapter is not just to give a new model for financial series and pricing, but I consider that I am actually proposing a new paradigm that helps to understand financial markets.

Finally, in Chapter 3 we study some aspects of the model presented in Chapter

2. For instance, asymptotic behaviour, the relation between the environment and its effect in the trajectories of the walk. Also, we explore the possibility to obtain a procedure to recover information about the environment using only the local time, and then, after discarding the implementation of the algorithm of Andreoletti for such model, we propose a conjecture. In summary, my contributions here are three; I give results regarding the recurrence of the new model, I build the concept of "potential" for the new model to study favourite sites, and I set a conjecture for the problem of recovering the environment.

Throughout this work, we show figures made in the software Geogebra and Xfig. All simulations were programmed in the software R. We write the pseudo-codes at the end of each chapter.

Chapter 1

Random walk in a random environment

For this Chapter, we focus in the Sinai walk $(X_n)_{n\geqslant 1}$, which is a discrete model, but it is important to know that there is an extensive study about random walks in random environment for continuous time, we refer the reader to (Brox, 1986) for more details.

The first studies in the literature about random walks in random environment were presented by Temkin (1972) in the field of metallurgy (from here the reference to use the words *quenched* or *annealed* to call to certain probabilities as we will see later) and Chernov (1967) with a model of DNA replication. Afterwards, Solomon, Kesten, Golosov, Sinai, Kozlov and Spitzer were the pioneers in this area of probability.

In 1982, Sinai proved that under certain conditions, the model presented unusual properties. Sinai demonstrated how the order of X_n is about $(\log n)^2$. Observe that this order is much slower than $\sqrt{n}$, the usual order of the symmetric random walk. It turns out that the presence of the random environment forces to the RWRE to stay for long periods of time in certain zones, called *valleys*. Furthermore, he proved that the process $\frac{\sigma^2 X_n}{(\log n)^2}$ converges in law to a non-degenerated random variable. Later in 1986, independently (Kesten, 1986) and (Golosov, 1986) described such a law explicitly.

1.1 Sinai's random walk

After introducing the definition of a random walk in a random environment, we are going to present the *Sinai model*. The main goal in this chapter will be to prove the recurrence theorem (Solomon, 1975) with detail and clearly using the Chung-Fuchs theorem in its proof. Later, we are going to prove that Sinai's walk remains bounded uniformly. Then, we will present the localization theorem (Sinai, 1982). Finally, we

are going to explain an interesting procedure that provide us information about the environment for a given trajectory of Sinai's walk (Andreoletti, 2011). At the end of the chapter, the appendix A has some results that we need for the proofs. In the appendix B, we give the algorithm of the simulations presented along this chapter.

Let $\mathbb{Z}$ be the set of all integer numbers, i.e., $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. We consider $\alpha = \{\alpha_i : i \in \mathbb{Z}\}$ a sequence of independent and identically distributed random variables (i.i.d) taking values in $(0,1)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Each realization is called an **environment**. That is, for each $\omega \in \Omega$, $\alpha = \{\alpha_i(\omega)\}_{i\in\mathbb{Z}}$ is a bi-infinite string of numbers between 0 and 1.

For every fixed sequence α, we define the random walk in the environment α as the Markov chain $\{X_n : n \geqslant 0\}$ taking values in $\mathbb{Z}$ with associated probability space $(\Omega_1, \mathcal{F}_1, P^\alpha)$ where $X_0 = 0$, and the transition probabilities are the following:

$$P^\alpha(X_{n+1} = y \mid X_n = x) = \begin{cases} \alpha_x & \text{if } y = x+1, \\ 1-\alpha_x & \text{if } y = x-1, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1.1}$$

In other words, for each point x in the state space $\mathbb{Z}$, there is a transition-probability associated α_x. Then, the particle moves one unit to the right with probability α_x or it moves one unit to the left with probability $1-\alpha_x$ as we can see in the Figure 1.1. We can say that a random walk in a random environment (RWRE) evolves with double randomness: the environment we call it the *extrinsic* randomness, and the one coming from P^α we call it *intrinsic* randomness.

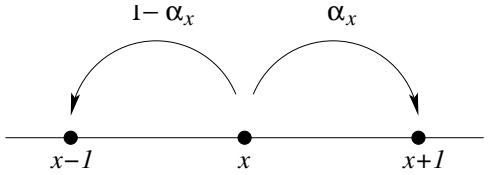

Figure 1.1: Transition probabilities in Sinai's walk.

Usually the probability measure P^α is referred as the *quenched law*. We use the notation $P_x^\alpha(\cdot)$ for the quenched probability such that $X_0 = x$, and E_x^α for the expectation associated to P_x^α. Also we denote by $\mathbb{E}$, the expectation associated to the measure $\mathbb{P}$. In addition, we are interested in the average of the quenched laws over all possible environments. We denote by $\mathbf{P}_x$ such law, and it is defined as

$$\mathbf{P}_x(A) := \mathbb{E}\, P_x^\alpha(A) = \int_\Omega P_x^{\alpha(\omega)}(A)\, \mathbb{P}(d\omega) \qquad \text{for all } A \in \mathcal{F}_1. \tag{1.1.2}$$

The measure $\mathbf{P}_x$ is called the *annealed law*. Observe that in general the RWRE is not a Markov chain under the annealed law, because the history of the RWRE modifies beforehand the law on the environment. In other words, the best way to approximate $P_x^\alpha(A)$ is by the law of large numbers, i.e., we run many repetitions of the experiment (that is, different environments) and the more often it reaches a state $j \in \mathbb{Z}$, the more information we have to approximate α_j depending the past of the walk in each experiment.

In fact, if the RWRE has a property A almost surely (a.s) with respect to (w.r.t) P_x^α (i.e., $P_x^\alpha(A) = 1$) and for almost all environments (i.e., $\omega \in \tilde{\Omega}$, where $\mathbb{P}(\tilde{\Omega}) = 1$), then the property A is true almost surely w.r.t $\mathbf{P}_x$.

As before, we consider $\alpha = \{\alpha_i : i \in \mathbb{Z}\}$ to be a sequence of i.i.d random variables. In addition, if the environment satisfies the following three conditions:

C.1 $\mathbb{E}\left(\log \dfrac{1-\alpha_0}{\alpha_0}\right) = 0,$

C.2 $0 < \sigma^2 = \mathbb{E}\left(\log \frac{1-\alpha_0}{\alpha_0}\right)^2 < \infty,$

C.3 There is a constant $0 < \beta < 1/2$ such that $\beta < \alpha_0 < 1-\beta$, $\mathbb{P}$-a.s,

then the process $(X_n)_{n\geqslant 0}$ is called **Sinai's walk**. The first condition is necessary to prove that the RWRE is a recurrent process. The second condition excludes the symmetric random walk, because when $\sigma^2 = 0$ and the condition **C.1** is true, then $\mathbb{P}(\alpha_0 = \frac{1}{2}) = 1$. Furthermore, the condition $\sigma^2 < \infty$ is necessary to apply the strong law of large numbers in some proofs. The condition **C.3** is called the *ellipticity condition* and it helps to control the fluctuations on the environment.

Note that the condition **C.3** implies that the random variable $\log \frac{1-\alpha_0}{\alpha_0}$ is bounded $\mathbb{P}$-a.s, because

$$\left|\log \frac{1-\alpha_0}{\alpha_0}\right| < \log \frac{1-\beta}{\beta} \quad \mathbb{P}\text{-a.s.} \tag{1.1.3}$$

1.2 About recurrence

First, we recall some basic definitions, see (Durrett, 1999) for instance. Let $\{Y_n\}_{n\geqslant 0}$ be a Markov chain on a state space S. Let $x, y \in S$, we assume that $Y_0 = x$. Denote by

$$T_y = \begin{cases} \inf\{n \geqslant 1 \mid Y_n = y,\} \\ +\infty, \ \textit{if such n does not exit,} \end{cases} \tag{1.2.1}$$

the first time that the Markov chain arrives to y given that it starts at the state x.

Definition 1.2.1. *Consider the following statements.*

a) A state x is said to communicate with the state y if

$$P_x(T_y < \infty) > 0.$$

b) We say two states x and y are intercommunicated if and only if x communicates with the state y, and y communicates with the state x.

c) If there is only one communication class, then the chain is called irreducible.

d) A state x is called transient if

$$P_x(Y_n = x \text{ for infinitely many } n) = 0.$$

e) A state x is called recurrent if

$$P_x(Y_n = x \text{ for infinitely many } n) = 1.$$

f) We say that an irreducible Markov chain is recurrent (transient) if all states of this chain are recurrent (transient).

Lemma 1.2.2. *(Durrett, 1999)[p. 12-20]*
Let $\{Y_0, Y_1, Y_2 \dots\}$ be a Markov chain with state space S. The following statements are equivalent.

i. The state x is recurrent.

ii. $P_x(T_x < \infty) = 1$,

iii. $P_x(Y_n = x \text{ for infinitely many } n) = 1$,

iv. $\sum_{n=1}^{\infty} P_x(Y_n = x) = \infty$.

Now, let us consider the simple symmetric random walk Y_n on $\mathbb{Z}$,

$$Y_n = Y_0 + \sum_{i=1}^{n} \xi_i,$$

where $Y_0 = x$, and $(\xi_i : i \geqslant 1)$ is a sequence of i.i.d random variables that taking values in $\{1, -1\}$ with $P(\xi_1 = 1) = 1/2 = P(\xi_1 = -1)$. It is well known that the symmetric random walk Y_n is a recurrent process. Then, it is natural to ask ourselves if the Sinai walk is also a recurrent process. Indeed, Solomon (1975) gave the necessary conditions to prove the result of recurrence.

First, we are going to present a required definition for the study of the Sinai walk.

Definition 1.2.3. *For every fixed α, we define the **potential** associated to the random environment α as*

$$V(x) = \begin{cases} \sum_{i=1}^{x} \log \frac{1-\alpha_i}{\alpha_i}, & \text{if } x = 1,2,\dots \\ 0, & \text{if } x = 0 \\ -\sum_{i=x+1}^{0} \log \frac{1-\alpha_i}{\alpha_i}, & \text{if } x = -1,-2,\dots \end{cases} \tag{1.2.2}$$

The potential $V(x)$ for a given environment α will be used to prove the recurrence. As well it is indispensable to analyse the asymptotic behaviour of the Sinai walk. This function was proposed and used in (Solomon, 1975).

Remark 1.2.4. *By hypothesis $\{\alpha_j\}_{j\in\mathbb{Z}}$ are i.i.d random variables and we know that the map $y \mapsto \log y$ is a measurable function, and $\log \frac{1-\alpha_o}{\alpha_o} \in L_1$ by (1.1.3). Then, the sequence $\{\log \frac{1-\alpha_j}{\alpha_j}\}_{j\in\mathbb{Z}}$ are i.i.d random variables with mean zero by the condition **C.1** and finite variance by the condition **C.2**. Therefore, $V(x)$ is really a two-sided random walk.*

It is important to note that the presence of the random environment forces the RWRE to stay for a long time in certain zones called *valleys*.

Remark 1.2.5. *Let us explore the idea behind a "valley". Observe that with the Definition 1.2.3 about the potential associated to an environment α, it is possible to write the values of α_x for any $x \in \mathbb{Z}$, explicitly as*

$$\alpha_x = \frac{e^{-(V(x)-V(x-1))}}{e^{-(V(x)-V(x-1))}+1}. \tag{1.2.3}$$

For $i > 0$, we can calculate

$$\begin{aligned} \frac{e^{-(V(i)-V(i-1))}}{e^{-(V(i)-V(i-1))}+1} &= \frac{e^{-\left(\sum_{x=1}^{i}\log\left(\frac{1-\alpha_x}{\alpha_x}\right)-\sum_{x=1}^{i-1}\log\left(\frac{1-\alpha_x}{\alpha_x}\right)\right)}}{e^{-\left(\sum_{x=1}^{i}\log\left(\frac{1-\alpha_x}{\alpha_x}\right)-\sum_{x=1}^{i-1}\log\left(\frac{1-\alpha_x}{\alpha_x}\right)\right)}+1} \\ &= \frac{e^{-\left(\log\left(\frac{1-\alpha_i}{\alpha_i}\right)\right)}}{e^{-\left(\log\left(\frac{1-\alpha_i}{\alpha_i}\right)\right)}+1} \\ &= \frac{\left(\frac{1-\alpha_i}{\alpha_i}\right)^{-1}}{\left(\frac{1-\alpha_i}{\alpha_i}\right)^{-1}+1} \\ &= \frac{\frac{\alpha_i}{1-\alpha_i}}{\frac{\alpha_i+1-\alpha_i}{1-\alpha_i}} \\ &= \alpha_i. \end{aligned}$$

However, the same expression can be used for the case $i \leqslant 0$, because

$$\begin{aligned}
\frac{e^{-(V(i)-V(i-1))}}{e^{-(V(i)-V(i-1))}+1} &= \frac{e^{V(i-1)-V(i)}}{e^{V(i-1)-V(i)}+1} \\
&= \frac{e^{-\sum_{x=i}^{0}\log\left(\frac{1-\alpha_x}{\alpha_x}\right)-\left(-\sum_{x=i+1}^{0}\log\left(\frac{1-\alpha_x}{\alpha_x}\right)\right)}}{e^{-\sum_{x=i}^{0}\log\left(\frac{1-\alpha_x}{\alpha_x}\right)-\left(-\sum_{x=i+1}^{0}\log\left(\frac{1-\alpha_x}{\alpha_x}\right)\right)}+1} \\
&= \frac{e^{-\log\frac{1-\alpha_i}{\alpha_i}}}{e^{-\log\frac{1-\alpha_i}{\alpha_i}}+1} \\
&= \alpha_i.
\end{aligned}$$

This is interesting because it tells us that

$$\begin{aligned}
\alpha_x &> \frac{1}{2} \Longleftrightarrow \\
2e^{-(V(x)-V(x-1))} &> e^{-(V(x)-V(x-1))}+1 \Longleftrightarrow \\
e^{-(V(x)-V(x-1))} &> 1 \Longleftrightarrow \\
V(x-1)-V(x) &> 0.
\end{aligned}$$

In other words, when the potential decreases is because the probabilities are greater than $\frac{1}{2}$, i.e., the random walk tends to go to the right. Similarly, we can obtain that $\alpha_x < \frac{1}{2} \Longleftrightarrow V(x-1) < V(x)$, that is, when the potential increases, the random walk tends to go to the left.

Due to Solomon, we have a complete criterion for describing when the RWRE is transitory or recurrent, for this section only analyse the recurrent case.

Theorem 1.2.6. *(Solomon, 1975)*

1. *If $\mathbb{E}(\log(\frac{1-\alpha_0}{\alpha_0})) < 0$ then X_n is transitory $(\boldsymbol{P}_0 - a.s.)$ and $\lim_{n\to\infty} X_n = +\infty$.*
2. *If $\mathbb{E}(\log(\frac{1-\alpha_0}{\alpha_0})) > 0$ then X_n is transitory $(\boldsymbol{P}_0 - a.s.)$ and $\lim_{n\to\infty} X_n = -\infty$.*
3. *If $\mathbb{E}(\log(\frac{1-\alpha_0}{\alpha_0})) = 0$ then X_n is recurrent $(\boldsymbol{P}_0 - a.s.)$ and $\limsup_{n\to\infty} X_n = +\infty$, $\liminf_{n\to\infty} X_n = -\infty$.*

Assume $a, x, b \in \mathbb{Z}$ such that $a < x < b$. Henceforth, we denote by

$$P_x^{\alpha}(T_a > T_b),$$

the quenched probability of paths that start at x and they arrive to b before reaching the state a. Respectively,

$$P_x^{\alpha}(T_a < T_b)$$

denotes the quenched probability of paths that start at x and they arrive to a before reaching the state b.

Lemma 1.2.7. *We can compute explicitly such probabilities:*

$$P_x^{\alpha}(T_a > T_b) = \frac{\sum_{i=a}^{x-1} \exp(V(i)-V(a))}{\sum_{i=a}^{b-1} \exp(V(i)-V(a))}, \tag{1.2.4}$$

$$P_x^{\alpha}(T_a < T_b) = \frac{\sum_{i=x+1}^{b} \exp(V(i)-V(b))}{\sum_{i=a+1}^{b} \exp(V(i)-V(b))} \tag{1.2.5}$$

Proof. Define $g(x) := P_x^{\alpha}(T_a > T_b)$. Indeed by the law of total probability, we are going to solve the following boundary value problem difference equation

$$g(x) = \alpha_x g(x+1) + (1-\alpha_x) g(x-1), \quad a < x < b \tag{1.2.6}$$

$$g(a) = 0 \tag{1.2.7}$$

$$g(b) = 1 \tag{1.2.8}$$

We multiply by $\alpha_x + (1-\alpha_x)$ on the left side of (1.2.6), and we obtain

$$\alpha_x g(x) + (1-\alpha_x) g(x) = \alpha_x g(x+1) + (1-\alpha_x) g(x-1).$$

That allows us to acquire the recurrence

$$g(x+1) - g(x) = \frac{1-\alpha_x}{\alpha_x} (g(x) - g(x-1)).$$

By taking $x = a+1$ and the initial condition of (1.2.7) in the above equation,

$$\begin{aligned} g(a+2) - g(a+1) &= \frac{1-\alpha_{a+1}}{\alpha_{a+1}} (g(a+1) - g(a)) \\ &= \frac{1-\alpha_{a+1}}{\alpha_{a+1}} g(a+1). \end{aligned}$$

Later, for $x = a+2$,

$$\begin{aligned} g(a+3) - g(a+2) &= \frac{1-\alpha_{a+2}}{\alpha_{a+2}} (g(a+2) - g(a+1)) \\ &= \frac{1-\alpha_{a+2}}{\alpha_{a+2}} \frac{1-\alpha_{a+1}}{\alpha_{a+1}} g(a+1). \end{aligned}$$

So on and so forth until $x = b-1$,

$$\begin{aligned} g(b) - g(b-1) &= \frac{1-\alpha_{b-1}}{\alpha_{b-1}} (g(b-1) - g(b-2)) \\ &= \frac{1-\alpha_{b-1}}{\alpha_{b-1}} \cdots \frac{1-\alpha_{a+1}}{\alpha_{a+1}} g(a+1). \end{aligned}$$

Thereupon using the final condition in (1.2.8), we have

$$1-g(b-1)=g(a+1)\prod_{j=a+1}^{b-1}\frac{1-\alpha_j}{\alpha_j}.$$

Adding in the previous calculations on both sides, the left side is a telescopic sum and therefore we obtain

$$1-g(a+1)=g(a+1)\sum_{i=a+1}^{b-1}\prod_{j=a+1}^{i}\frac{1-\alpha_j}{\alpha_j}, \tag{1.2.9}$$

or equivalently

$$1=g(a+1)\left(\sum_{i=a+1}^{b-1}\prod_{j=a+1}^{i}\frac{1-\alpha_j}{\alpha_j}+1\right).$$

That is[1]

$$g(a+1)=\left(\sum_{i=a}^{b-1}\prod_{j=a+1}^{i}\frac{1-\alpha_j}{\alpha_j}\right)^{-1}. \tag{1.2.10}$$

In general, when we add from $a+1$ to x, we write the recurrence as

$$g(x)-g(a+1)=g(a+1)\sum_{i=a+1}^{x-1}\prod_{j=a+1}^{i}\frac{1-\alpha_j}{\alpha_j}.$$

Therefore,

$$\begin{aligned}
g(x) &= g(a+1)\left(\sum_{i=a+1}^{x-1}\prod_{j=a+1}^{i}\frac{1-\alpha_j}{\alpha_j}+1\right)\\
&= \frac{\sum_{i=a}^{x-1}\prod_{j=a+1}^{i}\frac{1-\alpha_j}{\alpha_j}}{\sum_{i=a}^{b-1}\prod_{j=a+1}^{i}\frac{1-\alpha_j}{\alpha_j}} && \text{by (1.2.10),}\\
&= \frac{\sum_{i=a}^{x-1}\exp(\sum_{j=a+1}^{i}\log\frac{1-\alpha_j}{\alpha_j})}{\sum_{i=a}^{b-1}\exp(\sum_{j=a+1}^{i}\log\frac{1-\alpha_j}{\alpha_j})} && \text{using exponential and logarithmic functions,}\\
&= \frac{\sum_{i=a}^{x-1}\exp(V(i)-V(a))}{\sum_{i=a}^{b-1}\exp(V(i)-V(a))} && \text{by definition of the potential } V(x).
\end{aligned}$$

Analogously we can deduce a formula for $P_x^{\alpha}(T_a<T_b)$. □

For each fixed α, at first sight the Sinai walk is an irreducible Markov chain. In accordance with the Definition 1.2.1, it is enough to prove that the state $x=0$ is recurrent to prove that the Sinai walk is a recurrent process.

[1] For any sequence of real numbers $(y_j)_j$, it is defined the empty product $\prod_{j=a+1}^{a} y_j := 1$ and the empty sum $\sum_{j=a+1}^{a} y_j := 0$.

