UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Poincaré recurrence times in stochastic mixing processes

Vitor Gustavo de Amorim

Tese de Doutorado do Programa Interinstitucional de Pós-Graduação em Estatística (PIPGEs)

ICMC USP SÃO CARLOS

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: ______________________

Vitor Gustavo de Amorim

Poincaré recurrence times in stochastic mixing processes

Thesis submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP and to the Department of Statistics – DEs-UFSCar – in accordance with the requirements of the Statistics Interagency Graduate Program, for the degree of Doctor in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Alexsandro Giacomo Grimbert Gallo
Co-advisor: Prof. Dr. Miguel Natalio Abadi

USP – São Carlos
January 2022

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

A524p

Amorim, Vitor
Poincaré recurrence times in stochastic mixing processes / Vitor Amorim; orientador Alexsandro Gallo; coorientador Miguel Abadi. -- São Carlos, 2022.
115 p.

Tese (Doutorado - Programa Interinstitucional de Pós-graduação em Estatística) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2022.

1. Hitting and return times. 2. Exponential approximation. 3. Mixing processes. 4. Waiting time spectrum. 5. Potential well. I. Gallo, Alexsandro , orient. II. Abadi, Miguel, coorient. III. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

Vitor Gustavo de Amorim

Tempos de recorrência de Poincaré em processos estocásticos com mistura

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Doutor em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística. *VERSÃO REVISADA*

Área de Concentração: Estatística

Orientador: Prof. Dr. Alexsandro Giacomo Grimbert Gallo
Coorientador: Prof. Dr. Miguel Natalio Abadi

USP – São Carlos
Janeiro de 2022

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Defesa de Tese de Doutorado do candidato Vitor Gustavo de Amorim, realizada em 17/02/2022.

Comissão Julgadora:

Prof. Dr. Alexsandro Giacomo Grimbert Gallo (UFSCar)

Profa. Dra. Maria Eulalia Vares (UFRJ)

Prof. Dr. Jefferson Antonio Galves (IME-USP)

Prof. Dr. Jean-René Chazottes (CNRS)

Profa. Dra. Ana Cristina Moreira Freitas (U.Porto)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.
O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa Interinstitucional de Pós-Graduação em Estatística.

I dedicate this thesis to the people who, with their love and teachings, have brought me here: my wife Graziele, my brothers Marcelo and Fabio and my sister Luciana, my mother Celi and my dearly missed and beloved father Antonio Carlos (in memoriam).

ACKNOWLEDGEMENTS

First, I would like to thank my advisor professor Sandro Gallo for his support, teachings and directions. My incursion into the world of the stochastic processes began in a excellent course taught by him. After that, he showed me all the ways in Poincaré's recurrence theory through very rich and productive discussions. I could not fail to mention that his support was fundamental for the writing of this thesis, with his never ending valuable contributions.

I also express my gratitude to my co-advisor professor Miguel Abadi. His deep knowledge in the area of recurrence times, as one of the leading researchers in the field, has brought countless contributions to my research. Several results of this thesis were inspired by the ideas present in his papers.

Many professors in the Statistics Interagency Graduate Program (PIPGEs) were also very important throughout my whole period of PhD studies. In particular, for their teaching, I would like to thank those who taught the courses I took: Renato Gava, Vicente Cancho, Mario de Castro and, again, Sandro Gallo.

The support of the staff of the Department of Statistics (DEs-UFSCAr) and of the Institute of Mathematics and Computer Sciences (ICMC-USP) was very important during these four years. I express my thanks to them for their help on several occasions.

My affiliation institution, the Federal Institute of São Paulo (IFSP), was one of the main responsible for making this PhD possible. I am very grateful to them for their financial support and for encouraging research and training through paid leave programs.

During this entire period, I had the honour of sharing hours of study and fun with my colleagues at PIPGEs. In particular, I thank my friends Ana Paula, Bruna, Caio, Claudia, Jardel and Oilson for their friendship and help in difficult times.

Last but not least, I would like to thank my beloved family for their unconditional love and never ending support. I thank my loving and caring mother, who was always there for me when I needed her, as well as my brothers Marcelo and Fabio and my sister Luciana. Without you I would not have come this far, I love you! Finally, I have no words to express my love and gratitude to my wife Graziele. During the entire PhD period, she was by my side, supporting me and doing everything in her power to help me, especially during this period of finishing the thesis. My love and gratitude for you are for life.

"It is by logic that we prove,
but by intuition that we discover."
(Henri Poincaré)

ABSTRACT

AMORIM, V. G. **Poincaré recurrence times in stochastic mixing processes**. 2022. 115 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

In the context of the discrete-time stochastic processes, this thesis presents new results on Poincaré recurrence theory. After a complete review of recent results, we present a new theorem on the exponential approximations for hitting and return times distributions. We show that the scaling parameter of the approximate distribution, called "potential well", brings fundamental informations about the structure of the target set. Moreover, we show that the asymptotic properties of the potential well influences several aspects of the recurrence times, such as limiting distributions and moments. Finally, we apply our results to obtain the waiting time spectrum as a function of the Rényi entropy for classes of processes not covered by previous works.

Keywords: Hitting and return times, Exponential approximation, Potential well, Waiting time spectrum, Mixing processes.

RESUMO

AMORIM, V. G. **Tempos de recorrência de Poincaré em processos estocásticos com mistura**. 2022. 115 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

Esta tese apresenta novos resultados na teoria de recorrência de Poincaré, no contexto de processos estocásticos em tempo discreto. Após uma revisão completa de resultados recentes, apresentamos um novo teorema sobre aproximação exponencial para distribuições de tempos de entrada e retorno. Mostramos que o parâmetro de escala da distribuição aproximada, chamado "poço de potencial", traz ao teorema de aproximação informações fundamentais sobre a estrutura do conjunto alvo. Além disso, mostramos que as propriedades assintóticas do poço de potencial influenciam vários aspectos dos tempos de recorrência, como as distribuições limite e seus momentos. Por fim, aplicamos nossos resultados para obter o espectro do tempo de espera como função da entropia de Rényi para classes de processos não cobertas por trabalhos anteriores.

Palavras-chave: Tempos de entrada e retorno, Aproximação exponencial, Poço de potencial, Espectro do tempo de espera, Processos misturadores.

CONTENTS

CHAPTER

1

INTRODUCTION

The Poincaré recurrence theory has been studied in several contexts throughout history, such as dynamical systems, extreme value theory, stochastic processes among others. The first classical result, known as Poincaré recurrence theorem (see (SHIELDS, 1996)), states that typical orbits of measure-preserving systems return infinitely many times to any set with positive measure.

In the context of stochastic processes, measure-preserving means stationarity and the Poincaré theorem states that if a process $\mathbf{X} = (X_m)_{m\in\mathbb{N}}$ is stationary and starts from a set A with positive probability, then it returns to A almost surely in a finite time. In the context of stationary ergodic processes, there is a version of the Poincaré theorem stating that the first time $\mathbf{X}$ enters the set A (the hitting time) is also almost surely finite (see (SHIELDS, 1996)).

In the 1940's, Kac (1947) presented the first quantitative result on this issue. For stationary ergodic processes, he provided an explicit and intuitive formula to the mean return time to A: it equals the reciprocal of the measure of the target set A. After Kac's Lemma, it is natural to seek for a similar result for hitting times. However, this task proved to be more difficult and, since there are no explicit formulas for the hitting and return times distributions, approximate distributions became the main goal.

It is worth mentioning that most part of the results in this thesis are stated for stochastic mixing processes. However, hitting and return times have been recently studied in the contexts of dynamical systems and the extreme value theory. We refer to (HAYDN, 2013), (FREITAS; FREITAS; TODD, 2010), and (LUCARINI *et al.*, 2016) for nice revisions (and for the lists of references therein) about this approach, since we will not consider this context throughout the text.

Exponential approximations

Back to stochastic processes, when A is a cylinder set, several works have shown that, under strong mixing conditions, the hitting time distribution can be well approximated by an exponential law with parameter depending on A. Similarly, the return time has as approximate distribution a convex combination between a Dirac measure and an exponential law. We refer to (ABADI; GALVES, 2000) for a review of the results before the 2000s. In Section 3.2, we bring a new review of recent results on this issue, focusing in the different parameters used in the literature for these approximations.

The Kac's Lemma suggests that the right parameter for the exponential approximation of the hitting time would be the measure of the set A. However, Galves and Schmitt (1997) showed that a correcting factor $\theta(A)$ is necessary. Recent works therefore focused on obtaining approximations of the form

$$\sup_{t\geq 0}\left|\mu(T_A>t)-e^{-\theta(A)\mu(A)t}\right|\leq \varepsilon(A),$$

where μ is the stationary measure of the process, T_A denotes the hitting time to A and $\varepsilon(A)$ is an error term which converges to zero as the length of the cylinder A diverges. In the case of the return times, the typical result is

$$\sup_{t\geq\tau(A)}\left|\mu(T_A>t|A)-\bar{\theta}(A)e^{-\theta(A)\mu(A)t}\right|\leq \varepsilon(A),$$

where $\tau(A)$ is the shortest possible return to A.

Recent results is this area tried to reach at least three types of progress:

1. To obtain a sharp uniform error term $\varepsilon(A)$ and an error term $\varepsilon(A,t)$ that decreases in t;
2. To expand the classes of processes for which these approximations hold;
3. To obtain the better scaling parameter $\theta(A)$.

In Chapter 3, we present a new hitting and return time theorem (Theorem 3.3.4) which brings contributions for all of them.

Concerning (1), we detail in Section 3.1 the question of "total variation distance" versus the "pointwise" approximations. Our result provides approximations for hitting and return times whose error term $\varepsilon(A,t)$ decays exponentially fast in t (a pointwise approximation). We also show that the uniform error term $\varepsilon(A)=\sup_t\varepsilon(A,t)$ is sharp for both approximations. Together with point (3), our approximations allow us to derive explicit formulas that approximate all the moments of the hitting and return times. As a consequence, one can also obtain the spectrum of the waiting time, which we do in Chapter 5.

Our result also brings contributions to (2): it is the first pointwise approximation result including processes with infinite alphabet and incomplete grammar (where not every cylinder has positive measure). In order to illustrate these and other results throughout the text, we present in Chapter 2 several useful examples of processes and new results on mixing properties of renewal processes.

Potential well

The contribution on item (3) has a crucial role in this thesis. In Section 3.2, we show that previous works presented different scaling parameters whose expressions were hardly explicit and even more hardly computable. Abadi and Vergne (2008) provided in their result for return times a new scaling parameter (the potential well), which is easy to compute and has a physical and intuitive meaning. Inspired by them, our result brings the first pointwise approximation for hitting time distribution using the potential well as scaling parameter. In the case of return times, we extend the main result of (ABADI; VERGNE, 2008) concerning (2).

It is worth mentioning that two cylinders with the same probability have the same mean return time (by Kac's Lemma). Nonetheless, their overlapping properties may lead to different approximations for hitting and return times distributions, since the short return must be taken into account. Consider the following simple example of an i.i.d. process on $\{0,1\}$. Note that the sequences $A = 0000011111$ and $B = 0101010101$ have the same probability, but B can return in two steps, while A cannot. The potential well is the probability, conditioned on starting the process from the target set, that the return does not occur in the first possible time. In other words, this parameter captures the overlapping properties of the target set.

The use of the potential well as scaling parameter also leads to other questions. If we know the asymptotic behaviour of the potential well, we can apply our theorem to obtain the limiting distributions. Moreover, its positivity has influence on the scale of the hitting time and return times distributions. In view of this, the explicitness of the potential well as scaling parameter is an essential feature.

In Chapter 4, we highlight this role of the potential well in recurrence times. The main result there gives conditions for its positivity and almost sure convergence to one. As a consequence, we provide approximate formulas and the asymptotic behaviour for all the moments of the hitting and return times, which can be seen as an asymptotic Kac's Lemma. We also dedicate a section to show how the potential well behaves for three specific classes of processes. In particular, we present some examples that show explicitly the effects on recurrence times and the usefulness of the potential well, an "easy to compute" scaling parameter. Finally, we show the opposite effect: for points where the potential well is arbitrarily close to zero, the return time distribution converges to a degenerated law and the order of the mean hitting time is bigger than $\mu(A)^{-1} = \mathbb{E}(T_A|A)$. Actually, all of the hitting time β-moments have a bigger order than $\mu(A)^{-\beta}$ in this case, as opposed to the case in which the potential is bounded away from zero.

Shortest return time

A fundamental quantity in the study of recurrence times is the shortest possible return time of a n-cylinder A_n, denoted by $\tau(A_n)$. It gives the first time that the probability of the return of A_n is positive. This quantity is directly related with overlapping properties of A_n and, therefore, with the potential well. Under assumptions of specification, finite alphabet and positive entropy, it is known that $\tau(A_n)/n$ converges almost surely to one ((AFRAIMOVICH; CHAZOTTES; SAUSSOL, 2003) and (SAUSSOL; TROUBETZKOY; VAIENTI, 2002)). However, most of our results do not assume complete grammar nor finite alphabet, and we can have in some cases $\tau(A_n) > n$. In this sense, we present at the end of Chapter 4 a result on upper bound for the shortest return time.

Waiting time spectrum

Another sequence of random variables of interest when studying recurrence times is the waiting time W_n, which is the first time that a process hits a randomly chosen n-cylinder $\left[x_0^{n-1}\right]$. An important result due to (WYNER; ZIV, 1989) and extended by (SHIELDS, 1996) shows a relationship between the waiting time and the measure of $\left[x_0^{n-1}\right]$. It states that

$$\lim_{n\to\infty} \frac{1}{n} \ln W_n = h_\mu, \quad \text{a.s.},$$

where h_μ is the Shannon entropy of μ. On the other hand, the Shannon–McMillan–Breiman Theorem states that

$$\lim_{n\to\infty} \frac{1}{n} \ln \mu\left(\left[x_0^{n-1}\right]\right)^{-1} = h_\mu, \quad \text{a.s.}$$

Thus, for typical cylinders and large n we have $W_n \approx \mu\left(\left[x_0^{n-1}\right]\right)^{-1}$.

This relationship, together with the hitting time theorem in (ABADI, 2004), motivated Chazottes and Ugalde (2005) to find a "strong approximation" of $\mathbb{E}\left(W_n^q\right)$ by $\mathbb{E}\left(\mu\left(\left[x_0^{n-1}\right]\right)^{-q}\right)$ for all $q \in \mathbb{R}$. In fact, they showed that this holds for all $q \geq -1$, while for all $q < -1$, $\mathbb{E}\left(W_n^q\right)$ is of the order of $\mathbb{E}\left(\mu\left(\left[x_0^{n-1}\right]\right)\right)$. Then, the authors applied this result to obtain the waiting time spectrum

$$\mathscr{W}(q) = \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}\left(W_n^q\right)$$

as a function of the Rényi entropy function, when it exists. As a consequence, they derived large deviations results for waiting time.

In Chapter 5, we present an extension of the strong approximation obtained by them, concerning the classes of processes for which we can apply the result. Namely, for $q < -1$ we prove that the approximation holds for any stationary process, including the ones with infinite alphabet. For $q \geq -1$, we require mixing conditions weaker than those required in (CHAZOTTES; UGALDE, 2005).

In order to obtain the existence of the waiting time spectrum for new classes of processes, we provide a result on the existence of the Rényi entropy. In particular, we analyse this existence in the classes of renewal processes and *infinite* Markov chains, where we illustrate and justify the extension of the results.

To conclude, it is worth mentioning that there is a work in progress in collaboration with Abadi, Chazottes and Gallo, where our approximation theorem, as well as arguments presented in Chapter 5, are also being used to make a corrected and simplified version of the paper (ABADI; CHAZOTTES; GALLO, 2019) on the return time spectrum.

CHAPTER

2

MIXING PROPERTIES IN STOCHASTIC PROCESSES

The main objects of interest of this thesis are the mixing stochastic processes on countable alphabets. In the present chapter we introduce this context. We will therefore give the basic notation, define the main mixing properties of interest, and give some examples that will be used along the thesis. We also give some new results concerning the mixing properties of renewal processes. The proofs are given in the last section.

2.1 The framework in mixing processes

2.1.1 Notation and basic definitions

Throughout the text, we will consider a finite or countable set $\mathscr{A}$, which we will call "alphabet". Denote by $\mathscr{X} = \mathscr{A}^{\mathbb{N}}$ the set of right-infinite sequences of symbols taken from $\mathscr{A}$. An element of $\mathscr{X}$ will be denoted by $\mathbf{x} = x_0 x_1 \cdots x_n \cdots = x_0^\infty$. The string (or word) $a_m^n = a_m a_{m+1} ... a_n$, where $a_i \in \mathscr{A}$ and $m \leq n$, is the sequence of symbols $a_m, a_{m+1}, \cdots, a_n$. By a^n we denote the string of n consecutive symbols $a \in \mathscr{A}$. We define the cylinder $[a_m^n]$ as the set $\{\mathbf{x} \in \mathscr{X}; x_m^n = a_m^n\}$. The shift operator $\sigma : \mathscr{X} \to \mathscr{X}$ shifts the point $x = (x_0, x_1, x_2, \cdots)$ to the left by one coordinate: $(\sigma(x))_i = x_{i+1}$, $i \geq 0$.

We denote by $\mathscr{F}$ the σ-algebra over $\mathscr{X}$ generated by all cylinders and by $\mathscr{F}_I$ we denote the σ-algebra generated by cylinders with coordinates in $I \subset \mathbb{N}$. For the special case in which $I = \{i, \dots, j\}$, $0 \leq i \leq j \leq \infty$, we use the notation $\mathscr{F}_i^j$.

Consider the probability space $(\mathscr{X}, \mathscr{F}, \mathbb{P})$. On this space, we consider a right-infinite stationary stochastic process taking values in $\mathscr{A}$, which we denote by $\mathbf{X} = (X_m)_{m\in\mathbb{N}}$. By stationarity, we mean

$$\mathbb{P}\left(X_k^{k+n-1} = a_0^{n-1}\right) = \mathbb{P}\left(X_0^{n-1} = a_0^{n-1}\right)$$

for all $k \geq 0$, $n \geq 1$ and $a_0^{n-1} \in \mathscr{A}^n$. This process univocally defines a measure μ on $(\mathscr{X}, \mathscr{F})$ through

$$\mu\left(\left[a_0^{n-1}\right]\right) = \mathbb{P}\left(X_0 = a_0^{n-1}\right), \quad \forall n \geq 1; a_0^{n-1} \in \mathscr{A}^n.$$

The stationarity of $\mathbf{X}$ implies that μ is σ-invariant, that is $\mu \circ \sigma^{-1} = \mu$.

According to necessities of the presentation, and when no ambiguity is possible on the process to which we refer, we will sometimes switch between the dynamical system approach, using directly μ, and the stochastic processes approach referring to $\mathbf{X}$. In the first case, we directly write $\mu(F)$ for $F \in \mathscr{F}$, while in the second case we will abuse the notation and directly identify μ and $\mathbb{P}$, writing $\mu(F)$ instead of $\mu(\mathbf{X} \in F)$.

For the sake of notation, it will be sometimes useful to identify the string a_m^n with the cylinder $[a_m^n]$, that is, $\mu\left(a_m^n\right)$ will mean $\mu\left([a_m^n]\right)$. The set of all n-cylinders with positive measure is denoted by $\mathscr{C}_n = \left\{a_0^{n-1} \in \mathscr{A}^n; \mu\left(a_0^{n-1}\right) > 0\right\}$. We say that a process or measure has complete grammar if $\mathscr{C}_n = \mathscr{A}^n$.

Given $B, C \in \mathscr{F}$ such that $\mu(C) > 0$, the conditional measure will be denoted by $\mu(B|C) = \mu_C(B)$. In particular, for a word $A \in \mathscr{C}_n$, $\mu_A(B)$ will denote $\mu\left(B|X_0^{n-1} = A\right)$.

2.1.2 Classical types of mixing

Several results of the thesis are stated under mixing conditions that we define now.

Definition 2.1.1. Let $\mathbf{X}$ be a stationary process with stationary measure μ. For all $n \geq 1$, we define the functions:

$$\begin{aligned}
\psi(n) &:= \sup_{i\in\mathbb{N}, A\in\mathscr{F}_0^i, B\in\mathscr{F}_{i+n}^\infty} \left|\frac{\mu_A(B)}{\mu(B)} - 1\right|, \quad \text{where } \mu(A)\mu(B) > 0,\\
\phi(n) &:= \sup_{i\in\mathbb{N}, A\in\mathscr{F}_0^i, B\in\mathscr{F}_{i+n}^\infty} \left|\mu_A(B) - \mu(B)\right|, \quad \text{where } \mu(A) > 0,\\
\alpha(n) &:= \sup_{i\in\mathbb{N}; A\in\mathscr{F}_0^i, B\in\mathscr{F}_{i+n}^\infty} \left|\mu(A \cap B) - \mu(A)\mu(B)\right|,
\end{aligned}$$

as well as

$$\beta(n) := \sup \frac{1}{2} \sum_{k=1}^{K} \sum_{j=1}^{J} \left|\mu(A_k \cap B_j) - \mu(A_k)\mu(B_j)\right|$$

where the supremum is taken over all $i \in \mathbb{N}$ and all the possible partitions $\{A_1, \ldots, A_K\}$ and $\{B_1, \ldots, B_J\}$ of $\mathscr{X}$ satisfying $A_k \in \mathscr{F}_0^i$ and $B_j \in \mathscr{F}_{i+n}^\infty$.

We say that $\mathbf{X}$ is α, β, ϕ or ψ-mixing if $\alpha(n)$, $\beta(n)$, $\phi(n)$ or $\psi(n)$ converges to 0 as n diverges, respectively. We can say alternatively that μ is an α-mixing measure, as well as β, ϕ or ψ-mixing.

Note that $\alpha(n)$, $\psi(n)$, $\phi(n)$ and $\beta(n)$ are decreasing sequences, since $\mathscr{F}_0^i \subset \mathscr{F}_0^{i+1}$ for every $i \geq 0$. We refer to (BRADLEY, 2005) for an exhaustive review of mixing properties of stochastic processes. Therein it is for instance explained that

$$\psi\text{-mixing} \Rightarrow \phi\text{-mixing} \Rightarrow \beta\text{-mixing} \Rightarrow \alpha\text{-mixing}.$$

since $2\alpha(n) \leq \beta(n) \leq \phi(n) \leq (1/2)\psi(n)$.

2.1.3 Basic examples

Let us discuss the mixing properties of some specific classes of processes. The most trivial example is the so called i.i.d. process, which is defined over a finite or countable alphabet $\mathscr{A}$ and trivially has complete grammar and function $\psi(n)$ identically equal to zero. The independence of events over disjoint sets of coordinates of this process leads to the stronger mixing properties among all the stationary processes.

When we consider some dependence on the past, the class of the Markov chains is the first natural choice. With regard to mixing properties, we need to analyse two separate cases: finite and infinite (countable) alphabet. For the first case, Theorems 3.1 and 3.3 in (BRADLEY, 2005) state that any stationary, finite-state, irreducible and aperiodic Markov chain is *exponentially* ψ-mixing, that is, $\psi(n) \leq Ce^{-cn}$ for some positive constants C and c and any n.

On the other hand, Theorem 3.2 in (BRADLEY, 2005) states that any stationary, infinite-state, irreducible and aperiodic Markov chain is β-mixing (and therefore α-mixing) in general, but the decay rate is not necessarily exponential in both cases.

As we will see in the following examples, under the same conditions as above, infinite-state Markov chains can also be ψ or ϕ-mixing depending on characteristics of the transition matrix Q. If this is the case, Theorem 3.3 in (BRADLEY, 2005) also ensures that the chain is exponentially ψ or ϕ-mixing, respectively. For instance, if $\sup_{a,b\in\mathscr{A}} Q(a,b) < 1$, then the chain is exponentially ϕ-mixing (we provide a proof in Section 2.4 since we could not find a reference).

Example 2.1.2 (A ψ-mixing countable Markov chain)**.** For some $p \in (0,1)$, consider the Markov chain $\mathbf{X}$ defined over the alphabet $\mathscr{A} = \mathbb{N}$ by the transition matrix:

$$Q(i,j) = \begin{cases} p(1-p)^j, & \text{if } i = 0 \\ p, & \text{if } i > 0 \text{ and } j = 0 \\ 1-p, & \text{if } i > 0 \text{ and } j = 1. \end{cases}$$

It is immediate to verify that $\mathbf{X}$ is irreducible and aperiodic. Furthermore, direct calculations provide a probability distribution π over $\mathbb{N}$ satisfying $\pi Q = \pi$, which induces a stationary measure μ on the chain.

Let us show that this chain satisfies $\psi(n) = 0$ for any $n \geq 2$, that is, $\mathbf{X}$ is ψ-mixing. In order to calculate $\psi(2)$, it is sufficient to consider the cylinders that generate the sigma-algebras $\mathscr{F}_0^i$ and $\mathscr{F}_{i+2}^\infty$ for every $i \geq 0$. Then,

$$\psi(2) = \sup_{i\geq 0, a_0^i \in \mathscr{F}_0^i, b_{i+2}^{i+2+k} \in \mathscr{F}_{i+2}^\infty} \left| \frac{\mu_{a_0^i}\left(b_{i+2}^{i+2+k}\right)}{\mu\left(b_{i+2}^{i+2+k}\right)} - 1 \right| = \sup_{a,b\in\mathbb{N}} \left| \frac{Q^2(a,b)}{\mu(b)} - 1 \right|.$$

By analysing separately the cases $b > 1$, $b = 1$ and $b = 0$, we conclude that $Q^2(a,b) = \mu(b)$ for all $a, b \in \mathbb{N}$, which implies $\psi(n) \leq \psi(2) = 0$ for all $n \geq 2$.

We present now a well-known class of Markov chains also called the house of cards. Besides its own interest for providing useful examples in the context of the recurrence times, this chain will also be used to define and derive mixing properties of binary renewal processes.

Example 2.1.3 (The house of cards Markov chain)**.** Let $(q_i)_{i\geq 0}$ be a real sequence satisfying $q_i \in (0,1)$ for all $i \geq 0$. We define the house of cards Markov chain $\mathbf{X}$ over the alphabet $\mathscr{A} = \mathbb{N}$ as the chain given by the transition matrix:

$$Q(i,j) = \begin{cases} q_i, & j = i+1 \\ 1 - q_i, & j = 0. \end{cases}$$

For all $n \geq 0$, we denote

$$\sigma_n = \prod_{i=0}^{n-1} q_i \quad \text{and} \quad \Sigma(n) = \sum_{k=n}^{\infty} \sigma_k,$$

where we use the convention $\sigma_0 = 1$.

It is immediate to verify that the chain is irreducible and aperiodic. Further, direct calculations show that $\mathbf{X}$ has a stationary measure μ if, and only if, σ_n is summable, which will be assumed from now on. In this case, μ satisfies:

$$\mu(0) = \frac{1}{1+\Sigma(1)} \quad \text{and} \quad \mu(k) = \mu(0)\sigma_k, \ \forall k \geq 1.$$

Note that, in order to define the house of cards Markov chain, it is sufficient to know one of the functions σ_n or $\Sigma(n)$, since $q_i = \sigma_{i+1}/\sigma_i$ and $\sigma_n = \Sigma(n) - \Sigma(n+1)$ for all $n \geq 0$.

As already mentioned, under the above conditions, Markov chains are always β-mixing. However, we can show that for any choice of the parameters, the house of cards cannot be ψ-mixing. Indeed, for any $n \geq 1$, we have:

$$\begin{aligned}\psi(n) &\geq \sup_{\ell \geq n} \left| \frac{\mu(X_n = \ell | X_0 = \ell - n)}{\mu(X_n = \ell)} - 1 \right| \\ &= \frac{1}{\mu(0)} \sup_{\ell \geq n} \left| \frac{\prod_{j=\ell-n}^{\ell-1} q_j}{\sigma_\ell} - 1 \right| \\ &= \frac{1}{\mu(0)} \sup_{\ell \geq n} \left| \frac{1}{\sigma_{\ell-n}} - 1 \right| \\ &= \infty.\end{aligned}$$

On the other hand, in Theorem 2.2.3 below, we provide criteria on the parameters of the process to determine whether it is exponentially ϕ-mixing or only β-mixing.

Another example that will be useful in the study of recurrence times is a class of processes similar to the house of cards, which we call the "lazy" house of cards. It is also defined over $\mathbb{N}$ with a real sequence $(q_i)_{i \geq 0}$, $q_i \in (0,1)$, but its transition matrix allows that the chain stays in the same state in each step.

Example 2.1.4 (The lazy house of cards Markov chain)**.** Given a real sequence $(q_i)_{i \geq 0}$ satisfying $q_i \in (0,1)$, we define the lazy house of cards $\mathbf{X}$ as the Markov chain over the alphabet $\mathbb{N}$ given by the transition matrix:

$$Q(0,0) = q_0 = 1 - Q(0,1) \quad \text{and} \quad Q(i,j) = \begin{cases} q_i, & j = i \\ \frac{1-q_i}{2}, & j = i+1 \text{ or } j = 0. \end{cases}$$

Direct computations show that if

$$\sum_{j=2}^{\infty} \frac{1}{2^{j-1}(1-q_j)} < \infty,$$

then the chain has a stationary measure μ. This condition holds, for instance, when $\sup_i\{q_i\} < 1$ or when $1 - q_i$ decreases at a polynomially rate. The chain is obviously irreducible and aperiodic, which means that it is also β-mixing. As for the house of cards Markov chains, we can show that the lazy house of cards cannot be ψ-mixing. On the other hand, if $0 < \inf_i\{q_i\} \leq \sup_i\{q_i\} < 1$, then Proposition 2.4.1 will give at least one case in which the chain is exponentially ϕ-mixing.

2.2 Mixing properties in renewal processes

In this section we present a class of processes that provides a large range of scenarios in recurrence times. The binary renewal processes will be used in the next three chapters as a class that justifies and illustrates the results. Next, we study their mixing properties.

2.2.1 Definition of renewal processes

One can define a renewal process as function of the house of cards Markov chain. From now on, we denote by $\mathbf{X} = (X_m)_{m\in\mathbb{N}}$ the Markov chain defined in Example 2.1.3 and by $\bar{\mu}$ we denote its law. A binary renewal process $\mathbf{Y} = (Y_m)_{m\in\mathbb{N}}$ is defined by $Y_m = \mathbb{1}\{X_m = 0\}$ for all $m \geq 0$. This process has a unique stationary measure μ obtained from the stationary measure $\bar{\mu}$ of the chain $\mathbf{X}$.

By direct computations, μ satisfies

$$1-\mu(0) = \mu(1) = \bar{\mu}(0) = \frac{1}{1+\Sigma(1)},$$

$$\mu\left(10^n\right) = \mu\left(0^n 1\right) = \mu(1)\sigma_n \quad \text{and}$$

$$\mu\left(0^n\right) = \mu(1)\Sigma(n),$$

for all $n \geq 1$, where the functions σ_n and $\Sigma(n)$ are the same as defined in Example 2.1.3. Recall that we are assuming $\sum_n \sigma_n < \infty$ and that one can completely define the process knowing only one of the functions σ_n or $\Sigma(n)$, or simply the sequence $(q_i)_{i\geq 0}$. Note that σ_n defines the inter-arrival distribution of the renewal process, since $\sigma_n = \mu_1\left(0^n\right) = \mu_1(T > n)$, where T denotes the distance between two consecutive occurrences of 1.

We can obtain from the Markov property the so called renewal property (see Lemma 2.4.2), which states that the dependence on the past in renewal processes is up to the last occurrence of a 1. The renewal property ensures that the distance between consecutive occurrences of 1's are i.i.d. copies of T.

Remark 2.2.1. As shown in Proposition 3.1 in (ABADI; CARDENO; GALLO, 2015), a stationary measure μ of a renewal process is reversible, which means that $\mu(a_0a_1\cdots a_{n-1}) = \mu(a_{n-1}a_{n-2}\cdots a_0)$ for any $a_0^{n-1} \in \{0,1\}^n$.

2.2.2 β, ϕ and ψ-mixing mixing renewal processes

We now present some results to provide a complete characterization of the mixing properties of renewal processes as a function of the parameters. Let us start with an elementary result which is a direct consequence of the definition of $\mathbf{Y}$ as function of $\mathbf{X}$.

Proposition 2.2.2. Let $\Delta_{\mathbf{X}}(n)$ be one of the functions given in Definition 2.1.1 (α, β, ϕ, ψ) associated with the house of cards Markov chain $\mathbf{X}$ and $\Delta_{\mathbf{Y}}(n)$ the same function associated with the renewal process $\mathbf{Y}$. Then

$$\Delta_{\mathbf{Y}}(n) \leq \Delta_{\mathbf{X}}(n).$$

In particular, if $\mathbf{X}$ is Δ-mixing, so does $\mathbf{Y}$.

In particular, this result ensures that a renewal process is always β-mixing. We therefore focus, in the remaining of this section, on identifying necessary and/or sufficient conditions ensuring ϕ and ψ-mixing.

Conditions for ϕ-mixing

We now present a theorem that states necessary and sufficient conditions on the parameters so that processes **X** and **Y** are (exponentially) ϕ-mixing.

Theorem 2.2.3. Let **X** and **Y** be the house of cards Markov chain and its respective renewal process. Then the following statements are equivalent:

(a) **X** is exponentially ϕ-mixing.

(b) **Y** is exponentially ϕ-mixing.

(c) **Y** is ϕ-mixing.

(d) There exist positive constants C and c such that for all $r,s \geq 0$ we have

$$\frac{\sigma_{r+s}}{\sigma_r}(= q_r \cdots q_{r+s-1}) \leq Ce^{-cs}.$$

(e) There exist $n_0 \geq 1$ and $R < 1$ such that $\min\{q_r, \cdots, q_{r+n_0-1}\} \leq R$ for all $r \geq 0$.

Remark 2.2.4 (Condition $\sigma_n \leq Ce^{-cn}$ is necessary but not sufficient)**.** As a direct consequence of the theorem, notice that if a renewal process **Y** or a house of cards Markov chain **X** is ϕ-mixing, then $\sigma_n \leq Ce^{-cn}$ for some positive constants C and c, and in particular $\Sigma(n) \leq Ce^{-cn}/(1-e^{-c})$. It is natural to wonder whether this assumption is also equivalent to ϕ-mixing, that is, to (d) and (e) above. It turns out that this is not the case, as will be shown by Example 2.2.10 at the end of this subsection.

The next direct corollary provides easy to check conditions, directly on the parameters $(q_i)_{i\geq 0}$.

Corollary 2.2.5. If $q_i \xrightarrow{i} 1$, then **X** and **Y** are not ϕ-mixing. On the other hand, if $\limsup_i q_i < 1$, both are exponentially ϕ-mixing.

It is natural to wonder what happens when (q_i) does not converge but $\limsup_i q_i = 1$. The end of this subsection provides several examples showing that anything could happen in terms of mixing properties.

Conditions for ψ-mixing

The next theorem presents the only difference between the mixing properties of the house of cards Markov chain and the renewal process. It establishes a sufficient and necessary condition on the parameters for a renewal process **Y** to be ψ-mixing, which cannot occur with the Markov chain $\mathbf{X}$, as we saw in Example 2.1.3. Note that in the case of ψ-mixing renewal processes, unlike Markov chains, the decay rate of the function ψ is not known.

Theorem 2.2.6. Let $\mathbf{Y}$ be a renewal process defined by the function σ_n. Then, $\mathbf{Y}$ is ψ-mixing if, and only if, there exists $n_0 \in \mathbb{N}$ such that

$$\sup_{r,s\in\mathbb{N}} \frac{\sigma_{r+s+n_0}}{\sigma_r\sigma_s} < \infty. \tag{2.1}$$

Let us also state a direct corollary providing an easy to check condition on the parameters $(q_i)_{i\geq 0}$ ensuring ψ-mixing.

Corollary 2.2.7. If $(q_i)_{i\geq 0}$ is decreasing, then the associated renewal process $\mathbf{Y}$ is ψ-mixing.

Some explicit examples illustrating the above results

As we saw in Corollary 2.2.5, when $q_i \longrightarrow 1$, the respective renewal process is not ϕ-mixing (neither ψ), while on the other hand, $\limsup_i q_i < 1$ implies exponential ϕ-mixing. It remains to investigate what could happen when $\liminf_i q_i < \limsup_i q_i = 1$. It turns out that we may have: (1) ψ-mixing; (2) ϕ-mixing without ψ-mixing; (3) β-mixing without ϕ-mixing. The next three examples illustrates these three possibilities.

Example 2.2.8 (ψ-mixing with $\liminf_i q_i < \limsup_i q_i = 1$)**.** Define $q_i = 1/2$ for even i and $q_i = e^{-(2/(i+1))^2}$ for odd i. If $n \in \mathbb{N}$ is even, we have

$$\left(\frac{1}{2}\right)^{n/2} e^{\frac{2}{n}-2} \leq \sigma_n = \left(\frac{1}{2}\right)^{n/2} e^{-\sum_{j=1}^{n/2}\frac{1}{j^2}} \leq \left(\frac{1}{2}\right)^{n/2} e^{\frac{2}{n+2}-1}.$$

Hence, for $r,s \geq 2$ even numbers, we obtain:

$$\frac{\sigma_{r+s}}{\sigma_r\sigma_s} \leq \frac{e^{2/(r+s+2)-1}}{e^{2/r+2/s-4}} \leq e^4.$$

When r is an odd number, we have $\sigma_{r+1} \leq \sigma_r \leq \sigma_{r-1}$, which allows us to get back to the previous case. Then either case satisfies condition (2.1).

