

Critical behaviour of random interval diffeomorphisms: quasi-stationary measures and escape times

A Thesis Submitted for the Degree of
Doctor of Philosophy of Imperial College London

by

Guillermo Olicón Méndez

Supervisors: Dr. Martin Rasmussen and Prof. Jeroen S. W. Lamb

Department of Mathematics
Imperial College London
October 2020

Declaration

I declare that the work here presented is entirely my own, and that any material which is based on the work of others it is fully acknowledged and cited.

Guillermo Olicón Méndez

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

*Millions and billions of keys, which never seem to end and this is the real truth,
that they would never stop and that the keyboard is endless.
If this keyboard has no end, then... There's no music you could play on it.
You had the wrong seat - that's the piano on which God plays.*

Alessandro Baricco, *Novecento*.

Acknowledgements

There are no words to express my gratitude to my supervisors Martin Rasmussen and Jeroen Lamb, for the uncountable and gratifying discussions we had over the process of my PhD studies, for always being critical in our meetings, and for their infinite support and patience since day one. This thesis is just a glimpse of the collaboration we built over the past years.

I would like to express my gratitude to Dmitry Turaev, for the many fruitful conversations over the past years, and for being an examiner of this thesis. My thanks also go to Denis Villemonais, for agreeing to be an examiner of the work here presented, and his valuable comments. I would like to thank Sebastian Van Strien, Ale Jan Homburg and Yuzuru Sato, for useful discussions.

I feel very lucky to have spent so much time around friendly colleagues at Imperial College London. I am very grateful with Kalle Timperi, So Takao, and Matheus Manzatto, with whom I had lengthy and passionate discussions on different topics related to Dynamical Systems. My infinite gratitude goes to Maximilian Engel, for the never ending mathematical and non-mathematical conversations, but also for his always friendly encouragement in the most difficult times.

This work would have never seen its end without the infinite support of the many friends I made over the past years. To Alfredo Lozano, for his unconditional friendship despite the distance. To Matthew Davies, James Cox, and Veronica Glove, for marking a big difference during my time in London since we met. To Gabriela Depetri, Michael Hartl, Kathinka Nohl, Frith Carlisle, Clara Ingram, Caroline Mathon, Adrian Pablo, Malika Sharma, Sofía Trejo, Alejandra Garza, and María José González: for their friendship, their support, the laughs, and the good fun.

To my colleague and good friend Andrew Clarke, for his friendship, and for the so many debates, ideas, and conversations we exchanged over the last years. I look forward to the many more of those awayting for us in the future.

To my parents, Adriana Méndez and Guillermo Olicón, and to my brother Daniel Olicón, for the trust they deposited on me. Their love and support always gave me courage and strength since day one. To Silvia Vivas, for her support and her trust. To Camerino Estrada, for the mezcal, his friendship, and his trust: those will always be with me forever.

To Laura Yaniz Estrada, for her infinite love and support, for her extraordinary strength required for putting up in the hardest times over the last four years, and for always holding hands even in the distance.

This work would not have been possible without the support of the Consejo Nacional de Ciencia y Tecnología (CONACYT), and the Department of Mathematics at Imperial College London.

Abstract

In this thesis we consider discrete-time dynamical systems in the interval perturbed with bounded noise near a bifurcation of a minimal invariant set, its relation with stationary and quasi-stationary measures, and its asymptotic behaviour as the parameters tend to the bifurcation point.

First, we study existence and uniqueness of stationary measures on minimal invariant sets, and its approximation with the transition probabilities of the system. We derive analogous results for the unique stationary density, and conclude that the system exhibits exponential decay of (annealed) correlations. We further study the shape of the stationary density near the boundary of its support, and relate it to the hyperbolicity of the the boundary point for the so-called *extremal maps*.

Secondly, we study a similar problem for quasi-stationary measures supported on a metastable set, when the parameters are close to a bifurcation point. We prove existence and uniqueness of quasi-stationary measures, and the exponential convergence of conditional transition probabilities of the system towards this measure. We study the asymptotic behaviour of the mean escape time from the metastable region as the parameter tends to the bifurcation point.

The results here presented relate the bifurcation theory of dynamical systems under bounded noisy perturbations, with their statistical features. To our knowledge these are among the first of their kind.

Contents

Declaration	3
Copyright	5
Acknowledgements	9
Abstract	11
1 Introduction	19
2 Stationary measures for random diffeomorphisms	31
2.1 Main result and structure of the chapter	32
2.2 Exponential convergence to the stationary measure	33
2.3 Exponential decay of correlations	36
2.4 Convergence of densities in uniform topology	37
3 Tails of stationary measures	45

3.1	Main result and structure of the chapter	46
3.2	Embedding of deterministic maps into flows	47
3.3	Nonhyperbolic boundaries of minimal invariant sets	51
3.3.1	Numerical explorations	58
3.4	Hyperbolic boundaries	61
3.4.1	Numerical explorations	65
4	Quasi-stationary measures for Random Diffeomorphisms	67
4.1	Main result and structure of the chapter	69
4.2	Conditional Koopman and Markov Operators	70
4.3	Conditional transfer operator	75
4.4	Exponential convergence to quasi-stationary measures	77
5	Bifurcations and mean escape times	83
5.1	Main result and structure of the chapter	85
5.2	Mean escape time	87
5.3	Approximation of saddle-node families via flows	88
5.4	Deterministic escape times in saddle-node bifurcations	92
5.5	Asymptotic behaviour for mean escape times	97
5.6	Numerical experiments	105

<i>CONTENTS</i>	15
6 Outlook	109
A Total variation distance	113
B Set-valued dynamics and minimal invariant sets	117
C Functional operators for random maps	121
C.1 Transition probabilities for random maps	121
C.2 The Markov operator and stationary measures	123
C.3 The stochastic Koopman operator	125
C.4 The transfer operator and stationary densities	126
D Some results on hyperbolic fixed points	131
D.1 Exponential convergence of conditional set-valued dynamics	133

List of Figures

1.1	A deterministic bifurcation diagram of a toy model	22
1.2	Flickering in a toy model	23
1.3	An example of bifurcation of minimal invariant sets	24
1.4	Topological bifurcation of minimal invariant sets by changing the noise strength . .	25
2.1	Proof of Theorem 4.15	34
3.1	Integration region in the proof of Lemma 3.8	53
3.2	Upper estimate by an integral in Theorem 3.1	56
3.3	Numerical exploration of Theorem 3.1 in the nonhyperbolic case	61
3.4	Numerical exploration of Theorem 3.1 in the hyperbolic case	66
4.1	Escaping from almost invariant sets	68
5.1	Topological bifurcation of minimal invariant sets induced by the extremal maps . .	84
5.2	Survival rate as a function of the noise strength σ	85

5.3	Bottlenecking in deterministic saddle-node bifurcations	96
5.4	Integration region in the proof of Lemma 5.12	99
5.5	Upper estimate by an integral in Theorem 5.1	103
5.6	Numerical exploration of Theorem 5.1	106
5.7	Linear fitting of numerical explorations in Theorem 5.1	107
B.1	Transient and intermittency bifurcations	120

Chapter 1

Introduction

In recent years the scientific community's attention has been strongly drawn towards analysing abrupt changes in a wide range of natural and social phenomena, many of which have a direct impact in human societies. For instance, and despite some skeptical opinions in this regard, it is undeniable that the human activity has had a massive influence in the environment in the last two centuries, which has driven the average temperatures in different regions of the world to rise constantly above preoccupaying levels. The *global warming* effect has already generated radical changes in the climate behaviour, which affects the general state of ecosystems, the behaviour of its biodiversity, and the welfare of its human communities, specially of those which have always lived under precarious conditions.

From a classical point of view, the bifurcation theory of dynamical systems has been quite succesful in describing qualitative changes that occur by varying the parameters of the system. However, many real-world phenomena like climate change as mentioned above, are under the influence of exogenous forces of random nature, which can trigger a massive impact in their dynamical behaviour. Abrupt changes in complex systems which are driven by external influences or forces are called *critical transitions* [51].

In this thesis we study a simple mechanism for *noise-induced transitions* to occur in a *random dynamical systems* framework, where we have additive noisy perturbations of an iterative deterministic map $T : \mathbb{R} \rightarrow \mathbb{R}$. In other words, we work with *random difference equations* of the form

$$x_{n+1} = T(x_n) + \sigma\omega_n, \tag{1.1}$$

where T is a sufficiently smooth map such that $T' > 0$ and the sequence $(\omega_n)_{n=0}^\infty$ is an independently identically distributed (i.i.d.) sequence of random variables in $\Delta = [-1, 1]$. Each ω_n is distributed according to a probability measure $d\nu = p \, d\text{Leb}$ on $(\Delta, \mathcal{B}(\Delta))$, where p is a Lipschitz-continuous function, and Leb denotes the Lebesgue measure on $[-1, 1]$. As usual, the time is represented by $n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, where $\mathbb{N}$ is the set of natural numbers. We call the parameter σ the *noise strength*, as it measures how big the noisy perturbations are.

The system (1.1) is an example of a random dynamical system in discrete time with *bounded noise*. An alternative and more analytically tractable class of random dynamical systems with bounded noise are the so-called *iterated function systems*, where at each unit of time a deterministic map is chosen randomly from a finite collection of functions [6]. On the other hand, *random differential equations* [34, 35] can be examples of continuous-time systems with bounded noise, of which differential equations forced by *dichotomous Markov noise* (also called *dichotomous flows*) [7] are particular cases.

Despite the simple formulation of system (1.1), it has proven to be a hard task to do rigorous analysis on this type of systems, even in the simplest cases. For instance, whilst for *stochastic differential equations* stationary solutions of the Fokker-Planck equation provide explicit formulas for the stationary density of the system, in the case for random difference equations it is a hard task to even obtain an explicit formula in the simplest examples, like the linear case $T(x) = ax$ with $a \in \mathbb{R}$.

Alternatively, many mathematical models have been constructed using stochastic differential equations, and have proven to be quite successful in explaining different phenomena, like transitions in chemical potentials. In this context, the so-called *Eyring-Kramers law* [27, 38] relates the *mean first-passage time* between minima in a potential landscape that a Brownian particle takes, with the geometrical features of the barrier that the particle must cross (see [8] and references therein). The Eyring-Kramers law is a significant improvement of the *Arrhenius law* [4], which was already formulated as an exponential expression of the mean first-passage time (originally stated as a reaction rate).

The use of unbounded noisy perturbations, like in most examples using stochastic differential equations, has little justification in certain applications, like in biological, chemical or ecological systems, where quantities are unlikely to be unbounded or even negative. The use of Gaussian noise is sometimes justified by means of the central limit theorem. However, we believe this perspective blurs out some rich mechanistic and dynamical features of the phenomena under study.

In this thesis we contribute to the qualitative and statistical theory of random interval diffeo-

morphisms. The work here presented lies in the intersection between the bifurcation theory of random dynamical systems, its statistical behaviour, and its functional analytical properties. In particular, we are interested in some statistical features near a bifurcation of the *minimal invariant sets* of the system. Roughly speaking, minimal invariant sets are regions in the state space from which it is impossible for the orbits of (1.1) to escape from (see Appendix B for a full characterisation of such sets). Here, a *topological bifurcation* [44] at σ^* of a minimal invariant set refers to its discontinuous variation in the Hausdorff metric [6]. Once the system undergoes a bifurcation of this type, such bifurcating minimal invariant set stops being invariant for values $\sigma > \sigma^*$, and the orbits of (1.1) have positive probability of escaping from it, having thus the possibility of exploring a distant region of the state space.

In order to introduce the questions we address in the present document, let us present a toy model which exemplifies the phenomenology described in the paragraph above.

A toy model

Let us consider the following deterministic map

$$T(x) = b \cdot \frac{1 - e^{-x}}{1 + e^{-x}}, \quad (1.2)$$

where $b > 2$ is fixed. This sigmoidal map is everywhere increasing, and satisfies $T(0) = 0$ and $T'(0) > 1$, so that it admits two asymptotically stable fixed points. When considering also an additive parameter $a \in \mathbb{R}$ along with the map T , the dynamical system generated by

$$x_{n+1} = a + T(x_n) \quad (1.3)$$

exhibits two bifurcations for critical values a_1 and a_2 (see Figure 1.1)

In order to find the bifurcation values a^* we look for pairs (x^*, a^*) such that $T'(x^*) = 1$ and $T(x^*) + a^* = x^*$, so that x^* is a nonhyperbolic fixed point for $a^* + T$. First, notice that

$$T'(x) = \frac{2be^{-x}}{(1 + e^{-x})^2}.$$

Therefore, $T'(x) = 1$ if and only if

$$e^{-2x} - 2(b - 1)e^{-x} + 1 = 0,$$

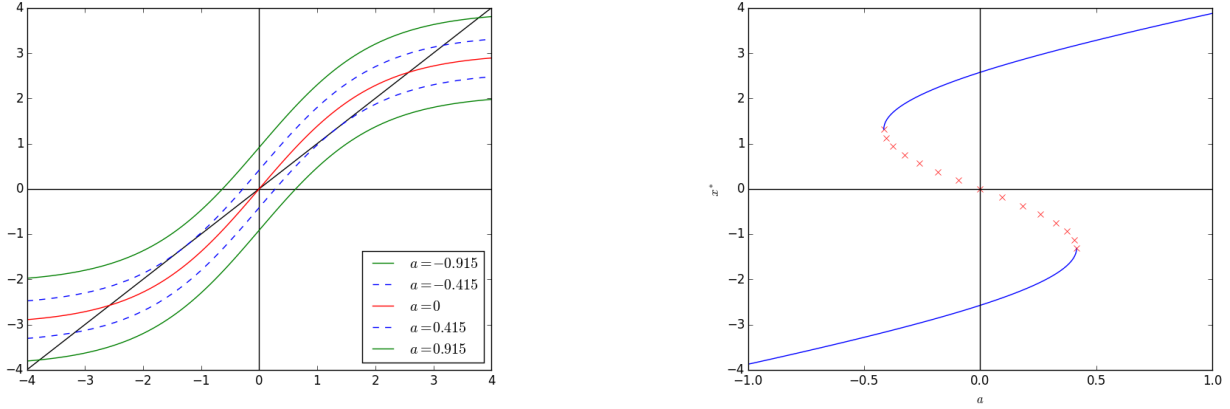


Figure 1.1: For $b = 3$, on the left the graph of T is shown in a red solid line, while in blue dotted lines the graphs of $a + T$ are presented for values of a where the system undergoes through a nondegenerate saddle-node bifurcation. Two examples past the bifurcation are shown in green. On the right, the bifurcation diagram is depicted.

which is a quadratic equation in $u = e^{-x}$. By solving the equation we obtain two solutions, $x_{1,2}$, given by

$$x_{1,2} = -\ln \left(b - 1 \pm \sqrt{(b-1)^2 - 1} \right),$$

where one of them is negative and the other one is positive. For this example, let us consider $x^* = -\ln \left(b - 1 + \sqrt{(b-1)^2 - 1} \right)$ to be the negative one. We now want to find $a^* > 0$ such that $T(x^*) + a^* = x^*$. By substituting the value of x^* we get

$$a^* = b \left(\frac{b-2 + \sqrt{(b-1)^2 - 1}}{b + \sqrt{(b-1)^2 - 1}} \right) - \ln \left(b - 1 + \sqrt{(b-1)^2 - 1} \right)$$

Since

$$T'(x) = \frac{2be^{-x}}{(1+e^{-x})^2}, \quad T''(x) = 2be^{-x} \frac{e^{-x} - 1}{(e^{-x} + 1)^3} = -\frac{T'(x)T(x)}{b},$$

then

$$T''(x^*) = -\frac{x^* - a^*}{b} = \frac{b-2 + \sqrt{(b-1)^2 - 1}}{b + \sqrt{(b-1)^2 - 1}}. \quad (1.4)$$

Therefore, we can write $T(x) + a$ as

$$a + T(x) = a - a^* + x + \alpha(x - x^*)^2 + o((x - x^*)^2),$$

where $\alpha = -(x^* - a^*)/(2b) > 0$. Hence, the system undergoes a *nondegenerate saddle-node bifurcation* at (x^*, a^*) [43].

Consider now the randomly perturbed system

$$x_{n+1} = a + T(x_n) + \sigma\omega_n,$$

as formulated in (1.1). A simple first approach in studying the random difference equation above consists of plotting some time series for different values of (a, σ) . By fixing the value of $\sigma > 0$, we observe that a qualitative change in the behaviour of the system can occur. In fact, there are values of $a < 0$ where the orbits of the system tend to stay in a region contained in $\mathbb{R}_-$, whereas for some values $a > 0$ the orbits remain confined in a region far away from the previous one and contained in $\mathbb{R}_+$. However, for intermediate values of $a \approx 0$, the orbits seem to hop between these two sites (see Figure 1.2).

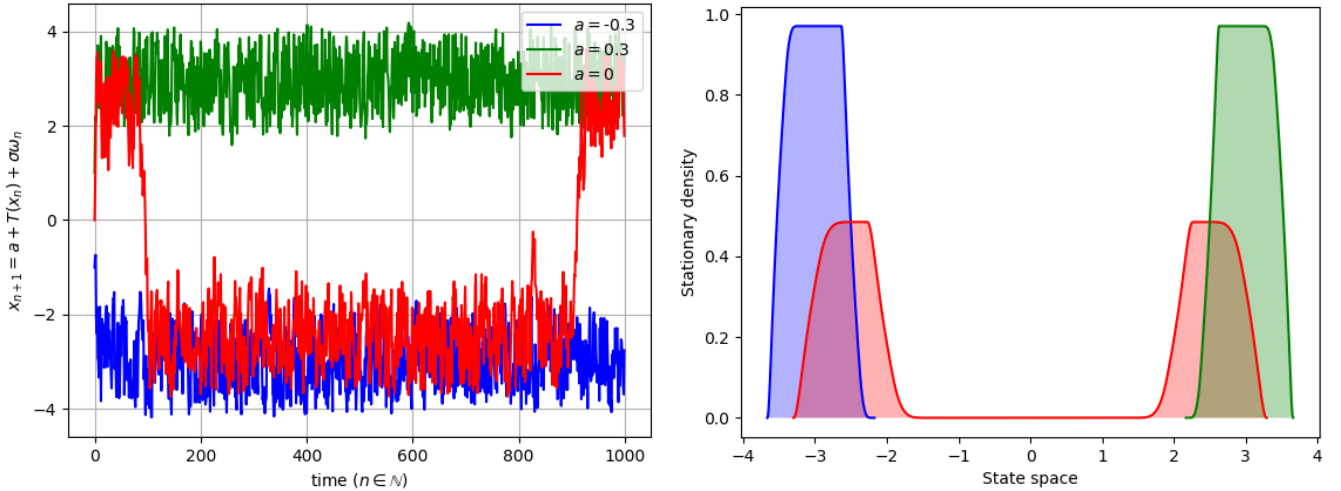


Figure 1.2: On the left, we plot three different orbits of the system $x_{n+1} = a + b \frac{1-e^{-x_n}}{1+e^{-x_n}} + \sigma\omega_n$, for values $a \in \{-0.3, 0, 0.3\}$. For $a = -0.3$ and $a = 0.3$, the blue and green time series respectively, remain bounded away from each other. On the other hand, when $a = 0$ the red series hops between both regions. On the right, we portray the corresponding stationary measure of the system, through their corresponding stationary density for each value of a . Here $b = 3$, and $\sigma = 0.515$ are fixed.

Due to the boundedness of the noise, the system cannot explore an arbitrarily large portion of the state space. Instead, the system eventually is confined in its *minimal invariant sets* (see

Appendix B). The intermittent behaviour we observe in Figure 1.2 may as well not be present at all. We observe in Figure 1.3 that a *flickering* or *intermittent* behaviour actually only happens whenever the noise strength σ is large enough. Otherwise, by changing the parameter a , one of the minimal invariant sets might simply disappear.

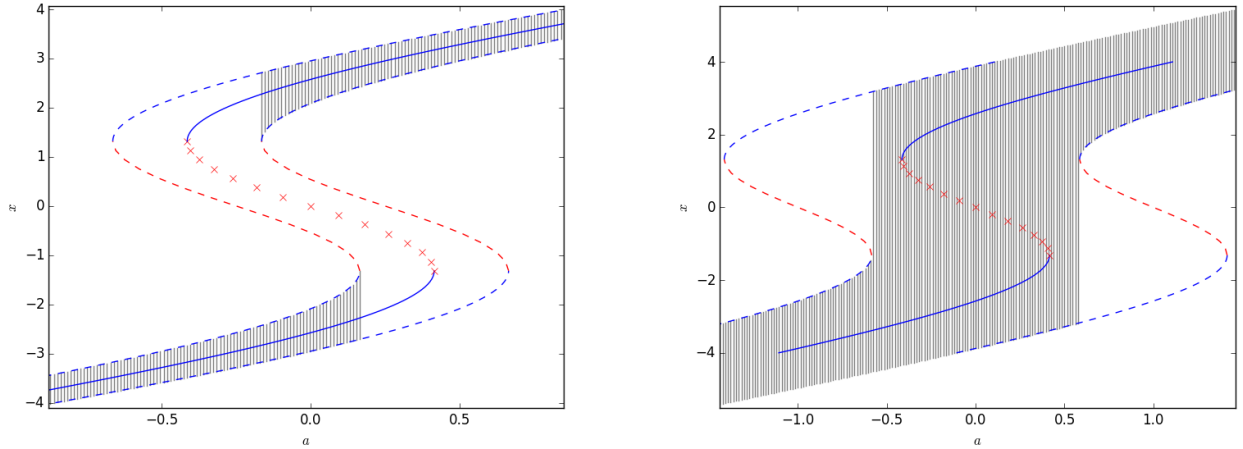


Figure 1.3: Different types of bifurcations of the minimal invariant sets can occur depending on the value of the parameters, which can be portrayed in a bifurcation diagram. In both examples here presented, the value of σ is fixed. On the left, we take σ less than a critical value σ^* so that there are no values of the parameter a where the minimal invariant sets “merge”. Instead, there are critical values of a for which the minimal invariant sets disappear. On the right, the scenario is totally the opposite, and there are values of a for which the minimal invariant sets get enlarged, and then an intermittent behaviour can happen. In the left we have thus portrayed an example of a *transient* bifurcation, while on the right we depict an *intermittency bifurcation* [59].

In this thesis we assume all deterministic parameters are fixed and consider $\sigma > 0$ to be our bifurcation parameter. In this case, minimal invariant sets grow by increasing σ . As anticipated in Figure 1.3, there can be a critical value $\sigma^* > 0$ for which this growth is discontinuous, see Figure 1.4. In this sense we agree with [20]: flickering or intermittent behaviours are present in highly stochastic systems.

In the rest of the thesis, this toy model will be used for numerical explorations. From now on, we fix the value $a = 0$ and explore different behaviours the system may exhibit when varying σ . In this sense, we are looking for *noise-induced phenomena*. From the discussion above we address the following questions:

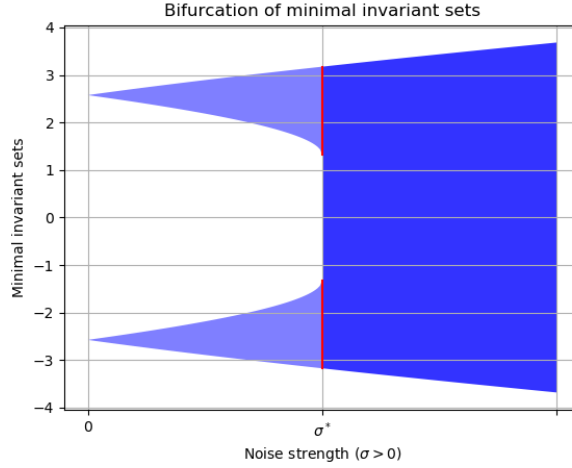


Figure 1.4: In this plot, we fix $a = 0$ and present the minimal invariant sets of the system for each $\sigma > 0$. Notice there is σ^* where two different minimal invariant sets seem to merge discontinuously into a bigger one. We will be interested in studying this type of bifurcation.

- The system admits a stationary measure. How can we approximate it? How does this translate to the stationary densities?
- Is there any feature in the stationary densities which reflect the type of bifurcation the system is undergoing through?
- Assume a region M were invariant for σ^* , but stops being so for any $\sigma > \sigma^*$. In average, what is the time the orbits of (1.1) remain confined in M before they escape from it?

In order to address these questions, let us consider the random dynamical system formalism induced by (1.1): consider

$$\Omega \equiv \Delta^{\mathbb{N}} := \{(\omega_n)_{n=0}^{\infty} \mid n \in \mathbb{N}_0, \omega_n \in \Delta\},$$

endowed with its Borel σ -algebra, $\mathcal{B}(\Delta^{\mathbb{N}})$ (generated by the family of cylinder sets), and the infinite product probability measure $\mathbb{P} : \mathcal{B}(\Omega) \rightarrow [0, 1]$ where for all $k \in \mathbb{N}$ and $A_1, \dots, A_k \in \mathcal{B}(\Delta)$,

$$\mathbb{P}(A_1 \times \dots \times A_k \times \Omega) = \nu(A_1) \cdots \nu(A_k).$$

By introducing the *shift map*, $\theta : \Omega \rightarrow \Omega$, given by

$$\theta(\omega_n)_{n=0}^{\infty} = (\omega_{n+1})_{n=0}^{\infty}.$$

The tuple $(\Omega, \mathcal{B}(\Omega), \mathbb{P}, \theta)$ is an *ergodic dynamical system*.

In this thesis we always work in a compact set $X \subset \mathbb{R}$, which typically is an invariant set. We define the function $\varphi : \mathbb{N}_0 \times X \times \Omega \rightarrow X$ given by

$$\varphi_\omega^n(x) = h_{\omega_{n-1}} \circ \cdots \circ h_{\omega_0}(x), \quad (1.5)$$

where $h_{\omega_i}(x) = T(x) + \sigma\omega_i$. For each $n \in \mathbb{N}$, the function $\varphi^n : X \times \Omega \rightarrow X$ is $\mathcal{B}(X) \otimes \mathcal{B}(\Omega), \mathcal{B}(X)$ -measurable, and for each $\omega \in \Omega$ the map $\varphi_\omega^n : \mathbb{R} \rightarrow \mathbb{R}$ is smooth. Moreover, the map φ satisfies the *cocycle property*: for all $n, m \in \mathbb{N}_0$, $\omega \in \Omega$ and $x \in X$,

$$\varphi_\omega^0(x) = x, \quad \varphi_\omega^{n+m}(x) = \varphi_{\theta^n \omega}^n \circ \varphi_\omega^m(x). \quad (1.6)$$

We say then, that φ is a *smooth random dynamical system* over $(\Omega, \mathcal{B}(\Omega), \mathbb{P}, \theta)$ [3].

The random dynamical system φ induces a (deterministic) dynamical system in the extended space $X \times \Omega$. This map $\Phi : X \times \Omega \rightarrow X \times \Omega$ is defined as $\Phi(x, \omega) = (\varphi_\omega(x), \theta\omega)$. The map Φ is called a *skew product flow*. We refer to such dynamical system as the pair $\Phi = (\varphi, \theta)$.

Stationary measures for random diffeomorphisms

In the first part of this thesis, we study *stationary measures* of (1.1). A stationary measure μ is a probability distribution which does not change with the evolution of the system. From a parallel point of view, stationary measures play the role of an equilibrium for the collective behaviour of the system (1.1).

If $\mathcal{M}_1(M)$ is the set of Borel probability measures on M , we consider the *Markov operator* $\mathcal{P} : \mathcal{M}_1(M) \rightarrow \mathcal{M}_1(M)$ (see Appendix C), which indicates the evolution of any initial distribution $\vartheta \in \mathcal{M}_1(M)$ with the system (1.1). Thus, a stationary measure μ is a fixed point, or an equilibrium measure, for $\mathcal{P}$.

There is a strong relationship between the minimal invariant sets of the system and its stationary measures which can be stated in the following theorem

Theorem A *Let M be a minimal invariant set for the system (1.1). Then there exists a unique stationary measure μ supported on $(M, \mathcal{B}(M))$, such that for any initial distribution ϑ , we have that $\mathcal{P}^n \vartheta \rightarrow \mu$ exponentially fast with respect to the total variation distance for probability measures (see Appendix A).*

This means that the evolution of the distributions along with the system (1.1) on a minimal invariant set necessarily stabilize asymptotically towards a stationary measure. In words of the Nobel laureate Ilya Prigogine, this mechanism is an example of how the system exhibits the *arrow of time*¹.

Because of this connection between minimal invariant sets and stationary measures, it is natural to search for features which reflect the bifurcation the system (1.1) is undergoing through in these special type of measures. One of those features is the behaviour of its tail, or equivalently the behaviour of its density, known as *stationary density*, near the boundaries of its support.

For this, consider the largest and smallest values that the random map $h(x; \omega) = T(x) + \sigma\omega$ can achieve. In other words, let $h_+(x) = T(x) + \sigma$, and $h_-(x) = T(x) - \sigma$ be the so-called *upper and lower extremal maps*, respectively. Minimal invariant sets $M = [x_-, x_+]$ are determined by the fixed points of h_+ , and h_- , where in fact x_+ and x_- are fixed points of h_+ and h_- , respectively. Hence their hyperbolicity and degeneracy play a fundamental role in the shape of any stationary density, as given in one of the main results of this thesis which may be stated as follows.

Theorem B *Let $M = [x_-, x_+]$ be a minimal invariant set of (1.1). If the stationary measure μ on $(M, \mathcal{B}(M))$ is such that $\mu = \phi d\text{Leb}$, then*

1. (Hyperbolic case) *If $T'(x_+) < 1$, then*

$$\phi(x) \sim (x_+ - x)^{\ln(x_+ - x)}, \quad \text{as } x \rightarrow x_+.$$

2. (Nonhyperbolic case) *If $T'(x_+) = 1$, and $r \geq 2$ is the first integer such that $T^{(r)}(x_+) \neq 0$, then*

$$\phi(x) \sim (x_+ - x)^{\frac{1}{(x_+ - x)^{r-1}}}, \quad \text{as } x \rightarrow x_+.$$

Not only it is an interesting problem in itself to estimate the tails of stationary measures [23]. In applications it could be relevant to know how likely the system visits a small neighbourhood of the boundaries. For example, in [42] an algorithm for anticipating bifurcations in random dynamical systems was proposed, which rely on having sufficient data precisely near the boundary points $\{x_-, x_+\}$ of the minimal invariant set. As it turns out due to the theorem above, it is extremely unlikely that the orbits of (1.1) visit a small neighbourhood of x_- or x_+ , so it takes an enormous amount of time to generate meaningful data for this purpose.

¹For an introduction on Prigogine's work on the topic, see [49].

Quasi-stationary measures for random diffeomorphisms

Recall that when increasing the noise strength σ , the system (1.1) may exhibit a bifurcation of its minimal invariant sets at a critical value $\sigma^* > 0$. As in Figure 1.4, assume that at $\sigma^* > 0$ we have two minimal invariant sets, which “merge” for $\sigma > \sigma^*$.

Observe that for any value $\sigma > \sigma^*$, the orbits may explore a wider region than the minimal invariant set they were confined in earlier for $\sigma = \sigma^*$. Let $[x_-(\sigma), x_+(\sigma)]$ be this large (unique) minimal invariant set of (1.1) for $\sigma > \sigma^*$, and $M^{\sigma^*} = [x_-(\sigma^*), x^*]$ the minimal invariant set for $\sigma = \sigma^*$ which is in the lower branch in Figure 1.4.

We focus our attention in the set $M^\sigma = [x_-(\sigma), x^*]$, which is a portion of the unique minimal invariant set cut off precisely at the point where the bifurcation occurred. For σ above but close to σ^* , the orbits escape from M^σ almost surely, but they do so in an extremely large amount of time. It is because of this eventual escape that no stationary measure exists on $(M^\sigma, \mathcal{B}(M^\sigma))$.

Instead, we look for another type of measure which plays a similar role, conditioning the system to remain in M^σ : the so-called *quasi-stationary measures*. These special type of measures are fixed points of the *conditional Markov operator*, $\bar{\mathcal{P}} : \mathcal{M}_1(M^\sigma) \rightarrow \mathcal{M}_1(M^\sigma)$, defined as

$$\bar{\mathcal{P}}\vartheta(A) = \frac{\mathcal{P}\vartheta(A \cap M^\sigma)}{\mathcal{P}\vartheta(M^\sigma)},$$

for all $A \in \mathcal{B}(M^\sigma)$ where $\mathcal{P}$ is the usual Markov operator as above, except that $\mathcal{P}\vartheta$ is no longer a probability measure on $(M^\sigma, \mathcal{B}(M^\sigma))$.

Quasi-stationary measures date back to [56], but it has not been until recent years that strong results on approximation of quasi-stationary measures were given [14, 13]. For an excellent introduction to the topic, see [16].

The following theorem is analogous to Theorem A, but conditioned to remain in M^σ .

Theorem C *Let M^σ be as above. Then, if $x_-(\sigma^*)$ is hyperbolic, there exists a unique quasi-stationary measure q supported on $(M^\sigma, \mathcal{B}(M^\sigma))$, such that for any initial distribution $\vartheta \in \mathcal{M}_1(M^\sigma)$, we have that $\bar{\mathcal{P}}^n \vartheta \rightarrow q$ exponentially fast with respect to the total variation distance for probability measures.*

The exponential convergence of $\overline{\mathcal{P}}^n \vartheta$ towards q allows the construction of the so-called *quasi-ergodic measures*. These measures have recently been used to define a conditional version of *Lyapunov exponents* [26], which provide an insight of the local properties of random dynamical systems.

Let $\tau(x, \omega)$ be the escape time from M^σ for the cocycle φ , that is

$$\tau(x, \omega) = \min\{n \in \mathbb{N} \mid \varphi_\omega^n(x) \notin M^\sigma\}.$$

The escape time τ is geometrically distributed when the initial distribution is precisely the quasi-stationary measure q . In other words, $\mathbb{P}_q[n < \tau] = \Lambda_\sigma^n$, for some $\Lambda_\sigma \in (0, 1)$. Hence, its mean value, $\mathbb{E}_q[\tau]$, satisfies

$$\mathbb{E}_q[\tau] = \frac{1}{1 - \Lambda_\sigma}.$$

In particular, when the bifurcation of the minimal invariant sets is due to a nondegenerate saddle-node bifurcation for the upper extremal map $h_+(x, \sigma) = T(x) + \sigma$, we may state the behaviour of Λ_σ as $\sigma \rightarrow \sigma^*$ in the following way.

Theorem D *Let $M^\sigma = [x_-(\sigma), x^*]$ be as above, where x^* is a nondegenerate nonhyperbolic fixed point for $h_+(x, \sigma^*)$. Then,*

$$1 - \Lambda_\sigma \sim (\sigma - \sigma^*)^{\frac{1}{\sqrt{\sigma - \sigma^*}}}.$$

Structure of the thesis

In Chapter 2 we use functional operator techniques in order to study the evolution of distributions and densities along with the system (1.1). For the basic background, the reader can consult Appendix C, where we made a self-contained summary of well known tools which we employ in this thesis: the Markov ($\mathcal{P}$), the transfer (L) and the stochastic Koopman operators (P). The main result in this chapter guarantees the existence and uniqueness of the stationary measure μ supported on each minimal invariant set of the system (1.1). Moreover, the evolution of distributions $\mathcal{P}^n \vartheta$ converge exponentially fast to μ with respect to the total variation norm, or equivalently, the evolution of densities $L^n g$ converge exponentially fast to the unique stationary density in L^1 . We prove as a direct consequence that the system exhibits exponential decay of correlations. Furthermore, we extend the exponential convergence of densities to the uniform convergence topology.

In Chapter 3, we describe asymptotically the shape of the stationary density supported on a minimal invariant set $[x_-, x_+]$ as $x \rightarrow x_{\pm}$. The proof uses an embedding theorem of maps into flows, which helps to derive asymptotic results for deterministic hitting-times. We confirm our results with some numerical explorations.

In Chapter 4, we introduce the conditional versions of the Markov ($\bar{\mathcal{P}}$), transfer ($\bar{L}$), and Koopman ($\bar{P}$) operators, and deduce their properties needed further in the chapter. The main goal is to prove that, near a bifurcating minimal invariant set, there is a unique quasi-stationary measure, which can be approximated exponentially fast with $\bar{\mathcal{P}}^n \vartheta$ for any initial distribution ϑ in total variation norm.

In Chapter 5, we prove the main result of this thesis, by giving the asymptotic order of Λ_{σ} as $\sigma \rightarrow \sigma^*$ near a bifurcation of a minimal invariant set due to a saddle-node bifurcation of the upper extremal map. To our knowledge, this is the first result of its kind, which highlights the complexity of random diffeomorphisms near bifurcation points.

In Chapter 6, we invite the reader to the discussion on the use of random dynamical systems for studying critical transitions of complex systems, and their potential early-warning signals. We also present some possible routes of research for future work on different mechanisms for bifurcations in low-dimensional random maps.

Chapter 2

Stationary measures for random diffeomorphisms

In this chapter we consider the random map $h : M \times \Delta \rightarrow M$, given by $h(x, \omega_0) = T(x) + \sigma\omega_0$, where $\Delta = [-1, 1]$, the deterministic map T is such that $T'(x) > 0$ for all $x \in M$, and $M = [x_-, x_+]$ is a minimal invariant interval for h (see Appendix B). The random dynamical system is defined by the difference equation

$$x_{n+1} = h(x_n, \omega_n), \tag{2.1}$$

where we consider $\omega = (\omega_n)_{n=0}^\infty$ to be an i.i.d. sequence of random variables distributed in Δ according to the probability measure $d\nu = p \, d\text{Leb}$. We further assume that $p : \Delta \rightarrow \mathbb{R}_+$ is a Lipschitz-continuous density, which is positive in $(-1, 1)$. This setting, for example, covers the case of uniformly distributed noisy perturbations, where $p(\omega_0) = 1/2$ for all $\omega_0 \in \Delta$.