Definition 1.2.8. *We have the following statements.*

*a. A **right** (left) **excursion** from a site x is defined as a random walk in which the first step is to the right (left).*

*b. We say that a process is **right** (left) **recurrent** if for any excursion to the right (left) of x, the particle returns to x with probability 1.*

Now we present the recurrent theorem, and we provide a proof based on the one in (Hughes, 1996) where we have incorporated the so-called Chung-Fuchs theorem.

Theorem 1.2.9. *Chung-Fuchs (Durrett, 2019)[p. 252-253]*
Consider S_n a random walk in dimension 1, $S_n = X_1 + X_2 + \cdots + X_n$, where $X_1, X_2, \ldots$ are i.i.d random variables. If $S_n/n \to 0$ in probability, then S_n is recurrent.

Theorem 1.2.10. *Recurrence.*
If $\mathbb{E}\left(\log \frac{1-\alpha_0}{\alpha_0}\right) = 0$, then the Sinai walk X_n is recurrent $\boldsymbol{P}_0 - a.s.$

Proof. Let α be a fixed environment. Then, by the law of total probability

$$P_0^{\alpha}(T_0 < \infty) = \alpha_0 P_1^{\alpha}(T_0 < \infty) + (1-\alpha_0) P_{-1}^{\alpha}(T_0 < \infty). \tag{1.2.11}$$

Consequently

$$P_0^{\alpha}(T_0 < \infty) = 1 \iff P_1^{\alpha}(T_0 < \infty) = 1 \text{ and } P_{-1}^{\alpha}(T_0 < \infty) = 1.$$

In other words by Lemma 1.2.2, X_n is quenched-recurrent if and only if X_n is right recurrent and left recurrent with the quenched law. Hence, if the recurrence is a quenched property for almost environment α, then it is annealed-recurrent a.s.

In what follows, we focus to prove that

$$\mathbb{E}\left(\log \frac{1-\alpha_0}{\alpha_0}\right) \geqslant 0 \iff X_n \text{ is right recurrent with the quenched law.} \tag{1.2.12}$$

Consider $a = 0$, $x = 1$ and $b > 1$ in the equation (1.2.4), i.e.,

$$\begin{aligned} 1 - P_1^{\alpha}(T_0 < T_b) &= P_1^{\alpha}(T_0 > T_b) \\ &= \frac{1}{\sum_{i=0}^{b-1} \exp(V(i))}. \end{aligned}$$

This implies that X_n is right recurrent with the law quenched if and only if

$$P_1^{\alpha}(T_0 < T_b) \to 1 \text{ as } b \to \infty \iff \sum_{i=0}^{b-1} \exp(V(i)) \to \infty \text{ as } b \to \infty. \tag{1.2.13}$$

Notice that

$$\sum_{i=0}^{b-1} \exp(V(i)) = 1 + \sum_{i=1}^{b-1} e^{V(i)}. \tag{1.2.14}$$

As we have said in the Remark 1.2.4, the potential $V(x)$ is a two-sided random walk. In particular, $V(x)$ for $x \geqslant 1$ is a random walk. Then, by the strong law of large numbers we know that

$$\frac{1}{i}V(i) = \frac{1}{i}\sum_{j=1}^{i} \log\frac{1-\alpha_j}{\alpha_j} \to \mathbb{E}\left(\log\frac{1-\alpha_0}{\alpha_0}\right) \quad \text{as } i \to \infty, \quad \mathbb{P}-\text{a.s.}$$

Indeed, if we first assume

$$\mathbb{E}\left(\log\frac{1-\alpha_0}{\alpha_0}\right) > 0, \tag{1.2.15}$$

then we consider the $i-$th term of the sum given in (1.2.14), $a_i = e^{V(i)} = e^{(\frac{1}{i}V(i))i}$, and we use the root test for convergence of series[2]. We calculate

$$L = \lim_{i\to\infty} e^{\frac{V(i)}{i}} = e^{\mathbb{E}(\log\frac{1-\alpha_0}{\alpha})} > 1;$$

this enables us conclude that

$$\sum_{i=1}^{b-1} e^{V(i)} \to \infty \quad \text{as } b \to \infty.$$

However, if we now assume

$$\mathbb{E}\left(\log\frac{1-\alpha_0}{\alpha_0}\right) = 0.$$

By the strong law of large numbers, we have that

$$\frac{1}{i}V(i) \to 0 \quad \text{as } i \to \infty \quad \mathbb{P}-\text{a.s.},$$

and therefore such convergence is true also in probability. Then, by the Theorem 1.2.9, $V(x)$ is recurrent, and as a consequence $V(x)$ visits infinitely many times both sides of the starting site with probability 1. For that reason $V(x) \geqslant 0$ for infinitely many x, and consequently

$$\sum_{i=1}^{b-1} e^{V(i)} \to \infty \quad \text{as } b \to \infty.$$

[2]Theorem. Let $\sum a_n$ be a series with positive terms, and let $L = \lim_{n\to\infty} \sqrt[n]{a_n}$.

i. If $L < 1$, then the series converges.

ii. If $L > 1$, then the series diverges.

iii. If $L = 1$, then the root test is inconclusive.

Thus by (1.2.13), we deduce that $P_1^{\alpha}(T_0 < T_b) \to 1$ as $b \to \infty$. In this way, we have shown that $\mathbb{E}\left(\log \frac{1-\alpha_0}{\alpha_0}\right) \geqslant 0$ implies that X_n is right recurrent with the quenched law for almost all environments.

Now, we must to prove that if X_n is right recurrent with the quenched law then $\mathbb{E}\left(\log \frac{1-\alpha_0}{\alpha_0}\right) \geqslant 0$, we use a proof by contraposition. We assume that $\mathbb{E}\left(\log \frac{1-\alpha_0}{\alpha_0}\right) < 0$, then by the strong law of large numbers and the root test for convergence of series, we know that

$$\mathbb{E}\left(\log \frac{1-\alpha_0}{\alpha_0}\right) < 0 \Rightarrow \sum_{i=1}^{b-1} e^{V(i)} < \infty \text{ as } b \to \infty,$$

so by (1.2.13), $P_1^{\alpha}(T_0 < T_b) \not\to 1$, in other words X_n is not right recurrent, which concludes the statement (1.2.12).

On the other hand, if the first step is to the left, we can write the proof analogously and deduce that $\mathbb{E}\left(\log \frac{1-\alpha_0}{\alpha 0}\right) \leqslant 0$ if and only if X_n is left recurrent with the quenched law for almost all environments.

Finally, we have proved that $\mathbb{E}\left(\log \frac{\alpha_0}{1-\alpha_0}\right) = 0$ if and only if X_n is right recurrent and left recurrent with the quenched law for almost all environments, thus X_n is recurrent with the quenched law for almost all environments. Consequently, X_n is recurrent $\mathbf{P}_0 -$ a.s. □

One can see in the Figure 1.2 a trajectory of Sinai's walk after n-steps time, with $n = 500,000$. In the Appendix B, we will present the algorithm to run the simulation, which has been programmed in the software R Studio (Core et al., 2013). For such simulation, we consider the quenched case with the environment α_x such that

$$\alpha_x = \begin{cases} 1/4 & \text{with probability } 1/2, \\ 3/4 & \text{with probability } 1/2. \end{cases} \tag{1.2.16}$$

1.3 Asymptotic behaviour

Let us to see the heuristic argument given in (Révész, 2005) about the order of Sinai's walk. Let α be a fixed environment. In accordance with the Lemma 1.2.7, we calculate for $b > 1$

$$P_1^{\alpha}(T_0 > T_b) = \left(\sum_{i=0}^{b-1} \exp(V(i))\right)^{-1}.$$

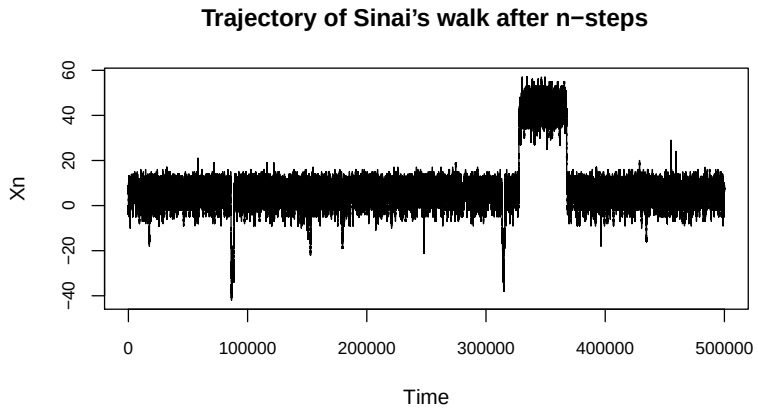

Figure 1.2: Simulation of a trajectory of Sinai's walk.

Aforementioned in the Remark 1.2.4, $V(x)$ is a two-sided random walk, where the typical size[3] of $\mid V(x) \mid$ is $O(\sqrt{x})$. If we define $\upsilon := P_1^{\alpha}(T_0 > T_b)$, then for b large enough

$$\upsilon \approx \exp(-\sqrt{b}). \tag{1.3.1}$$

Let N be the random variable which counts the number of visits of Sinai's walk to the state 0 before visiting the state b given that it starts at 1. Then

$$\begin{aligned} E_1^{\alpha}(N) &= \sum_{j=1}^{\infty} j \cdot P_1^{\alpha}(N=j) \\ &= \sum_{j=1}^{\infty} j\upsilon(1-\upsilon)^j \\ &= \upsilon(1-\upsilon)\sum_{j=1}^{\infty} j(1-\upsilon)^{j-1} \\ &= \upsilon(1-\upsilon)\frac{1}{(1-\upsilon)^2} \\ &= \frac{1-\upsilon}{\upsilon}. \end{aligned}$$

In addition, by using (1.3.1) we deduce that[4]

$$\frac{1-\upsilon}{\upsilon} \sim \exp(\sqrt{b}).$$

Hence

$$E_1^{\alpha}(N) \sim \exp(\sqrt{b}).$$

[3]We say that $x(n) = O(y(n))$ if there exist $C > 0$ and n_0 such that for all $n \geqslant n_0$, $\mid x(n) \mid \leqslant Cy(n)$.

[4]Let us recall that $x(n)$ and $y(n)$ have the same order, denote by $x(n) \sim y(n)$, if $\lim_{n\to\infty} \frac{x(n)}{y(n)} = 1$.

In other words, for n large the Sinai walk takes at least $e^{\sqrt{n}}$-steps reaches the state n by first time. That allows us to say that the displacement after n-steps does not exceed $(\log n)^2$.

The asymptotic behaviour of this random walk was described in (Sinai, 1982), where he demonstrated how $X_n \approx (\log n)^2$. Indeed that order is slower than $\sqrt{n}$, the usual order of the symmetric random walk. The idea of Sinai to analyse the asymptotic behaviour of X_n was to approach the random walk $V(x)$ with the Brownian motion in order to study typical behaviour. Let us now present some essential ideas of this procedure.

For a given environment α and $n \geqslant 1$, one defines the **scaled potential** as follows

$$V^n(t) = \begin{cases} \frac{1}{\log n}\sum_{i=0}^{[(\log n)^2 t]} \log\frac{1-\alpha_i}{\alpha_i} & \text{if } t = \frac{1}{\log^2 n}, \frac{2}{\log^2 n}, \dots \\ 0 & \text{if } t = 0 \\ -\frac{1}{\log n}\sum_{i=[(\log n)^2 t+1]}^{0} \log\frac{1-\alpha_i}{\alpha_i} & \text{if } t = \frac{-1}{\log^2 n}, \frac{-2}{\log^2 n}, \dots \end{cases} \tag{1.3.2}$$

and by linear interpolation we extend $V^n(t)$ for all $t \in \mathbb{R}$. Due to the time-space scaling, the Donsker invariance principle tell us that the process $(V^n(t) : t \in \mathbb{R})$ converges weakly to a two-sided Brownian motion.

For each $n \geqslant 1$ and $a, x, b \in \mathbb{Z}$ such that $a < x < b$, we can rewrite the Lemma 1.2.7 to obtain the equations in terms of $V^n(t)$ as

$$P_x^\alpha(T_a > T_b) = \frac{\sum_{i=a}^{x-1} \exp(\log n[V^n(i\cdot\log^{-2} n) - V^n(a\cdot\log^{-2} n)])}{\sum_{i=a}^{b-1} \exp(\log n[V^n(i\cdot\log^{-2} n) - V^n(a\cdot\log^{-2} n)])}, \tag{1.3.3}$$

$$P_x^\alpha(T_a < T_b) = \frac{\sum_{i=x+1}^{b} \exp(\log n[V^n(i\cdot\log^{-2} n) - V^n(b\cdot\log^{-2} n)])}{\sum_{i=a+1}^{b} \exp(\log n[V^n(i\cdot\log^{-2} n) - V^n(b\cdot\log^{-2} n)])}. \tag{1.3.4}$$

Definition 1.3.1. *For each $n \geqslant 1$ and given a fixed environment α. We call to the triplet $\{A^n, b^n, C^n\}$ a valley if*

$$\begin{aligned} V^n(A^n) &= \max_{A^n \leqslant t \leqslant b^n} V^n(t), \\ V^n(C^n) &= \max_{b^n \leqslant t \leqslant C^n} V^n(t), \\ V^n(b^n) &= \min_{A^n \leqslant t \leqslant C^n} V^n(t). \end{aligned}$$

The depth of the valley $\{A^n, b^n, C^n\}$ is defined as

$$d([A^n, C^n]) := \min\{V^n(A^n) - V^n(b^n), V^n(C^n) - V^n(b^n)\}.$$

For n large enough, one may consider

- $C_0^n := \min\{t = \frac{k}{(\log n)^2},\ k \in \mathbb{Z}^+ :\ V^n(t) \geqslant 1\}$,
- $A_0^n := \max\{t = \frac{k}{(\log n)^2},\ k \in \mathbb{Z}^- :\ V^n(t) \geqslant 1\}$,
- b_0^n such that $V^n(b_0^n) = \min_{A_0^n \leqslant t \leqslant C_0^n} V^n(t)$.

The triplet

$$\{A_0^n, b_0^n, C_0^n\} \tag{1.3.5}$$

is a valley with depth greater than or equal to 1. We also define the *right refinement* operation as follows. First, we consider the segment $[b_0^n, C_0^n]$ and seek two times C_1^n and b_1^n such that

$$V^n(C_1^n) - V^n(b_1^n) = \max_{b_0^n \leqslant t_1 \leqslant t_2 \leqslant C_0^n} \{V^n(t_1) - V^n(t_2)\}.$$

Now, it is possible to obtain a second right refinement seeking two times C_2^n, b_2^n such that

$$V^n(C_2^n) - V^n(b_2^n) = \max_{b_0^n \leqslant t_1 \leqslant t_2 \leqslant C_1^n} \{V^n(t_1) - V^n(t_2)\}.$$

Successively, we can construct by r-right refinement operations the set of maxima and minima points

$$\mathcal{C}_r^+ := \{b_0^n,\ C_r^n,\ b_r^n,\ \ldots, C_2^n,\ b_2^n,\ C_1^n,\ b_1^n,\ C_0^n\}.$$

Observe the Figure 1.3 to see an example of the right refinement operation. If we denote by $\delta_0 := V^n(C_0^n) - V^n(b_0^n)$, $\delta_1 := V^n(C_1^n) - V^n(b_1^n), \ldots, \delta_r := V^n(C_r^n) - V^n(b_r^n)$, it is clear the relation

$$\delta_0 > \delta_1 > \cdots > \delta_r \geqslant 0.$$

Analogously we define the *left refinement* operation. Given the segment $[A_0^n, b_0^n]$ we seek two times $\tilde{b}_1^n$ and A_1^n such that

$$V^n(A_1^n) - V^n(\tilde{b}_1^n) = \max_{A_0^n \leqslant t_1 \leqslant t_2 \leqslant b_0^n} \{V^n(t_2) - V^n(t_1)\}.$$

In a similar way, we construct by $\tilde{r}$-left refinement operations the set of maxima and minima points

$$\mathcal{C}_{\tilde{r}}^- := \{A_0^n,\ \tilde{b}_1^n,\ A_1^n, \ldots, \tilde{b}_{\tilde{r}}^n,\ A_{\tilde{r}}^n, b_0^n\}.$$

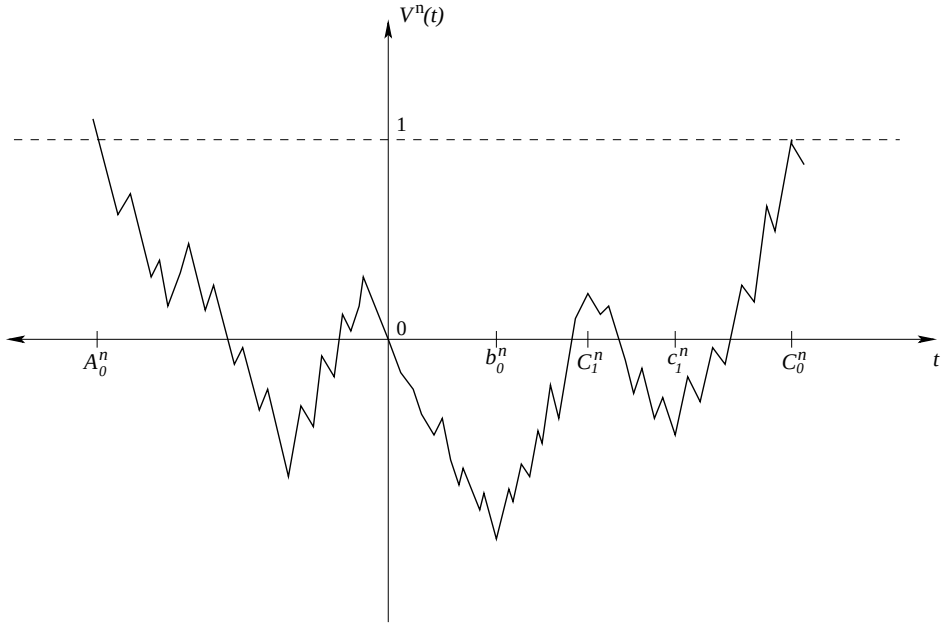

Figure 1.3: Example of a right refinement.

In what follows, we present the main ideas to prove how Sinai's walk, on a time window, stays trapped in a certain range described. The proof of the next lemma is based on a strong invariance principle for sum of independent random variables (Einmahl, 2009) and certain known results about the Brownian motion. The proof will be explain in Appendix A1.

Lemma 1.3.2. *Given $\eta > 0$, there exists n_0 large enough such that for all $n > n_0$, there exists a set $C_n \subset \Omega$ such that $\mathbb{P}(C_n) \geqslant 1-\eta$ and the following happens. For any choice $\omega \in C_n$, the corresponding valley $\{A_0^n, b_0^n, C_0^n\}$ described in (1.3.5) admits a finite number of r-right and r′-left refinements such that it exits the so-called* ***basic valley*** *$\{\tilde{M}_0, m_0, M_0\}$ with the following properties:*

i) $0 \in [\tilde{M}_0, M_0]$,

ii) $d([\tilde{M}_0, M_0]) \geqslant 1+\delta$, *for some $\delta > 0$ small enough.*

iii) For any other refinement $\{\tilde{M}_0', M_0'\}$ of the valley $\{\tilde{M}_0, M_0\}$, then $d([\tilde{M}_0', M_0']) < 1-\delta$.

iv) $\min_{\substack{t_1,t_2 \in [\tilde{M}_0, M_0] \\ t_1 \neq t_2}} | V^n(t_1) - V^n(t_2) | > \delta$.

v) $| \tilde{M}_0 | + | M_0 | \leqslant K$, *for some $K > 0$.*

Now, let us explain know the process remains inside a certain range.