Example 2.2.9 (ϕ-mixing without ψ-mixing and with $\liminf_i q_i < \limsup_i q_i = 1$)**.** Let $q_i = 1/2$ for even i and $q_i = i/(i+2)$ for odd i. We have ϕ-mixing by Theorem 2.2.3-(e), but we do not have ψ-mixing, since for n even we have $\sigma_n = \dfrac{1}{2^{n/2}(n+1)}$, which gives us

$$\begin{aligned}
\sup_{r,s\in\mathbb{N}} \frac{\sigma_{r+s+n_0}}{\sigma_r\sigma_s} &\geq \sup_{r,s \text{ even}} \frac{\sigma_{r+s+2n_0}}{\sigma_r\sigma_s} \\
&= \frac{1}{2^{n_0}} \sup_{r,s \text{ even}} \frac{(r+1)(s+1)}{(r+s+2n_0+1)} \\
&= \infty
\end{aligned}$$

for all $n_0 \in \mathbb{N}$.

Example 2.2.10 (β-mixing without ϕ-mixing and with $\liminf_i q_i < \limsup_i q_i = 1$)**.** If $(q_i)_{i\geq 0}$ is given by

$$q_i = \begin{cases} e^{-i}, & \text{if } i = 2^k, k \geq 1 \\ \left(\frac{i+1}{i+2}\right)^2, & \text{if } i \neq 2^k \end{cases},$$

then, given $R < 1$ and $n_0 \in \mathbb{N}$, there exist infinitely many $r \in \mathbb{N}$ such that for all $j = r, \cdots, r+n_0-1$ we have $q_j = \left(\dfrac{j+1}{j+2}\right)^2 > R$, which contradicts Item (e) of Theorem 2.2.3. This implies that $\mathbf{X}$ and $\mathbf{Y}$ are not ϕ-mixing.

On the other hand, for all $n \geq 2$ there exist $k \geq 1$ such that $n/2 \leq 2^k \leq n$, which implies

$$\sigma_n \leq q_{2^k} = e^{-2^k} \leq e^{-n/2}.$$

giving therefore an example illustrating Remark 2.2.4.

Our last two examples consider cases in which $\sup q_i < 1$ and are applications of Theorem 2.2.6 and Corollary 2.2.7.

Example 2.2.11 (ϕ-mixing without ψ-mixing and with $\sup q_i < 1$)**.** If we take $q_i = p\,(i+1/i+2)$, where $p \in (0,1)$, then $\sigma_n = p^n/(n+1)$ and for any $n_0 \in \mathbb{N}$ we have:

$$\sup_{r,s\in\mathbb{N}} \frac{\sigma_{r+s+n_0}}{\sigma_r \sigma_s} = p^{n_0} \sup_{r,s\in\mathbb{N}} \frac{(r+1)(s+1)}{(r+s+n_0+1)} = \infty,$$

that is, the condition (2.1) does not hold. Nevertheless, we have $\sup_i\{q_i\} = p < 1$, which implies ϕ-mixing by Corollary 2.2.5.

Example 2.2.12 (ψ-mixing with $\sup q_i < 1$)**.** If $q_i = p\,(i+2/i+1)$ for $p \in (0, 1/2)$, then Corollary 2.2.7 ensures that this process is ψ-mixing.

2.3 Further mixing properties

We conclude this chapter by presenting three further mixing properties that will be used later in Chapter 5. The first two, ψ^* and ψ' are related to ψ-mixing and are well-known in the literature. The third, ψ_g-regularity, was introduced recently by (ABADI; CARDENO, 2015) and is less known.

2.3.1 ψ^* and ψ'-mixing

Let us start defining the functions ψ' and ψ^*; we refer to (BRADLEY, 2005) for more details about them.

Definition 2.3.1. Let μ be a stationary measure. We define the functions:

$$\psi'(n) := \inf_{i\in\mathbb{N}, A\in\mathscr{F}_0^i, B\in\mathscr{F}_{i+n}^\infty} \frac{\mu(A\cap B)}{\mu(A)\mu(B)} \quad \text{and} \quad \psi^*(n) := \sup_{i\in\mathbb{N}, A\in\mathscr{F}_0^i, B\in\mathscr{F}_{i+n}^\infty} \frac{\mu(A\cap B)}{\mu(A)\mu(B)}$$

We say that μ is ψ'-mixing (resp. ψ^*-mixing) if $\psi'(n) \xrightarrow{n} 1$ (resp. $\psi^*(n) \xrightarrow{n} 1$).

Note that $\psi'(n)$ is increasing while $\psi^*(n)$ is decreasing. It is known that a stationary process is ψ-mixing if, and only if, it is ψ' *and* ψ^*-mixing. Furthermore, each one of these mixing conditions imply ϕ-mixing (see inequalities (1.11)-(1.18) and Theorem 4.1 in (BRADLEY, 2005)).

The next proposition gives conditions to obtain ψ^* or ψ' for renewal process, which will be used in Chapter 5.

Proposition 2.3.2. Consider a renewal process defined by $(q_i)_{i\geq 0}$.

(a) If $\sup_{r,s\in\mathbb{N}} \dfrac{\sigma_{r+s}}{\sigma_r\sigma_s} < \infty$, then the process is ψ-mixing and satisfies $\psi^*(1) < \infty$.

(b) If $\inf_{r,s\in\mathbb{N}} \dfrac{\sigma_{r+s}(1-q_{r+s})}{\sigma_r\sigma_s} > 0$, then the process is ϕ-mixing and satisfies $\psi'(1) > 0$.

2.3.2 ψ_g-regularity

Here, we present this new mixing property and compare it with the classical mixing conditions presented in Definition 2.1.1, especially in the context of the renewal processes.

Abadi and Cardeno (2015) introduced the so called ψ_g-regular condition, where $g \in \mathbb{N}$. Loosely speaking, unlike the classical mixing conditions, to get the ψ_g regularity one fixes a gap of length $g+1$ and evaluates the supremum and the infimum of the ratio $\mu_A(B)/\mu(B)$ over all cylinders with positive measure and fixed length. Then, it is required that the supremum (resp. infimum) does not grow (resp. decrease) exponentially fast as the length of the cylinders diverges. Namely, for a stationary process $\mathbf{X}$ with law μ, we define the functions:

$$\psi_g^+(m,n) = \sup_{x_0^{m-1}\in\mathscr{C}_m, y_0^{n-1}\in\mathscr{C}_n} \frac{\mu\left(X_0^{m-1} = x_0^{m-1}; X_{m+g}^{m+g+n-1} = y_0^{n-1}\right)}{\mu\left(x_0^{m-1}\right)\mu\left(y_0^{n-1}\right)} \quad \text{and}$$

$$\psi_g^-(m,n) = \inf_{x_0^{m-1}\in\mathscr{C}_m, y_0^{n-1}\in\mathscr{C}_n} \frac{\mu\left(X_0^{m-1} = x_0^{m-1}; X_{m+g}^{m+g+n-1} = y_0^{n-1}\right)}{\mu\left(x_0^{m-1}\right)\mu\left(y_0^{n-1}\right)}.$$

Note that there exists a relationship between the above functions and the mixing functions presented in Definition 2.3.1, since $\psi'(g+1) = \inf_{m,n}\{\psi_g^-(m,n)\}$ and $\psi^*(g+1) = \sup_{m,n}\{\psi_g^+(m,n)\}$.

Definition 2.3.3. We say that a process $\mathbf{X}$ with law μ is ψ_g-regular if there exists $g \in \mathbb{N}$ such that for every $a \geq 0$ and $b \in \mathbb{R}$ we have $0 < \psi_g^-(n, an+b) \leq \psi_g^+(n, an+b) < \infty$ and

$$\lim_{n\to\infty} \frac{\ln\left(\psi_g^-(n,an+b)\right)}{n} = \lim_{n\to\infty} \frac{\ln\left(\psi_g^+(n,an+b)\right)}{n} = 0.$$

There is a simple relationship between ψ_g-regular and ψ-mixing, which we state in the next proposition.

Proposition 2.3.4. If a process $\mathbf{X}$ is ψ-mixing, then it is ψ_g-regular.

Remark 2.3.5. Abadi and Cardeno (2015) provided in their Example 5.2 useful testing functions depending on σ_n and $\Sigma(n)$ that allow one to determine whether a renewal process is ψ_0-regular. Applying these criteria, they presented two examples of (exponentially) ϕ-mixing renewal processes that are ψ_g-regular. Furthermore, they showed that every renewal process defined by $\sigma_n = n^{-s}$ with $s > 1$ is ψ_0-regular. Notice that in this case we have $q_i \longrightarrow 1$, which means that the process is not ϕ-mixing by Corollary 2.2.5. This might suggest that the class of ψ_g-regular processes is at least as general as β-mixing processes. However, as we show in the next example, there exists at least one exponentially ϕ-mixing renewal process which is not ψ_g-regular.

Example 2.3.6. Consider a renewal process $\mathbf{Y}$ defined by the sequence $(q_i)_{i\geq 0}$ given by $q_0 = q$ and:

$$q_i = \begin{cases} p & \text{if } 2^k \leq i < 2^{k+1} \text{ for even } k \\ q & \text{if } 2^k \leq i < 2^{k+1} \text{ for odd } k \end{cases},$$

where $0 < p < q < 1$. Then $\mathbf{Y}$ is exponentially ϕ-mixing by Corollary 2.2.5. On the other hand, observe that for any $g \in \mathbb{N}$, we have

$$\begin{aligned} \psi_g^+(n,n) &\geq \frac{\mu\left(X_0^{2n+g-1} = 0^{2n+g}\right)}{\mu\left(0^n\right)\mu\left(0^n\right)} \\ &= \frac{\Sigma(2n+g)}{\mu(1)\Sigma(n)\Sigma(n)} \\ &\geq \frac{\sigma_{2n+g}}{\mu(1)D^2\sigma_n\sigma_n}, \end{aligned}$$

where D is a positive constant and the last inequality follows from Item (d) of Theorem 2.2.3 and (2.2). Consider now the subsequence $n_k = 2^k$ for odd k. Direct calculations show that $\sigma_{n_k} = p^{(2n_k-1)/3} q^{(n_k+1)/3}$. Therefore,

$$\begin{aligned} \frac{\sigma_{2n_k+g}}{\mu(1)D^2\sigma_{n_k}\sigma_{n_k}} &= \frac{q_{n_k}\cdots q_{2n_k-1}}{\mu(1)D^2\sigma_{n_k}} q_{2n_k}\cdots q_{2n_k+g-1} \\ &\geq \frac{p^g}{\mu(1)D^2}\cdot\left(\frac{q}{p}\right)^{(2n_k-1)/3}, \end{aligned}$$

which gives us:

$$\limsup_{n\to\infty} \frac{\ln\left(\psi_g^+(n,n)\right)}{n} \geq \frac{2}{3}\ln\left(\frac{q}{p}\right) > 0.$$

Thus, $\mathbf{Y}$ is not ψ_g-regular.

2.4 Proofs

Proposition 2.4.1. If the transition matrix Q of a stationary, irreducible and aperiodic Markov chain $\mathbf{X}$ satisfies $\sup_{a,b\in\mathscr{A}} Q(a,b) < 1$, then $\mathbf{X}$ is exponentially ϕ-mixing.

Proof. According to Theorem 3.4 in (BRADLEY, 2005), we just need to show that $\phi(1) < 1$. In this sense, it is sufficient to consider the cylinders that generate the sigma-algebras $\mathscr{F}_0^k$ and $\mathscr{F}_{k+1}^\infty$ for every $k \geq 0$. Thus, consider $A = \left[a_0^k\right] \in \mathscr{F}_0^k$ and $B = \left[b_{k+1}^{k+\ell}\right] \in \mathscr{F}_{k+1}^\infty$.

If we denote $\lambda = \sup_{a,b,\in\mathscr{A}} Q(a,b)$, we have

$$\mu_A(B) = Q(a_k, b_{k+1}) \prod_{j=k+1}^{k+\ell-1} Q(b_j.b_{j+1}) \leq \lambda.$$

Therefore, since $\mu_A(B) \in [0,\lambda]$ and $\mu(B) \in [0,1]$, we have

$$|\mu_A(B) - \mu(B)| \leq \max\{\lambda, 1-\lambda\} < 1,$$

which implies $\phi(1) < 1$. □

Consider a house of cards Markov chain $\mathbf{X}$ and the associated renewal process $\mathbf{Y}$ defined by $Y_m = \mathbb{1}\{X_m = 0\}$ for all $m \geq 0$. In the following proofs, for $0 \leq i \leq j \leq \infty$ we use the notation $\mathscr{F}_i^j$ for the sigma-algebra generated by $(X_i, \cdots, X_j)$ and $\mathscr{G}_i^j$ for the sigma-algebra generated by $(Y_i, \cdots, Y_j)$.

Lemma 2.4.2 (Renewal property)**.** Consider a renewal process $\mathbf{Y}$ defined by $Y_m = \mathbb{1}\{X_m = 0\}$, where $\mathbf{X}$ is a house of cards Markov chain. Let μ be the stationary measure of $\mathbf{Y}$. Given $k \geq 1$, if $A \in \mathscr{G}_0^{k-1}$ and $B \in \mathscr{G}_{k+1}^\infty$, then

$$\mu(B|Y_k = 1, A) = \mu(B|Y_k = 1)$$

Proof. Let $\bar{\mu}$ be the stationary measure of $\mathbf{X}$. We apply the Markov property to obtain:

$$\begin{aligned}
\mu(B|Y_k = 1, A) &= \bar{\mu}\left(\mathbb{1}^{-1}(B)|X_k = 0, \mathbb{1}^{-1}(A)\right) \\
&= \bar{\mu}\left(\mathbb{1}^{-1}(B)|X_k = 0\right) \\
&= \mu(B|Y_k = 1).
\end{aligned}$$

□

Lemma 2.4.3. If μ is the stationary measure of a renewal process $\mathbf{Y}$, then for all $j \geq 0$, $n \geq 2$, $A \in \mathscr{G}_0^j$ and $B \in \mathscr{G}_{j+n}^\infty$ we have

$$\mu\left(A, Y_{j+1}^{j+n-1} \neq 0^{n-1}, B\right) \leq \mu(1)^{-1}\mu(A)\mu(B).$$

Proof. We apply the renewal property to obtain:

$$\begin{aligned}
\mu\left(A, Y_{j+1}^{j+n-1} \neq 0^{n-1}, B\right) &= \sum_{i=0}^{n-2} \mu\left(A, Y_{j+n-1-i}^{j+n-1} = 10^i, B\right) \\
&= \sum_{i=0}^{n-2} \mu\left(A, Y_{j+n-1-i} = 1\right) \mu\left(Y_{j+n-i}^{j+n-1} = 0^i, B | Y_{j+n-1-i} = 1\right) \\
&\leq \mu(A) \sum_{i=0}^{n-2} \frac{\mu\left(Y_{j+n-1-i}^{j+n-1} = 10^i, B\right)}{\mu(1)} \\
&= \mu(1)^{-1} \mu(A) \mu\left(Y_{j+1}^{j+n-1} \neq 0^{n-1}, B\right) \\
&\leq \mu(1)^{-1} \mu(A) \mu(B).
\end{aligned}$$

□

Proof of Proposition 2.2.2. Since $Y_m = \mathbb{1}\{X_m = 0\}$ for all $m \geq 0$ and the indicator function is mensurable, for every $i \leq j$ we have $\mathscr{G}_i^j \subset \mathscr{F}_i^j$. Then, any supremum taken over $(\mathbb{N}, \mathscr{G}_0^i, \mathscr{G}_{i+n}^\infty)$ must be less than or equal to the same supremum taken over $(\mathbb{N}, \mathscr{F}_0^i, \mathscr{F}_{i+n}^\infty)$. This ends the proof. □

Proof of Theorem 2.2.3. The statement $(a) \Rightarrow (b)$ is a direct consequence of the Proposition 2.2.2; the statement $(b) \Rightarrow (c)$ is immediate. So we start by proving $(c) \Rightarrow (d)$.

For $r \geq 0$ and $s \geq 1$ we take $A = \{Y_0^r = 10^r\}$ and $B = \left\{Y_{r+\lfloor s/2 \rfloor+1}^{r+s} = 0^{\lceil s/2 \rceil}\right\}$. Then we get:

$$\frac{\sigma_{r+s}}{\sigma_r} = \mu_A\left(Y_{r+1}^{r+s} = 0^s\right) \leq \mu_A(B) \leq \mu\left(0^{\lceil s/2 \rceil}\right) + \phi\left(\lfloor s/2 \rfloor + 1\right).$$

Since $\mathbf{Y}$ is ϕ-mixing, there exist $s_0 \in \mathbb{N}$ large enough such that

$$R_0 = \mu\left(0^{\lceil s_0/2 \rceil}\right) + \phi\left(\lfloor s_0/2 \rfloor + 1\right) < 1.$$

For $s \geq s_0$, we write $s = qs_0 + r$ where $q \in \mathbb{N}$ and $0 \leq r < s_0$. Thus,

$$\frac{\sigma_{r+s}}{\sigma_r} \leq q_r \cdots q_{r+s_0-1} q_{r+s_0} \cdots q_{r+2s_0-1} \cdots q_{r+qs_0-1} \leq R_0^q \leq (R_0)^{s/s_0-1}.$$

By adjusting the constant C to cover the case $s < s_0$, we get (d).

Now we assume that (d) holds and take $n_0 \geq 1$ such that $R = C^{1/n_0} e^{-c} < 1$. Then for all $r \geq 0$, we have

$$q_r \cdots q_{r+n_0-1} \leq Ce^{-cn_0} = R^{n_0},$$

which implies $\min\{q_r, \cdots, q_{r+n_0-1}\} \leq R$.

Finally, let us prove that $(e) \Rightarrow (a)$. According to Theorem 3.4 in (BRADLEY, 2005), for $\mathbf{X}$ to be exponentially ϕ-mixing, we just need to show that $\phi(n_0) < 1$. In this sense, it is sufficient to consider cylinders that generate $\mathscr{F}_0^i$ and $\mathscr{F}_{i+n_0}^{\infty}$ for every $i \geq 0$. Thus, let $A_i = \{X_0^i = a_0^i\}$ such that $\bar{\mu}(A_i) > 0$ and $B_i = \left\{X_{i+n_0}^{i+n_0+k} = b_0^k\right\}$. Then, we have

$$\bar{\mu}(B_i) = \bar{\mu}(b_0)\prod_{j=0}^{k-1} Q(b_j, b_{j+1}) \quad \text{and} \quad \bar{\mu}_{A_i}(B_i) = \bar{\mu}\left(X_{i+n_0} = b_0 | X_i = a_i\right)\prod_{j=0}^{k-1} Q(b_j, b_{j+1}),$$

which implies

$$\phi(n_0) = \sup_{i \geq 0, A_i, B_i} |\bar{\mu}_{A_i}(B_i) - \bar{\mu}(B_i)| \leq \sup_{r,s \in \mathbb{N}} |\bar{\mu}\left(X_{n_0} = r | X_0 = s\right) - \bar{\mu}(r)|.$$

Consider the case in which $r < n_0$. Then

$$\begin{aligned} &\bar{\mu}(0)\sigma_{n_0} \leq \bar{\mu}(0)\sigma_r = \bar{\mu}(r) \leq \bar{\mu}(0) \\ \Rightarrow &|\bar{\mu}\left(X_{n_0} = r | X_0 = s\right) - \bar{\mu}(r)| \leq \max\{\bar{\mu}(0), 1 - \bar{\mu}(0)\sigma_{n_0}\}. \end{aligned}$$

On the other hand, if $r \geq n_0$, then

$$\bar{\mu}\left(X_{n_0} = r | X_0 = s\right) = \begin{cases} 0, & \text{if } s \neq r - n_0 \\ \prod_{j=r-n_0}^{r-1} q_j & \text{if } s = r - n_0. \end{cases}$$

Thus, using (e), we get $\bar{\mu}\left(X_{n_0} = r | X_0 = s\right) \leq R$. This gives us

$$|\bar{\mu}\left(X_{n_0} = r | X_0 = s\right) - \bar{\mu}(r)| \leq \max\{R, \bar{\mu}(0)\},$$

since $\bar{\mu}(r) = \bar{\mu}(0)\sigma_r \leq \bar{\mu}(0)$.

Therefore we conclude that

$$\phi(n_0) \leq \max\{R, \bar{\mu}(0), 1 - \bar{\mu}(0)\sigma_{n_0}\} < 1,$$

which ends the proof. $\square$

Proof of Corollary 2.2.5. If $q_i \longrightarrow 1$, we apply Item (e) of Theorem 2.2.3 to conclude that $\mathbf{X}$ and $\mathbf{Y}$ are not ϕ-mixing. On the other hand, note that $\limsup_i q_i < 1$ implies $S = \sup_i\{q_i\} < 1$, which satisfies the same condition (e) with $n_0 = 1$. $\square$

The next lemma constitutes a part of the proof of Theorem 2.2.6.

Lemma 2.4.4. Let $(q_i)_{i \geq 0}$ be a real valued sequence satisfying $0 < q_i < 1$ for all $i \geq 0$ and $\sigma_n = q_0 \cdots q_{n-1} \xrightarrow{n} 0$. Suppose that there exists $n_0 \in \mathbb{N}$ such that

$$\sup_{r,s \in \mathbb{N}} \frac{\sigma_{r+s+n_0}}{\sigma_r \sigma_s} < \infty.$$

Then, there exist $n_1 \geq 1$ and $R < 1$ satisfying $\min\{q_r, \cdots, q_{r+n_1-1}\} \leq R$ for all $r \geq 0$.

Proof. By hypothesis, there exists a constant $M>1$ such that

$$\frac{\sigma_{r+s+n_0}}{\sigma_r\sigma_s}\le M,\quad \forall r,s,\in\mathbb{N}.$$

On the other hand, we can get $s_0\in\mathbb{N}$ such that $\sigma_{s_0}<M^{-2}$. Thus, for all $r\in\mathbb{N}$ we have

$$\begin{aligned} M^2\frac{\sigma_{r+s_0+n_0}}{\sigma_r} &< \frac{\sigma_{r+s_0+n_0}}{\sigma_r\sigma_{s_0}}\le M \\ \Rightarrow q_r\cdots q_{r+s_0+n_0-1} &< \frac{1}{M}. \end{aligned}$$

Now, suppose that the lemma's conclusion does not hold. Then, there exist $r_0\in\mathbb{N}$ such that $\min\{q_{r_0},\cdots,q_{r_0+s_0+n_0-1}\}>M^{-1/(s_0+n_0)}$. Thus, we must have

$$\frac{1}{M}<q_{r_0}\cdots q_{r_0+s_0+n_0-1}<\frac{1}{M},$$

which is a contradiction. □

Proof of Theorem 2.2.6. Let us first show that condition (2.1) is sufficient for $\mathbf{Y}$ to be ψ-mixing. According to Lemma 2.4.4, this condition implies that there exist $n_1\ge 1$ and $R<1$ such that $\min\{q_r,\cdots,q_{r+n_1-1}\}\le R$ for all $r\ge 0$. Define $n_2:=\max\{n_0+1,2n_1+1)\}$.

In order to apply Theorem 4.1 in (BRADLEY, 2005), we will show that $\psi'(n_2+1)>0$ and $\psi^*(n_2+1)<\infty$ (see Definition 2.3.1), which implies that $\mathbf{Y}$ is ψ-mixing. In this sense, it is sufficient to consider cylinders that generate $\mathscr{G}_0^i$ and $\mathscr{G}_{i+n_2+1}^\infty$ for every $i\ge 0$.

Let us prove that $\psi'(n_2+1)>0$. We start by noting that, by Lemma 2.4.4, condition (2.1) implies condition (e) (and therefore (d)) of Theorem 2.2.3, which gives us

$$\sigma_n\le\Sigma(n)=\sigma_n\left(\sum_{j=0}^{\infty}\frac{\sigma_{n+j}}{\sigma_n}\right)\le\sigma_n\left(\sum_{j=0}^{\infty}Ce^{-cj}\right)=D\sigma_n \tag{2.2}$$

where $D=C/(1-e^{-c})$.

We divide the proof in three cases.

1) Consider the case $A=\{Y_0^i=0^{i+1}\}$ and $B=\left\{Y_{i+n_2+1}^{i+n_2+\ell}=0^\ell\right\}$. Then, applying (2.2) we get

$$\frac{\mu(A\cap B)}{\mu(A)\mu(B)}\ge\frac{\mu(A,Y_{i+1}^{i+n_2}=1^{n_2},B)}{\mu(A)\mu(B)}=\frac{\sigma_{i+1}(1-q_0)^{n_2-1}\sigma_\ell}{\mu(1)\Sigma(i+1)\Sigma(\ell)}\ge\frac{(1-q_0)^{n_2-1}}{D^2\mu(1)}.$$

2) Now we consider the case $A=\{Y_0^i=0^{i+1}\}$ and $B=\left\{Y_{i+n_2+1}^{i+n_2+\ell}=0^k1y_0^{\ell-k-2}\right\}$. By the definition of n_2, there exists $j_0 = j_0(k) := \min_{0\le j\le n_2-1}\{j; q_{j+k} \le R\}$. In this case we have

$$\begin{aligned}
\frac{\mu(A\cap B)}{\mu(A)\mu(B)} &\ge \frac{\mu(A, Y_{i+1}^{i+n_2} = 1^{n_2-j_0}0^{j_0}, B)}{\mu(A)\mu(B)} \\
&\ge \frac{(1-q_0)^{n_2-j_0-1}}{D\mu(1)} q_k \cdots q_{k+j_0-1}(1-q_{k+j_0}) \\
&\ge \frac{(1-q_0)^{n_2-1}}{D\mu(1)} R^{j_0}(1-R) \\
&\ge \frac{(1-q_0)^{n_2-1}}{D\mu(1)} R^{n_2-1}(1-R).
\end{aligned}$$

This argument can also be applied to the case $A = \left\{Y_0^i = x_0^{i-h-1}10^h\right\}$ and $B = \left\{Y_{i+n_2+1}^{i+n_2+\ell} = 0^\ell\right\}$, since μ is reversible (see Remark 2.2.1).

3) At last, we consider the case $A = \left\{Y_0^i = x_0^{i-h-1}10^h\right\}$ and $B = \left\{Y_{i+n_2+1}^{i+n_2+\ell} = 0^k 1 y_0^{\ell-k-2}\right\}$. In the same way as above, there exist $g_0 = g_0(h) := \min_{0\le g\le \lfloor n_2/2\rfloor -1}\{g; q_{h+g} \le R\}$ and $g_1 = g_1(k) := \min_{0\le g\le \lfloor n_2/2\rfloor -1}\{g; q_{k+g} \le R\}$, which gives us

$$\begin{aligned}
\frac{\mu(A\cap B)}{\mu(A)\mu(B)} &\ge \frac{\mu(A, Y_{i+1}^{i+n_2} = 0^{g_0}1^{n_2-g_0-g_1}0^{g_1}, B)}{\mu(A)\mu(B)} \\
&= \frac{(1-q_0)^{n_2-g_0-g_1-1}}{\mu(1)} q_h \cdots q_{h+g_0-1}(1-q_{h+g_0}) q_k \cdots q_{k+g_1-1}(1-q_{k+g_1}) \\
&\ge \frac{(1-q_0)^{n_2-1}}{\mu(1)} R^{g_0+g_1}(1-R)^2 \\
&\ge \frac{(1-q_0)^{n_2-1}}{\mu(1)} R^{n_2-2}(1-R)^2.
\end{aligned}$$

Thus we conclude that $\psi'(n_2+1) > 0$.

Let us prove now that $\psi^*(n_2+1) < \infty$. Since ψ^* is decreasing, it is sufficient to prove that $\psi^*(n_0+1) < \infty$. Applying Lemma 2.4.3, we have that for all $i \ge 0$, $A \in \mathscr{G}_0^i$ and $B \in \mathscr{G}_{i+n_0+1}^\infty$:

$$\begin{aligned}
\frac{\mu(A\cap B)}{\mu(A)\mu(B)} &= \frac{\mu(A, Y_{i+1}^{i+n_0} \ne 0^{n_0}, B)}{\mu(A)\mu(B)} + \frac{\mu(A, Y_{i+1}^{i+n_0} = 0^{n_0}, B)}{\mu(A)\mu(B)} \\
&\le \mu(1)^{-1} + \frac{\mu(A, Y_{i+1}^{i+n_0} = 0^{n_0}, B)}{\mu(A)\mu(B)},
\end{aligned}$$

which means that the supremum over the last term determines whether $\psi^*(n_0+1) < \infty$ or not.

In order to obtain an upper bound for this supremum, consider the same three cases as above and denote $L = \sup_{r,s\in\mathbb{N}} \frac{\sigma_{r+s+n_0}}{\sigma_r\sigma_s} < \infty$.

1) $A = \left\{Y_0^i = 0^{i+1}\right\}$ and $B = \left\{Y_{i+n_0+1}^{i+n_0+\ell} = 0^\ell\right\}$:

$$\frac{\mu(A, Y_{i+1}^{i+n_0} = 0^{n_0}, B)}{\mu(A)\mu(B)} = \frac{\Sigma(i+n_0+\ell+1)}{\mu(1)\Sigma(i+1)\Sigma(\ell)} \leq \frac{D\sigma_{i+n_0+\ell+1}}{\mu(1)\sigma_{i+1}\sigma_\ell} \leq \frac{DL}{\mu(1)}.$$

2) $A = \left\{Y_0^i = 0^{i+1}\right\}$ and $B = \left\{Y_{i+n_0+1}^{i+n_0+\ell} = 0^k 1 y_0^{\ell-k-2}\right\}$ (which includes the case $A = \left\{Y_0^i = x_0^{i-h-1}10^h\right\}$ and $B = \left\{Y_{i+n_0+1}^{i+n_0+\ell} = 0^\ell\right\}$, since μ is reversible):

$$\frac{\mu(A, Y_{i+1}^{i+n_0} = 0^{n_0}, B)}{\mu(A)\mu(B)} = \frac{\sigma_{i+n_0+k+1}}{\mu(1)\Sigma(i+1)\sigma_k} \leq \frac{\sigma_{i+n_0+k+1}}{\mu(1)\sigma_{i+1}\sigma_k} \leq \frac{L}{\mu(1)}$$

3) $A = \left\{Y_0^i = x_0^{i-h-1}10^h\right\}$ and $B = \left\{Y_{i+n_0+1}^{i+n_0+\ell} = 0^k 1 y_0^{\ell-k-2}\right\}$:

$$\frac{\mu(A, Y_{i+1}^{i+n_0} = 0^{n_0}, B)}{\mu(A)\mu(B)} = \frac{\sigma_{h+n_0+k}(1-q_{h+n_0+k})}{\mu(1)\sigma_h\sigma_k} \leq \frac{L}{\mu(1)}$$

Therefore, $\psi^*(n_0+1) < \infty$ and $\mathbf{Y}$ is ψ-mixing.

On the other hand, we assume that $\mathbf{Y}$ is ψ-mixing. Since $\mathbf{Y}$ is also ϕ-mixing, Theorem 2.2.3 implies that condition (d) holds, which also gives us the inequalities in (2.2).

Furthermore, if $\mathbf{Y}$ is ψ-mixing, then we have $\psi^*(n) \xrightarrow{n} 1$ (see (BRADLEY, 2005)), which means that there exists $n_0 \in \mathbb{N}$ such that $\psi^*(n_0+1) < \infty$.

Hence, if we take $A = \left\{Y_0^r = 10^r\right\}$ and $B = \left\{Y_{r+n_0+1}^{r+n_0+s} = 0^s\right\}$, we get:

$$\begin{aligned}
\psi^*(n_0+1) &\geq \sup_{r,s\in\mathbb{N}} \frac{\mu(A\cap B)}{\mu(A)\mu(B)} \\
&\geq \sup_{r,s\in\mathbb{N}} \frac{\mu(A, Y_{r+1}^{r+n_0} = 0^{n_0}, B)}{\mu(A)\mu(B)} \\
&= \sup_{r,s\in\mathbb{N}} \frac{\sigma_{r+s+n_0}}{\mu(1)\sigma_r\Sigma(s)} \\
&\geq \frac{1}{D\mu(1)} \sup_{r,s\in\mathbb{N}} \frac{\sigma_{r+s+n_0}}{\sigma_r\sigma_s}.
\end{aligned}$$

This ends the proof. □

Proof of Corollary 2.2.7. By hypothesis, we have

$$\frac{\sigma_{r+s}}{\sigma_r\sigma_s} = \prod_{j=0}^{s-1} \frac{q_{r+j}}{q_j} \leq 1$$

and the ψ-mixing follows from Theorem 2.2.6. □

Proof of Proposition 2.3.2. If the condition in (a) holds, Theorem 2.2.6 ensures that $\mathbf{Y}$ is ψ-mixing. Furthermore, the condition in (b) implies that there exists $L_1 \in (0,1)$ such that

$$L_1 = \inf_{r,s\in\mathbb{N}} \frac{\sigma_{r+s}(1-q_{r+s})}{\sigma_r\sigma_s} \leq \inf_{r\in\mathbb{N}} \frac{\sigma_r(1-q_r)}{\sigma_r\sigma_0},$$

which implies $q_r \leq 1-L_1$ for all $r\in\mathbb{N}$. Hence, by Corollary 2.2.5, $\mathbf{Y}$ is exponentially ϕ-mixing.

Now, let us show that conditions in (a) and (b) imply $\psi^*(1)<\infty$ and $\psi'(1)>0$, respectively. First, we denote $L_2 = \sup_{r,s\in\mathbb{N}} \frac{\sigma_{r+s}}{\sigma_r\sigma_s}$.

Observe that by Theorem 2.2.3 and (2.2), there exists a positive constant D such that $\sigma_n \leq \Sigma(n) \leq D\sigma_n$. Now, let us consider three cases. If $a_0^{m+n-1} = 0^{m+n}$, we have

$$\begin{aligned}
&\frac{\sigma_{m+n}(1-q_{m+n})}{\mu(1)D^2\sigma_m\sigma_n} \leq \frac{\Sigma(m+n)}{\mu(1)\Sigma(m)\Sigma(n)} \leq \frac{D\sigma_{m+n}}{\mu(1)\sigma_m\sigma_n}\\
\Rightarrow &\frac{L_1}{\mu(1)D^2} \leq \frac{\mu\left(a_0^{m+n-1}\right)}{\mu\left(a_0^{m-1}\right)\mu\left(a_m^{m+n-1}\right)} \leq \frac{DL_2}{\mu(1)}
\end{aligned}$$

For the case $a_0^{m-1} = 0^m$ and $a_m^{m+n-1} = 0^k1a_{m+k+1}^{m+n-1}$, we have

$$\begin{aligned}
&\frac{\sigma_{m+k}(1-q_{m+k})}{\mu(1)D\sigma_m\sigma_k} \leq \frac{\sigma_{m+k}}{\mu(1)\Sigma(m)\sigma_k} \leq \frac{\sigma_{m+k}}{\mu(1)\sigma_m\sigma_k}\\
\Rightarrow &\frac{L_1}{\mu(1)D} \leq \frac{\mu\left(a_0^{m+n-1}\right)}{\mu\left(a_0^{m-1}\right)\mu\left(a_m^{m+n-1}\right)} \leq \frac{L_2}{\mu(1)}.
\end{aligned}$$

Note that the inequalities above include the case $a_0^{m-1} = a_0^{m-k-2}10^k$ and $a_m^{m+n-1} = 0^n$, since μ is reversible (see Remark 2.2.1).

At last, when $a_0^{m-1} = a_0^{m-k-2}10^k$ and $a_m^{m+n-1} = 0^j1a_{m+j+1}^{m+n-1}$, we have

$$\frac{L_1}{\mu(1)} \leq \frac{\sigma_{k+j}(1-q_{k+j})}{\mu(1)\sigma_k\sigma_j} = \frac{\mu\left(a_0^{m+n-1}\right)}{\mu\left(a_0^{m-1}\right)\mu\left(a_m^{m+n-1}\right)} \leq \frac{L_2}{\mu(1)}$$

Therefore, the condition in (a) implies $\psi^*(1)<\infty$ and the condition in (b) implies $\psi'(1)>0$. $\square$

Proof of Proposition 2.3.4. Since the process is ψ-mixing, given $\varepsilon>0$, there exists $g_0\in\mathbb{N}$ large enough such that

$$1-\varepsilon \leq \psi'(g_0+1) \leq \psi_{g_0}^-(m,n) \leq \psi_{g_0}^+(m,n) \leq \psi^*(g_0+1) \leq 1+\varepsilon,$$

for all $m,n\in\mathbb{N}$, which implies that the process is ψ_{g_0}-regular. $\square$

CHAPTER

3

EXPONENTIAL APPROXIMATION FOR RECURRENCE TIMES IN MIXING PROCESS

The first section is devoted to definitions and some initial considerations on exponential approximations for recurrence times. In Section 3.2, we present a review of classical and recent results on hitting and return times of cylinder sets, focusing on stochastic mixing processes and results that hold for *all points* of the space $\mathscr{X}$. Then, in Section 3.3, we present in Theorem 3.3.4 the main result of the chapter, giving new exponential approximations for hitting and return time distribution. This result was published (ABADI; AMORIM; GALLO, 2021). In the last section, we present some technical results and the proofs of the main theorems.

3.1 Preliminary definitions and considerations

In this and the next section, we present a review of the literature. It is based on (ABADI; AMORIM; GALLO, 2021) and (ABADI; GALVES, 2000). At first, let us define the two main random variables of interest of this chapter: the *hitting time* and *return time* to a string $A \in \mathscr{C}_n$.

Definition 3.1.1. The hitting time of a point $x \in \mathscr{X}$ to a cylinder $A = \left[a_0^{n-1}\right]$ is defined by

$$T_A(x) = \inf\left\{k \geq 1; \sigma^k(x) \in A\right\}.$$

Alternatively, one can define the hitting time to A as a random variable that is a function of a process $\mathbf{X}$:

$$T_A = \inf\left\{k \geq 1; X_k^{k+n-1} = a_0^{n-1}\right\}.$$

When $x \in \left[a_0^{n-1}\right]$, we call $T_A(x)$ the return time to A.