The main objective of this chapter is to prove there exists a unique stationary measure for (2.1) supported on M , and that transition probabilities converge uniformly (with respect to the initial distribution) exponentially fast to the stationary measure in the total variation norm.

The uniqueness of the stationary measure implies this measure is ergodic. This type of measures is one of the most important tools employed for studying the statistical behaviour of random dynamical systems, as given by Birkhoff's ergodic theorem. By virtue of this theorem, time averages of observables along orbits coincide with the space average with respect to an ergodic measure in a full-measure set (for a comprehensive introduction to the topic see [37]).

2.1 Main result and structure of the chapter

The existence and uniqueness of stationary measures (and hence ergodic) for random diffeomorphisms supported on minimal invariant sets M have already been proved in [1, 59]. However, the monotonicity property imposed on T provides with sufficient conditions to prove that for each $\sigma > 0$ the transition probabilities $(\mathbb{P}_x^n)_{n=0}^\infty$, converge uniformly and exponentially fast towards the stationary measure, where $\mathbb{P}_x^n(A) = \int_\Delta \mathbb{1}_A(h(x, \omega_0)) d\nu(\omega_0)$ (cf. Appendix C).

The main objective of this chapter is to prove this exponential convergence holds, as stated in our main theorem below. Recall that (2.1) induces a set valued map $F(x) = T(x) + \sigma\Delta$, which is extended to the whole family of compact subsets of M (cf. Appendix B). We also consider the topology in the space of probability measures induced by the total variation norm, $\|\cdot\|_{TV}$ (see Appendix A).

Theorem 2.1 *For each $\sigma > 0$, there exists a unique stationary measure μ on $(M, \mathcal{B}(M))$ for the system (2.1). Furthermore, for each $n_0 \in \mathbb{N}$ such that $F^{n_0}(x_-)$ intersects $F^{n_0}(x_+)$ in a nondegenerate interval I , and for each closed interval $B \subset I$, bounded away from ∂I , there exists $D > 0$ such that*

$$\|\mathbb{P}_x^n - \mu\|_{TV} \leq (1 - D \cdot \text{Leb}(B))^{\frac{n}{n_0}}, \quad (2.2)$$

for all $n \in \mathbb{N}$ and $x \in M$.

Because of the uniform convergence (2.2) the cocycle is said to be *uniformly ergodic* [46].

We give a proof of Theorem 2.1 in Section 2.2. The proof is based on the *Doebelin's minorization theorem* (see [46]). It is because of this last theorem that the rate of convergence $(1 - D \cdot \text{Leb}(B))^{1/n_0}$ in (2.2) can be given explicitly.

A direct consequence of Theorem 2.1 is that the sequence $(L^n g)_{n=0}^\infty$ converges exponentially fast to the unique stationary density ϕ in L^1 (see Remark A.2 in Appendix A), where L is the transfer operator, and g is any density on M . In Section 2.3, we use this fact to prove that the so-called *annealed correlations* decay exponentially fast, and we prove this fact in Theorem 2.7. This property has already been shown in [59] by means of the spectral gap of the transfer operators. However, the method here presented has the advantage of giving a precise estimate on the rate of convergence.

Finally, in Section 2.4, we extend the exponential convergence of $(L^n g)$ to the uniform topology,

for any density g on M . The result follows essentially from the fact that the sequence $(L^n g)_{n=2}^\infty$ is uniformly Lipschitz-continuous, for each density g .

2.2 Exponential convergence to the stationary measure

We denote by $\mathcal{M}_1(M)$ the set of probability measures on $(M, \mathcal{B}(M))$, and $\mathcal{D}(M)$ the space of densities on M , that is functions $g : M \rightarrow \mathbb{R}_+$ such that $\|g\|_{L^1} = 1$. Recall that

$$\text{supp } g = \overline{\{x \in \mathbb{R} \mid g(x) > 0\}}.$$

In order to prove Theorem 2.1, we use the following result.

Theorem 2.2 (Doeblin's minorization theorem [46, Theorem 16.2.4]) *Assume the transition probabilities $\mathbb{P}_x^n$, with $x \in M$ and $n \in \mathbb{N}$, of system (2.1) satisfy*

Condition (D): *there exist $n_0 \in \mathbb{N}$ and a nonzero measure ρ on $(M, \mathcal{B}(M))$ such that for all $x \in M$ and $A \in \mathcal{B}(M)$*

$$\mathbb{P}_x^{n_0}(A) \geq \rho(A). \quad (2.3)$$

Then there exists a unique probability measure μ on $(M, \mathcal{B}(M))$ such that for all $n \in \mathbb{N}$

$$\|\mathbb{P}_x^n - \mu\|_{TV} \leq (1 - \rho(M))^{\frac{n}{n_0}}. \quad (2.4)$$

Moreover, μ is the unique stationary measure on $(M, \mathcal{B}(M))$.

Condition (D) is also known as *Doeblin's minorization condition*. Since (2.3) is uniform in x , assuming Condition (D) is fulfilled then for all $\vartheta \in \mathcal{M}_1(M)$ we have that $\mathbb{P}_\vartheta^{n_0}(A) \geq \rho(A)$, for all $A \in \mathcal{B}(M)$. Also, if (2.3) is fulfilled for some $n_0 \in \mathbb{N}$, then it is satisfied for all $n \geq n_0$. This follows from the properties of the transition probabilities, since if $n \geq n_0$, then

$$\mathbb{P}_x^n(A) = \int \mathbb{P}_y^{n_0}(A) d\mathbb{P}_x^{n-n_0}(y) \geq \rho(A) \cdot \mathbb{P}_x^{n-n_0}(M) = \rho(A)$$

Remark 2.3 *Because of the uniform bound in (2.4) with respect to x , we have that whenever Condition (D) is satisfied, then*

$$\|\mathbb{P}_\vartheta^n - \mu\|_{TV} \leq (1 - \rho(M))^{\frac{n}{n_0}} \quad (2.5)$$

for any initial probability measure ϑ on $(M, \mathcal{B}(M))$. Indeed, from (2.4) we have that for all $x \in M$ and $A \in \mathcal{B}(M)$

$$\mu(A) - (1 - \vartheta(M))^{n/n_0} \leq \mathbb{P}_x^n(A) \leq \mu(A) + (1 - \vartheta(M))^{n/n_0}.$$

Integrating over M with respect to ϑ gives (2.5).

We now proceed to prove Theorem 2.1. Our proof follows some ideas from [9].

Proof of Theorem 2.1. From Appendix C, the transition probability $\mathbb{P}_x^n$ is absolutely continuous with respect to Leb, for all $x \in M$ and $n \in \mathbb{N}$, with continuous density $k_n : M \times M \rightarrow \mathbb{R}_+$ for all $n \geq 2$ (see Proposition C.13).

Also $\text{supp } k_n(x, \cdot) = F^n(x)$, where the set-valued dynamics is given by $F^n(x) = [h_-^n(x), h_+^n(x)]$. Since $h_+^n(x_-) \rightarrow x_+$ and $h_-^n(x_+) \rightarrow x_-$ as $n \rightarrow \infty$, there is $n_0 \geq 2$ such that $h_-^{n_0}(x_+) < h_+^{n_0}(x_-)$, see Figure 2.1.

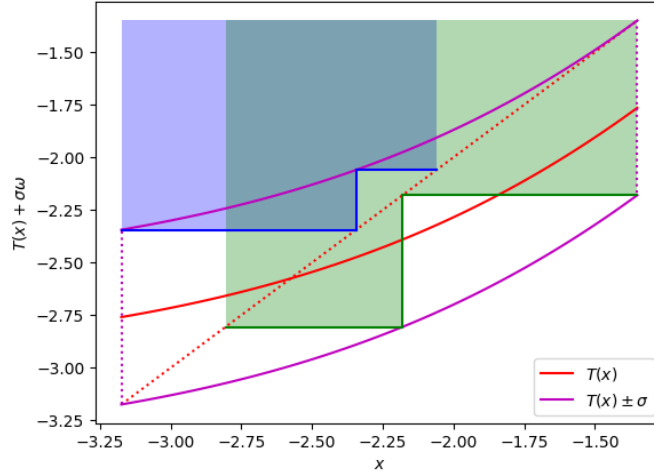


Figure 2.1: In blue we represent the set $F^{n_0}(x_-)$, while in green we portray $F^{n_0}(x_+)$. Notice that they indeed intersect each other in a nondegenerate interval (pictured in a darker colour). Here, the steps it takes for this to happen is precisely $n_0 = 2$, although any $n \geq n_0$ would do the job.

Let $B \subset I = [h_-^{n_0}(x_+), h_+^{n_0}(x_-)]$ be any closed interval bounded away from $\partial I = \{h_-^{n_0}(x_+), h_+^{n_0}(x_-)\}$. By monotonicity of h_+ and h_- , for any $x \in M$ we have $h_+^{n_0}(x) > h_+^{n_0}(x_-)$ and $h_-^{n_0}(x) < h_-^{n_0}(x_+)$.

This implies that $B \subset \bigcap_{x \in M} \text{int } F^{n_0}(x)$. Let

$$D := \inf_{x \in M, y \in B} k_{n_0}(x, y) > 0,$$

and consider the measure $\rho(A) = D \cdot \text{Leb}(A \cap B)$. Therefore,

$$\mathbb{P}_x^{n_0}(A) = \int_A k_{n_0}(x, y) dy \geq \int_{A \cap B} k_{n_0}(x, y) dy \geq \rho(A).$$

Condition (D) is then satisfied and the statement follows from Theorem 2.2. $\square$

Remark 2.4 *Observe that the closer B is to I its length $\text{Leb}(B)$ becomes larger, but the constant D is closer to 0. On the other hand the optimization of D would be obtained by choosing B very close to its maximum, but this would imply that $\text{Leb}(B)$ is very small. An optimal rate of convergence would depend on the specific map T and the density p .*

Let $\phi \in \mathcal{D}(M)$ be the density of the unique stationary measure μ , and consider the transfer operator $L : L^1(M) \rightarrow L^1(M)$ (cf. Definition C.9). The function ϕ is the unique fixed point of L in $\mathcal{D}(M)$, so it is the unique stationary density (see Proposition C.12). As a direct consequence of Theorem 2.1, starting with an arbitrary density g on M the sequence $(L^n g)_{n=0}^\infty$ converges exponentially fast to the unique stationary density in L^1 as proved in the following.

Corollary 2.5 *For all $n \in \mathbb{N}$ and every density $g \in \mathcal{D}(M)$, we have*

$$\|L^n g - \phi\|_{L^1} \leq 2(1 - \rho(M))^{\frac{n}{n_0}}, \quad (2.6)$$

where n_0 and ρ are given in Theorem 2.1, and ϕ is the unique stationary density.

Proof. Let ϑ be the probability measure such that $d\vartheta = g d\text{Leb}$. Therefore, $\mathbb{P}_\vartheta^n$ has density $L^n g$. From Proposition A.1 and (2.2) we have

$$\|L^n \psi - \phi\|_{L^1} = 2\|\mathbb{P}_\vartheta^n - \mu\|_{TV} \leq 2(1 - \vartheta(M))^{n/n_0}.$$

$\square$

2.3 Exponential decay of correlations

In this section we denote the Lebesgue space $L_\mu^1(M) \equiv L^1(M, \mathcal{B}(M), \mu)$, where μ is the unique stationary measure. Recall that $L^1(M) \equiv L^1(M, \mathcal{B}(M), \text{Leb})$ and $L^\infty(M) \equiv L^\infty(M, \mathcal{B}(M), \text{Leb})$. Consider the stochastic Koopman operator $P : L^\infty(M) \rightarrow L^\infty(M)$ (see Definition C.7).

We make use of Corollary 2.5 in order to prove that system (2.1) exhibits exponential decay of annealed correlations. Let ϕ be the unique stationary density. The following definitions are taken from [2].

Definition 2.6 (Annealed correlations) *The annealed correlation of non-zero functions $f \in L^\infty(M)$ against $g \in L_\mu^1(M)$ with respect to $\mu \times \mathbb{P}$ is*

$$\text{Cor}_{\mu \times \mathbb{P}}(f, g, n) = \frac{1}{\|f\|_\infty \|g\|_{L_\mu^1}} \left| \int (P^n f \cdot g) d\mu - \int f d\mu \cdot \int g d\mu \right|. \quad (2.7)$$

We say the system (2.1) exhibits *annealed decay of correlations* for observables in $L^\infty(M)$ against observables in $L_\mu^1(M)$, with respect to the measure $\mu \times \mathbb{P}$, if for every $f \in L^\infty(M)$ and every $g \in L_\mu^1(M)$ we have $\text{Cor}_{\mu \times \mathbb{P}}(f, g, n) \rightarrow 0$ as $n \rightarrow \infty$.

The proof of the exponential decay of correlations is taken almost identically from that of Theorem A and B in [2], although the setting is somewhat different. We include the proof for the reader's convenience.

Theorem 2.7 (Exponential annealed decay of correlations) *For all $f \in L^\infty(M)$, $g \in L_\mu^1(M)$, and $n \geq 2$, we have*

$$\text{Cor}_{\mu \times \mathbb{P}}(f, g, n) \leq 2(1 - D \cdot \text{Leb}(M))^{\frac{n-1}{n_0}}, \quad (2.8)$$

holds, where the constant $D > 0$ and $n_0 \in \mathbb{N}$ are given in Theorem 2.1.

Proof. Let $\kappa = 1 - D \cdot \text{Leb}(M)$, so that from Corollary 2.5

$$\|L^n g - \phi\|_{L^1} \leq 2\kappa^{n/n_0},$$

for all $n \in \mathbb{N}$. Recall also that for $n \in \mathbb{N}$,

$$P^n f(x) = \int k_n(x, y) f(y) dy,$$

where $k_n(x, y)$ is the density of $\mathbb{P}_x^n$ (see Proposition C.8). Recall also from Appendix C that $L(k_n(x, \cdot))(y) = k_{n+1}(x, y)$ for all $n \in \mathbb{N}$. Then,

$$\begin{aligned} \left| \int P^n f(x) \cdot g(x) d\mu(x) - \int f(x) d\mu(x) \int g(x) d\mu(x) \right| &= \left| \int \left(\int f(y) [k_n(x, y) - \phi(y)] dy \right) g(x) d\mu(x) \right| \\ &\leq \|f\|_\infty \int \left(\int |L^{n-1}(k(x, \cdot))(y) - \phi(y)| dy \right) |g(x)| d\mu(x) \leq 2\|f\|_\infty \|g\|_{L_\mu^1} \cdot \kappa^{(n-1)/n_0}. \end{aligned}$$

□

2.4 Convergence of densities in uniform topology

In this section we extend Corollary 2.5 to the uniform topology, this is we prove $(L^n g)_{n=0}^\infty$ converges to the stationary density ϕ exponentially fast with respect to $\|\cdot\|_\infty$, for any $g \in L^1(M)$.

From (C.16) in Appendix C, we know that the explicit formula for the transfer operator is given by

$$Lg(x) = \frac{1}{\sigma} \int_{h_+^{-1}(x) \vee x_-}^{h_-^{-1}(x) \wedge x_+} p\left(\frac{x - T(y)}{\sigma}\right) g(y) dy, \quad (2.9)$$

where p is the Lipschitz-continuous density of the noise distribution ν . In order to determine the upper and lower limits of the integral in the expression above, we must find x_1 and x_2 such that $h_-^{-1}(x_1) = x_+$ and $h_+^{-1}(x_2) = x_-$. In other words,

$$x_1 = h_-(x_+) = T(x_+) - \sigma = x_+ - 2\sigma,$$

$$x_2 = h_+(x_-) = T(x_-) + \sigma = x_- + 2\sigma.$$

Observe that there are essentially two cases:

Case 1: $x_- < x_2 \leq x_1 < x_+$

$$Lg(x) = \begin{cases} \frac{1}{\sigma} \int_{x_-}^{h_-^{-1}(x)} p\left(\frac{x - T(y)}{\sigma}\right) g(y) dy, & x \in [x_-, x_2] \\ \frac{1}{\sigma} \int_{h_+^{-1}(x)}^{h_-^{-1}(x)} p\left(\frac{x - T(y)}{\sigma}\right) g(y) dy, & x \in (x_2, x_1) \\ \frac{1}{\sigma} \int_{h_+^{-1}(x)}^{x_+} p\left(\frac{x - T(y)}{\sigma}\right) g(y) dy, & x \in [x_1, x_+] \end{cases} \quad (2.10)$$

Case 2: $x_- < x_1 \leq x_2 < x_+$

$$Lg(x) = \begin{cases} \frac{1}{\sigma} \int_{x_-}^{h_-^{-1}(x)} p\left(\frac{x-T(y)}{\sigma}\right) g(y) dy, & x \in [x_-, x_1] \\ \frac{1}{\sigma} \int_{x_-}^{x_+} p\left(\frac{x-T(y)}{\sigma}\right) g(y) dy, & x \in (x_1, x_2) \\ \frac{1}{\sigma} \int_{h_+^{-1}(x)}^{x_+} p\left(\frac{x-T(y)}{\sigma}\right) g(y) dy, & x \in [x_2, x_+] \end{cases} \quad (2.11)$$

We use these formulas in the next proposition to prove some regularization properties of the transfer operator. Consider the Lipschitz constant of a real-valued Lipschitz-continuous function on a metric space (X, d) defined as

$$\text{Lip}(f) = \sup_{x, y \in X; x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}.$$

Proposition 2.8 *For all $g \in L^\infty(M)$, we have Lg is Lipschitz-continuous. Moreover, there exist $A, B > 0$ such that*

$$\text{Lip}(Lg) \leq A\|g\|_\infty + B\|g\|_{L^1}. \quad (2.12)$$

Proof. We prove this proposition only for Case 1 above (see equation (2.10)), as the proof for Case 2 is done in a similar way. We do several intermediate calculations in order to prove (2.12).

Claim 1: *there exist $K_1, K_2 > 0$ such that for all $x, z \in [x_-, x_2]$,*

$$|Lg(x) - Lg(z)| \leq (K_1\|g\|_\infty + K_2\|g\|_{L^1})|x - z|$$

Indeed, assume without loss of generality that $x \leq z$. By monotonicity of h_-^{-1} it follows that

$$Lg(x) - Lg(z) = \frac{1}{\sigma} \int_{x_-}^{h_-^{-1}(x)} \left[p\left(\frac{x-T(y)}{\sigma}\right) - p\left(\frac{z-T(y)}{\sigma}\right) \right] g(y) dy - \frac{1}{\sigma} \int_{x_-}^{h_-^{-1}(z)} p\left(\frac{z-T(y)}{\sigma}\right) g(y) dy.$$

Hence,

$$|Lg(x) - Lg(z)| \leq \frac{\text{Lip}(p)\|g\|_{L^1}}{\sigma^2} \cdot |x - z| + \frac{\|p\|_\infty\|g\|_\infty}{\sigma} \cdot |h_-^{-1}(x) - h_-^{-1}(z)|.$$

Claim 1 follows by taking $K_1 = \|p\|_\infty (\sigma \cdot \inf_{x \in M} h'_-(h_-^{-1}(x)))^{-1}$ and $K_2 = \text{Lip}(p) \cdot \sigma^{-2}$.

Claim 2: *there exist $K_3, K_4 > 0$ such that for all $x, z \in [x_2, x_1]$,*

$$|Lg(x) - Lg(z)| \leq (K_3 \|g\|_\infty + K_4 \|g\|_{L^1}) |x - z|$$

In order to prove this, assume again without loss of generality that $x \leq z$. Since

$$\begin{aligned} Lg(x) - Lg(z) &= \\ &= \frac{1}{\sigma} \int_{h_+^{-1}(x)}^{h_+^{-1}(z)} p\left(\frac{x-T(y)}{\sigma}\right) g(y) dy + \frac{1}{\sigma} \int_{h_+^{-1}(z)}^{h_-^{-1}(x)} \left[p\left(\frac{x-T(y)}{\sigma}\right) - p\left(\frac{z-T(y)}{\sigma}\right) \right] g(y) dy + \\ &\quad + \frac{1}{\sigma} \int_{h_-^{-1}(x)}^{h_-^{-1}(z)} p\left(\frac{z-T(y)}{\sigma}\right) g(y) dy, \end{aligned}$$

it follows that

$$|Lg(x) - Lg(z)| \leq \frac{\|p\|_\infty \|g\|_\infty}{\sigma} \cdot |h_+^{-1}(x) - h_+^{-1}(z)| + \frac{\text{Lip}(p) \|g\|_{L^1}}{\sigma^2} \cdot |x - z| + \frac{\|p\|_\infty \|g\|_\infty}{\sigma} \cdot |h_-^{-1}(x) - h_-^{-1}(z)|.$$

The claim follows by taking $K_3 = \|p\|_\infty [(\sigma \cdot \inf_{x \in M} h'_-(h_-^{-1}(x)))^{-1} + (\sigma \cdot \inf_{x \in M} h'_+(h_+^{-1}(x)))^{-1}]$ and $K_4 = \text{Lip}(p) \cdot \sigma^{-2}$

Claim 3: *there exist $K_5, K_6 > 0$ such that for all $x, z \in [x_1, x_+]$*

$$|Lg(x) - Lg(z)| \leq (K_5 \|g\|_\infty + K_6 \|g\|_{L^1}) |x - z|.$$

Indeed, this claim follows similarly as Claim 1.

Claim 4: *inequality (2.12) holds for some $A, B > 0$.*

Let $A = 3 \max\{K_1, K_3, K_5\}$, and $B = 3 \max\{K_2, K_4, K_6\}$. Assume for instance that $x \in [x_-, x_2]$ and $z \in [x_1, x_+]$. Then,

$$\begin{aligned} |Lg(x) - Lg(z)| &\leq |Lg(x) - Lg(x_2)| + |Lg(x_2) - Lg(x_1)| + |Lg(x_1) - Lg(z)| \leq \\ &\leq (K_1 \|g\|_\infty + K_2 \|g\|_{L^1})(x_2 - x) + (K_3 \|g\|_\infty + K_4 \|g\|_{L^1})(x_1 - x_2) + (K_5 \|g\|_\infty + K_6 \|g\|_{L^1})(z - x_1) \\ &\leq (A \|g\|_\infty + B \|g\|_{L^1}) |x - z| \end{aligned}$$

The same follows similarly for $x \in [x_-, x_2]$ and $z \in [x_2, x_1]$, or for the case $x \in [x_2, x_1]$ and $z \in [x_1, x_+]$. $\square$

Let us proceed to improve Corollary 2.5 by showing that in fact we have exponential convergence with respect to the supremum norm $\|\cdot\|_\infty$. Since $Lg \in C^0(M)$ for any $g \in L^1(M)$, and Lg is Lipschitz-continuous whenever $g \in C^0(M)$, without loss of generality we can restrict ourselves to consider g to be a Lipschitz-continuous density on M . We need the following definition.

Definition 2.9 (Uniform Lipschitz-continuous sequence) *We say that a sequence $(g_n)_{n=0}^\infty$ of real-valued functions is uniformly Lipschitz-continuous on $A \subset \mathbb{R}$ if there exists $K > 0$ such that*

$$|g_n(x) - g_n(y)| \leq K|x - y|, \quad (2.13)$$

for all $x, y \in A$ and $n \in \mathbb{N}$.

In particular, a uniformly Lipschitz-continuous sequence $(g_n)_{n=0}^\infty$ in $A \subset \mathbb{R}$ is *uniformly equicontinuous*.

Furthermore, if a uniformly Lipschitz-continuous sequence $(g_n)_{n=0}^\infty$ converges to g in L^1 then g_n converges uniformly [29]. This result gives no information about the rate of convergence with respect to $\|\cdot\|_\infty$, for which we prove the following theorem by following closely the arguments in [29].

Lemma 2.10 *Let $(g_n)_{n=1}^\infty$ be a uniformly Lipschitz-continuous sequence of real-valued functions on $[a, b]$ such that g_n converges exponentially fast to g in L^1 . Then g_n converges exponentially fast to g with respect to $\|\cdot\|_\infty$.*

Proof. Let $C > 0$ and $\kappa \in (0, 1)$ such that $\|g_n - g\|_{L^1} \leq C\kappa^n$, for all $n \in \mathbb{N}$. Thus, for each $n, m \in \mathbb{N}$,

$$\|g_n - g_m\|_{L^1} \leq C(\kappa^n + \kappa^m).$$

We claim that for each $x \in [a, b]$ and $n, m \in \mathbb{N}$ there is $y \equiv y_{n,m}^x \in [a, b]$ which satisfies that $|x - y| \leq \sqrt{C}(\kappa^n + \kappa^m)^{1/2}$, and $|g_n(y) - g_m(y)| \leq \sqrt{C}(\kappa^n + \kappa^m)^{1/2}$. Indeed, assume on the contrary that there is $x_0 \in [a, b]$ and $n, m \in \mathbb{N}$ such that for every $y \in B_{r_0}(x_0) \cap [a, b] \equiv B$, where $r_0 = \sqrt{C}(\kappa^n + \kappa^m)^{1/2}$,

$$|g_n(y) - g_m(y)| > r_0.$$

Therefore,

$$C(\kappa^n + \kappa^m) \geq \|g_n - g_m\|_{L^1} \geq \int_B |g_n(y) - g_m(y)| dy > r_0 \cdot \text{Leb}(B) \geq C(\kappa^n + \kappa^m),$$

which is a contradiction and the claim follows.

Finally, for each $x \in M$ and $n, m \in \mathbb{N}$, and by considering the uniform Lipschitz-continuity of $(g_n)_{n=1}^\infty$ and the point $y \equiv y_{n,m}^x$ obtained above, we get

$$\begin{aligned} |g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(y)| + |g_n(y) - g_m(y)| + |g_m(y) - g_m(x)| \\ &\leq K|x - y| + \sqrt{C}(\kappa^n + \kappa^m)^{1/2} + K|x - y| \\ &\leq \sqrt{C}(2K + 1)(\kappa^n + \kappa^m)^{1/2} \leq \sqrt{C}(2K + 1)(\kappa^{n/2} + \kappa^{m/2}). \end{aligned} \tag{2.14}$$

Since (2.14) is valid for all $x \in [a, b]$, $(g_n)_{n=0}^\infty$ is a uniformly Cauchy sequence, implying that g_n converges uniformly to g . Furthermore, given $\epsilon > 0$, for all m large enough $\|g_m - g\|_\infty < \epsilon$. Hence, for all $x \in [a, b]$ we obtain

$$|g_n(x) - g(x)| \leq |g_n(x) - g_m(x)| + |g_m(x) - g(x)| \leq \sqrt{C}(2K + 1)(\kappa^{n/2} + \kappa^{m/2}) + \epsilon$$

By taking limits when $m \rightarrow \infty$, and since ϵ was arbitrarily small,

$$\|g_n - g\|_\infty \leq \sqrt{C}(2K + 1)\kappa^{n/2}.$$

□

The exponential converge of $L^n g$ for any Lipschitz-continuous density g supported on M follows directly from Lemma 2.10 as seen in the following statement.

Theorem 2.11 *For each Lipschitz-continuous density $g : M \rightarrow \mathbb{R}_+$, supported on a minimal invariant set $M = [x_-, x_+]$, the sequence $(L^n g)_{n=0}^\infty$ converges exponentially fast to the unique stationary density ϕ with respect to $\|\cdot\|_\infty$.*

Proof. Since $L^n g \rightarrow \phi$ exponentially fast in L^1 as given in Corollary 2.5, because of Lemma 2.10 it is sufficient to prove $(L^n g)_{n=1}^\infty$ is uniformly Lipschitz.

Claim: *the transfer operator L seen as $L : L^1(M) \rightarrow C^0(M)$ is bounded.*

Indeed, let $g \in L^1(M)$ and $x \in M$, then

$$|Lg(x)| \leq \frac{1}{\sigma} \int_{h_+^{-1}(x) \vee x_-}^{h_-^{-1}(x) \wedge x_+} p\left(\frac{x - T(y)}{\sigma}\right) |g(y)| dy.$$

The claim follows since $\|Lg\|_\infty \leq \frac{\|p\|_\infty}{\sigma} \cdot \|g\|_{L^1}$.

Let $n \geq 2$. Because of Proposition 2.8 there are $A, B > 0$ such that

$$\text{Lip}(L^n g) \leq A\|L^{n-1}g\|_\infty + B\|L^{n-1}g\|_{L^1}.$$

Since L preserves integrals, then $\|Lg\|_{L^1} \leq \|g\|_{L^1}$ for all $g \in L^1(M)$. Because of this and the claim above we have that for all $n \geq 2$,

$$\text{Lip}(L^n g) \leq \left(\frac{A \cdot \|p\|_\infty}{\sigma} + B\right) \|L^{n-2}g\|_{L^1}.$$

Hence, since $\|L\| = 1$ (this is for $L : L^1(M) \rightarrow L^1(M)$), then

$$\text{Lip}(L^n g) \leq \left(\frac{A\|p\|_\infty}{\sigma} + B\right) \|g\|_{L^1}.$$

Therefore, $\text{Lip}(L^n g) \leq K$ for all $n \in \mathbb{N}$, where $K = \max\{\text{Lip}(g), \text{Lip}(Lg), A\|p\|_\infty/\sigma + B\}$. $\square$

We conclude this chapter by proving a regularizing property of the transfer operator L . For this, let us denote by $C_\partial^k(M)$ the space of real-valued C^k functions $f : M \rightarrow \mathbb{R}$ such that they vanish in ∂M along with all its derivatives up to order k , that is $f^{(j)}(x_+) = f^{(j)}(x_-) = 0$ for $j = 0, \dots, k$.

First, recall Lg is continuous for all $g \in L^1(M)$. Moreover, for $x = x_+$, we get $h_+^{-1}(x_+) = h_-^{-1}(x) \wedge x_+ = x_+$, so that $Lg(x_+) = 0$, and analogously $Lg(x_-) = 0$. Hence $Lg \in C_\partial^0(M)$ for all $g \in L^1(M)$. We extend this regularizing property to the spaces $C_\partial^r(M)$ as stated in the following proposition.

Proposition 2.12 *Assume T is a C^r -diffeomorphism and p is a C^r density in $(-1, 1)$ for some $r \geq 1$, then for all $g \in C_\partial^k(M)$ we have $Lg \in C_\partial^{k+1}(M)$ for each $k \in \{0, \dots, r-1\}$.*

Proof. Again, we do prove this statement for Case 1 of the expression of Lg , see (2.10). Let $x \in (x_2, x_1)$, and assume $g \in C_\partial^0(X)$. Then, if T is a C^1 -diffeomorphism, and p is a C^1 density,

$$\frac{d}{dx} Lg(x) = \frac{p(-1)g(h_-^{-1}(x))(h_-^{-1})'(x) - p(1)g(h_+^{-1}(x))(h_+^{-1})'(x)}{\sigma} + \frac{1}{\sigma^2} \int_{h_+^{-1}(x)}^{h_-^{-1}(x)} p'\left(\frac{x - T(y)}{\sigma}\right) g(y) dy.$$

Therefore, since $h_-^{-1}(x_1) = x_+$, and $g(x_+) = 0$,

$$\lim_{x \rightarrow x_1^-} \frac{d}{dx} Lg(x) = -\frac{p(1)g(h_+^{-1}(x_1))(h_+^{-1})'(x_1)}{\sigma} + \frac{1}{\sigma^2} \int_{h_+^{-1}(x_1)}^{x_+} p' \left(\frac{x_1 - T(y)}{\sigma} \right) g(y) dy.$$

On the other hand, if $x \in (x_1, x_+)$,

$$\frac{d}{dx} Lg(x) = -\frac{p(1)g(h_+^{-1}(x))(h_+^{-1})'(x)}{\sigma} + \frac{1}{\sigma^2} \int_{h_+^{-1}(x)}^{x_+} p' \left(\frac{x - T(y)}{\sigma} \right) g(y) dy.$$

Thus, $\lim_{x \rightarrow x_1^+} \frac{d}{dx} Lg(x) = \lim_{x \rightarrow x_1^-} \frac{d}{dx} Lg(x)$. Analogously, $\lim_{x \rightarrow x_2^+} \frac{d}{dx} Lg(x) = \lim_{x \rightarrow x_2^-} \frac{d}{dx} Lg(x)$, so that $Lg(x)$ is continuously differentiable. From the last formula it follows that $\frac{d}{dx} Lg(x_+) = 0$. In a similar way, $\frac{d}{dx} Lg(x_-) = 0$. The rest of the statement follows by induction. $\square$

Since any stationary density $\phi(X) \rightarrow \mathbb{R}$ satisfies that $L\phi = \phi$, we have this immediate consequence.

Corollary 2.13 *If T is a C^r -diffeomorphism, and p is a C^r density in $(-1, 1)$, then any stationary density supported on M is C^r . Moreover, ϕ is flat at x_+ and x_- , that is $\phi^{(k)}(x_+) = \phi^{(k)}(x_-) = 0$ for $k = 0, \dots, r$.*

It is the main purpose of the upcoming chapter to deduce a more precise statement on the shape of ϕ near x_+ and x_- .

Chapter 3

Tails of stationary measures

In the previous chapter we proved the existence and uniqueness of the stationary measure μ supported on a minimal invariant set $M = [x_-, x_+]$ for the system

$$x_{n+1} = T(x_n) + \sigma\omega_n \equiv h(x_n, \omega_n), \quad (3.1)$$

along with the exponential convergence of $\mathbb{P}_x^n$ in total variation norm towards μ for all $x \in M$. Recall that $d\mu = \phi d\text{Leb}$, where ϕ is the unique stationary density. We now turn our attention to the asymptotic behaviour of the stationary density ϕ as $x \rightarrow x_{\pm}$.

Recall that minimal invariant sets are determined by the fixed points of the extremal maps $h_+(x) = T(x) + \sigma$ and $h_-(x) = T(x) - \sigma$. In this section we address different scenarios depending on the type of fixed point x_+ is for the upper extremal map h_+ . We focus only on the behaviour of ϕ near x_+ , as the case for x_- is completely analogous. Without loss of generality let us assume that $x_+ = 0$. We address both cases when 0 is either a hyperbolic or a nonhyperbolic fixed point.

In this chapter we assume T is a C^{r+1} -diffeomorphism, for some $r \geq 2$ such that

$$h_+(x) = \lambda x + (-1)^r \alpha x^r + o(x^r), \quad (3.2)$$

where $\lambda \in (0, 1]$, $\alpha > 0$, and $r \geq 2$ is the first natural number for which $T^{(r)}(0) \neq 0$. Notice that if $\lambda < 1$, then 0 is a hyperbolic fixed point of h_+ , and nonhyperbolic otherwise. We still consider $(\omega_n)_{n=0}^{\infty}$ to be an i.i.d. sequence of random variables, but we only consider the case for which ω_n is uniformly distributed in $[-1, 1]$ for all $n \in \mathbb{N}$.

The main goal in this chapter is to provide asymptotic orders for $\phi(x)$ as $x \rightarrow x_+$. We show that when x_+ is nonhyperbolic, there exist $\gamma_1, \gamma_2 > 0$ such that for each $\epsilon > 0$ there is $x_0 < 0$ such that

$$|x|^{\frac{\gamma_2 + \epsilon}{|x|^{r-1}}} \leq \phi(x) \leq |x|^{\frac{\gamma_1 - \epsilon}{|x|^{r-1}}},$$

for all $x \in (x_0, 0)$. On the other hand, when x_+ is hyperbolic, there exist $\gamma_3, \gamma_4 > 0$ such that for each $\epsilon > 0$ there is $x_0 < 0$ such that for all $x \in (x_0, 0)$,

$$|x|^{-\gamma_3 \ln |x|} \leq \phi(x) \leq |x|^{-\gamma_4 \ln |x|}.$$

3.1 Main result and structure of the chapter

Let us formulate precisely the main theorem of this chapter

Theorem 3.1 (Order of ϕ as $x \rightarrow 0$) *Let the deterministic map T be a C^{r+1} -diffeomorphism for some $r \geq 2$, and $\phi : M \rightarrow \mathbb{R}_+$ be the unique stationary density supported on the minimal invariant set $M = [x_-, x_+]$, such that $h_+(x) = \lambda x + (-1)^r \alpha x^r + o(x^r)$ for some $\alpha > 0$ and $\lambda \in (0, 1]$. Then*

1. (Nonhyperbolic case) *If $\lambda = 1$, we have*

$$\frac{r}{\alpha(r-1)} \leq \liminf_{x \rightarrow 0^-} \frac{|x|^{r-1} \ln \phi(x)}{\ln |x|} \leq \limsup_{x \rightarrow 0^-} \frac{|x|^{r-1} \ln \phi(x)}{\ln |x|} \leq \frac{2r}{\alpha(r-1)}. \quad (3.3)$$

2. (Hyperbolic case) *If $\lambda \in (0, 1)$, we get*

$$\frac{1}{\ln \lambda} \leq \liminf_{x \rightarrow 0} \frac{\ln \phi(x)}{(\ln |x|)^2} \leq \limsup_{x \rightarrow 0} \frac{\ln \phi(x)}{(\ln |x|)^2} \leq \frac{1}{2 \ln \lambda}. \quad (3.4)$$

Notice it has already been proved in [59] that stationary densities ϕ supported on a minimal invariant set $M = [x_-, x_+]$ are flat at the boundaries of M (see also Corollary 2.13). Our results further develop this statement for random interval diffeomorphisms by giving upper and lower bounds for the asymptotic behaviour of these densities near the boundaries of M . We conclude that it has no precise exponential behaviour near $x_{\pm}$ of the form $\phi(x) \approx e^{-\gamma/|x-x_{\pm}|^p}$ for any $p > 0$, like those expressions that appear in the classical theory of *large deviation principles* [21].

We first focus on the nonhyperbolic scenario for x_+ . In Section 3.2 we find a unique (local) flow ψ_t and its associated vector field $G : \mathbb{R} \rightarrow \mathbb{R}$ for which its time-one map coincides with h_+ , that is $h_+ = \psi_1$ near $x_+ = 0$. Whenever such flow exists, we say that the map h_+ is *embedded into the flow* ψ_t . The idea of embedding maps into flows dates back to Takens [54] in the case where the embeddable map is a C^∞ -diffeomorphism. More recently, such techniques have also been exploited for studying bifurcations of flows arising from *saddle-node limit cycles* [53].