Theorem 1.3.3. *Given $\eta > 0$. Under the assumptions of the Lemma 1.3.2, let $\alpha = \alpha(\omega)$ be a realization of the environment with $\omega \in C_n$ and we consider $\{\tilde{M}_0, m_0, M_0\}$ the basic valley. Then,*

$$P_0^{\alpha}(X_k \in [\tilde{M}_0 \cdot (\log n)^2,\ M_0 \cdot (\log n)^2] \ \text{ for all } \ 0 \leqslant k \leqslant n) \to 1, \ \text{ as } n \to \infty.$$

Proof. Denote by T_y^x the first time that the walk arrives to the state y given that it starts at x, i.e.,

$$T_y^x := \begin{cases} \inf\{k \geqslant 1 : \ X_k = y, \ \ \text{given } \ X_0 = x\}, \\ \\ +\infty, \ \text{such } k \text{ does not exist.} \end{cases}$$

Hereafter, for brevity we write $t' = t \cdot (\log n)^2$ for all $t = k/(\log n)^2$, $k \in \mathbb{Z}$. Then,

$$\begin{aligned} & P_0^{\alpha}(X_k \in [\tilde{M}_0 \cdot (\log n)^2, M_0 \cdot (\log n)^2] \ \text{ for all } \ 0 \leqslant k \leqslant n) \\ = \ & P_0^{\alpha}\left(\bigcap_{k=0}^{n} \{X_k \in [\tilde{M}_0 \cdot (\log n)^2, M_0 \cdot (\log n)^2]\}\right) \\ = \ & P_0^{\alpha}\left(\bigcap_{k=0}^{n} \{X_k \in [\tilde{M}_0', M_0']\}\right). \end{aligned}$$

Observe that

$$\begin{aligned} \bigcap_{k=0}^{n} \{X_k \in [\tilde{M}_0', M_0']\} &= \left\{\bigcap_{k=0}^{n} \{X_k \in [\tilde{M}_0', M_0']\}, T_{m_0'}^0 \leqslant n\right\} \bigcup \left\{\bigcap_{k=0}^{n} \{X_k \in [\tilde{M}_0', M_0']\}, T_{m_0'}^0 > n\right\} \\ &\supseteq \left\{\bigcap_{k=0}^{n} \{X_k \in [\tilde{M}_0', M_0']\}, T_{m_0'}^0 \leqslant n\right\} \\ &\supseteq \{T_{\tilde{M}_0'-1}^{m_0'} \wedge T_{M_0'+1}^{m_0'} > n\} \setminus \{\{T_{\tilde{M}_0'-1}^0 < T_{m_0'}^0\} \cup \{T_{M_0'+1}^0 < T_{m_0'}^0\}\}. \end{aligned}$$

Thereby

$$P_0^{\alpha}\left(\bigcap_{k=0}^{n} \{X_k \in [\tilde{M}_0', M_0']\}\right) \geqslant P_{m_0'}^{\alpha}(T_{\tilde{M}_0'-1}^{m_0'} \wedge T_{M_0'+1}^{m_0'} > n) - P_0^{\alpha}(T_{\tilde{M}_0'-1}^0 < T_{m_0'}^0) - P_0^{\alpha}(T_{M_0'+1}^0 < T_{m_0'}^0). \tag{1.3.6}$$

Here, without loss of generality one can assume $m_0 < 0$ and consequently

$$P_0^{\alpha}(T_{\tilde{M}_0'-1}^0 < T_{m_0'}^0) = 0. \tag{1.3.7}$$

Now, by (1.3.3) we obtain

$$P_0^{\alpha}(T_{M_0'+1}^0 < T_{m_0'}^0) = \frac{\displaystyle\sum_{i=m_0'}^{-1} \exp(\log n[V^n(i \cdot \log^{-2} n) - V^n(m_0' \cdot \log^{-2} n)])}{\displaystyle\sum_{i=m_0'}^{M_0'} \exp(\log n[V^n(i \cdot \log^{-2} n) - V^n(m_0' \cdot \log^{-2} n)])}.$$

Notice that all the terms in the denominator are non-negative and therefore

$$\sum_{i=m_0'}^{M_0'} \exp(\log n[V^n(i\cdot\log^{-2}n)-V^n(m_0'\cdot\log^{-2}n)])$$
$$\geqslant \exp(\log n[V^n(M_0'\cdot\log^{-2}n)-V^n(m_0'\cdot\log^{-2}n)]),$$

i.e.,

$$\frac{1}{\sum_{i=m_0'}^{M_0'}\exp(\log n[V^n(i\cdot\log^{-2}n)-V^n(m_0'\cdot\log^{-2}n)])}$$
$$\leqslant \frac{1}{\exp(\log n[V^n(M_0'\cdot\log^{-2}n)-V^n(m_0'\cdot\log^{-2}n)])}.$$

Thus,

$$\begin{aligned}
P_0^\alpha(T^0_{M_0'+1}<T^0_{m_0'}) &\leqslant \exp(-\log n[V^n(M_0'\cdot\log^{-2}n)-V^n(m_0'\cdot\log^{-2}n)]) && (1.3.8)\\
&\quad\cdot\sum_{i=m_0'}^{-1}\exp(\log n[V^n(i\cdot\log^{-2}n)-V^n(m_0'\cdot\log^{-2}n)])\\
&= \sum_{i=m_0'}^{-1}\exp(-\log n[V^n(M_0'\cdot\log^{-2}n)-V^n(m_0'\cdot\log^{-2}n)] && (1.3.9)\\
&\quad+\log n[V^n(i\cdot\log^{-2}n)-V^n(m_0'\cdot\log^{-2}n)])\\
&= \sum_{i=m_0'}^{-1}\exp(-\log n[V^n(M_0'\cdot\log^{-2}n)-V^n(i\cdot\log^{-2}n)])\\
&\leqslant |\,m_0'\,|\max_{m_0'\leqslant i\leqslant -1}\{\exp(-\log n[V^n(M_0'\cdot\log^{-2}n)-V^n(i\cdot\log^{-2}n)])\}\\
&\leqslant |\,m_0'\,|\exp(-\delta\log n) && (1.3.10)\\
&\leqslant K(\log n)^2\frac{1}{n^\delta}. && (1.3.11)
\end{aligned}$$

For the inequality (1.3.10) we use that according to the Lemma 1.3.2, the smallest difference is larger than δ. Also by the part *v*) of the same Lemma, it follows (1.3.11). Thus

$$P_0^\alpha(T^0_{M_0'+1}<T^0_{m_0'})\to 0,\ \text{as } n\to\infty.$$

Moreover by (1.3.6) and (1.3.7), we now only need to prove that

$$P^\alpha_{m_0'}(T^{m_0'}_{M_0'-1}\wedge T^{m_0'}_{M_0'+1}>n)\to 1,\ \text{as } n\to\infty. \tag{1.3.12}$$

Denote by $T_1^{x\to x}$ the first return time to x, $T_1^{x\to x}:=T_x^x$ and for $j\geqslant 2$, $T_j^{x\to x}$ the j-th return time i.e.,

$$T_j^{x\to x}:=\begin{cases}\inf\{k>T_{j-1}^{x\to x}:\ X_k=x\},\\ \\ +\infty,\ \text{such } k \text{ does not exist.}\end{cases}$$

Define $\tau_1^x := T_1^{x\to x}$ and for $j \geqslant 2$, $\tau_j^x := T_j^{x\to x} - T_{j-1}^{x\to x}$, in this way, $\sum_{j=1}^n \tau_j^x$ is the time that the walk takes to return n-times to x. Hence

$$\begin{aligned} P^\alpha_{m_0'}(T^{m_0'}_{\tilde{M}_0'-1} \wedge T^{m_0'}_{M_0'+1} > n) &= P^\alpha_{m_0'}(T^{m_0'}_{\tilde{M}_0'-1} \wedge T^{m_0'}_{M_0'+1} > n, \sum_{j=1}^n \tau_j^{m_0'} > n) \quad (1.3.13)\\ &\geqslant P^\alpha_{m_0'}(T^{m_0'}_{\tilde{M}_0'-1} \wedge T^{m_0'}_{M_0'+1} > \sum_{j=1}^n \tau_j^{m_0'}) \quad (1.3.14)\\ &= \left(P^\alpha_{m_0'}\left(T^{m_0'}_{\tilde{M}_0'-1} \wedge T^{m_0'}_{M_0'+1} \geqslant T^{m_0'}_{m_0'}\right)\right)^n \quad (1.3.15)\\ &= \left(1 - P^\alpha_{m_0'}\left(T^{m_0'}_{\tilde{M}_0'-1} \wedge T^{m_0'}_{M_0'+1} < T^{m_0'}_{m_0'}\right)\right)^n, \end{aligned}$$

where the equality (1.3.13) is true because the event $\{\sum_{j=1}^n \tau_i^{m_0'} > n\}$ has probability 1, and we have (1.3.15) by the strong Markov property.

In addition, by applying the law of total probability we can obtain that

$$\begin{aligned} P^\alpha_{m_0'}(T^{m_0'}_{\tilde{M}_0'-1} \wedge T^{m_0'}_{M_0'+1} < T^{m_0'}_{m_0'}) &= \alpha_{m_0'} P^\alpha_{m_0'+1}(T^{m_0'+1}_{M_0'+1} < T^{m_0'+1}_{m_0'})\\ &+ (1-\alpha_{m_0'}) P^\alpha_{m_0'-1}(T^{m_0'-1}_{\tilde{M}_0'-1} < T^{m_0'-1}_{m_0'}). \end{aligned}$$

Furthermore by (1.3.3), we compute

$$\begin{aligned} P^\alpha_{m_0'+1}(T^{m_0'+1}_{M_0'+1} < T^{m_0'+1}_{m_0'}) &= \frac{1}{\sum_{i=m_0'}^{M_0'} \exp(\log n[V^n(i\cdot\log^{-2} n) - V^n(m_0'\cdot\log^{-2} n)]}\\ &\leqslant \frac{1}{e^{\log n[V^n(M_0'\cdot\log^{-2} n) - V^n(m_0'\cdot\log^{-2} n)]}}\\ &\leqslant \frac{1}{n^{1+\delta}}. \quad (1.3.16) \end{aligned}$$

Let us explain (1.3.16). Using Lemma 1.3.2 part *ii*), we can bound the depth of the valley,

$$\begin{aligned} V^n(M_0'\cdot\log^{-2} n) - V^n(m_0'\cdot\log^{-2} n) &\geqslant d([\bar{M}_0, M_0])\\ &\geqslant 1+\delta, \end{aligned}$$

whereby

$$\begin{aligned} \exp(\log n[V^n(M_0'\cdot\log^{-2} n) - V^n(m_0'\cdot\log^{-2} n)]) &\geqslant \exp((1+\delta)\log n)\\ &= \frac{1}{n^{1+\delta}}. \end{aligned}$$

On the other hand, in a similar way but using now (1.3.4) we obtain the bound

$$\begin{aligned} P^{\alpha}_{m'_0-1}(T^{m'_0-1}_{\tilde{M}'_0-1} < T^{m'_0-1}_{m'_0}) &= \frac{1}{\sum_{i=\tilde{M}'_0}^{m'_0} \exp(\log n[V^n(i\cdot\log^{-2}n) - V^n(m'_0\cdot\log^{-2}n)]} \\ &\leqslant \frac{1}{e^{\log n[V^n(\tilde{M}'_0\cdot\log^{-2}n)-V^n(m'_0\cdot\log^{-2}n)]}} \\ &\leqslant \frac{1}{n^{1+\delta}}. \end{aligned}$$

This yields

$$\begin{aligned} P^{\alpha}_{m'_0}(T^{m'_0}_{\tilde{M}'_0-1} \wedge T^{m'_0}_{M'_0+1} \geqslant T^{m'_0}_{m'_0}) &\leqslant \alpha_{m'_0}\left(\frac{1}{n^{1+\delta}}\right) + (1-\alpha_{m'_0})\left(\frac{1}{n^{1+\delta}}\right) \\ &= \frac{1}{n^{1+\delta}}. \end{aligned}$$

Therefore

$$\left(1 - P^{\alpha}_{m'_0}(T_{\tilde{M}'_0-1} \wedge T_{M'_0+1} \geqslant T^{m'_0}_{m'_0})\right)^n \geqslant \left(1 - \frac{1}{n^{1+\delta}}\right)^n.$$

By using the inequality

$$(1-x)^n \geqslant 1 - nx, \text{ for all } 0 \leqslant x \leqslant 1 \text{ and } n \geqslant 1,$$

we obtain that

$$\begin{aligned} \left(1 - \frac{1}{n^{1+\delta}}\right)^n &\geqslant 1 - n\frac{1}{n^{1+\delta}} \\ &= 1 - \frac{1}{n^{\delta}}. \end{aligned}$$

That enables us to conclude (1.3.12).

□

In 1982, Sinai proved the so-called *localization theorem*, which shows that the asymptotic behaviour of the normalized process converges in $\mathbf{P}_0$-probability to a non-degenerated random variable.

Theorem 1.3.4. *(Sinai, 1982) Assume the conditions **C.1**, **C.2** and **C.3**. Then, there exists a random variable $m_n = m_n(\alpha)$ of the random environment such that for any $\delta > 0$,*

$$\lim_{n\to\infty} \boldsymbol{P}_0\left(\left|\frac{\sigma^2 X_n}{(\log n)^2} - m_n\right| < \delta\right) = 1.$$

Moreover, m_n has a limit distribution i.e.,

$$\lim_{n\to\infty} \mathbb{P}(m_n \leqslant x) = G(x).$$

Consequently the distribution of $\sigma^2 X_n/(\log n)^2$ converges to the same distribution $G(x)$ under $\mathbf{P}_0$.

The following idea is an interpretation of the Theorem 1.3.4. Given an environment α and a trajectory of the Sinai walk in a window time $[0,n]$, the random variable m_n which depends only on the environment α (specifically on the potential). It can be interpreted as the location of the bottom of the deepest valley in the sites of the state space where the Sinai walk has visited until the time n.

Remark 1.3.5. *Consider the Theorem 1.3.4. We assume that we have information about the trajectory of the Sinai walk. We know that the Sinai walk will spend much time around the deepest valley, that is, if the trajectory at time n is visiting any another state (for instance, another minimum of the potential) we know that it is likely that in the near future the walk reaches the state where is the deepest valley.*

Afterwards (Kesten, 1986) and (Golosov, 1986) proved independently that the previous limit coincides with the distribution of a certain functional of the standard Brownian motion, with density function

$$G'(x) = \frac{2}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\exp\left(-\frac{(2k+1)^2\pi^2}{8}\mid x\mid\right).$$

1.4 Recovering the random environment

In what follows, we describe an algorithm that allows us to recover the environment α from samples of the paths. First we introduce the definition of local time for stochastic processes in discrete time.

Definition 1.4.1. *The local time $\mathcal{L}$ at k ($k\in\mathbb{Z}$) within time window $[1,T]$ of the random walk X_n is defined as*

$$\mathcal{L}(k,T) = \sum_{i=1}^{T} 1_{\{X_i=k\}}.$$

In other words, the local time in the site k is the number of times that the process crosses the site k in a finite time window. Furthermore, we need to consider the following definitions:

$$\begin{aligned}\mathcal{L}^*(n) &= \max_{k\in\mathbb{Z}}(\mathcal{L}(k,n)),\\ \mathbb{F}_n &= \{k\in\mathbb{Z}:\ \mathcal{L}(k,n)=\mathcal{L}^*(n)\},\\ k_n^* &= \inf\{\mid k\mid:\ k\in\mathbb{F}_n\}.\end{aligned}$$

We say that $\mathcal{L}^*(n)$ is the maximum number of times that the walk visits the same state within time window $[0,n]$. On the other hand, the set $\mathbb{F}_n$ is called the set of the *favourite sites*. Moreover k_n^* is the smallest favourite site in absolute value, thereby in according with the Definition 1.2.1, $T_{k_n^*}$ denotes the first time that the walk X_n reaches the state k_n^*.

1.4.1 The procedure of Andreoletti

Consider $k,l \in \mathbb{Z}$ and $c_0 > 0$. Define for each $n > 1$,

$$\begin{aligned} V_{k,l}^n &= 1 - \frac{1}{\log n}(V(k) - V(l)), && (1.4.1)\\ \hat{V}_k^n &= \frac{\log(\mathcal{L}(k,n))}{\log n}, && (1.4.2)\\ u_n &= \frac{c_0 \log(\log(\log(n)))}{\log n}, \\ \mathbb{L}_n^\gamma &= \left\{ l \in \mathbb{Z} : \sum_{j=T_{k_n^*}}^{n} 1_{\{X_j = l\}} \geqslant (\log n)^\gamma \right\}. \end{aligned}$$

For l fixed, $k \mapsto V_{k,l}^n$ is a function that only depends on the environment, because the potential $V(k)$ depends only on α. In the paper (Andreoletti, 2011), it was proposed $\hat{V}_k^n$ as a good estimator for the function V_{k,m_n}^n and u_n as an error function. The author explains how the algorithm works on the set $\mathbb{L}_n^\gamma$, which is a random subset of $\mathbb{Z}$ that contains the sites where the local time after the instant $T_{k_n^*}$ is large enough.

Here we present his main result.

Theorem 1.4.2. *(Andreoletti, 2011)*
*Assume the conditions **C.1**, **C.2** and **C.3**. There exists constants c_0, c_1, c_2 and c_3 such that for all $\gamma > 6$, there exists n_0 such that for all $n > n_0$, there exists $G_n \subset \Omega_1$ with $\mathbb{P}(G_n) \geqslant 1 - \phi_1(n)$ and*

$$\inf_{\alpha \in G_n} P_0^\alpha \left(\bigcap_{k \in \mathbb{L}_n^\gamma} \{ \mid \hat{V}_k^n - V_{k,m_n}^n \mid < u_n \} \right) \geqslant 1 - \phi_2(n), \qquad (1.4.3)$$

where

$$\begin{aligned} \phi_1(n) &= \frac{c_1 \gamma \log(\log n)}{\log n}, \\ \phi_2(n) &= \frac{c_2 (\log(\log n))^2}{(\log n)^{\gamma/2-2}} + \frac{c_3 (\log(\log n))^8}{(\log n)^{\gamma-6}}. \end{aligned}$$

The theorem tell us that there exists a subset of environments G_n with large probability, where we can approximate (with large probability) to V^n_{k,m_n} through the estimator $\hat{V}^n_k$ on the states sufficiently visited by the Sinai walk, i.e., the states that belong to $\mathbb{L}^{\gamma}_n$.

However to run the algorithm given by (Andreoletti, 2011), the author explains that it is possible to approximate m_n with k^*_n, which only depends on the local time. Consequently, we can program the algorithm with k^*_n instead of m_n.

We programmed such algorithm in the statistical software R (Core et al., 2013) and we have written the pseudo-code in the Appendix B1. For instance, we took the trajectory of the Figure 1.2. Such path shows at first sight two sites frequently visited, around of the state $x = 7$ and $x = 50$. In addition, can see in Figure 1.4 the potential $V(x)$ given by the Definition 1.2.2 associated with the environment α with which was generated the trajectory. In fact, we observe that the potential have a deeper minimum around of the state $x = 7$.

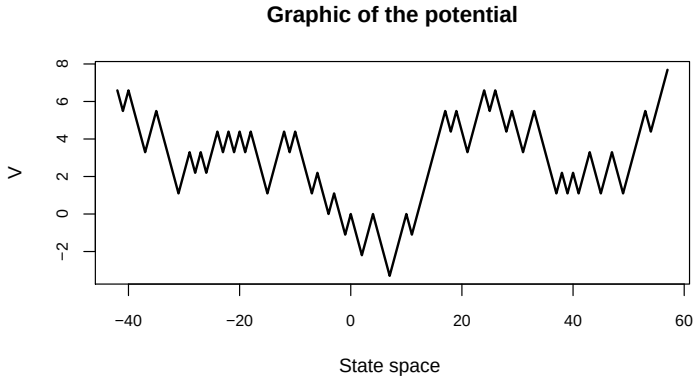

Figure 1.4: Potential $V(x)$ associated to the environment α as in (1.2.16), with which was generated the trajectory shown in Figure 1.2.

Once we have the trajectory of the Sinai walk, we apply the procedure of Andreoletti, which result is showed in Figure 1.5. In Appendix B1, we denoted by `Vnke`, `Vnk` and `un` instead of V^n_{k,k^*_n}, $\hat{V}^n_k$ and u_n, respectively.

In the practice, when we have a trajectory of Sinai's walk. We wish to recover the potential and to know where it will be more likely to find the random walk for a larger time window. In the paper (Andreoletti, 2011) the author makes the comment about when it is possible to recover the potential. He explains that if $0 \in \mathbb{L}^{\gamma}_n$, i.e., 0 has local time large enough, we can obtain explicitly from the function of the potential

$$\hat{V}^n_k \approx V^n_{k,k^*_n} = 1 - \frac{1}{\log n}(V(k) - V(k^*_n)), \tag{1.4.4}$$

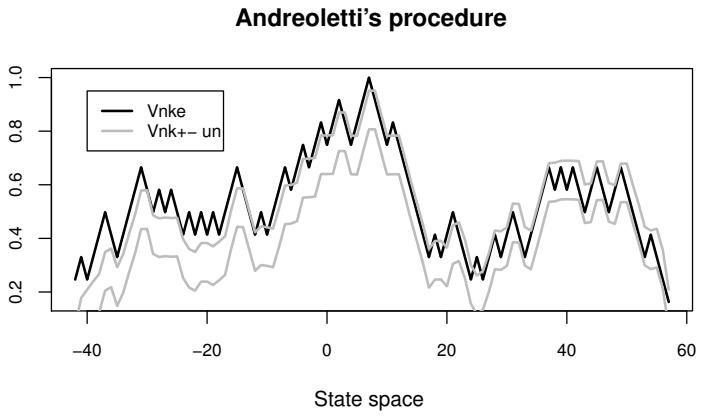

Figure 1.5: Simulation of Andreoletti's algorithm given a trajectory of Sinai's random walk.

substituting $k = 0$ and by definition $V(0) = 0$, then

$$\begin{aligned} V(0) - V(k_n^*) &= \log n[1 - V_{0,k_n^*}^n] \\ -V(k_n^*) &\approx \log n[1 - \hat{V}_0^n] \\ V(k_n^*) &\approx -\log n\left(1 - \frac{\log(\mathcal{L}(0,n))}{\log n}\right) \\ &= \log(\mathcal{L}(0,n)) - \log n. \end{aligned} \tag{1.4.5}$$

Therefore thanks to the differences in (1.4.4) and (1.4.5), we can approximate the value of $V(k)$ for $k \in \mathbb{L}_n^\gamma$ as

$$\begin{aligned} V(k) &\approx \log n\left(1 - \frac{\log(\mathcal{L}(k,n))}{\log n}\right) + V(k_n^*) \\ &= \log n - \log(\mathcal{L}(k,n)) + \log(\mathcal{L}(0,n)) - \log n \\ &= \log\left(\frac{\mathcal{L}(0,n)}{\mathcal{L}(k,n)}\right). \end{aligned} \tag{1.4.6}$$

For the preceding example, the trajectory that we have considered makes sufficient visits to the state $x = 0$. So, we can apply the previous algorithm to recover the potential and it is a good estimation as we can see in Figure 1.6, where we have denoted to $V(k)$ as `Vk` given by (1.4.6).