Thus, our main interest is to obtain informations about $\mu(T_A > t)$ and $\mu_A(T_A > t)$, that is, the hitting and return times distributions, respectively. The first classical result about it, known as Poincaré recurrence theorem, ensures that for any stationary process the return time is almost surely finite (see for instance (SHIELDS, 1996)), in other words, $\mu_A(T_A < \infty) = 1$.

There is a similar result for the hitting time, but we need to add the condition of ergodicity to the process. We say that a stationary measure μ (or process $\mathbf{X}$) is ergodic if, for $A \in \mathscr{F}$, $\sigma^{-1}(A) = A$ implies $\mu(A) = 0$ or $\mu(A) = 1$. For instance, it is a well known fact that every α-mixing process is ergodic. An extension of Poincaré recurrence theorem for ergodic processes says that, as well as return time, the hitting time satisfies $\mu(T_A < \infty) = 1$.

The first quantitative result on return time is the well known Kac's Lemma, which provides for ergodic processes an exact and intuitive formula to the mean return time:

$$\mathbb{E}(T_A|A) = \frac{1}{\mu(A)}.$$

However, we do not have a similar result for hitting time. Since there are no explicit expressions for $\mu(T_A > t)$ and $\mu_A(T_A > t)$, even for the case of an i.i.d. process and words A with $n > 1$, a natural interest arises to obtain approximated distributions or even limit distributions when n diverges.

We are interested in the case where we fix *any* point $x \in \mathscr{X}$ and consider $A = A_n(x) = x_0^{n-1}$ as target set. For ergodic and aperiodic processes, we will prove in Lemma 3.4.1 that the measure of $A_n(x)$ vanishes as n diverges. The nested sequence of sets $A_n(x)$, $n \geq 1$ is therefore called a *sequence of rare events*.

When dealing with rare events, and sufficiently strong mixing conditions, it is a well known fact that $\mu(T_A > t)$ has approximate exponential distributions with parameter depending on A, as well as $\mu_A(T_A > t)$ is well approximated by a convex combination between a Dirac function and an exponential law, as we can see in (ABADI; GALVES, 2000), (ABADI, 2004), (ABADI; VERGNE, 2008), (ABADI; SAUSSOL, 2011), (ABADI; CARDENO; GALLO, 2015). On the other hand, it is important to mention that this is not the case for ergodic processes in general (see (LACROIX, 2002)).

Furthermore, for a *typical* cylinder, $\mu\left(T_A > \mu(A)^{-1}t\right)$ and $\mu_A\left(T_A > \mu(A)^{-1}t\right)$ both converge to an exponential law with parameter one, that is, this convergence holds almost surely, as showed in (COLLET; GALVES; SCHMITT, 1999), (ABADI; SAUSSOL, 2016) and (ABADI; CARDENO; GALLO, 2015), for instance.

Indeed, Kac's Lemma suggests that the best exponential approximation for $\mu\left(T_A > t\right)$ and $\mu_A\left(T_A > t\right)$ would be $e^{-\mu(A)t}$. However, since the paper (GALVES; SCHMITT, 1997), it is known that a correction factor $\theta(A)$ must be used to get the convergence to the exponential law at any point. Naturally, one expects that for a typical cylinder we have

$\theta(A) \approx 1$.

In other words, the approximate distributions obtained in previous works involve the exponential law with parameter $\mu(A)\theta(A)$, where $A \in \mathscr{C}_n$. These works present in general two main types of approximations: a total variation distance type, which we will call Type 1, and a pointwise type, called Type 2. More specifically, these approximations have the form:

- ***Type 1***: *Total variation distance.*
  - Hitting times
  $$\sup_{t>0} \left| \mu(T_A > t) - e^{-\mu(A)\theta(A)t} \right| \leq \varepsilon(A),$$
  - Return times
  $$\sup_{t>0} \left| \mu_A(T_A > t) - \bar{\theta}(A) e^{-\mu(A)\theta(A)t} \right| \leq \varepsilon(A),$$

  where $\varepsilon(A) \xrightarrow{n} 0$.

- ***Type 2***: *Pointwise.* For any $t > 0$,
  - Hitting times
  $$\left| \mu(T_A > t) - e^{-\mu(A)\theta(A)t} \right| \leq \varepsilon(A,t),$$
  - Return times
  $$\left| \mu_A(T_A > t) - \bar{\theta}(A) e^{-\mu(A)\theta(A)t} \right| \leq \varepsilon(A,t),$$

  where $\varepsilon(A,t) \longrightarrow 0$ when n or t diverges.

Note that return time approximation indicates that

$$\mathbb{E}(T_A|A) \approx \bar{\theta}(A) \frac{1}{\mu(A)\theta(A)}.$$

The parameters θ and $\bar{\theta}$ need not to be equal, but in view of Kac's Lemma, the last expression suggests that $\theta \approx \bar{\theta}$. The approximations also indicate that

$$\mathbb{E}(T_A) \approx \frac{1}{\mu(A)\theta(A)}.$$

Thus, if we can compute $\theta(A)$, these approximations allow us to show whether the scaling of occurrence of return time is close to hitting time or not, depending on whether the parameter $\theta(A)$ approaches 0 or not. In some cases, the computation of $\theta(A)$ can even give us the limit distributions.

3.2 A review of related results

We present below a review of results concerning hitting and return time approximations in the context of mixing processes and whose results hold for all points. In each case, we highlight the characteristics of the scaling parameter used and the type of approximation, since our results present advances in this sense.

As far as we know Aldous and Brown (1993) published the first paper concerning exponential approximations for hitting time to rare events that holds for all points. They provided a Type 1 approximation in the context of the reversible and finite-state Markov chain and the scaling parameter they used is the inverse of the mean hitting time, which is expected when we deal with exponential law. However, this is an inaccessible parameter and does not bring information about the value of $\mathbb{E}(T_A)$.

In (GALVES; SCHMITT, 1997), the authors provide a Type 1 approximation for hitting time in summable ψ-mixing processes. This was the first paper to give an explicit formula for the scaling parameter, denoted there by $\lambda(A)$. Nonetheless, it has no intuitive significance and it is not easy to compute, since its calculation depends itself on knowing the hitting time distribution up to a large scale (but smaller than the reciprocal of the measure of the observable).

In the context of ψ-mixing and summable ϕ-mixing processes with finite alphabet, Abadi (2001) presented Type 1 approximations for the hitting time. The paper gives two different scaling parameters, one of which is similar to $\lambda(A)$, which depends on the scale of the order of $\mu(A)^{-1}$. In order to overcome this issue, the same paper provides another scaling parameter $\zeta(A)$, for which the calculation depends on the short return of the string A, making it easier to compute. However, its use leads to a larger error term for the approximation.

In the same context as above, Abadi (2004) obtained a Type 2 approximation theorem for the hitting time. The scaling parameter was still similar to $\lambda(A)$, but the error term $\varepsilon(A,t)$ in this case decreases exponentially fast in t. The paper also provides an approximate version of the Kac's Lemma for the rescaled mean hitting time $\mathbb{E}(\mu(A)T_A)$, which depends on the scaling parameter. Furthermore, a uniform sharp error term was provided, whose calculation also depends on the large scale of $\mu(A)^{-1}$.

Abadi (2006) presented Type 1 approximations for α-mixing processes with countable alphabets. The author used $\lambda(A)$ as scaling parameter and a new parameter $\zeta'(A)$ as the weight of the convex combination for the return time (the $\bar{\theta}(A)$ mentioned above). Both parameters depend on recurrence properties at the scale $\mu(A)^{-1}$. Nonetheless, the paper also introduces an intuitive and easily computable parameter, which will be called *potential well* and denoted by $\rho(A)$. He manages to prove that, for exponentially α-mixing processes (when $\alpha(n)$ vanishes exponentially fast), λ and ζ can be well approximated by

ρ.

In (ABADI; VERGNE, 2008), the authors provided the first Type 2 approximation for the return time. In the context of ϕ-mixing processes and finite alphabet, they used for the first time the potential well $\rho(A)$ as scaling parameter and weight of the convex combination at the same time, that is, $\bar{\theta} = \theta = \rho$. The error term $\varepsilon(A,t)$ presented there decreases exponentially fast in t and the uniform error term $\varepsilon(A)$ is sharp. The authors also gave an approximation for the rescaled moments $\mathbb{E}_A\left((\mu(A)T_A)^\beta\right)$ of the return time and extended a result in (HIRATA; SAUSSOL; VAIENTI, 2000), stating that hitting and return times are arbitrarily close if, and only if, the rescaled return time converges to an exponential law with parameter one. In all of their results, the process is assumed to have complete grammar.

Abadi and Saussol (2011) obtained Type 1 approximations for hitting and return times for α-mixing processes with finite or countable alphabet. The parameter used also was similar to $\lambda(A)$, but, as well as Abadi and Vergne (2008), they proved the convergence of the return time to convex combination involving exponential law using $\bar{\theta} = \theta$. Later, the authors proved in (ABADI; SAUSSOL, 2016) that, if $\alpha(n)$ decreases algebraically, then the parameter $\lambda(A)$ converges almost surely to one. As a consequence, we have the almost sure convergence of the rescaled hitting and return times to the parameter one exponential law.

Finally, Abadi, Cardeno and Gallo (2015) worked on the specific class of binary renewal processes, which is β-mixing in general (weaker than ϕ), and provided Type 1 approximations for hitting and return times using the potential well ρ as scaling parameter. The simplicity and intuitiveness of the potential well helped the authors to apply the renewal property and obtain an exact formula to calculate $\rho(A_n(x))$ and its limit for any point $x \in \mathscr{X}$. In other words, their results allow us to obtain the limit distributions of hitting and return times as a function of the parameters of the process.

3.3 Main result: hitting and return time approximation theorem

In this section, we present the main result of this chapter: a new hitting and return time approximation theorem, which was published in (ABADI; AMORIM; GALLO, 2021). In the context of Type 2 approximation, we use the potential well as a scaling parameter to obtain exponential approximations for hitting and return times whose error $\varepsilon(A,t)$ term decreases exponentially fast in t. We also show, with an example, the sharpness of the uniform error term $\varepsilon(A)$. In order to have $\sup_{A\in\mathscr{C}_n}\varepsilon(A) \xrightarrow{n} 0$, we assume that the process is ψ-mixing alone or ϕ-mixing with some further technical assumption. Furthermore, unlike similar previous results, our theorem holds for process with infinite alphabet and without complete grammar.

In the case of return times, the result is an extension and a correction of the main theorem in (ABADI; VERGNE, 2008). The extension follows from the fact that we include processes with incomplete grammar and infinite alphabet and we also get a specific and better error term for ψ-mixing processes. On the other hand, we show that the error term presented by Abadi and Vergne is incorrect for small values of t, namely, for $t \leq (2\mu(A))^{-1}$.

Before we can state our main theorem, we need to present some related concepts. This is done in Subsection 3.3.1. The main results will be stated in Subsection 3.3.2.

3.3.1 Overlapping properties, periodicity and potential well

The overlapping properties of the string a_0^{n-1} can strongly influence the distributions of hitting and return time to a_0^{n-1}. For instance, consider a binary renewal process presented in the previous chapter defined by the function $\Sigma(n) = \frac{1}{n}$. If we take $A = a_0^{n-1} = 0^n$, we get

$$\mu_A(T_A = 1) = \frac{\mu\left(0^{n+1}\right)}{\mu\left(0^n\right)} = \frac{n}{n+1} \longrightarrow 1. \tag{3.1}$$

In other words, when n diverges, the density of the distribution of the return time is highly concentrate on $t = 1$, due to the fact that A might overlap itself at $t = 1$. The hitting time distribution is also affected by overlap. For stationary processes we have $\mu(T_A = j) = \mu(A, T_A > j-1)$ (see Lemma 3.4.2). Thus, still using the same example as above with $A = 0^n$,

$$\begin{aligned}
\mu(T_A \leq t) &= \sum_{j=1}^{t} \mu(T_A = j) \\
&= \mu(A) \sum_{j=1}^{t} \mu_A(T_A > j-1) \\
&\leq \mu(A) + (t-1)\mu(A)\mu_A(T_A > 1).
\end{aligned} \tag{3.2}$$

If we take $t = \mu(A)^{-1} + 1$, we get $\mu\left(T_A \leq \mu(A)^{-1}\right) \longrightarrow 0$ when n diverges. In other words, the high probability of return in $t = 1$ implies a low probability of an "early" hitting time in the Kac's scale.

Therefore, in order to obtain approximate distributions of hitting and return time, the string's overlapping properties need to be taken into account. This leads us to define two fundamental concepts: the shortest possible return and the potential well.

Definition 3.3.1. The shortest possible return to a set $A \in \mathscr{F}$ is defined as

$$\tau(A) = \inf\left\{k \geq 1 : \mu_A\left(\sigma^{-k}(A)\right) > 0\right\}.$$

Alternatively, given a string $a_0^{n-1} \in \mathscr{C}_n$ generated by a stationary process $\mathbf{X}$, the shortest return can be defined as

$$\tau\left(a_0^{n-1}\right) = \inf\left\{k \geq 1 : \mu\left(X_k^{k+n-1} = a_0^{n-1} | X_0^{n-1} = a_0^{n-1}\right) > 0\right\}.$$

In the case where $A = A_n(x)$, we use the notations $\tau(A_n)$ and $\tau\left(x_0^{n-1}\right)$. Note that when the grammar is complete, we obviously have $1 \leq \tau(A_n) \leq n$. If it is not the case, we can also obtain upper bounds to $\tau(A_n)$, as we do in Chapter 4.

We observe that the traditional definition of the shortest return, namely $\tau(A) = \inf\left\{k \geq 1; A \cap \sigma^{-k}(A) \neq \emptyset\right\}$, has a topological meaning. However, we adopt the definition involving μ_A in order to deal with processes without complete grammar, since in this case we might have some $A \in \mathscr{C}_n$ and $k \in \mathbb{N}$ satisfying $A \cap \sigma^{-k}(A) \neq \emptyset$ and $\mu\left(A \cap \sigma^{-k}(A)\right) = 0$. For instance, consider the string $A = 12 \cdots n$ generated by the house of cards Markov chain (see Example 2.1.3).

The shortest possible return $\tau(A_n(x))$ has itself independent interest of study in the literature and we refer to (SAUSSOL; TROUBETZKOY; VAIENTI, 2002), (AFRAIMOVICH; CHAZOTTES; SAUSSOL, 2003), (ABADI; CARDENO, 2015), (ABADI; VAIENTI, 2008), (HAYDN; VAIENTI, 2010), (ABADI; LAMBERT, 2013) and (ABADI; GALLO; RADA-MORA, 2018) for several results on asymptotic concentration, large deviations and fluctuations of τ.

Obviously, by definition, $\mu_A(T_A \geq \tau(A)) = 1$. If $T_A(x) > \tau(A)$ for some point $x \in A$, we say that x *escapes* from A. The potential well $\rho(A)$ is defined as the proportional measure of points of A that escape from A.

Definition 3.3.2. The potential well of a set A is defined by

$$\rho(A) = \mu_A(T_A > \tau(A)).$$

For the case where $A = A_n(x)$, we use the notations $\rho(x_0^{n-1})$ and $\rho(A_n(x))$. The potential well $\rho(A)$ is an intuitive and "easy to compute" parameter, which also has a physical meaning. It represents the "energy" the system needs to leave the set A. In other words, $1 - \rho(A)$ represents the height of the barrier the system needs to cross to leave A. It also represents stability characteristics of the set; for instance, $\rho(A) \approx 1$ means that A is an unstable state.

Another aspect we need to know when studying recurrence times is the second order periodicity of a string $A \in \mathscr{C}_n$, which has influence on the size of the uniform error term $\varepsilon(A)$. It is also related to the overlapping properties of the string, but in a different sense: we look at the possible returns up to the end of the string, that is, up to $n-1$.

Consider the cylinder A and denote $\tau(A) = k$. We write $n = qk + r$, where $q \in \mathbb{N}$ and $0 \leq r < k$. When $k < n$, we can verify that $[A] \cap \sigma^{-mk}[A] \neq \emptyset$ for all $1 \leq m \leq q$ and $[A] \cap \sigma^{-j}[A] = \emptyset$ for all $mk \neq j < qk$. That is, up to the index qk, the cylinder overlaps itself only in multiples of k. However, the *first* return time cannot occur in $2k, \cdots, qk$, since the overlap in these times implies $T_A = k$.

On the other hand, it is immediate to verify that the *first* return may occur in the set

$$\mathscr{R}(A) = \left\{ j \in \{qk+1, ..., qk+r-1\} : \mu_A\left(\sigma^{-j}(A)\right) > 0 \right\},$$

which we call the set of the second order periods of A.

Thus, when $\tau(A) < n$, the first return of a string A has positive probability of occurring only at the times belonging to

$$\{\tau(A)\} \cup \{\mathscr{R}(A)\} \cup \left\{k \geq n; \mu_A\left(\sigma^{-k}(A)\right) > 0\right\}.$$

Obviously, when $\tau(A) \geq n$, we do not have second order periods and the first return can occur with positive probability at the times in $\{\tau(A)\} \cup \left\{k > \tau(A); \mu_A\left(\sigma^{-k}(A)\right) > 0\right\}$.

We set n_A as the first possible return to A after $\tau(A)$:

$$n_A = \min\left\{ j : \mu_A\left(\sigma^{-\tau(A)}(A^c) \cap \sigma^{-j}(A)\right) > 0 \right\}.$$

Notice that we may have $n_A > n$, when the grammar is not complete.

In order to illustrate these ideas under complete grammar, we refer to a concrete example and more details in (ABADI; VERGNE, 2008). For the general case, we present the following example.

Example 3.3.3. Consider the house of cards Markov chain defined in Example 2.1.3 in Chapter 2. Recall that it is defined over the infinite alphabet $\mathbb{N}$ and does not have complete grammar, since several transitions are forbidden due to the sparse nature of Q. Consider the strings $A = 00010001000$ and $B = k(k+1)\cdots(k+n-1)$ (for some $k \in \mathbb{N}$ and $n \geq 1$) generated by this Markov chain in the stationary measure. Then, we see that $\tau(A) = 4$, $\mathscr{R}(A) = \{9, 10\}$, and $n_A = 9$, while, on the other hand, $\tau(B) = k+n \geq n$, $\mathscr{R}(B) = \emptyset$, and $n_B = k+n+1$.

3.3.2 New Type 2 approximations for hitting and return times

Theorem 3.3.4 below was published in (ABADI; AMORIM; GALLO, 2021) and presents Type 2 approximations for hitting and return time under ϕ and ψ-mixing conditions with the potential well as the scaling parameter and an explicit error term.

As a Type 2 approximation, the error terms satisfy $\varepsilon(A,t) \longrightarrow 0$ when n or t diverges, and we additionally have that the convergence in t takes place exponentially fast. Furthermore, we have an error term uniform in t, that is, $\varepsilon(A) = \sup_t \varepsilon(A,t)$. We start defining it for the cases ψ and ϕ.

Denote by $A^{(k)}$ the string of the last k symbols of $A_n(x) = x_0^{n-1}$, that is, $A^{(k)} = x_{n-k}^{n-1}$. When $j \geq n$, we use the convention $\mu\left(A^{(j)}\right) = \mu\left(A^{(n)}\right) = \mu(A)$.

Note that by definition $\psi(g)$ is not necessarily finite, for $g \geq 1$. Thus, for ψ-mixing measures, we define:

$$g_0 = g_0(\psi) := \inf\{g \geq 1 : \psi(g) < \infty\} - 1. \tag{3.3}$$

Now, for the uniform error terms, define:

$$(a)\ \varepsilon_\psi(A) := n\mu\left(A^{(n_A - g_0)}\right) + \psi(n), \tag{3.4}$$

$$(b)\ \varepsilon_\phi(A) := \inf_{1 \leq w \leq n^*}\left\{(n+\tau(A))\mu\left(A^{(w)}\right) + \phi(\min\{n_A - w + 1, n\})\right\}, \tag{3.5}$$

where $n^* := \min\{n, n_A\}$.

Observe that we always have $n_A \geq n/2$, then ε_ψ is well defined for all $n > 2g_0$

We use ε to denote either ε_ψ or ε_ϕ when the argument or statement is general.

Theorem 3.3.4. Consider a stationary measure μ on $(\mathscr{X}, \mathscr{F})$ enjoying either ϕ-mixing with $\sup_{A \in \mathscr{C}_n} \mu(A)\tau(A) \xrightarrow{n} 0$ or ψ-mixing. There exist five positive constants C_i, $i = 1, \ldots, 5$, and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and all $A \in \mathscr{C}_n$, the following inequalities hold.

- For all $t \geq 0$:

$$\left|\mu(T_A > t) - e^{-\rho(A)\mu(A)t}\right| \leq \begin{cases} C_1\left(\tau(A)\mu(A) + t\mu(A)\varepsilon(A)\right) & t \leq [2\mu(A)]^{-1} \\ C_2\,\mu(A)t\varepsilon(A)e^{-\mu(A)t(\rho(A) - C_3\varepsilon(A))} & t > [2\mu(A)]^{-1}. \end{cases}$$

- For all $t \geq \tau(A)$:

$$\left|\mu_A(T_A > t) - \rho(A)e^{-\rho(A)\mu(A)(t-\tau(A))}\right|$$

$$\leq \begin{cases} C_4\,\varepsilon(A) & t \leq [2\mu(A)]^{-1} \\ C_5\,\mu(A)t\varepsilon(A)e^{-\mu(A)t(\rho(A) - C_3\varepsilon(A))} & t > [2\mu(A)]^{-1}. \end{cases}$$

Theorem 3.3.4 and its proof were inspired by the ideas of Theorem 4.1 in (ABADI; VERGNE, 2008). Nonetheless, our result is the first Type 2 approximation for the hitting time using the potential as scaling parameter. Moreover, we state both hitting and return times approximations in the broader context of countable alphabets and with not necessarily complete grammar, besides providing a specific and smaller error term for the ψ-mixing case.

The proof of Theorem 3.3.4 is given in Section 3.4 after some preliminary results. We now present some important and complementary remarks about it.

Remark 3.3.5. Although Theorem 3.3.4 provides Type 2 approximations for hitting and return times distributions, it might be useful to have an uniform error term in t, which is, in this case, $C\varepsilon(A)$, for some constant C. Indeed, in the case of hitting time and $t \leq (2\mu(A))^{-1}$, for n large enough, Lemma 3.4.4 implies $\tau(A)\mu(A) \leq 2\varepsilon_\psi(A)$. Thus, in both cases ϕ and ψ-mixing, we have $\tau(A)\mu(A)+t\mu(A)\varepsilon(A) \leq C'\varepsilon(A)$, for some constant C'. Obviously, the same holds when we analyse the case $t > (2\mu(A))^{-1}$ and the return time.

Remark 3.3.6. Under the conditions of Theorem 3.3.4, we can show that $\sup_{A\in\mathscr{C}_n}\varepsilon(A) \xrightarrow{n} 0$. Indeed, in the context of ϕ-mixing processes, Lemma 3.4.3 gives us $\mu(A) \leq Ce^{-cn}$ for all $n \geq 1$ and $A \in \mathscr{C}_n$ and some positive constants C and c. On the other hand, since $n_A \geq n/2$, we get for all $n > 2g_0$:

$$\varepsilon_\psi(A) \leq Cne^{-c(n/2-g_0)} + \psi(n)$$

and hence $\varepsilon_\psi \longrightarrow 0$, uniformly.

For the ϕ-mixing error term, we consider two cases. If $\tau(A) \leq 2n$, we take $w = \lceil n/4 \rceil$ and obtain

$$\varepsilon_\phi(A) \leq 3Cne^{-cn/4} + \phi(\lfloor n/4 \rfloor),$$

which ensures that $\varepsilon_\phi \longrightarrow 0$. This first case occurs, for instance, if one has a complete grammar. On the other hand, if $\tau(A) > 2n$, we take $w = n$. Note that $2n < \tau(A) < n_A$. Thus,

$$\varepsilon_\phi(A) \leq (n+\tau(A))\mu(A) + \phi(n) \leq Cne^{-cn} + \phi(n) + \sup_{A\in\mathscr{C}_n}\tau(A)\mu(A) \longrightarrow 0,$$

since the last term in the sum converges to zero by hypothesis.

Remark 3.3.7. Under certain conditions, we can drop the hypothesis $\sup_{A\in\mathscr{C}_n}\mu(A)\tau(A) \xrightarrow{n} 0$. For instance, under ϕ-mixing and complete grammar, Lemma 3.4.3 ensures $\tau(A)\mu(A) \leq Cne^{-cn}$ for all $A \in \mathscr{C}_n$. Another way is to assume that μ is summable ϕ-mixing, as shown in Theorem 4.4.3 and Remark 4.4.6.

Remark 3.3.8. Although the approximations under ψ-mixing are less general, the simplicity of the uniform error term $\varepsilon_\psi(A)$, when compared to $\varepsilon_\phi(A)$, justifies its statement. On the other hand, for each one of the cases, ψ and ϕ, the error term is the same (except for a constant) for $t > [2\mu(A)]^{-1}$ for hitting and return times approximations. The difference occurs for $t \leq [2\mu(A)]^{-1}$, which is due the the conditional measure μ_A and to the overlapping properties of the string A.

Remark 3.3.9. For application purposes, explicit values were given to each constant involved in Theorem 3.3.4, and can be found in (ABADI; AMORIM; GALLO, 2021).

Example 3.3.10. Let us show with a simple example the sharpness of the error term in the hitting and return time approximations given by Theorem 3.3.4. Consider an i.i.d. process $(X_m)_{m\in\mathbb{N}}$ on $\mathscr{A}$ and take $b \in \mathscr{A}$ such that $\mu(b) = p$. Consider the point $x = b^\infty \in \mathscr{A}^\mathbb{N}$ and the string $A = A_n(x) = b\cdots b$.

On the one hand, direct calculations give us: $\mu(A) = p^n$, $\tau(A) = 1$, $\rho(A) = 1-p$, $n_A = n$ and $\mu_A(T_A > n-1) = \rho(A) = 1-p$. Furthermore, an i.i.d. process is ψ-mixing with

$\psi(n) = 0$ for all $n \geq 1$, which implies $\varepsilon(A) = np^n$. By direct substitution we have for each $n \geq 2$:

$$\begin{aligned}\left|\mu_A(T_A > n-1) - \rho(A)e^{-\rho(A)\mu(A)((n-1)-\tau(A))}\right| &= \left|(1-p) - (1-p)e^{-(1-p)p^n((n-1)-1)}\right| \\ &= (1-p)\left(1 - e^{-(1-p)p^n(n-2)}\right),\end{aligned}$$

which implies that the exact error in the approximation for the return time at $n-1$ is of order np^n, just as stated by Theorem 3.3.4.

On the other hand, applying (3.2) for $t = n-1$, we obtain

$$\begin{aligned}\mu(T_A \leq n-1) &= \mu(A) + \mu(A)\sum_{j=2}^{n-1}\mu_A(T_A > j-1) \\ &= \mu(A)(1 + (n-2)\rho(A)) \\ &= p^n(1 + (n-2)(1-p)).\end{aligned}$$

Thus,

$$\left|\mu(T_A > n-1) - e^{-\rho(A)\mu(A)(n-1)}\right| = \left|1 - p^n(1 + (n-2)(1-p)) - e^{-(1-p)p^n(n-1)}\right|.$$

Since $1 - e^{-(1-p)p^n(n-1)} \sim (1-p)p^n(n-1)$[1], the exact error term above is of the order p^{n+1}. Furthermore, Theorem 3.3.4 states that, for $t = n-1$, the error term for the hitting time approximation is of the order

$$\varepsilon(A, n-1) = \tau(A)\mu(A) + (n-1)\mu(A)\varepsilon(A) = p^n + (n-1)np^{2n} \sim p^n.$$

Therefore,

$$\left|\mu(T_A > n-1) - e^{-\rho(A)\mu(A)(n-1)}\right| \sim \varepsilon(A, n-1).$$

Remark 3.3.11. In Theorem 3.3.4, the error term for the return time and for $t \leq [2\mu(A)]^{-1}$ corrects Theorem 4.1 in (ABADI; VERGNE, 2008). The previous example shows that our error term cannot be sharpened in general, and that their statement is indeed incorrect.

Remark 3.3.12. The house of cards Markov chain, under the conditions of Theorem 2.2.3, provides an example of a summable ϕ-mixing process over an infinite alphabet and without complete grammar. In other words, this is a process where Theorem 3.3.4 can be applied, unlike previous results.

We emphasize that, as well as in (ABADI; VERGNE, 2008), the potential well has a crucial role in the hitting and return time distributions, since it is the scaling parameter and the weight of the convex combination. Thus, the asymptotic behaviour and the positivity of $\rho(A)$ become fundamental in recurrence times. In this sense, we devoted Chapter 4 to present several results on this issue.

[1] For two real and positive-valued sequences $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$, the notation $a_n \sim b_n$ will mean that there exist positive constants C_1, C_2 and $n_0 \in \mathbb{N}$ such that $n \geq n_0$ implies $C_1 \leq \frac{a_n}{b_n} \leq C_2$.

3.4 Proofs

We say that a stationary measure μ over $(\mathscr{X},\mathscr{F})$ is *periodic* when there exists a periodic point $x = x_0^{p-1}x_0^{p-1}\cdots \in \mathscr{X}$ such that $\mu\left(\left\{x,\sigma(x),\cdots,\sigma^{(p-1)}(x)\right\}\right) = 1$. Otherwise, we call the measure *aperiodic*.

Lemma 3.4.1. If μ is an ergodic and aperiodic measure, then $\mu(x) = 0$ for any point $x = x_0^\infty \in \mathscr{X}$. In particular, the statement holds for α-mixing measures.

Proof. Given a point $x \in \mathscr{X}$, consider the set $B = B(x) = \bigcup_{j\geq 0}\left\{\sigma^j(x)\right\}$. It follows directly that

$$\sigma^{-1}(B) = \bigcup_{j\geq 0}\sigma^{-1}\left(\sigma^j(x)\right) \supset \bigcup_{j\geq 0}\left\{\sigma^{j-1}(x)\right\} \supset B.$$

Applying Lemma I.2.1 in (SHIELDS, 1996) we must have $\mu(B) = 0$ or $\mu(B) = 1$. Now, if x is periodic, we have $B = \left\{x,\sigma(x),\cdots,\sigma^{(p-1)}(x)\right\}$, which implies $\mu(x) \leq \mu(B) = 0$, since the process is aperiodic. On the other hand, note that for any $j \geq 0$ we have $\mu\left(\sigma^j(x)\right) = \mu\left(\sigma^{-1}\left(\sigma^j(x)\right)\right) \geq \mu\left(\sigma^{j-1}(x)\right)$. Therefore, we get for aperiodic x:

$$1 \geq \mu(B) = \sum_{j\geq 0}\mu\left(\sigma^{-1}\left(\sigma^j(x)\right)\right) \geq \sum_{j\geq 0}\mu(x) \Rightarrow \mu(x) = 0.$$

For the second statement, we have to prove that strong mixing implies ergodicity and aperiodicity of μ. The first fact is well-known (see (SHIELDS, 1996) for instance). For the second fact, we proceed by contradiction. Suppose that μ is periodic. Then, there exists $x \in \mathscr{X}$ such that $\mu(B(x)) = 1$ and any point $y \in B(x)$ can be written as $y = y_0^{p-1}y_0^{p-1}\cdots$. Thus, for all $j \neq kp$ with $k \geq 1$, we have $\mu\left(\sigma^{-p-j}\left(y_0^{p-1}\right)\cap\left[y_0^{p-1}\right]\right) = 0$. Hence, the α-mixing condition implies

$$\mu\left(y_0^{p-1}\right)^2 = 0 \Rightarrow \mu(y) \leq \mu\left(y_0^{p-1}\right) = 0,$$

for any $y \in B(x)$, which is a contradiction. □

Lemma 3.4.2. For any stationary process $\mathbf{X}$ with law μ, $A = a_0^{n-1} \in \mathscr{C}_n$ and $j \geq 1$ we have $\mu(T_A = j) = \mu(A, T_A > j-1)$.

Proof. Define the binary process $(Y_m)_{m\in\mathbb{N}}$ by $Y_m = \mathbb{1}\left\{X_m^{m+n-1} = a_0^{n-1}\right\}$. Note that $\mathbf{Y}$ is also a stationary and denote its measure by ν. Then we have

$$\begin{aligned}
\mu(T_A = j) &= \nu\left(Y_1^j = 0^{j-1}1\right)\\
&= \nu\left(Y_1^{j-1} = 0^{j-1}\right) - \nu\left(Y_1^j = 0^j\right)\\
&= \nu\left(Y_1^{j-1} = 0^{j-1}\right) - \nu\left(Y_0^{j-1} = 0^j\right)\\
&= \nu\left(Y_0^{j-1} = 10^{j-1}\right)\\
&= \mu(A, T_A > j-1).
\end{aligned}$$

□

The statement of Theorem 3.3.4 is for ϕ and ψ and for hitting and return times. The case of return times under ϕ-mixing was already done by Abadi and Vergne (2008) considering complete grammar and finite alphabet. Our proof follows their method, but we made adaptations and extensions to hitting time and to return time under ψ-mixing. Adaptations made to cover the incomplete grammar and the infinite alphabet for return times under ψ can be done in a similar way for the case case ϕ, which we will not present.

The following lemma plays a fundamental role in Theorems 3.3.4 and 4.2.1. It was originally proven in (ABADI, 2001) assuming the summability of the function ϕ, an assumption that can be dropped.

Lemma 3.4.3. Let μ be a ϕ-mixing measure. Then, there exists positive constants C and c such that for all $n \geq 1$ and all $A \in \mathscr{C}_n$, one has:

$$\mu(A) \leq Ce^{-cn}.$$

Proof. We denote by $\lambda = \sup\{\mu(a) : a \in \mathscr{A}\} < 1$. Consider a positive integer k_0, and for all $n \geq k_0$, write $n = k_0 q + r$, with $1 \leq q \in \mathbb{N}$ and $0 \leq r < k_0$. Suppose $A = a_0^{n-1}$, and apply the ϕ-mixing property to obtain:

$$\begin{aligned}\mu(A) &\leq \mu\left(\bigcap_{j=0}^{q-1}\left\{\sigma^{-jk_0}(a_{jk_0})\right\}\right) \leq \mu\left(\bigcap_{j=0}^{q-2}\left\{\sigma^{-jk_0}(a_{jk_0})\right\}\right)(\phi(k_0)+\mu(a_{(q-1)k_0}))\\ &\leq \mu\left(\bigcap_{j=0}^{q-2}\left\{\sigma^{-jk_0}(a_{jk_0})\right\}\right)(\phi(k_0)+\lambda).\end{aligned}$$

Iterating this argument, one concludes:

$$\mu(A) \leq (\phi(k_0)+\lambda)^q.$$

Since $\phi(k) \xrightarrow{k} 0$, there exists $k_0 \in \mathbb{N}$ such that $\phi(k_0)+\lambda < 1$. Thus, for $n \geq k_0$ and observing that $q = \frac{n-r}{k_0} > \frac{n}{k_0} - \frac{k_0-1}{k_0}$:

$$\mu(A) \leq (\phi(k_0)+\lambda)^{-(k_0-1)/k_0}\left((\phi(k_0)+\lambda)^{1/k_0}\right)^n.$$

This covers the case $n \geq k_0$. By eventually enlarging the constant C, one covers the case $n < k_0$. This ends the proof. □

For a complete grammar, one has $\tau(A_n(x)) \leq n$. Since we do not assume this, we need the following lemma, which provides upper bounds for $\tau(A)$ when μ is ψ-mixing or summable ϕ-mixing.

Lemma 3.4.4. Consider μ a ψ-mixing or summable ϕ-mixing measure. Then, there exists $n_2 \in \mathbb{N}$ such that for all $n \geq n_2$ and $A \in \mathscr{C}_n$,

- $\tau(A) \le 2n$, for ψ;
- $\tau(A) \le -\dfrac{2}{\mu(A)\ln\mu(A)} + n$, for summable ϕ.

Proof. We start with the case ψ. For n large enough, we have $\psi(n) < 1$, which implies:

$$\mu\left(A \cap \sigma^{-2n}(A)\right) \ge \mu(A)^2(1-\psi(n)) > 0$$

Since $\tau(A)$ is the smallest positive integer such that $\mu\left(A \cap \sigma^{-\tau(A)}(A)\right) > 0$, one has $\tau(A) \le 2n$.

Now consider the ϕ-mixing case. The summability of ϕ ensures that for g large enough, we have $\phi(g) \le 1/(g\ln g)$. Thus:

$$\mu\left(A \cap \sigma^{-g-n}(A)\right) \ge \mu(A)\left(\mu(A) - \phi(g)\right) \ge \mu(A)\left(\mu(A) - \frac{1}{g\ln g}\right).$$

Take $g = -\dfrac{2}{\mu(A)\ln\mu(A)}$. The rightmost parenthesis above becomes:

$$\mu(A)\left[1 - \frac{1}{2}\frac{-\ln\mu(A)}{(\ln(2) - \ln\mu(A) - \ln(-\ln\mu(A)))}\right]$$

which is positive for n large enough. □

We present below an useful inequality concerning hitting time distribution for any stationary process. It will be used in several following proofs. For all $t \ge 1$, we have:

$$\mu(T_A \le t) = \mu\left(\bigcup_{j=1}^{t} \sigma^{-j}(A)\right) \le \sum_{j=1}^{t} \mu\left(\sigma^{-j}(A)\right) = t\mu(A).$$

The remaining results of this subsection hold for $n \ge n'$, where $n' = 1$ for the case of ϕ-mixing and:

$$n' := \inf\{n > 2g_0 : \psi(n) < 1\} \tag{3.6}$$

for the ψ-mixing case (see (3.3) for the definition of g_0).