The problem of determining the speed of convergence towards the equilibrium x_+ under the flow ψ_t turns out to be equivalent to describing the first-hitting time

$$t_x^y := \inf\{t \in \mathbb{R} \mid \psi_t(x) > y\},$$

as $y \rightarrow 0$. Moreover, since the map h_+ is embedded in the flow ψ_t , its asymptotic behaviour coincides with that of the discrete first-hitting time

$$n_x^y = \min\{n \in \mathbb{N}_0 \mid h_+^n(x) > y\}$$

In Section 3.3 we proceed to prove the first part of Theorem 3.1. Furthermore, we show that a similar asymptotic behaviour holds for the tail of the stationary measure, $\mathcal{T}(x) = \mu([x, 0])$, as $x \rightarrow 0$. We confirm the theoretical bounds (3.3) with some numerical explorations.

By the end of the chapter, in Section 3.4 we address the second part of Theorem 3.1 for the case 0 is a hyperbolic fixed point. Notice how the asymptotic approximations (3.4) for the stationary density are of different nature with respect to the nonhyperbolic scenario. The proof follows in a similar fashion as in the case for nonhyperbolic fixed points, but with elementary techniques.

3.2 Embedding of deterministic maps into flows

In the rest of the present chapter, we will need asymptotic estimates for the first-hitting times

$$n_x^y := \min\{n \geq 0 \mid h_+^n(x) > y\} \tag{3.5}$$

as $y \rightarrow 0$. Notice n_x^y is the first time in the orbit of h_+ starting at x that enters the interval $(y, 0)$. However, in the case when x_+ is nonhyperbolic for h_+ it seems to be a hard task to understand the asymptotic behaviour of n_x^y as $y \rightarrow 0$ just from the properties of the map h_+ itself. Instead, we study it via the first-hitting times of a flow into which h_+ is embedded.

Observe that the upper extremal map h_+ is a C^{r+1} -diffeomorphism, since the deterministic map T is a C^{r+1} -diffeomorphism too. Thus, the existence and uniqueness of a C^{r+1} flow into which we can embed h_+ into in a neighbourhood of 0 is given by a generalization due to Yoccoz [57].

Theorem 3.2 (Takens embedding theorem [57, Appendix 3]) *let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a C^{r+1} -diffeomorphism for some $r \geq 1$ which preserves orientation. If $h - Id$ is not flat¹ at 0, where $Id(x) = x$, there exists an open neighbourhood of 0, U , and a vector field $G : U \rightarrow \mathbb{R}$ of class C^r such that its associated flow ψ_t satisfies $\psi_1 = h$ in U .*

The next statement gives a simple expression for the vector field defining the flow we embed h_+ into.

Proposition 3.3 *Let $h_+(x) = x + (-1)^r \alpha x^r + o(x^r)$ for $x \leq 0$, and ψ_t be a flow with associated vector field $G : U \subset \mathbb{R} \rightarrow \mathbb{R}$, where U is an open interval containing 0, such that $h_+ = \psi_1$ in U . Then*

$$G(x) = (-1)^r \alpha x^r + o(x^r). \quad (3.6)$$

Proof. let $G : U \subset \mathbb{R} \rightarrow \mathbb{R}$ and U as given in Theorem 3.2 with associated flow ψ_t such that $\psi_1 = h_+$ in U , where U is an open neighbourhood of 0. Observe that since 0 is an equilibrium point for ψ_t , then $\psi_t(0) = 0$ for all $t > 0$ and therefore $G(0) = 0$. Let $b_1, b_2, \dots, b_r$ be functions of t , and $a_1, a_2, \dots, a_r \in \mathbb{R}$ such that

$$\psi_t(x) = b_1(t)x + b_2(t)x^2 + \dots + b_r(t)x^r + o(x^r),$$

$$G(x) = a_1x + a_2x^2 + \dots + a_rx^r + o(x^r).$$

Since for each $x \in U$, the function $t \mapsto \psi_t(x)$ is a solution of the ordinary differential equation $\dot{x} = G(x)$, then

$$\dot{b}_1x + \dot{b}_2x^2 + \dots + \dot{b}_rx^r + o(x^r) = a_1(b_1x + \dots + b_rx^r) + \dots + a_r(b_1x + \dots + b_rx^r)^r + o(x^r). \quad (3.7)$$

Thus, by solving for the coefficients in the linear term in x , we obtain

$$b_1(t) = b_1(0)e^{a_1t}.$$

¹Recall a C^k function is flat at x if all its derivatives of order less than or equal to k vanish at x , where $k \in \mathbb{N}$.

Since $\psi_0(x) = x$ and $\psi_1(x) = x + (-1)^r \alpha x^r + o(x^r)$ for all x , then $b_1(0) = 1$ and $a_1 = 0$, implying that $b_1 \equiv 1$. Substituting this in (3.7) we obtain for the x^2 coefficients that b_2 satisfies $\dot{b}_2 = a_2$, where $b_2(0) = 0$ and $b_2(1) = \alpha$ if $r = 2$, and $b_2(1) = 0$ if $r \geq 3$. Hence,

$$b_2(t) = \begin{cases} \alpha t, & \text{if } r = 2 \\ 0, & \text{if } r > 2 \end{cases}.$$

In general, $b_1 \equiv 1$, $b_k \equiv 0$ for $k = 2, \dots, r-1$, $a_j = 0$ for $j = 1, \dots, r-1$, $a_r = (-1)^r \alpha$, and $b_r(t) = (-1)^r \alpha t$, for which

$$\begin{aligned} \psi_t(x) &= x + (-1)^r \alpha t x^r + o(x^r), \\ G(x) &= (-1)^r \alpha x^r + o(x^r) \end{aligned}$$

□

We now study the first-hitting time of ψ_t to the open interval $(y, 0)$ as $y \rightarrow 0$. Let

$$t_x^y := \inf\{t \geq 0 \mid \psi_t(x) \in (y, 0)\}.$$

In other words, whenever $x < y$, the time t_x^y is the only $t > 0$ such that $\psi_t(x) = y$, and equals 0 if $x \geq y$. In fact, for $x < y$,

$$t_x^y = \int_x^y \frac{ds}{G(s)}, \quad (3.8)$$

where $G : U \subset \mathbb{R} \rightarrow \mathbb{R}$ is given as in (3.6).

Lemma 3.4 (Scaling law for t_x^y) *for all negative x in U ,*

$$\lim_{y \rightarrow 0^-} |y|^{r-1} \cdot t_x^y = \frac{1}{\alpha(r-1)}. \quad (3.9)$$

Proof. For simplicity, we prove the case where r is even, as the case when it is odd just carries a (-1) factor along the proof as the one here presented.

Observe that $G(x) > 0$ for all $x < 0$ and $x \in U$, as otherwise there would exist a fixed point for h_+ which is not $x_+ = 0$. Since $G(s) = \alpha s^r + o(s^r)$, then for each $\epsilon \in (0, \alpha)$ there is $z_\epsilon \in U$ such that for all $s \in [z_\epsilon, 0)$,

$$(\alpha - \epsilon)s^r < G(s) < (\alpha + \epsilon)s^r.$$

From (3.8), for $x \in [z_\epsilon, 0)$ and $y \in (x, 0)$ we obtain

$$\frac{y^{r-1} - x^{r-1}}{(\alpha + \epsilon)(r-1)x^{r-1}y^{r-1}} \leq t_x^y \leq \frac{y^{r-1} - x^{r-1}}{(\alpha - \epsilon)(r-1)x^{r-1}y^{r-1}}.$$

Multiplying times $|y|^{r-1}$ we get

$$\frac{1}{(\alpha + \epsilon)(r - 1)} \leq \liminf_{y \rightarrow 0} |y|^{r-1} t_x^y \leq \limsup_{y \rightarrow 0} |y|^{r-1} t_x^y \leq \frac{1}{(\alpha - \epsilon)(r - 1)} \quad (3.10)$$

On the other hand, for points $x \in U$ such that $x < z_\epsilon$, notice that $t_x^y = t_x^{z_\epsilon} + t_{z_\epsilon}^y$, so in general for all nonzero $x \in U$ inequality (3.10) holds. Moreover, since it is true for all $\epsilon > 0$, the result follows. $\square$

Recall that n_x^y is the first time the orbit of h_+ starting at x hits the open interval $(y, 0)$. Formally speaking,

$$n_x^y = \min\{n \in \mathbb{N}_0 \mid h_+^n(x) > y\}.$$

The scaling law for n_x^y follows directly from Lemma 3.4 as given in the next statement.

Corollary 3.5 (Scaling law for n_x^y) *For each $x \in M \setminus \{0\}$*

$$\lim_{y \rightarrow 0^-} |y|^{r-1} \cdot n_x^y = \frac{1}{\alpha(r - 1)} \quad (3.11)$$

Proof. Assume without loss of generality that h_+ is embedded into ψ_t in $(\tilde{x}, 0]$ for some $\tilde{x} < 0$. Then, for all $x \in (\tilde{x}, 0)$ and $y \in (x, 0)$,

$$t_x^y \leq n_x^y \leq t_x^y + 1.$$

Thus, for all $x \in (\tilde{x}, 0)$, the limit (3.11) is satisfied. Furthermore, for any other $x \in M \setminus (\tilde{x}, 0]$ and $y \in (\tilde{x}, 0)$, let $N_x := n_{\tilde{x}}^y \in \mathbb{N}$. Then $n_x^y = N_x + n_{x_{N_x}}^y$, where $x_{N_x} = h_+^{N_x}(x)$. Thus, since $x_{N_x} \in [\tilde{x}, 0)$,

$$\lim_{y \rightarrow 0^-} |y|^{r-1} \cdot n_x^y = \lim_{y \rightarrow 0^-} |y|^{r-1} \cdot n_{x_{N_x}}^y = \frac{1}{\alpha(r - 1)}$$

$\square$

Remark 3.6 *Notice that for fixed negative $x \in U$, the function $t_x : (x, 0) \rightarrow (0, \infty)$ defined simply as $t_x(y) := t_x^y$ is a diffeomorphism which is strictly increasing. Moreover, its inverse y_x is precisely*

$$y_x(t) = \psi_t(x).$$

Therefore, by a change of variables, we can conclude that for any negative $x \in U$,

$$\lim_{t \rightarrow \infty} |\psi_t(x)|^{r-1} \cdot t = \frac{1}{\alpha(r-1)}.$$

In particular, by taking the sequence $t_n = n$, and since h_+ is embedded into ψ_t we obtain

$$\lim_{n \rightarrow \infty} n \cdot |h_+^n(x)|^{r-1} = \frac{1}{\alpha(r-1)}.$$

For any other $x \in M \setminus \{0\}$, if $n \geq N_x := n_{\tilde{x}}$ for some $\tilde{x} \in U$, then $n = N + m$ for some $m \in \mathbb{N}_0$. Therefore, since $h_+^N(x) \in (\tilde{x}, 0)$,

$$\lim_{n \rightarrow \infty} n |h_+^n(x)|^{r-1} = \lim_{m \rightarrow \infty} (N + m) |h_+^{N+m}(x)|^{r-1} = \lim_{m \rightarrow \infty} m |h_+^m(h_+^N(x))|^{r-1} = \frac{1}{\alpha(r-1)}. \quad (3.12)$$

3.3 Nonhyperbolic boundaries of minimal invariant sets

We now study the behaviour of the stationary density ϕ as $x \rightarrow 0^-$, where 0 is a nonhyperbolic fixed point of h_+ . Recall that ϕ is a fixed point of the transfer operator L . Thus, ϕ satisfies the functional equation (see C.17 in Appendix C)

$$\phi(x) = \frac{1}{2\sigma} \int_{h_+^{-1}(x) \vee x_-}^{h_+^{-1}(x) \wedge 0} \phi(y) dy.$$

Observe that $h_+^{-1}(x) \geq 0$ if and only if $x \geq h_-(0) = T(0) - \sigma = -2\sigma$. Thus, for $x \geq -2\sigma$

$$\phi(x) = \frac{1}{2\sigma} \int_{h_+^{-1}(x)}^0 \phi(y) dy. \quad (3.13)$$

Let us show that the stationary density is decreasing near 0. This property does not depend on the hyperbolicity of 0.

Proposition 3.7 *The stationary density is decreasing in $(-2\sigma, 0)$.*

Proof. The result is a direct consequence of (3.13). Observe that since ϕ is C^1 in $(-2\sigma, 0)$, then

$$\phi'(x) = \frac{d}{dx} \left[\frac{1}{2\sigma} \int_{h_+^{-1}(x)}^0 \phi(y) dy \right] = -\frac{\phi(h_+^{-1}(x))}{2\sigma \cdot h'_+(h_+^{-1}(x))} < 0.$$

□

Let us proceed with the following lemma, which lays the groundwork for the proof of Theorem 3.1. Again, this lemma holds regardless of the hyperbolicity of 0.

Lemma 3.8 *There exist constants $K_1, K_2 > 0$ such that for every $x \in (h_+(-2\sigma), 0)$ we have*

$$K_1(x - h_+^{-1}(x))^2 \cdot \phi(h_+^{-1}(x)) \leq \phi(x) \leq K_2(x - h_+^{-1}(x)) \cdot \phi(h_+^{-1}(x)) \quad (3.14)$$

Proof. We split the proof in two parts, where we prove each of the inequalities in (3.14).

Part 1: *upper bound in (3.14).*

It follows from (3.13) and the monotonicity of ϕ that for $x > -2\sigma$,

$$\phi(x) = \frac{1}{2\sigma} \left[\int_{h_+^{-1}(x)}^x \phi(y) dy + \int_x^0 \phi(y) dy \right] \leq \left[\frac{x - h_+^{-1}(x)}{2\sigma} \right] \phi(h_+^{-1}(x)) + \frac{|x|}{2\sigma} \phi(x).$$

Therefore,

$$\left(1 - \frac{|x|}{2\sigma}\right) \phi(x) \leq \left[\frac{x - h_+^{-1}(x)}{2\sigma} \right] \phi(h_+^{-1}(x)),$$

for which we conclude that

$$\phi(x) \leq \left(\frac{x - h_+^{-1}(x)}{2\sigma - |x|} \right) \cdot \phi(h_+^{-1}(x)).$$

The upper bound in (3.14) follows by taking $K_2 = 1/(2\sigma - |h_+(-2\sigma)|)$.

Part 2: *lower bound in (3.14).*

Observe that formula (3.13) can be iterated if $h_+^{-1}(x) > -2\sigma$, that is if $x > h_+(-2\sigma)$, which holds by hypothesis. Hence, we get

$$\phi(x) = \frac{1}{(2\sigma)^2} \int_{h_+^{-1}(x)}^0 \int_{h_+^{-1}(y)}^0 \phi(z) dz dy. \quad (3.15)$$

Observe that the integration region in (3.15) (see Figure 3.1) is the same as

$$\{(y, z) \mid h_+^{-1}(x) \leq y \leq h_+(z); h_+^{-2}(x) \leq z \leq 0\}.$$

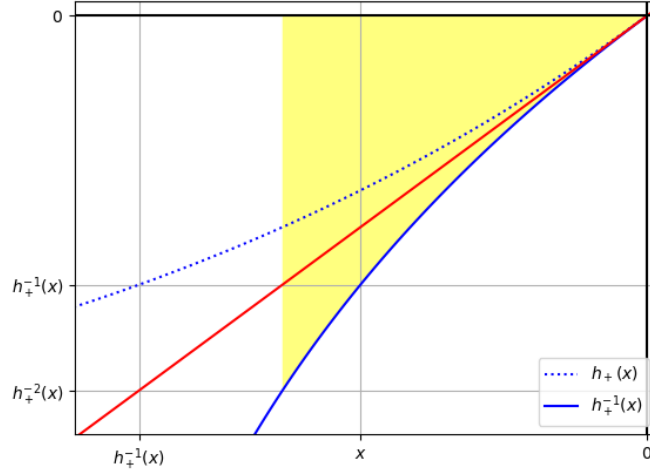


Figure 3.1: The shaded region represents the integration region in (3.15). The order of integration can be interchanged, by modifying the integration limits.

Therefore, rewriting (3.15) by exchanging the order of integration, we get

$$\phi(x) = \frac{1}{(2\sigma)^2} \int_{h_+^{-2}(x)}^0 [h_+(z) - h_+^{-1}(x)] \phi(z) dz.$$

Recall that ϕ is decreasing in the interval $(-2\sigma, 0)$. Hence, by reducing the interval of integration,

$$\phi(x) \geq \frac{1}{(2\sigma)^2} \int_{h_+^{-2}(x)}^{h_+^{-1}(x)} [h_+(z) - h_+^{-1}(x)] \phi(z) dz \geq \frac{\phi(h_+^{-1}(x))}{(2\sigma)^2} \int_{h_+^{-2}(x)}^{h_+^{-1}(x)} [h_+(z) - h_+^{-1}(x)] dz$$

By the mean value theorem,

$$h_+(z) = h_+^{-1}(x) + h'_+(\xi)(z - h_+^{-2}(x))$$

for some $\xi \in (h_+^{-2}(x), z)$. Since h'_+ is continuous and positive, then we have

$$\phi(x) \geq M_1 \cdot (h_+^{-1}(x) - h_+^{-2}(x))^2 \cdot \phi(h_+^{-1}(x)). \quad (3.16)$$

where $M_1 := \inf_{x \in [h_+^{-1}(-2\sigma), 0]} h'_+(x)/(8\sigma^2) > 0$.

Again, by the mean value theorem, there is $\eta \in (h_+^{-2}(x), h_+^{-1}(x))$ such that

$$h_+^{-1}(x) - h_+^{-2}(x) = \frac{x - h_+^{-1}(x)}{h'_+(\eta)} \geq \frac{x - h_+^{-1}(x)}{\sup_{z \in (h_+^{-1}(-2\sigma), 0)} h'_+(z)},$$

which we can use in (3.16) so that

$$\phi(x) \geq K_1 \cdot (x - h_+^{-1}(x))^2 \cdot \phi(h_+^{-1}(x)), \quad (3.17)$$

where $K_1 = M_1 / \sup_{z \in (h_+^{-1}(-2\sigma), 0)} h'_+(z)$. $\square$

We now employ the previous lemma in the case 0 is a nonhyperbolic fixed point for the upper extremal map h_+ . As seen in (3.14), it is necessary to estimate the factor $x - h_+^{-1}(x)$. Recall h_+ is a C^{r+1} -diffeomorphism for some $r \geq 2$, and we assume that $x_+ = 0$ is a nonhyperbolic fixed point for h_+ for which

$$h_+(x) = x + (-1)^r \alpha x^r + o(x^r) = x + \alpha |x|^r + o(x^r),$$

where α is a positive constant.

We give a similar expansion for h_+^{-1} for $x \in M$. Indeed, assume

$$h_+^{-1}(x) = a_1 x + a_2 x^2 + \cdots + a_r x^r + o(x^r).$$

Since $h_+(h_+^{-1}(x)) = x$ for all x , then

$$x = a_1 x + a_2 x^2 + \cdots + a_{r-1} x^{r-1} x^{r-1} + (a_r + (-1)^r \alpha) x^r + o(x^r).$$

Hence, $a_1 = 1$, $a_j = 0$ for $j = 2, \dots, r-1$, and $a_r = (-1)^{r+1} \alpha$, so that

$$h_+^{-1}(x) = x + (-1)^{r+1} \alpha x^r + o(x^r) = x - \alpha |x|^r + o(x^r).$$

Let us proceed then with the proof of the first part of Theorem 3.1.

Proof of Theorem 3.1 (nonhyperbolic case). Let $\epsilon \in (0, \alpha)$ be an arbitrarily small constant. Then there exists a negative $x_0 \in (h_+(-2\sigma), 0)$ such that for all $x \in [x_0, 0]$,

$$(\alpha - \epsilon)|x|^r \leq x - h_+^{-1}(x) \leq (\alpha + \epsilon)|x|^r. \quad (3.18)$$

Again, we split the proof in two parts, one for each inequality in (3.3).

Part 1: upper bound in (3.14).

From the lower bound in (3.14) from Lemma 3.8, and from (3.18), for all $x \in (x_0, 0)$ we have

$$\phi(x) \geq \tilde{K}_1 |x|^{2r} \phi(h_+^{-1}(x)),$$

where $\tilde{K}_1 = K_1(\alpha - \epsilon)$. We can use the same inequality for $\phi(h_+^{-1}(x))$ if $h_+^{-1}(x) > x_0$. In such a case,

$$\phi(x) \geq \tilde{K}_1^2 |x|^{2r} |h_+^{-1}(x)|^{2r} \phi(h_+^{-1}(x)).$$

In general, given $x > x_0$, we can iterate the previous procedure at most $n_{x_0}^x - 1$ times. Thus, for $n \in \{1, \dots, n_{x_0}^x - 1\}$,

$$\phi(x) \geq \tilde{K}_1^n \cdot |x|^{2r} \cdots |h_+^{-n+1}(x)|^{2r} \cdot \phi(h_+^{-n}(x)).$$

However, since $|h_+^{-j}(x)| \geq |x|$ for all $j \in \mathbb{N}$, we conclude that

$$\phi(x) \geq \tilde{K}_1^n \cdot |x|^{2rn} \cdot \phi(h_+^{-n}(x)). \quad (3.19)$$

Since ϕ is decreasing in $[-2\sigma, 0]$, it follows by considering $n = n_{x_0}^x - 1$

$$\phi(x) \geq \tilde{K}_1^n \cdot |x|^{2rn} \cdot \phi(h_+(x_0)).$$

By taking natural logarithm, and multiplying times $|x|$ we obtain

$$|x| \ln \phi(x) \geq |x| (n_{x_0}^x - 1) \cdot (\ln \tilde{K}_1 + 2r \ln |x|) + |x| \ln \phi(h_+^{-1}(x_0)).$$

Diving by $\ln |x|$ (which is negative for x close to 0) and taking $\limsup$ as $x \rightarrow 0^-$, because of Corollary 3.5 we obtain

$$\limsup_{x \rightarrow 0^-} \frac{|x| \ln \phi(x)}{\ln |x|} \leq \frac{2r}{\alpha(r-1)}. \quad (3.20)$$

Part 2: lower bound in (3.14).

Consider now the upper bound in (3.14) given in Lemma 3.8 so that for $x \in (x_0, 0)$, due to (3.18),

$$\phi(x) \leq \tilde{K}_2 \cdot |x|^r \cdot \phi(h_+^{-1}(x)),$$

where $\tilde{K}_2 = K_2(\alpha + \epsilon)$. Again, by iterating the previous inequality n times for $n \leq n_{x_0}^x - 1$

$$\phi(x) \leq \tilde{K}_2^n \cdot |x|^r \cdots |h_+^{-n+1}(x)|^r \cdot \phi(h_+^{-n}(x)). \quad (3.21)$$

By taking natural logarithm, and since ϕ is decreasing in $[-2\sigma, 0]$,

$$\ln \phi(x) \leq r \cdot \sum_{j=0}^{n-1} \ln |h_+^{-j}(x)| + n \ln K_2 + \ln \phi(h_+^{-1}(x_0)). \quad (3.22)$$

We now estimate from above the sum appearing in the last inequality. Observe that

$$S_n := \sum_{j=0}^{n-1} \ln |h_+^{-j}(x)| = \sum_{j=1}^n \frac{\ln |h_+^{-j}(x)| \cdot \Delta h_+^{-j}}{h_+^{-j}(x) - h_+^{-j-1}(x)} = \sum_{j=0}^{n-1} \frac{\ln |h_+^{-j}(x)| \cdot \Delta h_+^{-j}}{\alpha(h_+^{-j}(x))^2 + o((h_+^{-j}(x))^2)},$$

where we denote $\Delta h_+^{-j} = h_+^{-j}(x) - h_+^{-j-1}(x)$. This last expression resembles a Riemann sum and we approximate it by an integral. For this, recall that $x - h_+^{-1}(x) \leq (\alpha + \epsilon)|x|^r$ for all $x \in (x_0, 0)$ as indicated in (3.18). Thus, for each $n \in \{1, 2, \dots, n_{x_0}^x - 1\}$ we have

$$S_n \leq \frac{1}{\alpha + \epsilon} \sum_{j=0}^{n-1} \frac{\ln |h_+^{-j}(x)|}{|h_+^{-j}(x)|^r} \Delta h_+^{-j} \quad (3.23)$$

Since the function $z \mapsto \ln z / z^r$ is increasing and negative for $z \in (0, 1)$ (see Figure 3.2), assuming without loss of generality that $x_0 > -1$, then

$$S_n \leq \frac{1}{\alpha + \epsilon} \int_{|x|}^{|h_+^{-n+1}(x)|} \frac{\ln s}{s^r} ds.$$

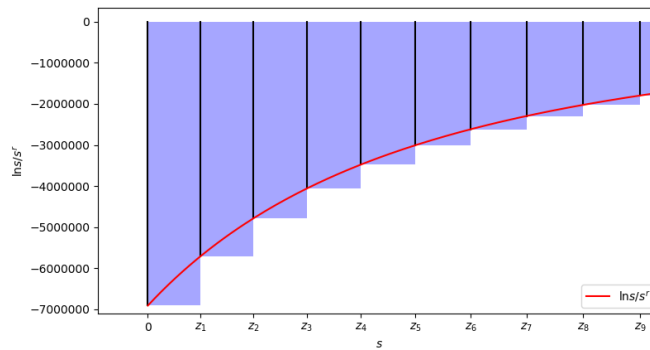


Figure 3.2: The function $z \mapsto \ln z$ is negative and increasing for $z \in (0, 1)$. The sum in the right hand side of (3.23) is shown as the blue region. Here, $z_i = |h_+^{-i}(x)|$ for $i \in \mathbb{N}$.

The integral above can be solved explicitly by doing the substitution $u = \ln s$ and integrating by parts afterwards. Therefore we obtain

$$S_n \leq \frac{1}{(\alpha + \epsilon)(r - 1)} \left[\frac{\ln |x|}{|x|^{r-1}} - \frac{\ln |h_+^{-n+1}(x)|}{|h_+^{-n+1}(x)|^{r-1}} + \frac{1}{(r - 1)|x|^{r-1}} - \frac{1}{(r - 1)|h_+^{-n+1}(x)|^{r-1}} \right].$$

Notice that for all $x > x_0$, we have $|h_+(x_0)| \leq |h_+^{-n_{x_0}+1}(x)| \leq |x_0|$. By multiplying times $|x|^{r-1}$ and dividing by $\ln |x|$, it follows that for all $\epsilon > 0$

$$\liminf_{x \rightarrow 0} \frac{|x|^{r-1} S_n}{\ln |x|} \geq \frac{1}{(\alpha + \epsilon)(r - 1)}, \quad (3.24)$$

so it remains valid for $\epsilon = 0$. Finally, from (3.22) for $n = n_{x_0}^x$, and From Corollary 3.5, the product $n_{x_0}^x |x|$ remains bounded as $x \rightarrow 0$. Because of (3.24) we obtain the lower limit

$$\liminf_{x \rightarrow 0} \frac{|x| \ln \phi(x)}{\ln |x|} \geq r \cdot \liminf_{x \rightarrow 0} \frac{|x| S_n}{\ln |x|} \geq \frac{r}{\alpha(r - 1)}.$$

□

Remark 3.9 Theorem 3.1 excludes the possibility of an exponential behaviour for ϕ of the form $\phi(x) \approx e^{-\frac{\gamma}{|x|^p}}$ for any $\gamma, p > 0$ as $x \rightarrow 0^-$. Indeed, from (3.3) we know there exists $\tilde{x} < 0$ such that for every $x \in (\tilde{x}, 0)$

$$\frac{r}{2\alpha(r - 1)} \leq \frac{|x|^{r-1} \ln \phi(x)}{\ln |x|} \leq \frac{3r}{\alpha(r - 1)}.$$

The left inequality implies that,

$$|x|^{r-1} \ln \phi(x) \leq \frac{r}{2\alpha(r - 1)} \ln |x| \rightarrow -\infty \text{ as } x \rightarrow 0^-.$$

On the other hand, from the right inequality we get that for any $p > r - 1$

$$|x|^p \ln \phi(x) \geq \frac{3r}{\alpha(r - 1)} |x|^{p-r+1} \ln |x| \rightarrow 0 \text{ as } x \rightarrow 0^-.$$

We can then conclude that

$$\lim_{x \rightarrow 0^-} |x|^p \ln \phi(x) = \begin{cases} -\infty & \text{if } 0 < p \leq r - 1 \\ 0 & \text{if } p > r - 1 \end{cases}. \quad (3.25)$$

Observe this abrupt transition in the behaviour of $|x|^p \ln \phi(x)$ as $x \rightarrow 0$ detects the type of tangency the graph of h_+ has with respect to the identity line at $x_+ = 0$.

Let us conclude this section with an analogous result for the tail of the stationary measure, μ . The *cumulative distribution function* of μ is defined as the function $x \mapsto \mu([x_-, x])$, and the tail of μ is defined as $\mathcal{T}(x) := \mu([x, x_+])$. The behaviour of $\mathcal{T}$ is a direct consequence of Theorem 3.1.

Corollary 3.10 *Let $M = [x_-, 0]$ be a minimal invariant set for (3.1) where the deterministic map T is a C^{r+1} diffeomorphism for $r \geq 1$, and let μ be the unique stationary measure supported on M , with tail function $\mathcal{T}(x) = \mu([x, 0])$. Then, if 0 is a nonhyperbolic fixed point for h_+ so that $h_+(x) = x + (-1)^r \alpha x^r + o(x^r)$ for some $\alpha > 0$, then*

$$\frac{r}{\alpha(r-1)} \leq \liminf_{x \rightarrow 0^-} \frac{|x|^{r-1} \ln \mathcal{T}(x)}{\ln |x|} \leq \limsup_{x \rightarrow 0^-} \frac{|x|^{r-1} \ln \mathcal{T}(x)}{\ln |x|} \leq \frac{2r}{\alpha(r-1)} \quad (3.26)$$

Proof. Let $\phi = d\mu/d\text{Leb}$ be the stationary density for μ . Then

$$\mu([x, 0]) = \int_x^0 \phi(y) dy = 2\sigma \cdot \left(\frac{1}{2\sigma} \int_x^0 \phi(y) dy \right) = 2\sigma \cdot \phi(h_+(x)). \quad (3.27)$$

Since $h_+(x) = x + (-1)^r \alpha x^r + o(x^r)$,

$$\frac{|x|^{r-1} \ln \phi(h_+(x))}{\ln |x|} = \frac{|h_+(x)|^{r-1} \ln \phi(h_+(x))}{\ln |h_+(x)|} \cdot \left(1 + \frac{\ln |1 + o(1)|}{\ln |x|} \right) \cdot \left(\frac{1}{1 + o(1)} \right)^{r-1}.$$

The result then follows from Theorem 3.1 by taking $\ln$ in (3.27), multiplying times $|x^{r-1}|$, dividing by $\ln |x|$, and taking limits. $\square$

3.3.1 Numerical explorations

Let us verify our previous results with a numerical approximation of the stationary density with a piecewise constant function. This procedure is commonly known as the *Ulam's method*. For this we consider a regular partition of the minimal invariant set $M = [x_-, x_+]$, given by

$$\mathcal{P} = \{x_0 = x_- < x_1 < \dots < x_N = x_+\},$$

where $\Delta x_i = x_i - x_{i-1}$ is constant for all $i = 1, 2, \dots, N$. Let us denote by $I_i = [x_{i-1}, x_i)$, for each $i = 1, \dots, N$. We consider the transition probabilities $\mathbb{P}_x$ for each $x \in M$ and the stochastic matrix $Q = (Q_{ij})_{i,j=1}^N$, where

$$Q_{ij} = \frac{1}{\text{Leb}(I_j)} \int_{I_j} \mathbb{P}_x(I_i) dx.$$

We approximate the integral above with the trapezoidal rule, so that

$$Q_{ij} \approx \frac{1}{\Delta x_j} \frac{(\mathbb{P}_{x_j}(I_i) + \mathbb{P}_{x_{j-1}}(I_i))\Delta x_j}{2} = \frac{\mathbb{P}_{x_j}(I_i) + \mathbb{P}_{x_{j-1}}(I_i)}{2}.$$

Since Q is stochastic, $\Lambda = 1$ is an eigenvalue and coincides with its spectral radius. It has an associated non-negative left eigenvector which can be approximated, for instance, as

$$v^{(n)} = \frac{(Q^T)^n v^{(0)}}{\|(Q^T)^n v^{(0)}\|_1},$$

for $n \in \mathbb{N}$ large enough, where $v^{(0)} \in \mathbb{R}^N$ is an arbitrary vector with positive entries, and $\|\cdot\|_1$ is the L^1 norm in $\mathbb{R}^N$. Observe that $v^{(n)}$ satisfies $\|v^{(n)}\| = 1$, for which it will be necessary to rescale it to have an adequate approximation of the stationary density. We consider then for n large

$$v^* = \frac{1}{\Delta x_1} v^{(n)},$$

and the numerical approximation of the stationary density ϕ , is given by

$$\phi_{ap}(x) = v_i^* \quad \text{if } x \in I_i.$$

Let us consider the random difference equation presented in Chapter 1,

$$x_{n+1} = b \left(\frac{1 - e^{-x}}{1 + e^{-x}} \right) + \sigma \omega_n \equiv T(x_n) + \sigma \omega_n$$

where $b > 2$, and $(\omega_n)_{n=0}^\infty$ is an i.i.d. sequence of random variables uniformly distributed in $[-1, 1]$.

In order to find a nonhyperbolic fixed point for some σ^* , we look for the pair (x_+, σ^*) such that $T'(x_+) = 1$ and $T(x_+) + \sigma = x_+$, so that x_+ is a non-degenerate saddle-node fixed point for h_+ . We did this already in Chapter 1, where

$$x_+ = \ln \left(b - 1 + \sqrt{(b-1)^2 - 1} \right) < 0,$$

$$\sigma^* = b \left(\frac{b - 2 + \sqrt{(b-1)^2 - 1}}{b + \sqrt{(b-1)^2 - 1}} \right) - \ln \left(b - 1 + \sqrt{(b-1)^2 - 1} \right).$$

Also, as we say in Chapter 1,

$$\alpha = -\frac{x^* - \sigma^*}{2b}.$$

Recall from Theorem 3.1 that for $r = 2$,

$$\gamma_1 \lesssim \frac{|x_+ - x| \ln \phi(x)}{\ln |x_+ - x|} \lesssim \gamma_2,$$

where $\gamma_1 = 2/\alpha$ and $\gamma_2 = 4/\alpha$, and $\alpha = T''(x_+)$. As an example, when multiplying times $\ln |x_+ - x|$ and dividing by $|x_+ - x|$ for x close to but less than x_+ , we obtain from the previous inequality

$$\ln \phi(x) \gtrsim \frac{\gamma_2 \ln |x_+ - x|}{|x_+ - x|}.$$

By multiplying times (-1) and taking logarithms we get

$$\ln \ln \left(\frac{1}{\phi(x)} \right) \lesssim \ln \gamma_2 + \ln \left(\frac{1}{x_+ - x} \right) + \ln \ln \left(\frac{1}{x_+ - x} \right).$$

In a similar fashion

$$\ln \ln \left(\frac{1}{\phi(x)} \right) \gtrsim \ln \gamma_1 + \ln \left(\frac{1}{x_+ - x} \right) + \ln \ln \left(\frac{1}{x_+ - x} \right).$$

The last two inequalities imply the curve $(-\ln \phi(x), -\ln |x_+ - x|)$ is roughly in between the curves $\ln \gamma_i + u + \ln u$ in $\log - \log$ scale. This is depicted in Figure 3.3 below.

Remark 3.11 *The numerical exploration for a nondegenerate nonhyperbolic fixed point (that is when $r = 2$) in the toy model suggests that*

$$\lim_{x \rightarrow x_+} \frac{|x_+ - x| \ln \phi(x)}{\ln |x_+ - x|} = \frac{4}{\alpha}.$$

In order to achieve this limit one should improve the upper bound in Lemma 3.8. So far inequality (3.14) is the best result we were able to obtain, and other attempts lead us to less accurate estimates. For instance, notice from the monotonicity of ϕ near $x_+ = 0$ and (3.13) that

$$\phi(x) \leq \frac{1}{2\sigma} |h_+^{-1}(x)| \cdot \phi(h_+^{-1}(x))$$

Applying the inequality above $n - 1$ times and taking logarithms we obtain

$$\ln \phi(x) \leq \sum_{j=1}^{n-1} \ln |h_+^{-j}(x)| + \dots$$

By taking $n = n_{x_0}^x$, we repeat the same procedure as in Part 2 in the proof of Theorem 3.1, in order to obtain

$$\frac{1}{\alpha(r-1)} \leq \liminf_{x \rightarrow 0} \frac{|x| \ln \phi(x)}{\ln |x|},$$

which is less accurate than the lower bound in (3.3). It remains as an open question to improve Lemma 3.8 in order to get a precise limit for ϕ as $x \rightarrow x_+$.

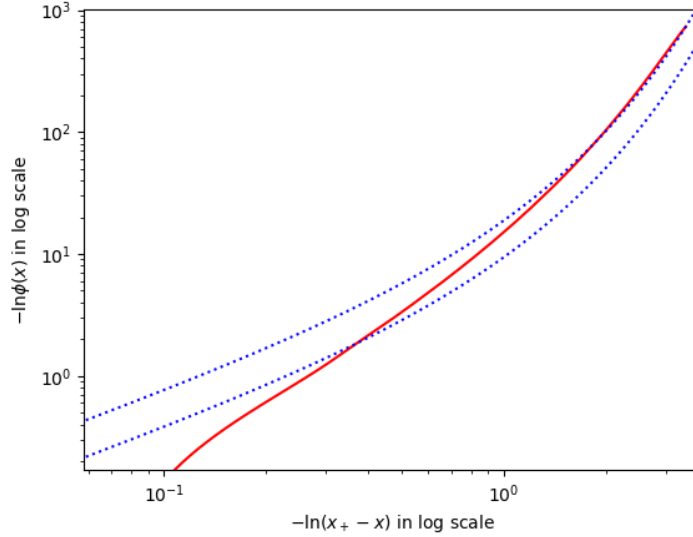


Figure 3.3: In red solid line is plotted $(-\ln|x_+ - x|, -\ln\phi(x))$ in log-log scale, whereas the dotted blue lines correspond to the theoretical bounds given in Theorem 3.1. Here we took a partition of $N = 10^5$ points, and we fixed the parameter $b = 3$.