When we have a trajectory of Sinai's random walk in a time window $[0,n]$***, using only the local time, if the state*** $0 \in \mathbb{L}_n^\gamma$***, we can give a good estimation of the potential associated with the environment which was generated the trajectory.***

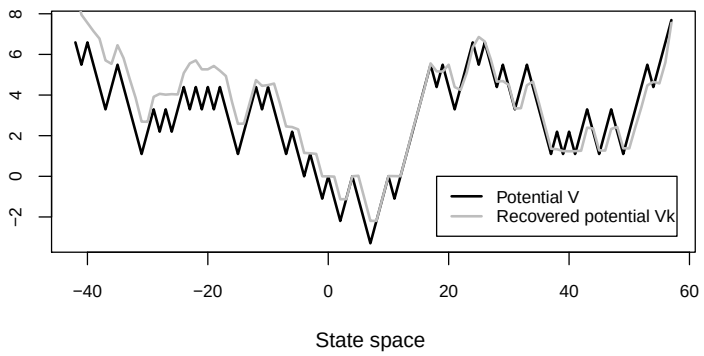

Figure 1.6: Recovered potential using (1.4.6).

Remark 1.4.3. *But what happens outside of $\mathbb{L}_n^\gamma$? Probably the trajectory of the Sinai walk presents other sites frequently visited, but not enough to belong to $\mathbb{L}_n^\gamma$. (Andreoletti, 2011) also comments on the case where $0 \notin \mathbb{L}_n^\gamma$, with this drawback we can not recover $V(k_n^*)$ with the mentioned argument. However, using the expression (1.2.3) and the Definition 1.2.3, it is possible to recover the environment and therefore the potential for states near to k_n^*. The problem is that while the states are far from k_n^*, its approximation will be worse and therefore the sum that appears in the potential $V(x)$ will accumulate probably lots of errors.*

Let us to explain with detail previous remark. From the equation (1.4.1), we know the approximated value of the differences once we have the values of $\hat{V}_k^n$ thanks to the local time, i.e.,

$$V(k) - V(k_n^*) = (\log n)(1 - \hat{V}_k^n). \tag{1.4.7}$$

Then we can calculate for all $i \in \mathbb{Z}$ by (1.2.3) the values

$$\begin{aligned} \alpha_i &= \frac{e^{-(V(i)-V(i-1))}}{e^{-(V(i)-V(i-1))}+1} \\ &= \frac{e^{-(V(i)-V(k_n^*))+(V(i-1)-V(k_n^*))}}{e^{-(V(i)-V(k_n^*))+(V(i-1)-V(k_n^*))}+1}. \end{aligned} \tag{1.4.8}$$

The analysis of Andreoletti tell us that the differences (1.4.7) are good when $k \in \mathbb{L}_n^\gamma$, and therefore the value of α_k given by (1.4.8) will be a good estimation when k is enough visited. Finally we can recover the potential. But if k is far from k_n^*, the value of α_k is not a good estimator of the original environment. Then by the definition of the potential, $V(x)$ will accumulate lots of errors. Thereby we could conclude a false behaviour of Sinai's random walk given that environment α.

1.4.2 Refinement of Andreoletti's algorithm

In what follows, we will explore another example of Sinai's random walk. In Figure 1.7 the trajectory does not have enough visits to the origin, its favourite site in the time window $[0, 500000]$ is the state $k_n^* = 96$, which is far from the origin.

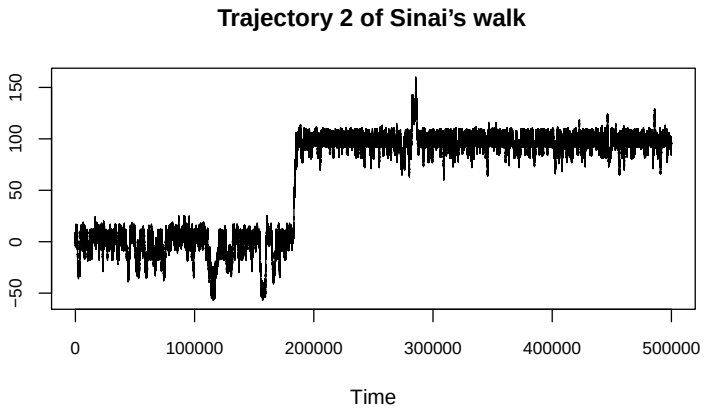

Figure 1.7: Second trajectory of Sinai's walk.

If we apply the mechanism to recover the potential in the same way as we did it for the Figure 1.2, we obtain the following approximation showed in Figure 1.8, which leads to lots of errors. For instance, when we see the recovered potential, we could believe that the minimum around the state $x = 20$ is deeper than the minimum around the state $x = 100$, and to conclude a false interpretation for the behaviour for other trajectories with that potential.

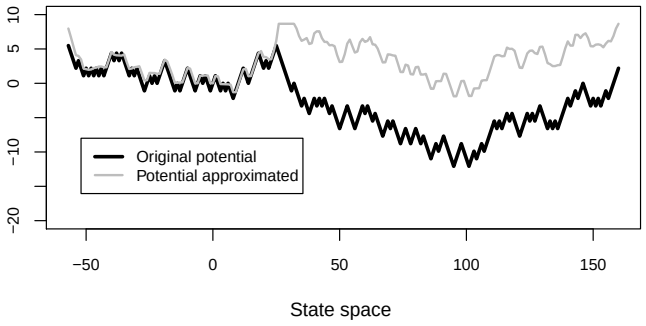

Figure 1.8: Recovered potential when we use Andreoletti's procedure in accordance with (1.4.6), but k_n^* is far from the origin.

On the other hand, we apply the mechanism of recovering of the potential in accordance with the Remark 1.4.3 and we can observe the result in Figure 1.9. Maybe the approximation is not the best because there are sites j where the local time is small, then we do not have enough information to recover α_j and neither the potential. However, this approximation respects the position where occurs the minimum of the original potential.

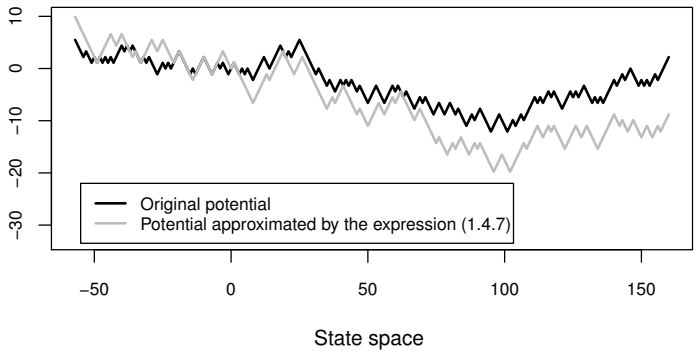

Figure 1.9: Potential generated by the recovery of each α_j through the formula (1.4.8).

In this subsection, we focus on improving the algorithm of Andreoletti. Let us explain this using the information obtained in Figure 1.7. Hereafter we use the following notation:

$$\begin{aligned}
n &= 500000 \\
(X_k)_{k=0,\dots,n} &= \text{the trajectory of Sinai's walk given in Figure 1.7} \\
S &= \text{the states visited by} (X_k)_{k=0,\dots,n} \\
T_{100} &= \min\{k : X_k = 100\} \\
n_2 &= n - T_{100} \\
(X'_k)_{k=0,\dots,n_2} &= (X_k)_{k=T_{100},\dots,n} \\
(Y_k)_{k=0\dots n_2} &= (X'_k)_{k=0,\dots,n_2} - 100 \\
S' &= \text{the states visited by } (Y_k)_{k=0,\dots,n_2}
\end{aligned}$$

We have considered the time window $[0,n]$ in two parts $[0,T_{100}-1]$ and $[T_{100},n]$. For that reason the trajectory $(Y_k)_{k=0,\dots,n_2}$ is such that

$$Y_0 = X'_0 - 100 = X_{T_{100}} - 100 = 100 - 100 = 0.$$

This is, we cut a part of the original trajectory and we translate it to the origin. Therefore as we mentioned before the favourite site of $(X_k)_{k=0,\dots,n}$ is $k_n^* = 96$, and now we have calculated that $k_{n_2}^* = -4$ is the smallest favourite site in absolute value of $(Y_k)_{k=0,\dots,n_2}$. For both trajectories, we know explicitly the sets

$$\begin{aligned} S &= \{-57,-56,\dots,160\}, \\ S' &= \{-40,-39,\dots,60\}. \end{aligned}$$

It is also possible to calculate the local time for each trajectory in its corresponding time window. We denote it by

$$\begin{aligned} \mathcal{L}(k,n) &= \text{the local time of } k \in S, \text{ for } (X_k)_{k=0,\dots,n} \\ \mathcal{L}'(k,n_2) &= \text{the local time of } k \in S', \text{ for } (Y_k)_{k=0,\dots,n_2}. \end{aligned}$$

When we apply the procedure of Andreoletti, we recall that he proposed the estimator $\hat{V}_k^n$ as an approximation of the function of the environment $V_{k,k_n^*}^n$. This is,

$$\frac{\log(\mathcal{L}(k,n))}{\log n} \approx 1 - \frac{1}{\log n}(V(k) - V(k_n^*)) \tag{1.4.9}$$

in accordance with (1.4.1) and (1.4.2). Moreover we remember that it is possible to compare the estimator $\hat{V}_k^n$ versus the function of the environment $V_{k,k_n^*}^n$ thanks to know initially the potential $V(k)$ for all $k \in \mathbb{Z}$ or for the case of our pseudo code, it is enough $k \in S$ (see Appendix B1).

The main idea is to improve the recovered potential. We are going to propose two ideas which give us a better numerical approximation.

Proposal 1 From the fact that the trajectory X_n spends a lot of time around the state $x = 100$ we can hope that the minimum of the potential in that site is deeper than any other point.

Let us add a special feature, as we have assumed that the trajectory $(Y_k)_{k=0,\dots,n_2}$ begins in 0, we are going to consider that the potential V begins in that origin. We know beforehand the environment α then it is not difficult to program the construction of the potential V with the values α_i associated with those states. To avoid creating a a misunderstanding, we denote V_2 that potential.

We apply Andreoletti's algorithm to the trajectory $(Y_k)_{k=0,\dots,n_2}$, i.e., for the time window $[0,n_2]$ and $k_{n_2}^* = -4$.

$$\begin{aligned} V_2(k) &\approx (\log n_2)\left(1 - \frac{\log(\mathcal{L}'(k,n_2))}{\log n_2}\right) + V_2(k_{n_2}^*) \\ &= \log n_2 - \log(\mathcal{L}'(k,n_2) + V_2(k_{n_2}^*). \end{aligned} \tag{1.4.10}$$

For this trajectory the value of $k^*_{n_2}$ is near to the new origin, then it is possible to take the approximation of $V_2(k^*_{n_2})$ with the local time $\mathcal{L}'(k,n_2)$ as

$$V_2(k^*_{n_2}) \approx \log(\mathcal{L}'(0,n_2)) - \log n_2. \tag{1.4.11}$$

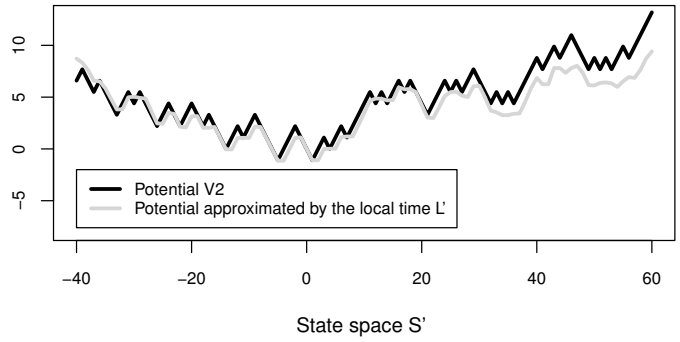

Figure 1.10: Recovered potential by (1.4.10) and (1.4.11) in the states S'.

Therefore we recover the potential V_2 for the states S' as we can observe in Figure 1.10, however we are interested in the original potential V. For that reason, we propose to move down the depth $V_2(k^*_{n_2}) \approx -3.2406$ on the recovered values initially by the methodology (1.4.6) in the states $S' + 100 = \{60, 61, \ldots, 160\}$. We can see such movement in Figure 1.11.

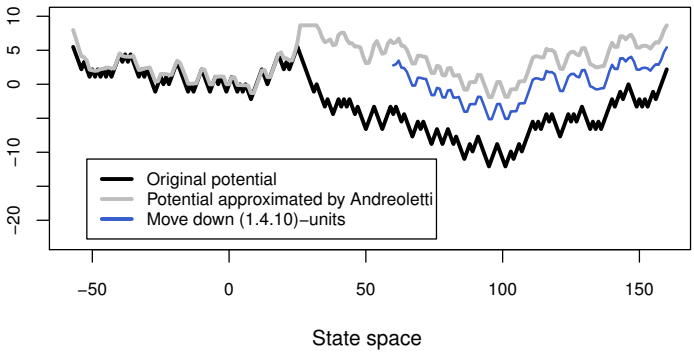

Figure 1.11: Recovered potential by (1.4.6) in the states S and moved down (1.4.11)-units on the states $S' + 100$

Although we do not have enough information in the states $[26,59]$, because the original trajectory has few visits in those states and then, the local time gives us little information. But with the previous procedure we only join the approximation for the states $[-57,25]$ with $[60,160]$ concatenating with the segment of the approximated trajectory by the algorithm of Andreoletti in the states $[26,59]$. The result is shown in Figure 1.12, it is far to be a good estimation when compared with the original potential. However it is closer to reality if initially we do not know the environment.

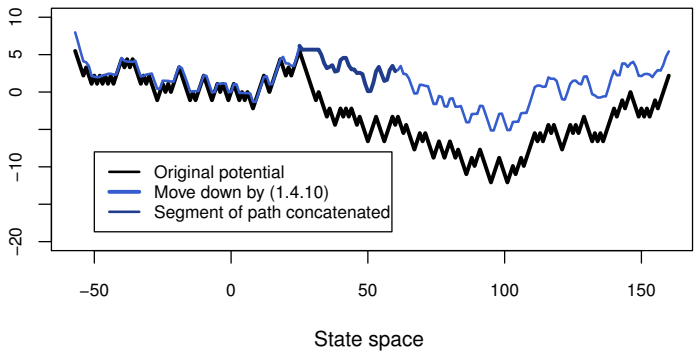

Figure 1.12: The trajectory of the segments moved and concatenated.

Proposal 2 As a continuation of the previous idea, we know that the depth in the state $k_n^* = 96$ should be large than the depth in any other state, specially deeper than other minima of the potential V.

We heuristically propose to make use only of the local time, we sum the depth obtained by $V_2(k_{n_2}^*)$ in (1.4.11) plus the depth given by Andreoletti's procedure $V(k_n^*)$ in (1.4.5). We denote such length as $\tilde{V}(k_n^*) := V_2(k_{n_2}^*) + V(k_n^*)$. Then the recovered potential by Andreoletti's procedure (1.4.6) on the states $S' + 100$, it is move down $\tilde{V}(k_n^*)$-units. Explicitly we know for the trajectories $(X_k)_{k=0,\dots,n}$ and $(Y_k)_{k=0,\dots,n_2}$, the values

$$\begin{aligned} \tilde{V}(k_n^*) &\approx (-3.2406) + (-4.45554) \\ &= -7.69614 \end{aligned} \tag{1.4.12}$$

In the same way that the procedure of the Proposal 1, we join the approximation for the states $[-57,25]$ with $[60,160]$ concatenating with the segment of the approximated trajectory of Andreoletti's method. It is possible to observe the result of such approximation to the potential in Figure 1.13.

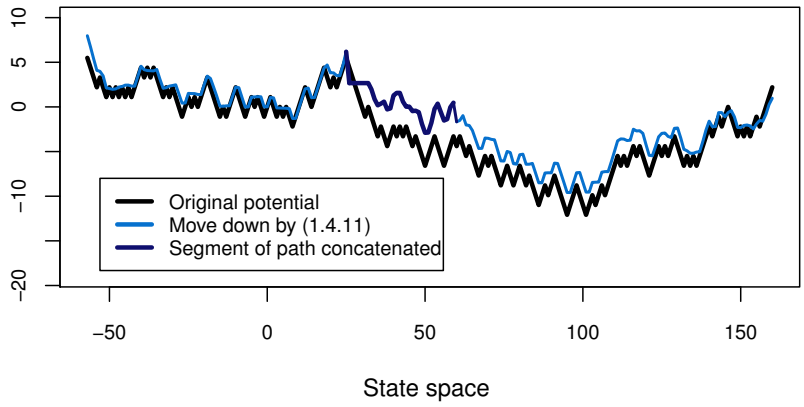

Figure 1.13: Original potential and the Proposal 2 using (1.4.12).

Summarise

It is essential the information about the environment α to know the behaviour of the random walk for the quenched case. Because thanks to the potential associated to the environment we can predict in which states the random walk spends a lot of time or which they will be little visited.

We assume $(X_k)_{k=0,\ldots,n}$ a trajectory of Sinai's random walk in a time window $[0,n]$. By the behaviour of the trajectory, we can possibly observe the following scenarios:

i) the random walk has the smallest favourite site k_n^* near to the origin. In this case, the mentioned procedure in (Andreoletti, 2011) works satisfactorily to recover the potential. Also if the random walk visits sufficiently other minimum, as we had seen applying the Remark 1.4.3 it is possible to recover the environment in a larger number of states. Thus we will have recovered a good estimation of the potential for a greater number of states.

ii) the smallest favourite site k_n^* is far from the origin, but we can observe other site frequently visited near to the state $x = 0$ (as in the presented case in Figure 1.7). That site do not belong to $\mathbb{L}_n^\gamma$, however we can make a cut in the path and a translation to the origin, then to count the local time in a new time window $[0,n_2]$ on the new visited states. Therefore we could apply the analogous to the Proposal 1 or Proposal 2 to improve the approximation of the original potential.

iii) other case, the smallest favourite site k_n^* is far from the origin and there is not another minimum frequently visited in the time window $[o,n]$, or if there is, it is

even further away. For that case, the algorithm given by (Andreoletti, 2011) does not provide a good estimation of the environment, and it is not feasible either apply a cut of the path and count the local time of the new path because the information about the environment is limited near to the origin, therefore by the Definition (1.2.3) is more difficult obtained a good estimation of the potential by this methodologies.

iv) finally, it is possible that the favourite site k_n^* is not in a valley sufficiently depth, and maybe there are two o more states visited by the random walk in a short time, i.e., the walk moves in short periods of time between those minima, see for example Figure 1.14. Then it is not a good idea to apply the procedure to cut the path and count the local time of the new path. However, the approximation with the local time given by (Andreoletti, 2011) is not bad in accordance with the simulations carried out. We can observe in Figure 1.15 the result of the procedure of Andreoletti, i.e., we apply (1.4.6) to recover the potential associated to the environment, which we generated the trajectory in Figure 1.14.

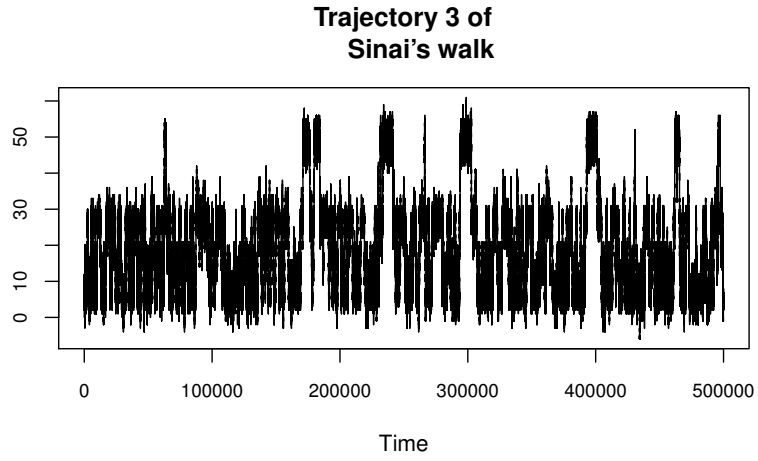

Figure 1.14: Example of a trajectory of Sinai's random walk.