Let us define $M := \psi(g_0+1)+1$.

Proposition 3.4.5. Let μ be a ψ-mixing measure. Then, for all $n \ge n'$, $A \in \mathscr{C}_n$ and $k \ge n_A$, the following inequality holds:

$$\mu_A(\tau(A) < T_A \le k) \le M(k - n_A + 1)\mu\left(A^{(n_A - g_0)}\right).$$

Proof. By definition of n_A, we first note that $\mu_A(\tau(A) < T_A \le k) = \mu_A(n_A \le T_A \le k)$. Consider the case in which $n_A \le n + g_0$. In this case, for $j \ge n_A$, one trivially has:

$$\{T_A = j\} \subset \sigma^{-j}(A) \subset \sigma^{-j-(n-(n_A-g_0))}\left(A^{(n_A-g_0)}\right).$$

Thus:

$$\mu_A(n_A \le T_A \le k) \le \mu_A\left(\bigcup_{n_A \le j \le k} \sigma^{-j-(n-(n_A-g_0))}\left(A^{(n_A-g_0)}\right)\right).$$

Note that A and the union on the right hand-side in the above inequality are separated by a gap of length g_0+1. By ψ-mixing, one concludes that the left-hand side is bounded by:

$$(\psi(g_0+1)+1)\mu\left(\bigcup_{n_A \le j \le k} \sigma^{-j-(n-(n_A-g_0))}\left(A^{(n_A-g_0)}\right)\right) \le M(k-n_A+1)\mu\left(A^{(n_A-g_0)}\right).$$

For $n_A > n+g_0$, recall first the convention in Section 3.3, which states that $\mu\left(A^{(n_A-g_0)}\right) = \mu(A)$. In a similar way to the first case,

$$\begin{aligned}\mu_A(n_A \le T_A \le k) &\le \mu_A\left(\bigcup_{n_A \le j \le k} \sigma^{-j}(A)\right)\\ &\le M(k-n_A+1)\mu(A).\end{aligned}$$

□

For the next proposition, recall that ε stands either for ε_ψ (3.4) or for ε_ϕ (3.5), according to the mixing property of the measure under consideration. Further, let us use the notation $T_A{}^{[i]} := T_A \circ \sigma^i$.

Proposition 3.4.6. Let μ be a ϕ or ψ-mixing measure. Then, for all $n \ge n'$, $A \in \mathscr{C}_n$ and $t \ge \tau(A)$:

$$|\mu_A(T_A > t) - \rho(A)\mu(T_A > t)| \le C\varepsilon(A),$$

where $C = 4$ for ε_ϕ and $C = 4(M+1)$ for ε_ψ.

Proof. The proof for ε_ϕ can be found in Proposition 4.1 Item (b) of (ABADI; VERGNE, 2008), and it remains valid even for a non-complete grammar and infinite alphabet. We observe that the error term defined therein is:

$$\varepsilon'(A) = \inf_{1 \le w \le n_A}\left\{(2n+\tau(A))\mu\left(A^{(w)}\right) + \phi(n_A - w + 1))\right\} \le 2\varepsilon_\phi(A),$$

which justifies $C = 4$ for this case.

Here, we prove the case ε_ψ in the same way. We start by assuming that $t \ge \tau(A) + 2n$. By the triangle inequality:

$$\begin{aligned}&|\mu_A(T_A > t) - \rho(A)\mu(T_A > t)|\\ &\le \left|\mu_A\left(T_A > \tau(A); T_A{}^{[\tau(A)]} > t - \tau(A)\right) - \right.\\ &\quad \left.\mu_A\left(T_A > \tau(A); T_A{}^{[\tau(A)+2n]} > t - \tau(A) - 2n\right)\right| \qquad (3.7)\\ &+ \left|\mu_A\left(T_A > \tau(A); T_A{}^{[\tau(A)+2n]} > t - \tau(A) - 2n\right) - \rho(A)\mu\left(T_A > t - \tau(A) - 2n\right)\right| \qquad (3.8)\\ &+ |\rho(A)\mu(T_A > t - \tau(A) - 2n) - \rho(A)\mu(T_A > t)|. \qquad (3.9)\end{aligned}$$

For the first modulus, by inclusion of sets:

$$(3.7) \le \mu_A\left(T_A > \tau(A); {T_A}^{[\tau(A)]} \le 2n\right) = \mu_A(\tau(A) < T_A \le \tau(A) + 2n).$$

If $n_A > \tau(A) + 2n$, the last term is equal to zero. Otherwise, we apply Proposition 3.4.5 to obtain:

$$\begin{aligned} \mu_A(\tau(A) < T_A \le \tau(A) + 2n) &\le M(\tau(A) + 2n - n_A + 1)\mu\left(A^{(n_A - g_0)}\right) \\ &\le 4Mn\mu\left(A^{(n_A - g_0)}\right) \end{aligned} \tag{3.10}$$

where the last inequality follows from Lemma 3.4.4.

By ψ-mixing, the modulus (3.8) is bounded by:

$$\rho(A)\mu(T_A > t - \tau(A) - 2n)\psi(n) \le \psi(n).$$

Note that the modulus is not needed for (3.9), and by inclusion:

$$\begin{aligned} (3.9) &\le \rho(A)\mu\left({T_A}^{[t-\tau(A)-2n]} \le \tau(A) + 2n\right) \\ &= \rho(A)\mu\left(T_A \le \tau(A) + 2n\right) \\ &\le (2n + \tau(A))\mu(A) \\ &\le 4n\mu\left(A^{(n_A - g_0)}\right) \end{aligned}$$

where the equality and second inequality follow from the stationarity of μ.

Therefore, for $t \ge \tau(A) + 2n$, the sum of (3.7), (3.8) and (3.9) is bounded by:

$$4Mn\mu\left(A^{(n_A - g_0)}\right) + \psi(n) + 4n\mu\left(A^{(n_A - g_0)}\right) \le 4(M+1)\varepsilon_\psi(A).$$

We now consider the case where $\tau(A) \le t < \tau(A) + 2n$. We have:

$$\begin{aligned} |\mu_A(T_A > t) - \rho(A)\mu(T_A > t)| &\le |\mu_A(T_A > t) - \rho(A)| + |\rho(A) - \rho(A)\mu(T_A > t)| \\ &\le \mu_A(\tau(A) < T_A \le \tau(A) + 2n) + t\mu(A) \\ &\le 4Mn\mu\left(A^{(n_A - g_0)}\right) + (\tau(A) + 2n)\mu\left(A^{(n_A - g_0)}\right) \\ &\le 4(M+1)\varepsilon_\psi(A). \end{aligned}$$

The last inequality follows from (3.10). The other inequalities are straightforward. This ends the proof. □

The next lemma establishes upper bounds for the tail distribution at the scale given by Kac's lemma, namely $1/\mu(A)$. For technical reasons, we actually choose the scale:

$$f_A := 1/(2\mu(A)).$$

Lemma 3.4.7. Let μ be a stationary measure. Then, for all $n \geq 1$, $A \in \mathscr{C}_n$, positive integer k, and $B \in \mathscr{F}_{kf_A}^{\infty}$, the following inequalities hold:

(a) $\mu(T_A > kf_A; B) \leq (\psi(n)+1)^k \mu(T_A > f_A - 2n)^k \mu(B)$,

(b) $\mu(T_A > kf_A; B) \leq (\mu(T_A > f_A - 2n) + \phi(n))^k (\mu(B) + \phi(n))$,

(c) $\mu_A(T_A > kf_A; B) \leq (\psi(n)+1)^k \mu(T_A > f_A - 2n)^{k-1} \mu(B)$.

Proof. We start by observing that $\{T_A > kf_A\} \subset \{T_A > kf_A - 2n\} \in \mathscr{F}_0^{kf_A - n}$. Thus, applying the ψ-mixing property, we get:

$$\mu(T_A > kf_A; B) \leq \mu(T_A > kf_A - 2n; B) \leq (\psi(n)+1)\mu(T_A > kf_A - 2n)\mu(B) \tag{3.11}$$

Furthermore:

$$\{T_A > kf_A - 2n\} = \left\{T_A > (k-1)f_A; {T_A}^{[(k-1)f_A]} > f_A - 2n\right\}.$$

Now, one can take in particular $B = \left\{{T_A}^{[(k-1)f_A]} > f_A - 2n\right\} \in \mathscr{F}_{(k-1)f_A}^{\infty}$ and then apply (3.11) with $k-1$ instead of k to get:

$$\begin{aligned}\mu(T_A > kf_A - 2n) &\leq (\psi(n)+1)\mu\left(T_A > (k-1)f_A - 2n\right)\mu\left({T_A}^{[(k-1)f_A]} > f_A - 2n\right)\\ &= (\psi(n)+1)\mu\left(T_A > (k-1)f_A - 2n\right)\mu\left(T_A > f_A - 2n\right).\end{aligned}$$

The equality follows by stationarity. Iterating this argument, one concludes that:

$$\mu(T_A > kf_A - 2n) \leq (\psi(n)+1)^{k-1}\mu(T_A > f_A - 2n)^k . \tag{3.12}$$

Applying the resulting inequality in (3.11), we get Statement (a). In a similar way, ϕ-mixing gives:

$$\mu(T_A > kf_A; B) \leq \mu(T_A > kf_A - 2n)(\mu(B) + \phi(n)).$$

Thus:

$$\begin{aligned}\mu(T_A > kf_A - 2n) &\leq \mu(T_A > (k-1)f_A - 2n)\left(\mu\left({T_A}^{[(k-1)f_A]} > f_A - 2n\right) + \phi(n)\right)\\ &\leq \left(\mu\left(T_A > f_A - 2n\right) + \phi(n)\right)^k . \end{aligned} \tag{3.13}$$

which ends the proof of (b).

The proof for (c) follows the same lines as Item (a), by observing that for $A, B \in \mathscr{F}_0^i$ and $C \in \mathscr{F}_{i+n}^{\infty}$, the ψ-mixing property implies $\mu_A(B; C) \leq \mu_A(B)\mu(C)(\psi(n)+1)$. $\square$

The next proposition is the key to the proof of Theorem 3.3.4, and the idea is the following. We work under the time scale f_A. When $t = kf_A,\ k \in \mathbb{N}$, then we simply cut t into k pieces of equal size f_A. Then, the case of general $t = kf_A + r, r < f_A$ is approximated by its integer part kf_A. Technically, this is done in (b) and (a), respectively.

Proposition 3.4.8. Let μ be a ϕ or ψ-mixing measure. Then, for all $n \geq n'$, $A \in \mathscr{C}_n$ and positive integer k, the following inequalities hold:
(a) For $0 \leq r \leq f_A$:

1. $|\mu(T_A > kf_A + r) - \mu(T_A > kf_A)\mu(T_A > r)| \leq C'(\psi(n)+1)^{k-1}\mu(T_A > f_A - 2n)^k \varepsilon_\psi(A)$
2. $|\mu(T_A > kf_A + r) - \mu(T_A > kf_A)\mu(T_A > r)| \leq C'(\mu(T_A > f_A - 2n) + \phi(n))^k \varepsilon_\phi(A)$
3. $|\mu_A(T_A > kf_A + r) - \mu_A(T_A > kf_A)\mu(T_A > r)| \leq C'((\psi(n)+1)\mu(T_A > f_A - 2n))^{k-1}\varepsilon_\psi(A)$.

(b) For $k \geq 1$:

1. $\left|\mu(T_A > kf_A) - \mu(T_A > f_A)^k\right| \leq C'\varepsilon_\psi(A)(k-1)(\psi(n)+1)^{k-2}\mu(T_A > f_A - 2n)^{k-1}$
2. $\left|\mu(T_A > kf_A) - \mu(T_A > f_A)^k\right| \leq C'\varepsilon_\phi(A)(k-1)(\mu(T_A > f_A - 2n) + \phi(n))^{k-1}$
3. $\left|\mu_A(T_A > kf_A) - \mu_A(T_A > f_A)\mu(T_A > f_A)^{k-1}\right| \leq C'\varepsilon_\psi(A)(k-1)((\psi(n)+1)\mu(T_A > f_A - 2n))^{k-2}$

where $C' = 2(M+1)$ for the cases involving ψ and $C' = 4$ for ϕ.

Proof. We will prove Items (a)-1 and (a)-2 together. Initially, consider the case in which $r < 2n$. In this case, for all $n \geq n'$, we have:

$$\begin{aligned}
&|\mu(T_A > kf_A + r) - \mu(T_A > kf_A)\mu(T_A > r)| \\
&\leq \left|\mu\left(T_A > kf_A, {T_A}^{[kf_A]} > r\right) - \mu(T_A > kf_A)\right| + \mu(T_A > kf_A)|1 - \mu(T_A > r)| \\
&\leq \mu\left(T_A > kf_A, {T_A}^{[kf_A]} \leq r\right) + \mu(T_A > kf_A)\mu(T_A \leq r).
\end{aligned} \tag{3.14}$$

By Lemma 3.4.7-(a) and (3.12), the last sum is bounded by:

$$\begin{aligned}
&((\psi(n)+1)\mu(T_A > f_A - 2n))^k \mu\left({T_A}^{[kf_A]} \leq r\right) + \mu(T_A > kf_A - 2n) r\mu(A) \\
&\leq ((\psi(n)+1)\mu(T_A > f_A - 2n))^k \mu(T_A \leq r) + (\psi(n)+1)^{k-1}\mu(T_A > f_A - 2n)^k r\mu(A) \\
&\leq (\psi(n)+1)^{k-1}\mu(T_A > f_A - 2n)^k r\mu(A)(M+1) \\
&\leq 2(M+1)(\psi(n)+1)^{k-1}\mu(T_A > f_A - 2n)^k \varepsilon_\psi(A)
\end{aligned} \tag{3.15}$$

which gives us (a)-1. To get (a)-2 for $r < 2n$, we apply Lemma 3.4.7-(b) and (3.13) in a similar way. Thus, (3.14) is bounded by:

$$(\mu(T_A > f_A - 2n) + \phi(n))^k(\mu(T_A \le r) + \phi(n)) + (\mu(T_A > f_A - 2n) + \phi(n))^k \mu(T_A \le r)$$
$$\le 4(\mu(T_A > f_A - 2n) + \phi(n))^k \varepsilon_\phi(A).$$

We now consider the case $r \ge 2n$. The triangle inequality gives us:

$$|\mu(T_A > kf_A + r) - \mu(T_A > kf_A)\mu(T_A > r)|$$
$$\le \left|\mu\left(T_A > kf_A; T_A{}^{[kf_A]} > r\right) - \mu\left(T_A > kf_A; T_A{}^{[kf_A+2n]} > r - 2n\right)\right| \tag{3.16}$$
$$+ \left|\mu\left(T_A > kf_A; T_A{}^{[kf_A+2n]} > r - 2n\right) - \mu\left(T_A > kf_A\right)\mu\left(T_A{}^{[kf_A+2n]} > r - 2n\right)\right| \tag{3.17}$$
$$+ \left|\mu\left(T_A > kf_A\right)\mu\left(T_A{}^{[kf_A+2n]} > r - 2n\right) - \mu\left(T_A > kf_A\right)\mu\left(T_A > r\right)\right|. \tag{3.18}$$

We proceed as in (3.7) and use Lemma 3.4.7-(a) to get:

$$\begin{aligned}(3.16) &\le \mu\left(T_A > kf_A; T_A{}^{[kf_A]} \le 2n\right)\\ &\le ((\psi(n)+1)(\mu(T_A > f_A - 2n))^k \mu(T_A \le 2n)\\ &\le 2n\mu(A)((\psi(n)+1)(\mu(T_A > f_A - 2n))^k.\end{aligned} \tag{3.19}$$

For the case ϕ, we apply Lemma 3.4.7-(b) and get:

$$(3.16) \le (\mu(T_A > f_A - 2n) + \phi(n))^k(2n\mu(A) + \phi(n)). \tag{3.20}$$

By ψ-mixing and (3.12):

$$\begin{aligned}(3.17) &\le \mu(T_A > kf_A - 2n)\psi(n)\\ &\le (\psi(n)+1)^{k-1}\mu(T_A > f_A - 2n)^k \psi(n)\end{aligned} \tag{3.21}$$

Applying ϕ-mixing and (3.13):

$$(3.17) \le (\mu(T_A > f_A - 2n) + \phi(n))^k \phi(n) \tag{3.22}$$

Finally, using the stationarity and the same arguments as above:

$$\begin{aligned}(3.18) &= \mu(T_A > kf_A)\mu\left(r - 2n < T_A \le r\right)\\ &\le 2n\mu(A)\mu(T_A > kf_A - 2n).\end{aligned} \tag{3.23}$$

Therefore, (3.12), (3.19), (3.21), and (3.23) give us:

$$(3.16) + (3.17) + (3.18) \le 2(M+1)(\psi(n)+1)^{k-1}\mu(T_A > f_A - 2n)^k \varepsilon_\psi(A)$$

and from (3.13), (3.20), (3.22), and (3.23), we get:

$$(3.16)+(3.17)+(3.18) \leq 4(\mu(T_A > f_A - 2n) + \phi(n))^k \varepsilon_\phi(A)$$

which ends the proof of (a)-1 and (a)-2.

For the proof of (a)-3, we write a similar triangle inequality as above:

$$\begin{aligned}
&|\mu_A(T_A > kf_A + r) - \mu_A(T_A > kf_A)\mu(T_A > r)| \\
&\leq \left|\mu_A\left(T_A > kf_A; T_A{}^{[kf_A]} > r\right) - \mu_A\left(T_A > kf_A; T_A{}^{[kf_A+2n]} > r - 2n\right)\right| \\
&+ \left|\mu_A\left(T_A > kf_A; T_A{}^{[kf_A+2n]} > r - 2n\right) - \mu_A\left(T_A > kf_A\right)\mu\left(T_A{}^{[kf_A+2n]} > r - 2n\right)\right| \\
&+ \mu_A\left(T_A > kf_A\right)\left|\mu\left(T_A{}^{[kf_A+2n]} > r - 2n\right) - \mu\left(T_A > r\right)\right|.
\end{aligned}$$

Then, we follow the same as we did for (a)-1, but applying Item (c) of Lemma 3.4.7 and using the ψ-mixing property:

$$|\mu_A(B;C) - \mu_A(B)\mu(C)| \leq \mu_A(B)\mu(C)\psi(n)$$

where $A, B \in \mathscr{F}_0^i$ and $C \in \mathscr{F}_{i+n}^\infty$. For the case $r < 2n$, we use:

$$\begin{aligned}
&|\mu_A(T_A > kf_A + r) - \mu_A(T_A > kf_A)\mu(T_A > r)| \\
&\leq \left|\mu_A\left(T_A > kf_A, T_A{}^{[kf_A]} > r\right) - \mu_A(T_A > kf_A)\right| + \mu_A(T_A > kf_A)|1 - \mu(T_A > r)|
\end{aligned}$$

and proceed as we did in (3.15), applying again Lemma 3.4.7-(c). This ends Item (a).

We now come to the proof of Items (b)-1 and (b)-2. For $k = 1$, we have an equality. For $k \geq 2$, we get:

$$\begin{aligned}
&\left|\mu(T_A > kf_A) - \mu(T_A > f_A)^k\right| \\
&= \left|\sum_{j=2}^{k} (\mu(T_A > jf_A) - \mu(T_A > (j-1)f_A)\mu(T_A > f_A))\,\mu(T_A > f_A)^{k-j}\right| \\
&\leq \sum_{j=2}^{k} |\mu(T_A > jf_A) - \mu(T_A > (j-1)f_A)\mu(T_A > f_A)|\,\mu(T_A > f_A)^{k-j}.
\end{aligned} \tag{3.24}$$

We put $r = f_A$ in Item (a)-1 to obtain (b)-1:

$$\begin{aligned}
(3.24) &\leq 2(M+1)\varepsilon_\psi(A)\sum_{j=2}^{k}(\psi(n)+1)^{j-2}\mu(T_A > f_A - 2n)^{j-1}\mu(T_A > f_A)^{k-j} \\
&\leq 2(M+1)\varepsilon_\psi(A)(k-1)(\psi(n)+1)^{k-2}\mu(T_A > f_A - 2n)^{k-1}.
\end{aligned}$$

Furthermore, we get the inequality (b)-2, under ϕ-mixing, proceeding similarly as above:

$$\begin{aligned}
(3.24) &\leq 4\varepsilon_\phi(A)\sum_{j=2}^{k}(\mu(T_A > f_A - 2n) + \phi(n))^{j-1}(\mu(T_A > f_A - 2n) + \phi(n))^{k-j} \\
&= 4\varepsilon_\phi(A)(k-1)(\mu(T_A > f_A - 2n) + \phi(n))^{k-1}
\end{aligned}$$

Finally, we prove (b)-3 applying (a)-3 as follows:

$$
\begin{aligned}
&\left|\mu_A(T_A > kf_A) - \mu_A(T_A > f_A)\mu(T_A > f_A)^{k-1}\right| \\
&\leq \sum_{j=2}^{k} |\mu_A(T_A > jf_A) - \mu_A(T_A > (j-1)f_A)\mu(T_A > f_A)|\, \mu(T_A > f_A)^{k-j} \\
&\leq 2(M+1)\varepsilon_\psi(A) \sum_{j=2}^{k} ((\psi(n)+1)\mu(T_A > f_A - 2n))^{j-2}\mu(T_A > f_A)^{k-j} \\
&\leq 2(M+1)\varepsilon_\psi(A)(k-1)((\psi(n)+1)\mu(T_A > f_A - 2n))^{k-2}.
\end{aligned}
$$

□

The next lemma is a classical result and we stated it without proof. Note that it is a discrete version of the mean value theorem, which follows with a straightforward computation.

Lemma 3.4.9. Given $a_1,...,a_n,b_1,...,b_n$ real numbers such that $0 \leq a_i, b_i \leq 1$, the following inequality holds:

$$
\left|\prod_{i=1}^{n} a_i - \prod_{i=1}^{n} b_i\right| \leq \sum_{i=1}^{n} |a_i - b_i| \left(\max_{1\leq i\leq n}\{a_i, b_i\}\right)^{n-1} \leq \sum_{i=1}^{n} |a_i - b_i|\,.
$$

Now we turn to the proof of the main result. Theorem 3.3.4 contains eight statements, each statement corresponding to the choice of:

- recurrence time: hitting or return,
- mixing property: ψ or ϕ,
- amplitude of t: smaller or larger than f_A.

Recall the definition of n' in (3.6). The proof of Theorem 3.3.4 holds for all $n \geq n_0$, where n_0 is explicitly given by:

$$
n_0 := \inf\left\{ m \geq n';\ \sup_{A\in\mathscr{C}_n} \mu(A)\tau(A) < 1/2,\ \forall n \geq m \right\} \tag{3.25}
$$

which is finite since $\sup_{A\in\mathscr{A}^n} \mu(A)\tau(A) \xrightarrow{n} 0$. Then, in particular, we have $\tau(A) < f_A$ for all $n \geq n_0$ and $A \in \mathscr{C}_n$.

Let us start the poof by the statements for small t's. Here, we assume that $1 \leq t \leq f_A := [2\mu(A)]^{-1}$.

Proof for hitting time, ϕ and ψ together. Recall that $\varepsilon(A)$ denotes $\varepsilon_\phi(A)$ or $\varepsilon_\psi(A)$, depending on whether the measure is ϕ or ψ-mixing.

Applying the inequality $|1-e^{-x}| \le x$ for $x \ge 0$, we obtain the statement for $1 \le t \le \tau(A)$ as follows:

$$\begin{aligned}\left|\mu(T_A > t) - e^{-\rho(A)\mu(A)t}\right| &\le |\mu(T_A > t) - 1| + \left|1 - e^{-\rho(A)\mu(A)t}\right| \\ &\le 2\tau(A)\mu(A). \end{aligned} \tag{3.26}$$

We consider now the case $\tau(A) < t \le f_A$. For positive $i \in \mathbb{N}$, define:

$$p_i = \frac{\mu_A(T_A > i-1)}{\mu(T_A > i-1)}.$$

Then:

$$\begin{aligned}\frac{\mu(T_A > t)}{\mu(T_A > \tau(A))} &= \prod_{i=\tau(A)+1}^{t} \frac{\mu(T_A > i)}{\mu(T_A > i-1)} = \prod_{i=\tau(A)+1}^{t} (1 - \mu(T_A = i | T_A > i-1)) \\ &= \prod_{i=\tau(A)+1}^{t} (1 - \mu(A)p_i) \end{aligned} \tag{3.27}$$

where we used Lemma 3.4.2 in the last equality.

Thus, for $\tau(A) < t \le f_A$, we apply (3.27) to obtain:

$$\begin{aligned}&\left|\mu(T_A > t) - e^{-\rho(A)\mu(A)t}\right| \\ &= \left|\mu(T_A > \tau(A)) \prod_{i=\tau(A)+1}^{t} (1 - \mu(A)p_i) - e^{-\rho(A)\mu(A)\tau(A)} \prod_{i=\tau(A)+1}^{t} e^{-\rho(A)\mu(A)}\right| \\ &\le \left|\mu(T_A > \tau(A)) - e^{-\rho(A)\mu(A)\tau(A)}\right| + \left|\prod_{i=\tau(A)+1}^{t} (1 - \mu(A)p_i) - \prod_{i=\tau(A)+1}^{t} e^{-\rho(A)\mu(A)}\right| \\ &\le 2\tau(A)\mu(A) + \sum_{i=\tau(A)+1}^{t} \left|1 - \mu(A)p_i - e^{-\rho(A)\mu(A)}\right| \end{aligned} \tag{3.28}$$

where the two inequalities follow from Lemma 3.4.9 and (3.26).

On the other hand, by the triangle inequality:

$$\left|1 - p_i\mu(A) - e^{-\rho(A)\mu(A)}\right| \le |p_i - \rho(A)|\,\mu(A) + \left|1 - \rho(A)\mu(A) - e^{-\rho(A)\mu(A)}\right|. \tag{3.29}$$

Since $|1 - x - e^{-x}| \le \frac{x^2}{2}$ for all $0 \le x \le 1$, by doing $x = \rho(A)\mu(A)$, we get:

$$\left|1 - \rho(A)\mu(A) - e^{-\rho(A)\mu(A)}\right| \le \frac{\rho(A)^2\mu(A)^2}{2} \le \frac{\varepsilon(A)\mu(A)}{2}.$$

Furthermore, still for $\tau(A) + 1 \le i \le f_A + 1$, Proposition 3.4.6 gives us:

$$|p_i - \rho(A)| = \left|\frac{\mu_A(T_A > i-1)}{\mu(T_A > i-1)} - \rho(A)\right| \le \frac{C\varepsilon(A)}{\mu(T_A > i-1)} \le 2C\varepsilon(A),$$

where, for the last inequality, we used:

$$\mu(T_A > i-1) = 1 - \mu(T_A \leq i-1) \geq 1 - (i-1)\mu(A) \geq 1 - f_A\mu(A) = \frac{1}{2}.$$

Thus, applying (3.29), we obtain for $\tau(A)+1 \leq i \leq f_A + 1$:

$$\left|1 - p_i\mu(A) - e^{-\rho(A)\mu(A)}\right| \leq (2C + 1/2)\,\varepsilon(A)\mu(A). \tag{3.30}$$

Therefore, (3.28) and (3.30) give us:

$$\begin{aligned}\left|\mu(T_A > t) - e^{-\rho(A)\mu(A)t}\right| &\leq 2\tau(A)\mu(A) + (2C+1/2)\,(t - \tau(A))\varepsilon(A)\mu(A) \\ &\leq (2C+1/2)\,[\tau(A)\mu(A) + t\mu(A)\varepsilon(A)] \end{aligned} \tag{3.31}$$

which concludes the statement of Theorem 3.3.4 for hitting time at small t's (with either ϕ or ψ). □

Proof for return time, ϕ and ψ together. We first note that the statement is trivial for $t = \tau(A)$, then we consider $t > \tau(A)$. By definition, we have $\mu_A(T_A > t) = p_{t+1}\mu(T_A > t)$. Then, we use again the triangle inequality to write:

$$\begin{aligned}&\left|\mu_A(T_A > t) - \rho(A)e^{-\rho(A)\mu(A)(t-\tau(A))}\right| \\ &\leq \mu(T_A > t)\,|p_{t+1} - \rho(A)| + \rho(A)\left|\mu(T_A > t) - e^{-\rho(A)\mu(A)(t-\tau(A))}\right|. \end{aligned} \tag{3.32}$$

As we saw before, the first modulus above is bounded by $2C\varepsilon(A)$. On the other hand, we apply (3.27) to obtain for $\tau(A) < t \leq f_A$:

$$\left|\mu(T_A > t) - e^{-\rho(A)\mu(A)(t-\tau(A))}\right| = \left|\mu(T_A > \tau(A)) \prod_{i=\tau(A)+1}^{t} (1 - \mu(A)p_i) - \prod_{i=\tau(A)+1}^{t} e^{-\rho(A)\mu(A)}\right|.$$

This is bounded, applying Lemma 3.4.9, by:

$$\begin{aligned}&|\mu(T_A > \tau(A)) - 1| + \left|\prod_{i=\tau(A)+1}^{t} (1 - \mu(A)p_i) - \prod_{i=\tau(A)+1}^{t} e^{-\rho(A)\mu(A)}\right| \\ &\leq \tau(A)\mu(A) + (2C+1/2)\,t\mu(A)\varepsilon(A)\end{aligned}$$

where the last inequality follows from (3.28) and (3.30). Finally, notice that $t\mu(A) \leq f_A\mu(A) = 1/2$ and $\tau(A)\mu(A) \leq 2\varepsilon(A)$ (use Lemma 3.4.4 for ψ). Therefore, we obtain from (3.32):

$$\left|\mu_A(T_A > t) - \rho(A)e^{-\rho(A)\mu(A)(t-\tau(A))}\right| \leq (3C + 9/4)\,\varepsilon(A) \tag{3.33}$$

This concludes the statement of Theorem 3.3.4 for the return time at small t's (with either ϕ or ψ). □

Now we turn to the proofs of the statements for large t's. The proof for the return time for $t > f_A$ was given in (ABADI; VERGNE, 2008) under ϕ-mixing, a finite alphabet, and a complete grammar. The proof still holds if one just assumes a countable alphabet and an incomplete grammar (recall Remark 3.3.6 for the uniform convergence to zero of the error term ε_ϕ). Thus, we focus on the hitting time under each mixing assumption and the return time only under ψ-mixing.

Proof of Theorem 3.3.4 for hitting times, for $t > f_A$. Write $t = kf_A + r$ with integer $k \geq 1$ and $0 \leq r < f_A$. Thus, we have:

$$\begin{aligned}\left|\mu(T_A > t) - e^{-\rho(A)\mu(A)t}\right| &\leq |\mu(T_A > kf_A + r) - \mu(T_A > kf_A)\mu(T_A > r)| && (3.34)\\ &+ \left|\mu(T_A > kf_A) - \mu(T_A > f_A)^k\right| \mu(T_A > r) && (3.35)\\ &+ \left|\mu(T_A > f_A)^k - e^{-\rho(A)\frac{k}{2}}\right| \mu(T_A > r) && (3.36)\\ &+ \left|e^{-\rho(A)\frac{k}{2}}\mu(T_A > r) - e^{-\rho(A)\mu(A)t}\right|. && (3.37)\end{aligned}$$

In order to get an upper bound for the sum of (3.34) and (3.35), we analyse the ψ and ϕ cases separately and start with the ψ-mixing. Applying Items (a)-1 and (b)-1 of Proposition 3.4.8, that sum is bounded by:

$$\begin{aligned}&\leq C'\varepsilon_\psi(A)(\psi(n)+1)^{k-1}\mu(T_A > f_A - 2n)^k\left(1 + (k-1)((\psi(n)+1)\mu(T_A > f_A - 2n))^{-1}\right)\\ &\leq 2(M+1)\varepsilon_\psi(A)\left((\psi(n)+1)\mu(T_A > f_A - 2n)\right)^k 2k\\ &\leq 8(M+1)\varepsilon_\psi(A)\mu(A)t\left((\psi(n)+1)\mu(T_A > f_A - 2n)\right)^k. && (3.38)\end{aligned}$$

where the last two inequalities are justified by $\mu(T_A > f_A - 2n)^{-1} \leq \mu(T_A > f_A)^{-1} \leq 2$ and $k \leq 2\mu(A)t$.

On the other hand, applying (3.31) with $t = f_A - 2n$, we get:

$$\begin{aligned}\left|\mu(T_A > f_A - 2n) - e^{-\rho(A)\mu(A)(f_A - 2n)}\right| &= \left|\mu(T_A > f_A - 2n) - e^{-\frac{\rho(A)}{2} + 2n\rho(A)\mu(A)}\right|\\ &\leq (2C + 1/2)\left(\tau(A)\mu(A) + (f_A - 2n)\mu(A)\varepsilon_\psi(A)\right)\\ &\leq (5C + 5/4)\,\varepsilon_\psi(A)\end{aligned}$$

where we use $\tau(A)\mu(A) \leq 2\varepsilon_\psi(A)$.

Furthermore, by the Mean Value Theorem (MVT):

$$\begin{aligned}\left|e^{-\frac{\rho(A)}{2} + 2n\rho(A)\mu(A)} - e^{-\frac{\rho(A)}{2}}\right| &\leq 2n\rho(A)\mu(A)e^{-\frac{\rho(A)}{2} + 2n\rho(A)\mu(A)}\\ &\leq 2n\mu(A)e^{2n\mu(A)} \leq \frac{11}{2}n\mu(A)\end{aligned}$$

since for $n \geq n_0$, we have $2n\mu(A) \leq 2\sup\mu(A)\tau(A) \leq 1$.

Thus, it follows that:

$$\begin{aligned}
&\left|(\psi(n)+1)\mu(T_A > f_A - 2n) - e^{-\frac{\rho(A)}{2}}\right| \\
&\leq \psi(n) + \left|\mu(T_A > f_A - 2n) - e^{-\frac{\rho(A)}{2}+2n\rho(A)\mu(A)}\right| + \left|e^{-\frac{\rho(A)}{2}+2n\rho(A)\mu(A)} - e^{-\frac{\rho(A)}{2}}\right| \\
&\leq (5C+27/4)\,\varepsilon_\psi(A).
\end{aligned}$$

Therefore:

$$((\psi(n)+1)\mu(T_A > f_A - 2n))^k \leq \left(e^{-\frac{\rho(A)}{2}} + (5C+27/4)\,\varepsilon_\psi(A)\right)^k.$$

Since $e^x - 1 \geq x \; \forall x \in \mathbb{R}$, by doing $K = (5C+27/4)\,e^{1/2}$, we get:

$$\begin{aligned}
&\left(e^{K\varepsilon_\psi(A)} - 1\right) \geq K\varepsilon_\psi(A) \geq (5C+27/4)\,\varepsilon_\psi(A) e^{\frac{\rho(A)}{2}} \\
&\Rightarrow e^{-\frac{\rho(A)}{2}}\left(e^{K\varepsilon_\psi(A)} - 1\right) \geq (5C+27/4)\,\varepsilon_\psi(A) \\
&\Rightarrow e^{-\frac{\rho(A)}{2}+K\varepsilon_\psi(A)} \geq (5C+27/4)\,\varepsilon_\psi(A) + e^{-\frac{\rho(A)}{2}}.
\end{aligned} \tag{3.39}$$

Now, using that $k = 2\mu(A)(t-r)$, we have:

$$\begin{aligned}
((\psi(n)+1)\mu(T_A > f_A - 2n))^k &\leq \left(e^{-\frac{\rho(A)}{2}+K\varepsilon_\psi(A)}\right)^k \\
&= e^{-\rho(A)\mu(A)t+\rho(A)\mu(A)r+2K\varepsilon_\psi(A)\mu(A)t-2K\varepsilon_\psi(A)\mu(A)r} \\
&\leq e^{-\mu(A)t\left(\rho(A)-2K\varepsilon_\psi(A)\right)} e^{\mu(A)r} \\
&\leq e^{1/2} e^{-\mu(A)t\left(\rho(A)-C_3\varepsilon_\psi(A)\right)}
\end{aligned} \tag{3.40}$$

where the last inequality follows from $e^{\mu(A)r} \leq e^{\mu(A)f_A}$.