3.4 Hyperbolic boundaries

In this subsection we address the scenario where the minimal invariant set $M = [x_-, x_+]$ is such that x_+ is a hyperbolic fixed point for h_+ . In other words, $h_+(x_+) = x_+$ and $h'_+(x_+) = \lambda$ with $\lambda \in (0, 1)$. As usual assume without loss of generality that $x_+ = 0$. In this case,

$$h_+(x) = \lambda x + o(x),$$

$$h_+^{-1}(x) = \frac{1}{\lambda}x + o(x),$$

As before, we derive asymptotic bounds for the stationary density ϕ as $x \rightarrow 0^-$. To achieve this goal, we follow the same strategy as in the previous section, in order to obtain an asymptotic expression for the first-hitting time $n_x^y = \min\{n \in \mathbb{N}_0 \mid h_+^n(x) > y\}$. This is done in a straightforward manner in the following lemma.

Lemma 3.12 *For all negative $x \in M$,*

$$\lim_{y \rightarrow 0^-} \frac{n_x^y}{\ln|y|} = \frac{1}{\ln \lambda}$$

Proof. Fix an arbitrary $\epsilon < \min\{\lambda, 1 - \lambda\}$. Thus, there is $x_0 < 0$ such that for all $x \in (x_0, 0)$,

$$(\lambda - \epsilon)x < h_+(x) < (\lambda + \epsilon)x.$$

More generally, since $h_+(x) > x$ for all $x \in M$, if $x \in (x_0, 0)$ then so is $h_+^n(x)$ for all $n \in \mathbb{N}$. Hence, for each $n \in \mathbb{N}$

$$(\lambda - \epsilon)^n x < h_+^n(x) < (\lambda + \epsilon)^n x. \quad (3.28)$$

Fix $x \in (x_0, 0)$ and let y be an arbitrary point in the interval $(x, 0)$. From (3.28), $n_x^y \leq n$ if $(\lambda - \epsilon)^n x > y$. The minimum value of n which satisfies the previous inequality is

$$n = \left\lfloor \frac{\ln |y| - \ln |x|}{\ln(\lambda - \epsilon)} \right\rfloor.$$

It follows that for all $x \in (x_0, 0)$,

$$n_x^y \leq \frac{\ln |y| - \ln |x|}{\ln(\lambda - \epsilon)}. \quad (3.29)$$

Analogously, $n_x^y \geq n$ if $(\lambda + \epsilon)^n x \leq y$. The maximum value of n for which the previous inequality holds is

$$n = \left\lceil \frac{\ln |y| - \ln |x|}{\ln(\lambda + \epsilon)} \right\rceil.$$

Hence, for all $x \in (x_0, 0)$,

$$n_x^y \geq \frac{\ln |y| - \ln |x|}{\ln(\lambda + \epsilon)}. \quad (3.30)$$

If we divide by $\ln |y|$, from (3.29) and (3.30) it follows that for all $x \in (x_0, 0)$

$$\frac{1}{\ln(\lambda - \epsilon)} \leq \liminf_{y \rightarrow 0^-} \frac{n_x^y}{\ln |y|} \leq \limsup_{y \rightarrow 0^-} \frac{n_x^y}{\ln |y|} \leq \frac{1}{\ln(\lambda + \epsilon)} \quad (3.31)$$

For any other $x \in [x_-, x_0]$, it is true that $n_x^y = n_x^{x_0} + n_{h_+(x_0)}^y$, so that by dividing by $\ln |y|$, inequality (3.31) is valid for all negative $x \in M$. Furthermore, since this is true for all $\epsilon > 0$,

$$\lim_{y \rightarrow 0^-} \frac{n_x^y}{\ln |y|} = \frac{1}{\ln \lambda}. \quad (3.32)$$

□

Let us proceed with the second part of Theorem 3.1.

Proof of Theorem 3.1 (hyperbolic case). We use again Lemma 3.8: there are $K_1, K_2 > 0$ such that for all $x > h_+(-2\sigma)$

$$K_1(x - h_+^{-1}(x))^2 \cdot \phi(h_+^{-1}(x)) \leq \phi(x) \leq K_2(x - h_+^{-1}(x)) \cdot \phi(h_+^{-1}(x)). \quad (3.33)$$

Let ϵ be an arbitrary positive number which is smaller than $(1 - \lambda)/\lambda$, and choose $x_0 < 0$ such that for all $x \in (x_0, 0)$

$$\left(\frac{1 - \lambda - \epsilon\lambda}{\lambda}\right) \cdot |x| \leq x - h_+^{-1}(x) \leq \left(\frac{1 - \lambda + \epsilon\lambda}{\lambda}\right) \cdot |x|, \quad (3.34)$$

$$\left(\frac{1}{\lambda} - \epsilon\right) \cdot |x| \leq |h_+^{-1}(x)| \leq \left(\frac{1}{\lambda} + \epsilon\right) \cdot |x| \quad (3.35)$$

We first obtain the lower bound in (3.4). From inequalities (3.33) and (3.34) we have

$$\phi(x) \geq \tilde{K}_1 \cdot |x|^2 \phi(h_+^{-1}(x)),$$

where $\tilde{K}_1 = K_1 \cdot (1 - \lambda - \epsilon\lambda)^2/\lambda^2 > 0$. More generally, if $x > h_+^n(x_0)$ for some $n \in \mathbb{N}$, by iterating the last inequality n times we have

$$\phi(x) \geq \tilde{K}_1^n \cdot |x|^2 \cdots |h_+^{-n+1}(x)|^2 \cdot \phi(h_+^{-n}(x)). \quad (3.36)$$

Notice that, as long as $x > h_+^{-n}(x_0)$, it follows from (3.35) that

$$|h_+^{-1}(x)|^j \geq \left(\frac{1}{\lambda} - \epsilon\right)^j |x|^j$$

for $j \leq n$. Hence, by substituting in (3.36) we obtain

$$\phi(x) \geq \tilde{K}_1^n \cdot \left(\frac{1}{\lambda} - \epsilon\right)^{n(n-1)} |x|^{2n} \cdot \phi(h_+^{-n}(x)).$$

Notice the largest n for which the previous lines hold is $n_{x_0}^x - 1$. Since ϕ is decreasing in $[-2\sigma, 0]$ (see Proposition 3.7), then $\phi(h_+^{-n}(x)) > \phi(h_+(x_0))$ for $n = n_{x_0}^x - 1$.

By taking logarithms, dividing by $(\ln |x|)^2$, and taking $\liminf_{x \rightarrow 0^-}$, it follows from Lemma 3.12 that

$$\liminf_{x \rightarrow 0^-} \frac{\ln \phi(x)}{(\ln |x|)^2} \geq \frac{1}{\ln \lambda} + \frac{\ln(1 - \epsilon\lambda)}{(\ln \lambda)^2}.$$

Since this holds for arbitrarily small ϵ , the lower bound in (3.4) holds.

For the upper bound in (3.4), we proceed analogously: for each $n \in \mathbb{N}$ such that $x > h_+^n(x_0)$, where x_0 is close enough to 0 so that (3.34) and (3.35), we have

$$\phi(x) \leq \tilde{K}_2 \cdot |x| \cdots |h_+^{-n+1}(x)| \cdot \phi(h_+^{-n}(x)),$$

where $\tilde{K}_2 = K_2 \cdot (1 - \lambda + \epsilon\lambda)/\lambda$. From (3.35), it follows that for each $j \leq n - 1$

$$|h_+^j(x)| \leq \left(\frac{1}{\lambda} + \epsilon\right)^j |x|.$$

Hence, by substituting it in the last inequality, we obtain

$$\phi(x) \leq \tilde{K}_1^n \left(\frac{1}{\lambda} + \epsilon\right)^{\frac{n(n-1)}{2}} \phi(h_+^{-n}(x)).$$

Again, for $n = n_{x_0}^x - 1$, by taking logarithms, dividing by $(\ln |x|)^2$, and taking $\limsup_{x \rightarrow 0^-}$ we obtain

$$\limsup_{x \rightarrow 0^-} \frac{\ln \phi(x)}{(\ln |x|)^2} \leq \frac{1}{2 \ln \lambda} + \frac{\ln(1 + \epsilon\lambda)}{2(\ln \lambda)^2}$$

due to Lemma 3.12. Since the last inequality is valid for any positive ϵ small enough, the upper bound in (3.4) holds. $\square$

Remark 3.13 *Just as in Remark 3.9, the stationary density does not exhibit a precise exponential behaviour of the form $\phi(x) \approx e^{-1/|x|^p}$ for any $p > 0$. Indeed,*

$$|x|^p \ln \phi(x) = (|x|^{p/2} \ln |x|)^2 \cdot \frac{\ln \phi(x)}{(\ln |x|)^2}.$$

From Theorem 3.1, the second factor remains bounded as $x \rightarrow 0^-$, whereas $|x|^{p/2} \ln |x| \rightarrow 0$ as $x \rightarrow 0^-$. Because of this, even though ϕ is flat at $x_+ = 0$ it has a subexponential behaviour near x_+ for any $p > 0$.

Let us prove an analogous result for the tail $\mathcal{T}(x) = \mu([x, 0])$, in the case 0 is a hyperbolic fixed point.

Corollary 3.14 *Let $M = [x_-, 0]$ be a minimal invariant set for (3.1) where the deterministic map T is a C^1 -diffeomorphism, and let μ be the unique stationary measure supported on M , with tail function $\mathcal{T}(x) = \mu([x, 0])$. Then if 0 is a hyperbolic fixed point for h_+ with $h'_+(0) = \lambda \in (0, 1)$, then*

$$\frac{1}{\ln \lambda} \leq \liminf_{x \rightarrow 0} \frac{\ln \mathcal{T}(x)}{(\ln |x|)^2} \leq \limsup_{x \rightarrow 0} \frac{\ln \mathcal{T}(x)}{(\ln |x|)^2} \leq \frac{1}{2 \ln \lambda}. \quad (3.37)$$

Proof. Observe that if $h_+(x) = \lambda x + o(x)$,

$$\frac{\ln \phi(h_+(x))}{(\ln |x|)^2} = \frac{\ln \phi(h_+(x))}{(\ln |h_+(x)|)^2} \cdot \left(1 + \frac{\ln |\lambda + o(1)|}{\ln |x|}\right)^2.$$

The result follows from (3.27). □

3.4.1 Numerical explorations

To conclude this chapter, we obtain a numerical approximation of the stationary density $\phi : M \rightarrow \mathbb{R}$ and we calculate $\ln \phi(x)/(\ln |x_+ - x|)^2$ for values of x close to x_+ .

In order to achieve this we approximate ϕ using Ulam's method as in subsection 3.3.1: let $N \in \mathbb{N}$ be a sufficiently large integer, and we split M into N subintervals of the same length. We construct the stochastic matrix $Q = (Q_{ij})_{i,j=1}^N$ for which its dominant left eigenvector gives the approximation of the stationary density.

For this example we consider the simple random difference equation

$$x_{n+1} = \lambda x_n + \sigma \omega_n,$$

with $\lambda \in (0, 1)$, where as before $(\omega_n)_{n=0}^\infty$ is an i.i.d. sequence of random variables which are uniformly distributed in $[-1, 1]$. Given any $\sigma > 0$, the fixed points for h_- and h_+ are $x_- = -\frac{\sigma}{1-\lambda}$ and $x_+ = \frac{\sigma}{1-\lambda}$. Observe that $h'_+(x_+) = \lambda$.

Let $\gamma_1 = 1/(\ln \lambda)$ and $\gamma_2 = 1/(2 \ln \lambda)$. From Theorem 3.1 we have that for x close to x_+ ,

$$\gamma_1 \lesssim \frac{\ln \phi(x)}{(\ln |x_+ - x|)^2} \lesssim \gamma_2.$$

Therefore,

$$\ln |\gamma_2| + 2 \ln \ln \frac{1}{|x_2 - x|} \lesssim \ln \ln \frac{1}{\phi(x)} \lesssim \ln |\gamma_1| + 2 \ln \ln \frac{1}{|x_2 - x|}.$$

The last expression tells us that the curve $(-\ln |x_+ - x|, -\ln \phi(x))$ in $\log - \log$ scale should not deviate much from the band given by $(|\gamma_i|(\ln |x_+ - x|)^2, -\ln \phi(x))$ as $x \rightarrow x_+$, where this band is defined by straight lines in $\log - \log$ scale. This is depicted in [Figure 3.4](#).

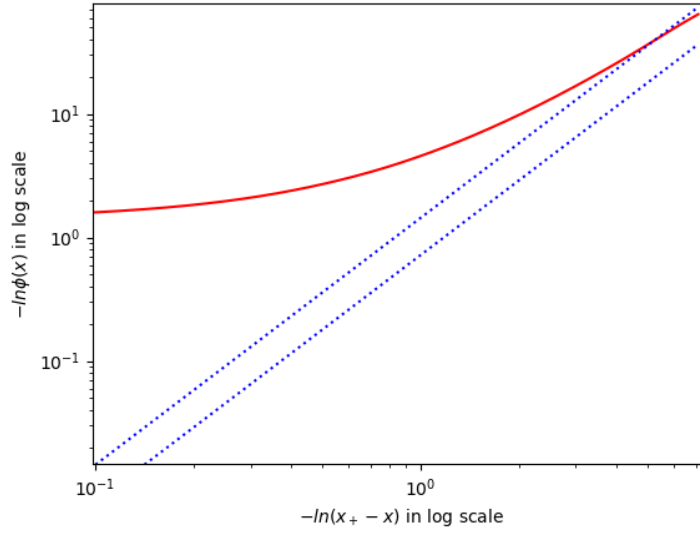


Figure 3.4: In the red solid line is plotted $(-\ln |x_+ - x|, -\ln \phi(x))$ in $\log - \log$ scale, whereas the dotted blue lines correspond to the theoretical bounds given in [Theorem 3.1](#). Here, we took $N = 7500$, and the parameters $\lambda = 0.25$ and $\sigma = 1$.

Chapter 4

Quasi-stationary measures for Random Diffeomorphisms

In this chapter we study the same random difference equation

$$x_{n+1} = T(x_n) + \sigma\omega_n \equiv h(x_n, \omega_n), \quad (4.1)$$

where $T : \mathbb{R} \rightarrow \mathbb{R}$ is a C^1 map with $T' > 0$, and $(\omega_n)_{n=0}^\infty$ is an independently and identically distributed sequence of random variables on $\Delta = [-1, 1]$, whose distribution $d\nu = p$ Leb on $(\Delta, \mathcal{B}(\Delta))$ is such that p is Lipschitz-continuous and positive on $(-1, 1)$. Let φ_ω^n be its associated cocycle map.

Let $M^\sigma = [x_-(\sigma), x_+(\sigma)]$ be a minimal invariant set of (4.1) for each $\sigma \in (0, \sigma^*]$. Assume that for some $\sigma = \sigma^*$, the fixed point $x_-(\sigma^*)$ is a hyperbolic fixed point of the lower extremal map $h_-(x, \sigma^*) = T(x) - \sigma^*$, so that for σ sufficiently close to σ^* the fixed point $x_-(\sigma)$ of $h_-(x, \sigma)$ is again hyperbolic. Consider then for $\sigma > \sigma^*$ the set $M^\sigma = [x_-(\sigma), x^*]$. Notice M^σ is no longer an invariant set for $\sigma > \sigma^*$. In fact, for all points in $(T^{-1}(x^* - \sigma), x^*]$, there is positive probability of escaping from M^σ in one step, see Figure 4.1.

It is important to note that if there is a region in the state space with positive probability of escaping from M^σ in one step, then the system cannot admit a *stationary measure* on M^σ . Because of this, we look instead for *quasi-stationary measures*, which are those that satisfy a stationary-like property when conditioning the system to remain in M^σ . Indeed, if q is a probability measure on

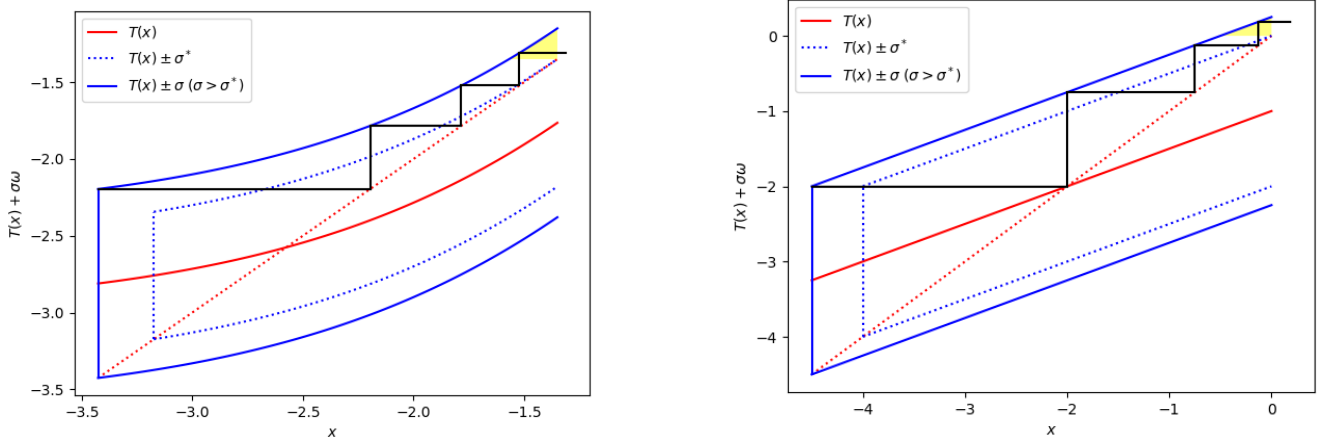


Figure 4.1: On the left, a saddle-node bifurcation occurs for the upper extremal map for some value $\sigma = \sigma^*$. In this case, there is a bifurcation of the minimal invariant set M^{σ^*} . The yellow area indicates the values of x which have positive probability to escape from M^σ in one step. In fact, all points escape from M^σ almost surely. On the right, a similar scenario occurs for a linear system, although in this case there is no bifurcation of minimal invariant sets.

$(M^\sigma, \mathcal{B}(M^\sigma))$, then q is a *quasi-stationary* measure if and only if for all $n \in \mathbb{N}$ and $A \in \mathcal{B}(M^\sigma)$

$$\mathbb{P}_q^n(A \mid n < \tau) := \frac{\int \mathbb{P}[\varphi_\omega(x), \dots, \varphi_\omega^{n-1}(x) \in M^\sigma; \varphi_\omega^n(x) \in A] dq(x)}{\int \mathbb{P}[\varphi_\omega(x), \dots, \varphi_\omega^n(x) \in M^\sigma] dq(x)} = q(A), \quad (4.2)$$

where

$$\tau(x; \omega) = \min\{n \in \mathbb{N} \mid \varphi_\omega^n(x) \notin M^\sigma\} \quad (4.3)$$

is the *escape-time* from M^σ of the cocycle φ_ω^n . Quasi-stationary measures can also be found in the literature as *conditionally stationary measures* [17, 59]. There is a close relationship between quasi-stationary measures and *conditionally invariant measures* for dynamical systems. In fact, q is a quasi-stationary measure if and only if $q \times \mathbb{P}$ is conditionally invariant for the dynamical system $\Phi = (\varphi, \theta)$. For a review on conditionally invariant measures for deterministic dynamics, see [22].

4.1 Main result and structure of the chapter

The main objective of this chapter is to prove there exists a unique quasi-stationary measure q^σ on $(M^\sigma, \mathcal{B}(M^\sigma))$, as given in the next statement.

Theorem 4.1 *Let $M^\sigma = [x_-(\sigma), x^*]$ be such that M^{σ^*} is a minimal invariant set of (4.1), where $x_-(\sigma^*)$ is a hyperbolic fixed point of the lower extremal map $h_-(x, \sigma^*)$. Then, there is $\delta > 0$ such that for each $\sigma \in (\sigma^*, \sigma^* + \delta)$ there exists a unique quasi-stationary measure $q^\sigma \in \mathcal{M}_1(M^\sigma)$. Furthermore, there exist $c > 0$ and $\gamma \in (0, 1)$ such that for all $x \in M^\sigma$ and $n \in \mathbb{N}$,*

$$\|\mathbb{P}_x^n(\cdot \mid n < \tau) - q(\cdot)\|_{TV} \leq c\gamma^n,$$

where $\mathbb{P}_x^n(A \mid n < \tau)$ is the n -step transition probability, conditioned to remain in M^σ up to time n .

Notice how this last theorem is the conditional analogous of Theorem 2.1. Hence, quasi-stationary measures in our setting capture the long term behaviour of the collective behaviour of the system when conditioning it to remain in M .

In Sections 4.2 and 4.3 we introduce the mathematical tools we employ in the rest of the chapter. In particular, in Section 4.2 we define the conditional versions of the stochastic Koopman operator and the Markov operator, while in Section 4.3 we introduce the conditional transfer operator. These play analogous roles to those presented in Appendix C.

In Section 4.4 we prove Theorem 4.1. For this purpose we make use of equivalent conditions to the exponential convergence of $\mathbb{P}_x^n(\cdot \mid n < \tau)$ towards the unique quasi-stationary measure introduced in [14]. We prove this theorem holds by following similar ideas in [13]

In [59] the existence and uniqueness of quasi-stationary measures were proved in a more general context by using perturbation techniques of the transfer operator, although no rate of convergence is provided. Also, in [15] a similar result to Theorem 4.1 was given for maps $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $\|T(x) - x\| < 1$ and when ω_n is uniformly distributed. Here, this last condition is not considered.

In the rest of the chapter we assume $\sigma > \sigma^*$ is fixed, and close to σ^* . We omit all the dependence on σ of the elements we work with (e.g. $q \equiv q^\sigma$, $h_+(x) \equiv h_+(x, \sigma)$, and so on) unless it may cause confusion. Since we restrict the analysis to the set M , we omit the explicit integration region whenever we integrate over M .

4.2 Conditional Koopman and Markov Operators

Consider the compact interval $M = [x_-, x^*]$ as introduced above, and consider the *escape time* from M , $\tau : M \times \Omega \rightarrow \mathbb{N}$, defined as

$$\tau(x; \omega) = \min\{n \in \mathbb{N} \mid \varphi_\omega^n(x) \notin M\}.$$

Note that τ satisfies that

$$\tau^{-1}(n, \infty] = \{(x, \omega) \mid \varphi_\omega^i(x) \in M \text{ for } i = 1, 2, \dots, n\} = \bigcap_{i=0}^n \varphi^{-i}(M),$$

for all $n \in \mathbb{N}$. By continuity of the random map h , the cocycle φ is continuous too, and thus the escape time is a $\mathcal{B}(X \times \Omega), \mathcal{B}(X)$ -measurable function.

Let us begin with a definition analogous to Definition C.7.

Definition 4.2 (Restricted Koopman operator) *The restricted (stochastic) Koopman operator is the linear operator $\hat{P} : L^\infty(M) \rightarrow L^\infty(M)$ defined as¹*

$$\hat{P}f = P(f\mathbf{1}_M)|_M, \tag{4.4}$$

It follows directly from the expression for the regular stochastic Koopman operator (see Proposition C.8) that for all $f \in L^\infty(M)$, $n \in \mathbb{N}$, and $x \in M$,

$$\hat{P}^n f(x) = \int f(\varphi_\omega^n(x)) \cdot \mathbf{1}_M(\varphi_\omega^n(x)) \cdots \mathbf{1}_M(\varphi_\omega(x)) \, d\mathbb{P}(\omega). \tag{4.5}$$

Also, because of Proposition C.8, we get

$$\hat{P}^n f(x) = \int f(y) \hat{k}_n(x, y) dy, \tag{4.6}$$

for all $n \in \mathbb{N}$, where

$$\hat{k}_1(x, y) := \mathbf{1}_M(y) \cdot k(x, y) = \frac{\mathbf{1}_M(y)}{\sigma} \cdot p\left(\frac{y - T(x)}{\sigma}\right) \tag{4.7}$$

¹Here, the function $f\mathbf{1}_M$ is the extension of f such that $f(x) = 0$ for all $x \notin M$.

which follows from (C.4), and

$$\hat{k}_{n+1}(x, y) = \int \hat{k}(x, z) \hat{k}_n(z, y) dz. \quad (4.8)$$

It is also straightforward to see that $\|\hat{P}\| \leq 1$.

Let us show that $(\hat{P}^n f)_{n=1}^\infty$ is a uniformly Lipschitz-continuous sequence for each $f \in L^\infty$.

Lemma 4.3 *there exists $D \equiv D^\sigma > 0$ such that for all $f \in L^\infty(M)$ and all $n \in \mathbb{N}$, we have*

$$\text{Lip}(\hat{P}^n f) \leq D \|f\|_\infty. \quad (4.9)$$

Proof. Let $x, y \in M$ and $f \in L^\infty(M)$. Since p and T are Lipschitz-continuous, it follows from (4.6) that

$$\begin{aligned} |\hat{P}f(x) - \hat{P}f(y)| &\leq \int |k(x, z) - k(y, z)| \cdot |f(z)| \cdot \mathbf{1}_M(z) dz \\ &\leq \frac{\text{Lip}(p)}{\sigma^2} \cdot \|f\|_\infty \cdot \int |T(x) - T(y)| \mathbf{1}_{F(x) \cap M}(z) dz \\ &\leq \frac{2\text{Lip}(p) \cdot \text{Lip}(T)}{\sigma} \|f\|_\infty |x - y|. \end{aligned}$$

The result for $n = 1$ follows by taking $D = 2\text{Lip}(p) \cdot \text{Lip}(T)/\sigma$.

Now, for $n > 1$, observe that

$$\text{Lip}(\hat{P}^{n+1}f) = \text{Lip}(\hat{P}\hat{P}^n f) \leq D \|\hat{P}^n f\|_\infty \leq D \|f\|_\infty,$$

since $\|\hat{P}\|_\infty \leq 1$. □

Not only does the operator $\hat{P}$ map the space $L^\infty(M)$ to the space of Lipschitz-continuous functions, but also leaves the space of strictly positive continuous functions invariant.

Lemma 4.4 *If $f \in L^\infty(M)$ is positive everywhere then so is $\hat{P}f$. In other words, the restricted Koopman operator $\hat{P}$ is a strictly positive operator.*

Proof. Recall that a positive function on a set with positive measure has positive integral. For $x \in M$ fixed, $f(y)k(x, y) > 0$ for all $y \in \text{int}(\hat{F}(x))$. Since $\hat{P}f(x) = \int_{\hat{F}(x)} f(y)k(x, y)dy$, the result follows. □

In order to calculate the probability of escaping from M in more than n steps, starting at some point $x \in M$, we can make use of iterates of $\hat{P}$. Indeed, denote such probability as

$$\mathbb{P}_x(n < \tau) := \mathbb{P}(\{\omega \mid \varphi_\omega^i(x) \in M \text{ for } i = 0, 1, \dots, n\}).$$

It then follows that

$$\mathbb{P}_x(n < \tau) = \hat{P}^n \mathbf{1}_M(x). \quad (4.10)$$

These probabilities of remaining in M are lower the closer we get to x^* . This is precisely formulated in the next proposition.

Proposition 4.5 *For each $n \in \mathbb{N}$ and σ above but close enough to σ^* , the functions $\psi_n(x) = \hat{P}^n \mathbf{1}_M(x) = \mathbb{P}_x[n < \tau]$ are Lipschitz-continuous, strictly positive, and non-increasing in x .*

Proof. Since $\psi_n(x) = \hat{P}^n \mathbf{1}_M(x)$ for all $n \in \mathbb{N}$, the Lipschitz-continuity follows from Lemma 4.3, whereas the strict positivity follows directly from Lemma 4.4.

We prove by induction that the functions ψ_n are non-increasing for each $n \in \mathbb{N}$. For $n = 1$, there is $y_1 \equiv y_1(\sigma) < x^*$ such that $T(y_1) + \sigma = x^*$ for all σ close to σ^* , that is $y_1 = T^{-1}(x^* - \sigma) \equiv h_+^{-1}(x^*)$. Then,

$$\psi_1(x) = \begin{cases} 1, & \text{if } x \leq y_1 \\ \int_{-1}^{\omega_x} p(\omega_0) d\omega_0, & \text{if } x > y_1, \end{cases} \quad (4.11)$$

where $\omega_x \in \Delta$ is the value for which $T(x) + \sigma\omega_x = x^*$. Since T is increasing, ψ_1 is non-increasing.

Assume now the result is true for $n \in \mathbb{N}$. Since $\psi_{n+1} = \hat{P}\psi_n$, then for all $x \in M$

$$\psi_{n+1}(x) = \int_{-1}^1 \psi_n(h_{\omega_0}(x)) \cdot \mathbf{1}_M(h_{\omega_0}(x)) p(\omega_0) d\omega_0$$

Let $x, y \in M$ such that $x \leq y$, therefore

$$\begin{aligned} \psi_{n+1}(y) - \psi_{n+1}(x) &= \int_{-1}^1 [\psi_n(h_{\omega_0}(y)) \mathbf{1}_M(h_{\omega_0}(y)) - \psi_n(h_{\omega_0}(x)) \mathbf{1}_M(h_{\omega_0}(x)) p(\omega_0)] d\omega_0 \\ &= \int_{-1}^1 [\psi_n(h_{\omega_0}(y)) - \psi_n(h_{\omega_0}(x))] \cdot \mathbf{1}_M(h_{\omega_0}(y)) \cdot p(\omega_0) d\omega_0 + \\ &\quad + \int_{-1}^1 [\mathbf{1}_M(h_{\omega_0}(y)) - \mathbf{1}_M(h_{\omega_0}(x))] \cdot \psi_n(h_{\omega_0}(x)) \cdot p(\omega_0) d\omega_0 \end{aligned}$$

The first term is non-positive by induction hypothesis, since $h_{\omega_0}(y) \geq h_{\omega_0}(x)$ for all $\omega_0 \in \Delta$. The second term is non-positive since $h_{\omega_0}(y) \in M$ implies $h_{\omega_0}(x) \in M$. $\square$

Remark 4.6 *Observe that for each $n \in \mathbb{N}$ there exists $\alpha_n > 0$ such that $\mathbb{P}_x[n < \tau] \geq \alpha_n$ for all $x \in M$. In fact, by the previous result, $\alpha_n = \mathbb{P}_{x^*}[n < \tau]$.*

As opposed to the classical stochastic Koopman operator, the measures

$$\hat{\mathbb{P}}_x^n(A) := \hat{P}^n \mathbb{1}_A(x), \quad (4.12)$$

are not necessarily probability measures, although they are finite measures of full measure less than 1. In order to obtain full measure 1 in M , we need to renormalize it. In other words, we define the conditional probabilities

$$\mathbb{P}_x^n(A \mid n < \tau) := \frac{\hat{\mathbb{P}}_x^n(A)}{\hat{\mathbb{P}}_x^n(M)} = \frac{\hat{P}^n \mathbb{1}_A(x)}{\hat{P}^n \mathbb{1}_M(x)} \quad (4.13)$$

The probability measures (4.13) quantify the chance of being in the set A at time n given that the system has remained in M up to that moment.

In an analogous way as in (C.8), we can start with an arbitrary initial distribution ϑ , as opposed to a point-mass distribution δ_x , so that

$$\mathbb{P}_\vartheta^n(A \mid n < \tau) := \frac{\hat{\mathbb{P}}_\vartheta^n(A)}{\hat{\mathbb{P}}_\vartheta^n(M)}, \quad (4.14)$$

where $\mathbb{P}_\vartheta = \int \mathbb{P}_x d\vartheta(x)$. Like in the classical sense, observe that for every $n \in \mathbb{N}$, the integral with respect to the measures $\hat{\mathbb{P}}_x^n$ is

$$\int f d\hat{\mathbb{P}}_x^n = \hat{P}^n f(x). \quad (4.15)$$

Expression (4.14) motivates the following definition.

Definition 4.7 (Quasi-stationary measure) *A probability measure q on $(M, \mathcal{B}(M))$ is a quasi-stationary measure for the system (4.1) if for all $A \in \mathcal{B}(M)$*

$$q(A) = \frac{\hat{\mathbb{P}}_q(A)}{\hat{\mathbb{P}}_q(M)}. \quad (4.16)$$

The right hand side of (4.16) defines a nonlinear operator $\bar{\mathcal{P}}$ on the space of probability measures on $(M, \mathcal{B}(M))$.

Definition 4.8 (Conditional Markov operator) *The conditional Markov operator is the operator $\bar{\mathcal{P}} : \mathcal{M}_1(M) \rightarrow \mathcal{M}_1(M)$ defined as*

$$\bar{\mathcal{P}}\vartheta(A) = \frac{\hat{\mathbb{P}}_\vartheta(A)}{\hat{\mathbb{P}}_\vartheta(M)}, \quad (4.17)$$

for all $A \in \mathcal{B}(M)$.

Notice that the conditional Markov $\bar{\mathcal{P}}$ is nonlinear, as opposed to the nonconditional case. However, it is closely related to the linear operator appearing in the numerator (4.17).

Definition 4.9 (Restricted Markov operator) *The restricted Markov operator is the linear operator $\hat{\mathcal{P}} : \mathcal{M}_1(M) \rightarrow \mathcal{M}(M)$, where $\mathcal{M}(M)$ is the space of finite measures on $(M, \mathcal{B}(M))$, defined as*

$$\hat{\mathcal{P}}\vartheta := \hat{\mathbb{P}}_\vartheta. \quad (4.18)$$

It follows from its definition that a measure q on $(M, \mathcal{B}(M))$ is a quasi-stationary measure if and only if it is a fixed point of the conditional Markov operator $\bar{\mathcal{P}}$. Equivalently, q is an eigenmeasure of the restricted Markov operator $\hat{\mathcal{P}}$, so that

$$\hat{\mathcal{P}}q = \left(\hat{\mathcal{P}}q(M) \right) \cdot q \equiv \Lambda q,$$

where its eigenvalue $\Lambda \equiv \Lambda^\sigma \in (0, 1)$.

From (4.13), q is a quasi-stationary measure if $q(A) = \hat{\mathbb{P}}_q(A \mid 1 < \tau)$ for all $A \in \mathcal{B}(M)$. In [16], q is defined to be a quasi-stationary measure if $q(A) = \hat{\mathbb{P}}_q^n(A \mid n < \tau)$ for all $n \in \mathbb{N}$. This shall follow from the following proposition.

Proposition 4.10 *For all $n \in \mathbb{N}$, and $\vartheta \in \mathcal{M}_1(M)$,*

$$\bar{\mathcal{P}}^n \vartheta = \frac{\hat{\mathcal{P}}^n \vartheta}{\hat{\mathcal{P}}^n \vartheta(M)} \equiv \hat{\mathbb{P}}_\vartheta^n(\cdot \mid n < \tau) \quad (4.19)$$

Proof. Let $\vartheta \in \mathcal{M}_1(M)$. By definition, (4.19) holds for $n = 1$, so in order to prove the statement by induction let us assume (4.19) holds for some $k \in \mathbb{N}$. Then,

$$\bar{\mathcal{P}}^{k+1}\vartheta = \frac{\hat{\mathcal{P}}\left(\frac{\hat{\mathcal{P}}^k\vartheta}{\hat{\mathcal{P}}^k\vartheta(M)}\right)}{\hat{\mathcal{P}}\left(\frac{\hat{\mathcal{P}}^k\vartheta}{\hat{\mathcal{P}}^k\vartheta(M)}\right)(M)} = \frac{\hat{\mathcal{P}}^{k+1}\vartheta}{\hat{\mathcal{P}}^{k+1}(M)}$$

□

Remark 4.11 Proposition 4.10 allows us to interpret the operator $\bar{\mathcal{P}}$ as the evolution of probability measures, by conditioning the system to remain in M at each step. On the other hand, $\hat{\mathcal{P}}$ quantifies the proportion of the initial distribution that remains in M after one step. In particular, this proportion remains constant at each step when we take a quasi-stationary measure as our initial distribution, and it is given by the eigenvalue Λ .

Notice that given any initial distribution $\vartheta \in \mathcal{M}_1(M)$, the probability measures $\bar{\mathcal{P}}^n\vartheta$ are absolutely continuous with respect to Leb . In the next section we address the problem of finding their densities.

4.3 Conditional transfer operator

In a similar way as for the nonconditional case, we study the evolution of densities via an operator defined on $L^1(M)$.

Definition 4.12 (Restricted transfer operator) *The restricted transfer operator $\hat{L} : L^1(M) \rightarrow L^1(M)$ is such that for any $g \in L^1(M)$ and any $f \in L^\infty(M)$ we have*

$$\int f(x) \cdot \hat{L}g(x)dx = \int \hat{P}f(x) \cdot g(x)dx. \quad (4.20)$$

It is possible to give an explicit formula for the action of $\hat{L}$ on $L^1(M)$. Indeed, given $g \in L^1(M)$ and arbitrary $f \in L^\infty(M)$ it follows from (4.6) and Fubini's theorem that

$$\int f(x) \cdot \hat{L}g(x)dx = \int \left(\int f(y)\hat{k}(x, y)dy \right) \cdot g(x)dx = \int f(y) \cdot \left(\int g(x)\hat{k}(x, y)dx \right) dy.$$

Since the last expression holds for all $f \in L^\infty(M)$, then

$$\hat{L}g(x) = \int \hat{k}(y, x)g(y)dy \quad (4.21)$$

Observe that the space of densities on M ,

$$\mathcal{D}(M) = \{g \in L^1(M) \mid g \geq 0; \|g\|_{L^1} = 1\},$$

is no longer invariant under the action of $\hat{L}$. In order for this to happen we renormalize the outcomes, as in the next definition.