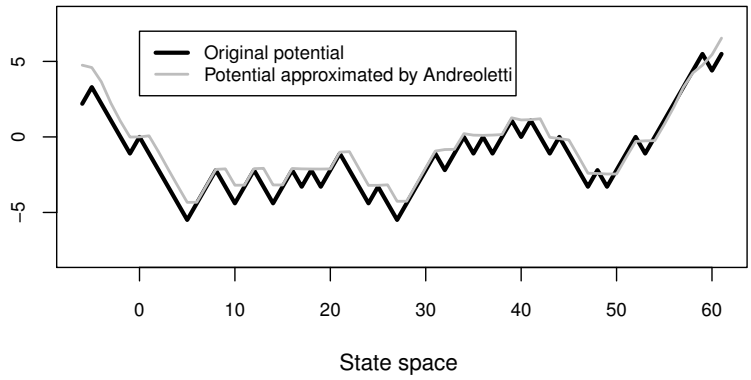

Figure 1.15: Original potential and recovered potential using (1.4.6).

1.5 Appendix A1

In order to explain why it is possible to obtain a space of probability large enough described in the Lemma 1.3.2, we present a summary of some results. For such a task, we sketch out the following proposition about the sum of random variables i.i.d., which is taken from (Einmahl, 2009); Corollary 2.2 for one-dimension.

Proposition 1.5.1. *Let $Y, Y_1, Y_2, \ldots$ be i.i.d mean zero random variables on a probability space $(\Omega, \mathcal{F}, P)$. Assume that $\sigma^2 := \mathbb{E}|Y|^2 < \infty$. Let $Z(0) = 0$ be and for $k \geqslant 1$, $Z(k) = \sum_{i=1}^{k} Y_i$. For each $n \geqslant 1$, define the partial sum process sequence $Z^n = \{Z^n(t)\}_{0 \leqslant t \leqslant 1}$, this is, $Z^n : \Omega \to (\mathcal{C}([0,1]), d)$ given by*

$$Z^n(t) = \begin{cases} Z(k) & \text{if } t = \frac{k}{n},\ 0 \leqslant k \leqslant n, \\ \text{linearly interpolated} & \text{otherwise.} \end{cases} \tag{1.5.1}$$

Assume the following statements:

Hypothesis 1. *For a sequence c_n such that $c_n/\sqrt{n}$ is eventually non-increasing.*

Hypothesis 2. *$\sum_{n=1}^{\infty} P\{|\, Y \,| \geqslant c_n\} < \infty$.*

Hypothesis 3. *There exist a $\gamma \in (1/3, 1)$ such that c_n/n^{γ} is eventually non-decreasing.*

Then a construction is possible such that

$$d\left(\frac{1}{c_n} Z^n, \frac{\sigma}{c_n} \tilde{W}^n\right) \xrightarrow{P} 0, \quad n \to \infty, \tag{1.5.2}$$

where $(W(t), t \geqslant 0)$ denote the Brownian motion and $\tilde{W}^n(t) = W(nt)$, $0 \leqslant t \leqslant 1$, $\xrightarrow{P}$ stand for convergence in probability, and d is the metric induced by the sup-norm on $\mathcal{C}([0,1])$.

Therefore we can apply the Proposition 1.5.1 to obtain the following. Given $\varepsilon > 0$ and $\delta > 0$, there exist N such that if $n \geqslant N$, then

$$P\left(\omega \in \Omega : \sup_{0\leqslant t\leqslant 1}\left|\frac{1}{c_n}Z^n(t) - \frac{\sigma}{c_n}\tilde{W}^n(t)\right| < \varepsilon\right) \geqslant 1-\delta. \tag{1.5.3}$$

In other words, there exist a subset of Ω large enough such that the trajectories of the rescaled walk approximates to trajectories of the rescaled Brownian motion. Of course as mentioned by (Einmahl, 2009) when $c_n = \sqrt{n}$, it follows the well-known Donsker's theorem.

Furthermore when defining the partial sum process sequence $Z^n : \Omega \to (\mathcal{C}([0,1]), d)$, the space $\mathcal{C}([0,1])$ may be replaced with $\mathcal{C}([0,T])$ for any $T > 0$ (consequently also with $\mathcal{C}([-T,T])$, and even with $\mathcal{C}([0,\infty))$ through the locally uniform convergence and therefore on $\mathcal{C}(\mathbb{R})$.

Let us explain with more detail. We shall show that we can apply the potential $V^n(t)$ defined by (1.3.2) to obtain the result of the Proposition 1.5.1.

First we write the definition described in (1.3.2) as

$$V^n(t) = \begin{cases} \frac{1}{\log n}\sum_{i=0}^{k}\log\frac{1-\alpha_i}{\alpha_i} & \text{if } t = \frac{k}{\log^2 n},\ k = 1,2\ldots \\ 0 & \text{if } t = 0 \\ -\frac{1}{\log n}\sum_{i=k+1}^{0}\log\frac{1-\alpha_i}{\alpha_i} & \text{if } t = \frac{k}{\log^2 n},\ k = -1,-2\ldots \\ \text{linearly interpolated} & \text{otherwise.} \end{cases}$$

We want to apply the Proposition 1.5.1. With that mindset we propose $c_n = \log n$. Thereby we take the random walk $V^n(t)$ with the spatial rescaling factor c_n and a scaling factor in the time $\frac{1}{(\log n)^2}$. Now, we will prove that the hypothesis for the Proposition 1.5.1 are true.

Hypothesis 1. The sequence $c_n/\sqrt{n}$ (the spatial scaling factor between the square root of the reciprocal of the scaling factor in the time) is eventually non-increasing. In fact, we have the quotient

$$\frac{\log n}{\sqrt{(\log n)^2}} = 1,$$

for all $n \in \mathbb{N}$.

Hypothesis 2. We know that the variable $\log\frac{1-\alpha_0}{\alpha_0}$ is bounded $\mathbb{P}$-a.s by (1.1.3), that is,

$$\left|\log\frac{1-\alpha_0}{\alpha_0}\right| < \log\frac{1-\beta}{\beta},$$

for some $0 < \beta < 1/2$. Note that for such β fixed, there exist N_0 large enough such that

$$\log \frac{1-\beta}{\beta} < \log N_0.$$

Therefore

$$\mathbb{P}\left(\left|\log \frac{1-\alpha_0}{\alpha_0}\right| \geqslant \log n\right) = 0 \quad \text{for all } n \geqslant N_0.$$

As a consequence

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\left|\log \frac{1-\alpha_0}{\alpha_0}\right| \geqslant \sqrt{n}\right)$$

$$= \sum_{n=1}^{N_0-1} \mathbb{P}\left(\left|\log \frac{1-\alpha_0}{\alpha_0}\right| \geqslant \sqrt{n}\right) + \sum_{n=N_0}^{\infty} \mathbb{P}\left(\left|\log \frac{1-\alpha_0}{\alpha_0}\right| \geqslant \sqrt{n}\right) < \infty.$$

Hypothesis 3. If we choose $\gamma = 1/2 \in (1/3, 1)$, the quotient of the spatial scaling factor between the reciprocal of the scaling factor in the time is a sequence non-increasing. Indeed

$$\frac{\log n}{((\log n)^2)^{\gamma}} = 1,$$

for all $n \in \mathbb{N}$.

Therefore we have the main result of the Proposition 1.5.1, i.e., given $\varepsilon > 0$ and $\delta > 0$, there exist N large enough such that if $n \geqslant N$. Then

$$P(\Omega_n) \geqslant 1 - \delta, \tag{1.5.4}$$

where $\Omega_n = \{\omega \in \Omega : d\left(V^n(t), \frac{\sigma}{\log n}\tilde{W}^n(t)\right) < \varepsilon$, and $(W(t), t \in \mathbb{R})$ denote a two-sided Brownian motion, $\tilde{W}^n(t) = W((\log n)^2 t)$, and d is the metric defined on $\mathcal{C}(\mathbb{R})$ as

$$d(f,g) = \sum_{i=1}^{n} \frac{1}{2^n} \min(1, \max_{-n \leqslant t \leqslant n} | f(t) - g(t) |).$$

Finally, we know some results for the Brownian motion (we refer the reader to (Mörters and Peres, 2010; Peres et al., 2001)), which we can apply on a space of probability large enough to the trajectories of the scaling potential in accordance with the inequality (1.5.4).

Lemma 1.5.2. *Define* $\tau = \{t > 0 : | W(t) | = 1\}$. *Then with probability 1,* $\tau < \infty$.

Lemma 1.5.3. *Given two disjoint closed time intervals, the maxima of Brownian motion on them are different almost surely.*

In other words, in the space of probability mentioned in (1.5.4), the trajectories of the scaling potential have certain properties refereed in the Lemma 1.3.2 as follow: the properties *i*) to *iii*) are true by construction of the valleys, while *iv*) is due to Lemma 1.5.3, i.e., the minima constructed in each refinement are different a.s. and *v*) is a consequence of the Lemma 1.5.2.

Appendix B1

Here we set forth the following algorithm in *R Studio* (Core et al., 2013) to obtain a trajectory of the Sinai walk for a given environment. We will use the symbol # to introduce comments.

Algorithm 1 Simulation of a trajectory of Sinai's random walk X_k

Step 1. We define the variable.

```
n=number of steps time.
```

 # We first fix an environment, which we call it `alphaE`*. To do this we use the instruction sample, which allows us to generate a sample of the specified size (we chose* $2n-1$*) from a data set (we have chosen* $\{1/4,$ *or* $3/4\}$*), either with or without replacement (with replacement for this case), where one can specify the probability weights for obtaining the elements of the vector being sampled.*

```
alphaE<-sample(c(1/4,3/4),2*n-1, replace=T,
prob=c(0.5,0.5))
```

 # We are going to define the function `g` *to associate the vector* `alphaE` *with the state Space* $\mathbb{Z}$ *restricted to the set* {-(n-1),...,-2,-1,0,1,2,...,n-1}.

```
g<-function(x){x+n}
```

Step 2. We generate the trajectory of Sinai's random walk. For that we use an auxiliary function, it denoted by `aux1`, which takes a value 1 or -1 depending of the probability weights assigned by the environment `alphaE`. Finally, `X` represents the trajectory of the Sinai walk.

```
aux1<-numeric(n)
X<-numeric(n+1)
X[1]<-0
for(i in 1:n)
  {
   aux1[i]<-sample(c(1,-1), 1, prob=c(alphaE[g(X[i])],
   1-alphaE[g(X[i])]))
   X[i+1]<-X[i]+aux1[i]
  }
```

Step 3. We plot the trajectory that we have generated.

```
plot(c(0:n), X, type="l", col="black",
xlim=range(c(0:n)),lwd=2,xlab="time")
title(main = "Trajectory of the Sinai walk after
n-steps ")
```

Algorithm 2 Andreoletti's algorithm

Step 1. We assume that we have a trajectory of the Sinai walk X_k. The first step is calculate the local time in each x in the state space. We shall be aware that such a task is not easy, the computational time is large, because we are counting how many times X_k reaches the state x in the finite time window $t = 0, 1, \ldots, n$. Note that we had generated an environment `alphaE` truncated down until the state $-(n-1)$ and at the top until the state $(n-1)$, this is, we need to run the algorithm of the local time $(2n-1)$-times. For that reason, one prefers run the algorithm only on the states visited by X_k. Here, we call `Lkn` to the value $\mathcal{L}(k,n)$.

```
M=max(X)
m=min(X)
S=c(m:M)
# S denotes the visited states by the random walk X.
L=length(S)
Lkn<-numeric(L)
for(j in 1:L)
 {
  indicatiorFunction<-numeric(length(c(2:(n+1))))
   for(i in 2:(n+1)
   {
  if(X[i]==S[j]){indicatiorFunction[i]<-1}
  else {indicatiorFunction[i]<-0}
   }
 Lkn[j]<-sum(indicatiorFunction)
 }
```

Step 2. We use the notation `Len` $= \mathcal{L}^*(n)$, `ke` $= k_n^*$ and `Tke` $= T_{k_n^*}$ in accordance with the definitions given in the Subsection 1.4

```
Len<-max(Lkn)
aux2<-which(Lkn==Len)
Fn=S[aux2]
if(min(Fn)>0){ke=min(abs(Fn))}
if(min(Fn)<0){ke=-min(abs(Fn))}
Tke<-(min(which(X==ke)))-1
```

Step 3. We seeking which states belongs to the set $\mathbb{L}_n^{\gamma}$, which we denote by `Lngam` and `gam` instead of γ for the pseudo code.

```
aux3<-numeric(L)
indicatiorFunction2<-numeric(length(c((Tke+1):(n+1))))
for(j in 1:L)
{
  for(i in (Tke+1):(n+1))
     {
      if(X[i]==S[j]){indicatorFunction2[i]<-1}
      else{indicatorFunction2[i]<-0}
     }
aux3[j]<-sum(indicatorFunction2)
}
gam=7
Lngam<-which(aux3>=((log(n)) ∧ gam))
```

Step 4. We define the function of the potential $V(x)$ and the estimator defined by Andreoletti. We use the notation `Vnke`, `Vnk` and `un` instead of V^n_{k,k_n^*}, $\hat{V}^n_k$ and u_n, respectively.

```
V=numeric(length(L))
# V denotes the potential only for the visited states.
z<-which(S==0)
V[z]=0
if(M>0){
 for(i in c(1:M)){
 suma1<-0
 for(j in 1:i){
 suma1<-suma1+log((1-E[g(j)])/E[g(j)])
 }
 V[i+z]<-suma1
 }}
if(m<0){
 for(i in c(m:(-1))){
 suma2<-0
 for(j in (i+1):0){
 suma2<-(suma2-log((1-E[g(j)])/E[g(j)]))
 }
```

```
 V[i+z]<-suma2
 }}
Vnke<-numeric(L)
 for(i in 1:L)
 {
  Vnke[i]<-1-(1/(log(n)))*(V[i]-V[which(S==ke)])
 }
Vnk<-numeric(L)
 for(j in 1:L)
 {
  Vnk[j]<-(log(Lkn[j]))/(log n)
 }
un<-(log(log(log(n))))/(log n)
```

Step 5. We plot the graphic that shows the function of the potential `Vnke` and the estimator `Vnk+un`, `Vnk-un`.

```
plot(S, Vnke, type="l",col="black",xlim=range(S),lwd=2,
xlab="States visited", ylab=" ")
title(main = "Andreoletti's approximation",col.main =
"black",cex.main = 1.1, font.main = 1)
legend(S[1],-0.5,legend=c("Vnke","Vnk±un"),
col=c("black","gray"), lwd=2, text.width = NULL)
lines(S, Vnk+un, col="gray")
lines(S, Vnk-un, col="gray")
```

 # In accordance with **??** *from the Andreoletti's procedure we can recover de potential through:*

```
Vk=numeric(L)
    for(i in 1:L)
    {
    Vk[i] = log ( Lkn[which(S==0)]/Lkn[i] )
    }
```

Algorithm 3 Recover the potential in accordance with the Subsection 1.4.2

```
# We denote by T100= T_100, n2= n_2, Xprime= X', Y= Y_k and Sprime= S'
T100=which(X==100)[1]
n2=n-T100
Xprime=X[T100:(n+1)]
Y=Xprime-100
mprime=min(Y)
Mprime=max(Y)
Sprime<-c(mprime:Mprime)
# We use the algorithm 2 to apply Andreoletti's procedure.
Lprime=length(Sprime)
Lknprime<-numeric(Lprime)
for(j in 1:Lprime)
 {
  indicatiorFunction3<-numeric(length(c(2:n2)))
   for(i in 2:n2)
   {
    if(Y[i]==Sprime[j]){indicatiorFunction3[i]<-1}
    else {indicatiorFunction3[i]<-0}
   }
 Lknprime[j]<-sum(indicatiorFunction3)
 }
# We denote by Vapprox the approximation given in (??).
Vapprox=numeric(L)
for(k in min(Sprime+100):max(Sprime+100))
 {
  ktilde[k]=k-100
  Vapprox[k]=log((n2/n)*(Lkn[z+0]/Lknprime[(ktilde[k])]))
  }
# We plot the graphic that shows the original potential, the approximated potential
by Andreoletti's procedure and the approximation for the states [60,160].
plot(L,V,col="black",lwd=3,type="l",
xlab="State space",ylab=" ",cex.axis=0.85,xlim=range(L))
```

```
lines(L[1:which(L==25)],Vk[1:which(L==25)],
col="azure4",lwd=3)
lines(Lprime+100, Vapprox, col="azure4", lwd=3)
legend(-52, -12, legend=c("Original potential",
"Potential approximated with Andreoletti's procedure",
"Potential approximated by (??)", col=c("black",
"gray", "azure4"), lwd=3:3 ,cex=0.7)
```

Chapter 2

Option pricing and random environments

As a motivation, in Chapter 1 we have studied the case of Sinai's random walk (Sinai, 1982), which is a random walk in random environment with certain conditions. The crux of that model is to observe how each realization of the environment modifies the trajectories of the walk. Such process showed very interesting properties and quite different behaviour of the symmetric random walk, because it evolves with double randomness: the environment and the intrinsic randomness.

Moreover in 1979, Cox, Ross and Rubinstein (Cox et al., 1979) presented one of the most useful methodologies in discrete-time option pricing: the binomial model. As is known, for a European call or put option this procedure yields in the limit the well-known Black-Scholes formula. The binomial model assumes (besides other hypotheses) that the movements of the stock price take one of only two possible values at each time period. In other words, the paths of the movements of the underlying asset are modelled with a usual random walk, which makes it easy to implement numerically.

An interesting variant would be to ask if there exist a model in which a random walk in random environment can be applied naturally in conjunction with the Cox, Ross, and Rubinstein mechanism? We are going in that direction, and we will present a model that combines both ideas. In the articles (Ganikhodjaev, 2013; Xiaoping, 2014) the authors worked a model with a random binary environment i.e., two environments. However, the idea of the Sinai model is totally different, and we find this idea more convenient for financial modelling because it helps to associate directly the environment to the behaviour of the process. For instance, suppose for a moment that we know or infer some information on the environment, then we might be able to use

this information to produce some forecast about the movements of the process. This kind of way of thinking is what makes Sinai's random walk special, and we wish to translate this way of thinking to the financial modelling.

2.1 Binomial model

As we have mentioned in the Introduction, Cox, Ross and Rubinstein (Cox et al., 1979) presented a discrete-time model for valuation of options. The construction of such a model is straightforward, and it makes possible to employ an efficient numerical algorithm. Before proceeding to present our model, we will describe the main ideas of the CRR model.

In the financial markets, a *derivative security* is a contract whose value depends on another asset called the *underlying asset*. Let us focus in the *European call option*, which is a contract where the holder gets the right, but not the obligation to buy the underlying asset (we consider a non-dividend-paying stock) at maturity date T, for a specific price K (it called *strike price*). Let $S(t)$ be the underlying asset price at time t. In this way, we only know $S(0)$. Therefore the payoff for the call option is the $\max(S(T)-K,0)$. The main goal is to give a fair price C that the holder should pay at time $t=0$ for the option.

There are certain hypothesis one needs to assume.

a. There are not transaction costs.

b. The players are allowed to perform the so-called short sells.

c. It is possible to buy or sell any amount of shares.

First, we divide the time T in N discrete periods, $t_k = \frac{kT}{N}$, $k = 0,1,\ldots,N$. The binomial model assumes that the underlying asset at any period of time takes two possible values, it goes up with a multiplicative factor u with probability p, or it goes down with a multiplicative factor d with probability $1-p$, where $0 < d < 1 < u$. We obtain a binomial tree that represents the possible movements of the prices after N periods.

In order to price the call option, we set up a hedging strategy that can help us to achieve the same return over the time. To do this, we construct a portfolio H with two stocks:

i. The risk-free bond $B(t)$ with compounded risk-free interest rate r, and

ii. the underlying asset $S(t)$.

This portfolio must be self-financing and replicate the cash flow of the call option under all possible scenarios. Self-financing means that one is not allowed to withdraw or put more money during the life of the portfolio, but we can just reallocate its composition.