Therefore, it follows from (3.38) that the sum of (3.34) and (3.35) is bounded by:

$$14(M+1)\varepsilon_\psi(A)\mu(A)te^{-\mu(A)t\left(\rho(A)-C_3\varepsilon_\psi(A)\right)}.$$

Consider now the case of ϕ-mixing. We apply Items (a)-2 and (b)-2 of Proposition 3.4.8 to get an upper bound for the sum of (3.34) and (3.35):

$$\begin{aligned}
&|\mu(T_A > kf_A + r) - \mu(T_A > kf_A)\mu(T_A > r)| + \left|\mu(T_A > kf_A) - \mu(T_A > f_A)^k\right| \mu(T_A > r) \\
&\leq 4\varepsilon_\phi(A)\,(\mu(T_A > f_A - 2n) + \phi(n))^k \left(1 + (k-1)\,(\mu(T_A > f_A - 2n) + \phi(n))^{-1}\right) \\
&\leq 4\varepsilon_\phi(A)\,(\mu(T_A > f_A - 2n) + \phi(n))^k\, 2k \\
&\leq 16\varepsilon_\phi(A)\mu(A)t\,(\mu(T_A > f_A - 2n) + \phi(n))^k.
\end{aligned}$$

Similarly to the ψ-mixing case, one obtains:

$$(\mu(T_A > f_A - 2n) + \phi(n))^k \leq e^{1/2} e^{-\mu(A)t\left(\rho(A)-C_3\varepsilon_\phi(A)\right)}$$

which implies in the ϕ-mixing case that the sum of (3.34) and (3.35) is bounded by:

$$27\varepsilon_\phi(A)\mu(A)te^{-\mu(A)t\left(\rho(A)-C_3\varepsilon_\phi(A)\right)}.$$

Now, we will treat the cases ψ and ϕ together to obtain upper bounds for (3.36) and (3.37). In order to get an upper bound for (3.36), we apply (3.31) with $t = f_A$:

$$\begin{aligned}\left|\mu(T_A > f_A) - e^{-\rho(A)\mu(A)f_A}\right| &= \left|\mu(T_A > f_A) - e^{-\frac{\rho(A)}{2}}\right| \\ &\leq (2C+1/2)\left(\tau(A)\mu(A) + f_A\mu(A)\varepsilon(A)\right) \\ &\leq (5C+5/4)\,\varepsilon(A). \end{aligned} \tag{3.41}$$

Thus, applying Lemma 3.4.9, we have:

$$\begin{aligned}&\left|\mu(T_A > f_A)^k - e^{-\rho(A)\frac{k}{2}}\right| \\ &\leq \sum_{i=1}^{k}\left|\mu(T_A > f_A) - e^{-\frac{\rho(A)}{2}}\right|\left(\max\left\{\mu(T_A > f_A), e^{-\frac{\rho(A)}{2}}\right\}\right)^{k-1}.\end{aligned}$$

The max is bounded using (3.41) by:

$$e^{-\frac{\rho(A)}{2}} + (5C+5/4)\,\varepsilon(A).$$

Naturally, the absolute value is also bounded by using (3.41), and we get that the above sum is bounded above by:

$$k\,(5C+5/4)\,\varepsilon(A)\left(e^{-\frac{\rho(A)}{2}} + (5C+5/4)\,\varepsilon(A)\right)^{k-1}. \tag{3.42}$$

Recalling that $k = 2\mu(A)(t-r)$ and proceeding as we did for (3.39) and (3.40), we get the following upper bound for (3.36):

$$2\,(5C+5/4)\,\varepsilon(A)\mu(A)t\,e^{-\mu(A)t(\rho(A)-C_3\varepsilon(A))}e \leq 7\,(4C+1)\,\varepsilon(A)\mu(A)te^{-\mu(A)t(\rho(A)-C_3\varepsilon(A))}.$$

To conclude the proof for the hitting time, we apply (3.31) with $t = r$ to bound (3.37) as follows:

$$\begin{aligned}\left|e^{-\rho(A)\frac{k}{2}}\mu(T_A > r) - e^{-\rho(A)\mu(A)t}\right| &= e^{-\rho(A)\mu(A)t+\rho(A)\mu(A)r}\left|\mu(T_A > r) - e^{-\rho(A)\mu(A)r}\right| \\ &\leq (2C+1/2)\,e^{-\rho(A)\mu(A)t+\mu(A)f_A}\left(\tau(A)\mu(A) + r\mu(A)\varepsilon(A)\right) \\ &\leq (2C+1/2)\left(\tau(A)\mu(A) + f_A\mu(A)\varepsilon(A)\right)e^{-\rho(A)\mu(A)t}e^{1/2} \\ &\leq (17C+5)\varepsilon(A)\mu(A)te^{-\mu(A)t(\rho(A)-C_3\varepsilon(A))}\end{aligned}$$

where the term $\mu(A)t$ follows from $1 = 2\mu(A)f_A \leq 2\mu(A)t$. □

Proof of Theorem 3.3.4 for the return time, for $t > f_A$ and under ψ-mixing. We use again the triangle inequality to write:

$$\begin{aligned}
&\left|\mu_A(T_A > t) - \rho(A)e^{-\rho(A)\mu(A)(t-\tau(A))}\right| \\
&\leq |\mu_A(T_A > kf_A + r) - \mu_A(T_A > kf_A)\mu(T_A > r)| && (3.43)\\
&+ \left|\mu_A(T_A > kf_A) - \mu_A(T_A > f_A)\mu(T_A > f_A)^{k-1}\right|\mu(T_A > r) && (3.44)\\
&+ \left|\mu_A(T_A > f_A)\mu(T_A > f_A)^{k-1} - \rho(A)e^{-\rho(A)\frac{k}{2}}\right|\mu(T_A > r) && (3.45)\\
&+ \rho(A)e^{-\rho(A)\frac{k}{2}}\left|\mu(T_A > r) - e^{-\rho(A)\mu(A)(r-\tau(A))}\right|. && (3.46)
\end{aligned}$$

Applying Items (a)-3 and (b)-3 of Proposition 3.4.8, the sum of (3.43) and (3.44) is bounded by:

$$\begin{aligned}
&2(M+1)\varepsilon_\psi(A)((\psi(n)+1)\mu(T_A > f_A - 2n))^{k-1}(1+(k-1)((\psi(n)+1)\mu(T_A > f_A - 2n))^{-1})\\
&\leq 2(M+1)\varepsilon_\psi(A)((\psi(n)+1)\mu(T_A > f_A - 2n))^{k-1}2k\\
&\leq 8(M+1)\varepsilon_\psi(A)\mu(A)t((\psi(n)+1)\mu(T_A > f_A - 2n))^{k-1}.
\end{aligned}$$

Replacing k by $k-1$ in (3.40), the last term is bounded above by:

$$8(M+1)\varepsilon_\psi(A)\mu(A)t\, e\, e^{-\mu(A)t\left(\rho(A)-C_3\varepsilon_\psi(A)\right)} \leq 22(M+1)\varepsilon_\psi(A)\mu(A)t\, e^{-\mu(A)t\left(\rho(A)-C_3\varepsilon_\psi(A)\right)}.$$

On the other hand, Lemma 3.4.9 gives us:

$$\begin{aligned}
(3.45) \leq &\left(\max\left\{\mu_A(T_A > f_A), \mu(T_A > f_A), e^{-\rho(A)/2}\right\}\right)^{k-1}\left(\left|\mu_A(T_A > f_A) - \rho(A)e^{-\rho(A)/2}\right|\right.\\
&\left.+\sum_{i=1}^{k-1}\left|\mu(T_A > f_A) - e^{-\rho(A)/2}\right|\right)
\end{aligned}$$

The last sum is bounded by $(5C+5/4)(k-1)\varepsilon_\psi(A)$ using (3.41). On the other hand, applying (3.33) with $t = f_A$ and the MVT, we obtain:

$$\begin{aligned}
&\left|\mu_A(T_A > f_A) - \rho(A)e^{-\rho(A)/2}\right|\\
&\leq \left|\mu_A(T_A > f_A) - \rho(A)e^{-\rho(A)/2+\rho(A)\mu(A)\tau(A)}\right| + \rho(A)\left|e^{-\rho(A)/2+\rho(A)\mu(A)\tau(A)} - e^{-\rho(A)/2}\right|\\
&\leq (3C+9/4)\varepsilon_\psi(A) + \rho(A)\mu(A)\tau(A)e^{-\rho(A)(1/2-\mu(A)\tau(A))}\\
&\leq (3C+17/4)\varepsilon_\psi(A)
\end{aligned}$$

since $e^{-\rho(A)(1/2-\mu(A)\tau(A))} \leq 1$ and $\mu(A)\tau(A) \leq 2\varepsilon_\psi(A)$ for $n \geq n_0$.

Furthermore, the last inequality implies:

$$\mu_A(T_A > f_A) \leq \rho(A)e^{-\rho(A)/2} + (3C+17/4)\varepsilon_\psi(A) \leq e^{-\rho(A)/2} + (5C+5/4)\varepsilon_\psi(A)$$

and by (3.41), we get:

$$\max\left\{\mu_A(T_A > f_A), \mu(T_A > f_A), e^{-\rho(A)/2}\right\} \leq e^{-\rho(A)/2} + (5C+5/4)\varepsilon_\psi(A).$$

Therefore, as we saw in (3.42), we have:

$$\begin{aligned}
(3.45) &\leq (5C+5/4)\varepsilon_\psi(A)k\left(e^{-\rho(A)/2} + (5C+5/4)\varepsilon_\psi(A)\right)^{k-1} \\
&\leq 2(5C+5/4)\varepsilon_\psi(A)\mu(A)t\, e^1\, e^{-\mu(A)t\left(\rho(A)-C_3\varepsilon_\psi(A)\right)} \\
&\leq (109M+116)\varepsilon_\psi(A)\mu(A)te^{-\mu(A)t\left(\rho(A)-C_3\varepsilon_\psi(A)\right)}.
\end{aligned}$$

Finally, by doing $t = r$ in (3.31) and applying the MVT once again, we get:

$$\begin{aligned}
&\left|\mu(T_A > r) - e^{-\rho(A)\mu(A)(r-\tau(A))}\right| \\
&\leq \left|\mu(\tau_A > r) - e^{-\rho(A)\mu(A)r}\right| + \left|e^{-\rho(A)\mu(A)r} - e^{-\rho(A)\mu(A)(r-\tau(A))}\right| \\
&\leq (2C+1/2)(\tau(A)\mu(A) + r\mu(A)\varepsilon_\psi(A)) + \rho(A)\mu(A)\tau(A)e^{-\rho(A)\mu(A)(r-\tau(A))} \\
&\leq (2C+1/2)(2\varepsilon_\psi(A) + f_A\mu(A)\varepsilon_\psi(A)) + (7/2)\varepsilon_\psi(A) \\
&\leq (5C+19/4)\varepsilon_\psi(A).
\end{aligned}$$

The third inequality follows from $e^{-\rho(A)\mu(A)(r-\tau(A))} \leq e^{\rho(A)\mu(A)\tau(A)} \leq e^{1/2}$, since $n \geq n_0$. Now, just note that $\rho(A)\mu(A)\tau(A) \leq 2\varepsilon_\psi(A)$.

Therefore, we finish the proof by obtaining the following upper bound:

$$\begin{aligned}
(3.46) &\leq (5C+19/4)\varepsilon_\psi(A)e^{\rho(A)\mu(A)r}e^{-\rho(A)\mu(A)t} \\
&\leq (5C+19/4)\varepsilon_\psi(A)2\mu(A)t\; e^{f_A\mu(A)}e^{-\mu(A)t(\rho(A)-C_3\varepsilon_\psi(A))} \\
&\leq (66M+82)\varepsilon_\psi(A)\mu(A)te^{-\mu(A)t(\rho(A)-C_3\varepsilon_\psi(A))}.
\end{aligned}$$

$\square$

CHAPTER

4

POTENTIAL WELL

In this chapter, we discuss the role of the potential well in recurrence times. We present some results, also published in (ABADI; AMORIM; GALLO, 2021), as well as other new and complementary results. Namely, we analyse the possible values of the parameter $\rho(A)$ in its range $[0,1]$ and show its influence on distributions and moments of recurrence times. In particular, we state some results on the positivity and the asymptotic behaviour of the potential well. Then, we derive useful corollaries concerning the moments of the recurrence times and the almost sure convergence of their distributions. We also dedicate a section for examples and related results on specific classes of processes. Finally, we finish the chapter with a theorem on the shortest return, which is directly related to the potential well. The proofs are given in the last section.

4.1 Role of the potential well

Let us first highlight the crucial role of the potential well in the recurrence times.

Limiting distributions

Theorem 3.3.4 shows that the potential can be used as scaling parameter for the approximate distributions of the recurrence times. Abadi, Cardeno and Gallo (2015) showed that this relationship extends to any renewal process. In these cases, when $\rho(A_n(x)) \xrightarrow{n} \rho(x)$, we have

$$\mu\left(T_{A_n(x)} > \mu(A_n(x))^{-1}t\right) \xrightarrow{n} e^{-\rho(x)t}, \quad \forall\, t \geq 0. \tag{4.1}$$

For return times, under conditions of Theorem 3.3.4, we have the convergence to the following convex combination:

$$\mu_{A_n(x)}\left(T_{A_n(x)} > \mu(A_n(x))^{-1}t\right) \xrightarrow{n} (1-\rho(x))\delta + \rho(x)e^{-\rho(x)t}, \quad \forall\, t \geq 0, \tag{4.2}$$

where $\delta = \mathbb{1}\{t=0\}$ is the Dirac measure. Furthermore, by Theorem 3.2 and Corollary 3.1 in (ABADI; CARDENO; GALLO, 2015), for all $x \in \mathscr{X} - \{0^\infty\}$ we have $\rho = \inf_{n\geq 1} \rho(A_n(x)) > 0$, and for any $t \geq \rho^{-1}$:

$$\mu_{A_n(x)}\left(T_{A_n(x)} > \mu(A_n(x))^{-1}t\right) \xrightarrow{n} \rho(x)e^{-\rho(x)t}.$$

When $x = 0^\infty$, we can have different regimes, as we will show in Example 4.3.9.

Scale of the recurrence times

Even when $\rho(A)$ does not converge, it plays a role in the scale of occurrence of events. As we will show in the next section, the relationship

$$\mathbb{E}(T_A) \sim \mu(A)^{-1} = \mathbb{E}_A(T_A)$$

depends on the positivity of $\rho(A)$[1]. In fact, we will show that, for α-mixing processes, whenever the potential well approximates to zero we have

$$\mathbb{E}(T_A) \gg \frac{1}{\mu(A)} = \mathbb{E}_A(T_A).$$

Actually the statement is more general and shows that the order of the β-moments of the hitting time, for $\beta > 0$, is bigger than $\mu(A)^{-\beta}$. In other words, this case shows the role of the potential well in a drastic change of scale of the hitting time. Furthermore, for any stationary process, the convergence of $\rho(A)$ to zero leads to the convergence of the return time distribution to a degenerated law.

This raises a natural question: what it takes in terms of mixing conditions to have uniform boundedness of $\rho(A)$ away from 0? We present in Theorem 4.2.1 two conditions to ensure the positivity of the potential well.

Positivity and almost sure convergence

Recall that Abadi and Saussol (2016) proved that, for α-mixing processes with algebraic decay of the function $\alpha(n)$[2], the rescaled hitting and return times distributions converge, almost surely, to an exponential law with parameter one. Thus, it also holds for ϕ and ψ-mixing process with algebraic decay rate. In this case, the result of Abadi em Saussol and Theorem 3.3.4 prove, indirectly, that the potential well converges almost surely to one, since the limit distributions must be equal in this cases. In fact, in Theorem 4.2.1 and Proposition 4.4.1 we prove that the same holds for ϕ-mixing without any assumption on the rate and for other general classes of processes. The first result was partially proved in (ABADI; AMORIM; GALLO, 2021).

Given a stationary measure μ on $(\mathscr{X}, \mathscr{F})$, the Shannon entropy of μ is defined by:

$$h_\mu := -\lim_{n\to\infty} \frac{1}{n} \sum_{A\in\mathscr{C}_n} \mu(A)\ln(\mu(A)).$$

[1] Recall that $a_n \sim b_n$ means that $\max\{a_n/b_n, b_n/a_n\}$ is bounded.
[2] Algebraic decay means that $\alpha(n)n^p \longrightarrow 0$ for some $p > 0$.

4.2 Main result and consequences

By the nature of the potential well, some aspects of Theorem 4.2.1 require an understanding of the asymptotic behaviour of the shortest return $\tau(A_n(x))$. Notice that it is also related to the hypothesis of Theorem 3.3.4, since we require that $\sup_A \mu(A)\tau(A) \xrightarrow{n} 0$. In Section 4.4, we present some results on this issue, which has its own interest. These results will be used in the proofs of some properties of the potential well that we will state in this section.

Theorem 4.2.1. Let $\mathbf{X}$ be a stationary process with law μ.

(a) Suppose that $\mathbf{X}$ satisfies one of the following two conditions:

(a1) $\mathbf{X}$ is a ϕ-mixing process over a finite or countable alphabet;

(a2) $\mathbf{X}$ is an α-mixing process over a finite alphabet satisfying $h_\mu > 0$ and $\alpha(j) \leq j^{-r}$ for some $r > 2$.

Then, $\rho(A_n(x)) \xrightarrow{n} 1$, almost surely.

(b) If μ is ψ-mixing or summable ϕ-mixing, there exists $n_0 \in \mathbb{N}$ such that

$$\inf_{n \geq n_0, A_n(x) \in \mathscr{C}_n} \rho\left(A_n(x)\right) > 0,$$

with $n_0 = 1$ when the process has a finite alphabet. In particular, there exists $\rho > 0$ such that *for any point* $x \in \mathscr{X}$, we have

$$\liminf_{n \to \infty} \rho\left(A_n(x)\right) \geq \rho.$$

Remark 4.2.2. In Proposition 4.4.1 we will prove that, if an ergodic process over a finite alphabet satisfies $h_\mu > 0$ and $\limsup_{n \to \infty} (1/n) \ln \tau(A_n(x)) < h_\mu$ almost surely, then $\rho(A) \longrightarrow 1$ almost surely. Thus, another way to obtain the Item (a) of Theorem 4.2.1 is to assume the so called specification property, defined in (SAUSSOL; TROUBETZKOY; VAIENTI, 2002). Under these conditions, the authors proved that $\tau(A_n(x))/n \longrightarrow 1$ almost surely, which implies $\limsup_{n \to \infty} (1/n) \ln \tau(A_n(x)) = 0$. Notice that complete grammar is a stronger assumption than specification property. On the other hand, in the next section we prove that for any Markov chain (including non ϕ-mixing), the potential well converges almost surely to one. We conjecture that this property extends beyond the conditions of Theorem 4.2.1.

Remark 4.2.3. We saw earlier that the positivity of the potential well does not hold in general for β-mixing process. However, in the next section we provide an example of a non ϕ-mixing Markov chain such that $\rho(A) \geq \rho > 0$ for all $A \in \mathscr{C}_n$ and $n \geq 1$. Even in process where $\inf_{x \in \mathscr{X}, n \geq 1} \rho(A_n(x)) = 0$, our examples illustrate that in some cases one can find a set $\mathscr{X}' \subset \mathscr{X}$ such that $\mu(\mathscr{X}') = 1$ and $\inf_{x \in \mathscr{X}', n \geq 1} \rho(A_n(x)) > 0$.

Moments of the hitting and return times

Inspired by Corollary 4.1 in (ABADI; VERGNE, 2008), Corollary 4.2.4 below provides an approximation for the β-moment ($\beta > 0$) of the rescaled hitting time $\mu(A)T_A$. In the case of return times, we extend their result for the cases involving incomplete grammar or infinite alphabet. Furthermore, we establish the approximation using the correct error term, as mentioned in Remark 3.3.11. As a direct consequence, we have the moments of the recurrence times for typical cylinders given by the gamma function and the order of the mean hitting time as the inverse of the measure of A. This result is also useful in the context of the waiting time spectrum, as we will see in Chapter 5.

Corollary 4.2.4. Let μ be a ψ-mixing or summable ϕ-mixing measure. Then, for any $\beta > 0$, there exist positive constants $C_1 = C_1(\beta)$, $C_2 = C_2(\beta)$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and all $A \in \mathscr{C}_n$, the following inequalities hold.

(a) $\left|\mu(A)^\beta \mathbb{E}\left(T_A^\beta\right) - \dfrac{\Gamma(\beta+1)}{\rho(A)^\beta}\right| \leq C_1 \varepsilon(A)$.

(b) $\left|\mu(A)^\beta \mathbb{E}_A\left(T_A^\beta\right) - \dfrac{\Gamma(\beta+1)}{\rho(A)^{\beta-1}}\right| \leq C_2 \varepsilon(A)^{\min\{1,\beta\}}$.

Furthermore, we have

(c) $\mathbb{E}\left(\left(\mu(A_n(x))T_{A_n(x)}\right)^\beta\right), \mathbb{E}_A\left(\left(\mu(A_n(x))T_{A_n(x)}\right)^\beta\right) \xrightarrow{n} \Gamma(\beta+1)$, almost surely.

(d) $\mathbb{E}\left(T_A^\beta\right) \sim \mu(A)^{-\beta} \sim \mathbb{E}_A\left(T_A^\beta\right)$ and each one of them grows at least exponentially fast.

Remark 4.2.5. Notice that the Corollary 4.2.4 implies that $\mathbb{E}\left((\mu(A)T_A)^\beta\right)$ and $\mathbb{E}_A\left((\mu(A)T_A)^\beta\right)$ are finite for all $\beta > 0$. In fact, to prove Items (a) and (b) we first need to prove this finiteness, which we did in Lemma 4.5.1. Furthermore, note that Item (b) provides a weak version of Kac's Lemma, as already mentioned in (ABADI; VERGNE, 2008).

Almost sure convergence

In the broad context of measuring preserving dynamical systems (which includes stationary processes), Hirata, Saussol and Vaienti (2000) proved that rescaled return time distribution $\mu_{A_n(x)}\left(T_{A_n(x)} > \mu(A_n(x))^{-1}t\right)$ converges to a parameter one exponential law if, and only if, the rescaled hitting and return time distributions are arbitrarily close. Later, Abadi and Saussol (2016) proved that the first condition (and therefore both) holds almost surely for α-mixing process with algebraic decay rate of $\alpha(n)$.

The next Corollary states that, under assumptions of Theorem 3.3.4, both conditions also hold almost surely, without any assumption on the rate of ψ or ϕ. It is worth mentioning that Theorem 4.4 in (BRADLEY, 2005) shows that, given a rate of convergence to zero, there exists a stationary process such that $\alpha(n)$ and $\psi(n)$ converges to zero at the same rate. That is, one can obtain ψ-mixing processes for which $\alpha(n)$ has a decay which is slower than algebraic.

Corollary 4.2.6. Consider a stationary measure μ on $(\mathscr{X}, \mathscr{F})$ enjoying either ϕ-mixing with $\sup_{A\in\mathscr{C}_n} \mu(A)\tau(A) \xrightarrow{n} 0$ or ψ-mixing. For each $x \in \mathscr{X}$, denote $A = A_n(x)$. Then for all $t \geq 0$, we have

(a) $\lim\limits_{n\to\infty} |\mu(T_A > t) - \mu_A(T_A > t)| = 0, \quad \text{a.s.}$

(b) $\left.\begin{matrix} \mu\left(T_A > \mu(A)^{-1}t\right) \\ \mu_A\left(T_A > \mu(A)^{-1}t\right) \end{matrix}\right\} \xrightarrow{n} e^{-t}, \quad \text{a.s.}$

Scale test

The next proposition shows another useful application of the potential well. Recall that Abadi and Saussol (2011) obtained an exponential approximation theorem for the hitting and return times with a scaling parameter $\lambda(A)$. Later, in (ABADI; SAUSSOL, 2016), the authors proved that $\lambda(A) \leq (1+o(1))\rho(A)$. Therefore, when $\rho(A)$ converges to zero, the return time distribution converges to a degenerated law, and the parameter of the exponential approximation for the hitting time distribution cannot be of the order of $\mu(A)^{-1}$. Since $\rho(A)$ is much easier to compute, it can be used for α-mixing process as a testing function to determine whether hitting and return times distributions have these properties or not. Actually, we prove that the same holds for any stationary process.

We also show that, when $\rho(A_n(x))$ is arbitrarily close to zero for some point $x \in \mathscr{X}$, the scale of the moments of the hitting and return times is bigger than $\mu(A)^{-\beta}$, $\beta > 1$. In particular the order of the mean hitting time is bigger than $\mu(A)^{-1} = \mathbb{E}_A(T_A)$. In other words, only the first moment of the return time has the order of the inverse of the measure.

Proposition 4.2.7. Consider a stationary measure μ on $(\mathscr{X}, \mathscr{F})$ and denote $A = A_n(x)$.

(a) If $\rho(A) \xrightarrow{n} 0$ for some $x \in \mathscr{X}$, then

$$\mu_A\left(T_A > \mu(A)^{-1}t\right) \xrightarrow{n} 0, \quad \forall\, t \geq 1.$$

If moreover $\tau(A)\mu(A) \xrightarrow{n} 0$, then

$$\mu\left(T_A > \mu(A)^{-1}t\right) \xrightarrow{n} 1, \ \forall t \geq 0 \quad \text{and} \quad \mu_A\left(T_A > \mu(A)^{-1}t\right) \xrightarrow{n} \begin{cases} 1, & \text{if } t = 0 \\ 0, & \text{if } t \in (0,1). \end{cases}$$

(b) If μ is α-mixing and $\liminf\limits_{n\to\infty} \rho(A) = 0$ for some $x \in \mathscr{X}$, then

$$\limsup_{n\to\infty} \mu(A)^{\beta}\mathbb{E}\left(T_A^{\beta}\right) = \infty, \ \forall \beta > 0 \quad \text{and} \quad \limsup_{n\to\infty} \mu(A)^{\beta}\mathbb{E}_A\left(T_A^{\beta}\right) = \infty, \ \forall \beta > 1.$$

Remark 4.2.8. As we will see in Proposition 4.3.8, any non ϕ-mixing renewal process satisfies $\liminf_n(\rho(A)) = 0$ and $\tau(A)\mu(A) \longrightarrow 0$ when we consider the point $x = 0^{\infty}$.

4.3 Examples

In this section we present some examples and results in order to analyse the asymptotic behaviour and the positivity of the potential well in some specific classes of processes, such as the ones that we presented in Chapter 2: i.i.d., Markov chains and binary renewal processes.

Abadi, Cardeno and Gallo (2015) provide a complete description of the potential well behaviour for renewal processes. Inspired by them, we give explicit formulas for the computation of the potential well and its limit in Propositions 4.3.1 and 4.3.3, for i.i.d. and Markov chains, respectively.

I.I.D. processes

We start with the simplest case of i.i.d. processes, for which we completely described the behaviour of the potential well in the next proposition. First, let us define the period of a point $x \in \mathscr{X}$ by $\tau(x) := \sup\{\tau(A_n(x)); n \geq 1\}$. When $\tau(x) < \infty$ we say that x is a periodic point, otherwise, we call it aperiodic.

Proposition 4.3.1. Consider an i.i.d. process over a finite or countable alphabet $\mathscr{A}$ with law μ. Then, for all $n \geq 1$ and $x \in \mathscr{X}$, we have

(a) $\rho(A_n(x)) = 1 - \displaystyle\prod_{j=0}^{\tau(A_n(x))-1} \mu(x_j)$.

(b) If $x = x_0^{p-1} x_0^{p-1} \cdots$ is a periodic point of period p, then, for all $n \geq p$:

$$\rho(A_n(x)) = 1 - \prod_{j=0}^{p-1} \mu(x_j).$$

(c) $\rho(A_n(x)) \longrightarrow 1$ if, and only if, x is aperiodic.

(d) $\displaystyle\inf_{n \geq 1, x \in \mathscr{X}} \rho(A_n(x)) = 1 - \lambda$, where $\lambda = \max\{\mu(a), a \in \mathscr{A}\}$.

Example 4.3.2. Consider an i.i.d. process with law μ over an alphabet $\mathscr{A}$. Proposition 4.3.1 shows that $\rho(x)$ exists and is positive for all $x \in \mathscr{A}^{\mathbb{N}}$. Thus, (4.1) and (4.2) always hold on this case. Note that, if x is aperiodic, the limiting distributions of hitting and return times are equal to the parameter one exponential law.

For instance, let $a, b \in \mathscr{A}$ such that $\mu(a) = p$ and $\mu(b) = q$ and consider the periodic point $x = abbaabba\cdots$. It is immediate to verify that for $n \geq 4$, we have:

$$\rho(A_n(x)) = 1 - p^2 q^2 = \rho(x).$$

In this case, since the process is ψ-mixing, the error term of the approximation is given by $\varepsilon(A) = Cnp^{n/2}q^{n/2}$.

Furthermore, for any $x \in \mathscr{A}^{\mathbb{N}}$, we can apply Corollary 4.2.4 to obtain:

$$\mu(A)^{\beta}\mathbb{E}\left(T_A^{\beta}\right) \xrightarrow{n} \frac{\Gamma(\beta+1)}{\rho(x)^{\beta}} \quad \text{and} \quad \mu(A)^{\beta}\mathbb{E}_A\left(T_A^{\beta}\right) \xrightarrow{n} \frac{\Gamma(\beta+1)}{\rho(x)^{\beta-1}}. \tag{4.3}$$

Markov chains

Recall from Chapter 2 the mixing properties of the stationary, irreducible and aperiodic Markov chains. In the case of the finite-state, the chain is exponentially ψ-mixing, which allows us to apply all the previous results. For the infinite-state case, we have β-mixing (and therefore α-mixing) in general, but the decay rate is not necessarily exponential in both cases. This means that Theorem 4.2.1 does not ensure the almost sure convergence to one nor the positivity of the potential well.

We therefore state the next proposition which is specific to the Markov case. It is a complete description of the potential well. Set $\kappa_A := \min\{\tau(A), n\}$, for all $n \geq 1$ and $A \in \mathscr{C}_n$.

Proposition 4.3.3. Consider a stationary, irreducible and aperiodic Markov chain $\mathbf{X}$ over a finite or countable alphabet $\mathscr{A}$. Suppose that $\mathbf{X}$ has law μ and is defined by the transition matrix Q. Then, for any $x \in \mathscr{X}$ and $n \geq 1$, we have:

(a) $\rho(A_n(x)) = 1 - Q^{\tau(A)-\kappa_A+1}(x_{\kappa_A-1}, x_0) \prod_{j=0}^{\kappa_A-2} Q(x_j, x_{j+1})$.

(b) If $x = x_0^{p-1} x_0^{p-1} \cdots$ is a periodic point of period p, then, for all $n \geq p$:

$$\rho(A_n(x)) = 1 - Q(x_{p-1}, x_0) \prod_{j=0}^{p-2} Q(x_j, x_{j+1}).$$

(c) $\rho(A_n(x)) \longrightarrow 1$ if, and only if, x is aperiodic.

(d) If $\lambda = \sup_{a,b \in \mathscr{A}} Q(a,b)$, then

$$\inf_{n \geq 1, x \in \mathscr{X}} \rho(A_n(x)) \geq 1 - \lambda,$$

where the equality holds if $\lambda = \sup_{a \in \mathscr{A}} Q(a,a)$.

(e) For all $x \in \mathscr{X}$,

$$\lim_{n \to \infty} \rho(A_n(x)) > 0.$$

Remark 4.3.4. Note that when $\mathscr{A}$ is finite, we must have $\lambda < 1$. If the alphabet is countable and $\lambda < 1$, we know from Proposition 2.4.1 that the Markov chain is exponentially ϕ-mixing. In this case, both Theorem 4.2.1 and Proposition 4.3.3-(d) ensure the positivity of the potential well. When $\lambda = 1$, the positivity in the infinite-state case depends on characteristics of the transition matrix. Although Item (e) states that $\rho(x) = \lim_n \rho(A_n(x)) > 0$ for all $x \in \mathscr{X}$, we may not have the positivity of $\rho(A)$ when we take the infimum over all points. The following examples show some possible situations.

Example 4.3.5 (The house of cards Markov chain)**.** Consider the house of cards Markov chain $\mathbf{X}$ defined in Example 2.1.3. Recall $\mathbf{X}$ cannot be ψ-mixing, but it can be exponentially ϕ-mixing or only β-mixing.

Positivity

In the case of the house cards, even when it is not ϕ-mixing, we can show the positivity of the potential well, which is due to the structure of Q. In order to prove this, suppose first that $A_n(x) = x_0^{n-1}$ has at least one symbol equal to zero, let us say $x_\ell = 0$. If $\ell \neq n-1$, then $Q(x_\ell, x_{\ell+1}) \leq \max\{q_0, 1-q_0\}$, since $x_{\ell+1}$ can be 0 or 1. Note that either $0 \leq \ell \leq \tau(A) - 1 < n-1$ or $\tau(A) \geq n$. In both cases, we obtain from Proposition 4.3.3:

$$\rho(A_n(x)) = 1 - Q^{\tau(A)-\kappa_A+1}(x_{\kappa_A-1}, x_0) \prod_{j=0}^{\kappa_A-2} Q(x_j, x_{j+1}) \geq 1 - \max\{q_0, 1-q_0\}. \tag{4.4}$$

Otherwise, when $\ell = n-1$ and $x_i \neq 0$, for $i < n-1$, then $\tau(A) \geq n$ and we must have $X_n^{\tau(A)} = 12\cdots(x_0-1)x_0$, since this is the shortest path (with positive measure) to return to x_0^{n-1}. Thus, we conclude that $Q^{\tau(A)-n+1}(x_{n-1}, x_0) \leq q_0 \leq \max\{q_0, 1-q_0\}$ and (4.4) also holds in this case.

On the other hand, if $A_n(x)$ has no zeros, it must have the form:

$$A_n(x) = k(k+1)\cdots(k+n-1),$$

for some $k \geq 1$. In this case, the shortest possible return occurs when, immediately after $A_n(x)$, we have the sequence $01\cdots k\cdots(k+n-1)$, which implies

$$\mu_A(T_A = \tau(A)) = (1 - q_{k+n-1})\sigma_{k+n} \leq q_0 \leq \max\{q_0, 1-q_0\}.$$

Thus, for any choice of $(q_i)_{i\geq 0}$, we have

$$\inf_{n\geq 1, x\in\mathscr{X}} \rho(A_n(x)) \geq 1 - \max\{q_0, 1-q_0\} > 0.$$

Further, when $1 - q_0 \geq q_0$, the above inequality turns into an equality, since the point $x = 0^\infty$ satisfies $\rho(A_n(x)) = q_0$ for all $n \geq 1$.

Note that, by Corollary 2.2.5, when $q_i \longrightarrow 1$, the chain $\mathbf{X}$ does not satisfy the hypothesis of Theorem 4.2.1, but satisfies its two conclusions.

Convergence

By Proposition 4.3.3, for any Markov chain and $x \in \mathscr{X}$ we have that $\rho(x)$ exists and is positive. This means that, when the chain is ϕ-mixing, (4.1) and (4.2) also hold in this case as well as the limiting moments given in (4.3).

For instance, consider $q_i = p(i+1)/(i+2)$ for $i \geq 0$ and $p \in (0,1)$, which gives us an exponentially ϕ-mixing chain. For the periodic point $x = 01\cdots k01\cdots k\cdots$ and $n \geq k$, we have

$$\rho(A_n(x)) = 1 - (1-q_k)\sigma_k = 1 - \frac{p^k(k(1-p)+2-p)}{(k+1)(k+2)} = \rho(x).$$

Example 4.3.6 (The lazy house of cards Markov chain)**.** Recall from Example 2.1.4 the "lazy" version of the house of cards Markov chain. Under certain conditions on the sequence $(q_i)_{i\geq 0}$, this chain is at least β-mixing, but it can be exponentially ϕ-mixing when, for instance, $\sup_i\{q_i\} < 1$. In this case, we have another example of process over infinite alphabet and without complete grammar for which we can apply Theorem 3.3.4 and further results of this chapter.

The case when $\rho(A) \approx 0$

When $\sup_i\{q_i\} = 1$, Proposition 4.3.3-(d) implies

$$\inf_{n\geq 1, x\in\mathscr{X}} \rho(A_n(x)) = 0,$$

which means that the chain cannot be ϕ-mixing in view of Theorem 4.2.1. Indeed, for any $n \in \mathbb{N}$, we have $\rho(A_n(x))$ arbitrarily close to zero when we take the infimum over the points $x \in \mathscr{X}$ satisfying $\tau(x) = 1$, since

$$\inf_{m\in\mathbb{N}} \rho\left(A_n\left(m^\infty\right)\right) = \inf_{m\in\mathbb{N}}(1-q_m) = 0, \quad \forall n \geq 1.$$

By the way, this case provides another example where the potential well strongly influences the hitting and return time distributions. Applying a similar argument as we did in Proposition 4.2.7, we can conclude:

- If $A = m^n$, then, for any $\beta > 0$, $\mu(A)^\beta\,\mathbb{E}\left(T_A^\beta\right)$ is arbitrarily large as m grows, which also holds for $\beta > 1$ and $\mu(A)^\beta \mathbb{E}_A\left(T_A^\beta\right)$. In particular, the more m grows, the larger is the difference between the mean hitting time and the mean return time.
- For each fixed $t > 0$, $\mu\left(T_A > \mu(A)^{-1}t\right)$ is arbitrarily close to one as m grows, while $\mu_A\left(T_A > \mu(A)^{-1}t\right)$ is arbitrarily close to zero.

Almost sure positivity

Let us show that the potential well is bounded away from zero at points with period greater than one. Notice that if the word $A_n(x)$ has at least one symbol zero, the same reasoning as we did in Example 4.3.5 leads to $\rho(A_n(x)) \geq 1 - \max\{q_0, 1-q_0\}$. Otherwise, if $A_n(x)$ does not have a symbol zero and satisfies $\tau(A_n(x)) > 1$, then $A_n(x)$ has at least one transition of the type $k \to k+1$, which implies:

$$\rho(A_n(x)) = 1 - Q^{\tau(A)-\kappa_A+1}(x_{\kappa_A-1}, x_0) \prod_{j=0}^{\kappa_A-2} Q(x_j, x_{j+1}) \geq 1 - Q(k, k+1) \geq \frac{1}{2}$$

Therefore, the lazy house of cards always satisfies

$$\inf_{n\geq 1, x:\tau(x)>1} \rho(A_n(x)) > 0.$$

Renewal processes

Finally, let us deal with the class of renewal processes. First we recall some results obtained by Abadi, Cardeno and Gallo (2015).

Theorem 4.3.7. (ABADI; CARDENO; GALLO, 2015). Let $\mathbf{Y}$ be a binary renewal process with law μ.

(a) If $x = 0^k 1 x_{k+1}^{\infty} \in \mathscr{X}$, then for any $n \geq k+1$ we have:

$$\rho(A_n(x)) = 1 - \mu_{10^k}\left(x_{\tau(A_n(x))-1} \cdots x_{k+2} x_{k+1} 1\right) \sigma_k.$$

(b) $\rho(A_n(x)) \xrightarrow{n} 1$ if, and only if, x is aperiodic.