Definition 4.13 (Conditional transfer operator) *The conditional transfer operator is the non-linear operator $\bar{L} : \mathcal{D}(M) \rightarrow \mathcal{D}(M)$ defined as*

$$\bar{L}g(x) = \frac{\hat{L}g(x)}{\int \hat{L}g(x)dx}. \quad (4.22)$$

Observe that if $\vartheta \in \mathcal{M}_1(M)$ and $d\vartheta = g \, d\text{Leb}$, then for all $A \in \mathcal{B}(M)$ it follows from (4.20)

$$\int_A \hat{L}g(x)dx = \int \hat{P}\mathbf{1}_A(x) \cdot g(x)dx = \int \hat{\mathbb{P}}_x(A)d\vartheta(x) = \hat{P}\vartheta.$$

This implies that $\bar{L}g$ is the density of $\bar{\mathcal{P}}\vartheta$. In general, for any $n \in \mathbb{N}$ and $g \in L^1(M)$, due to Proposition 4.10 we have

$$\bar{L}^n g(x) = \frac{\hat{L}^n g(x)}{\int \hat{L}^n g(x)dx}. \quad (4.23)$$

In particular, for all $n \in \mathbb{N}$, the densities $\bar{k}_n$ of $\hat{\mathbb{P}}_x^n(\cdot \mid n < \tau)$ satisfy

$$\bar{k}_n(x, y) = \frac{\hat{k}_n(x, y)}{\int \hat{k}_n(x, y)dy}, \quad (4.24)$$

where $\hat{k}_n$ is given by (4.8), and

$$\hat{k}_{n+1}(x, y) = \hat{L}^n(k(x, \cdot))(y) = \int \hat{k}(x, z)\hat{k}_n(x, z) \, dz$$

The corresponding notion of equilibrium for densities is given by the following definition.

Definition 4.14 (Quasi-stationary density) *A function $\hat{\phi} \in \mathcal{D}(M)$ is a quasi-stationary density if it is a fixed point for $\bar{L}$.*

Note that $\hat{\phi}$ is a quasi-stationary density if and only if it is an eigenfunction of $\hat{L}$, where its eigenvalue is $\Lambda = \mathbb{P}_q(M) = \int \phi(x)dx$. Equivalently, $\hat{\phi}$ is a quasi-stationary density if and only if the measure $dq = \hat{\phi} d\text{Leb}$ is a quasi-stationary measure.

4.4 Exponential convergence to quasi-stationary measures

Recall we assume that there exists $\sigma^* > 0$ such that $M^* \equiv M^{\sigma^*} = [x_-(\sigma^*), x^*]$ is a minimal invariant set for (4.1) at $\sigma = \sigma^*$, where its left boundary $x_-(\sigma^*)$ is a hyperbolic fixed point for $h_-(x, \sigma^*)$ with $T'(x_-(\sigma^*)) < 1$. Moreover, since $T'(x_-(\sigma^*)) = \lambda(\sigma^*) \in (0, 1)$, by the implicit function theorem there is $x_- : [\sigma^*, \sigma^* + \delta_1] \rightarrow \mathbb{R}$ such that

$$\begin{aligned} T(x_-(\sigma)) - \sigma &= x_-(\sigma), \\ \lambda(\sigma) &= T'(x_-(\sigma)) \in (0, 1), \end{aligned} \tag{4.25}$$

for all $\sigma \in [\sigma^*, \sigma^* + \delta_1]$, where x_- is C^1 and λ is continuous and bounded away from 1. This last property implies that $x_- \equiv x_-(\sigma)$ is an attracting hyperbolic fixed point for $h_-(\cdot, \sigma)$ for each $\sigma \in [\sigma^*, \sigma^* + \delta_1]$. Moreover,

$$x'_-(\sigma) = -\frac{1}{1 - \lambda(\sigma)} < 0,$$

so x_- is decreasing.

When considering the set $M = [x_-, x^*]$, observe that the only route for escaping from M is through its right boundary x^* , that is $h_{\omega_0}(x) \geq x_-$ for all $x \in M$ and $\omega_0 \in [-1, 1]$.

In order to prove the main theorem of this chapter we employ the following auxiliary result.

Theorem 4.15 (Existence and uniqueness of quasi-stationary measures, [14]) *Assume that $\tau < \infty$ $\mathbb{P}_x$ -a.s. for every $x \in M$, and for all $n \in \mathbb{N}$ and $x \in M$ we have $\mathbb{P}_x(n < \tau) > 0$. Then, there exists a probability measure q on M and two constants $c > 0$ and $\gamma \in (0, 1)$ such that, for all $x \in M$ and $n \in \mathbb{N}$,*

$$\|\mathbb{P}_x^n(\cdot \mid n < \tau) - q(\cdot)\|_{TV} \leq c\gamma^n$$

if and only if the following condition holds: there exist a probability measure $\nu_0 \in \mathcal{M}_1(M)$, $n_0 \in \mathbb{N}$, and constants $c_1, c_2 > 0$ such that

(A) for all $x \in M$ and $B \in \mathcal{B}(M)$

$$\mathbb{P}_x^{n_0}(B \mid n_0 < \tau) \geq c_1 \nu_0(B), \quad (4.26)$$

(B) for all $x \in M$ and $m \in \mathbb{N}$

$$\mathbb{P}_{\nu_0}(m < \tau) \geq c_2 \hat{\mathbb{P}}_x(m < \tau). \quad (4.27)$$

In this case q is the unique quasi-stationary measure for the system (4.1).

We shall refer to the conditions (A) and (B) above as the *Champagnat-Villemonais conditions*.

Let us consider the escape-time from M for the deterministic map $h_+(x) = T(x) + \sigma$, that is the function $n_+ : M \rightarrow \mathbb{N}$ defined as

$$n_+(x) = \min\{n \geq 0 \mid h_+^n(x) > x^*\},$$

and let $N_0 = n_+(x_-)$. We use N_0 to prove in the next lemma that the orbits of (4.1) escape $\mathbb{P}_x$ -a.s. for all $x \in M$.

Lemma 4.16 *For all $x \in M$, we have $\mathbb{P}_x(\tau = \infty) := \mathbb{P}[\omega \in \Omega \mid \tau(x; \omega) = \infty] = 0$.*

Proof. Since h_+ is continuous and M^c is open, there is $\epsilon > 0$ small such that

$$I^{N_0} \times \Omega \subset \{\omega \in \Omega \mid \tau(x_-, \omega) = N_0\},$$

where $I = [1 - \epsilon, 1]$. Hence, $\mathbb{P}_{x_-}(N_0 = \tau) \geq \epsilon^{N_0}$. Furthermore, because of Proposition 4.5, for all $x \in M$

$$\mathbb{P}_x(N_0 < \tau) \leq \mathbb{P}_{x_-}(N_0 < \tau) = 1 - \mathbb{P}_{x_-}(\tau = N_0) \leq 1 - \epsilon^{N_0} < 1.$$

It follows that

$$\mathbb{P}_x(2N_0 < \tau) = \hat{P}^{N_0}(P^{N_0} \mathbf{1}_M)(x) \leq (1 - \epsilon^{N_0}) \hat{P}^{N_0} \mathbf{1}_M(x) \leq (1 - \epsilon^{N_0})^2.$$

In general, for each $k \in \mathbb{N}$,

$$\mathbb{P}_x(kN_0 < \tau) \leq (1 - \epsilon^{N_0})^k.$$

Therefore

$$\mathbb{P}_x(\tau = \infty) = \lim_{k \rightarrow \infty} \mathbb{P}_x[kN_0 < \tau] = 0.$$

□

We proceed now with the proof of the main theorem. Our techniques are based on those in [13].

Proof of Theorem 4.1. Notice that Lemma 4.16 states that τ is finite $\mathbb{P}_x$ -a.s. for all $x \in M$. On the other hand, Remark 4.6 guarantees that for all $n \in \mathbb{N}$ and $x \in M$, the probabilities $\mathbb{P}_x(n < \tau)$ are positive. By Theorem 4.15, it suffices to show that the Champagnat-Villemonais conditions hold.

We split the proof of this theorem into four steps. First, we prove a technical result which is used later in the proof. Later on, we consider $\delta > 0$ for which we can find $n_0 \in \mathbb{N}$ needed in Condition (A). Afterwards, we show that Condition (A) in Theorem 4.15 is fulfilled for some $c_1 > 0$, and for some probability measure ν_0 on $(M, \mathcal{B}(M))$, with n_0 as found before. Finally, we prove that Condition (B) is fulfilled for ν_0 , and some $c_2 > 0$. The rest of the statement follows then by Theorem 4.15.

Claim 1: for each $n \in \mathbb{N}$ there exists $D'' \equiv D''_n$ such that for all $m \geq n$,

$$\text{Lip}(\hat{P}^m \mathbf{1}_M) \leq D'' \|\hat{P}^m \mathbf{1}_M\|_\infty$$

Indeed, let $n \in \mathbb{N}$ and $m \geq n$. Then, for all $x \in M$,

$$\hat{P}^m \mathbf{1}_M(x) = \int \hat{P}^n \mathbf{1}_M(y) d\hat{\mathbb{P}}_x^{m-n}(y) \geq \alpha_n \cdot \hat{\mathbb{P}}_x^{m-n}(M) = \alpha_n \cdot \hat{P}^{m-n} \mathbf{1}_M(x),$$

where $\alpha_n = \mathbb{P}_{x^*}^n(n < \tau)$ as in Remark 4.6. In other words,

$$\|\hat{P}^{m-n} \mathbf{1}_M\|_\infty \leq \frac{1}{\alpha_n} \|\hat{P}^m \mathbf{1}_M\|_\infty \quad (4.28)$$

Therefore, from (4.28),

$$\text{Lip}(\hat{P}^m \mathbf{1}_M) = \text{Lip}(\hat{P}^n \hat{P}^{m-n} \mathbf{1}_M) \leq D \|\hat{P}^{m-n} \mathbf{1}_M\|_\infty \leq \frac{D}{\alpha_n} \|\hat{P}^m \mathbf{1}_M\|_\infty$$

where $D > 0$ is as in Lemma 4.3. The claim follows by taking $D'' = D/\alpha_n$.

Claim 2: *there exists $\delta > 0$ such that for each $\sigma \in (\sigma^*, \sigma^* + \delta)$ there is $n_0 \equiv n_0(\sigma)$ for which*

$$h_-^n(x^*) - x_- < \alpha_1^n/D, \quad (4.29)$$

for all $n \geq n_0$, where $\alpha_1 = \mathbb{P}_{x^}[n < \tau]$ like in Remark 4.6, and $D > 0$ is given in Lemma 4.3.*

Since x_- is a hyperbolic fixed point for which $\lambda(\sigma) = T'(x_-) \in (0, 1)$ for all $(\sigma^*, \sigma^* + \delta_1)$ (see 4.25), there are $C \equiv C_\sigma > 0$ and $\rho \equiv \rho(\sigma) \in (0, 1)$ such that

$$h_-^n(x^*) - x_- \leq C\rho^n,$$

for all $n \in \mathbb{N}$, where $\rho \in (0, 1)$ for all $\sigma \in [\sigma^*, \sigma^* + \delta_1]$ and it depends continuously on σ (see Proposition D.3 in Appendix D).

Since $\alpha_1 \equiv \alpha_1(\sigma) \rightarrow 1$ as $\sigma \rightarrow \sigma^*$, and $\rho(\sigma) \in (0, 1)$ for all $\sigma \in [\sigma^*, \sigma^* + \delta_1]$, there exists $\delta \in (0, \delta_1)$ such that for all $\sigma \in (\sigma^*, \sigma^* + \delta)$, we have $\rho < \alpha_1$.

Then, for $n_0 \geq N_0 + 1$ large enough we have that

$$C\rho^n < \frac{\alpha_1^n}{D},$$

for all $n \geq n_0$. The claim follows.

Claim 3: *let $I = [h_-^{n_0-1}(x^*), x^*]$ and $\nu_0(A) = \frac{\text{Leb}(A \cap I)}{\text{Leb}(I)}$. Then, there exists $c_1 > 0$ such that for all $A \in \mathcal{B}(M)$, and all $x \in M$,*

$$\hat{\mathbb{P}}_x^{n_0}(A \mid n_0 < \tau) \geq c_1 \nu_0(A).$$

Notice that for each $n \in \mathbb{N}$,

$$\hat{\mathbb{P}}_x^n(A \mid n < \tau) = \int_A \bar{k}_n(x, y) dy,$$

where $\bar{k}_n(x, y) = \hat{k}(x, y) / \int_M \hat{k}(x, y) dy$. Since $\bar{k}_n(x, \cdot)$ is supported on the *conditional set-valued map* $\hat{F}^n(x) := F(x) \cap M$, and for $n = n_0$ we have $\hat{F}^{n_0}(x) = [h_-^{n_0}(x), x^*]$, then we have, $I \subset F^{n_0}(x)$ for all $x \in M$. By continuity, $\bar{k}_{n_0}(x, y) > 0$ for all $x \in M$ and $y \in I$. Therefore,

$$\mathbb{P}_x^{n_0}(A \mid n_0 < \tau) \geq \int_{A \cap I} \bar{k}_{n_0}(x, y) dy \geq \inf_{x \in M, y \in I} \bar{k}_{n_0}(x, y) \cdot \text{Leb}(A \cap I) \equiv \tilde{c}_1 \text{Leb}(A \cap I)$$

Then, we have

$$\mathbb{P}_x^{n_0}(A \mid n_0 < \tau) \geq c_1 \cdot \nu_0(A), \quad (4.30)$$

where $\nu_0(A) = \text{Leb}(A \cap I)/\text{Leb}(I)$ and $c_1 = \tilde{c}_1 \cdot \text{Leb}(I)$. The claim follows.

Claim 4: *there exists $c_2 > 0$ such that for all $x \in M$ and $m \in \mathbb{N}$,*

$$\hat{P}_{\nu_0}(m < \tau) \geq c_2 \hat{\mathbb{P}}_x(m < \tau).$$

For each $m \in \mathbb{N}$, let us define the non-negative function $g : M \rightarrow \mathbb{R}_+$ as

$$g_m(x) = \frac{\hat{P}^m \mathbf{1}_M(x)}{\|\hat{P}^m \mathbf{1}_M\|_\infty}. \quad (4.31)$$

By Claim 1, there is $D'' \equiv D''_{n_0}$ such that for all $m \geq n_0$

$$|g_m(x) - g_m(y)| = \frac{|\hat{P}^m \mathbf{1}_M(x) - \hat{P}^m \mathbf{1}_M(y)|}{\|\hat{P}^m \mathbf{1}_M\|_\infty} \leq D''|x - y|,$$

so $\text{Lip}(g_m) \leq D''$. By compactness, g_m attains its maximum in M (which is 1). Moreover, $\hat{P}^m \mathbf{1}_M$ is nonincreasing (see Proposition 4.5), so $g_m(x_-) = 1$. By Lipschitz-continuity, for all $x \in M$,

$$g_m(x) = 1 - (g_m(x_-) - g_m(x)) \geq 1 - D''(x - x_-).$$

Let $f(x) = \max\{1 - D''(x - x_-), 0\}$. Since g_m is non-negative, we conclude that $g_m \geq f$. Notice that f is continuous, and it is positive in $(x_-, x_- + 1/D'')$. By the choice of n_0 , as given in Claim 2, then $h_-^{n_0}(x_+) < x_- + \frac{1}{D''}$ and then $I \cap (x_-, x_- + \frac{1}{D''})$ is an open interval. Therefore, for all $m \geq n_0$ and $x \in M$,

$$\hat{\mathbb{P}}_{\nu_0}(m < \tau) = \int \hat{P}^m \mathbf{1}_M d\nu_0 = \|\hat{P}^m \mathbf{1}_M\|_\infty \int g_m d\nu_0 \geq \tilde{c}_2 \hat{\mathbb{P}}_x(m < \tau),$$

where $\tilde{c}_2 = \int f d\nu_0 > 0$.

For $m < n_0$, and for all $x \in M$,

$$\mathbb{P}_{\nu_0}(m < \tau) = \int \hat{P} \mathbf{1}_M d\nu_0 \geq \alpha_1^m \geq \alpha_1^{n_0} \geq \alpha_1^{n_0} \hat{\mathbb{P}}_x(m < \tau).$$

The result follows by taking $c_2 = \min\{\tilde{c}_2, \alpha_1^{n_0}\}$. □

Corollary 4.17 *For each $\sigma \in (\sigma^*, \sigma^* + \delta)$ and each density g supported on M , the sequence $(\bar{L}^n g)_{n=0}^\infty$ converges exponentially fast in L^1 as $n \rightarrow \infty$ to the unique quasi-stationary density $\hat{\phi} \equiv \phi^\sigma$.*

Proof. This follows directly from Theorem 4.1, and Proposition A.1. □

Chapter 5

Bifurcations and mean escape times

In the previous chapter we considered a minimal invariant set $M^* = [x_-(\sigma^*), x^*]$ at $\sigma = \sigma^*$, for the system

$$x_{n+1} = h(x_n, \omega_n), \quad (5.1)$$

where $h(x, \omega_0) = T(x) + \sigma\omega_0$, and the sequence $(\omega_n)_{n=0}^\infty$ is an i.i.d. sequence of random variables. We also considered a parametrized family of sets $M^\sigma = [x_-(\sigma), x^*]$, where $x_-(\sigma)$ is the unique fixed point of the lower extremal map $h_-(x, \sigma) = T(x) - \sigma$. We proved the existence and uniqueness of a quasi-stationary measure q^σ on $(M^\sigma, \mathcal{B}(M^\sigma))$, for all σ close enough but above σ^* , for the cocycle φ_ω^n defined by the random difference equation (5.1).

We assume in the rest of this chapter that the deterministic map T is C^∞ , although only a finite degree of differentiability is required for the results presented here. We further assume that each ω_n in the sequence $(\omega_n)_{n=0}^\infty$ is uniformly distributed on $[-1, 1]$.

The main objective of this chapter is to study how likely the orbits of the system (5.1) remain in M^σ , under the assumption that the upper extremal map $h_+(x, \sigma) := T(x) + \sigma$ exhibits a nondegenerate *saddle-node bifurcation* at $(x, \sigma) = (x^*, \sigma^*)$. In other words, the map h_+ satisfies

$$h_+(x^*, \sigma^*) = x^*, \quad \frac{\partial h_+}{\partial x}(x^*, \sigma^*) = 1, \quad \frac{\partial^2 h_+}{\partial x^2}(x^*, \sigma^*) = 2\alpha > 0. \quad (5.2)$$

Note that this is the simplest condition for a *topological bifurcation* of the minimal invariant set M^* to occur [42] (see Figure 5.1).

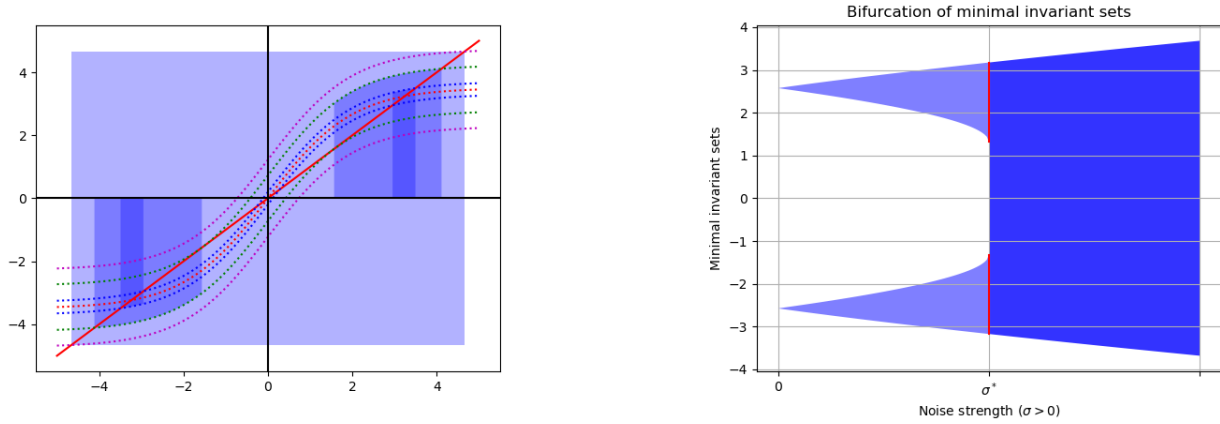


Figure 5.1: On the left, the upper and lower extremal maps for different values of σ are presented, where the shaded regions indicate the minimal invariant sets by projecting them onto the horizontal axis. On the right, the minimal invariant sets are presented as a function of σ . Notice that at a critical value σ^* the union of minimal invariant sets changes discontinuously (in the Hausdorff metric).

For this purpose, we consider the escape time, $\tau : M^\sigma \times \Omega \rightarrow \mathbb{N}$, introduced in Section 4.2 (see (4.18)), and defined as

$$\tau(x; \omega) = \min\{n \in \mathbb{N} \mid \varphi_\omega^n(x) \notin M^\sigma\}. \quad (5.3)$$

When taking the quasi-stationary measure q^σ as initial distribution, it can be proved that τ is geometrically distributed [16, 59], that is there exists $\Lambda^\sigma \equiv \Lambda \in (0, 1)$ such that for all $n \in \mathbb{N}$

$$\mathbb{P}_{q^\sigma}[n < \tau] := \int_M \mathbb{P}_x^n(n < \tau) dq^\sigma(x) = (\Lambda^\sigma)^n. \quad (5.4)$$

It follows then that the *mean escape time* with initial distribution q^σ , $\mathbb{E}_{q^\sigma}[\tau]$, is calculated as

$$\mathbb{E}_{q^\sigma}[\tau] := \int_M \int_\Omega \tau(x; \omega) d\mathbb{P}(\omega) dq^\sigma(x) = \frac{1}{1 - \Lambda^\sigma}. \quad (5.5)$$

In probabilistic terms, $\mathbb{E}_{q^\sigma}[\tau] = \int_M \mathbb{E}_x[\tau] dq^\sigma(x)$.

As it turns out, the rate Λ^σ is precisely the eigenvalue of the restricted Markov operator $\hat{\mathcal{P}}$ introduced in the previous chapter, for which its associated eigenmeasure is the unique quasi-stationary measure, q^σ . Equivalently, it is the eigenvalue for the operator $\hat{L}$ for the quasi-stationary density $\hat{\phi}^\sigma$ (see Figure 5.2).

Since $\Lambda^\sigma \rightarrow 1$ as $\sigma \rightarrow \sigma^*$, then $\mathbb{E}_{q^\sigma}[\tau]$ diverges as $\sigma \rightarrow \infty$, implying that in average the orbits of the system take a long time to abandon M^σ , for which M^σ is called *nearly* or *almost invariant set* [18], or *metastable set* [10]. It is our goal in this chapter to describe in more detail how $\mathbb{E}_{q^\sigma}[\tau]$ diverges as $\sigma \rightarrow \sigma^*$.

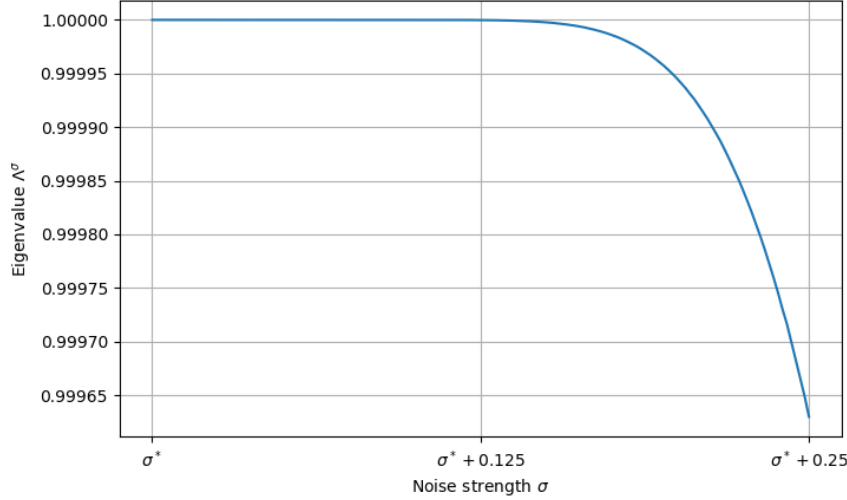


Figure 5.2: The eigenvalue Λ^σ . It does have a flat behaviour at $\sigma = \sigma^*$. We investigate this feature in this chapter.

5.1 Main result and structure of the chapter

The behaviour of $\mathbb{E}_{q^\sigma}[\tau]$, or equivalently the behaviour of Λ^σ , as $\sigma \rightarrow \sigma^*$, is closely related to that of $\hat{\phi}^\sigma(x^*)$ for σ close to σ^* . The way $\hat{\phi}^\sigma(x^*) \rightarrow \hat{\phi}^{\sigma^*}(x^*) = 0$ as $\sigma \rightarrow \sigma^*$ is given in the main result of this chapter stated as follows.

Theorem 5.1 (Asymptotic order of $\hat{\phi}^\sigma(x^*)$ near σ^*) *Let $\hat{\phi}^\sigma$ be the unique quasi-stationary density supported on $M^\sigma = [x_-(\sigma), x^*]$ for $\sigma > \sigma^*$, where $x_-(\sigma)$ is the unique fixed point of the lower extremal map $h_-(x, \sigma) = T(x) - \sigma$, and the upper extremal map $h_+(x, \sigma) = T(x) + \sigma$*

undergoes a nondegenerate saddle-node bifurcation at $(x, \sigma) = (x^*, \sigma^*)$, satisfying (5.2). Then,

$$\frac{\pi}{2\sqrt{\alpha}} \leq \liminf_{\sigma \rightarrow \sigma^*} \frac{\sqrt{\sigma - \sigma^*} \cdot \ln \hat{\phi}^\sigma(x^*)}{\ln(\sigma - \sigma^*)} \leq \limsup_{\sigma \rightarrow \sigma^*} \frac{\sqrt{\sigma - \sigma^*} \cdot \ln \hat{\phi}^\sigma(x^*)}{\ln(\sigma - \sigma^*)} \leq \frac{\pi}{\sqrt{\alpha}} \quad (5.6)$$

This last theorem implies that for each $r > 0$, there is $\delta > 0$ such that for all $\sigma \in (\sigma^*, \sigma^* + \delta)$,

$$(\sigma - \sigma^*)^{\gamma_1/\sqrt{\sigma - \sigma^*}} \leq \hat{\phi}^\sigma(x^*) \leq (\sigma - \sigma^*)^{\gamma_2/\sqrt{\sigma - \sigma^*}},$$

where $\gamma_1 = \frac{\pi}{2\sqrt{\alpha}} + r$, and $\gamma_2 = \frac{\pi}{\sqrt{\alpha}} - r$. Notice that the functions which serve as upper and lower bound are flat at σ^*

We start Section 5.2 by proving in Proposition 5.2 that the escape time τ is geometrically distributed, when taking q^σ as the initial distribution, as given by (5.4). As a direct consequence of this, in Proposition 5.3 we prove that (5.5) holds. Hence, the mean escape time can be studied via the eigenvalue Λ^σ , which in turn is related to the quasi-stationary density ϕ^σ .

We proceed in Section 5.3 by revisiting the idea of approximating discrete dynamical systems with the time-one map of a flow. Here, we make use of the *partial embedding* techniques developed in [36, Chapter X], and conclude in Proposition 5.5 that we can approximate the upper extremal maps with the time-one map of some flows ψ_t^σ , up to an error of order $o((\sigma - \sigma^*)^2)$. The expression of their generating vector fields is given in Proposition 5.6.

Later, in Section 5.4 we consider the parametrized family of flows, $(\psi_t^\sigma)_{t \in \mathbb{R}_+}$ mentioned above, its discretization $(\psi_n^\sigma)_{n \in \mathbb{N}_0}$, and the dynamical system $(h_+^n)_{n \in \mathbb{N}_0}$ it approximates as above. We analyze the first time their orbits abandon the set $\{x \leq x^*\}$, denoted by $t_x(\sigma)$, $m_x(\sigma)$, and $n_x(\sigma)$, respectively. We show in Proposition 5.7, Corollary 5.8, and Lemma 5.9, that

$$t_x(\sigma), m_x(\sigma), n_x(\sigma) \sim (\sigma - \sigma^*)^{-1/2}.$$

These scaling laws have been called in the literature as *bottlenecking* or *saddle-node ghosts* [39]. The scaling law for the escape time of h_+ , $n_x(\sigma)$, has already been found in [25], using complex variable techniques. The method here presented, however, is essentially different, and we present it in full detail.

In Section 5.5, we prove Theorem 5.1 in full detail. The method employed is similar to that used in Section 3.3, but the calculations are somewhat different. Furthermore, as a direct consequence of Theorem 5.1, we prove an analogous asymptotic result for the escape rate Λ^σ in Corollary 5.14.

Last, we portray Corollary 5.14 with numerical explorations. Since Λ^σ is flat at σ^* (see Figure 5.2), we look for logarithmic-like scales in order to visualize the differences between the theoretical predictions and the numerical approximations.

5.2 Mean escape time

In the rest of the chapter we omit the dependence on σ of all the elements we work with, whenever it causes no confusion. We also omit the integration region in $\mathbb{R}$, as we always assume it is M , unless stated otherwise.

Let us consider the unique quasi-stationary measure q on $(M, \mathcal{B}(M))$ for the system (5.1). We first show that the escape time τ , as given in (5.3), is geometrically distributed when we take q as the initial distribution. All the results here presented can be found in the standard literature on quasi-stationary measures, although for the reader's convenience we present a proof for them.

Proposition 5.2 (Geometric distribution of the escape time [16, Theorem 2.2]) *Let Λ be the eigenvalue of q for the restricted Markov operator $\hat{\mathcal{P}}$ (cf. Section 4.2). Then, for all $n \in \mathbb{N}$, we have*

$$\mathbb{P}_q(n < \tau) = \Lambda^n, \quad (5.7)$$

Proof. Recall that $\hat{\mathcal{P}}q = \Lambda q$, and that $\mathbb{P}_q(n < \tau) := \int \mathbb{P}_x(n < \tau) dq(x)$. Then, since $\mathbb{P}_x(n < \tau) = \hat{P}^n \mathbf{1}_M(x)$, where $\hat{P} : L^\infty(M) \rightarrow L^\infty(M)$ is the restricted Koopman operator (see 4.4), we obtain

$$\mathbb{P}_q(n < \tau) = \int \hat{P}^n \mathbf{1}_M(x) dq(x) = \int \hat{\mathbb{P}}_x^n(M) dq(x) = (\hat{\mathcal{P}}^n q)(M) = \Lambda^n q(M) = \Lambda^n \quad \square$$

Because of the previous result, the eigenvalue Λ is called the *rate of survival*. Moreover, Λ can be related to the mean escape time $\mathbb{E}_q[\tau]$ as given in the next proposition.

Proposition 5.3 (Mean escape time [59, Lemma 5.6]) *Let Λ be the eigenvalue of q for the restricted Markov operator $\hat{\mathcal{P}}$. Then*

$$\mathbb{E}_q[\tau] = \frac{1}{1 - \Lambda} \quad (5.8)$$

Proof. Recall that by definition $\mathbb{E}_q[\tau] = \int_M \int_\Omega \tau(x; \omega) d\mathbb{P}(\omega) dq(x)$. From Proposition 5.3, it follows that $\mathbb{P}_q[\tau = n] = \mathbb{P}_q[n - 1 < \tau] - \mathbb{P}_q[n < \tau] = \Lambda^{n-1} - \Lambda^n$. Since τ takes values only in $\mathbb{N}$, we get

$$\mathbb{E}_q[\tau] = \sum_{n=1}^{\infty} n \cdot \int \mathbb{P}_x[\tau = n] dq(x) = \sum_{n=1}^{\infty} n(\Lambda^{n-1} - \Lambda^n) = (1 - \Lambda) \sum_{n=1}^{\infty} n\Lambda^{n-1} = \frac{1}{1 - \Lambda} \quad \square$$

From Proposition 5.2, $\Lambda = \mathbb{P}_q[1 < \tau]$, and recall that M^{σ^*} is a minimal invariant set, so that $\Lambda \rightarrow 1$ as $\sigma \rightarrow \sigma^*$. This in turn implies that $\mathbb{E}_q[\tau] \rightarrow \infty$ as $\sigma \rightarrow \sigma^*$. We devote the rest of the chapter to describe the speed at which these limits occur.

5.3 Approximation of saddle-node families via flows

In this section, we aim to follow the same ideas behind Takens embedding theorem (see Theorem 3.2), in the presence of a nondegenerate saddle-node bifurcation. Let us consider the upper extremal map of (5.1), $h_+(x, \sigma) = T(x) + \sigma$, and assume without loss of generality that $x^* = 0$ is a nondegenerate nonhyperbolic fixed point for $h_+(\cdot, \sigma^*)$. For simplicity, by renaming $\epsilon = \sigma - \sigma^*$, we denote as $h_\epsilon(x) := h_+(x, \sigma^* + \epsilon) = \epsilon + T(x) + \sigma^*$. Therefore, (5.2) is translated as

$$h_0(0) = 0, \quad h'_0(0) = 1, \quad h''_0(0) = 2\alpha > 0.$$

The map h_ϵ can then be written as

$$h_\epsilon(x) = \epsilon + x + \alpha x^2 + o(x^2).$$

The family $(h_\epsilon)_{\epsilon \geq 0}$ is called a *saddle-node family*, in the sense that it is a transversal perturbation of a nondegenerate saddle-node map, this one being h_0 . Despite h_0 being embeddable into a flow ψ_t^0 , as given in Theorem 3.2, for $\epsilon > 0$ we cannot guarantee h_ϵ is embeddable into a flow ψ_t^ϵ in such a way that $\psi_t^\epsilon \rightarrow \psi_t^0$ in the C^r -topology¹, for $r \geq 2$. Instead, we look for a so-called *partial embedding* [36] by employing the following lemma.

¹It is possible to provide with an embedding of a parametrized family of maps into a parametrized family of flows. However, one can only guarantee convergence in C^1 -topology, the most, as $\epsilon \rightarrow 0$, in a neighbourhood of 0. Convergence in C^r -topology, for $r \geq 2$, is guaranteed in compact intervals bounded away from 0, see [33, Proposition 6.1].

Lemma 5.4 ([36, Lemma X.4.3]) *For any $k \in \mathbb{N}$ there exists a natural number $N = N(k)$ such that whenever h is C^N (both in (x, ϵ)), there exists a C^k vector field whose time-one map of the associated flow, ψ_t^ϵ , satisfies*

$$|h_\epsilon(x) - \psi_1^\epsilon(x)| \leq B\epsilon^k, \quad (5.9)$$

for some $B > 0$, and for all $x \in [-x_1, x_1]$ and $\epsilon \in [0, \epsilon_1]$ with x_1 and ϵ_1 small enough.

Notice that for $\epsilon = 0$, the maps ψ_1^0 and h_0 coincide in $[-x_1, x_1]$, so the map h_0 is embeddable.

Let us consider the escape times, $n_x(\sigma^* + \epsilon)$ and $m_x(\sigma^* + \epsilon)$ of the deterministic dynamical systems $(h_\epsilon^n)_{n \geq 0}$ and $(\psi_n^\epsilon)_{n \geq 0}$, respectively. In the rest of the section we denote them as

$$\begin{aligned} n_x(\sigma^* + \epsilon) \equiv n_x(\epsilon) &= \min\{n \geq 0 \mid h_\epsilon^n(x) > 0\}, \\ m_x(\sigma^* + \epsilon) \equiv m_x(\epsilon) &= \min\{n \geq 0 \mid \psi_n^\epsilon(x) > 0\}. \end{aligned} \quad (5.10)$$

In the next proposition, for each x close enough to 0, we give bounds on the discrepancy between h_ϵ^n and ψ_n^ϵ , for as long as $n \leq \min\{n_x(\epsilon), m_x(\epsilon)\}$. These bounds will be useful later on when we study $n_x(\epsilon)$ as $\epsilon \rightarrow 0$.

Proposition 5.5 *Let $h_\epsilon = \epsilon + x + \alpha x^2 + o(x^2)$ as above. Then, there exist constants $B, x_1, \epsilon_1 > 0$, and a flow ψ_t^ϵ for each $\epsilon \in [0, \epsilon_1]$, which is C^3 in $(-x_1, x_1) \times (0, \epsilon_1)$, such that*

$$h_\epsilon(x) = \psi_1^\epsilon(x) + r_\epsilon^1(x), \quad |r_\epsilon^1(x)| \leq B\epsilon^3, \quad (5.11)$$

for all $x \in [-x_1, x_1]$. Furthermore, provided both x_1, ϵ_1 are small enough there is $\kappa > 0$ such that for each $x \in [-x_1, 0]$, $\epsilon \in [0, \epsilon_1]$, and $n \leq \min\{n_x(\epsilon), m_x(\epsilon)\}$, we have

$$h_\epsilon^n(x) = \psi_n^\epsilon(x) + r_\epsilon^n(x), \quad |r_\epsilon^n(x)| \leq \frac{B}{\kappa} [(1 + \kappa\epsilon)^n - 1] \epsilon^2. \quad (5.12)$$

Proof. Since T is a C^∞ map, observe that (5.11) follows directly from Lemma 5.4, by taking $k = 3$. Therefore, since x^* is a nondegenerate nonhyperbolic fixed point satisfying (5.2), and ψ_t^ϵ coincides with h_+ when $\epsilon = 0$, we can assume without loss of generality that for all $(x, \epsilon) \in [-x_1, x_1] \times [0, \epsilon_1]$, we have $\frac{\partial^2}{\partial x^2} \psi_1^\epsilon(x) > 0$.