Moreover, it is often desirable that u and d are constant through time and such that $ud = 1$. However the method of replicating portfolio works even if these conditions fail. The extreme case could be when u and d are not the same constant over time and they also depend on the branch of the tree. More exactly, if $S(j,k)$ denotes the price of the underlying asset at the node (j,k), then it follows the movements

$$\begin{array}{ccc} & & S(2j,k+1) = u_j^{(k)}\, S(j,k) \\ & \nearrow^{\alpha_j^{(k)}} & \\ S(j,k) & & \\ & \searrow_{1-\alpha_j^{(k)}} & \\ & & S(2j+1,k+1) = d_j^{(k)}\, S(j,k), \end{array} \tag{2.1.1}$$

where $k = 0,1,\ldots,N-1$ is the period and $j = 0,1,\ldots,(2^k-1)$ is the branch in each period. The price $S(j,k)$ goes up a multiplicative factor $u_j^{(k)}$ with probability $\alpha_j^{(k)}$, or it goes down a multiplicative factor $d_j^{(k)}$ with probability $1-\alpha_j^{(k)}$. It is possible to have a sample space of 2^N paths for the tree of prices. The construction of the self-financing replicating portfolio is described as follows. First of all, let $V(j,k)$ be the value of the portfolio at the node (j,k), then at time N it should hold that

$$V(j,N) = \max(S(j,N)-K,0) \quad \text{for all } j = 0,1,\ldots,2^N-1.$$

Then we adjust the portfolio at each node and we work the backward induction to obtain the price for the replicating portfolio at time t_0. To do this, we must find h and b from the equations

$$\begin{aligned} h\, u_j^{(k)} S(j,k) + b\, R &= V(2j,k+1), \\ h\, d_j^{(k)} S(j,k) + b\, R &= V(2j+1,k+1), \end{aligned}$$

for each $k = N-1,\ldots,1,0$ and $j = 0,1,\ldots,(2^k-1)$, where $R = B(t_1)$ is the bond price at time t_1. Consequently we obtain the solution

$$\begin{aligned} h &= \frac{V(2j,k+1)-V(2j+1,k+1)}{(u_j^{(k)}-d_j^{(k)})S(j,k)}, \\ b &= \frac{V(2j,k+1)(u_j^{(k)}-d_j^{(k)}-1)+V(2j+1,k+1)}{R(u_j^{(k)}-d_j^{(k)})}. \end{aligned} \tag{2.1.2}$$

Therefore, we set up a portfolio of h shares of underlying asset and b dollars in risk-free bond. Thus the value of these operations is

$$V(j,k) = h\,S(j,k) + b. \tag{2.1.3}$$

In order to apply this methodology, a fundamental requirement is that the model should be free of arbitrage. It means that is not possible to construct a trading strategy that begins without money at time t_0, and it has a positive probability of making money at time T. It turns out that such property holds if and only if $d_j^{(k)} < R < u_j^{(k)}$, $k = 0,1,\ldots,N-1$ and $j = 0,1,\ldots,k$. In particular if the factors u and d are constants over time, then $d < R < u$ is a condition of no arbitrage. When we have this condition the price of both financial assets, the call option and the portfolio H, must be equal at time t_0, i.e.,

$$C = V(0,0).$$

Remark 2.1.1. *Notice that it is ok if the tree obtained through binomial model does not join its nodes as in the case of the constant factors u and d. Because each node and its two next branches are a simple representation of binomial model itself.*

2.1.1 Predicament: why it is not possible to use the Sinai model for option pricing?

The methodology described above gives rise to powerful formula to price contingent claims. In the end, if u and d are constant through time and $d = 1/u$, the price of the option is given by

$$C = \frac{1}{e^{rT}} \sum_{i=0}^{N} \binom{N}{i} (p^*)^i (q^*)^{N-i} \max(u^i d^{N-i} S(0) - K, 0), \tag{2.1.4}$$

with $p^* = \frac{R-d}{u-d}$ and $q^* = \frac{u-R}{u-d}$. It is well-know that when $u = e^{\sigma\sqrt{\frac{T}{N}}}$ and $d = 1/u$, where σ denotes the volatility (constant), the price under this model converges to the price given by the Black-Scholes formula. The striking paradigm of this formulation is the conclusion that C is actually an expectation under a new probability measure specified by p*,i.e., $C = \mathrm{E}^*(\frac{G(S_T)}{e^{rT}})$, where $G(S_T) = \max(S_T - k, 0)$ is the payoff function. For a general reference about options one can see (Hull, 2009; Pliska, 1997).

In accordance with the solutions given by (2.1.2), the probabilities $\alpha_j^{(k)}$ described in (2.1.1) does not play any role. Indeed, if we try to use Sinai's model for the movements of the prices with random probabilities $\alpha_j^{(k)}$, the solutions in (2.1.2) are still the

same, i.e., the randomness of the probabilities to go up or down would be obsolete. In other words, it is ok that different investors consider different probabilities to go up or down in the movements of the stock price because that probability p does not play any role in the price for the option.

2.2 The new model

Our goal is to propose a model which is capable to produce a fair price using the binomial CRR scheme, and at the same time that it incorporates the concept of a random environment presented in the market. An effort of this kind was proposed by (Ganikhodjaev, 2013), and then by (Xiaoping, 2014) where they worked a model with a random binary environment, i.e., two environments.

However it turns out that such a model has a different nature than Sinai's walk and than our model. Indeed the main feature of Sinai's walk is that always uses the same previously fixed probability every time it goes through the same site. This means that once the environment is given, the walk moves using a fixed probability α_x at each site x. That is important because this feature is what brings the new phenomena in the random walk. The idea of our proposal is to give a random walk whose upwards or downwards movements are itself random. We could say that for some environments and for some sites the position would move upwards with a greater step than the step going downwards, and for the other sites the other way round.

Let us describe it precisely. Given $\Delta > 0$, we define the lattice $\mathbb{Z}_\Delta$ as

$$\mathbb{Z}_\Delta := \{\ldots, -2\Delta, -\Delta, 0, \Delta, 2\Delta, \ldots\}.$$

We consider $\alpha = \{\alpha_x : x \in \mathbb{Z}_\Delta\}$ a collection of i.i.d random variables

$$\alpha_x = \begin{cases} \Delta & \text{with probability } q, \\ -\Delta & \text{with probability } 1-q, \end{cases} \tag{2.2.1}$$

with $q \in [0,1]$. We call any realization of α a **random environment**. For each α fixed and $x \in \mathbb{Z}_\Delta$, we define the random walk in the random environment α as the Markov chain $\{X_n : n \geqslant 0\}$ with state space $\mathbb{Z}_\Delta$ on a probability space $(\Gamma^\alpha, \mathcal{G}^\alpha, P^\alpha)$, which we are going to describe in subsection 3.2. We consider $X_0 = 0$ and the movements under P^α are the following

$$(X_{n+1} \mid X_n = x) = \begin{cases} x + \alpha_x & \text{with probability } p, \\ x - \beta\alpha_x & \text{with probability } 1-p, \end{cases} \tag{2.2.2}$$

where $\beta \in \mathbb{N}$ and $p \in (0,1)$ is fixed. Notice that β tells us if the jump to the left or to the right is longer or shorter, whereas α_x tells us exactly towards which direction such a jump is indeed longer. Observe that if we consider the parameters $p = 1/2$ and $\beta = 1$ in the description of the model (2.2.2), we recover the symmetric random walk. In Figure 2.1, we can observe the behaviour of the walk X_n each time that visits a state $x \in \mathbb{Z}_\Delta$, given an environment α. We can say that for each $x \in \mathbb{Z}_\Delta$, there exists a random variable α_x. If $\alpha_x = \Delta$, then the particle moves one step of length Δ to the right with probability p or it moves $\beta-$steps of length Δ to the left with probability $1-p$. If the random variable $\alpha_x = -\Delta$, then the particle moves $\beta-$steps of length Δ to the right with probability $1-p$ or it moves one step to the left with probability p.

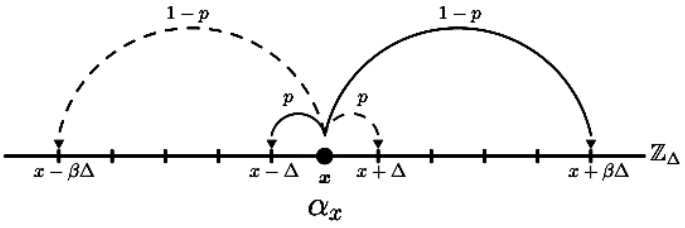

Figure 2.1: Dynamic of a particle given an environment α at any $x \in \mathbb{Z}_\Delta$.

In Figure 2.2, we can see one simulation of each one of the three different models: The symmetric random walk, Sinai's model and our model (all of them with $n = 500,000$ time steps and on a lattice $\mathbb{Z}_\Delta$, $\Delta = 0.05$).

The features of each model are the following:

(i) For the symmetric walk, we consider the parameters that we have mentioned $p = 1/2$ and $\beta = 1$.

(ii) For Sinai's walk, $X_0 = 0$, $P^\alpha(X_{n+1} = x + \Delta \mid X_n = x) = \alpha_x = 1 - P^\alpha(X_{n+1} = x - \Delta \mid X_n = x)$ for $n \geqslant 1$ and $x \in \mathbb{Z}_\Delta$. Here we consider the environment α defined by (1.2.16).

(iii) Lastly, for our model we use the description in (2.2.2) with the parameters $p = 1/2$ and $\beta = 2$.

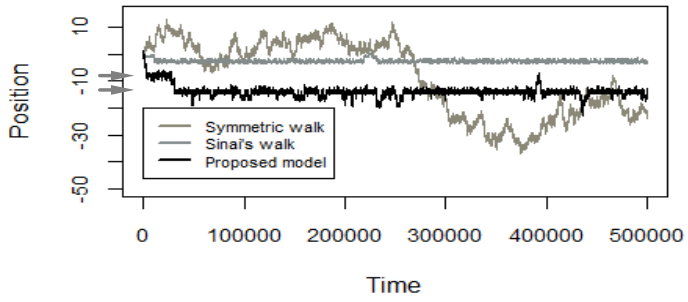

Figure 2.2: Simulation of paths from different random walks.

Notice that our model has a different erratic behaviour than the symmetric walk. In fact such behaviour resembles Sinai's walk, however, its growth is not too slow as the Sinai model.

2.2.1 Connection between the environment and a quenched path

Now we would ask ourselves if there is a tool as the so-called *potential*, that we described above in the Chapter 1 for Sinai's random walk. It turns out that it is possible to have such a tool. To achieve that, we propose a two-sided random walk W' as follows. Given an environment α, let W' be the following function on the state space $\mathbb{Z}_\Delta$:

$$W'(x) = \begin{cases} \sum_{i=\Delta}^{x} \alpha_i & \text{if } x = \Delta, 2\Delta, 3\Delta, \dots \\ 0, & \text{if } x = 0 \\ -\sum_{i=x+\Delta}^{0} \alpha_i, & \text{if } x = -\Delta, -2\Delta, -3\Delta \dots \end{cases} \tag{2.2.3}$$

We discovered numerically that the minima of W' tell us where the walk spends more time. Somehow, W' may play a similar role than W for the Sinai walk.

In Figure 2.2, we have observed that the path of our model stays more time around the states -8 and -14, marked with little arrows on the left side. On the other hand, we can observe in Figure 2.3 that W' has a minimum around of the states -8 (i.e., -160Δ) and -14 (i.e., -280Δ).

Therefore given an environment α, we can study the behaviour of the two-sided random walk W' and, try to infer the behaviour of the random walk X_n in the random environment α.

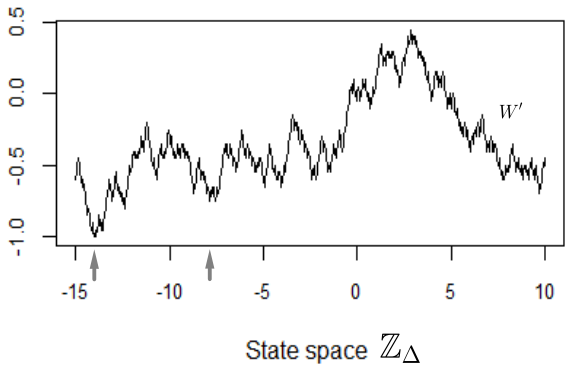

Figure 2.3: Two-sided random walk W' associated to the environment α.

2.2.2 Probability measures on paths

For each environment α fixed, it would be natural to define the quenched and annealed probability in an analogous way as it has been defined for Sinai's walk. However it is important to highlight the fact that the model (2.2.2) associates to each environment α a different space of paths Γ^α. In Subsection 2.3.1 we have an example to visualize this phenomenon.

In fact for our purpose, we will consider the random walk in random environment X_n on a finite time window $T := \{0,1,2,\dots,N\}$ for some $N > 0$. Thereby, we equipped Γ^α with the discrete σ-algebra denoted by $\mathcal{G}^\alpha$. Furthermore it will be sufficient to truncate the sequence of the environment α, because the walk $\{X_n\}_{n\in T}$ visits only a finite number of states. To do this, we define

$$\mathbb{Z}_{2N\Delta} := \{-2N\Delta,\dots,-2\Delta,-\Delta,0,\Delta,2\Delta,\dots,2N\Delta\} \tag{2.2.4}$$

and from now on we take the finite sequence of random variables $\alpha = \{\alpha_x\}_{x\in\mathbb{Z}_{2N\Delta}}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

In accordance with the notation used in Subsection 1.1, we also call $P_x^\alpha(\cdot)$ the *quenched probability* such that $X_0 = x$, E_x^α the expectation with respect to P_x^α and $\mathbb{E}$ the expectation associated to the measure $\mathbb{P}$. This enables us to take the average over the quenched probabilities with respect to the environment distribution to obtain the *annealed probability* denoted by $\mathbf{P}_x$. To define it explicitly, we first consider the space of all trajectories of length N, denoted by $\mathcal{X} = \cup_\alpha \Gamma^\alpha$, and we provide it with the discrete σ-algebra $\sigma(\mathcal{X})$. Therefore we can calculate for any set $A \in \sigma(\mathcal{X})$,

$$\mathbf{P}_x(A) := \int_\Omega P_x^{\alpha(\omega)}(A)\,\mathbb{P}(d\omega) = \mathbb{E}\,P_x^\alpha(A). \tag{2.2.5}$$

Remark 2.2.1. *Our proposed model is not a Markov chain with the annealed measure.*

2.3 Valuation

At this moment we are in position to describe the pricing methodology using our model. We will focus on a call option.

For each fixed environment α, we generate a discrete tree of prices based on the walk X_n given by the model (2.2.2). As it is usually the case, we consider

$$S(n) = S(0)\exp(X_n), \quad n = 0, 1, \ldots, N \tag{2.3.1}$$

as the model for prices. Thus, at any time period $S(n)$ takes two possible values depending the environment α. This tree has a special feature: at each step time the nodes may or may not join into another node, in contrast to the classical binomial model. However, as we have mentioned in Remark 2.1.1, that does not represent a problem for applying the CRR procedure.

In the CRR model we construct a hedging strategy with a self-financing portfolio with two stocks: a risk-free bond $B(t)$ with compounded risk-free interest rate r and the underlying asset $S(t)$ by the model (2.3.1). Then, we look for the portfolio that replicates the payoff of the call option over all possible scenarios. This enables us to compute the value of the call option for that environment α, which is the value of the replicating portfolio. Hereafter we can call it the *quenched price* C_α. In other words, *if we would know the environment that moves market's prices, the quenched price will be the correct price of the option*.

The following lemma helps us to ensure the existence of C_α.

Lemma 2.3.1. *If the following condition is true, then the quenched model (2.3.1) is free arbitrage:*

$$\exp(-\Delta) < R < \exp(\Delta), \tag{2.3.2}$$

where $R = B(\frac{T}{N})$.

Proof. The condition (2.3.2) implies directly

$$\exp(-2\Delta) < \exp(-\Delta) < R < \exp(\Delta) < \exp(2\Delta).$$

This is, for any value that the underlying asset takes in each period with the model (2.3.1)-(2.2.2), it is possible to apply an argument of arbitrage (similarly to the classical binomial model). □

However, when we do not know the random environment that is behind the real prices of the underlying asset, we may find an annealed expectation

$$C = \int_\Omega C_{\alpha(\omega)}\, \mathbb{P}(d\omega), \tag{2.3.3}$$

in order to explore what the price of the option could be.

2.3.1 Numerical exercise

This Subsection summarises the idea of the model through the following exercise. We consider a risk asset such that $S(0) = 100$ dollars. Fix $\Delta = 0.02$ and we generate an environment α. The Figure 2.4 shows the binomial tree associated to α with all possible scenarios of X_n after 5 periods.

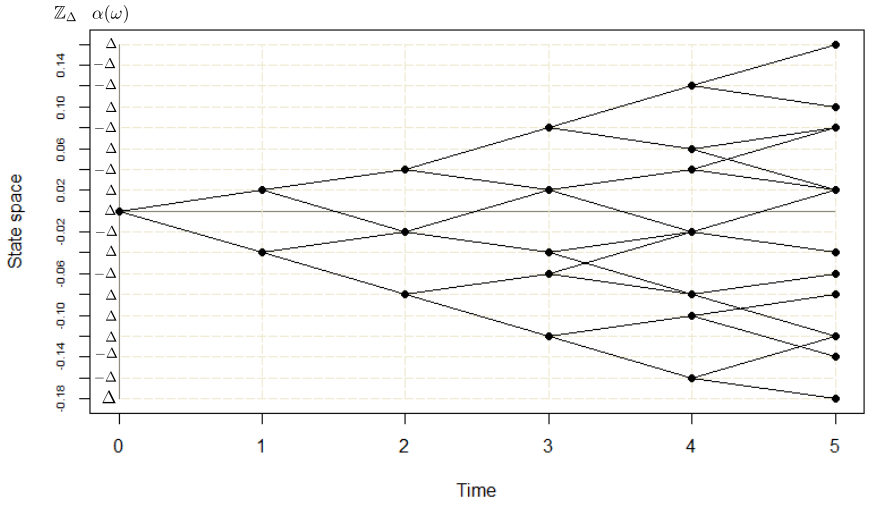

Figure 2.4: Example of a binomial tree given an environment α generated by (2.2.1) with $p = q = 0.5$, which can be observed in the ordinate-axis.

We present in Figure 2.5, the tree of prices using the model (2.3.1). To do this we have considered $N = 5$, $T = 1$, $r = 0.09$, $K = 101$ and $R = 1.018163$, and we construct the replicating portfolio. The value of such portfolio is displayed in Figure 2.6, i.e., the quenched price $C_\alpha = V_\alpha(0,0)$.

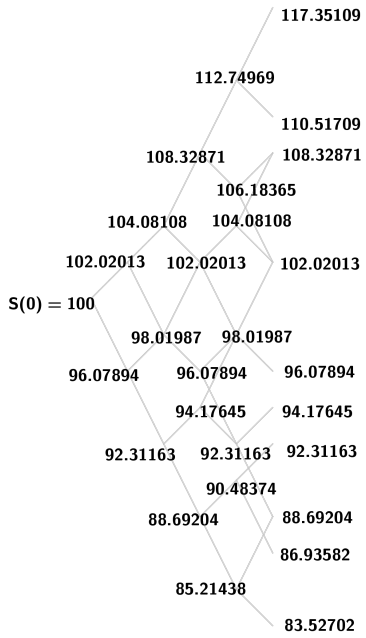

Figure 2.5: Tree of prices associated to the environment α.

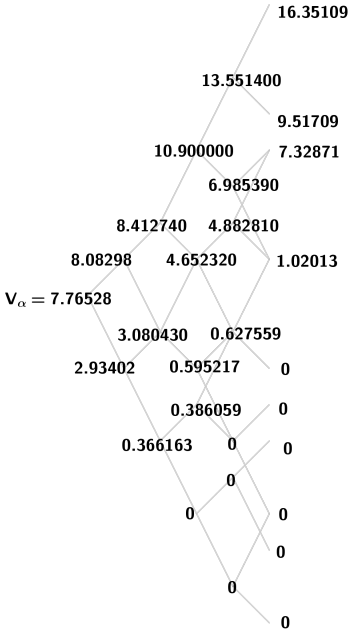

Figure 2.6: Tree of prices for the replicating portfolio.

2.4 Empirical studio

Finally, we show an example to valuate a call option. We have taken the historical prices from the web site (Mexder, 2018) based on the Options Bulletin issued on April 30, 2018 and we consider options on BMV IPC Futures with the information presented in Table 2.1[1].

In Figure 2.7, we simulate three paths of underlying asset using our model (2.3.1)-(2.2.2) with the parameter $\Delta = 0.00347$, which was empirically adjusted. Each one of the paths were generated with its own environment. Also, we can observe the real trajectory of the underlying asset price during the lifetime of the contract.

Let us carry out our valuation. First we obtain from the data the values $T = 0.4$ and $R = 1.000125$. Then to run the algorithm, we consider $N = 310$ discrete periods during the lifetime of the contract and the fixed parameter Δ. The program was run with a sample of 300 random environments. We carried out our program with the

[1]We obtained the value for the underlying asset on April 30, 2018 from the web site (Investing, 2018) and the approximate value for the compounded risk-free interest rate from the web site (Banxico, 2018).

Contract	Call option
Underlying asset	BMV IPC Future
Maturity date	September 21, 2018
Strike price	$K = 57,500$
Initial price	$S_0 = 48358.16$
TIIE Compounded risk-free interest rate	$r = 9.7376\%$

Table 2.1: Information of the contract.