(c) There exists $C \in (0,1)$ such that any periodic point $x \neq 0^\infty$ satisfies

$$\rho(A_n(x)) \xrightarrow{n} \rho(x) \in [C, 1).$$

(d) $\rho(0^n) = \dfrac{\sigma_n}{\Sigma(n)} = \dfrac{1}{\sum_{j=0}^{\infty} \frac{\sigma_{n+j}}{\sigma_n}}$.

In the next proposition, we add two properties to Theorem 4.3.7.

Proposition 4.3.8. Consider a binary renewal process $\mathbf{Y}$ defined from a sequence $(q_i)_{i\geq 0}$. Then:

(a) $\inf\limits_{n\geq 1; A\neq 0^n} \rho(A) \geq 1 - \max\{q_0, 1-q_0\}$ and the equality holds when $1 - q_0 \geq q_0$.

(b) $\inf\limits_{n\geq 1} \rho(0^n) > 0$ if, and only if, $\mathbf{Y}$ is ϕ-mixing.

Example 4.3.9. Let us show the different regimes provided for the class of the renewal processes.

ρ converges to zero

Consider a renewal process defined by the function $\Sigma(n) = \frac{1}{n}$ and take $A = 0^n$. In this case, we have

$$\rho(A) = 1 - \frac{\mu\left(0^{n+1}\right)}{\mu\left(0^n\right)} = \frac{1}{n+1},$$

which means that the process is not ϕ-mixing, by Proposition 4.3.8.

Notice that this case where we have $\inf_{n,x} \rho(A_n(x)) = 0$ is different from Example 4.3.6, since here $\rho(0^n) \xrightarrow{n} 0$ and there $\rho\left(m^n\right) \xrightarrow{n} 1 - q_m > 0$ (with $q_i \longrightarrow 1$), although in both cases we cannot obtain a positive and uniform lower bound for the potential well.

This example satisfies the conditions of Proposition 4.2.7. In this case, one can obtain an explicit inequality involving $\mathbb{E}\left(T_A^\beta\right)$ and $\mu(A)^{-\beta}$. Applying Theorem 3.1 in (ABADI; CARDENO; GALLO, 2015) together with the Markov inequality we get

$$\begin{aligned}\mathbb{E}\left(T_A^\beta\right) &\geq (\mu(A)\rho(A))^{-\beta}\mu\left(T_A > (\mu(A)\rho(A))^{-1}\right)\\ &\geq \mu(A)^{-\beta}(n+1)^\beta\left(e^{-1} - C\varepsilon(A)\right)\\ &\gg \mu(A)^{-\beta},\end{aligned}$$

since $\varepsilon(A) \xrightarrow{n} 0$. Notice that, for $\beta = 1$, the convergence of $\rho(A)$ to zero makes the mean hitting time to A much larger than its mean return time.

ρ is positive and does not converge

The point $x = 0^\infty$ can also provide a case in which $\mu\left(T_A > \mu(A)^{-1}t\right)$ does not converge. Consider the renewal processes defined by the sequence $(q_i)_{i\geq 0}$ given by $q_i = p$ for even i and $q_i = q$ for odd i, where $0 < p < q < 1$. Then, direct computations give

$$\rho(0^n) = \frac{1}{1 + q_n + q_n q_{n+1} + \cdots} = \begin{cases} \frac{1-pq}{1+p}, & \text{if } n \text{ is even.}\\ \frac{1-pq}{1+q}, & \text{if } n \text{ is odd.}\end{cases}$$

Applying Theorem 3.1 in (ABADI; CARDENO; GALLO, 2015), we get

$$e^{-\rho(0^n)t} - C\varepsilon(0^n) \leq \mu\left(T_{0^n} > \mu\left(0^n\right)^{-1}t\right) \leq e^{-\rho(0^n)t} - C\varepsilon(0^n),$$

which means that, for all $t \geq 0$, $\mu\left(T_{0^n} > \mu\left(0^n\right)^{-1}t\right)$ oscillates between arbitrarily small neighbourhoods of $e^{-\frac{(1-pq)t}{1+p}}$ and $e^{-\frac{(1-pq)t}{1+q}}$, since $\varepsilon(0^n) \longrightarrow 0$. Note that Corollary 4.2.4 implies $\mathbb{E}(T_{0^n}) \sim \mu(0^n)^{-1} = \mathbb{E}_{0^n}(T_{0^n})$, since this process is ψ-mixing by Theorem 2.2.6.

ρ converges to a positive number

The last regime involving the point $x = 0^\infty$ occurs when there exists $\rho(x) \in (0,1)$. For instance, when $q_i \xrightarrow{i} L \in (0,1)$, we will prove in Lemma 5.4.5 that $\rho(0^n) \xrightarrow{n} 1 - L$. In this case, (4.1) and (4.2) hold. On the other hand, for any other point $x \in \mathscr{X} - \{0^\infty\}$, Theorem 4.3.7 says that $\rho(x)$ exists and is positive. For instance, for the periodic point $x = 1100111001\cdots$, we have $\rho(x) = 1 - (1-q_0)^2(1-q_2)\sigma_2$.

4.4 Shortest return time

We present in this section some considerations and two results on the shortest possible return time. We start by noting that for any $x \in \mathscr{X}$, $\tau(A_n(x))$ is an increasing sequence. Then we obviously have $\tau(A_n(x)) \xrightarrow{n} \tau(x) = \sup\{\tau(A_n(x)); n \geq 1\}$. Recall that a periodic point is one that satisfies $\tau(x) = p < \infty$. It is immediate to verify that x is periodic if, and only if, there exists $x_0^{p-1} \in \mathscr{A}^p$ such that

$$x = x_0^{p-1}x_0^{p-1}x_0^{p-1}\cdots.$$

Naturally, in the context of complete grammar, we have $\tau(A_n(x)) \leq n$. On the other hand, the string $A = (k, \cdots, k+n-1)$ generated by a house of cards Markov chain satisfies $\tau(A_n(x)) = k+n$, showing that the shortest return can be larger than n.

However, we can get upper bounds for $\tau(A_n(x))$. For instance, applying (3.2) with $t = \tau(A)$, we get

$$1 \geq \mu(T_A \leq \tau(A)) = \tau(A)\mu(A), \tag{4.5}$$

that is, $\tau(A) \leq \mu(A)^{-1}$, which holds for any stationary process. In ergodic and aperiodic processes, the above inequality must be strict, since the equality would lead to existence of a periodic point x such that $\mu(x) > 0$ (see Lemma 3.4.1). Notice that, in view of the Shannon–McMillan–Breiman theorem, the above inequality implies

$$\limsup_{n\to\infty} \frac{1}{n} \ln \tau(A_n(x)) \leq h_\mu, \quad \text{a.s.},$$

which gives us an exponential upper bound for $\tau(A)$.

This bound is related to the potential well. The next proposition shows that, if the above inequality is strict, then $\rho(A)$ converges to one almost surely.

Proposition 4.4.1. Let $\mathbf{X}$ be an ergodic process over a finite alphabet with law μ satisfying $h_\mu > 0$ and

$$\limsup_{n\to\infty} \frac{1}{n} \ln \tau(A_n(x)) < h_\mu, \quad \text{a.s.} \tag{4.6}$$

Then, $\rho(A_n(x)) \xrightarrow{n} 1$ almost surely.

Remark 4.4.2. Notice that Proposition 4.4.1 implies that, for ergodic processes with positive entropy and finite alphabet, we have $\rho(A) \longrightarrow 1$ almost surely whenever the set $\{x \in \mathscr{X}; \limsup_n (1/n)\ln(\tau(A_n(x)) = h_\mu\}$ has measure zero.

The next theorem provides some useful conditions and respective upper bounds for the shortest return time. In particular, we give three classes of processes that satisfy (4.6).

Theorem 4.4.3. Consider a stationary process with law μ.

(a) For any $C > 1$ there exists n_0 such that for all $n \geq n_0$ and $A \in \mathscr{C}_n$, we have

- $\tau(A) \leq Cn$, if μ is ψ-mixing.
- $\tau(A) \leq \dfrac{C\mu(A)^{-1}}{\ln(\mu(A)^{-1})} + n$, if is summable ϕ-mixing.
- $\tau(A) \leq \dfrac{\mu(A)^{-1}}{\ln(\mu(A)^{-1})} + n$, if μ is α-mixing satisfying $\alpha(j) \leq j^{-r}$ for some $r > 2$.

(b) Suppose that $h_\mu > 0$. Then, if any of the conditions (b1), (b2) or (b3) given below holds, there exists $\delta \in (0, h_\mu]$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \ln \tau(A_n(x)) \leq h_\mu - \delta, \quad \text{a.s.}$$

(b1) μ is ψ-mixing;

(b2) μ is ϕ-mixing over a finite alphabet such that $\phi(j) \leq j^{-s}$ for some $s > 1$;

(b3) μ is α-mixing over a finite alphabet such that $\alpha(j) \leq j^{-r}$ for some $r > 2$.

Remark 4.4.4. There are results in the literature indicating that $\tau(A)$ is of a much lower order than $e^{n(h_\mu - \delta)}$. Recall from Remark 4.2.2 that Saussol, Troubetzkoy and Vaienti (2002) and Afraimovich, Chazottes and Saussol (2003) showed that, under certain conditions, $\tau(A_n(x))/n \longrightarrow 1$ almost surely (note that for periodic points, $\tau(A_n(x))/n \longrightarrow 0$). However, if we drop the specification property or the finite alphabet or analyse the "all points" case, the order of $\tau(A_n(x))$ is an open question. As far as we know, there is no known example of a process with $\tau(A_n(x))$ that has exponential order for some x.

Remark 4.4.5. Item (a) of Theorem 4.4.3 were stated for ψ and ϕ with $C = 2$ in Lemma 3.4.4, as a preliminary result for the proof of Theorem 3.3.4. The choice of constant 2 in both cases is due to technical purposes, for the sake of simplicity. As stated in Theorem 4.4.3, it can be replaced by any constant larger than one. Note that we cannot take $C = 1$ in the ψ-mixing case. If we consider an irreducible aperiodic finite Markov chain, where the transition matrix has at least one entry equal to zero, then we can get $A_n(x)$ such that $\tau(A_n(x)) > n$ for infinitely many values of n. The optimality of this bound for the ϕ-mixing or α-mixing cases is an open question.

Remark 4.4.6. Note that Lemma 3.4.3 and Theorem 4.4.3 imply that $\sup_{A\in\mathscr{C}_n} \tau(A)\mu(A) \xrightarrow{n} 0$ for the cases ψ and summable ϕ. For the case α, this convergence holds if $\sup_{A\in\mathscr{C}_n} n\mu(A)$ converges to zero, which is not always the case. When the alphabet is finite, Item (a) of Theorem 4.4.3 and the Shannon–McMillan–Breiman theorem imply the almost sure convergence of $\tau(A_n(x))\mu(A_n(x))$ to zero.

4.5 Proofs

Recall that if $A = a_0^{n-1}$, $A^{(k)}$ denotes the string a_{n-k}^{n-1} of the last k letters of A.

Proof of Theorem 4.2.1. (a) Consider first the condition (a1). Let $\mathscr{B} = \{x \in \mathscr{X}; \tau(x) = \infty\}$ be the set of aperiodic points of $\mathscr{X}$. For $x \in \mathscr{B}$, denote $A_n = A_n(x)$, and consider the case $\tau(A_n) < n$. Then, we have:

$$\begin{aligned} 1-\rho(A_n) &= \mu_{A_n}\left(\sigma^{-n}\left(A_n^{(\tau(A_n))}\right)\right) \\ &\le \mu_{A_n}\left(\sigma^{-n-\lfloor\tau(A_n)/2\rfloor}\left(A_n^{(\lceil\tau(A_n)/2\rceil)}\right)\right) \\ &\le \mu\left(A_n^{(\lceil\tau(A_n)/2\rceil)}\right) + \phi\left(\lfloor\tau(A_n)/2\rfloor+1\right). \end{aligned}$$

Since $x \in \mathscr{B}$, we have $\tau(A_n) \xrightarrow{n} \infty$, which implies that the last expression converges to zero.

For the case $\tau(A_n) \ge n$, we use a similar argument:

$$\begin{aligned} 1-\rho(A_n) &= \mu_{A_n}\left(\sigma^{-\tau(A_n)}(A_n)\right) \\ &\le \mu_{A_n}\left(\sigma^{-\tau(A_n)-\lfloor n/2\rfloor}\left(A_n^{(\lceil n/2\rceil)}\right)\right) \\ &\le \mu\left(A_n^{(\lceil n/2\rceil)}\right) + \phi\left(\lfloor n/2\rfloor+1\right) \end{aligned}$$

which also converges to zero. Therefore, $\rho(A_n) \xrightarrow{n} 1$. We conclude the proof in this case by noting that $\mathscr{X} - \mathscr{B}$ is a countable set, and thus, $\mu(\mathscr{B}) = 1$.

The case involving α-mixing immediately follows from Proposition 4.4.1 and Theorem 4.4.3.

(b) By Remark 4.4.6, for ψ-mixing or summable ϕ-mixing measures, there exists $n_0 \ge 1$ such that:

$$\forall n \ge n_0,\ \forall A \in \mathscr{A}^n,\ \ \mu(A)^{-1} > \tau(A).$$

Now, since $\mu_A(T_A > j), j \ge 1$ is a nonincreasing sequence, the potential well is larger than or equal to the arithmetic mean of the subsequent $\mu(A)^{-1} - \tau(A)$ elements:

$$\begin{aligned} \rho(A) = \mu_A(T_A > \tau(A)) &\ge \frac{1}{\mu(A)^{-1} - \tau(A)} \sum_{j=\tau(A)}^{\mu(A)^{-1}-1} \mu_A(T_A > j) \\ &\ge \frac{1}{\mu(A)^{-1}} \sum_{j=\tau(A)}^{\mu(A)^{-1}-1} \mu_A(T_A > j) \\ &= \sum_{j=\tau(A)}^{\mu(A)^{-1}-1} \mu(A; T_A > j) \\ &= \sum_{j=\tau(A)}^{\mu(A)^{-1}-1} \mu(T_A = j+1). \end{aligned} \tag{4.7}$$

In the last equality, we used Lemma 3.4.2. By (4.7), one obtains:

$$\begin{aligned} \rho(A) &\ge \mu(T_A \le \mu(A)^{-1}) - \mu(T_A \le \tau(A)) \\ &= \mu(T_A \le \mu(A)^{-1}) - \tau(A)\mu(A) \end{aligned} \tag{4.8}$$

where the equality follows by stationarity and the definition of $\tau(A)$.

Applying again Remark 4.4.6, we know that $\tau(A)\mu(A) \xrightarrow{n} 0$ uniformly. Thus, it is enough to find a strictly positive lower bound for $\mu(T_A \leq \mu(A)^{-1})$. Let:

$$N = \sum_{j=1}^{\mu(A)^{-1}} \mathbb{1}_A \circ \sigma^j$$

which counts the number of occurrences of A up to $\mu(A)^{-1}$. By the so-called second moment method,

$$\mu(T_A \leq \mu(A)^{-1}) = \mu(N \geq 1) \geq \frac{\mathbb{E}(N)^2}{\mathbb{E}(N^2)}. \tag{4.9}$$

Stationarity gives $\mathbb{E}(N) = 1$. It remains to prove that $\mathbb{E}(N^2)$ is bounded above by a constant. Expanding N^2, using stationarity and $\mathbb{E}(N) = 1$, we obtain:

$$\mathbb{E}(N^2) = 1 + 2\sum_{j=1}^{\mu(A)^{-1}} (\mu(A)^{-1} - j)\,\mu(A \cap \sigma^{-j}(A)). \tag{4.10}$$

Let us first consider the ϕ-mixing case. For $j \geq n$, mixing gives $\mu(A \cap \sigma^{-j}(A)) \leq \mu(A)^2 + \mu(A)\phi(j-n+1)$. Thus,

$$\sum_{j=n}^{\mu(A)^{-1}} (\mu(A)^{-1} - j)\,\mu(A \cap \sigma^{-j}(A)) \leq \frac{1}{2} + \sum_{\ell=0}^{\mu(A)^{-1}-n} \phi(\ell+1) \tag{4.11}$$

where we used $\mu(A)^{-1} - j \leq \mu(A)^{-1}$ to get the last term.

For $1 \leq j \leq n-1$, as before $A^{(j)} \subset A^{(\lceil j/2 \rceil)}$; thus:

$$\begin{aligned}
\mu\left(A \cap \sigma^{-j}(A)\right) &= \mu\left(A \cap \sigma^{-n}\left(A^{(j)}\right)\right) \\
&\leq \mu\left(A \cap \sigma^{-n-\lfloor j/2 \rfloor}\left(A^{(\lceil j/2 \rceil)}\right)\right) \\
&\leq \mu(A)\left(\mu\left(A^{(\lceil j/2 \rceil)}\right) + \phi(\lfloor j/2+1 \rfloor)\right) \\
&\leq \mu(A)\left(Ce^{-c\lceil j/2 \rceil} + \phi(\lfloor j/2+1 \rfloor)\right).
\end{aligned}$$

Therefore,

$$\sum_{j=1}^{n-1} (\mu(A)^{-1} - j)\,\mu(A \cap \sigma^{-j}(A)) \leq \sum_{j=1}^{n-1} \left(Ce^{-c\lceil j/2 \rceil} + \phi(\lfloor j/2+1 \rfloor)\right). \tag{4.12}$$

Therefore, by (4.11) and (4.12), the summability of ϕ concludes the proof for the ϕ-mixing case.

If μ is ψ-mixing, we separate the sum in (4.10) into three parts. First, recall the definition of g_0 in Section 3.3. For $1 \leq j \leq g_0$, we bound the sum as follows:

$$\sum_{j=1}^{g_0} (\mu(A)^{-1} - j)\,\mu(A \cap \sigma^{-j}(A)) \leq \sum_{j=1}^{g_0} \mu(A)^{-1}\mu(A) = g_0.$$

For $g_0+1 \le j \le g_0+n-1$, we have by ψ-mixing:

$$(\mu(A)^{-1}-j)\,\mu(A\cap\sigma^{-j}(A)) \le \mu(A)^{-1}\mu\left(A\cap\sigma^{-n-g_0}\left(A^{(j-g_0)}\right)\right) \le M\mu(A)^{-1}\mu(A)\mu\left(A^{(\ell)}\right)$$

where we denoted $\ell = j-g_0$. Thus:

$$\sum_{j=g_0+1}^{g_0+n-1}(\mu(A)^{-1}-j)\,\mu(A\cap\sigma^{-j}(A)) \le M\sum_{\ell=1}^{n-1}Ce^{-c\ell}.$$

Finally, applying ψ-mixing again,

$$\sum_{j=g_0+n}^{\mu(A)^{-1}}(\mu(A)^{-1}-j)\,\mu(A\cap\sigma^{-j}(A)) \le M\sum_{j=n+g_0}^{\mu(A)^{-1}}\mu(A)^{-1}\mu(A)^2 \le M,$$

concluding the proof of the ψ-mixing case.

We conclude by noting that for every $n<n_0$ and $A\in\mathscr{C}_n$, we have $\rho(A)>0$. Thus if the alphabet $\mathscr{A}$ is finite, the set $\{\rho(A) : A\in\mathscr{C}_n, n<n_0\}$ is finite and has a strictly positive infimum, which implies that the infimum of $\{\rho(A_n)\}$ over all $n\ge 1$ and $A_n\in\mathscr{C}_n$ is also strictly positive. □

Lemma 4.5.1. Let μ be a ϕ-mixing measure. There exists $n_0\in\mathbb{N}$ such that for $n\ge n_0$ and $A\in\mathscr{C}_n$, the following statements hold.

(a) There exist $t_0>0$ and constants $C,c>0$ such that for all $t\ge t_0$, $\mu(T_A>t)\le Ce^{-ct}$.

(b) For the same t_0 as Item (a) and $t\ge t_0$, we have $\mu_A(T_A>t)\le C'e^{-ct}$, for some $C'>0$.

(c) For any $\beta>0$, we have $\mathbb{E}\left((\mu(A)T_A)^\beta\right)$, $\mathbb{E}_A\left((\mu(A)T_A)^\beta\right)<\infty$.

Proof. (a) First, we take a fixed $\xi\in(0,1)$ and $n_0\in\mathbb{N}$ such that $\phi(n)\le\xi/2$ for all $n\ge n_0$. Since μ is ergodic, for each $n\ge n_0$ and $A\in\mathscr{C}_n$, we can obtain $t_0\gg n$ such that $\mu(T_A>t_0-2n)\le\xi/2$. On the other hand, for $t\ge t_0$ we write $t=kt_0+r$, where $k\ge 1$ and $0\le r<t_0$. Then we have

$$\begin{aligned}\mu(T_A>t)\le\mu(T_A>kt_0-2n)&=\mu\left(T_A>(k-1)t_0; T_A^{[(k-1)t_0]}>t_0-2n\right)\\&\le\mu\left(T_A>(k-1)t_0-2n; T_A^{[(k-1)t_0]}>t_0-2n\right)\end{aligned}$$

By ϕ-mixing definition and stationarity we get

$$\mu(T_A>kt_0-2n)\le\mu\left(T_A>(k-1)t_0-2n\right)\left(\mu\left(T_A>t_0-2n\right)+\phi(n)\right).$$

Iterating over k, the last term is bounded by

$$\left(\mu\left(T_A>t_0-2n\right)+\phi(n)\right)^k\le\xi^{\frac{t-r}{t_0}}\le\xi^{t/t_0-1}$$

and just take $C = \xi^{-1}$ and $c = -\ln\left(\xi^{1/t_0}\right)$.

(b) Similarly to item (a), one obtains

$$\mu_A(T_A > t) \leq \mu_A(T_A > kt_0 - 2n) \leq \mu_A\left(T_A > (k-1)t_0 - 2n; T_A^{[(k-1)t_0]} > t_0 - 2n\right). \tag{4.13}$$

Note that, when $A, B \in \mathscr{F}_0^i$ and $C \in \mathscr{F}_{i+n}^{\infty}$, the definition of $\phi(n)$ implies $\mu_A(B;C) \leq \mu_A(B)(\mu(C) + \phi(n))$. Thus, the rightmost term in (4.13) is bounded by

$$\mu_A(T_A > (k-1)t_0 - 2n)(\mu(T_A > t_0 - 2n) + \phi(n)) \leq \xi^{k-1} \leq \xi^{t/t_0 - 2}$$

and we take $C' = \xi^{-2}$.

(c) For any $\beta > 0$, it follows from Item (a) that

$$\begin{aligned}
\mathbb{E}\left((\mu(A)T_A)^\beta\right) &= \beta \int_0^\infty t^{\beta-1} \mu\left(T_A > \mu(A)^{-1}t\right) dt \\
&\leq \beta \int_0^{\mu(A)t_0} t^{\beta-1} dt + C\beta \int_{\mu(A)t_0}^{\infty} t^{\beta-1} e^{-c\mu(A)^{-1}t} dt \\
&\leq (\mu(A)t_0)^\beta + \frac{C\,\Gamma(\beta+1)}{(c\mu(A)^{-1})^\beta}.
\end{aligned}$$

Applying Item (b) and the same reasoning as above, we also obtain

$$\mathbb{E}_A\left((\mu(A)T_A)^\beta\right) \leq (\mu(A)t_0)^\beta + \frac{C'\,\Gamma(\beta+1)}{(c\mu(A)^{-1})^\beta}.$$

□

Proof of Corollary 4.2.4. (a) Consider a random variable $Y \geq 0$ with distribution

$$\mu(Y > t) = e^{-\rho(A)t}.$$

Then,

$$\mathbb{E}\left(Y^\beta\right) = \beta \int_0^\infty t^{\beta-1} e^{-\rho(A)t} dt = \frac{\Gamma(\beta+1)}{\rho(A)^\beta}.$$

On the other hand, we apply Item (c) of Lemma 4.5.1 to get

$$\begin{aligned}
\left|\mathbb{E}\left((\mu(A)T_A)^\beta\right) - \mathbb{E}\left(Y^\beta\right)\right| &= \beta \left| \int_0^\infty t^{\beta-1} \left(\mu\left(T_A > \mu(A)^{-1}t\right) - \mu(Y > t)\right) dt \right| \\
&\leq \beta \int_0^\infty t^{\beta-1} \left| \mu\left(T_A > \mu(A)^{-1}t\right) - e^{-\rho(A)t} \right| dt.
\end{aligned}$$

Applying Theorem 3.3.4 and Remark 3.3.5, we have a constant C such that the last term is bounded by

$$C\beta\left(\int_0^{1/2} t^{\beta-1}\varepsilon(A) dt + \int_{1/2}^\infty \varepsilon(A) t^\beta e^{-t(\rho(A) - C_3\varepsilon(A))} dt\right) \leq \frac{C\varepsilon(A)}{2^\beta} + \frac{C\beta\Gamma(\beta+1)\varepsilon(A)}{(\rho/2)^{\beta+1}},$$

where we apply Theorem 4.2.1-(b) to obtain for n large enough $\rho(A) - C_3\varepsilon(A) \geq \rho/2$. This ends the proof of Item (a).

(b) Consider a random variable $Z \geq 0$ with distribution

$$\mu(Z > t) = \begin{cases} 1 & t < \tau(A)\mu(A), \\ \rho(A)e^{-\rho(A)t} & t \geq \tau(A)\mu(A) \end{cases}.$$

Then,

$$\begin{aligned} \mathbb{E}\left(Z^\beta\right) &= \beta \int_0^\infty t^{\beta-1}\mu(Z>t)dt \\ &= \beta\left(\int_0^{\tau(A)\mu(A)} t^{\beta-1}dt + \rho(A)\int_{\tau(A)\mu(A)}^\infty t^{\beta-1}e^{-\rho(A)t}dt\right) \\ &\leq (\tau(A)\mu(A))^\beta + \frac{\Gamma(\beta+1)}{\rho(A)^{\beta-1}}. \end{aligned}$$

Note that in this case we also have

$$\mathbb{E}\left(Z^\beta\right) \geq \beta\rho(A)\int_0^\infty t^{\beta-1}e^{-\rho(A)t}dt = \frac{\Gamma(\beta+1)}{\rho(A)^{\beta-1}},$$

which implies

$$\left|\mathbb{E}\left(Z^\beta\right) - \frac{\Gamma(\beta+1)}{\rho(A)^{\beta-1}}\right| \leq (\tau(A)\mu(A))^\beta.$$

On the other have, we have

$$\left|\mathbb{E}_A\left((\mu(A)T_A)^\beta\right) - \frac{\Gamma(\beta+1)}{\rho(A)^{\beta-1}}\right| \leq \left|\mathbb{E}_A\left((\mu(A)T_A)^\beta\right) - \mathbb{E}\left(Z^\beta\right)\right| + (\tau(A)\mu(A))^\beta.$$

Since $\tau(A)\mu(A) \leq 2\varepsilon(A)$ (see Remark 3.3.5), we just have to bound the first modulus in the last sum.

Now, we apply Lemma 4.5.1 in a similar way to Item (a):

$$\left|\mathbb{E}_A\left((\mu(A)T_A)^\beta\right) - \mathbb{E}\left(Z^\beta\right)\right| \leq \beta\int_0^\infty t^{\beta-1}\left|\mu_A\left(T_A > \mu(A)^{-1}t\right) - \mu(Z>t)\right|dt. \tag{4.14}$$

When $t < \tau(A)\mu(A)$, the modulus in the above integral is equal to zero. For $t \geq \tau(A)\mu(A)$, it is bounded by

$$\begin{aligned} &\left|\mu_A\left(T_A > \mu(A)^{-1}t\right) - \rho(A)e^{-\rho(A)(t-\tau(A)\mu(A))}\right| + \rho(A)e^{-\rho(A)t}\left|e^{\rho(A)\mu(A)\tau(A)} - 1\right| \\ &\leq \varepsilon(A,t) + e^{-\rho(A)t}e^{\rho(A)\mu(A)\tau(A)}\rho(A)\mu(A)\tau(A), \end{aligned} \tag{4.15}$$

where $\varepsilon(A,t)$ is the error term given by Theorem 3.3.4 and the second term follows from the Mean Value Theorem.

Let us compute the integral in (4.14) separately for the two last terms. For the first one, we get

$$\begin{aligned}\int_{\tau(A)\mu(A)}^{\infty} t^{\beta-1}\varepsilon(A,t)dt &\le C\varepsilon(A)\int_{\tau(A)\mu(A)}^{1/2} t^{\beta-1}dt + C'\varepsilon(A)\int_{1/2}^{\infty} t^{\beta}e^{-t(\rho(A)-C_3\varepsilon(A))}dt \\ &\le \frac{C\varepsilon(A)}{\beta 2^{\beta}} + \frac{C'\varepsilon(A)\Gamma(\beta+1)}{(\rho(A)-C_3\varepsilon(A))^{\beta+1}} \\ &\le \varepsilon(A)\left(\frac{C}{\beta 2^{\beta}} + \frac{C'\Gamma(\beta+1)}{(\rho/2)^{\beta+1}}\right),\end{aligned}$$

where we used again $\rho(A) - C_3\varepsilon(A) \ge \rho/2$.

Finally, consider the second term in (4.15). Since $\rho(A)\mu(A)\tau(A) \longrightarrow 0$ and $\tau(A)\mu(A) \le 2\varepsilon(A)$, we have for large n:

$$e^{-\rho(A)t}e^{\rho(A)\mu(A)\tau(A)}\rho(A)\mu(A)\tau(A) \le 4\varepsilon(A)e^{-\rho(A)t}.$$

Therefore, the integral for the second term (4.15) is bounded by

$$4\varepsilon(A)\int_0^{\infty} t^{\beta-1}e^{-\rho(A)t}dt = \frac{4\varepsilon(A)\Gamma(\beta)}{\rho(A)^{\beta}} \le \frac{4\varepsilon(A)\Gamma(\beta)}{\rho^{\beta}},$$

which ends the proof of Item (b).

We finish the proof by noting that Items (c) and (d) follows directly from (a), (b), Theorem 4.2.1 and Lemma 3.4.3. □

Proof of Corollary 4.2.6. (a) For $t < \tau(A)$, we have $\mu_A(T_A > t) = 1$ and, applying (3.2), we have $\mu(T_A > t) = 1 - t\mu(A)$, which implies

$$|\mu(T_A > t) - \mu_A(T_A > t)| = t\mu(A) \le \tau(A)\mu(A).$$

On the other hand, for $t \ge \tau(A)$, Proposition 3.4.6 gives us

$$\begin{aligned}|\mu(T_A > t) - \mu_A(T_A > t)| &\le |\rho(A)\mu(T_A > t) - \mu_A(T_A > t)| + \mu(T_A > t)\,|1-\rho(A)| \\ &\le C\varepsilon(A) + |1-\rho(A)|.\end{aligned}$$

Thus, by Theorem 4.2.1 we get

$$|\mu(T_A > t) - \mu_A(T_A > t)| \le \max\{\tau(A)\mu(A), C\varepsilon(A) + |1-\rho(A)|\} \xrightarrow{n} 0, \quad \text{a.s.}$$

(b) Applying Theorems 3.3.4 and 4.2.1, one has for all $t \ge 0$:

$$\mu\left(T_A > \mu(A)^{-1}t\right) \xrightarrow{n} e^{-t}, \quad \text{a.s.}$$

We finish the proof by noting that Item (a) ensures the same convergence for the rescaled return time $\mu_A\left(T_A > \mu(A)^{-1}t\right)$. □

Proof of Proposition 4.2.7. (a) First we consider the case of the return time for $t \geq 1$. In this case, we have $t \geq 1 \geq \tau(A)\mu(A)$ (see (4.5)). Then, $\mu_A\left(T_A > \mu(A)^{-1}t\right) \leq \rho(A) \longrightarrow 0$. Now, for each $t \in (0,1)$, we use $\tau(A)\mu(A) \longrightarrow 0$ to get $n_0 \in \mathbb{N}$ such $t > \tau(A)\mu(A)$ for all $n \geq n_0$, which also gives us the above convergence to zero.

Consider now the hitting time. For each $t > 0$ we take n_0 satisfying the same as above. Then (3.2) gives us

$$\mu\left(T_A \leq \mu(A)^{-1}t\right) \leq \tau(A)\mu(A) + \left(\mu(A)^{-1}t - \tau(A)\right)\mu(A)\rho(A),$$

which, by the hypotheses, converges to zero for each $t \geq 0$.

(b) Consider the scaling parameter $\lambda(A)$ defined by Abadi and Saussol (2011). Applying their Theorem 1 and the Markov inequality we get large $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and $\beta > 0$:

$$\mathbb{E}\left(T_A^\beta\right) \geq (\mu(A)\lambda(A))^{-\beta}\mu\left(T_A > (\mu(A)\lambda(A))^{-1}\right) \geq (\mu(A)\lambda(A))^{-\beta}\frac{e^{-1}}{2}.$$

According to Corollary 6 in (ABADI; SAUSSOL, 2016), if $\liminf_n \rho(A) = 0$, then $\liminf_n \lambda(A) = 0$, which means that $\limsup_n \lambda(A)^{-\beta} = \infty$. The conclusion for hitting times follows from the above inequalities. Concerning return time, applying the same theorem and the Markov inequality for conditional measure, we get

$$\mu(A)^\beta \mathbb{E}_A\left(T_A^\beta\right) \geq \lambda(A)^{-\beta+1}\frac{e^{-1}}{2},$$

which ends the proof. □

Proof of Proposition 4.3.1. Items (a) and (b) directly follow by

$$1 - \rho(A) = \mu_A(T_A = \tau(A)) = \mu_A\left(\sigma^{-n}\left(A^{(\tau(A))}\right)\right) = \prod_{j=n-\tau(A)}^{n-1}\mu(x_j) = \prod_{j=0}^{\tau(A)-1}\mu(x_j),$$

where the last equality follows from definition of $\tau(A)$.

(c) If $\rho(A)$ is periodic with period p, then the above formula implies $\rho(A) \longrightarrow 1 - \prod_{j=0}^{p-1}\mu(x_j) < 1$, which means that $\rho(A)$ converges to one only in aperiodic points. On the other hand, if $x \in \mathscr{X}$ is aperiodic, we have

$$\rho(A) = 1 - \prod_{j=0}^{\tau(A)-1}\mu(x_j) = 1 - \mu\left(A^{(\tau(A))}\right) \longrightarrow 1.$$

(d) Applying again Item (a),

$$\rho(A_n(x)) \geq 1 - \mu(x_0) \geq 1 - \lambda.$$

Notice that $\lambda = \max\{\mu(a), a \in \mathscr{A}\} < 1$ is well defined even when the alphabet is infinite, since $\sum_{a\in\mathscr{A}}\mu(a) = 1$. Thus, there exists $b \in \mathscr{A}$ such that $\mu(b) = \lambda$. Then, for $x = b^\infty$ we have $\rho(A_n(x)) = 1 - \lambda$ for all $n \geq 1$. □

Proof of Proposition 4.3.3. (a) Recall that $\kappa_A = \min\{n, \tau(A)\}$. Consider first the case $\kappa_A = \tau(A)$ and denote $A = A_n(x)$. Then

$$\begin{aligned}1-\rho(A) &= \mu_A\left(\sigma^{-n}\left(A^{(\tau(A))}\right)\right)\\ &= Q(x_{n-1}, x_{n-\tau(A)}) \prod_{j=n-\tau(A)}^{n-2} Q(x_j, x_{j+1})\\ &= Q(x_{\tau(A)-1}, x_0) \prod_{j=0}^{\tau(A)-2} Q(x_j, x_{j+1}),\end{aligned}$$

where the last equality follows from the fact that the last $\tau(A)$ letters of the word A has the same symbols as the first $\tau(A)$ letters (not necessarily in the same order).

Suppose now that $\kappa_A = n$. Then

$$1-\rho(A) = \mu_A\left(\sigma^{-\tau(A)}(A)\right) = Q^{\tau(A)-n+1}(x_{n-1}, x_0) \prod_{j=0}^{n-2} Q(x_j, x_{j+1}).$$

Item (b) directly follows from (a). To prove Item (c), note that if $x = x_0^{p-1} \cdots$ is periodic, then $\mu\left(x_0^{p-1} x_0\right) > 0$, since we are considering only the strings $A_n(x)$ with positive measure. Thus, in this case we have $Q(x_{p-1}, x_0) \prod_{j=0}^{p-2} Q(x_j, x_{j+1}) > 0$ and $\rho(A_n(x)) \longrightarrow \rho(x) < 1$.

On the other hand, if x is aperiodic, then $\kappa_A \xrightarrow{n} \infty$ and Item (a) gives us:

$$1-\rho(A_n(x)) = Q^{\tau(A)-\kappa_A+1}(x_{\kappa_A-1}, x_0) \prod_{j=0}^{\kappa_A-2} Q(x_j, x_{j+1}) \leq \frac{\mu\left(x_0^{\kappa_A-1}\right)}{\mu(x_0)} \xrightarrow{n} 0.$$

(d) When $\tau(A) = \kappa_A = 1$, by Item (a) we have that $\rho(A_n(x)) \geq 1 - Q(x_0, x_0) \geq 1 - \lambda$. If $\tau(A) > 1$, then $\rho(A_n(x)) \geq 1 - Q(x_0, x_1) \geq 1 - \lambda$.