In the rest of the proof we omit the dependence on ϵ for h_ϵ^n , ψ_n^ϵ , and r_ϵ^n . Let $M_x(\epsilon) = \min\{n_x(\epsilon), m_x(\epsilon)\}$, and assume $M_x(\epsilon) \geq 2$. Since $h^2(x) = \psi_1(h(x)) + r^1(h(x))$, by the mean value theorem, there exists ξ_2 between $\psi_1(x)$ and $h(x)$ such that

$$h_2(x) = \psi_2(x) + \psi_1'(\xi_2)r^1(x) + r^1(h(x)).$$

Since $\psi_1'' > 0$, then $\psi_1'(\xi_2) < \psi_1'(0)$. Moreover, since $h'(0) = 1$ for all ϵ , assuming ϵ_1 is small enough we have that for some $\kappa > 0$

$$0 < 1 - \kappa\epsilon \leq \psi_1'(0) \leq 1 + \kappa\epsilon, \quad (5.13)$$

for all $\epsilon \in [0, \epsilon_1]$. Therefore,

$$h_2(x) \leq \psi_2(x) + \psi_1'(0)r^1(x) + r^1(h(x)) \leq \psi_2(x) + (2 + \kappa\epsilon)B\epsilon^3 = \psi_2(x) + \frac{B}{\kappa}[(1 + \kappa\epsilon)^2 - 1]\epsilon^2.$$

In general,

$$h^n(x) = \psi_1(h^{n-1}(x)) + r^1(h^{n-1}(x)).$$

By the mean value theorem, there exists ξ_n in between $h^{n-1}(x)$ and $\psi_{n-1}(x)$, such that

$$h^n(x) = \psi_n(x) + \psi_1'(\xi_n)r^{n-1}(x) + r^1(h^{n-1}(x)).$$

If $n \leq M_x(\epsilon)$, assuming (5.12) holds for $n - 1$, and due to (5.11) and (5.13), it follows that

$$h^n(x) \leq \psi_n(x) + \frac{B(1 + \kappa\epsilon)}{\kappa}[(1 + \kappa\epsilon)^{n-1} - 1]\epsilon^2 + B\epsilon^3 \leq \frac{B}{\kappa}[(1 + \kappa\epsilon)^n - 1]\epsilon^2$$

In a similar way, $h^n(x) \geq \psi_n(x) - \frac{B}{\kappa}[(1 + \kappa\epsilon)^n - 1]\epsilon^2$, whenever $n \leq M_x(\epsilon)$. Together with the previous inequality, (5.12) holds. $\square$

In the next proposition, we give an expression for the vector field, $G_\epsilon(x)$, which generates the flow ψ_t^ϵ obtained in the last proposition.

Proposition 5.6 *Let $h_\epsilon(x) = \epsilon + x + \alpha x^2 + o(x^2)$ as before, and let ψ_t^ϵ be the flow obtained in Proposition 5.5. Then, the generating vector field of ψ_t^ϵ is given by*

$$G_\epsilon(x) = G_0(x) + g_1(x)\epsilon + \eta(x, \epsilon), \quad (5.14)$$

for all x, ϵ close enough to $(0, 0)$, where $G_0(x) = \alpha x^2 + o(x^2)$, $g_1(x) = 1 + o(1)$, and $\eta(x, \epsilon) = o(\epsilon)$.

Proof. Since ψ_t^ϵ is C^3 , then we can consider the expression (5.14), where G_0 and g_1 are to be determined. Recall that $G_0(x)$ generates the flow ψ_t^0 . Since $h_0(x) = x + \alpha x^2 + o(x^2)$ is embedded in ψ_t^0 , by Proposition 3.3, we have $G_0(x) = \alpha x^2 + o(x^2)$.

For simplicity, let us omit the dependence on ϵ of ψ_t^ϵ . In order to find g_1 above, we do as follows: since ψ_1 is the time-one map of the flow associated to G_ϵ , then for all x close to 0 we have

$$\int_x^{\psi_1(x)} \frac{ds}{G_\epsilon(s)} = 1.$$

By taking derivatives with respect to x , we get

$$\psi_1'(x)G_\epsilon(x) - G_\epsilon(\psi_1(x)) = 0. \quad (5.15)$$

The last expression is an example of the so-called *transfer equation* [36], where G_ϵ is to be determined.

Recall from (5.11) in Proposition 5.5 that $\psi_1(x) = h_\epsilon(x) - r_1^\epsilon(x)$, where $r_1^1(x) = o(\epsilon)$. By the mean value theorem, there is ξ between $h_\epsilon(x)$ and $\psi_1(x)$ such that

$$G_\epsilon(\psi_1(x)) = G_\epsilon(h_\epsilon(x)) - G'_\epsilon(\xi)r_1^1(x)$$

Hence, by substituting this expression in (5.15), we get

$$h'_\epsilon(x)G_\epsilon(x) - G_\epsilon \circ h_\epsilon(x) = G_\epsilon(x) \cdot \frac{d}{dx}r_1^\epsilon(x) - G'_\epsilon(\xi)r_1^\epsilon(x). \quad (5.16)$$

Notice the right hand side of (5.16) is $o(\epsilon)$.

We now substitute (5.14) into (5.16), so that

$$h'_\epsilon(x) [G_0(x) + g_1(x)\epsilon + o(\epsilon)] - G_0(h_\epsilon(x)) - g_1(h_\epsilon(x))\epsilon + o(\epsilon) = o(\epsilon).$$

Observe that since $h_\epsilon(x) = \epsilon + h_0(x)$, then $h'_\epsilon(x) = h'_0(x)$. Also,

$$\begin{aligned} G_0(h_\epsilon(x)) &= G_0(h_0(x)) + G'_0(h_0(x))\epsilon + o(\epsilon), \\ g_1(h_\epsilon(x)) &= g_1(h_0(x)) + g'_1(h_0(x))\epsilon + o(\epsilon). \end{aligned} \quad (5.17)$$

Hence, by substituting in the equation above, we obtain

$$h'_0(x)G_0(x) - G_0 \circ h_0(x) + [h'_0(x)g_1(x) - g_1 \circ h_0(x) - G'_0 \circ h_0(x)]\epsilon = o(\epsilon).$$

Since $\psi_1^0 = h_0$, we have that $h'_0(x)G_0(x) - G_0 \circ h_0(x) = 0$. Moreover, g_1 satisfies the transfer equation

$$h'_0(x)g_1(x) - g_1(h_0(x)) = g'_0(h_0(x)). \quad (5.18)$$

From the last equation we look for $g_1(0)$. However, notice that by substituting $x = 0$ above we obtain no new information. Instead we take derivatives in (5.18) and then we evaluate at $x = 0$. Thus,

$$h_0''(x)g_1(x) + h_0'(x)g_1'(x) - g_0''(h_0(x))h_0'(x) = g_1'(h_0(x))h_0'(x).$$

Since $h_0(0) = 0$, $h_0'(0) = 1$, and $h_0''(0) = g_0''(0) = 2\alpha$ we get $g_1(0) = 1$, so that $g_1(x) = 1 + o(1)$. $\square$

Propositions 5.5 and 5.6 constitute the cornerstone for estimating escape times for h_ϵ and the one-time map ψ_1^ϵ . We address this problem in the following section.

5.4 Deterministic escape times in saddle-node bifurcations

In the present section we give an asymptotic result for the behaviour of the escape time, $n_x(\epsilon)$, for the map h_ϵ . Let us first analyze the escape time, $t_x(\sigma^* + \epsilon) \equiv t_x(\epsilon)$, for the approximating flow ψ_t^ϵ obtained in the previous section. This escape time is defined as

$$t_x(\epsilon) := \inf\{t \geq 0 \mid \psi_t^\epsilon(x) > 0\}. \quad (5.19)$$

Equivalently, for any $x < 0$, the escape time $t_x(\epsilon)$ is the unique $t \geq 0$ such that

$$\psi_{t_x(\epsilon)}(x) = 0. \quad (5.20)$$

We prove that $t_x(\epsilon) \sim \epsilon^{-1/2}$, as given in the next proposition.

Proposition 5.7 (Scaling law for $t_x(\epsilon)$) *Let $\psi_t^\epsilon(x)$ be the flow defined by the ODE*

$$\dot{x} = G_\epsilon(x),$$

where $G_\epsilon(x)$ is given as in Proposition 5.6. Then, there exists $x_2 > 0$ such that for all $x \in (-x_2, 0)$ we have

$$\lim_{\epsilon \rightarrow 0} \sqrt{\epsilon} \cdot t_x(\epsilon) = \frac{\pi}{2\sqrt{\alpha}}. \quad (5.21)$$

Proof. Given the one-dimensional ODE

$$\dot{x} = G_\epsilon(x),$$

the escape time $t_x(\epsilon)$ is given by

$$t_x(\epsilon) = \int_x^0 \frac{ds}{G_\epsilon(s)}.$$

Recall that $G_\epsilon(x) = G_0(x) + g_1(x)\epsilon + \eta(x, \epsilon)$, where $g_1(x) = 1 + o(1)$ and $\eta(x, \epsilon) = o(\epsilon)$. Let us fix $x_2 > 0$ such that $g_1(x) \geq 1/2$ for all $x \in (-x_2, 0)$. Furthermore, for each $x \in (-x_2, 0)$ and $\delta \in (0, \min\{\alpha, 1\})$, there are positive $y \equiv y(\delta) < x_2$, and $\epsilon_0 \equiv \epsilon_0(\delta, x)$ small enough so that for all $x \in [-y, 0)$ and $\epsilon \in (0, \epsilon_0)$ we have

$$(\alpha - \delta)x^2 + (1 - \delta)\epsilon \leq G_\epsilon(x) \leq (\alpha + \delta)x^2 + (1 + \delta)\epsilon.$$

Therefore,

$$\int_x^0 \frac{ds}{(\alpha + \delta)s^2 + (1 + \delta)\epsilon} \leq t_x(\epsilon) \leq \int_x^0 \frac{ds}{(\alpha - \delta)s^2 + (1 - \delta)\epsilon}. \quad (5.22)$$

But

$$\int_x^0 \frac{ds}{(\alpha \pm \delta)s^2 + (1 \pm \delta)\epsilon} = \frac{1}{\sqrt{(\alpha \pm \delta)(1 \pm \delta)\epsilon}} \arctan \left(\frac{\sqrt{\alpha \pm \delta}}{\sqrt{(1 \pm \delta)\epsilon}} |x| \right).$$

Hence, for all $x \in (-y, 0)$,

$$\frac{\pi}{2\sqrt{(\alpha - \delta)(1 - \delta)}} \leq \liminf_{\epsilon \rightarrow 0} \sqrt{\epsilon} t_x(\epsilon) \leq \limsup_{\epsilon \rightarrow 0} \sqrt{\epsilon} t_x(\epsilon) \leq \frac{\pi}{2\sqrt{(\alpha + \delta)(1 + \delta)}} \quad (5.23)$$

We now extend the previous limits to the interval $(-x_2, 0)$. Indeed, let $x \in (-x_2, y)$ and consider its escape time $t_x(\epsilon)$. Notice $t_x(\epsilon) = t_x^y(\epsilon) + t_y(\epsilon)$, where

$$t_x^y(\epsilon) := \inf\{t \geq 0 \mid \psi_t^\epsilon(x) > y\}.$$

Let $\epsilon_1 \equiv \epsilon_1(x) < \epsilon_0$ be small enough so that $\eta(x, \epsilon) < \frac{1}{4}\epsilon$. Then, for all $x \in (-x_2, -y)$ and $\epsilon \in (0, \epsilon_1)$,

$$t_x^y(\epsilon) = \int_x^y \frac{ds}{G_\epsilon(s)} \leq \int_x^y \frac{ds}{G_0(x) + (1/4)\epsilon} \leq \int_x^y \frac{ds}{G_0(x)} = t_x^y(0).$$

It follows that for all $x \in (-x_2, 0)$, and all $\epsilon \in (0, \epsilon_1)$, we have $t_y(\epsilon) \leq t_x(\epsilon) \leq t_x^y(0) + t_y(\epsilon)$. Then, (5.23) holds for all $x \in (-x_2, 0)$, and δ small enough. Therefore, the proposition holds. $\square$

It is a direct consequence of the previous proposition that the same scaling law holds for the escape time $m_x(\epsilon)$. Recall that $M^* = M^{\sigma^*} \subset M^\sigma$ for all σ above but close to σ^* .

Corollary 5.8 (Scaling law for $m_x(\epsilon)$) *Let ψ_t^ϵ be the flow in Proposition 5.7 and $m_x(\epsilon)$ the escape time of $(\psi_n^\epsilon)_{n \geq 0}$ (see 5.10). Then, for all $x \in M^* \setminus \{0\}$ we have*

$$\lim_{\epsilon \rightarrow 0} \sqrt{\epsilon} \cdot m_x(\epsilon) = \frac{\pi}{2\sqrt{\alpha}}. \quad (5.24)$$

Proof. For all $\epsilon > 0$ and $x < 0$,

$$t_x(\epsilon) < m_x(\epsilon) \leq t_x(\epsilon) + 1.$$

Therefore, from Proposition 5.7 it follows that for all $x \in [-x_2, 0)$, where $x_2 > 0$ is small enough, the limit (5.24) holds.

Now, consider $x \in M^*$ such that $x < -x_2$. Then, $m_x(\epsilon) = m_x^{x_2}(\epsilon) + m_{x_2}(\epsilon)$, where

$$m_x^{x_2}(\epsilon) := \min\{n \geq 0 \mid \psi_n^\epsilon(x) > y\}.$$

Since $0 \leq m_x^{x_2}(\epsilon) \leq m_x^{x_2}(0)$, for all $\epsilon > 0$, then for all $x \in M^* \setminus \{0\}$,

$$\frac{\pi}{2\sqrt{\alpha}} = \lim_{\epsilon \rightarrow 0} \sqrt{\epsilon} m_{x_2}(\epsilon) \leq \lim_{\epsilon \rightarrow 0} \sqrt{\epsilon} m_x(\epsilon) \leq \lim_{\epsilon \rightarrow 0} (\sqrt{\epsilon} m_x^{x_2}(\epsilon) + \sqrt{\epsilon} m_{x_2}(\epsilon)) = \frac{\pi}{2\sqrt{\alpha}}$$

□

We now turn our attention to the escape time $n_x(\epsilon)$ for the map h_ϵ . For simplicity we omit again the dependence on ϵ in h_ϵ , ψ_n^ϵ , and r_ϵ^n , unless it may cause confusion. We now prove that $n_x(\epsilon)$ exhibits the same scaling law as $m_x(\epsilon)$ and $t_x(\epsilon)$

Lemma 5.9 (Scaling law for $n_x(\epsilon)$) *for all $x \in M^* \setminus \{0\}$,*

$$\lim_{\epsilon \rightarrow 0} \sqrt{\epsilon} n_x(\epsilon) = \frac{\pi}{2\sqrt{\alpha}}. \quad (5.25)$$

Proof. We split the proof in three parts. First, we show that for each $x \in M^* \setminus \{0\}$, there is $\beta > 0$ such that $(1 + M\epsilon)^{m_x(\epsilon)} \leq e^{\beta\sqrt{\epsilon}}$. Afterwards, we prove that for each $x \in (-x_1, 0)$, where x_1 is as in Proposition 5.5, and all $\epsilon > 0$ sufficiently small, $m_x(\epsilon) - 1 \leq n_x(\epsilon) \leq m_x(\epsilon) + 1$. Finally, we prove (5.25) holds for all $x \in M^* \setminus \{0\}$.

Claim 1: *let $x \in M^* \setminus \{0\}$, and $\epsilon_0 > 0$. Then, there exists $\beta \equiv \beta(x, \epsilon_0) > 0$ such that*

$$(1 + M\epsilon)^{m_x(\epsilon)} \leq e^{\beta\sqrt{\epsilon}}, \quad (5.26)$$

for all $\epsilon \in (0, \epsilon_0)$.

Indeed, we use the binomial expansion, so that

$$(1 + \kappa\epsilon)^{m_x(\epsilon)} = \sum_{j=0}^{m_x(\epsilon)} \binom{m_x(\epsilon)}{j} \kappa^j \epsilon^j \leq \sum_{j=0}^{m_x(\epsilon)} \frac{(\sqrt{\epsilon} m_x(\epsilon))^j \cdot (\kappa \sqrt{\epsilon})^j}{j!}.$$

By Corollary 5.8, there exists $\tilde{\beta} > 0$ such that $\sqrt{\epsilon} m_x(\epsilon) < \tilde{\beta}$ for all $\epsilon > 0$. Therefore,

$$(1 + \kappa\epsilon)^{m_x(\epsilon)} \leq \sum_{j=0}^{m_x(\epsilon)} \frac{(\tilde{\beta} \kappa \sqrt{\epsilon})^j}{j!} \leq e^{\beta \sqrt{\epsilon}},$$

where $\beta = \tilde{\beta} \kappa$.

Claim 2: let $x_1 > 0$ as in Proposition 5.5. Then, for each $x \in [-x_1, 0)$ there exists $\tilde{\epsilon} > 0$ such that for all $\epsilon \in (0, \tilde{\epsilon})$,

$$m_x(\epsilon) - 1 \leq n_x(\epsilon) \leq m_x(\epsilon) + 1. \quad (5.27)$$

Let $\tilde{\epsilon} > 0$ be such that for all $\epsilon \in (0, \tilde{\epsilon})$

$$-B \left(e^{\beta \sqrt{\epsilon}} - 1 \right) \epsilon > \max\{h_\epsilon^{-1}(0), \psi_{-1}^\epsilon(0)\}. \quad (5.28)$$

Notice that such $\tilde{\epsilon}$ exists since $(e^{\beta \sqrt{\epsilon}} - 1) \rightarrow 0$ as $\epsilon \rightarrow 0$, while both $(h_\epsilon^{-1})'(0)$ and $(\psi_{-1}^\epsilon)'(0)$ tend to -1 as $\epsilon \rightarrow 0$.

Let $\epsilon \in (0, \tilde{\epsilon})$, and assume first that $n_x(\epsilon) \geq m_x(\epsilon)$. We prove that $h^{m_x(\epsilon)}(x) > h^{-1}(0)$, so that $h^{m_x(\epsilon)+1} > 0$, which implies $m_x(\epsilon) + 1 \geq n_x(\epsilon)$. Since we assumed that $n_x(\epsilon) \geq m_x(\epsilon)$, then $M_x(\epsilon) = m_x(\epsilon)$. Therefore, 5.12 in Proposition 5.5 holds for $n = m_x(\epsilon)$. Hence,

$$h^{m_x(\epsilon)}(x) = \psi_{m_x(\epsilon)}(x) + r^{m_x(\epsilon)}(x) > r^{m_x(\epsilon)}(x),$$

by definition of $m_x(\epsilon)$. Furthermore, by (5.12), Claim 1 above, and the choice of $\tilde{\epsilon}$, we have

$$h^{m_x(\epsilon)}(x) \geq -B \left(e^{\beta \sqrt{\epsilon}} - 1 \right) \epsilon^2 > h^{-1}(0).$$

Thus, $n_x(\epsilon) \leq m_x(\epsilon) + 1$.

Analogously, we now assume that $m_x(\epsilon) \geq n_x(\epsilon)$, and we prove that $\psi_{n_x(\epsilon)}(x) > \psi_{-1}(0)$. By (5.12) for $n = n_x(\epsilon)$, we have

$$\psi_{n_x(\epsilon)}(x) > -B \left[(1 + \kappa\epsilon)^{n_x(\epsilon)} - 1 \right] \epsilon^2 \geq -B \left[(1 + \kappa\epsilon)^{m_x(\epsilon)} - 1 \right] \epsilon^2 \geq -B \left(e^{\beta\sqrt{\epsilon}} - 1 \right) \epsilon^2 \geq \psi_{-1}(0).$$

Hence, $n_x(\epsilon) + 1 \geq m_x(\epsilon)$, and the claim follows.

Claim 3: for all $x \in M^* \setminus \{0\}$, the limit (5.27) holds.

Notice that (5.27) holds for all $x \in [-x_1, 0)$, due to Claim 2 and Corollary 5.8. For $x \leq -x_1$, observe that $n_{x_1}(\epsilon) \leq n_x(\epsilon) \leq n_x^{x_1}(0) + n_{x_1}(\epsilon)$, where

$$n_x^{x_1}(\epsilon) := \min\{n \geq 0 \mid h_\epsilon^n(x) > x_1\}.$$

The result follows by multiplying times $\sqrt{\epsilon}$ and taking limits as $\epsilon \rightarrow 0$. $\square$

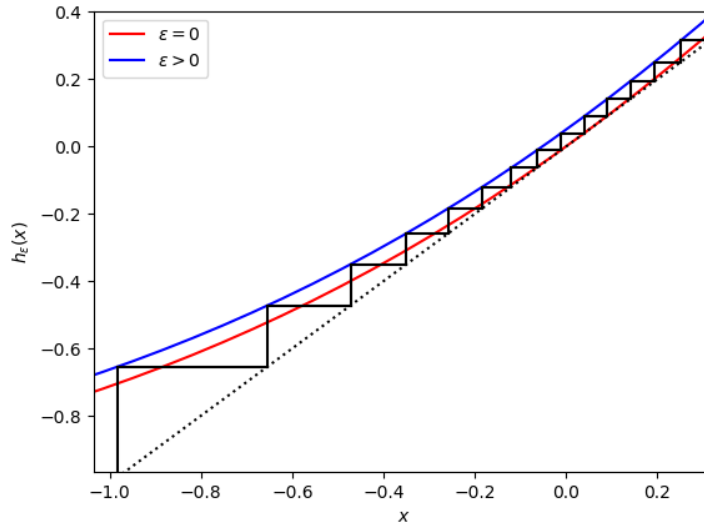


Figure 5.3: A bottleneck phenomenon. Even though the system eventually goes beyond 0, it does take a long time whenever ϵ is close to 0. Because of Lemma 5.25, this escape time goes like $1/\sqrt{\epsilon}$.

The previous result will be used in the following section for giving asymptotic estimates on how the quasi-stationary density $\hat{\phi}$ behaves at x^* as $\sigma \rightarrow \sigma^*$.

5.5 Asymptotic behaviour for mean escape times

In this section, the main goal is to analyse the behaviour of the unique quasi-stationary density $\hat{\phi}$ at x^* as $\sigma \rightarrow \sigma^*$ of the system (5.1), assuming x^* is a nonhyperbolic fixed point as given by (5.2). Recall that $\hat{\phi}^\sigma$ is supported on $M^\sigma = [x_-(\sigma), x^*]$, where $T(x_-(\sigma)) - \sigma = x_-(\sigma)$ for all σ close to σ^* . Like in previous sections, we assume without loss of generality that $x^* = 0$.

Recall that the quasi-stationary density $\hat{\phi}^\sigma$ is the unique eigenfunction of the restricted transfer operator $\hat{L} : L^1(M^\sigma) \rightarrow L^1(M^\sigma)$, introduced in Section 4.3, which as in (4.21) is given by

$$\hat{L}g(x) = \int \hat{k}(y, x)g(y)dy,$$

for all $g \in L^1(M^\sigma)$, where

$$\hat{k}(x, y) = \frac{1}{2\sigma} \mathbb{1}_{M^\sigma \cap F(x)}(y),$$

since we assume $(\omega_n)_{n=0}^\infty$ is an i.i.d. sequence of random variables, where each ω_n is distributed according to $d\nu = \frac{1}{2}d\text{Leb}$ on $[-1, 1]$. Its corresponding eigenvalue is $\Lambda^\sigma = \int \hat{L}\hat{\phi}^\sigma d\text{Leb}$. Thus, the quasi-stationary density $\hat{\phi}^\sigma$ satisfies the functional equation

$$\hat{\phi}^\sigma(x) = \frac{1}{2\sigma\Lambda^\sigma} \int_{h_+^{-1}(x) \vee x_-(\sigma)}^{h_-^{-1}(x) \wedge 0} \hat{\phi}^\sigma(y)dy, \quad (5.29)$$

where $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. Recall that the upper and lower extremal maps are the maps $h_+(x; \sigma) = T(x) + \sigma$ and $h_-(x; \sigma) = T(x) - \sigma$, respectively. Observe that h_+ can be written as

$$h_+(x) \equiv h_+(x; \sigma) = (\sigma - \sigma^*) + x + \alpha x^2 + o(x^2).$$

In the following we omit all the dependences on σ , where it causes no confusion.

We proceed by using similar techniques as in Section 3.3: we first find lower bound estimates for $\hat{\phi}(0)$ in terms of $\hat{\phi}(h_+^{-1}(0))$. To fix some groundwork, we prove the following simple result.

Proposition 5.10 *For each negative $x_0 > \max\{-2\sigma^*, x_-(\sigma^*) + 2\sigma^*\}$, there is $\sigma_0 > \sigma^*$ such that*

$$\hat{\phi}(x) = \frac{1}{2\sigma\Lambda} \int_{h_+^{-1}(x)}^0 \hat{\phi}(y)dy, \quad (5.30)$$

for all $x \in (x_0, 0)$, and $\sigma \in (\sigma^*, \sigma_0)$.

Proof. Notice that $-2\sigma^* = T(0) - \sigma^* \geq T(0) - \sigma$, for all $\sigma \geq \sigma^*$. Therefore, whenever $x > -2\sigma^*$, we have $h_-^{-1}(x) > 0$.

On the other hand, observe that $x_-(\sigma^*) + 2\sigma^* = T(x_-(\sigma^*)) + \sigma^*$. Hence, given $x_0 > x_-(\sigma^*) + 2\sigma^*$, there is $\sigma_0 > \sigma^*$ such that for all $\sigma \in (\sigma^*, \sigma_0)$, we have $T(x_-(\sigma)) + \sigma < x_0$. Thus, $x_-(\sigma) < h_+^{-1}(x)$ for all $x \in (x_0, 0)$, and $\sigma \in (\sigma^*, \sigma_0)$. The result follows by substituting the upper and lower limit in (5.29). $\square$

In the following, we prove a monotonicity property of $\hat{\phi}$ near 0, for each $\sigma \in (\sigma^*, \sigma_0)$.

Proposition 5.11 *For each negative $x_0 > \max\{-2\sigma^*, x_-(\sigma^*) + 2\sigma^*\}$, there is $\sigma_0 > \sigma^*$ such that the quasi-stationary density $\hat{\phi}$ is decreasing in $(x_0, 0)$ for all $\sigma \in (\sigma^*, \sigma_0)$.*

Proof. Let σ_0 be as in Proposition 5.10. Since for each $\sigma \in (\sigma^*, \sigma_0)$, the quasi-stationary density $\hat{\phi}$ is continuous, it follows from (5.30) that $\hat{\phi}$ is continuously differentiable in $(x_0, 0)$. Therefore,

$$\frac{\partial}{\partial x} \hat{\phi}(x) = -\frac{\hat{\phi}(h_+^{-1}(x))}{2\sigma\Lambda \cdot h'_+(h_+^{-1}(x))} < 0.$$

$\square$

Let us proceed then with the following Lemma.

Lemma 5.12 *Let $x_0 > \max\{-2\sigma^*, T(x_-(\sigma^*)) + \sigma^*\}$, and $\sigma_0 > \sigma^*$ as in Proposition 5.11. Assume without loss of generality that $\sigma_0 - \sigma^*$ is small enough so that for all $\sigma \in (\sigma^*, \sigma_0)$, $h_+(x_0) < 0$ and $|x_0| < 2\sigma\Lambda$. Then, there exist $M_1, M_2 > 0$ such that for all $\sigma \in (\sigma^*, \sigma_0)$, and all $x \in [h_+(x_0), 0]$,*

$$M_1(x - h_+^{-1}(x))^2 \cdot \hat{\phi}(h_+^{-1}(x)) \leq \hat{\phi}(x) \leq M_2(x - h_+^{-1}(x)) \cdot \hat{\phi}(h_+^{-1}(x)). \quad (5.31)$$

Proof. We split the proof into two parts, one for each inequality in (5.31).

Part 1: lower bound of (5.31).

By the choice of x_0 and σ_0 , formula (5.30) holds for $\hat{\phi}(x)$. Moreover, since for all $z \in [h_+^{-1}(x), 0]$ we have that $z \in (x_0, 0)$, then we can apply (5.30) twice so that

$$\hat{\phi}(x) = \frac{1}{(2\sigma\Lambda)^2} \int_{h_+^{-1}(x)}^0 \int_{h_+^{-1}(y)}^0 \hat{\phi}(z) \, dz dy. \quad (5.32)$$

Observe from Figure 5.4 that the integration region in (5.32) is the same as

$$\{(y, z) \mid h_+^{-1}(x) \leq y \leq h_+(z) \wedge 0; h_+^{-2}(x) \leq z \leq 0\}.$$

Hence,

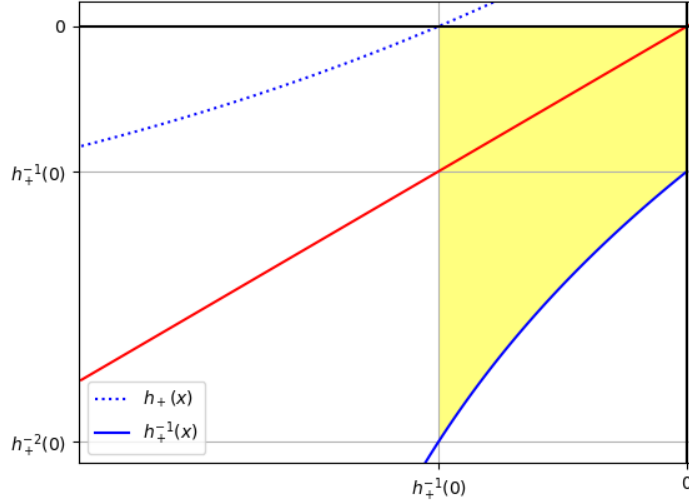


Figure 5.4: The integration region in (5.32).

$$\hat{\phi}(x) = \frac{1}{(2\sigma\Lambda)^2} \int_{h_+^{-2}(x)}^0 \int_{h_+^{-1}(x)}^{h_+(z) \wedge 0} \hat{\phi}(z) dy dz = \frac{1}{(2\sigma\Lambda)^2} \int_{h_+^{-2}(x)}^0 [(h_+(z) \wedge 0) - h_+^{-1}(x)] \hat{\phi}(z) dz. \quad (5.33)$$

Since $\hat{\phi}$ is decreasing near 0, we bound from below in (5.33) by replacing the upper integration limit, so that

$$\hat{\phi}(x) \geq \frac{\hat{\phi}(h_+^{-1}(x))}{(2\sigma\Lambda)^2} \int_{h_+^{-2}(x)}^{h_+^{-1}(x)} [h_+(z) - h_+^{-1}(x)] dz. \quad (5.34)$$

Since $h_+(z) - h_+^{-1}(x) = h_+(z) - h_+(h_+^{-2}(x))$, due to the mean value theorem, there is $\tilde{M}_1 > 0$ such that

$$h_+(z) - h_+^{-1}(x) \geq \tilde{M}_1(z - h_+^{-2}(x)).$$

By using this inequality in (5.34) we obtain

$$\hat{\phi}(x) \geq \frac{\tilde{M}_1 \cdot \hat{\phi}(h_+^{-1}(x))}{(2\sigma\Lambda)^2} \int_{h_+^{-2}(x)}^{h_+^{-1}(x)} (z - h_+^{-2}(x)) dz \geq M_1(x - h_+^{-1}(x))^2 \cdot \hat{\phi}(h_+^{-1}(x)), \quad (5.35)$$

where $M_1 = \tilde{M}_1/(8\sigma_0^2)$.

Part 2: upper bound of (5.31).

We split the integral in (5.30) as

$$\hat{\phi}(x) = \frac{1}{2\sigma\Lambda} \left[\int_{h_+^{-1}(x)}^x \hat{\phi}(y) dy + \int_x^0 \hat{\phi}(y) dy \right].$$

By the monotonicity of $\hat{\phi}$,

$$\hat{\phi}(x) \leq \frac{x - h_+^{-1}(x)}{2\sigma\Lambda} \hat{\phi}(h_+^{-1}(x)) + \frac{|x|}{2\sigma\Lambda} \hat{\phi}(x).$$

Hence,

$$\hat{\phi}(x) \leq M_2(x - h_+^{-1}(x)) \cdot \hat{\phi}(h_+^{-1}(x)),$$

where $M_2 = (2\sigma^* \inf_{\sigma \in (\sigma^*, \sigma_0)} \Lambda - |x_0|)^{-1}$. $\square$

Remark 5.13 By the mean value theorem, for each $x \in [h_+^2(x_0), 0]$, there is $\xi \in (h_+^{-1}(x), x)$ such that

$$x - h_+^{-1}(x) = h_+^{-1}(h_+(x)) - h_+^{-1}(x) = \frac{h_+(x) - x}{h'_+(\xi)}.$$

Whenever x_0 is close enough to 0, we have that $h''_+(x) > 0$ for all $x \in [x_0, 0]$, so that $h'(x) < 1$. Hence,

$$x - h_+^{-1}(x) \geq h_+(x) - x = (\sigma - \sigma^*) + \alpha x^2 + o(x^2).$$

Furthermore, assuming x_0 is close enough to 0, we obtain

$$x - h_+^{-1}(x) \geq (\sigma - \sigma^*),$$

for all $x \in [x_0, 0]$. Using the last inequality in (5.31), we get

$$\hat{\phi}(x) \geq M_1(\sigma - \sigma^*)^2 \cdot \hat{\phi}(h_+^{-1}(x)). \quad (5.36)$$

We now proceed to prove Theorem 5.1. We use iteratively the previous Lemma in order to obtain the main result.

Proof of Theorem 5.1. Recall the inequalities we aim to prove are

$$\frac{\pi}{2\sqrt{\alpha}} \leq \liminf_{\sigma \rightarrow \sigma^*} \frac{(\sigma - \sigma^*)^{1/2} \ln \hat{\phi}(0)}{\ln(\sigma - \sigma^*)} \leq \limsup_{\sigma \rightarrow \sigma^*} \frac{(\sigma - \sigma^*)^{1/2} \ln \hat{\phi}(0)}{\ln(\sigma - \sigma^*)} \leq \frac{\pi}{\sqrt{\alpha}}. \quad (5.37)$$

Let $x_0 < 0$ close enough to 0 and $\sigma_0 > \sigma^*$ so that (5.36) in Remark 5.13 holds for all $x \in [x_0, 0]$ and $\sigma \in (\sigma^*, \sigma_0)$. We split the proof in two parts, one for each inequality above.

Part 1: upper bound in (5.37).

From (5.36) in Remark 5.13, it follows that

$$\hat{\phi}(0) \geq M_1(\sigma - \sigma^*)^2 \cdot \hat{\phi}(h_+^{-1}(0)).$$

If $h_+^{-1}(0) > h_+(x_0)$, then we use (5.36) again so that

$$\hat{\phi}(0) \geq M_1^2(\sigma - \sigma^*)^4 \cdot \hat{\phi}(h_+^{-2}(x)).$$

In general, we can repeat this inequality for as long as $h_+^{-n}(0) \geq x_0$. So, if $n_{x_0} \equiv n_{x_0}(\sigma)$ is the escape time for the upper extremal map h_+ , starting at x_0 , then for each $\sigma \in (\sigma^*, \sigma_0)$ we can iterate this inequality for $n_{x_0} - 1$ times. Therefore,

$$\hat{\phi}(0) \geq M_1^{n_{x_0}-1}(\sigma - \sigma^*)^{2(n_{x_0}-1)} \cdot \hat{\phi}(h_+^{-n_{x_0}+1}(0)) \geq M_1^{n_{x_0}-1}(\sigma - \sigma^*)^{2(n_{x_0}-1)} \cdot \hat{\phi}(h_+(x_0)).$$

By taking logarithms, multiplying times $(\sigma - \sigma^*)^{1/2}$, and dividing by $\ln(\sigma - \sigma^*)$, we obtain

$$\frac{(\sigma - \sigma^*)^{1/2} \ln \hat{\phi}(0)}{\ln(\sigma - \sigma^*)} \leq 2(\sigma - \sigma^*)^{1/2}(n_{x_0} - 1) + \frac{(\sigma - \sigma^*)^{1/2}(n_{x_0} - 1) \ln M_1}{\ln(\sigma - \sigma^*)} + \frac{(\sigma - \sigma^*)^{1/2} \ln \hat{\phi}(h_+(x_0))}{\ln(\sigma - \sigma^*)}$$

Because of Lemma 5.9,

$$\limsup_{\sigma \rightarrow \sigma^*} \frac{(\sigma - \sigma^*)^{1/2} \ln \hat{\phi}(0)}{\ln(\sigma - \sigma^*)} \leq \frac{\pi}{\sqrt{\alpha}}.$$

Part 2: lower bound in (5.37).

From the upper bound in (5.31) in Lemma 5.12, we have that

$$\hat{\phi}(0) \leq M_2 |h_+^{-1}(0)| \cdot \hat{\phi}(h_+^{-1}(0)).$$

Like in the previous part, we repeat this inequality n times, for as long as $h_+^{-n}(0) > x_0$. By the monotonicity of $\hat{\phi}$, for $n = n_{x_0} - 1$,

$$\hat{\phi}(0) \leq M_2^n (h_+^{-1}(0)) \cdots (h_+^{-n+2}(0) - h_+^{-n+1}(0)) \cdot \hat{\phi}(x_0).$$

By taking logarithms, we obtain

$$\ln \hat{\phi}(0) \leq \sum_{j=0}^{n_{x_0}-2} \ln (h_+^{-j}(0) - h_+^{-j-1}(0)) + (n_{x_0} - 1) \ln M_2 + \ln \hat{\phi}(x_0). \quad (5.38)$$

First notice that, because of Lemma 5.9,

$$\lim_{\sigma \rightarrow \sigma^*} \frac{\sqrt{\sigma - \sigma^*} \cdot (n_{x_0} - 1)}{\ln(\sigma - \sigma^*)} \ln M_2 = 0.$$

Also,

$$\lim_{\sigma \rightarrow \sigma^*} \frac{\sqrt{\sigma - \sigma^*}}{\ln(\sigma - \sigma^*)} \ln \hat{\phi}(x_0) = 0.$$

Therefore, in order to obtain the lower bound in (5.37), it is enough to analyse the sum in the right hand side of (5.38). Let $S = \sum_{j=0}^{n_{x_0}-2} \ln (h_+^{-j}(0) - h_+^{-j-1}(0))$, and let $\Delta h_+^{-j} := h_+^{-j}(0) - h_+^{-j-1}(0)$. Then,

$$S = \sum_{j=0}^{n-2} \frac{\ln (h_+^{-j}(0) - h_+^{-j-1}(0))}{h_+^{-j}(0) - h_+^{-j-1}(0)} \Delta h_+^{-j}. \quad (5.39)$$

We proceed to bound this last expression with an integral.