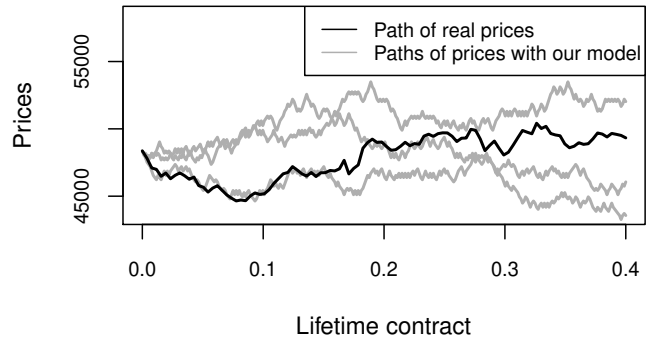

Figure 2.7: Example of paths for the underlying asset with the model (2.3.1)-(2.2.2)

statistical software R-Studio (Core et al., 2013) and we present in the Appendix a pseudo code of the program. The Table 2.2 describes the results obtained, even the obtained price by the usual binomial model with the same data T, r, S_0, K, N-periods and taking the implicit volatility $\sigma = 0.13641151$.

Min.	111.6467
Median.	115.1538
Mean.	115.1544
Max.	118.5428
Current price in the market	115
Price with the usual binomial model	114.4251

Table 2.2: Summary Data Quenched prices.

Moreover, we obtain a standard deviation SD= 1.289028. Informally speaking, one may propose that the theoretical value (2.3.3) belongs to the interval

$$[114.9628, 115.3461]$$

with 99 per cent of confidence. Because applying for example, the Shapiro-Wilk normality test we obtained the p-value $= 0.2858$. In Figure 2.8 we observe the histogram with the result obtained by the simulations taking independents environment and 310 periods for each path.

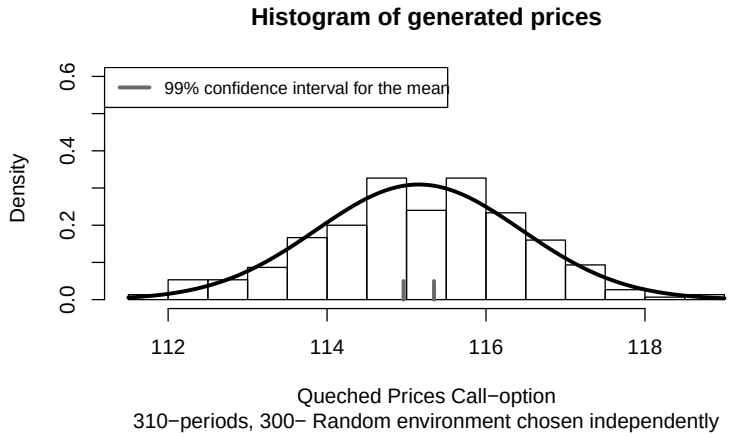

Figure 2.8: Histogram to approximate the theoretical value (2.3.3).

Lastly, we present in Table 2.3 the same experiment with different values for Δ, we show the result of each interval with 95 per cent of confidence for such Δ-values. This is, we can observe the sensitivity to parameter variation Δ in the model 2.2.2.

The simulations took independent environments for each Δ-value and 310 periods for each path as the presented example with $\Delta = 0.00347$. Such task took - days in a computer with the following features: AMD A8-7410 APU processor and 6 GB RAM memory.

2.5 Appendix A2

Here we present the algorithm in *R Studio* to obtain a trajectory of the model (2.2.2) for a given environment.

Δ-values	Interval with 99 per cent of confidence
0.00342	[107.5242, 107.9101]
0.00343	[109.0098, 109.3539]
0.00344	[110.4576, 110.8074]
0.00345	[111.9654, 112.3257]
0.00346	[113.3880, 113.7576]
0.00347	**[114.9628, 115.3461]**
0.00348	[116.5470, 116.9244]
0.00349	[117.9993, 118.4082]
0.00350	[119.7281, 120.0993]
0.00351	[121.2098, 121.5969]
0.00352	[122.7791, 123.1843]

Table 2.3: Summary Data Quenched prices.

Algorithm 4 Simulation of a path of the model (2.2.2).

Step 1. We define the variables.

```
t = number of steps time.
n = number of states on ℤ_Δ that we will label with the
random variables i.i.d (the environment).
delta = the size of the jump for the walk.
```

 # First we fix an environment, which we call it `E`*. It is a sample of size* $2n-1$ *from a data set* {`delta, -delta`}.

```
alphaX<-delta*(2*rbinom((2*n-1),1,0.5)-1)
```

 # We are going to define the function `g` *to associate the vector* `E` *with the state space* $\mathbb{Z}_\Delta$ *restricted to the following set*

$$\{-(n-1)\Delta,\ldots,-2\Delta,-\Delta,0,\Delta,2\Delta,\ldots,(n-1)\Delta\}.$$

```
Dx<-function(x){alphaX[(x/delta)+n]}
```

`X` *denotes the trajectory of the model (2.2.2).*

```
X<-numeric(t+1)
X[1]<-0
aux<-numeric(t)
```

```
for(i in 1:t)
{
aux[i]<-sample(c(Dx(X[i]),-2*Dx(X[i])),1,prob=c(0.5,0.5))
X[i+1]<-X[i]+aux[i]
}
# S denotes the trajectory of the model described by (2.3.1)
S<-numeric(t+1)
 for(i in 1:(t+1))
{
S[i]<-S0*exp(X[i])
}
```

Algorithm 5 Applying the technique of replicating portfolios with random environments

Step 1. Define the local variables.

`Ti = Time of live of the contract.`

`delta = parameter for the model (2.2.2).`

`n = number of the random variables described in` **Step 2.**

`N = Number of periods during the time [0,Ti]`

`S0 = Initial price of the underlying asset.`

`K = Strike price.`

`r = Compounded risk-free interest rate.`

`R = Discount factor in each period.`

`P = Settlement Price.`

 # Let `g` *be the Payoff for the Call option.*

`NR = number of sample size of random environments`

Step 2. We run the algorithm to obtain a quenched price given each one of `NR`-different environments. We will write each price in a coordinate in a vector that we will denote by `Pr`.

```
Pr<-numeric(NR)
for (y in 1:NR) {
```

`E` *represents a random environment. As follows:* `2*n-1` *samples are chosen independently from Bernoulli variables which it chooses one value between* `delta` *or* `-delta` *with probability 1/2.*

Remember we are going to generate random variables from `-(n-1)` *until* `(n-1)` *instead of* $\mathbb{Z}$ = {*...,-n,-(n-1),...,0,...,n-1,n,..*}

```
 E<-delta*(2*rbinom((2*n-1),1,0.5)-1)
```

`Dx` *assigns to each state x the random variable correspondent from the environment* `E`.

```
Dx<-function(x)E[(x/delta)+n]
```

Let `C` *be the matrix which save in each column the values generated in the tree step by step.*

We begin with a square matrix 2x2, `C[1,1]`=X_0, `C[1,2]` *and* `C[2,2]` *are the two possible values after one-step,* `C=0` *otherwise.*

```
C<-matrix( 0, nrow = 2, ncol=2)
C[1,1]=0
C[1,2]=Dx(C[1,1])
C[2,2]=-2*Dx(C[1,1])
```

Then the following algorithm updates its last column.

We take `l` *a vector where the i-th coordinate represents the number of rows of the matrix* `C` *in the i-th column.*

```
l<-numeric(N+1)
l[1]=1
l[2]=2
for(j in 3:(N+1))
{
```

`M` *is the column vector, which coordinates are the two possible values of the two last possible generated values by the model given the environment* `E`.

```
M=matrix(0, nrow = (2*(l[(j-1)])), ncol= 1)
for(i in 1:(l[j-1]))
 {
M[(2*i-1),1]=round(C[i,(j-1)]+Dx(C[i,(j-1)]),9)
M[(2*i),1]=round(C[i,(j-1)]-2*Dx(C[i,(j-1)]),9)
 }
```

`B` *only leaves the values that are not repeated in the vector* `M`.

```
B=unique(M)
l[j]=length(B)
di=(l[j])-length(C[,(j-1)])
```

the variable `di` *tell us how many zeros we need to fill the matrix* `C`.

```
if(di>0)
```

```
{
Z=matrix(0, nrow =((l[j])-length(C[,(j-1)])),
ncol=(j-1))
C=cbind(rbind(C,Z),B)}
if(di==0){C=cbind(C,B)}
if(di<0)
  {
Z=matrix(0,nrow=(length(C[,(j-1)])-l[j]), ncol=1)
C=cbind(C,rbind(B,Z))
  }
# rbind: To join two datasets vertically.
# cbind: To join two datasets horizontally.
}
```

Step 3. Now we are going to apply the usual model $S(t) = So * exp(X(t))$, i.e., we are going to obtain the price tree given the environment `E`.

```
S=matrix(0,nrow = nrow(C), ncol=(N+1))
S[1,1]=S0
for(j in 2:(N+1))
 {
 for(i in 1:(l[j])){
 S[i,j]=S0*exp(C[i,j]) }
 }
```

Step 4. `Q` is the matrix which (i,j)-coordinate is the value of the call-option using replicating portfolios.

`ncol=N+1`, *because it is a tree with N-periods, while* `nrow=l[N+1]`, *it is the number the rows at the end of* `N`*-steps.*

```
Q<-matrix(0, nrow =l[N+1], ncol=(N+1))
for(i in 1:(l[N+1])){
Q[i,N+1]=g(S[i,N+1])}
```

we know that the coordinates in the last column are equal to their intrinsic values.

```
for(j in N:1){
```

Option value calculated at each preceding node.

```
for(i in 1:(l[j])){
Ma=matrix(0, nrow = 2, ncol=2)
Ma[1,1]=R
```

```
Ma[2,1]=R
Ma[1,2]=S0*exp(C[i,j]+Dx(C[i,j]))
Ma[2,2]=S0*exp(C[i,j]-2*Dx(C[i,j]))
# Ma is the matrix to solve the system of equations in each step. We remember that
the model of self-financing and replicating portfolios work backwards through
each period.
a=which(signif(S[,(j+1)],digits=7)==signif(S0*
exp(C[i,j]+Dx(C[i,j])), digits = 7))
b=which(signif(S[,(j+1)],digits=7)==signif(S0*
exp(C[i,j]-2*Dx(C[i,j])), digits = 7))
# a,b looking for the coordinates produced by the model in each step.
Ve=matrix(0, nrow = 2, ncol=1)
Ve[1,1]=Q[a,(j+1)]
Ve[2,1]=Q[b,(j+1)]
# Ve are the prices for the option in the binomial tree.
W<-solve(Ma,Ve)
# W are the solution for the sistem of equations Ma x = Ve.
Q[i,j]=W[1,1]+W[2,1]*S0*exp(C[i,j])
# Q[i,j] is the value of the portfolio in that node of the tree.
# We need Wi[1,1]-shares of riskless bond and Wi[2,1]-shares of the under-
lying asset for buy or sell depending of the sign.
}
}
# Finally we fill the vector of price for each environment
Pr[y]=Q[1,1]
# Q[1,1] is price for the Call option.
}
```

Chapter 3

More about the model

3.1 Behaviour of the model

We consider the model (2.2.2) introduced in Chapter 2. We recall that we have defined it as a Markov chain for a fixed environment α. It is worth investigating, what kind of Markov chain is it? Given the definition 1.2.1, we will use the next lemma.

Lemma 3.1.1. *(Durrett, 1999)[p. 12-20]*
Let $\{Y_0, Y_2, \dots\}$ be a Markov chain with state space S. The following statements are equivalent.

i. *The state x is transient.*

ii. $P_x(T_x < \infty) < 1$,

iii. $P_x(Y_n = x \text{ for infinitely many } n) = 0$,

iv. $\sum_{n=1}^{\infty} P_x(Y_n = x) < \infty$.

We will explore some unrealistic but informative cases. In what follows, we focus in a particular case of the environment in the model (2.2.2). Given $\Delta > 0$, we consider $\alpha = \{\alpha_x : x \in \mathbb{Z}_\Delta\}$ such that

$$\alpha_x = \Delta \text{ for all } x. \tag{3.1.1}$$

Moreover, we consider the parameters $p = 0.5$ and $\beta = 2$ as before. The random walk X_n in the environment α is taking the following movements:

$$(X_{n+1} \mid X_n = x) = \begin{cases} x + \Delta & \text{with probability } 0.5, \\ x - 2\Delta & \text{with probability } 0.5. \end{cases} \tag{3.1.2}$$

Easily we can see that for an environment α as (3.1.1) all trajectories after n-steps are equally probable. Furthermore the random walk X_n in accordance with (3.1.2) visits all states with probability strictly positive, i.e., all states in $\mathbb{Z}_\Delta$ are accessible. So that, the Markov chain is irreducible.

To illustrate that random walk, in Figure 3.1 we can observe the superposition of all trajectories after 7-steps and a particular trajectory in blue color.

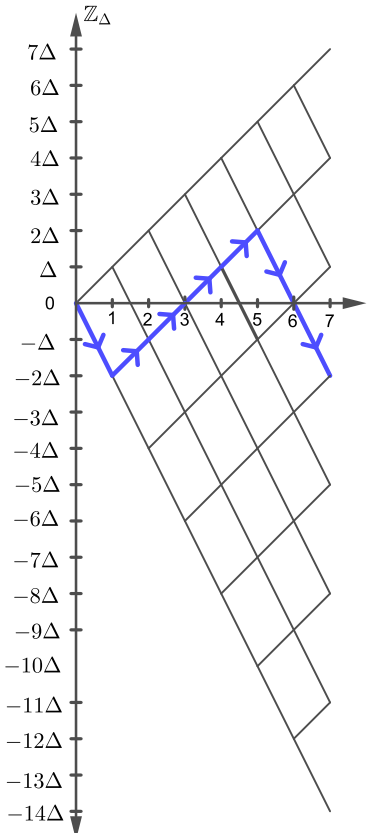

Figure 3.1: Network generated given the environment α described in (3.1.1).

Proposition 3.1.2. *Given the environment α as (3.1.1), the random walk X_n is transient.*

Proof. According to the Lemma 3.1.1 because the Markov chain is irreducible, it is enough to show that $\sum_{n=1}^{\infty} P_x^\alpha(X_n = x) < \infty$ only for one state, for instance $x = 0$.

First we observe that for $n > 0$

$$P_0^\alpha(X_n = 0) = \begin{cases} 0 & if \quad n \neq 3k \ \ (k \geqslant 1), \\ \\ > 0 & if \quad n = 3k \ \ (k \geqslant 1). \end{cases}$$

Then

$$\sum_{n=1}^{\infty} P_0^{\alpha}(X_n = 0) = \sum_{k=1}^{\infty} P_0^{\alpha}(X_{3k} = 0).$$

We calculate explicitly the probability of return in $3k$-steps as

$$P_0^{\alpha}(X_{3k} = 0) = \frac{\binom{3k}{k}}{2^{3k}},$$

for $k = 1, 2, 3, \ldots$. To calculate this formula, we can think the walk as an arrangement, i.e., for each step to **down**, we need to **up** 2-steps. Thus to return to the origin after 3-steps, we can order an arrange (**up**,**down**,**down**) by means of combinations $\binom{3}{1}$. In general after $3k$-steps, exactly $\binom{3k}{k}$ trajectories return to the origin of a total of 2^{3k}.

By using the Stirling formula[1] one can calculate

$$\begin{aligned}
\sum_{k=1}^{\infty} P_0^{\alpha}(X_{3k} = 0) &= \sum_{k=1}^{\infty} \frac{\binom{3k}{k}}{2^{3k}} \\
&= \sum_{k=1}^{\infty} \frac{(3k)!}{k!\,(2k)!\,2^{3k}} \\
&= \frac{\sqrt{2\pi(3k)}\,(3k)^{3k}\,e^{-3k}}{\sqrt{2\pi k}\,k^k e^{-k}\,\sqrt{2\pi(2k)}\,(2k)^{2k}\,e^{-2k}\,2^{3k}} \\
&= \frac{\sqrt{3}}{\sqrt{4\pi}} \sum_{k=1}^{\infty} \frac{3^{3k}}{\sqrt{k}\,2^{5k}}. \qquad (3.1.3)
\end{aligned}$$

By the ratio test, we calculate the limit of the following quotient

$$\frac{\frac{3^{3(k+1)}}{2^{5(k+1)}\sqrt{k+1}}}{\frac{3^{3k}}{2^{5k}\sqrt{k}}} = \frac{27}{32}\frac{\sqrt{k}}{\sqrt{k+1}} \to \frac{27}{32} < 1, \quad \text{as } k \to \infty.$$

Therefore the serie given by (3.1.3) converges and in consequence, the state $x = 0$ is transient under the environment α given by (3.1.1). □

Furthermore for that environment α, we can obtain more information about the random walk X_n.

Lemma 3.1.3. *Given the environment α as (3.1.1), then*

$$\lim_{n\to\infty} X_n = -\infty \quad a.s.$$

[1] Stirling formula: $n! \approx \sqrt{2\pi n} n^n e^{-n}$ for n large.

Proof. We know the possible movements of the random walk X_n in accordance with (3.1.2). Then by the strong law of large numbers,

$$\begin{aligned}\lim_{n\to\infty}\frac{X_n}{n} &= \mathbb{E}(X_1) \quad \text{a.s.}\\ &= \frac{1}{2}(\Delta)+\frac{1}{2}(-2\Delta)\\ &< 0.\end{aligned}$$

Since any realization of $X_n(\omega)$ is a sequence of real numbers. We denote $\lim_{n\to\infty}\dfrac{X_n(\omega)}{n} = b < 0$, b not necessarily finite. This is, given $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$

$$\left|\frac{X_n(\omega)}{n} - b\right| < \varepsilon.$$

In other words for all $n \geqslant N$, it follows that

$$-\varepsilon < \frac{X_n(\omega)}{n} - b < \varepsilon,$$

$$n(b-\varepsilon) < X_n(\omega) < n(b+\varepsilon).$$

By Archimedean property we can find ε such that $\varepsilon < |b|$ and so $b+\varepsilon < 0$.
Hence

$$\lim_{n\to\infty} X_n = -\infty \text{ a.s.}$$

□

Opposite to the previous situation, if we assume an environment $\alpha = \{\alpha_x : x \in \mathbb{Z}_\Delta\}$ such that

$$\alpha_x = -\Delta \quad \text{for all } x, \tag{3.1.4}$$

and we take the parameters $p = 0.5$ and $\beta = 2$. The model of the random walk has the following displacement

$$(X_{n+1} \mid X_n = x) = \begin{cases} x-\Delta & \text{with probability } 0.5,\\ x+2\Delta & \text{with probability } 0.5.\end{cases}$$

Using a similar argument as before, it follows the next analogous results.

Proposition 3.1.4. *The random walk under the environment α as (3.1.4) is also a transient Markov chain.*

Lemma 3.1.5. *Given the environment α described in (3.1.4), then*

$$\lim_{n\to\infty} X_n = \infty \quad a.s.$$

Summary

For the two particular environments (3.1.1) and (3.1.4) that we have studied. We obtained information about transitivity and the behaviour about a random walk in such environments. It is interesting to know that for the environment (3.1.1) a particle there will go to $-\infty$ a.s., while a particle given the environment (3.1.4) will go to ∞.

Why do not we calculate explicitly the probability to return to the origin given the environment (2.2.1)? Because each environment changes the network, we can not count easily the number of paths that return to the origin after n-steps. On the other hand, a used technique as in Sinai's walk about the probability to return to the origin is complicated because to establish a difference equation with the model (2.2.2) the coefficients are constant given an environment, however, the jump (grade of the equation) is random.

It is a further work to find a result about what conditions on the environment allow us to classify the behaviour of the random walk for the model (2.2.2).

3.2 Connecting the environment and the local time

As a continuation of the section 2.2.1, we remember the definition of the function of the environment

$$W'(x) = \begin{cases} \sum_{i=\Delta}^{x} \alpha_i & \text{if } x = \Delta, 2\Delta, 3\Delta, \dots \\ 0, & \text{if } x = 0 \\ -\sum_{i=x+\Delta}^{0} \alpha_i, & \text{if } x = -\Delta, -2\Delta, -3\Delta \dots \end{cases}$$

It is our purpose to show as this function W' has a connection between the environment and the behaviour of the random walk of the model (2.2.2), because we realized that the walk spends a lot of local time in the minima of the function W'.

We begin analysing about what happens when W' decreases and then, what happens when W' increasing. It is worth to see how the changes of direction of the function W' represent changes in the environment due to the extrinsic randomness.

First we assume that W' decreasing i.e., $W'(x-\Delta) > W'(x)$ for $x \in \mathbb{Z}_\Delta$.
For $x > 0$, if we assume $W'(x-\Delta) > W'(x)$ then

$$\begin{aligned} \sum_{i=\Delta}^{x-\Delta} \alpha_i &> \sum_{i=\Delta}^{x} \alpha_i \\ &= \sum_{i=\Delta}^{x-\Delta} \alpha_i + \alpha_x, \end{aligned}$$

which implies by the definition 2.2.1 of α_x,

$$0 > \alpha_x \Leftrightarrow \alpha_x = -\Delta. \tag{3.2.1}$$

Whereas for $x < 0$, if $W'(x-\Delta) > W'(x)$ then

$$\begin{aligned} -\sum_{i=x}^{0} \alpha_i &> -\sum_{i=x+\Delta}^{0} \alpha_i \\ -\alpha_x - \sum_{i=x+\Delta}^{0} \alpha_i &> -\sum_{i=x+\Delta}^{0} \alpha_i, \end{aligned}$$

that is

$$-\alpha_x > 0 \Leftrightarrow \alpha_x < 0 \Leftrightarrow \alpha_x = -\Delta. \tag{3.2.2}$$

Lastly, if $x = 0$,

$$W'(0-\Delta) > W'(0) \Rightarrow -\alpha_0 > 0,$$

i.e., $\alpha_x = -\Delta$.