On the other hand, suppose that $\lambda = \sup_{a \in \mathscr{A}} Q(a, a)$ a consider the set $\mathscr{P} = \{x \in \mathscr{X}; \tau(x) = 1\}$ of the points with period one. Then,

$$1-\lambda \leq \inf_{n\geq 1, x \in \mathscr{X}} \rho(A_n(x)) \leq \inf_{n \geq 1, x \in \mathscr{P}} \rho(A_n(x)) = \inf_{a \in \mathscr{A}} (1 - Q(a, a)) = 1 - \lambda.$$

(e) It is a direct consequence of Items (b) and (c). □

Proof of Proposition 4.3.8. (a) Consider the string $A = 0^k 1 a_{k+1}^{n-1} \in \mathscr{C}_n$. Then, by Theorem 4.3.7, we have

$$\rho(A) \geq 1 - \mu_{10^k}(a_{\tau(A)-1}) \sigma_k.$$

If $k \geq 1$, then $\mu_{10^k}(a_{\tau(A)-1}) \sigma_k \leq \sigma_k \leq q_0$. Otherwise, for $k = 0$, we get $\mu_1(a_{\tau(A)-1}) \leq \max\{q_0, 1 - q_0\}$, since $a_{\tau(A)-1}$ must be equal to 0 or 1. In both cases, one obtain

$$\rho(A) \geq 1 - \max\{q_0, 1 - q_0\}.$$

Furthermore, when $1-q_0 \geq q_0$, we have the equality, since the point $x = 1^\infty$ satisfies $\rho(A_n(x)) = q_0$ for all $n \geq 1$. (b) According to Theorem 2.2.3, if $\mathbf{Y}$ is ϕ-mixing, then there exist positive constants C and c such that for all $n, j \geq 0$:

$$\frac{\sigma_{n+j}}{\sigma_n} \leq Ce^{-cj}.$$

Thus, for all $n \geq 1$:

$$\rho(0^n) = \mu_{0^n}(1) = \frac{\sigma_n}{\Sigma(n)} = \frac{1}{\sum_{j=0}^{\infty} \frac{\sigma_{n+j}}{\sigma_n}} \geq \frac{1-e^{-c}}{C}.$$

On the other hand, suppose that $\mathbf{Y}$ is not ϕ-mixing. Given $\varepsilon > 0$, we can apply Theorem 2.2.3-(e) to conclude that there exist infinitely many $n_0, j_0 \in \mathbb{N}$ such that

$$\begin{cases} q_{n_0}, q_{n_0+1}, \cdots, q_{n_0+j_0-1} > 1-\varepsilon \quad \text{and} \\ (1-\varepsilon)^{j_0} < 1/2. \end{cases}$$

Therefore,

$$\begin{aligned}
\rho(0^{n_0}) &= \frac{1}{1+q_{n_0}+q_{n_0}q_{n_0+1}+\cdots} \\
&\leq \frac{1}{1+q_{n_0}+q_{n_0}q_{n_0+1}+\cdots+q_{n_0}\cdots q_{n_0+j_0-1}} \\
&\leq \frac{1}{1+(1-\varepsilon)+\cdots+(1-\varepsilon)^{j_0}} \\
&= \frac{\varepsilon}{1-(1-\varepsilon)^{j_0+1}} \\
&\leq 2\varepsilon.
\end{aligned}$$

Since there exist infinitely many $n_0 \in \mathbb{N}$ satisfying the above inequalities, we conclude that

$$0 = \liminf_{n\geq 1} \rho(0^n) = \inf_{n\geq 1} \rho(0^n),$$

which ends the proof. $\square$

Proof of Proposition 4.4.1. For each $x \in \mathscr{X}$, denote

$$s(x) = \limsup_{n\to\infty} \frac{1}{n} \ln \tau(A_n(x)).$$

The hypothesis ensures that for almost every x, we can take $\varepsilon > 0$ satisfying $s(x)+\varepsilon < h_\mu$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\frac{1}{n} \ln \tau(A_n(x)) < s(x) + \varepsilon.$$

Thus,

$$\begin{aligned}
1-\rho(A_n) &= \mu_{A_n}(T_{A_n} \leq \tau(A_n)) \\
&\leq \mu_{A_n}\left(T_{A_n} \leq e^{n(s(x)+\varepsilon)}\right)
\end{aligned}$$

Applying Lemma 4 in (ABADI; SAUSSOL, 2016), we get

$$\lim_{n\to\infty}(1-\rho(A_n))=0, \quad \text{a.s.}$$

□

Proof of Theorem 4.4.3. (a) As we said in Remark 4.4.5, the cases ψ and ϕ were already proved in Lemma 3.4.4. Let us prove the case α. Note that, for $A\in\mathscr{C}_n$, we have

$$\mu\left(A,\sigma^{-j-n}(A)\right)\geq\mu(A)^2-\alpha(j).$$

Consider $j=\dfrac{\mu(A)^{-1}}{\ln(\mu(A)^{-1})}$. The hypothesis about α decay gives us

$$\alpha(j)\leq\left(\mu(A)\ln\left(\mu(A)^{-1}\right)\right)^r,$$

which implies

$$\mu\left(A,\sigma^{-j-n}(A)\right)\geq\mu(A)^2\left(1-\mu(A)^{r-2}\left(\ln\left(\mu(A)^{-1}\right)\right)^r\right).$$

Notice that the function $f(x)=x^{r-2}(\ln(1/x))^r$ converges to zero when x goes to zero, which means that for n large enough we have $\mu\left(A,\sigma^{-j-n}(A)\right)>0$. By definition, $\tau(A)$ is the smallest positive integer k such that $\mu\left(A,\sigma^{-k}(A)\right)>0$, then we must have $\tau(A)\leq j+n$.

(b) For the ψ-mixing case, we have:

$$0\leq\limsup_{n\to\infty}\frac{1}{n}\ln(\tau(A_n(x))\leq\limsup_{n\to\infty}\frac{1}{n}\ln(Cn)=0,$$

where the second inequality follows from Item (a). Thus, the statement holds with $\delta=h_\mu$ and for all points $x\in\mathscr{X}$.

Suppose now that (b2) holds and take $\delta,\varepsilon\in(0,h_\mu)$ satisfying

$$s>\frac{h_\mu+\varepsilon}{h_\mu-\delta}.$$

If we take $j=e^{n(h_\mu-\delta)}$, then the Shannon–McMillan–Breiman Theorem ensures that, for almost every $x\in\mathscr{X}$ and for all n large enough:

$$\mu_{A_n(x)}\left(\sigma^{-j-n}(A_n(x))\right)\geq\mu(A_n(x))-\phi(j)\geq e^{-n(h_\mu+\varepsilon)}-e^{-sn(h_\mu-\delta)}>0,$$

since $s(h_\mu-\delta)>h_\mu+\varepsilon$. Thus, we conclude that $\tau(A_n(x))\leq e^{n(h_\mu-\delta)}$ for almost every x and large n, which implies the statement.

Finally, when we have the condition (b3), we take $\delta,\varepsilon\in(0,h_\mu)$ satisfying

$$r>\frac{2h_\mu+\varepsilon}{h_\mu-\delta},$$

Then, we follow the same as the previous case by noting that

$$\mu\left(A_n(x), \sigma^{-j-n}(A_n(x)\right) \geq \mu(A_n(x))^2 - \alpha(j),$$

which is positive for almost every x, large n and $j = e^{n(h_\mu - \delta)}$.

$\square$

CHAPTER 5

WAITING TIME SPECTRUM

The waiting time, of order n, between $\mathbf{X}$ and $\mathbf{Y}$ is the first time $\mathbf{Y}$ displays the first n coordinates of $\mathbf{X}$. In this chapter we use the hitting time approximation theorem to obtain a *strong approximations* for the moments of the waiting time, when both processes are independent and follows the same stationary law μ. In order to apply Theorem 3.3.4, we still work in the context of ϕ-mixing processes, but we also obtain some related results for non ϕ-mixing renewal processes. Applying the strong approximation theorem, we discuss the relationship between the waiting time spectrum and the Rényi entropy, for which we give some new conditions of existence. Then, we obtain the waiting time spectrum as a function of the Rényi entropy for some classes of processes not covered in previous works. The proofs are given in the last section.

5.1 Strong approximation for waiting times

The waiting time is a sequence of random variables defined over the space $\mathscr{X} \times \mathscr{X}$. Loosely speaking, we independently choose two points $x, y \in \mathscr{X}$ and we look at the first time that the point y reproduces the first n coordinates of x.

Definition 5.1.1. For each $n \geq 1$, we define the waiting time of order n as a function $W_n : \mathscr{X} \times \mathscr{X} \longrightarrow \mathbb{N} \cup \{\infty\}$ given by:

$$W_n(x,y) := \inf\left\{k \geq 1; y_k^{k+n-1} = x_0^{n-1}\right\}.$$

From now on, whenever we consider a stationary measure μ on $\mathscr{X}$, we will denote by $\pi = \mu \times \mu$ the respective product measure on the space $\mathscr{X} \times \mathscr{X}$, and $\mathbb{E}_\pi$ will denote the expectation with respect to π.

Wyner and Ziv (1989) proved that, if μ is β-mixing, then

$$\lim_{n\to\infty} \frac{1}{n} \ln W_n(x,y) = h_\mu, \quad \text{in probability,}$$

where h_μ is the entropy of μ. Shields (1996) proved that this convergence holds almost surely. In view of the Shannon–McMillan–Breiman Theorem, we conclude that $W_n(x,y) \sim \mu\left(x_0^{n-1}\right)^{-1}$ almost surely[1], that is, the time required for a point y to output the string x_0^{n-1} is typically of the order of the reciprocal of its measure. Moreover, our Corollary 4.2.4 states that for any cylinder $A \in \mathscr{C}_n$ we have $\mathbb{E}\left(T_A^\beta\right) \sim \mu(A)^{-\beta}, \beta > 0$. These facts indicate that we may have, at least for q's in some subset of $\mathbb{R}$,

$$\mathbb{E}_\pi\left(W_n^q\right) \sim \mathbb{E}_\mu\left(\mu\left(X_0^{n-1}\right)^{-q}\right) = \sum_{x_0^{n-1}} \mu\left(x_0^{n-1}\right)^{1-q}.$$

This is what Chazottes and Ugalde (2005) called *strong approximation*. They proved that this holds for $q > -1$ when μ is a Bowen-Gibbs measure on finite alphabet and with potential of summable variation (which is a particular case of ψ-mixing measure). However, they showed that, for $q < -1$, the behaviour changes and $\mathbb{E}_\pi\left(W_n^q\right) \sim \mathbb{E}_\mu\left(\mu\left(X_0^{n-1}\right)\right)$. As a consequence, the authors showed that $\frac{1}{n}\ln W_n(x,y)$ and $-\frac{1}{n}\ln\left(\mu(A_n(x)\right)$ have different large deviations in general.

In the next theorem, we provide an extension of their result. For $q < -1$, our result holds for any stationary process, while for $q > -1$ we prove the approximation for any ψ-mixing or summable ϕ-mixing processes. Furthermore, our result also holds for countable alphabet. Moreover, we correct their statement, and we show that for $q = -1$, the strong approximation has an extra factor; as a consequence, $\mathbb{E}_\pi\left(W_n^{-1}\right) \nsim \mathbb{E}_\mu\left(\mu\left(X_0^{n-1}\right)\right)$. Nevertheless, in the next section we will show that this "discontinuity" does not affect the waiting time spectrum.

Theorem 5.1.2. Let μ a stationary measure. Then:

(a) $\mathbb{E}_\pi\left(W_n^q\right) \sim \sum_{A \in \mathscr{C}_n} \mu(A)^2$ for all $q < -1$.

(b) If μ is ψ-mixing or summable ϕ-mixing, then

- $\mathbb{E}_\pi\left(W_n^q\right) \sim \sum_{A \in \mathscr{C}_n} \mu(A)^{1-q}$ for all $q > -1$;
- $\mathbb{E}_\pi\left(W_n^{-1}\right) \sim \sum_{A \in \mathscr{C}_n} \mu(A)^2 \ln\left(\mu(A)^{-1}\right)$.

Remark 5.1.3. In Chapter 2 we presented some classes of ϕ-mixing processes which are not covered by Theorem 3.1 in (CHAZOTTES; UGALDE, 2005) when $q > -1$. For instance, Example 2.1.2 gives a ψ-mixing and *infinite* Markov chain, while Examples 2.2.11 and 2.1.3 (when satisfies Theorem 2.2.3) present ϕ-mixing processes that are not ψ-mixing. For the case $q < -1$, our result obviously holds for a much more general class of processes.

[1] Recall that $a_n \sim b_n$ means that $\max\{a_n/b_n, b_n/a_n\}$ is bounded.

Remark 5.1.4. When we are able to prove that the Rényi entropy (see Definition 5.2.1) exists, Theorem 5.1.2 allows to get the so-called cumulant generating function of $(W_n)_n$, also called "waiting time spectrum" in dynamical systems. This in turn can be used, together with classical results in (PLACHKY; STEINEBACH, 1975), to get the large deviation rate function of $\frac{\ln W_n}{n}$ as the Legendre-Fenchel transform. This was done by Chazottes and Ugalde (2005) for waiting time, but also more recently by Abadi, Chazottes and Gallo (2019) for the return time, in both works for Bowen-Gibbs measures. The next section is dedicated to finding conditions under which the waiting time spectrum exists.

Remark 5.1.5. The asymptotic behaviour of the potential well also has an influence on the strong approximation of the waiting time. As we can see in the proof of Theorem 5.1.2, the upper bound of the strong approximation for $q > 0$ comes from the following inequality:

$$I(A) = \int_0^\infty t^{q-1}\mu\left(T_A > \frac{t}{\mu(A)}\right) dt \leq C_q,$$

for all large $n \in \mathbb{N}$, $A \in \mathscr{C}_n$ and some positive constant C_q that depends on q. In Section 5.4, we prove that we can have $\limsup_n I(A_n(x)) = \infty$ when $\liminf_n(\rho(A_n(x)) = 0$ for some periodic point $x \in \mathscr{X}$. For instance, in renewal processes we have this situation whenever the process is non ϕ-mixing (see Proposition 4.3.8), while the ϕ-mixing case, which is always summable, is covered by Theorem 5.1.2.

In the same way, for $q \in (-1,0)$, the lower bound comes from:

$$J(A) = \int_{\mu(A)}^\infty t^{-1-|q|}\mu\left(T_A < \frac{t}{\mu(A)}\right) dt \geq C'_q > 0.$$

In this case, we also provide a proof that $\liminf_n \rho(A_n(x)) = 0$ implies $\liminf_n J(A_n(x)) = 0$ for some periodic point $x \in \mathscr{X}$.

When dealing with renewal processes, we cannot apply Theorem 5.1.2 in general by the assumption of mixing properties. In the next proposition, we provide a weak version of this theorem for $q \in (-1,0]$, which will allow us to obtain, in the next section, the spectrum of the waiting time for $q \leq 0$ for a class of non ϕ-mixing processes.

Proposition 5.1.6. Let $\mathbf{Y}$ be a renewal process defined by the sequence $(q_i)_{i\geq 0}$ such that $q_i \xrightarrow{i} 1$. Then, for all $q \in (-1,0]$ we have

$$C\sigma_n^2 \sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} \leq \mathbb{E}_\pi\left(W_n^q\right) \leq C_q \sum_{A\in\mathscr{C}_n} \mu(A)^{1-q},$$

where $C > 0$ and C_q is a positive constant depending on q. Furthermore,

$$\frac{1}{n}\ln\left(\sigma_n^2\right) \xrightarrow{n} 0.$$

Remark 5.1.7. Notice that the class of renewal processes provides examples where $\mathbb{E}_\pi\left(W_n^q\right) = \infty$ for all $q > s$ for some $s > 0$. For instance, consider a renewal process defined by $\Sigma(n) = n^{-s}$ with $s > 0$. In this case we have $q_i \longrightarrow 1$ and hence the process is not ϕ-mixing. If we take the string $A = 1^n$, then

$$\mu(T_A \geq t) \geq \mu\left(0^t\right) = \frac{\mu(1)}{t^s}.$$

Applying (5.3), we get for all $q \geq s$:

$$\begin{aligned}
\mathbb{E}_\pi\left(W_n^q\right) &\geq q\mu(A)^{1-q}\int_0^\infty t^{q-1}\mu\left(T_A \geq \frac{t}{\mu(A)}\right)dt \\
&\geq q\mu(1)\mu(A)^{1-q+s}\int_0^\infty t^{q-s-1}dt \\
&= \infty.
\end{aligned}$$

Thus, we cannot obtain, for $q > 0$, a result similar to the Proposition 5.1.6. This suggests that we depend on ϕ-mixing condition to obtain the waiting time spectrum for $q > 0$.

5.2 Rényi entropy and waiting time spectrum

The strong approximation theorem 5.1.2 provides a relationship between the waiting time spectrum and a function known as Rényi entropy. Under the conditions of the theorem, the existence of the waiting time spectrum is equivalent to the existence of the Rényi entropy. In this section we discuss some conditions and present new results on this issue. Let us define these functions.

Definition 5.2.1. Consider a stationary process $\mathbf{X}$ with law μ. For each non-zero $q \in \mathbb{R}$, we define for $n \geq 1$:

$$H_n(q) = \frac{1}{nq}\ln\sum_{A\in\mathscr{C}_n}\mu(A)^{1-q} = \frac{1}{nq}\ln\mathbb{E}_\mu\left(\mu\left(X_0^{n-1}\right)^{-q}\right).$$

The generalized Rényi entropy of the measure μ is defined by

$$H_R(q) = \lim_{n\to\infty} H_n(q)$$

as long as the limit exists. We also define

$$H_R^+(q) = \limsup_{n\to\infty} H_n(q) \quad \text{and} \quad H_R^-(q) = \liminf_{n\to\infty} H_n(q).$$

At last, we define $H_R(0) = h_\mu$, which always exists. This definition is motivated by the following fact:

$$\lim_{q\to 0} H_n(q) = -\frac{1}{n}\sum_{A\in\mathscr{C}_n}\mu(A)\ln(\mu(A)).$$

Notice that for the sequence $H_n(q)$ to be well-defined, we must have

$$\sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} < \infty, \tag{5.1}$$

for all $n \geq 1$, which is always the case when we have a finite alphabet $\mathscr{A}$. On the other hand, if $\mathscr{A}$ is infinite, we have for any $q \leq 0$:

$$\sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} \leq \sum_{A\in\mathscr{C}_n} \mu(A) = 1.$$

However, when $q \geq 1$ we get

$$\sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} \geq \sum_{A\in\mathscr{C}_n} 1 = \infty,$$

since $\mathscr{C}_n$ is infinite (for each state $a_0 \in \mathscr{A}$ we can get a word a_0^{n-1} with positive measure). For the case $q \in (0,1)$, we will present examples in the next section showing that the sum in (5.1) may be finite or not, depending on the characteristics of the measure μ.

Let us define now the waiting time spectrum.

Definition 5.2.2. For $n \geq 1$ and $q \in \mathbb{R}$, we define

$$\mathscr{W}_n(q) := \frac{1}{n} \ln \mathbb{E}_\pi\left(W_n^q\right).$$

The waiting time spectrum is the function $\mathscr{W}(q) : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$\mathscr{W}(q) := \lim_{n\to\infty} \mathscr{W}_n(q)$$

provided the limit exists.

We state in the next corollary a description of the relationship between the functions $H_R(q)$ and $\mathscr{W}(q)$, which is a consequence of Theorem 5.1.2 and Proposition 5.1.6.

Corollary 5.2.3. Consider a stationary measure μ. Under the conditions of Theorem 5.1.2 and Proposition 5.1.6, the following statements hold:

(a) For each $q > -1$, $H_R(q)$ exists if, and only if, $\mathscr{W}(q)$ exists. In this case we have $\mathscr{W}(q) = qH_R(q)$.

(b) If $H_R(-1)$ exists, then $\mathscr{W}(q)$ exists for all $q < -1$. Furthermore, if, for some $\delta > 0$, $H_R(q)$ exists for all $q \in (-1-\delta, -1+\delta)$, then $\mathscr{W}(-1)$ also exists. In both cases we have $\mathscr{W}(q) = -H_R(-1)$ for all $q \leq -1$.

(c) If $H_R(q)$ exists on $(-1-\delta, \gamma)$ for some $\gamma > -1$, then $\mathscr{W}(q)$ is a continuous function on $(-\infty, \gamma)$ and a convex function on $(-1, \gamma)$.

The above corollary raises the natural question of whether Rényi entropy exists. In fact, several recent papers have proven its existence under different conditions: (GRASSBERGER; PROCACCIA, 1984), (PITTEL, 1985), (SZPANKOWSKI, 1996), (LUCZAK; SZPANKOWSKI, 1997) and (HAYDN; VAIENTI, 2010).

Let us highlight some of them that are related to our case. In (CHAZOTTES; UGALDE, 2005), we can see that $H_R(q)$ exists for all $q \in \mathbb{R}$ when μ is a Bowen-Gibbs measure with potential of summable variation (which is a particular case of ψ-mixing). The authors used this existence and the strong approximation to obtain large deviation results for the waiting time.

Ko (2012) proved that $H_R(q)$ exists for all $q < 0$ for a class of processes called simple mixing, which is equivalent to ψ^*-mixing (see Definition 2.3.1) in the context of ergodic mixing process. As we saw in Chapter 2, this class is more general than ψ-mixing and less general than ϕ-mixing.

In order to obtain large deviations for the shortest possible return function, Abadi and Cardeno (2015) showed that, if a process is ψ_g-regular (see Definition 2.3.3), then $H_R(q)$ exists for all $q \leq 0$. Recall from Chapter 2 that ψ_g-regular includes every ψ-mixing processes, as well as some β-mixing processes that are not ψ nor ϕ-mixing.

The next theorem provides a new condition for the existence of the Rényi entropy and brings at least three new contributions: 1) Unlike previous results, the theorem applies to processes with infinite alphabet (when (5.1) holds); 2) Our conditions include processes that are not covered by Bowen-Gibbs measure nor ψ^*-mixing; 3) We prove the existence *for all* $q \in \mathbb{R}$.

Recall from 2.3.1 the definition of the functions $\psi'(n)$ and $\psi^*(n)$.

Theorem 5.2.4. Consider a stationary measure μ over $\left(\mathscr{A}^{\mathbb{N}}, \mathscr{F}\right)$, where $\mathscr{A}$ is a finite or countable alphabet. If μ satisfies either $\psi'(1) > 0$ or $\psi^*(1) < \infty$, then $H_R(q)$ exists for all $q \in \mathbb{R}$ satisfying (5.1).

Remark 5.2.5. As a consequence of Theorems 5.1.2, 5.2.4 and Corollary 5.2.3, we have new conditions that ensure the existence and identifies the waiting time spectrum for all $q \in \mathbb{R}$. For instance, we will present at the end of this section some cases of ϕ-mixing infinite Markov chains and renewal processes that satisfy these conditions. Notice that the conditions $\psi'(1) > 0$ and $\psi^*(1) < \infty$ imply ϕ-mixing, but not every ϕ-mixing process is covered by them.

We present now three classes of ϕ-mixing renewal processes for which the Rényi entropy exists for all $q \in \mathbb{R}$. We also show the finiteness of Rényi entropy for $q \leq 1$ and provide a condition for the finiteness when $q > 1$.

Theorem 5.2.6. Consider a stationary renewal process $\mathbf{Y}$ defined by the sequence $(q_i)_{i\geq 0}$. The following statements hold.

(a) If any of the following three conditions holds, then $\mathbf{Y}$ is ϕ-mixing and $H_R(q)$ exists for all $q \in \mathbb{R}$.

(a1) $\sup_{r,s\in\mathbb{N}} \frac{\sigma_{r+s}}{\sigma_r\sigma_s} < \infty$

(a2) $\inf_{r,s\in\mathbb{N}} \frac{\sigma_{r+s}(1-q_{r+s})}{\sigma_r\sigma_s} > 0$

(a3) $q_i \xrightarrow{i} L \in (0,1)$

(b) For all $q \leq 1$, $H_R^+(q)$ and $H_R^-(q)$ are finite. For $q > 1$, the finiteness holds if, and only if, there exist $c, C > 0$ and $n_0 \in \mathbb{N}$ such that $\sigma_n(1-q_n) \geq Ce^{-cn}$ for all $n \geq n_0$.

5.3 Examples

The following examples show that, compared to previous results, our contributions expand the classes of processes for which one can obtain the waiting time spectrum.

Renewal processes

The class of renewal processes can provide a wide variety of cases to illustrate our results. For instance, if we take $\sigma_n = n^{-s}$ for some $s > 1$, we get an example of a *non ϕ-mixing* process for which one can obtain $\mathscr{W}(q)$ for all $q \leq 0$. Indeed, Abadi and Cardeno (2015) showed in their Example 5.2 that this process is ψ_0-regular, which means that $H_R(q)$ exists for all $q \leq 0$. Thus, we can apply Theorem 5.1.2 and Proposition 5.1.6 to show the existence of the spectrum as a function of Rényi entropy.

On the other hand, the renewal process defined by $q_i = p(i+1)/(i+2)$, where $p \in (0,1)$, obviously satisfies the condition (a3) of Theorem 5.2.6. Moreover, as we saw in Example 2.2.11, this process is exponentially ϕ-mixing. Then, we can apply Theorems 5.1.2 and 5.2.6 to conclude that $\mathscr{W}(q)$ exists in function of $H_R(q)$ and is finite for all $q \in \mathbb{R}$. The finiteness comes from the fact that $\sigma_n(1-q_n) \geq (p/2)^n(1-p)$, then we can apply Item (b) of Theorem 5.2.6. Notice that this case is not ψ-mixing nor ψ^*-mixing (or simple mixing) since for all $n \geq 1$:

$$\psi^*(n) \geq \sup_{r,s\in\mathbb{N}} \frac{\mu(10^{r+n+s}1)}{\mu(10^r)\mu(0^s1)} \geq \frac{p^n(1-p)}{\mu(1)} \sup_{r,s\in\mathbb{N}} \frac{(r+1)(s+1)}{(r+n+s+1)} = \infty.$$

One can also obtain a ψ-mixing renewal process such that $\mathscr{W}(q)$ exists and is finite *for all $q \in \mathbb{R}$*, but is not Bowen-Gibbs measure. For this, it is enough to take (q_i) defined by $q_i = p$ for even i and $q_i = q$ for odd i, where $0 < p < q < 1$. In this case, direct calculations show that:

$$\sup_{r,s\in\mathbb{N}} \frac{\sigma_{r+s}}{\sigma_r\sigma_s} = \frac{q}{p}.$$

Thus, this is a ψ-mixing process (by Theorem 2.2.6) and we can apply condition (a1). However, we can show that the potential function is not continuous in this case and hence does not have summable variations (see (CHAZOTTES; UGALDE, 2005) for details).

Markov chains

The class of the finite Markov chains is included in the case of the Bowen-Gibbs measure, which was solved by Chazottes and Ugalde (2005), but their results hold for finite alphabet only. On the other hand, our results also provide the waiting time spectrum for some infinite Markov chains. Recall from the beginning of this section that $H_n(q)$ is well-defined for $q \leq 0$ and is not defined for $q \geq 1$. The case $q \in (0,1)$ depends on the characteristics of the transition matrix of the chain.

For instance, if we consider the house of cards Markov chain, we know from Example 2.1.3 that it is not ψ-mixing (neither ψ^*-mixing), but it can be exponentially ϕ-mixing under the conditions of Theorem 2.2.3. In this case, the strong approximation holds for all $q \in \mathbb{R}$. Furthermore, if we assume $S = \sup_i\{q_i\} < 1$, then we can show that $\mathscr{W}(q)$ exists and is finite for all $q < 1$. At the end of Section 5.4, we present a proof for this.

On the other hand, consider the infinite Markov chain defined in the Example 2.1.2. As we showed, this chain is ψ-mixing, and hence the strong approximation theorem holds for all $q \in \mathbb{R}$. Furthermore, we can show that for all $q < 1$, we have

$$p^{(1-q)n} = \mu\left(0^n\right)^{1-q} \leq \sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} = a_n(q) < \infty,$$

Direct computations also show that this chain satisfies $\psi'(1) > 0$ *and* $\psi^*(1) < \infty$. That is, this is a ψ-mixing process over an infinite alphabet such that $\mathscr{W}(q)$ exists and is finite for all $q < 1$.

5.4 Proofs

Proof of Theorem 5.1.2. (a) First, we observe that for all $q<0$ we have

$$\begin{aligned}
\mathbb{E}_\pi\left(W_n^q\right) &= \sum_{A\in\mathscr{C}_n} \mu\left(A\right)\mathbb{E}_\mu\left(T_A^{-|q|}\right) \\
&= \sum_{A\in\mathscr{C}_n} \mu(A)\int_0^1 \mu\left(T_A^{-|q|}>s\right)ds \\
&= |q|\sum_{A\in\mathscr{C}_n} \mu(A)^{1+|q|}\int_{\mu(A)}^{\infty} t^{-1-|q|}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt, \qquad (5.2)
\end{aligned}$$

where the last equality follows from the substitution $t=\mu(A)s^{-1/|q|}$.

Since μ is stationary, we have $\mu(T_A<t)\leq t\mu(A)$ and

$$\int_{\mu(A)}^{\infty} t^{-1-|q|}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt \leq \int_{\mu(A)}^{\infty} t^{-|q|}dt = \frac{\mu(A)^{1-|q|}}{|q|-1}.$$

On the other hand, note that for all $t\geq 2\mu(A)$ we have $\mu(T_A<t/\mu(A))\geq\mu(T_A=1)=\mu(A)$. Then

$$\int_{\mu(A)}^{\infty} t^{-1-|q|}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt \geq \mu(A)\int_{2\mu(A)}^{\infty} t^{-1-|q|}dt = \frac{\mu(A)^{1-|q|}}{|q|2^{|q|}}.$$

Therefore, applying (5.2), one obtain, for $q<-1$

$$\frac{1}{2^{|q|}}\sum_{A\in\mathscr{C}_n}\mu(A)^2 \leq \mathbb{E}_\pi\left(W_n^q\right) \leq \frac{|q|}{|q|-1}\sum_{A\in\mathscr{C}_n}\mu(A)^2.$$

(b) We divide the proof into three cases, starting with $q\in(-1,0]$. The statement is trivial for $q=0$. For $q\in(-1,0)$, we have to proof that the integral of (5.2) is between two positive constants. We apply again the inequality $\mu(T_A<t)\leq t\mu(A)$ to get the upper bound for any stationary measure μ:

$$\begin{aligned}
\int_{\mu(A)}^{\infty} t^{-1-|q|}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt &= \left(\int_{\mu(A)}^{1}+\int_{1}^{\infty}\right)t^{-1-|q|}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt \\
&\leq \int_{\mu(A)}^{1} t^{-|q|}dt + \int_1^{\infty} t^{-1-|q|}dt \\
&= \frac{1-\mu(A)^{1-|q|}}{1-|q|}+\frac{1}{|q|} \\
&\leq \frac{1}{(1-|q|)|q|}.
\end{aligned}$$

For the lower bound, we suppose that μ is ψ-mixing or summable ϕ-mixing. Applying Lemma 3 in (ABADI, 2001), for all $t\leq(2\mu(A))^{-1}$, we get a positive constant K such that $\mu(T_A\leq t)\geq Kt\mu(A)$. Thus,

$$\int_{\mu(A)}^{\infty} t^{-1-|q|}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt \geq K\int_{\mu(A)}^{1/2} t^{-|q|}dt \geq \frac{K}{2^{2-|q|}(1-|q|)}.$$

This finishes this case.

Consider now $q>0$ and note that

$$\begin{aligned}\mathbb{E}_\pi\left(W_n^q\right) &= \sum_{A\in\mathscr{C}_n}\mu(A)\,\mathbb{E}_\mu\left(T_A^q\right) \\ &= \sum_{A\in\mathscr{C}_n}\mu(A)\int_0^\infty \mu\left(T_A^q>s\right)ds \\ &= q\sum_{A\in\mathscr{C}_n}\mu(A)^{1-q}\int_0^\infty t^{q-1}\mu\left(T_A>\frac{t}{\mu(A)}\right)dt,\end{aligned} \tag{5.3}$$

where the last equality follows from substitution $t=s^{1/q}\mu(A)$.

We apply again the inequality $\mu(T_A>t)\geq 1-t\mu(A)$ to obtain for any stationary measure μ:

$$\int_0^\infty t^{q-1}\mu\left(T_A>\frac{t}{\mu(A)}\right)dt \geq \int_0^1 t^{q-1}(1-t)dt = \frac{1}{q(q+1)}.$$

On the other hand, applying Corollary 4.2.4, we get

$$\mathbb{E}_\mu\left(T_A^q\right)\leq\left(\frac{\Gamma(q+1)}{\rho(A)^q}+C_1\varepsilon(A)\right)\mu(A)^{-q}.$$

Notice that $\varepsilon(A)\xrightarrow{n}0$ and $\rho(A)\geq\rho>0$ by Theorem 4.2.1. Thus, the upper bound follows from (5.3).

At last, let us consider $q=-1$. In this case, (5.2) becomes

$$\mathbb{E}_\pi\left(W_n^{-1}\right)=\sum_{A\in\mathscr{C}_n}\mu(A)^2\int_{\mu(A)}^\infty t^{-2}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt.$$

For any stationary measure μ, the upper bound is obtained as follows

$$\begin{aligned}\int_{\mu(A)}^\infty t^{-2}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt &= \left(\int_{\mu(A)}^1+\int_1^\infty\right)t^{-2}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt \\ &\leq \int_{\mu(A)}^1 t^{-1}dt+\int_1^\infty t^{-2}dt \\ &\leq 2\ln\left(\mu(A)^{-1}\right).\end{aligned}$$

To finish the proof, when μ is ψ-mixing or summable ϕ-mixing, we apply again Lemma 3 (ABADI, 2001):

$$\int_{\mu(A)}^\infty t^{-2}\mu\left(T_A<\frac{t}{\mu(A)}\right)dt\geq K\int_{\mu(A)}^{1/2}t^{-1}dt\geq\frac{K}{2}\ln\left(\mu(A)^{-1}\right).$$

□

Remark 5.4.1. In the proof of Theorem 5.1.2, we can see that, for $q \in [-1,0]$, the upper bound of the strong approximation is obtained by simply assuming the stationarity of μ; the mixing conditions were used only for the lower bound. On the other hand, for $q > 0$ only the upper demands mixing properties, since stationarity is sufficient for the lower bound.

In the next Lemma, we prove the statements of Remark 5.1.5.

Lemma 5.4.2. Consider a stationary measure μ on $(\mathscr{X}, \mathscr{F})$. If there exists a periodic point $x \in \mathscr{X}$ satisfying $\liminf_n \rho(A_n(x)) = 0$, then:

(a) For all $q > 0$, $\limsup_{n\to\infty} \int_0^\infty t^{q-1} \mu\left(T_{A_n(x)} > \frac{t}{\mu(A_n(x))}\right) dt = \infty$.

(b) For all $q \in (-1,0)$, $\liminf_{n\to\infty} \int_{\mu(A)}^\infty t^{-1-|q|} \mu\left(T_A < \frac{t}{\mu(A)}\right) dt = 0$.

Proof. (a) Denote by $A = A_n(x)$ and $\tau(x) = p$. For each $M > 0$, we will prove that there exist infinitely many $n \in \mathbb{N}$ such that:

$$I(A) = \int_0^\infty t^{q-1} \mu\left(T_{A_n(x)} > \frac{t}{\mu(A_n(x))}\right) dt > M.$$

Applying (3.2) for each fixed $t \geq p\mu(A)$ we get

$$\liminf_{n\to\infty} \mu\left(T_A \leq \frac{t}{\mu(A)}\right) \leq \liminf_{n\to\infty} (p\mu(A) + t\rho(A)) = 0. \tag{5.4}$$

Now, we fix $\varepsilon > 0$ and t_0 large enough such that $(1-\varepsilon)t_0^q/q > M$. By (5.4), we can get infinitely many $n \in \mathbb{N}$ such that:

$$\mu\left(T_A > \frac{t_0}{\mu(A)}\right) > 1 - \varepsilon.$$

Therefore,

$$I(A) \geq (1-\varepsilon) \int_0^{t_0} t^{q-1} dt > M.$$

(b) Given $\varepsilon > 0$, take $t_0 > 0$ satisfying $t_0^{-|q|}/|q| < \varepsilon/3$. Note that there exist infinitely many $n \in \mathbb{N}$ such that $\rho(A)t_0^{1-|q|}/(1-|q|) < \varepsilon/3$ and $p\mu(A)^{1-|q|}/|q| < \varepsilon/3$. Then, applying (3.2) again, we get:

$$\begin{aligned}
\int_{\mu(A)}^\infty t^{-1-|q|} \mu\left(T_A < \frac{t}{\mu(A)}\right) dt &\leq \int_{\mu(A)}^{t_0} t^{-|q|-1} (p\mu(A) + t\rho(A)) dt + \int_{t_0}^\infty t^{-|q|-1} dt \\
&\leq \frac{p\mu(A)^{1-|q|}}{|q|} + \rho(A)\frac{t_0^{1-|q|}}{1-|q|} + \frac{t_0^{-|q|}}{|q|} \\
&< \varepsilon.
\end{aligned}$$

$\square$

The next lemma is part of the proof of Proposition 5.1.6.

Lemma 5.4.3. Let $\mathbf{Y}$ be a renewal process with law μ. Then for all $A \in \mathscr{C}_n$, we have for some positive constant C:

$$\mu\left(T_A \leq \mu(A)^{-1}\right) \geq C\sigma_n^2$$

Proof. When $A = 0^n$, the inequality follows directly from

$$\mu\left(T_{0^n} \leq \mu\left(0^n\right)^{-1}\right) \geq \mu\left(T_{0^n} = 1\right) = \mu\left(0^n\right) \geq \mu(1)\sigma_n.$$

Now, consider $A \in \mathscr{C}_n - \{0^n\}$. In the proof of Theorem 4.2.1, we showed in (4.9) that

$$\mu\left(T_A \leq \mu(A)^{-1}\right) \geq \frac{1}{\mathbb{E}\left(N^2\right)}, \tag{5.5}$$

where

$$\mathbb{E}\left(N^2\right) = 1 + 2\sum_{j=1}^{\mu(A)^{-1}} \left(\mu(A)^{-1} - j\right)\mu\left(A \cap \sigma^{-j}(A)\right).$$

This holds for any stationary measure so we can use it here.