Let us denote $\epsilon = \sigma - \sigma^*$ as before. Since $h_+(x) = \epsilon + x + \alpha x^2 + o(x^2) \equiv \epsilon + h_0(x)$, then $h_+^{-1}(x) = h_0^{-1}(x - \epsilon)$. Let $\delta \in (0, \alpha)$, and assume $|x_0|$ is small enough and σ_0 is close enough to σ^* , so that for all $x \in [x_0, 0]$ and $\sigma \in (\sigma^*, \sigma_0)$

$$x - \epsilon - (\alpha + \delta)(x - \epsilon)^2 \leq h_+^{-1}(x) \leq x - \epsilon - (\alpha - \delta)(x - \epsilon)^2.$$

Therefore, it follows from (5.39) that

$$S \leq \sum_{j=0}^{n-2} \frac{\ln [\epsilon + (\alpha + \delta)(\epsilon + h_+^{-j}(0))^2]}{\epsilon + (\alpha - \delta)(\epsilon + h_+^{-j}(0))^2} \Delta h_+^{-j}. \quad (5.40)$$

Since the function $s \mapsto \frac{\ln(\epsilon + (\alpha + \delta)s^2)}{\epsilon + (\alpha - \delta)s^2}$ is increasing and negative for all $s \in (0, x_0)$ (assuming again $|x_0|$ is small enough, see Figure 5.5), then

$$S \leq \int_0^{|h_+(x_0)|} \frac{\ln(\epsilon + (\alpha + \delta)(\epsilon + s)^2)}{\epsilon + (\alpha - \delta)(\epsilon + s)^2} ds$$

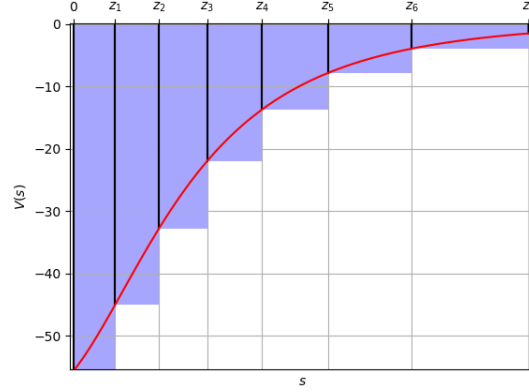


Figure 5.5: The graph of $V(s) = \ln(\epsilon + (\alpha + \delta)(\epsilon + s)^2) / (\epsilon + (\alpha - \delta)(\epsilon + s)^2)$ is depicted in a red solid line. The area above the graph is less than the blue-shaded area, representing the sum in (5.40), where $z_i = |h_+^{-i}(0)|$

Since $\ln(\epsilon + (\alpha + \delta)(\epsilon + s)^2) = \ln \epsilon + \ln(1 + (\alpha + \delta)(s + \epsilon)^2/\epsilon)$, then

$$S \leq \int_{\epsilon}^{\epsilon + |h_+(x_0)|} \frac{\ln \epsilon}{\epsilon + (\alpha - \delta)s^2} ds + \int_{\epsilon}^{\epsilon + |h_+(x_0)|} \frac{\ln(1 + (\alpha + \delta)s^2/\epsilon)}{\epsilon + (\alpha - \delta)s^2} ds \equiv I_1 + I_2.$$

The first integral equals

$$I_1 = \ln \epsilon \sqrt{\frac{1}{\epsilon(\alpha - \delta)}} \left[\arctan \left(\frac{(\epsilon + |h_+(x_0)|)\sqrt{\alpha - \delta}}{\epsilon} \right) - \arctan \left(\sqrt{\epsilon(\alpha - \delta)} \right) \right].$$

Hence,

$$\lim_{\epsilon \rightarrow 0} \frac{\sqrt{\epsilon} \cdot I_1}{\ln \epsilon} = \frac{\pi}{2\sqrt{\alpha - \delta}}. \quad (5.41)$$

On the other hand, for the second integral I_2 we have

$$I_2 = \int_{\epsilon}^{\epsilon + |h_+(x_0)|} \frac{\ln(1 + (\alpha + \delta)s^2/\epsilon)}{\epsilon + (\alpha - \delta)s^2} ds \leq \frac{1}{\alpha - \delta} \int_{\epsilon}^{\epsilon + |h_+(x_0)|} \frac{\ln(1 + (\alpha + \delta)s^2/\epsilon)}{s^2} ds.$$

The right hand side of the last expression is solvable when integrating by parts. Indeed, by taking $u = \ln(1 + (\alpha + \delta)s^2)$, $dv = s^{-2}ds$, we obtain

$$I_2 \leq \left[\frac{\ln \left(1 + \frac{\alpha + \delta}{\epsilon} s^2 \right)}{(\alpha - \delta)s} + \frac{2}{\sqrt{\epsilon(\alpha - \delta)}} \arctan \left(s \sqrt{\frac{\alpha - \delta}{\epsilon}} \right) \right]_{\epsilon}^{\epsilon + |h_+(x_0)|}$$

Due to l'Hôpital's rule, we have that

$$\lim_{\epsilon \rightarrow 0} \frac{\sqrt{\epsilon}}{\ln \epsilon} \cdot \left(\frac{\ln(1 + (\alpha + \delta)\epsilon)}{(\alpha - \delta)\epsilon} \right) = \lim_{\epsilon \rightarrow 0} \frac{\sqrt{\epsilon}}{\ln \epsilon} \cdot \left(\frac{\ln(1 + (\alpha + \delta)(\epsilon + z_1)^2/\epsilon)}{(\alpha - \delta)(\epsilon + z_1)} \right) = 0.$$

Hence, it follows that

$$\lim_{\epsilon \rightarrow 0} \frac{\sqrt{\epsilon} I_2}{\ln \epsilon} = 0. \quad (5.42)$$

Because of (5.41) and (5.42), we get

$$\liminf_{\epsilon \rightarrow 0} \frac{\sqrt{\epsilon} \ln \hat{\phi}(0)}{\ln \epsilon} \geq \liminf_{\epsilon \rightarrow 0} S \geq \lim_{\epsilon \rightarrow 0} (I_1 + I_2) = \frac{\pi}{2\sqrt{\alpha - \delta}}.$$

The last expression is true for all δ small enough, and thus the lower bound in (5.37) holds. $\square$

The following result is analogous to Theorem 5.1 but for the eigenvalue, Λ^σ , of the conditional transfer operator, Λ^σ , for the quasi-stationary density $\hat{\phi}^\sigma$.

Corollary 5.14 *Let $\hat{\phi}^\sigma$ be the unique quasi-stationary density supported on $M^\sigma = [x_-(\sigma), x^*]$, and Λ^σ its associated eigenvalue for the operator $\hat{L}$. If x^* is a nondegenerate nonhyperbolic fixed point for the map $x \mapsto T(x) + \sigma^*$, then*

$$\frac{\pi}{2\sqrt{\alpha}} \leq \liminf_{\sigma \rightarrow \sigma^*} \frac{(\sigma - \sigma^*)^{1/2} \ln(1 - \Lambda^\sigma)}{\ln(\sigma - \sigma^*)} \leq \limsup_{\sigma \rightarrow \sigma^*} \frac{(\sigma - \sigma^*)^{1/2} \ln(1 - \Lambda^\sigma)}{\ln(\sigma - \sigma^*)} \leq \frac{\pi}{\sqrt{\alpha}} \quad (5.43)$$

Proof. Recall that $\Lambda^\sigma = \mathbb{P}_{q^\sigma}[1 < \tau]$. Therefore, $1 - \Lambda^\sigma = \int_M P \mathbf{1}_{\mathbb{R} \setminus M^\sigma}(x) dq^\sigma(x)$. However, note that for all points $x \in [x_-, h_+^{-1}(x^*)]$, the probability of leaving M^σ is 0. On the other hand, for all $x \in [h_+^{-1}(x^*), x^*]$,

$$P \mathbf{1}_{\mathbb{R} \setminus M^\sigma}(x) = \frac{T(x) + \sigma - x^*}{2\sigma}.$$

Therefore,

$$1 - \Lambda^\sigma = \int_{h_+^{-1}(0)}^{x^*} \left(\frac{T(x) + \sigma - x^*}{2\sigma} \right) \hat{\phi}^\sigma(x) dx \quad (5.44)$$

We will bound from above and from below the previous expression in order to get (5.43). We start by noting that, since T is increasing and since $T(x^*) + \sigma^* = x^*$, we have

$$1 - \Lambda^\sigma \leq \frac{T(x^*) + \sigma - x^*}{2\sigma} \int_{h_+^{-1}(0)}^0 \hat{\phi}^\sigma(x) dx = (\sigma - \sigma^*) \hat{L} \hat{\phi}^\sigma(0) = \Lambda^\sigma(\sigma - \sigma^*) \hat{\phi}^\sigma(0).$$

By taking logarithms, multiplying times $(\sigma - \sigma^*)^{1/2}$, diving by $\ln(\sigma - \sigma^*)$ and taking $\liminf_{\sigma \rightarrow \sigma^*}$ on both sides, we obtain the left inequality in (5.43) due to Theorem 5.1.

For the upper bound, we use Proposition 5.11, so that $\hat{\phi}^\sigma$ is decreasing near 0 for all σ small enough. Hence, since $T(x) + \sigma > x$,

$$1 - \Lambda^\sigma \geq \hat{\phi}^\sigma(0) \cdot \int_{h_+^{-1}(0)}^0 \frac{T(x) + \sigma}{2\sigma} dx \geq \hat{\phi}^\sigma(0) \cdot \int_{h_+^{-1}(0)}^0 \frac{x}{2\sigma} \geq \frac{|h_+^{-1}(0)|^2}{4\sigma} \cdot \hat{\phi}^\sigma(0).$$

Recall that $h_+^{-1}(x) = x - (\sigma - \sigma^*) + o(x - (\sigma - \sigma^*))$. Then, $|h_+^{-1}(0)|^2 = (\sigma - \sigma^*)^2 + o((\sigma - \sigma^*)^2)$. Therefore,

$$1 - \Lambda^\sigma \geq \frac{(\sigma - \sigma^*)^2 + o((\sigma - \sigma^*)^2)}{16\sigma} \cdot \hat{\phi}^\sigma(0).$$

The right inequality in (5.43) follows from analogous arguments as the previous step. $\square$

5.6 Numerical experiments

In the present section we present some plots obtained numerically for the behaviour of the largest eigenvalue Λ^σ of the conditional transfer operator $\hat{L}$ for the toy model presented in Chapter 1.

Because of Theorem 4.15, it was shown in [14] that for every $x \in M^* = [x_-(\sigma^*), x^*]$

$$\Lambda^\sigma = \lim_{n \rightarrow \infty} \frac{\mathbb{P}_x[n+1 < \tau]}{\mathbb{P}_x[n < \tau]} = \lim_{n \rightarrow \infty} \frac{\hat{P}^{n+1}\mathbf{1}(x)}{\hat{P}^n\mathbf{1}(x)}. \quad (5.45)$$

Hence, we take an arbitrary point $x \in M^*$, and $n \in \mathbb{N}$ large to approximate Λ^σ as $\frac{\hat{P}^{n+1}\mathbf{1}(x)}{\hat{P}^n\mathbf{1}(x)}$.

As usual, we consider the toy model

$$x_{n+1} = b \left(\frac{1 - e^{-x}}{1 + e^{-x}} \right) + \sigma \omega_n \equiv T(x_n) + \sigma \omega_n,$$

where $b > 2$, and $(\omega_n)_{n=0}^\infty$ is an i.i.d. sequence, uniformly distributed on $[-1, 1]$. Recall from Subsection 3.3.1 that the upper extremal map undergoes a saddle-node bifurcation at

$$x^* = \ln \left(b - 1 + \sqrt{(b-1)^2 - 1} \right),$$

$$\sigma^* = b \left(\frac{b-2+\sqrt{(b-1)^2-1}}{b+\sqrt{(b-1)^2-1}} \right) - \ln \left(b - 1 + \sqrt{(b-1)^2 - 1} \right).$$

Recall also that $\alpha = |x^* - \sigma^*|/2b$. Therefore, from (5.43) we have

$$\gamma \leq \liminf_{\sigma \rightarrow \sigma^*} \frac{(\sigma - \sigma^*)^{1/2} \ln(1 - \Lambda^\sigma)}{\ln(\sigma - \sigma^*)} \leq \limsup_{\sigma \rightarrow \sigma^*} \frac{(\sigma - \sigma^*)^{1/2} \ln(1 - \Lambda^\sigma)}{\ln(\sigma - \sigma^*)} \leq 2\gamma, \quad (5.46)$$

where $\gamma = \frac{\pi\sqrt{b}}{\sqrt{2|x^* - \sigma^*|}}$.

For σ close to σ^* , from (5.46) we get

$$\frac{2\gamma \ln(\sigma - \sigma^*)}{(\sigma - \sigma^*)^{1/2}} \lesssim \ln(1 - \Lambda) \lesssim \frac{\gamma \ln(\sigma - \sigma^*)}{(\sigma - \sigma^*)^{1/2}} \quad (5.47)$$

Therefore, in a semilog plot we should be able to compare it with the functions appearing on each side of the approximate inequalities in (5.47), see Figure 5.6.

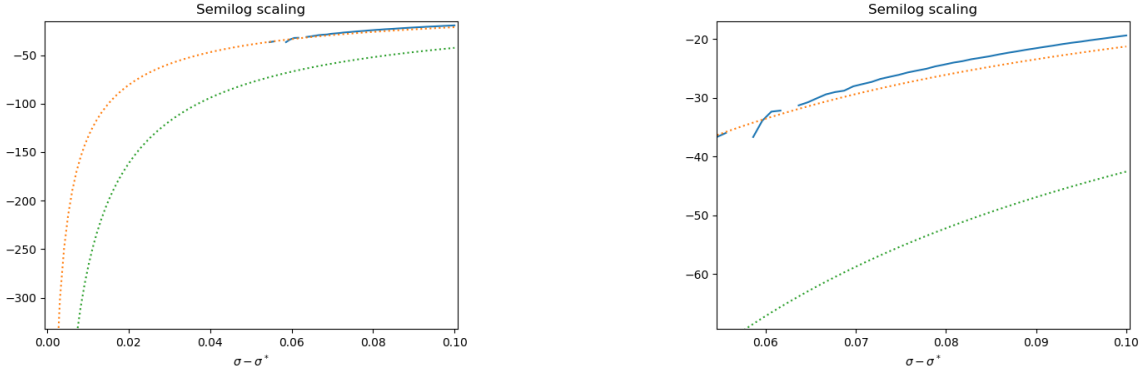


Figure 5.6: In both subfigures the numerically approximated value of $\ln(1 - \Lambda)$ is presented in a solid blue line. In dotted lines, one can observe the functions $\frac{k\gamma \ln(\sigma - \sigma^*)}{\sqrt{\sigma - \sigma^*}}$ where $k = 1$ is given in orange and $k = 2$ is the green one. The smallest the value of $\sigma - \sigma^*$, that is the more we move to the left, the more the solid blue line approaches the band defined by the dotted lines.

We can also visualize this in a different scale which is more compatible with straight line visualization in the appropriate scale. For example, from the right approximate inequality in (5.47) we can multiply times -1 and then take logarithms in order to get

$$\ln(\gamma) - \frac{1}{2} \ln(\sigma - \sigma^*) + \ln(-\ln(\sigma - \sigma^*)) \lesssim \ln(-\ln(1 - \Lambda)).$$

An analogous expression follows from the left approximation of the left hand side in (5.47). Thus by adding $\ln(-\ln(\sigma - \sigma^*))$ we obtain

$$\ln(\gamma) - \frac{1}{2} \ln(\sigma - \sigma^*) \lesssim \ln\left(\frac{\ln(1 - \Lambda)}{\ln(\sigma - \sigma^*)}\right) \lesssim \ln(2\gamma) - \frac{1}{2} \ln(\sigma - \sigma^*). \quad (5.48)$$

In Figure 5.7 it can be seen how $\ln\left(\frac{\ln(1-\Lambda)}{\ln(\sigma-\sigma^*)}\right)$ compares to the lines $\ln(k\gamma) - \frac{1}{2}\ln(\sigma - \sigma^*)$ in logarithmic scale, for $k = 1, 2$.

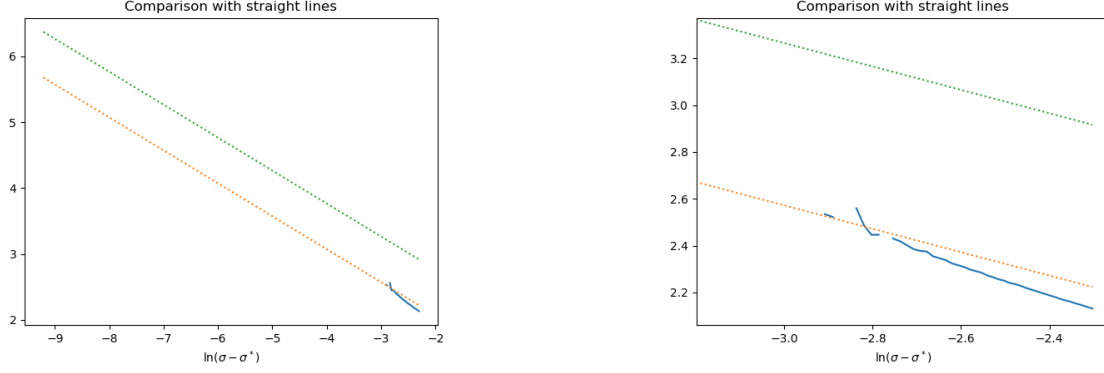


Figure 5.7: Here we plot as a blue solid line $\ln(1 - \Lambda)/\ln(\sigma - \sigma^*)$ as a function of $\sigma - \sigma^*$, both in a logarithmic scale. In dotted lines, we present the lines $\ln(k\gamma) - (1/2)\ln(\sigma - \sigma^*)$ where $k = 1$ is given in orange and $k = 2$ is the green one. Again, the more negative $\ln(\sigma - \sigma^*)$ becomes, the closer the solid line remains to the band defined by dotted lines.

Remark 5.15 *Just like in Remark 3.11, the results here presented are the best obtained. Similar calculations to those performed in 3.11 lead to less accurate bounds in Theorem 5.1.*

Chapter 6

Outlook

As we mentioned early in this thesis, radical changes in complex systems which are driven by noisy, or more generally by nonautonomous perturbations, are called *critical transitions* [51]. This type of behaviour has proven to be ubiquitous in several disciplines, being particularly important in ecology due to its scientific, ethical, and truth to be told, economical relevance. Back in the 70s, Holling already considered ecological systems as systems undergoing a constant transient state. In his words [32]:

An equilibrium centred view is essentially static and provides little insight into the transient behaviour of systems that are not near the equilibrium. Natural, undisturbed systems are likely to be continually in a transient state; they will be equally so under the influence of man. As man's numbers and economic demands increase, his use of resources shifts equilibrium states and moves populations away from equilibrium.

The study of critical transitions in complex systems requires indeed nonequilibrium analysis under the influence of exogenous perturbations. Examples of this are precisely *noise-induced phenomena* in random dynamical systems.

In many models of critical transitions, the system usually remains in a certain range of states, until it reaches a *tipping point* and drifts towards a different set of states into which it tends to remain for a long time as well. As many concepts in complex systems theory, the term “tipping point” is sometimes loosely employed, or not well defined, for which in [55] the authors suggest to restrict its use only for “any situation where accelerating change caused by a positive feedback

drives the system to a new state”. This exemplifies the need of addressing critical transitions in a context-dependent manner, where one should put emphasis on the mechanisms from which these radical changes emerge.

In this thesis we studied partially one of those mechanisms: for one-dimensional random dynamical systems with bounded noise, a *flickering* or *intermittent* behaviour may occur due to a bifurcation of a minimal invariant set. This allows the system to explore a wider range of states which will eventually be reached under the influence of the random perturbations. This is a mechanism for *noise-induced tipping* (or simply N-tipping), where “...noisy fluctuations result in the system departing from a neighbourhood of a quasi-static attractor” [5]. Here, we only studied the mean time for escaping from a metastable set M , while a complete picture of this phenomenon would include a comparison between the mean time the orbits take to escape from M and the mean time they take to return to M .

Intermittent behaviour has been studied in the context of deterministic dynamics as a phenomenon present in different deterministic bifurcations [48, 33]. More recently, it has been shown that *on-off intermittency* in families of iterated function systems under the presence of a zero Lyapunov exponent [30].

In a similar fashion, it is of utter importance to study *early warning signals* of critical transitions based on the specific setting one is working with. These types of indicators are quantitative or qualitative characteristics in the time-series of a dynamical model or from real-world data which anticipate the occurrence (or its possibility) of an abrupt change in the system. Examples of such indicators are *critical slowing down* (measured as a decrease in the speed of recovery under perturbations, an increase in autocorrelations, and rising variance), *skewness* of fluctuations, and *flickering* (see [52] and references therein). Its numerical exploration and mathematical analysis have mostly been done for certain families of *stochastic differential equations* under the presence of *weak noise* [40, 41].

In the context of discrete-time random dynamical systems with bounded noise, an algorithm was designed to anticipate bifurcations [42], which needs sufficient information near the boundary of the minimal invariant sets. As proved in Theorem 3.1, the time needed to gather data near the boundary points is enormously large. Notice how Theorem 3.1 provides with an indicator of bifurcations of minimal invariant sets (for instance, when increasing the noise strength σ to a point where one of the boundaries of the minimal invariant set becomes nonhyperbolic). It is yet to be explored whether these results can be applied to real-world data.

Classical early-warning signals have their flaws when put in practice. In [11], the authors warn

of the present bias when analysing historical data; and in [12], they remark the importance of a model-based analysis as different dynamical behaviour can exhibit similar effects in their time-series. In this regard, in [50] the authors study an early-warning signal for the transition between a normal human cardiac rhythm and an alternating rhythm, modelled with a period-doubling bifurcation in a simple map perturbed with noise. Interestingly, in [31] a Ricker map is studied under the influence of bounded noise, which presents no essential change in the variance or skewness prior to an abrupt change. We hope the work presented in this thesis contributes to the discussion from the somewhat overlooked perspective of dynamical systems perturbed with bounded noise.

We leave for future work the exploration of other noise-induced phenomena related to bifurcations in the underlying deterministic dynamical system, like the *flip bifurcation* where the deterministic system exhibits a period-doubling bifurcation.

Appendix A

Total variation distance

Given two probability measures, μ_1 and μ_2 , on a measurable space $(M, \mathcal{B})$, we define the *total variation distance* between them as

$$\|\mu_1 - \mu_2\|_{TV} := \sup_{A \in \mathcal{B}} |\mu_1(A) - \mu_2(A)|. \quad (\text{A.1})$$

We are interested in the measurable space $(M, \mathcal{B})$, where M is a minimal invariant set for the system (2.1), and $\mathcal{B} \equiv \mathcal{B}(M)$ its Borel σ -algebra.

The use of the norm notation in (A.1) is not an arbitrary one. In general, one can consider the space of bounded signed measures on $(M, \mathcal{B})$, which we denote as $\mathcal{S}_b(M)$. Given $\mu \in \mathcal{S}_b(M)$, it has a unique Jordan decomposition $\mu = \mu_+ - \mu_-$, where both μ^+, μ^- are mutually singular nonnegative and bounded measures on $(M, \mathcal{B}(M))$ (see Theorem 3.4 in [28]). The *total variation measure* of μ is then defined as $|\mu| = \mu_+ + \mu_-$, and $\|\mu\| := |\mu|(X)$. The function $\mu \rightarrow \|\mu\|$ is a norm in $\mathcal{S}_b(M)$ (see for instance Proposition 7.16 in [28], where this result is derived in more generality). The difference of two probability measures, $\mu_1 - \mu_2$ is a signed measure for which $\|\mu_1 - \mu_2\| = 2\|\mu_1 - \mu_2\|_{TV}$.

The notion of convergence in total variation implies *strong convergence of measures*. Recall that a sequence of probability measures $(\mu_n)_{n=1}^\infty$ on $(M, \mathcal{B}(M))$ converge in the strong sense to μ , denoted simply as $\mu_n \rightarrow \mu$, if $\mu_n(A) \rightarrow \mu(A)$ for all $A \in \mathcal{B}(M)$. We can then interpret the total variation norm as a uniform strong convergence over the whole σ -algebra $\mathcal{B}(M)$. We can characterize the total variation distance between probability measures by comparing integrals of measurable and bounded functions, or by Jordan decompositions, as given in the following well

known statement (see for instance Chapter 4 in [45] for discrete state spaces).

Proposition A.1 *Given two probability measures μ_1 and μ_2 on $(M, \mathcal{B}(M))$,*

$$\|\mu_1 - \mu_2\|_{TV} = (\mu_1 - \mu_2)^+(M) = \frac{1}{2} \sup_{\|f\|_\infty \leq 1} \left| \int f d\mu_1 - \int f d\mu_2 \right|, \quad (\text{A.2})$$

where $\mu_1 - \mu_2 = (\mu_1 - \mu_2)^+ - (\mu_1 - \mu_2)^-$ is its Jordan decomposition, and the supremum on the right is taken over all measurable functions with values on the interval $[-1, 1]$. Moreover, if μ_i ($i = 1, 2$) are absolutely continuous with respect to the Lebesgue measure with densities ϕ_i , respectively, then

$$\|\mu_1 - \mu_2\|_{TV} = \frac{1}{2} \|\phi_1 - \phi_2\|_{L^1}. \quad (\text{A.3})$$

Proof. Recall from the Jordan decomposition theorem that there exist measurable sets M_+ , M_- such that $M = M_+ \cup M_-$ with $M_+ \cap M_- = \emptyset$, and $\mu_1 - \mu_2$ can be uniquely decomposed as $\mu_1 - \mu_2 = (\mu_1 - \mu_2)^+ - (\mu_1 - \mu_2)^-$, where $(\mu_1 - \mu_2)^\pm$ are non-negative measures and $(\mu_1 - \mu_2)^\pm(M_\mp) = 0$. Then,

$$\|\mu_1 - \mu_2\|_{TV} \geq \mu_1(M_+) - \mu_2(M_+) = (\mu_1 - \mu_2)^+(M). \quad (\text{A.4})$$

On the other hand, for each $\epsilon > 0$ there is $A \in \mathcal{B}(M)$ such that

$$\|\mu_1 - \mu_2\|_{TV} - \epsilon \leq \mu_1(A) - \mu_2(A) \leq (\mu_1 - \mu_2)^+(A) \leq (\mu_1 - \mu_2)^+(M). \quad (\text{A.5})$$

Since ϵ was arbitrarily small, from (A.4) and (A.5) we get the first equality in (A.2).

For the second equality in (A.2) let us note first that since $\mu_1(M) - \mu_2(M) = 0$, then $(\mu_1 - \mu_2)^+(M_+) = (\mu_1 - \mu_2)^-(M_-)$. By taking $f = \mathbb{1}_{M_+} - \mathbb{1}_{M_-}$,

$$\frac{1}{2} \sup_{\|f\|_\infty \leq 1} \left| \int f d\mu_1 - \int f d\mu_2 \right| \geq \frac{1}{2} [(\mu_1 - \mu_2)(M_+) + (\mu_1 - \mu_2)(M_-)] = (\mu_1 - \mu_2)^+(M). \quad (\text{A.6})$$

For the reverse inequality,

$$\frac{1}{2} \sup_{\|f\|_\infty \leq 1} \left| \int f d\mu_1 - \int f d\mu_2 \right| \leq \frac{1}{2} \sup_{\|f\|_\infty \leq 1} \int |f| d|\mu_1 - \mu_2| \leq (\mu_1 - \mu_2)^+(M), \quad (\text{A.7})$$

where $|\mu_1 - \mu_2| = (\mu_1 - \mu_2)^+ + (\mu_1 - \mu_2)^-$. Inequalities (A.6) and (A.7) imply the second equality in (A.2).

In the case the probability measures μ_1 and μ_2 are absolutely continuous with respect to the Lebesgue measure with densities ϕ_1 and ϕ_2 respectively, we can take $M_+ = \{x \in M \mid \phi_1(x) - \phi_2(x) \geq 0\}$ and $M_- = M \setminus M_+$. Therefore, by taking again $f = \mathbb{1}_{M_+} - \mathbb{1}_{M_-}$,

$$\frac{1}{2} \sup_{\|f\|_\infty \leq 1} \left| \int f d\mu_1 - \int f d\mu_2 \right| \geq \frac{1}{2} \int [(\phi_1 - \phi_2)^+ + (\phi_1 - \phi_2)^-] dLeb = \frac{1}{2} \|\phi_1 - \phi_2\|_{L^1}. \quad (\text{A.8})$$

Conversely,

$$\frac{1}{2} \sup_{\|f\|_\infty \leq 1} \left| \int f d\mu_1 - \int f d\mu_2 \right| \leq \frac{1}{2} \sup_{\|f\|_\infty \leq 1} \int |f| |\phi_1 - \phi_2| dLeb \leq \frac{1}{2} \|\phi_1 - \phi_2\|_{L^1}. \quad (\text{A.9})$$

Therefore, (A.8) and (A.9) imply (A.3). $\square$

Remark A.2 *The second part of A.1 implies that whenever a sequence of absolutely continuous probability measures with respect to Lebesgue converge at a certain rate, the limit is also absolutely continuous and their densities converge at the same rate in L^1 .*

Appendix B

Set-valued dynamics and minimal invariant sets

Consider the random difference equation

$$x_{n+1} = T(x_n) + \sigma\omega_n. \quad (\text{B.1})$$

The random difference equation above induces a set-valued map as follows: given $x \in \mathbb{R}$, we define

$$F(x) := T(x) + \sigma\Delta = \{T(x) + \sigma\omega_0 \mid \omega_0 \in [-1, 1]\}.$$

In other words, $F(x)$ is the set of all possible outcomes under the map h_x . We extend the map F to the space of compact sets of $\mathbb{R}$, $\mathcal{K}(\mathbb{R})$, as

$$F(K) = \bigcup_{x \in K} F(x), \quad (\text{B.2})$$

for all $K \in \mathcal{K}(\mathbb{R})$.

Notice that $F : \mathcal{K}(\mathbb{R}) \rightarrow \mathcal{K}(\mathbb{R})$. Indeed, let $K \subset \mathbb{R}$ be a compact set, and $(x_n)_{n=1}^\infty$ be a sequence in $F(K)$. For each x_n , there are $y_n \in K$ and $\omega_n \in [-1, 1]$ such that $x_n = h(y_n, \omega_n)$ for all $n \in \mathbb{N}$. By compactness of $K \times \Delta$, the sequence $(y_n, \omega_n)_{n=1}^\infty$ has a convergent subsequence $(y_{n_k}, \omega_{n_k}) \rightarrow (y^*, \omega^*) \in K \times \Delta$. Therefore, by the continuity of the random map h , we have $x_{n_k} \rightarrow x^*$, where $x^* = h(y^*, \omega^*)$.

The following definition is crucial for understanding the collective behaviour of the system (B.1).

Definition B.1 (Invariant sets [44]) *A compact set $X \in \mathcal{K}(\mathbb{R})$ is F -invariant (or invariant for the system (B.1)), if $F(X) = X$. We also say $M \in \mathcal{K}(\mathbb{R})$ is a minimal invariant set for F (or for the system (B.1)) if it is an invariant set with no proper compact subsets which is invariant.*

We always assume we restrict (1.1) to an invariant set $X \subset \mathbb{R}$, so that we redefine the random map as $h : X \times \Delta \rightarrow X$. It is well known that (B.1) admits a finite number of minimal invariant sets, where each of them support an ergodic measure [1, 24, 59].

By varying the parameters of the system, like the noise strength $\sigma > 0$, minimal invariant sets can merge or disappear. Assume the random map h depends on a parameter $\lambda \in \Lambda$, where Λ is some metric space, and denote by $\mathcal{M}_\lambda \in \mathcal{K}(X)$ the union of the minimal invariant sets of (B.1) for such value λ . Let F_λ be the corresponding set-valued map (B.2).

Definition B.2 (Topological bifurcation [44]) *We say that F_λ exhibits a topological bifurcation of minimal invariant sets at λ^* if for any open neighbourhood V of λ^* , there does not exist a family of homeomorphisms $(H_\lambda)_{\lambda \in V}$, $H_\lambda : \mathbb{R} \rightarrow \mathbb{R}$, such that H_λ depends continuously on λ , and $H_\lambda(\mathcal{M}_\lambda) = \mathcal{M}_{\lambda^*}$ for all $\lambda \in V$.*

In [42], it is only considered minimal invariant sets $M = [x_-, x_+]$ of one-dimensional set-valued maps, where x_- and x_+ are fixed points of the upper and lower extremal maps $h_\pm(x) = T(x) \pm \sigma$. In other words, $h_+(x_+) = x_+$, and $h_-(x_-) = x_-$. We prove in the next theorem this is indeed the only possible scenario.

Theorem B.3 (Characterization of minimal invariant sets) *Let M be a compact subset of X . Then M is a minimal invariant set for the set-valued map F if and only if $M = [x_-, x_+]$, where*

$$\textbf{(M1)} \quad h_-(x_-) = x_-, \quad h_+(x_+) = x_+ ,$$

$$\textbf{(M2)} \quad h_+(x) > x \text{ for all } x \in [x_-, x_+),$$

$$\textbf{(M3)} \quad h_-(x) < x \text{ for all } x \in (x_-, x_+].$$

Proof. We split the proof in four parts. First we show that if M is F -invariant, then it is the finite union of F -invariant (compact) intervals. Later on, we show that if $M = [x_-, x_+]$ is an invariant

interval, then $h(x)_+ = x_+$ and $h_-(x)$. Afterwards, we prove that **(M1)**, **(M2)**, and **(M3)** imply M is a minimal invariant set, and finally we prove the converse holds.

Claim 1: *let M be F -invariant. Then M is the finite union of disjoint F -invariant compact intervals.*

Indeed, let K be a compact subset of X . Every connected component of $F(K)$ must be an interval contained in X of length at least 2σ . Since X is bounded, there can only be a finite number of connected components. Since $F(K)$ is compact, and each connected component is bounded away from each other, they must be closed and thus compact.

Now, assume $M = \bigcup_{i=1}^n I_i$, where each I_i is an interval. Denote $I_i = [a_i, b_i]$ for $i = 1, 2, \dots, n$, and assume that $b_i < a_{i+1}$ for $i = 1, \dots, n-1$. Observe that it is enough to prove I_1 is F -invariant, since in such a case $\tilde{M} = M \setminus I_1$ is F -invariant too. Then, let us assume first that there is $x \in I_1$ such that $F(x) \cap I_j \neq \emptyset$ for some $j \neq 1$. Then, $F(x) \subset I_j$ as otherwise $F(x)$ would intersect $X \setminus M$. However, by monotonicity of h_+ , there is $y < x$ such that $h_+(y) = b_1$, and by continuity there is $z \in (y, x)$ such that $h_+(z) \notin M$, leading to a contradiction. Therefore $F(I_1) \subset I_1$. In a similar fashion, it follows that $F(x) \cap I_1 = \emptyset$ for all $x \in M \setminus I_1$. However, for each $x \in I_1$ there is $y \in M$ such that $x \in F(y)$, for which such y must be in I_1 . Therefore, $F(I_1) = I_1$.

Claim 2: *let $M = [x_-, x_+]$ be F -invariant. Then x_+ and x_- are fixed points of h_+ and h_- , respectively.*

By the invariance of M , $F(x_+) \subset M$ and thus $h_+(x_+) \leq x_+$. Assume on the contrary that $h_+(x_+) < x_+$. Then, by the monotonicity of h_+ , we get $h_+(x) \leq h_+(x_+) < x_+$ for all $x \in M$. Then M is not F -invariant, which is a contradiction. Analogously, $h_-(x_-) = x_-$.

Claim 3: *if $M = [x_-, x_+]$ for which conditions **(M1)**, **(M2)**, and **(M3)** hold, then M is a minimal invariant set.*

Assume **(M1)**, **(M2)** and **(M3)** hold. It is straightforward to see that $h_+(x) \leq x_+$ for all $x \in M$, since otherwise $h_+(x_+) > x_+$ by the monotonicity of h_+ . In an analogous way $x_- \leq h_-(x)$ for all $x \in M$. Hence, $F(M) \subset M$. Now, because of **(M1)**, **(M2)** and **(M3)**, and the intermediate value theorem, for each $x \in M$, there is $\omega_0 \in \Delta$ such that $h(x; \omega_0) = x$ and $x \in F(x)$. Therefore $M \subset F(M)$, and thus M is F -invariant.

We proceed to prove that M has no proper F -invariant sets. By Claim 1, assume without loss of generality that $I = [a, b] \subset M$ is an F -invariant interval. By Claim 2, $h_-(a) = a$ and $h_+(b) = b$,

and because of **(M1)**, **(M2)** and **(M3)** we have $a = x_-$ and $b = x_+$, so that $I = M$.

Claim 4: if M is a minimal invariant set, then **(M1)**, **(M2)**, and **(M3)** hold.

By minimality and Claim 1, M is a compact interval, say $M = [x_-, x_+]$. Observe that Claim 2 already guarantees that **(M1)** is satisfied.