That is for any $x \in \mathbb{Z}_\Delta$, the knowledge about increments or decreases in the function W' gives us information on which values take the environment in each state x.

Remark 3.2.1. *Assume the model (2.2.2) given an environment α described in (2.2.3). Let W' be the function on the environment α and $x, y \in \mathbb{Z}_\Delta$, with $x < y$. The following statement are true.*

a) W' is strictly decreasing in $[x,y] \subset \mathbb{Z}_\Delta$ if and only if $\alpha_i = -\Delta$ for all $i \in [x,y]$.

b) W' is strictly increasing in $[x,y] \subset \mathbb{Z}_\Delta$ if and only if $\alpha_i = \Delta$ for all $i \in [x,y]$.

We can explain it as: if an environment α shows a sequence of states x_k where their random variables associated take the values $\alpha_{x_k} = -\Delta$ and subsequently it shows other sequence of states y_k where their random variable associated take the values $\alpha_{y_k} = \Delta$, then the function of the environment W' changes from be decreasing to increasing.

In other words, the changes of the function W generate a phenomenon similar to the valleys (for the Sinai's random walk), because with the parameters that we have considered $p = 0.5$, $q = 0.5$ and $\beta = 2$, we know that all paths are equally likely. Therefore we can see in Figure ref mesh3 as a large part of the trajectories is concentrated around the state where it occurs the minimum of the function W' (that is, in the change of decreasing to increasing). This yields that when a random walk passes through a minimum, the environment makes the walk frequently touches those points around the minimum.

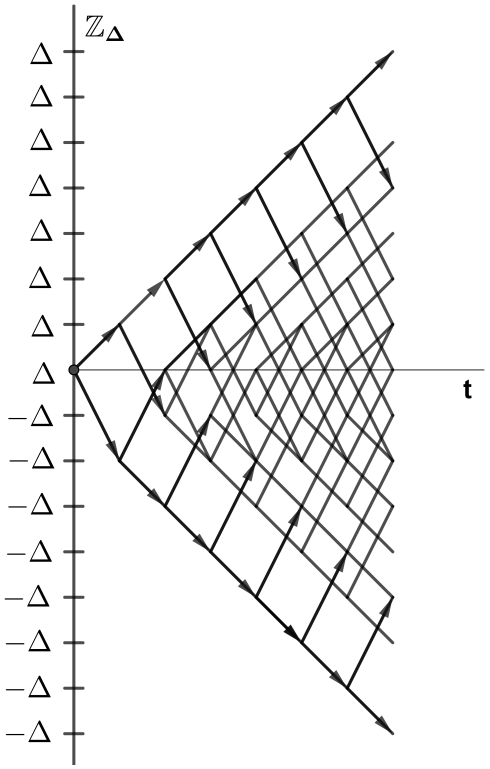

Figure 3.2: Network generated by the sequence α_{x_k} followed by the sequence α_{y_k}.

One may consider the Figure 3.3 as a trajectory of the model (2.2.2) in a time window $[0, 500000]$, where the parameters are $\Delta = 0.05$, $p = 0.5$, $q = 0.5$ and $\beta = 2$. At first sight in that Figure, we can observe that the walk visits frequently two sites in the state space $\mathbb{Z}_\Delta$, around the site $-5 = -100\Delta$ and $13 = 260\Delta$.

As we have already pointed out the function W' proposed in (2.2.3) is a good tool that help us to identify the sites more visited by the random walk with the model (2.2.2). The simulation presented in Figure 3.4 shows that there is a minimum around of the states $-5 = -100\Delta$ and $13 = 260\Delta$.

According to the Definition 1.4.1 about the local time, we show in Figure 3.5 the local time of the trajectory given in Figure 3.3 by means of a program in R (Core et al., 2013). It is possible to observe that precisely the state $13.35 = 267\Delta$ is the most visited state, with exactly 16764 visits of the random walk with the model (2.2.2). Which coincides with the information analysed through the function W'.

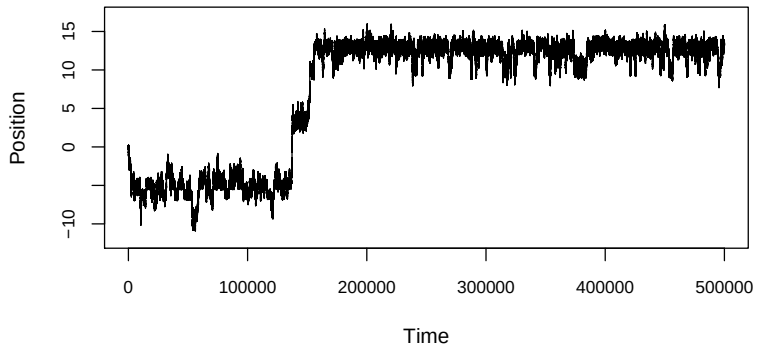

Figure 3.3: Dynamic of a particle given an environment α as (2.2.1).

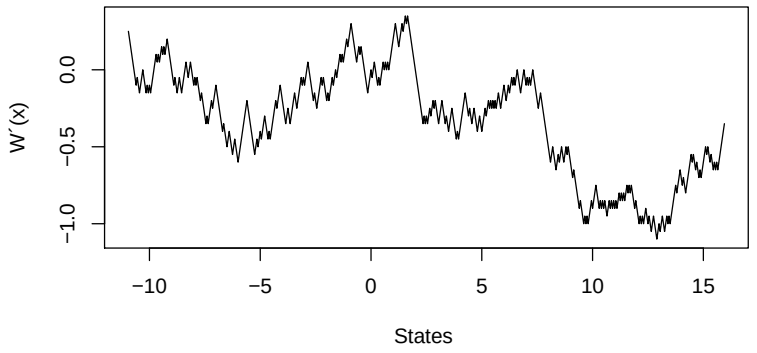

Figure 3.4: Function W' associated to the environment α, which was generated the trajectory of the Figure 3.3.

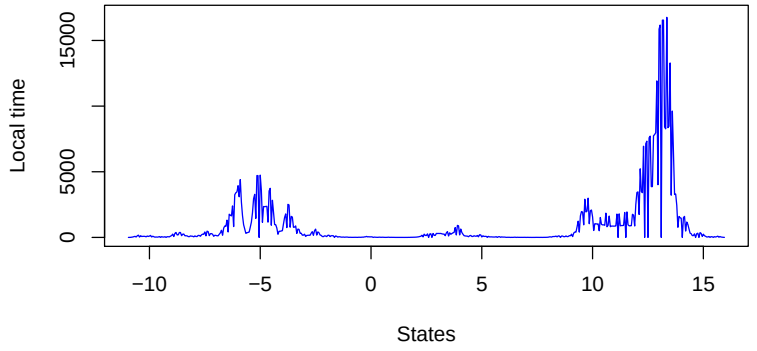

Figure 3.5: Local time associated to the trajectory showed in Figure 3.3

3.3 Recovering the environment

For the study of the model (2.2.2) with $p \in [0,1]$, i.e.,

$$(X_{n+1} \mid X_n = x) = \begin{cases} x + \alpha_x & \text{with probability } p, \\ x - \beta\alpha_x & \text{with probability } 1-p, \end{cases}$$

we want to obtain relevant information about the environment $\alpha = \{\alpha_x : x \in \mathbb{Z}_\Delta\}$. In the same way that the analysis of Sinai's random walk, if we know the environment, we can know with high probability the behaviour of the trajectories of the random walk given that environment.

It is not surprising that the described method in (1.4.6) does not work to recover the environment of the model (2.2.2), because its features are different of the Sinai model in spite of both are random walks in random environment.

In Figure 3.6 we can see a trajectory of the model (2.2.2) with the following parameters: $\Delta = 1$, $p = 0.5$, $q = 0.5$ and $\beta = 2$. However we have mentioned that the method to recover the potential (1.4.6) is far from to be a good estimation of the environment as we can appreciate in Figure 3.7. It is not enough to consider the logarithm of the local time of the trajectory, but maybe under certain modification we can improve that approximation.

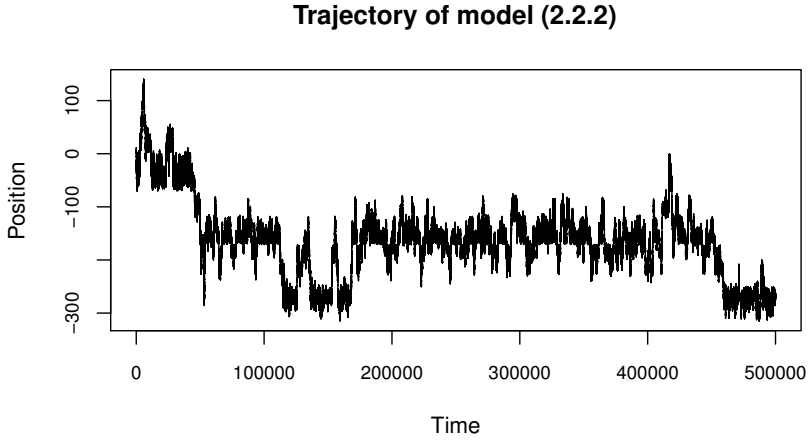

Figure 3.6: Trajectory of the model (2.2.2) with parameters $\Delta = 1$, $p = 0.5$ y $q = 0.5$

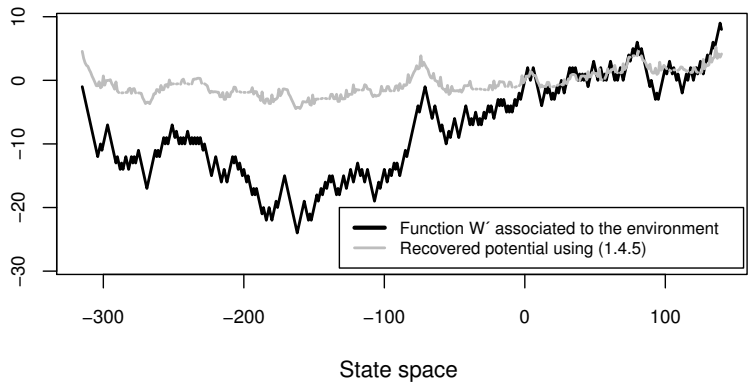

Figure 3.7: Function W' associated to the environment and the method (1.4.6) applied to the trajectory in Figure 3.6.

When we present in Chapter 2 the Figure 2.2, we observe that the order of the model (2.2.2) is bigger than the order of the Sinai walk i.e., $(\log n)^2$ and smaller than the order of the symmetric random walk i.e., $n^{\frac{1}{2}}$. Hence we think that it could be useful the following numerical modification to recover information about the environment.

Remember the methodology described in (1.4.6) for the Sinai walk, that is

$$V(k) \approx \log(\mathcal{L}(0,n)) - \log(\mathcal{L}(k,n)),$$

which we know that it is a good approximation of the potential if the smallest favourite site in absolute value is near to the origin. Instead of such approximation, we propose for the model (2.2.2) to apply the local time to recover a sketch of the environment in the following manner: For each $k \in \mathbb{Z}_{\Delta}$, we take

$$W'(k) \approx (\log \mathcal{L}(0,n))^{3/2} - (\log \mathcal{L}(k,n))^{3/2}. \tag{3.3.1}$$

We have run the simulation in R (Core et al., 2013) applying the local time with the proposal given in (3.3.1) to the trajectory given in Figure 3.6 and we show the result in Figure 3.8. The recovered potential is not good for all states, however it could provide us more information than the obtained in Figure 3.7.

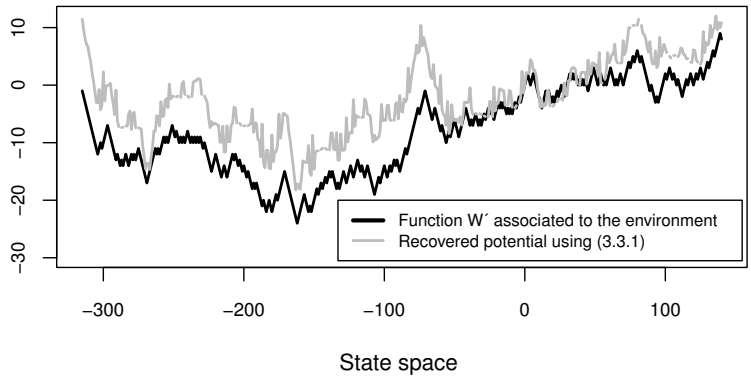

Figure 3.8: Function W' associated to the environment and a numeric proposal given by (3.3.1).

When we apply this procedure we can deduce that if we consider a random walk in that environment, the walk will spend a lot of time in the states round $160 = 160\Delta$, and it is very likely than the walk visit the states round $-270 = -270\Delta$ and also spends a lot of time. We could hope certain behaviour of any walk in that environment.

We remember that originally we proposed the model (2.2.2) with the idea of include the sense of *random environment* in an applied problem of valuation of options (Chapter 2, section 2.3). For that we called *quenched price* C_α to the price of the portfolio that replicates the payoff of the call option over all possible scenarios given the environment α. However, we clarify that since we do not know the environment behind the real prices of the underlying asset, our idea was to propose an interval of prices to make a weighting of quenched prices, as we did it in the numerical example of the section 2.3.1.

We let open the following issue: *if we would know the environment that moves the market's prices, the quenched price will be the correct price of the option*. Therefore our interest in the search of a mechanism to recover the potential and then, the environment, if we assume know a trajectory of the random walk in a random environment as the model (2.2.2).

3.4 Appendix A3

We give the pseudo code for the applications of the Chapter 3.

Algorithm 6 Simulation of the proposal described in (3.3.1) to recover information about the environment.

Step 1. We generated a trajectory `X` with the model (2.2.2), which we assume in accordance with the same parameters described in Appendix A2, i.e., the values `t`, `n`, `delta`, the variables `alphaX` and the function `Dx`.

```
X<-numeric(t+1)
X[1]<-0
aux<-numeric(t)
for(i in 1:t)
{
aux[i]<-sample(c(Dx(X[i]),-2*Dx(X[i])),1,prob=c(0.5,0.5))
X[i+1]<-X[i]+aux[i]
}
```

Step 2. We calculate the function of the environment W' (denote by `Wprime`) associated to `alphaX`.

```
M=max(Y)
m=min(Y)
S<-c(m:M)
# S denotes the visited states by the random walk X.
L=length(S)
Wprime=numeric(L)
nz<-which(S==0)
if(M>0) {
c1=numeric(M/delta)
for(i in 1:(M/delta)){
c1[i]=i*delta
}
Wprime[nz]=0
for(i in c1){
suma1<-0
for(j in seq(delta,i,by=delta)){
suma1<-(suma1+alphaX[g(j)])
}
Wprime[(i/delta)+nz]<-suma1
}}
if(m<0){
```

```
c2=numeric(abs(m/delta))
for(i in 1:abs(m/delta))
{
c2[i]=m+(i-1)*delta
}
for(i in c2){
suma2<-0
for(j in seq((i+delta),0,by=delta)){
suma2<-(suma2-alphaX[g(j)])
}
V[(i/delta)+nz]<-suma2
}}
```

Step 3. We count the local time for each state `k∈S`.

```
TLoc<-numeric(L)
for(j in 1:L){
indicatiorFunction<-numeric(length(c(2:(t+1))))
for(i in 2:(t+1)){
if(X[i]==S[j]){indicatiorFunction[i]<-1}
else{indicatiorFunction[i]<-0}
}
TLoc[j]<-sum(indicatiorFunction)
}
```

 # We apply the numeric proposal given by (3.3.1)

```
Wk=numeric(L)
for(i in 1:L){
Wk[i]=(log(TLoc[which(L==0)]))^(1.5)-(log(TLoc[i]))^(1.5)
}
```

 # We plot the graphics of the approximation `Wk` *and the original function* `Wprima`

```
plot(L,V,col="black",lwd=2,type="l", ylab="",
xlab="State space ", cex.axis=0.85,xlim=range(L),
ylim=range(c(-30:12))), lines(L,Vk, col="gray", lwd=2)
```

Conclusions & further research

Chapter 1. In the first part of this work we study the well-known random walk in a random environment, the Sinai walk. We study its properties, its behaviour and the essential differences compared with the symmetric random walk. Also we define the so-called *potential* associated to an environment, this allowed us to stablish the relation between the local time of the random walk and the minima of the potential.

We described the procedure of Andreoletti to recover the potential given a trajectory of the Sinai walk using only the local time. We explain why that algorithm is good when the site k_n^* (the smallest favourite site in absolute value) is near to the origin, and we propose one idea based on the algorithm of Andreoletti to improve the approximation for the case when k_n^* is not near the origin.

We explain possible scenarios of a trajectory of Sinai's walk and we comment if Andreoletti's procedure is recommended to recover the environment or how we can apply our proposal to recover the environment.

Chapter 2. In this chapter, we propose a new model to incorporate the existence of a random environment in the financial markets. Such a model allows to price contingent claims using the classical CRR procedure. Our aim is that this new model can bring new perspective for financial modelling and pricing.

To the best of our knowledge, we are proposing the first ideas and the possibility to merge the CRR procedure and the concept of the random environment.

A number of open questions and issues for future research are given as follows.

1. There might be different statistical procedures to estimate the parameter Δ, p, q and β.

2. Notice that the possible price of a contingent claim depends on whether one has more information of the environment (a quenched price) or one may not have any information at all (as in the expectation given by (2.3.3)). Actually, one may try to run a procedure to infer information of the environment based on the observations of the prices. Such issue was addressed for Sinai's model in

(Andreoletti, 2011).

In Chapter 3. Here we study the new model (2.2.2) proposed in Chapter 2. We analyse two unlikely, but informative environments to describe and understand the behaviour of the random walk in that environments, and we show as for these instances that the walk is transient.

Also we propose a function W' associated to the environment, which works in a similar way to the potential for the Sinai random walk. Given a trajectory with the model (2.2.2), the local time in a finite time window is large in the sites where it occurs the minima of the function W'.

In addition given a trajectory with the model (2.2.2), and we have study the algorithm of Andreoletti. By comparing the order of the Sinai walk and the symmetric random walk, we propose a numerical procedure to approximate the environment associated.

It is a completely open problem to study the existence of an invariance principle of our model, such as the Donsker theorem for the Brownian motion.

Bibliography

Andreoletti, P. (2011). On the estimation of the potential of Sinai's RWRE. *Brazilian Journal of Probability and Statistics*, 25(2):121–144.

Banxico (2018). Banco de México. `http://www.banxico.org.mx/`.

Brox, T. (1986). A one-dimensional diffusion process in a Wiener medium. *The Annals of Probability*, 118(4):1206–1218.

Core, T. R. et al. (2013). R: A language and environment for statistical computing.

Cox, J. C., Ross, S. A., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of financial Economics*, 7(3):229–263.

Durrett, R. (1999). *Essentials of stochastic processes*, volume 1. New York: Springer.

Durrett, R. (2019). *Probability: theory and examples*, volume 49. Cambridge university press.

Einmahl, U. (2009). A new strong invariance principle for sums of independent random vectors. *Journal of Mathematical Sciences*, 163(4):311–327.

Ganikhodjaev, N & Bayram, K. (2013). Random Cox-Ross-Rubinstein model and plain vanilla options. *World Applied Sciences Journal*, 21(1):84–87.

Golosov, A. O. (1986). On limiting distributions for a random walk in a critical one-dimensional random environment. *Russian Mathematical Surveys*, 41(2):199–200.

Hughes, B. D. (1996). *Random walks and random environments*, volume 2. Oxford University Press.

Hull, J. C. (2009). *Options futures and other derivatives*. Prentice-Hall International. 7th ed.

Investing (2018). Options Bulletin on BMV IPC Futures. https://mx.investing.com/indices/ipc-historical-data.

Kesten, H. (1986). The limit distribution of Sinai's random walk in random environment. *Physica A: Statistical Mechanics and its Applications*, 138(1-2):299–309.

Mexder (2018). Mercado Mexicano de Derivados. http://www.mexder.com.mx.

Mörters, P. and Peres, Y. (2010). *Brownian motion*, volume 30. Cambridge University Press.

Peres, Y. et al. (2001). An invitation to sample paths of brownian motion. *Lecture notes (unpublished)*.

Pliska, S. R. (1997). *Introduction to Mathematical Finance: Discrete Time Models*. Blackwell Publishing, New York.

Révész, P. (2005). *Random walk in random and non-random environments*. World Scientific.

Sinai, Y. G. (1982). The limiting behavior of a one-dimensional random walk in a random medium. *Theory of Probability & Its Applications*, 27(2):256–268.

Solomon, F. (1975). Random walks in a random environment. *The Annals of Probability*, 3:1–31.

Xiaoping, Hu & Jie, C. (2014). Randomized binomial tree and pricing of american-style options. *Mathematical Problems in Engineering*, 2014.