We obtain an upper bound for the above sum by separating it into two parts. First, consider the case $n < j \leq \mu(A)^{-1}$ and denote $A = 0^\ell 1\underline{a}10^k$, where $\underline{a}$ is the binary sequence between the first and the last 1 of the string A. Then, applying Lemma 2.4.3, we get

$$\begin{aligned}
\mu\left(A \cap \sigma^{-j}(A)\right) &\leq \mu(1)^{-1}\mu(A)^2 + \mu\left(A, Y_n^{j-1} = 0^{j-n}, \sigma^{-j}(A)\right) \\
&= \mu(1)^{-1}\mu(A)^2 + \mu\left(0^\ell 1\underline{a}1\right)\sigma_{k+j-n+\ell}(1 - q_{k+j-n+\ell})\mu\left(\underline{a}10^k|1\right) \\
&\leq \mu(1)^{-1}\mu(A)^2\left(1 + \frac{\sigma_{k+j-n+\ell}}{\sigma_k\sigma_\ell}\right) \\
&\leq \mu(1)^{-1}\mu(A)^2\left(1 + \sigma_n^{-2}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{j=n+1}^{\mu(A)^{-1}} \left(\mu(A)^{-1} - j\right)\mu\left(A \cap \sigma^{-j}(A)\right) &\leq \mu(1)^{-1}\sum_{j=n+1}^{\mu(A)^{-1}} \left(\mu(A)^{-1}\right)\mu(A)^2\left(1 + \sigma_n^{-2}\right) \\
&\leq \mu(1)^{-1}\left(1 + \sigma_n^{-2}\right).
\end{aligned}$$

On the other hand, we obviously have

$$\sum_{j=1}^{n} \left(\mu(A)^{-1} - j\right)\mu\left(A \cap \sigma^{-j}(A)\right) \leq \sum_{j=1}^{n} \left(\mu(A)^{-1}\right)\mu(A)\mu_A\left(\sigma^{-j}(A)\right) \leq n \leq \sigma_n^{-2},$$

where the last inequality follows from the fact that σ_n is summable.

Therefore, (5.5) implies

$$\mu\left(T_A \leq \mu(A)^{-1}\right) \geq \frac{1}{1 + 2\left(\sigma_n^{-2} + \mu(1)^{-1}\left(1 + \sigma_n^{-2}\right)\right)} \geq C\sigma_n^2.$$

□

Proof of Proposition 5.1.6. The upper bound of the inequality was proved in Theorem 5.1.2 for any stationary measure. In order to obtain the lower bound, we just apply (5.2) and Lemma 5.4.3 as follows:

$$\begin{aligned}
\mathbb{E}_\pi\left(W_n^q\right) &= |q| \sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} \int_{\mu(A)}^{\infty} t^{-1-|q|}\mu\left(T_A \leq \frac{t}{\mu(A)}\right) dt \\
&\geq |q| \sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} \int_{1}^{\infty} t^{-1-|q|}\mu\left(T_A \leq \mu(A)^{-1}\right) dt \\
&\geq C\sigma_n^2 \sum_{A\in\mathscr{C}_n} \mu(A)^{1-q}.
\end{aligned}$$

We finish the proof by noting that the hypothesis $q_i \longrightarrow 1$ and the Cesàro's Mean Theorem imply $(1/n)\ln\left(\sigma_n^2\right) \xrightarrow{n} 0$. □

Proof of Corollary 5.2.3. The proof if Item (a) follows directly from Theorem 5.1.2 and Proposition 5.1.6. To prove Item (b), observe that if $H_R(-1)$ exists, then Theorem 5.1.2 ensures that $\mathscr{W}(q) = -H_R(-1)$ for all $q < -1$. Let us prove the statement for $\mathscr{W}(-1)$.

Note that, since $\mathscr{W}_n(q)$ is increasing in q, then for all $q \in (-1-\delta, -1)$ and $q' \in (-1, -1+\delta)$ we have

$$\mathscr{W}_n(q) \leq \mathscr{W}_n(-1) \leq \mathscr{W}_n(q').$$

Thus, applying Item (a) and the statement for $q < -1$, we get

$$-H_R(-1) \leq \liminf_{n\to\infty} \mathscr{W}_n(-1) \leq \limsup_{n\to\infty} \mathscr{W}_n(-1) \leq q' H_R(q').$$

Now, consider the function $h : (-1-\delta, -1+\delta) \longrightarrow \mathbb{R}$ given by $h(q) = qH_R(q)$, which is well-defined by our hypothesis on $H_R(q)$. By Hölder's inequality, we obtain for $\alpha \in (0,1)$ and $q, q' \in (-1-\delta, -1+\delta)$:

$$\begin{aligned}
h(\alpha q + (1-\alpha)q') &= \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}_\mu \left(\mu\left(X_0^{n-1}\right)^{-q\alpha} \mu\left(X_0^{n-1}\right)^{-q'(1-\alpha)} \right) \\
&\leq \lim_{n\to\infty} \frac{1}{n} \ln \left(\mathbb{E}_\mu \left(\mu\left(X_0^{n-1}\right)^{-q} \right) \right)^{\alpha} \left(\mathbb{E}_\mu \left(\mu\left(X_0^{n-1}\right)^{-q'} \right) \right)^{1-\alpha} \\
&= \alpha h(q) + (1-\alpha) h(q').
\end{aligned}$$

Therefore, h is a convex function and hence a continuous function on $(-1-\delta, -1+\delta)$, which implies

$$\limsup_{n\to\infty} \mathscr{W}_n(-1) \leq \lim_{q'\to -1} h(q') = -H_R(-1).$$

Thus we conclude that $\mathscr{W}(-1) = \lim_{n\to\infty} \mathscr{W}_n(-1) = -H_R(-1)$.

(c) If $H_R(q)$ exists on $(-1-\delta, \gamma)$, then by Items (a) and (b):

$$\mathscr{W}(q) = \begin{cases} -H_R(-1), & \text{if } q \leq -1 \\ h(q), & \text{if } -1 < q < \gamma. \end{cases}$$

Thus, Item (b) shows that $\mathscr{W}(q)$ is convex (and therefore continuous) on $(-1,\gamma)$. Obviously, the continuity extends to $q \leq -1$. $\square$

The following Lemma is a direct consequence of the Fekete's Subadditive Lemma. It states that, if a real sequence satisfies $\delta_{m+n} \leq \delta_m + \delta_n$ or $\delta_{m+n} \geq \delta_m + \delta_n$ for all $m,n \geq 1$, then δ_n/n converges to $\inf_{n\geq 1} \frac{\delta_n}{n}$ or $\sup_{n\geq 1} \frac{\delta_n}{n}$, respectively. We will use a slight variation of this result which we state below. To obtain this result, its enough to take the sequence $\lambda_n = \ln(K\delta_n)$ and apply the Fekete's Lemma.

Lemma 5.4.4 (Fekete's Lemma)**.** Let $(\delta_n)_{n\geq 1}$ be a real valued sequence. Suppose that $\delta_n > 0$ for all $n \geq 1$ and that there exists a constant $K > 0$ such that $\delta_{m+n} \leq K(\delta_m\delta_n)$ or $\delta_{m+n} \geq K(\delta_m\delta_n)$ for all $m,n \geq 1$. Then the sequence $\left(\dfrac{\ln(\delta_n)}{n}\right)$ converges to $\inf\limits_{n\geq 1} \dfrac{\ln(K\delta_n)}{n}$ or $\sup\limits_{n\geq 1} \dfrac{\ln(K\delta_n)}{n}$, respectively.

Proof of Theorem 5.2.4. Let us start by assuming that $\psi'(1) > 0$. Then, there exists a positive constant K such that for all $m,n \geq 1$ and $a_0^{m+n-1} \in \mathscr{C}_{m+n}$ we have:

$$\begin{aligned}
& \mu\left(a_0^{m+n-1}\right) \geq K\mu\left(a_0^{m-1}\right)\mu\left(a_m^{m+n-1}\right) \\
\Rightarrow & \begin{cases} \mu\left(a_0^{m+n-1}\right)^{1-q} \geq K^{1-q}\mu\left(a_0^{m-1}\right)^{1-q}\mu\left(a_m^{m+n-1}\right)^{1-q}, & q \leq 1 \\ \mu\left(a_0^{m+n-1}\right)^{1-q} \leq K^{1-q}\mu\left(a_0^{m-1}\right)^{1-q}\mu\left(a_m^{m+n-1}\right)^{1-q}, & q > 1 \end{cases} \\
\Rightarrow & \begin{cases} \displaystyle\sum_{a_0^{m+n-1}\in\mathscr{C}_{m+n}} \mu\left(a_0^{m+n-1}\right)^{1-q} \geq K^{1-q} \sum_{a_0^{m+n-1}\in\mathscr{C}_{m+n}} \mu\left(a_0^{m-1}\right)^{1-q}\mu\left(a_m^{m+n-1}\right)^{1-q}, & q \leq 1 \\ \displaystyle\sum_{a_0^{m+n-1}\in\mathscr{C}_{m+n}} \mu\left(a_0^{m+n-1}\right)^{1-q} \leq K^{1-q} \sum_{a_0^{m+n-1}\in\mathscr{C}_{m+n}} \mu\left(a_0^{m-1}\right)^{1-q}\mu\left(a_m^{m+n-1}\right)^{1-q}, & q > 1. \end{cases}
\end{aligned} \tag{5.6}$$

Now, observe that

$$\sum_{a_0^{m+n-1}\in\mathscr{C}_{m+n}} \mu\left(a_0^{m-1}\right)^{1-q}\mu\left(a_m^{m+n-1}\right)^{1-q} = \left(\sum_{a_0^{m-1}\in\mathscr{C}_m} \mu\left(a_0^{m-1}\right)^{1-q}\right)\left(\sum_{a_0^{n-1}\in\mathscr{C}_n} \mu\left(a_0^{n-1}\right)^{1-q}\right),$$

where the case in which $\mathscr{A}$ is infinite follows from Mertens' Theorem (see (RUDIN, 1976)).

Thus, if we denote $\delta_n = \sum\limits_{a_0^{n-1}\in\mathscr{C}_n} \mu\left(a_0^{n-1}\right)^{1-q}$, we obtain from (5.6) that for all $m,n \geq 1$:

$$\begin{cases} \delta_{m+n} \geq K^{1-q}\delta_m\delta_n, & q \leq 1 \\ \delta_{m+n} \leq K^{1-q}\delta_m\delta_n, & q > 1. \end{cases}$$

Applying Lemma 5.4.4, we conclude that $\left(\dfrac{\ln(\delta_n)}{n}\right)$ converges and $H_R(q)$ exists.

Now we assume that $\psi^*(1) < \infty$. Then, we can get a positive constant K' such that for all $m,n \geq 1$ and $a_0^{m+n-1} \in \mathscr{C}_{m+n}$ we have

$$\mu\left(a_0^{m+n-1}\right) \leq K' \mu\left(a_0^{m-1}\right) \mu\left(a_m^{m+n-1}\right),$$

and the proof follows the same way as above. $\square$

The next two lemmas are a part of the proof of Theorem 5.2.6. For the first Lemma, consider a renewal process defined by the sequence $(q_i)_{i\geq 0}$ and suppose that $q_i \xrightarrow{i} L \in (0,1)$. We define the function $\varphi : \{0,1\}^\infty \longrightarrow \mathbb{R}$ given by

$$\varphi\left(x_0^\infty\right) := \begin{cases} \ln(q_k), & \text{if } x_0^\infty = 0^{k+1} 1 x_{k+2}^\infty \\ \ln(1-q_k), & \text{if } x_0^\infty = 10^k 1 x_{k+2}^\infty \\ \ln(L), & \text{if } x_0^\infty = 0^\infty \\ \ln(1-L), & \text{if } x_0^\infty = 10^\infty \end{cases}.$$

We also define for each $a_0^{n-1} \in \{0,1\}^n$:

$$S\left(a_0^{n-1}\right) := \exp\left(\sup_{x \in \left[a_0^{n-1}\right]} \sum_{k=0}^{n-1} \varphi\left(x_k^\infty\right)\right).$$

Lemma 5.4.5. Let $(Y_m)_{m\in\mathbb{N}}$ be a renewal process defined by the sequence $(q_i)_{i\geq 0}$ and suppose that $q_i \longrightarrow L \in (0,1)$. Then, there exists a sequence $(C_n)_{n\geq 1}$ such that $C_n \xrightarrow{n} 0$ and

$$e^{-nC_n} \leq \frac{\mu\left(a_0^{n-1}\right)}{S\left(a_0^{n-1}\right)} \leq e^{nC_n}$$

for all $a_0^{n-1} \in \{0,1\}^n$.

This lemma could be proved using arguments of ergodic theory for equilibrium states. Here, we preferred to make the calculations by hand.

Proof. First, we define the two following sequences for $n \geq 0$:

$$R_n := 1 + \sum_{j=0}^{\infty} \prod_{k=n}^{n+j} q_k = 1 + q_n + q_n q_{n+1} + \cdots$$

$$V_n := \sup_{j\geq n} \left\{ \left|\ln\left(\frac{R_n^{-1}}{1-q_j}\right)\right|, \left|\ln\left(\frac{R_n^{-1}}{1-L}\right)\right|, \left|\ln\left(\frac{1-R_n^{-1}}{q_j}\right)\right|, \left|\ln\left(\frac{1-R_n^{-1}}{L}\right)\right| \right\}.$$

Let us prove that $(R_n)^{-1} \longrightarrow 1-L$ and hence $V_n \longrightarrow 0$. Take an arbitrary $\varepsilon > 0$ such that $0 < L-\varepsilon < L+\varepsilon < 1$. Then we can get n_0 such that $L-\varepsilon < q_n < L+\varepsilon$ for all $n \geq n_0$ and

$$\begin{aligned} & \tfrac{1}{1+(L+\varepsilon)+(L+\varepsilon)^2+\cdots} < R_n^{-1} < \tfrac{1}{1+(L-\varepsilon)+(L-\varepsilon)^2+\cdots} \\ \Rightarrow \quad & 1-L-\varepsilon < R_n^{-1} < 1-L+\varepsilon \end{aligned}$$

for all $n \geq n_0$, which gives us the two convergences.

On the other hand, we have for any $a_0^{n-1} \in \{0,1\}^n$ and $x \in \left[a_0^{n-1}\right]$:

$$\frac{\mu\left(a_0^{n-1}\right)}{\mu\left(a_1^{n-1}\right)e^{\varphi(x)}} = \begin{cases} 1 & \text{if } a_0^{n-1} = 0^{k+1}1a_{k+2}^{n-1} \text{ or } a_0^{n-1} = 10^k1a_{k+2}^{n-1},\ k \geq 0 \\ \frac{R_{n-1}^{-1}}{1-L} \text{ or } \frac{R_{n-1}^{-1}}{1-q_j} & \text{if } a_0^{n-1} = 10^{n-1},\ x = 10^{\infty} \text{ or } x = 10^j1x_{j+2}^{\infty}, j \geq n-1 \\ \frac{1-R_{n-1}^{-1}}{L} \text{ or } \frac{1-R_{n-1}^{-1}}{q_j} & \text{if } a_0^{n-1} = 0^n,\ x = 0^{\infty} \text{ or } x = 0^{j+1}1x_{j+2}^{\infty},\ j+1 \geq n. \end{cases}$$

Furthermore, notice that for any of the above cases, we have

$$e^{-V_{n-1}} \leq \frac{\mu\left(a_0^{n-1}\right)}{\mu\left(a_1^{n-1}\right)e^{\varphi(x)}} \leq e^{V_{n-1}}.$$

Let us justify one of the cases and the others follow the same way. Just note that:

$$e^{-V_{n-1}} \leq e^{\ln\left(\frac{R_{n-1}^{-1}}{1-q_j}\right)} \leq e^{V_{n-1}}.$$

Similarly, we can show that

$$\begin{gathered} e^{-V_{n-2}} \leq \frac{\mu\left(a_1^{n-1}\right)}{\mu\left(a_2^{n-1}\right)e^{\varphi(\sigma(x))}} \leq e^{V_{n-2}} \\ e^{-V_{n-3}} \leq \frac{\mu\left(a_2^{n-1}\right)}{\mu\left(a_3^{n-1}\right)e^{\varphi(\sigma^2(x))}} \leq e^{V_{n-3}} \\ \vdots \\ e^{-V_0} \leq \frac{\mu(a_{n-1})}{e^{\varphi(\sigma^{n-1}(x))}} \leq e^{V_0}. \end{gathered}$$

Now we define

$$C_n = \frac{1}{n}\sum_{j=0}^{n-1} V_n,$$

which converges to zero since $V_n \longrightarrow 0$. Multiplying all the inequalities above, we get

$$e^{-nC_n} \leq \frac{\mu\left(a_0^{n-1}\right)}{\exp\left(\sum_{k=0}^{n-1}\varphi\left(x_k^{\infty}\right)\right)} \leq e^{nC_n}.$$

Since the above inequalities hold for all $x \in \left[a_0^{n-1}\right]$, we conclude that

$$e^{-nC_n} \leq \frac{\mu\left(a_0^{n-1}\right)}{S\left(a_0^{n-1}\right)} \leq e^{nC_n}.$$

□

Lemma 5.4.6. Let $(Y_m)_{m\in\mathbb{N}}$ be a renewal process defined by the sequence $(q_i)_{i\geq 0}$ and suppose that $q_i \longrightarrow L \in (0,1)$. Then, for all $q \in \mathbb{R}$, the following limit exists.

$$\lim_{n\to\infty}\frac{1}{n}\ln\sum_{a_0^{n-1}\in\mathscr{C}_n} S\left(a_0^{n-1}\right)^{1-q}$$

Proof. For $n \geq 1$, define the sequence

$$\delta_n = \sum_{a_0^{n-1} \in \mathscr{C}_n} S\left(a_0^{n-1}\right)^{1-q} = \sum_{a_0^{n-1} \in \mathscr{C}_n} \exp\left((1-q) \sup_{x \in \left[a_0^{n-1}\right]} \sum_{k=0}^{n-1} \varphi\left(x_k^\infty\right)\right).$$

On the other hand, we have

$$\begin{aligned}
&\sup_{x \in \left[a_0^{m-1}\right]} \left\{\sum_{k=0}^{m-1} \varphi\left(x_k^\infty\right)\right\} + \sup_{y \in \left[a_m^{m+n-1}\right]} \left\{\sum_{k=0}^{n-1} \varphi\left(y_k^\infty\right)\right\} \\
&= \sup_{x \in \left[a_0^{m-1}\right]; y \in \left[a_m^{m+n-1}\right]} \left\{\sum_{k=0}^{m-1} \varphi\left(x_k^\infty\right) + \sum_{k=0}^{n-1} \varphi\left(y_k^\infty\right)\right\} \\
&= \sup_{x \in \left[a_0^{m-1}\right]; y \in \{y_m^{m+n-1} = a_m^{m+n-1}\}} \left\{\sum_{k=0}^{m-1} \varphi\left(x_k^\infty\right) + \sum_{k=m}^{m+n-1} \varphi\left(y_k^\infty\right)\right\} \\
&\geq \sup_{z \in \left[a_0^{m+n-1}\right]} \left\{\sum_{k=0}^{m+n-1} \varphi\left(z_k^\infty\right)\right\}.
\end{aligned}$$

Thus, for $q < 1$ one has:

$$\begin{aligned}
&\delta_{m+n} \\
&= \sum_{a_0^{m+n-1} \in \mathscr{C}_{n+m}} \exp\left((1-q) \sup_{z \in \left[a_0^{m+n-1}\right]} \sum_{k=0}^{m+n-1} \varphi\left(z_k^\infty\right)\right) \\
&\leq \sum_{a_0^{m+n-1} \in \mathscr{C}_{n+m}} \exp\left((1-q) \sup_{x \in \left[a_0^{m-1}\right]} \left\{\sum_{k=0}^{m-1} \varphi\left(x_k^\infty\right)\right\}\right) \exp\left((1-q) \sup_{y \in \left[a_m^{m+n-1}\right]} \left\{\sum_{k=0}^{n-1} \varphi\left(y_k^\infty\right)\right\}\right) \\
&= \sum_{a_0^{m-1} \in \mathscr{C}_m} \exp\left((1-q) \sup_{x \in \left[a_0^{m-1}\right]} \left\{\sum_{k=0}^{m-1} \varphi\left(x_k^\infty\right)\right\}\right) \sum_{a_0^{n-1} \in \mathscr{C}_n} \exp\left((1-q) \sup_{y \in \left[a_0^{n-1}\right]} \left\{\sum_{k=0}^{n-1} \varphi\left(y_k^\infty\right)\right\}\right) \\
&= \delta_m \delta_n.
\end{aligned}$$

Therefore, we apply Fekete's Lemma 5.4.4 to conclude that $\lim_{n\to\infty} \frac{\ln(\delta_n)}{n}$ exists for all $q < 1$. By the same reasons as above, the case $q \geq 1$ gives us $\delta_{m+n} \geq \delta_m \delta_n$, which ends the proof. □

Proof of Theorem 5.2.6. (a) The statements for conditions (a1) and (a2) follow directly from Proposition 2.3.2 and Theorem 5.2.4. Suppose that (a3) holds. Then the process is ϕ-mixing, since $\limsup_i q_i = L < 1$ (see Corollary 2.2.5). On the other hand, Lemma 5.4.5 gives us

$$H_R(q) = \lim_{n\to\infty} \frac{1}{nq} \ln \sum_{A \in \mathscr{C}_n} \mu(A)^{1-q} = \lim_{n\to\infty} \frac{1}{nq} \ln \sum_{A \in \mathscr{C}_n} S(A)^{1-q},$$

which exists by Lemma 5.4.6.

(b) When $q \leq 1$, it is immediate to verify that

$$\left(\mu(1)(1-q_0)^{n-1}\right)^{1-q} = \mu\left(1^n\right)^{1-q} \leq \sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} \leq 2^n$$

Thus,

$$qH_R^-(q), qH_R^+(q) \in \left[(1-q)\ln(1-q_0), \ln(2)\right].$$

For the case $q > 1$, assume that $\sigma_n(1-q_n) \geq Ce^{-cn}$. In this case,

$$\mu\left(0^n\right) \geq \mu(1)\sigma_n(1-q_n) \geq \mu(1)Ce^{-cn}.$$

Furthermore, if $A \neq 0^n$, then we can write for some $k \geq 2$:

$$A = 0^{\ell_1}10^{\ell_2}1\cdots 10^{\ell_{k-1}}10^{\ell_k},$$

where $\ell_i \geq 0$, $i = 1,\cdots,k$, so that

$$\mu(A) = \mu(1)\sigma_{\ell_1}\sigma_{\ell_2}(1-q_{\ell_2})\cdots\sigma_{\ell_{k-1}}(1-q_{\ell_{k-1}})\sigma_{\ell_k} \geq \mu(1)C^k e^{-c\sum_i \ell_i}.$$

In other words, letting $n_0(A)$ the number of symbols 0 in the string A and $n_1(A) = n - n_0(A)$, we have

$$\mu(A) \geq \mu(1)C^k e^{-cn_0} \geq \mu(1)\min\{1, C^n\}e^{-cn}$$

Thus, for all $A \in \{0,1\}^n$, we have $\mu(A) \geq \mu(1)\min\{1,C^n\}e^{-cn}$. Therefore, for $q > 1$, we obtain

$$2^n \leq \sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} \leq 2^n(\mu(1)\min\{1,C^n\})^{1-q}e^{-c(1-q)n},$$

which implies that $qH_R^+(q)$ and $qH_R^-(q)$ are finite.

Finally, let us show that, for $q > 1$, the condition $\sigma_n(1-q_n) \geq Ce^{-cn}$ is necessary to obtain the finiteness of $H_R^+(q)$. Suppose that for every $D, d > 0$ there exist infinitely many $n \in \mathbb{N}$ such that $\sigma_n(1-q_n) \leq De^{-dn}$. In this case we would have, for infinitely many n,

$$\sum_{A\in\mathscr{C}_n} \mu(A)^{1-q} \geq \mu\left(10^{n-2}1\right)^{1-q} \geq \mu(1)^{1-q}D^{1-q}e^{(q-1)d(n-2)},$$

that is, $H_R^+(q) \geq (q-1)d$ for every $d > 0$.

□

Proof of the example involving Markov chains. Let μ be a stationary measure of a house of cards Markov chain satisfying $S = \sup_i\{q_i\} < 1$. First, let us show that $\psi'(1) > 0$. We compute for all $m, n \geq 1$ and $a_0^{m+n-1} \in \mathscr{C}_{m+n}$:

$$\frac{\mu\left(a_0^{m+n-1}\right)}{\mu\left(a_0^{m-1}\right)\mu\left(a_m^{m+n-1}\right)} = \frac{Q(a_{m-1}, a_m)}{\mu(a_m)}.$$

Further, notice that for all $j, k \in \mathbb{N}$ we have

$$\frac{Q(j,k)}{\mu(k)} = \begin{cases} \frac{1-q_j}{\mu(0)}, & \text{if } k = 0 \\ \frac{1}{\mu(0)\sigma_j}, & \text{if } k = j+1 \end{cases} \geq \frac{1-S}{\mu(0)},$$

which implies $\psi'(1) > 0$.

It remains to show that $H_R(q)$ is finite and that (5.1) holds for all $q \in (0,1)$. In the following, we show both ones at the same time. For the lower bound we have

$$\sum_{a_0^{n-1} \in \mathscr{C}_n} \mu\left(a_0^{n-1}\right)^{1-q} \geq \mu\left(0^n\right)^{1-q} = \left(\mu(0)(1-q_0)^{(n-1)}\right)^{1-q},$$

which implies

$$qH_R(q) \geq (1-q)\ln(1-q_0)$$

for all $q < 1$.

On the other hand, consider a word $A = a_0^{n-1} \in \mathscr{C}_n$. Observe that for each $a_0 \in \mathbb{N}$ there exist exactly 2^{n-1} words $a_0^{n-1} \in \mathscr{C}^n$. To justify this statement, note that there exists 2^{n-1} ways to choose the zeros in the word a_1^{n-1}. Given a_0 and the zeros, there is only one way to complete the word with positive measure.

Furthermore, if A has at most $n/2$ zeros, then the product $\prod_{j=0}^{n-2} Q(a_j, a_{j+1})$ has at least $n/2 - 1$ factors of type q_i, which means that it is bounded above by $S^{n/2-1}$. In the same way, if A has more than $n/2$ zeros, then $\prod_{j=0}^{n-2} Q(a_j, a_{j+1}) \leq \max\{q_0, (1-q_0)\}^{n/2-1}$. Hence, if we denote $M = \max\{S, (1-q_0)\}^{n/2-1}$, we get

$$\begin{aligned}
\sum_{a_0^{n-1} \in \mathscr{C}_n} \mu\left(a_0^{n-1}\right)^{1-q} &= \sum_{a_0^{n-1} \in \mathscr{C}_n} \left(\mu(a_0) \prod_{j=0}^{n-2} Q(a_j, a_{j+1})\right)^{1-q} \\
&\leq \left(\mu(0) M^{n/2-1}\right)^{1-q} \sum_{a_0 \in \mathbb{N}} 2^{n-1} (\sigma_{a_0})^{1-q} \\
&= K_q \left(\mu(0) M^{n/2-1}\right)^{1-q} 2^{n-1},
\end{aligned}$$

where

$$K_q = \sum_{a_0 \in \mathbb{N}} (\sigma_{a_0})^{1-q} \leq \sum_{k \in \mathbb{N}} \left(S^{1-q}\right)^k < \infty.$$

Therefore,

$$qH_R(q) \leq \frac{(1-q)\ln M}{2} + \ln 2.$$

Thus, Theorem 5.2.4 ensures that $H_R(q)$ exists for all $q < 1$ and, by Theorem 5.1.2, $\mathscr{W}(q)$ also exists and is finite. $\square$

BIBLIOGRAPHY

ABADI, M. Exponential approximation for hitting times in mixing processes. **Mathematical Physics Electronic Journal**, v. 7, 2001. Citations on pages 44, 53, 101, and 102.

______. Sharp error terms and neccessary conditions for exponential hitting times in mixing processes. **Annals of Probability**, v. 32, n. 1A, p. 243–264, 2004. Available: <https://doi.org/10.1214/aop/1078415835>. Citations on pages 20, 42, and 44.

______. Hitting, returning and the short correlation function. **Bulletin of the Brazilian Mathematical Society**, v. 37, n. 4, p. 593–609, 2006. Available: <https://doi.org/10.1007/s00574-006-0030-1>. Citation on page 44.

ABADI, M.; AMORIM, V.; GALLO, S. Potential well in poincaré recurrence. **Entropy**, v. 23, n. 3, 2021. Available: <https://www.mdpi.com/1099-4300/23/3/379>. Citations on pages 41, 45, 48, 50, 69, and 70.

ABADI, M.; CARDENO, L. Rényi entropies and large deviations for the first match function. **IEEE Transactions on Information Theory**, v. 61, n. 4, p. 1629–1639, 2015. Available: <10.1109/TIT.2015.2406695>. Citations on pages 31, 32, 33, 47, 98, and 99.

ABADI, M.; CARDENO, L.; GALLO, S. Potential well spectrum and hitting time in renewal processes. **Journal of Statistical Physics**, v. 159, n. 5, p. 1087 – 1106, 2015. Available: <https://doi.org/10.1007/s10955-015-1216-y>. Citations on pages 28, 42, 45, 69, 70, 74, 78, and 79.

ABADI, M.; CHAZOTTES, J.-R.; GALLO, S. The complete l^q-spectrum and large deviations for return times for equilibrium states with summable potentials. **arXiv**, 2019. Available: <https://arxiv.org/abs/1902.03441>. Citations on pages 21 and 95.

ABADI, M.; GALLO, S.; RADA-MORA, E. A. The shortest possible return time of β-mixing processes. **IEEE Transactions on Information Theory**, v. 64, n. 7, p. 4895–4906, 2018. Available: <10.1109/TIT.2017.2757494>. Citation on page 47.

ABADI, M.; GALVES, A. Inequalities for the occurrence times of rare events in mixing processes: the state of the art. **Markov Process. Related Fields**, Polymat, v. 7, n. 1, p. 97–112, 2000. Citations on pages 18, 41, and 42.

ABADI, M.; LAMBERT, R. The distribution of the short-return function. **Nonlinearity**, v. 26, n. 5, p. 1143–1162, 2013. Available: <https://doi.org/10.1088/0951-7715/26/5/1143>. Citation on page 47.

ABADI, M.; SAUSSOL, B. Hitting and returning to rare events for all alpha-mixing processes. **Stochastic Processes and their Applications**, v. 121, n. 2, p. 314–323, 2011. Available: <https://doi.org/10.1016/j.spa.2010.11.001>. Citations on pages 42, 45, 73, and 88.

_____. Almost sure convergence of the clustering factor in α-mixing processes. **Stochastics and Dynamics**, v. 16, n. 3, 2016. Available: <https://doi.org/10.1142/S0219493716600169>. Citations on pages 42, 45, 70, 72, 73, 88, and 91.

ABADI, M.; VAIENTI, S. Large deviations for short recurrence. **Discrete and Continuous Dynamical Systems-Series A**, v. 21, n. 3, p. 729–747, 2008. Citation on page 47.

ABADI, M.; VERGNE, N. Sharp error terms for return time statistics under mixing conditions. **Journal of Theoretical Probability**, v. 22, n. 1572-9230, p. 18–37, 2008. Available: <http://dx.doi.org/10.1007/s10959-008-0199-x>. Citations on pages 19, 42, 45, 46, 48, 49, 51, 53, 55, 64, and 72.

AFRAIMOVICH, V.; CHAZOTTES, J.-R.; SAUSSOL, B. Pointwise dimensions for poincaré recurrences associated with maps and special flows. **Discrete & Continuous Dynamical Systems**, v. 9, n. 2, p. 263–280, 2003. Available: <http://dx.doi.org/10.3934/dcds.2003.9.263>. Citations on pages 20, 47, and 81.

ALDOUS, D.; BROWN, M. Inequalities for rare events in time-reversible markov chains ii. **Stochastic Processes and their Applications**, v. 44, n. 1, p. 15–25, 1993. Available: <https://www.sciencedirect.com/science/article/pii/030441499390035З>. Citation on page 44.

BRADLEY, R. C. Basic properties of strong mixing conditions. a survey and some open questions. **Probability Surveys**, Institute of Mathematical Statistics and Bernoulli Society, v. 2, n. none, p. 107–144, 2005. Available: <https://doi.org/10.1214/154957805100000104>. Citations on pages 24, 25, 31, 32, 34, 35, 37, 39, and 72.

CHAZOTTES, J.-R.; UGALDE, E. Entropy estimation and fluctuations of hitting and recurrence times for gibbsian sources. **Discrete & Continuous Dynamical Systems-B**, v. 5, n. 3, p. 565–586, 2005. Available: <http://dx.doi.org/10.3934/dcdsb.2005.5.565>. Citations on pages 20, 94, 95, 98, and 100.

COLLET, P.; GALVES, A.; SCHMITT, B. Repetition times for gibbsian sources. **Nonlinearity**, v. 12, n. 4, p. 1225–1237, 1999. Available: <https://doi.org/10.1088/0951-7715/12/4/326>. Citation on page 42.

FREITAS, A. C. M.; FREITAS, J. M.; TODD, M. Hitting time statistics and extreme value theory. **Probability Theory and Related Fields**, v. 147, n. 3-4, p. 675–710, 2010. Available: <https://doi.org/10.1142/S0219493716600169>. Citation on page 17.

GALVES, A.; SCHMITT, B. Inequalities for hitting times in mixing dynamical systems. **Random Comput. Dynam.**, v. 5, n. 4, p. 337–348, 1997. Citations on pages 18, 42, and 44.

GRASSBERGER, P.; PROCACCIA, I. Dimensions and entropies of strange attractors from a fluctuating dynamics approach. **Physica D: Nonlinear Phenomena**, v. 13, p. 34–54, 1984. Citation on page 98.

HAYDN, N. Entry and return times distribution. **Stochastics and Dynamics**, 2013. Available: <https://arxiv.org/abs/1306.4475v1>. Citation on page 17.

HAYDN, N.; VAIENTI, S. The rényi entropy function and the large deviation of short return times. **Ergodic Theory and Dynamical Systems**, v. 30, n. 1, p. 159–179, 2010. Available: <https://doi.org/10.1017/S0143385709000030>. Citations on pages 47 and 98.

HIRATA, M.; SAUSSOL, B.; VAIENTI, S. Statistics of return times: A general framework and new applications. **Communications in Mathematical Physics**, v. 206, 2000. Citations on pages 45 and 72.

KAC, M. On the notion of recurrence in discrete stochastic processes. **Bulletin of the American Mathematical Society**, American Mathematical Society, v. 53, n. 10, p. 1002 – 1010, 1947. Available: <https://doi.org/>. Citation on page 17.

KO, M. **Rényi Entropy and Recurrence**. Phd Thesis (PhD Thesis) — Faculty of the USC Graduate School - Universty of Southern California, 2012. Citation on page 98.

LACROIX, Y. Possible limit laws for entrance times of an ergodic aperiodic dynamical system. **Israel Journal of Mathematics**, v. 132, p. 253–263, 12 2002. Citation on page 42.

LUCARINI, V.; FARANDA, D.; FREITAS, A. C. G. M. M.; FREITAS, J. M. M.; HOLLAND, M.; KUNA, T.; NICOL, M.; TODD, M.; VAIENTI, S. **Extremes and Recurrence in Dynamical Systems**. [S.l.]: John Wiley & Sons: Hoboken, 2016. Citation on page 17.

LUCZAK, T.; SZPANKOWSKI, W. A suboptimal lossy data compression based on approximate pattern matching. **IEEE Transactions on Information Theory**, v. 43, n. 5, p. 1439–1451, 1997. Citation on page 98.

PITTEL, B. Asymptotical growth of a class of random trees. **The Annals of Probability**, Institute of Mathematical Statistics, v. 13, n. 2, p. 414–427, 1985. Available: <https://doi.org/10.1214/aop/1176993000>. Citation on page 98.

PLACHKY, D.; STEINEBACH, J. G. A theorem about probabilities of large deviations with an application to queuing theory. **Periodica Mathematica Hungarica**, v. 6, p. 343–345, 1975. Citation on page 95.

RUDIN, W. **Principles of Mathematical Analysis**. [S.l.]: McGraw-Hill, Inc., 1976. Citation on page 106.

SAUSSOL, B.; TROUBETZKOY, S.; VAIENTI, S. Recurrence, dimensions, and lyapunov exponents. **Journal of Statistical Physics**, v. 106, n. 3, p. 623–634, 2002. Available: <https://doi.org/10.1023/A:1013710422755>. Citations on pages 20, 47, 71, and 81.

SHIELDS, P. C. **The Ergodic Theory of Discrete Sample Paths**. [S.l.]: American Mathematical Society, 1996. Citations on pages 17, 20, 42, 52, and 94.

SZPANKOWSKI, W. A generalized suffix tree and its (un)expected asymptotic behaviors. **SIAM Journal on Computing**, v. 22, 1996. Citation on page 98.

WYNER, A. D.; ZIV, J. Some asymptotic properties of the entropy of a stationary ergodic data source with applications to data compression. **IEEE Transactions on Information Theory**, v. 35, n. 6, p. 1250–1258, 1989. Available: <https://doi.org/10.1109/18.45281>. Citations on pages 20 and 93.

ICMC USP
SÃO CARLOS