Notice that since $h_-(x_-) = x_-$, then $h_+(x_-) > x_-$. By continuity, $h_+(x) > x$ for all x close enough to x_- . Let $\tilde{x} = \inf\{x \in M \mid h_+(x) = x\}$, which is bounded away from x_- . Therefore, $\tilde{x} \in (x_-, x_+]$, and by definition the compact interval $\tilde{M} = [x_-, \tilde{x}]$ satisfies conditions **(M1)**, **(M2)** and **(M3)**. Because of Claim 3, $\tilde{M} \subset M$ is a minimal invariant set. Thus, $\tilde{M} = M$, $\tilde{x} = x_+$, and therefore **(M2)** follows. In an analogous way, condition **(M3)** holds. $\square$

Because of Theorem B.3 the bifurcation of minimal invariant sets can be visualized in a similar way as for bifurcation diagrams for deterministic systems. In fact, we make use of the deterministic bifurcation diagrams of both the upper and lower extremal maps, h_+ and h_- , for the toy model in Chapter 1, where we exhibit *transient bifurcations* and *intermittent bifurcations* of the minimal invariant sets [59], see Figure B.1

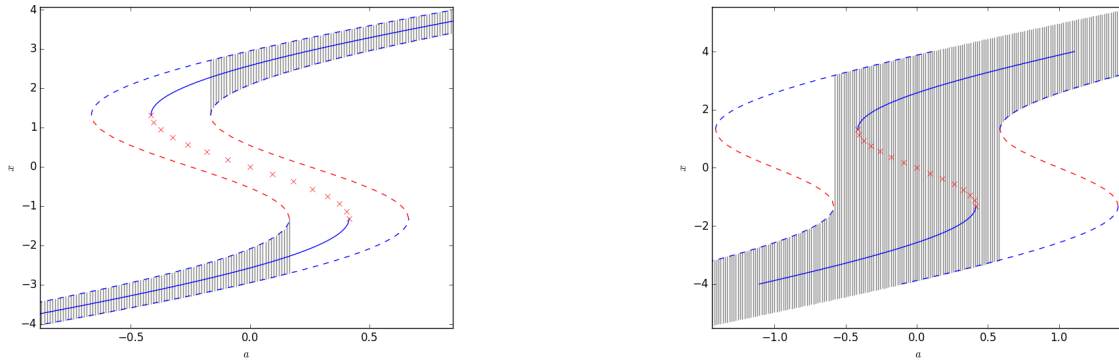


Figure B.1: here we present two different types of bifurcations of minimal invariant sets for the random map $h(x, \omega_0) = a + T(x) + \sigma\omega_0$, where the deterministic map T is as in (1.2), and we have a as a bifurcation parameter in the horizontal axis. On the left, the minimal invariant sets on the upper and lower branches disappear at critical values of the parameter a , exhibiting then *transient bifurcations*. On the right, minimal invariant sets on each of the upper and lower branches are enlarged at critical values of a , exhibiting then *intermittency bifurcations*.

As mentioned above, minimal invariant sets are relevant from a statistical point of view since they support the ergodic measures of (B.1), see Chapter 2.

Appendix C

Functional operators for random maps

In this appendix we provide with a brief but self-contained introduction to the functional operators employed in this thesis. As in the rest of the thesis, we work with random difference equations of the form

$$x_{n+1} = T(x_n) + \sigma\omega_n, \tag{C.1}$$

where $(\omega_n)_{n=0}^\infty$ is an i.i.d. sequence of random variables in $\Delta = [-1, 1]$, according to a probability measure $d\nu = p \, d\text{Leb}$, where p is Lipschitz-continuous.

C.1 Transition probabilities for random maps

Let us assume there is $X \subset \mathbb{R}$ an invariant set for (C.1), as defined in Appendix B, in which we restrict the system. In the following we omit the integration region, as we always assume is X , unless stated otherwise. Denote by $\mathcal{M}_1(X)$ the set of probability measures on $(X, \mathcal{B}(X))$. Recall that system (C.1) induces a cocycle map, φ_ω^n .

When starting at a point $x \in X$, the chance of falling into a measurable set $A \in \mathcal{B}(X)$ is quantified according to the following definitions.

Definition C.1 (Transition probabilities) *The (1-step) transition probability of (C.1), start-*

ing at $x \in X$, is the probability measure $\mathbb{P}_x \in \mathcal{M}_1(X)$, given as

$$\mathbb{P}_x(A) = \int_{\Delta} \mathbf{1}_A(h(x; \omega)) \, d\nu(\omega), \quad (\text{C.2})$$

for all $A \in \mathcal{B}(X)$. More generally, for $n \in \mathbb{N}$ and $x \in X$, the n -step transition probability of (C.1) starting at x is the probability measure

$$\mathbb{P}_x^n(A) = \int_{\Omega} \mathbf{1}_A(\varphi_{\omega}^n(x)) \, d\mathbb{P}(\omega), \quad (\text{C.3})$$

for all $A \in \mathcal{B}(X)$.

Remark C.2 Not only $\mathbb{P}_x$ is a probability measure on $(X, \mathcal{B}(X))$, but for each $A \in \mathcal{B}(X)$ the function $x \mapsto \mathbb{P}_x(A)$ is measurable.

Since the noise distribution ν is absolutely continuous with respect to Lebesgue, then so is $\mathbb{P}_x$ for all $x \in X$. Indeed, let $A = (a, b) \subset X$ be an open interval contained in X , then

$$\mathbb{P}_x(A) = \int_{-1}^1 \mathbf{1}_A(T(x) + \sigma\omega) g(\omega) \, d\omega = \frac{1}{\sigma} \int_A \mathbf{1}_{F(x)}(y) p\left(\frac{y - T(x)}{\sigma}\right) \, dy$$

Recall that p is already supported on $[-1, 1]$, so $\frac{y - T(x)}{\sigma} \in [-1, 1]$ if and only if $y \in F(x) = T(x) + \sigma\Delta$. This way, we can define the transition density, k , as the function

$$\begin{aligned} k : X \times X &\rightarrow \mathbb{R}_+, \\ k(x, y) &= \frac{1}{\sigma} p\left(\frac{y - T(x)}{\sigma}\right). \end{aligned} \quad (\text{C.4})$$

For each $x \in X$ fixed, the function $k(x, \cdot)$ is supported on the set-valued map $F(x)$.

The previous formula for the transition probability $\mathbb{P}_x$ is actually valid for any $A \in \mathcal{B}(X)$, as given by the following proposition.

Proposition C.3 For all $A \in \mathcal{B}(X)$ and $n \geq 2$,

$$\mathbb{P}_x(A) = \int_A k(x, y) \, dy, \quad (\text{C.5})$$

$$\mathbb{P}_x^n(A) = \int_A k_n(x, y) dy, \quad (\text{C.6})$$

where

$$k_n(x, y) = \int_X k(x, z) k_{n-1}(z, y) dz \quad (\text{C.7})$$

Proof. As shown above, formula (C.5) is valid for all A open interval in X . The family of open subintervals of X is a π -system, meaning it is closed under finite intersections, that generates $\mathcal{B}(X)$. On the other hand, observe that the right hand side of (C.5) is a measure itself, as a function of $A \in \mathcal{B}(X)$. Since they coincide in a π -system, then they are equal in the whole σ -algebra generated by such π -system, this being $\mathcal{B}(X)$. The result for $\mathbb{P}_x^n$ follows by induction. $\square$

We extend Definition C.1 to an arbitrary initial distribution. Let $\vartheta \in \mathcal{M}_1(X)$ an arbitrary probability measure. Then, for all $n \in \mathbb{N}$ and $A \in \mathcal{B}(X)$

$$\mathbb{P}_\vartheta^n(A) := \int \mathbb{P}_x^n(A) d\vartheta(x), \quad (\text{C.8})$$

where $\mathbb{P}_\vartheta^1 \equiv \mathbb{P}_\vartheta$. With the notation (C.8), notice that

$$\mathbb{P}_x^n(A) = \mathbb{P}_{\delta_x}^n(A),$$

where δ_x is the *Dirac measure* concentrated at x , that is the measure such that $\delta_x(A) = 1$ if $x \in A$ and $\delta_x(A) = 0$ otherwise.

C.2 The Markov operator and stationary measures

Transition probabilities with arbitrary distributions $\vartheta \in \mathcal{M}_1(X)$ as given in (C.8) motivate the following operator

Definition C.4 (Markov operator) We call the Markov operator of system (C.1) to the linear operator $\mathcal{P} : \mathcal{M}_1(X) \rightarrow \mathcal{M}_1(X)$ defined as

$$\mathcal{P}\vartheta = \mathbb{P}_\vartheta, \quad (\text{C.9})$$

for all $A \in \mathcal{B}(X)$.

Observe that $\mathcal{P}^n \vartheta = \mathbb{P}_\vartheta^n$. Indeed, for any $n \in \mathbb{N}$ and $\vartheta \in \mathcal{M}_1(x)$,

$$\mathcal{P}^n \vartheta(A) = \int \mathbb{P}_x(A) d(\mathcal{P}^{n-1} \vartheta)(x) = \int \mathbb{P}_x^n(A) d\vartheta(x) = \mathbb{P}_\vartheta^n(A),$$

for all $A \in \mathcal{B}(X)$. Because of this, $(\mathcal{P}^n \vartheta)_{n=0}^\infty$ is interpreted as the evolution of the distribution ϑ with the evolution of the system (C.1). A notion of equilibrium for distributions is then given in the following definition.

Definition C.5 (Stationary measure) *A measure $\mu \in \mathcal{M}_1(X)$ is a stationary measure of the system (C.1) (or, for the induced cocycle φ_ω^n), if it is a fixed point of the Markov operator $\mathcal{P}$.*

It is well known that μ is a stationary measure if and only if the product measure $\mu \times \mathbb{P}$ is an invariant measure for the skew-product system $\Phi = (\varphi, \theta)$ on $(X \times \Omega, \mathcal{B}(X) \otimes \mathcal{B}(\Omega))$ [37, 59]. Furthermore,

Definition C.6 (Ergodic measure) *We say that a stationary measure $\mu \in \mathcal{M}_1(X)$ for (1.1) is ergodic if the measure $\mu \times \mathbb{P}$ on $(X \times \Omega, \mathcal{B}(X) \otimes \mathcal{B}(\Omega))$ is ergodic in the classical sense, that is if every invariant set for Φ has either full or null $\mu \times \mathbb{P}$ measure.*

Just like in the deterministic case, if μ is an ergodic measure then by means of *Birkhoff's ergodic theorem*, and considering $\pi_X : \Omega \times X \rightarrow X$ the canonical projection on X , then for each $g \in L^1(\mu)$ we have

$$\lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} g(\varphi_\omega^j(x)) = \lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} g \circ \pi_X(\Phi^j(\omega, x)) = \int_{\Omega \times X} g \circ \pi_X d(\mathbb{P} \times \mu) = \int_X g d\mu,$$

for $(\mathbb{P} \otimes \mu)$ -a.e. $(\omega, x) \in \Omega \times X$.

The existence of stationary measures can be proved in great generality by extending the Markov operator to the vector space of finite signed measures on $(X, \mathcal{B}(X))$. Here, $\mathcal{M}_1(X)$ is a nonempty, convex, and invariant subspace, which is compact in the *weak*-topology* due to *Alaouglu's theorem* [28]. Thus, the existence of fixed points is a direct consequence of the *Tychonoff-Schauder fixed point theorem*¹. Another strategy for proving the existence of stationary measures is to use the *Krylov-Bogoliubov* method [37].

¹For a comprehensive reference on fixed-point theorems, see [58].

The method we employ in Chapter 2 suits the setting here presented, but it is stronger in the sense that not only guarantees the existence, but also its uniqueness, and an exponential rate of convergence of $(\mathcal{P}^n \vartheta)_{n=0}^\infty$ in TV -norm (see Appendix A).

C.3 The stochastic Koopman operator

In deterministic dynamics, when we have a map $T : X \rightarrow X$, the *Koopman operator* applied to a measurable and bounded function f is just the composition $f \circ T$. This operator takes its name after the French-American mathematician Bernard Koopman, and in recent years it has proved to be extremely useful in the machine learning of dynamical systems (see [47] and references therein).

In random dynamics, since we have a family of maps $(h_{\omega_0})_{\omega_0 \in \Delta}$, we can take the average of all of their associated Koopman operators with respect to the probability measure ν on $(\Delta, \mathcal{B}(\Delta))$. More precisely, let us denote the space $L^\infty(X, \mathcal{B}(X), \text{Leb})$ simply as $L^\infty(X)$,

Definition C.7 (Stochastic Koopman operator [19]) *The stochastic Koopman operator of the system (C.1) is the operator $P : L^\infty(X) \rightarrow L^\infty(X)$ such that for all $f \in L^\infty(X)$,*

$$Pf(x) = \int_{\Delta} (f \circ h_{\omega}) (x) d\nu(\omega). \quad (\text{C.10})$$

It is straightforward to see that $\|Pf\|_\infty \leq \|f\|_\infty$, implying that P is a bounded (continuous) operator and $\|P\| \leq 1$. In fact, since $P\mathbf{1}_X = \mathbf{1}_X$ then $\|P\| = 1$.

From (C.2), one can see that

$$\mathbb{P}_x(A) = P\mathbf{1}_A(x).$$

This simple observation is the base for extending properties of the probability measures $\mathbb{P}_x$ to the action of the operator P . For example, we can use the kernel k for calculating Pf as given in the following Proposition.

Proposition C.8 *For all $f \in L^\infty(X)$ and $n \in \mathbb{N}$,*

$$P^n f(x) = \int f(y) k_n(x, y) dy, \quad (\text{C.11})$$

where k_n is the transition density (C.7)

Proof. Take $n = 1$. Equation (C.11) is valid for indicator functions $\mathbb{1}_A$ with $A \in \mathcal{B}(X)$ due to (C.5). Since P is linear, then it is valid for simple functions as well. For general $f \geq 0$, consider a sequence of increasing simple functions, f_n , converging to f . By the monotone convergence theorem [28], the result follows. For $f \in L^\infty(X)$, decompose f in its positive and negative part, f_+ and f_- and apply the formula (C.11) to each of them. For $n \geq 2$, the result follows from Fubini's theorem and an inductive argument. $\square$

It is a direct consequence from Proposition C.8 that, for all $f \in L^\infty$,

$$Pf(x) = \int f(y) d\mathbb{P}_x(y).$$

Moreover, for all $n, m \in \mathbb{N}_0$

$$P^{n+m}f(x) = \int P^n f(y) d\mathbb{P}_x^m(y).$$

C.4 The transfer operator and stationary densities

Recall that for all $x \in X$, the transition probability $\mathbb{P}_x$ is absolutely continuous with respect to Leb, with density $k(x, y)$ as given in (C.4). Hence $\mathbb{P}_\vartheta \ll \text{Leb}$, for all $\vartheta \in \mathcal{M}_1(X)$.

Let $\mathcal{D}(X)$ be the space of probability densities supported on X , that is $g \in \mathcal{D}(X)$ if and only if $g : X \rightarrow \mathbb{R}_+$, and $\|g\|_{L^1} = 1$. For each $g \in \mathcal{D}(X)$, consider the probability measure $\vartheta_g(A) = \int_A g(x)dx$. By the Radon-Nykodym theorem [28], we can associate to g , the unique density function in $L^1(X)$ of the measure $\mathcal{P}\vartheta_g$, which we denote as Lg . This construction can be extended to any $g \in L^1(X)$ by linearity, as given in the following definition.

Definition C.9 (Transfer operator) *The transfer operator (or the Ruelle-Perron-Frobenius operator) of the system (C.1) is the operator $L : L^1(X) \rightarrow L^1(X)$, such that for each $g \in L^1(X)$ we have*

$$\int_A Lg(x)dx = \int \mathbb{P}_x(A)g(x)dx, \tag{C.12}$$

for all $A \in \mathcal{B}(X)$.

Analogously as for the Markov operator, the sequence $(L^n g)_{n=0}^\infty$ can be interpreted as the

evolution of densities with the system (C.1). In particular, for fixed $x \in X$,

$$L(k(x, \cdot))(y) = \int k(z, y)k(x, z)dz = k_2(x, y).$$

furthermore, for $x \in X$ fixed and all $n \in \mathbb{N}$,

$$L(k_n(x, \cdot))(y) = k_{n+1}(x, y). \quad (\text{C.13})$$

There is a close relationship between the stochastic Koopman operator and the transfer operator, as they both act like adjoint operators. We prove this in the next theorem.

Theorem C.10 (Duality between P and L , [59]) *For any $g \in L^1(X)$ and $f \in L^\infty(M)$ we have*

$$\int [f(x) \cdot Lg(x)] dx = \int [Pf(x) \cdot g(x)] dx. \quad (\text{C.14})$$

Proof. Let $g \in L^1(X)$. Notice that (C.14) is true when $f = \mathbb{1}_A$ for arbitrary $A \in \mathcal{B}(X)$ due to (C.12). Moreover, by linearity this equality remains true for any measurable simple function. Thus, by using the monotone convergence theorem it also holds for any non-negative, measurable and bounded function. Again by linearity and by splitting into positive and negative parts, it remains true for any $f \in L^\infty$. $\square$

It is straightforward to see that $L(\mathcal{D}(X)) \subset \mathcal{D}(X)$. Indeed, let $g \in \mathcal{D}$, then from (C.12) we have that for all $A \in \mathcal{B}(X)$,

$$\int_A Lg(x) dx \geq 0.$$

Thus, $Lg \geq 0$ almost everywhere. Furthermore, L preserves integrals since

$$\int Lg(x) dx = \int \mathbb{P}_x(X)g(x)dx = \int g(x) dx.$$

Therefore $L(\mathcal{D}(X)) \subset \mathcal{D}(X)$.

Definition C.5 provides a notion of equilibrium for probability measures under the Markov operator $\mathcal{P}$. We introduce an equivalent notion for densities under the transfer operator L .

Definition C.11 (Stationary density) *We say that $\phi \in \mathcal{D}(X)$ is a stationary density if it is a fixed point of the transfer operator L .*

The relationship between stationary measures and stationary densities is given in the next result.

Proposition C.12 *A probability measure $\mu \in \mathcal{M}_1(X)$ is a stationary measure if and only if $\phi = d\mu/d\text{Leb}$ is a stationary density.*

Proof. Let ϕ be the density function of μ . Then, for every $A \in \mathcal{B}(X)$ we have

$$\int_A \phi \, d\text{Leb} = \mu(A) = \int \mathbb{P}_x(A) \, d\mu(x) = \int \mathbb{P}_x(A) \phi(x) dx.$$

By (C.12) we obtain $L\phi = \phi$. The converse is proved analogously. $\square$

An explicit formula for L in the context of random diffeomorphisms was given in [59]. Observe that if $g \in L^1(X)$, for any $A \in \mathcal{B}(X)$ we get,

$$\int_A Lg(x) \, dx = \int_X g(x) \int_A k(x, y) \, dy \, dx = \int_A \left(\int_X g(x) k(x, y) dx \right) dy.$$

Since this holds for all $A \in \mathcal{B}(X)$ we have that for *Leb*-a.e. $x \in X$,

$$Lg(x) = \int_X k(y, x) g(y) dy. \quad (\text{C.15})$$

Recall that $k(x, y) > 0$ if and only if $y \in F(x)$, where F is the set-valued map associated to (2.1). Thus $y \in [h_-(x), h_+(x)]$ if and only if $x \in F^*(y) = [h_+^{-1}(y), h_-^{-1}(y)]$. Hence, the limits in the last formula for Lg can be restricted to

$$Lg(x) = \frac{1}{\sigma} \int_{h_+^{-1}(x) \vee x_-}^{h_-^{-1}(x) \wedge x_+} p\left(\frac{x - T(y)}{\sigma}\right) g(y) dy, \quad (\text{C.16})$$

where $X = [x_-, x_+]$, since it is invariant for (1.1). Here we denote $a \vee b = \max\{a, b\}$, and $a \wedge b = \min\{a, b\}$.

In the particular case where the sequence $(\omega_n)_{n=0}^\infty$ is i.i.d. uniformly on $[-1, 1]$, then $p \equiv 1/2$ and

$$Lg(x) = \frac{1}{2\sigma} \int_{T^{-1}(x-\sigma) \vee x_-}^{T^{-1}(x+\sigma) \wedge x_+} g(y) dy. \quad (\text{C.17})$$

Let us finish this appendix by proving that the transition densities $k_n : X \times X \rightarrow \mathbb{R}_+$ are continuous, for the setting we deal with in this thesis.

Proposition C.13 *For each $n \geq 2$, the function k_n is continuous in $X \times X$.*

Proof. We first prove that k_2 is continuous. Recall that for each $x \in X$, the density $k(x, \cdot)$ is supported on $F(x) = [h_-(x), h_+(x)]$. Therefore,

$$k_2(x, y) = \int_{h_-(x)}^{h_+(x)} k(x, z)k(z, y)dz$$

Let us take without loss of generality $x < x_0$, where $x, x_0 \in X$ are close enough to each other, so that

$$h_-(x) \leq h_-(x_0) \leq h_+(x) \leq h_+(x_0)$$

Then, for any $y \in X$

$$\begin{aligned} |k_2(x, y) - k_2(x_0, y)| &\leq \\ &\leq \|k\|_\infty^2 [h_-(x_0) - h_-(x) + h_+(x_0) - h_+(x)] + \frac{1}{\sigma} \int_{h_-(x_0)}^{h_+(x)} \left| p\left(\frac{z-T(x)}{\sigma}\right) - p\left(\frac{z-T(x_0)}{\sigma}\right) \right| dz \leq \\ &\leq \left[\|k\|_\infty^2 (\text{Lip}(h_-) + \text{Lip}(h_+)) + \frac{\|k\|_\infty \text{Lip}(p) \text{Lip}(T) \text{Leb}(X)}{\sigma^2} \right] |x - x_0|. \end{aligned}$$

As a summary, there exists $K_1 > 0$ such that for all $y \in X$,

$$|k_2(x, y) - k_2(x_0, y)| \leq K_1 |x - x_0|. \quad (\text{C.18})$$

Consider the continuous functions $H_+(y) = h_+^{-1}(y) \vee x_-$ and $H_-(y) = h_-^{-1}(y) \wedge x_-$. Then, from (C.16) we get

$$k_2(x, y) = \int_{H_+(y)}^{H_-(y)} k(x, z)k(z, y)dz.$$

Analogously to the previous part, for all $x \in X$, and $y, y_0 \in X$ close enough to each other,

$$|k_2(x, y) - k_2(x, y_0)| \leq K_2 (|H_+(y) - H_+(y_0)| + |H_-(y) - H_-(y_0)|) + K_3 |y - y_0|, \quad (\text{C.19})$$

for some $K_2, K_3 > 0$. The continuity of k_2 follows from (C.18) and (C.19).

The statement for arbitrary $n \geq 2$ is proved in a similar fashion. $\square$

Appendix D

Some results on hyperbolic fixed points

In this appendix we develop some standard results on the exponential attractiveness of hyperbolic fixed points of C^1 maps $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$, where h is a local diffeomorphism, that is the Jacobian matrix Dh is invertible for each $x \in \mathbb{R}^n$.

Recall that if (X, d) is a metric space and $h : X \rightarrow X$ has a fixed point $z \in X$, then z is *attracting* (in X) if for all $x \in X$ we have $d(h^n(x), z) \rightarrow 0$ as $n \rightarrow \infty$. The fixed point z is *exponentially attracting* in X if for each $x \in X$ there are $C \equiv C(x) > 0$ and $\rho \equiv \rho(x) \in (0, 1)$ such that

$$d(h^n(x), z) \leq C\rho^n.$$

In this case, we call ρ the rate of convergence. We say that z is *uniformly exponentially attracting* if both C and ρ can be chosen independently of $x \in X$. The fixed point z is called *(uniformly) locally exponentially attracting* if there exists $r > 0$ such that z is uniformly exponentially attracting in $B_r(z)$.

In the following proposition we extend the local exponential attractiveness property of a fixed point of a continuous map $h : X \rightarrow X$ to the whole space X , if X is compact. For this we need the fixed point $z \in X$ to be attracting.

Proposition D.1 *Let $h : X \rightarrow X$ be continuous, where X is a compact metric space. Assume $z \in X$ is a locally exponentially attracting fixed point for h , which is also attracting. Then, it is uniformly exponentially attracting.*

Proof. Since z is locally exponentially attracting, there are $r, C > 0$ and $\rho \in (0, 1)$ such that for all $x \in V = B_r(z)$

$$d(h^n(x), z) \leq C\rho^n. \quad (\text{D.1})$$

On the other hand, since z is attracting, we have $X = \bigcup_{n=0}^{\infty} h^{-n}(V)$. By compactness, there is $N \in \mathbb{N}$ such that $X = \bigcup_{n=0}^N h^{-n}(B_r(z))$. Thus, for all $n > N$ and $x \in X$ we have

$$d(h^n(x), z) = d(h^{n-N}(h^N(x)), z) \leq C\rho^{n-N} = \tilde{C}\rho^n,$$

where $\tilde{C} = C/\tilde{\lambda}^N$.

For $n \leq N$, and since $\text{diam}(X) = \sup_{x,y \in X} d(x, y)$,

$$d(h^n(x), z) \leq \text{diam}(X) \leq \left(\frac{\text{diam}(X)}{\rho^N} \right) \rho^n.$$

The result follows. $\square$

Let us explore the case in which $X = \mathbb{R}^n$, and assume $z \in \mathbb{R}^n$ is a fixed point of a C^1 map $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Assume further that z is hyperbolic and the spectral radius of the Jacobian matrix $\lambda = r(Dh(z))$ satisfies $\lambda < 1$. It is well known that in this case z is locally exponentially attracting. More precisely,

Proposition D.2 *For each $\rho \in (\lambda, 1)$, there is $r > 0$ such that for all $x \in B_r(z)$ and $n \in \mathbb{N}$*

$$\|h^n(x) - z\| \leq r \cdot \rho^n, \quad (\text{D.2})$$

Proof. Since $x \mapsto Dh(x)$ is continuous, and since $\|Dh(z)\| = \lambda$, for each $\epsilon > 0$ there is $r > 0$ such that

$$\|Dh\|_{\infty} := \sup_{x \in B_r(z)} \|Dh(x)\| \leq \lambda + \epsilon.$$

Let us take $\epsilon = \rho - \lambda$. Then, for all $x \in B_r(z)$

$$\|h^n(x) - z\| = \|h^n(x) - h(z)\| \leq \|Dh\|_{\infty}^n \cdot \|x - z\| \leq r \cdot \rho^n$$

$\square$

We now extend Proposition D.2 to the case when we have a parametrized family of smooth maps on $\mathbb{R}^n$. Consider a C^1 function $h : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^n$, where we shall think of $\mathbb{R}^k$ as a parameter space.

Proposition D.3 *Assume there exists $(z^*, \sigma^*) \in \mathbb{R}^n \times \mathbb{R}^k$ such that z^* is a hyperbolic fixed point for the map $h(\cdot, \sigma^*)$ for which $\lambda^* = r(Dh_{\sigma^*}(z^*)) < 1$. Then, there is an open neighbourhood $U \subset \mathbb{R}^k$ of σ^* and a C^1 function $z : U \rightarrow \mathbb{R}^n$ such that $z(\sigma)$ is a locally exponentially attracting fixed point of $h(\cdot, \sigma)$. Moreover, one can choose a rate of convergence $\rho \equiv \rho(\sigma)$ to be continuous with respect to σ in U .*

Proof. It is a direct consequence of the implicit function theorem that there exists a C^1 function $z : U \rightarrow \mathbb{R}^n$, where $U := B_r(\sigma^*) \subset \mathbb{R}^k$ for some $r > 0$, such that for all $\sigma \in U$

$$z(\sigma^*) = z^*, \quad T_\sigma(z(\sigma)) = z(\sigma).$$

Furthermore, $\lambda(\sigma) := r(Dh_\sigma(z(\sigma)))$ is continuous. Therefore, we can assume without loss of generality that $r > 0$ is small enough so that $\lambda(\sigma) \in (0, 1)$ for all $\sigma \in U$. Then, for each $\sigma \in U$ there is a neighbourhood $V \equiv V_\sigma \subset \mathbb{R}^n$ of $z(\sigma)$ in which $z(\sigma)$ is uniformly exponentially attracting.

Moreover, from Proposition D.2 we can take $\rho(\sigma) = \frac{1+\lambda(\sigma)}{2} \in (\lambda(\sigma), 1)$ which is continuous.

□

D.1 Exponential convergence of conditional set-valued dynamics

Consider the set $M \equiv M^\sigma = [x_-(\sigma), x^*]$ from Chapter 4, where $M^* \equiv M^{\sigma^*}$ is invariant for the random map

$$h(x; \omega_0) = T(x) + \sigma\omega,$$

where T is C^1 and $T' > 0$. Recall that $x_-(\sigma)$ is hyperbolic and exponentially attracting for all $\sigma \in [\sigma^*, \sigma^* + \delta_1]$ for some $\delta_1 > 0$.

The set valued-map, $F \equiv F^\sigma : \mathcal{K}(M) \rightarrow \mathcal{K}(M)$ induced by the random map h , where $\mathcal{K}(M)$ is the space of nonempty compact subsets of M , is given as

$$\hat{F}(x) := \hat{F}(\{x\}) = [h_-(x), h_+(x) \wedge x^*]$$

$$\hat{F}(K) = \bigcup_{x \in K} \hat{F}(x),$$

for all $K \in \mathcal{K}(M)$ Since M is compact, the space $\mathcal{K}(M)$ is a compact metric space when endowed with the *Hausdorff distance* [6], where the Hausdorff distance, d_H , is defined as

$$d_H(A, B) = \max\{\sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A)\},$$

for all $A, B \in \mathcal{K}(M)$.

Let $n_+ : M \rightarrow \mathbb{N}$ be the escape time from M of the upper extremal map $h_+(x) = T(x) + \sigma$, and consider $N_0 \equiv N_0^\sigma = n_+(x_-)$. Then, for any $n \geq N_0$ and $x \in M$ we have that $\hat{F}^n(x) = [h_-^n(x), x^*]$. It follows that $\hat{F}^n(x) \rightarrow M$ exponentially fast with respect to the Hausdorff distance.

Proposition D.4 *For each $\sigma \in (\sigma^*, \sigma^* + \delta_1)$ there exists $C \equiv C(\sigma) > 0$ and $\tilde{\lambda} \equiv \tilde{\lambda}(\sigma)$ such that for all $n \geq N_0 = n_+(x_-)$, we have $d_H(M, \hat{F}^n(x)) \leq C\tilde{\lambda}^n$, where d_H is the Hausdorff metric in the space of compact subsets of M .*

Proof. Note that for $n \geq N_0$

$$d_H(M, \hat{F}^n(x)) = |h_-^n(x) - x_-|.$$

Since x_- is locally exponentially attracting, and since it is attracting in M , by Proposition D.2, it is uniformly exponentially attracting in M .

□

Bibliography

- [1] V. Araújo. Attractors and time averages for random maps. In *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis*, volume 17, pages 307–369. Elsevier, 2000.
- [2] V. Araújo and H. Aytaç. Decay of correlations and laws of rare events for transitive random maps. *Nonlinearity*, 30(5):1834, 2017.
- [3] L. Arnold. Random dynamical systems. In *Dynamical systems*, pages 1–43. Springer, 1995.
- [4] S. Arrhenius. On the reaction velocity of the inversion of cane sugar by acids. In *Selected readings in chemical kinetics*, pages 31–35. Elsevier, 1967.
- [5] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox. Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1962):1166–1184, 2012.
- [6] M. F. Barnsley. *Fractals everywhere*. Academic press, 2014.
- [7] I. Bena. Dichotomous markov noise: exact results for out-of-equilibrium systems. *International Journal of Modern Physics B*, 20(20):2825–2888, 2006.
- [8] N. Berglund. Kramers’ law: Validity, derivations and generalisations. *Markov Processes and Related Fields*, 19(3):459–490, 2013.
- [9] R. Bhattacharya and M. Majumdar. *Random dynamical systems: theory and applications*. Cambridge University Press, 2007.
- [10] A. Bittracher, P. Koltai, S. Klus, R. Banisch, M. Dellnitz, and C. Schütte. Transition manifolds of complex metastable systems. *Journal of nonlinear science*, 28(2):471–512, 2018.

- [11] C. Boettiger and A. Hastings. Early warning signals and the prosecutor’s fallacy. *Proceedings of the Royal Society B: Biological Sciences*, 279(1748):4734–4739, 2012.
- [12] C. Boettiger and A. Hastings. Quantifying limits to detection of early warning for critical transitions. *Journal of the Royal Society Interface*, 9(75):2527–2539, 2012.
- [13] N. Champagnat, K. A. Coulibaly-Pasquier, and D. Villemonais. Criteria for exponential convergence to quasi-stationary distributions and applications to multi-dimensional diffusions. In *Séminaire de Probabilités XLIX*, pages 165–182. Springer, 2018.
- [14] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q -process. *Probability Theory and Related Fields*, 164(1-2):243–283, 2016.
- [15] N. Champagnat and D. Villemonais. General criteria for the study of quasi-stationarity. *arXiv preprint arXiv:1712.08092*, 2017.
- [16] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions: Markov chains, diffusions and dynamical systems*. Springer Science & Business Media, 2012.
- [17] F. Colonius. Conditionally stationary measures for random diffeomorphisms. *Procedia IU-TAM*, 6:151–159, 2013.
- [18] F. Colonius, A. J. Homburg, and W. Kliemann. Near invariance and local transience for random diffeomorphisms. *Journal of Difference Equations and Applications*, 16(2-3):127–141, 2010.
- [19] N. Črnjarić-Žić, S. Maćešić, and I. Mezić. Koopman operator spectrum for random dynamical systems. *Journal of Nonlinear Science*, pages 1–50, 2019.
- [20] V. Dakos, E. H. van Nes, and M. Scheffer. Flickering as an early warning signal. *Theoretical ecology*, 6(3):309–317, 2013.
- [21] O. Dembo and A. Zeitouni. Large deviations techniques and applications, 1998.
- [22] M. F. Demers and L.-S. Young. Escape rates and conditionally invariant measures. *Nonlinearity*, 19(2):377, 2005.
- [23] J. Diebolt. Estimates of the tail of the stationary density function of certain nonlinear autoregressive processes. *Journal of Theoretical Probability*, 4(4):655–667, 1991.
- [24] J. L. Doob. *Stochastic processes*, volume 101. New York Wiley, 1953.

- [25] J. Duarte, C. Januário, N. Martins, and J. Sardanyés. Scaling law in saddle-node bifurcations for one-dimensional maps: a complex variable approach. *Nonlinear Dynamics*, 67(1):541–547, 2012.
- [26] M. Engel, J. Lamb, and M. Rasmussen. Conditioned lyapunov exponents for random dynamical systems. *Transactions of the American Mathematical Society*, 372(9):6343–6370, 2019.
- [27] H. Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115, 1935.
- [28] G. B. Folland. *Real analysis: modern techniques and their applications*, volume 40. John Wiley & Sons, 1999.
- [29] W. F. Ford and J. A. Pennline. When does convergence in the mean imply uniform convergence? *The American Mathematical Monthly*, 114(1):58–60, 2007.
- [30] M. Gharaei and A. J. Homburg. Random interval diffeomorphisms. *Discrete & Continuous Dynamical Systems-Series S*, 10(2), 2017.
- [31] A. Hastings and D. B. Wysham. Regime shifts in ecological systems can occur with no warning. *Ecology letters*, 13(4):464–472, 2010.
- [32] C. S. Holling. Resilience and stability of ecological systems. *Annual review of ecology and systematics*, 4(1):1–23, 1973.
- [33] A. J. Homburg and T. Young. Intermittency in families of unimodal maps. *Ergodic theory and dynamical systems*, 22(1):203–226, 2002.
- [34] A. J. Homburg and T. R. Young. Hard bifurcations in dynamical systems with bounded random perturbations. *Regular and Chaotic Dynamics*, 11(2):247–258, 2006.
- [35] A. J. Homburg and T. R. Young. Bifurcations of random differential equations with bounded noise on surfaces. *Topological methods in nonlinear analysis*, 35(1):77, 2010.
- [36] Y. Ilyashenko and W. Li. *Nonlocal bifurcations*. Number 66. American Mathematical Soc., 1999.
- [37] Y. Kifer. *Ergodic theory of random transformations*, volume 10. Birkhauser, 1986.
- [38] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

- [39] C. Kuehn. Scaling of saddle-node bifurcations: degeneracies and rapid quantitative changes. *Journal of Physics A: Mathematical and Theoretical*, 42(4):045101, 2008.
- [40] C. Kuehn. A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics. *Physica D: Nonlinear Phenomena*, 240(12):1020–1035, 2011.
- [41] C. Kuehn. A mathematical framework for critical transitions: normal forms, variance and applications. *Journal of Nonlinear Science*, 23(3):457–510, 2013.
- [42] C. Kuehn, G. Malavolta, and M. Rasmussen. Early-warning signals for bifurcations in random dynamical systems with bounded noise. *Journal of Mathematical Analysis and Applications*, 464(1):58–77, 2018.
- [43] Y. A. Kuznetsov. *Elements of applied bifurcation theory*, volume 112. Springer Science & Business Media, 2013.
- [44] J. Lamb, M. Rasmussen, and C. Rodrigues. Topological bifurcations of minimal invariant sets for set-valued dynamical systems. *Proceedings of the American Mathematical Society*, 143(9):3927–3937, 2015.
- [45] D. A. Levin, Y. Peres, and E. L. Wilmer. Markov chains and mixing times. *American Mathematical Soc., Providence*, 2009.
- [46] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.
- [47] I. Mezic. Koopman operator, geometry, and learning. *arXiv preprint arXiv:2010.05377*, 2020.
- [48] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Communications in Mathematical Physics*, 74(2):189–197, 1980.
- [49] I. Prigogine. The end of certainty: Time, chaos and the new laws of nature. 1997.
- [50] T. Quail, A. Shrier, and L. Glass. Predicting the onset of period-doubling bifurcations in noisy cardiac systems. *Proceedings of the National Academy of Sciences*, 112(30):9358–9363, 2015.
- [51] M. Scheffer. *Critical transitions in nature and society*, volume 16. Princeton University Press, 2009.
- [52] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.

- [53] L. P. Shilnikov and D. V. Turaev. A new simple bifurcation of a periodic orbit of blue sky catastrophe type. *Translations of the American Mathematical Society-Series 2*, 200:165–188, 2000.
- [54] F. Takens. Normal forms for certain singularities of vectorfields. In *Annales de l'institut Fourier*, volume 23, pages 163–195, 1973.
- [55] E. H. van Nes, B. M. S. Arani, A. Staal, B. van der Bolt, B. M. Flores, S. Bathiany, and M. Scheffer. What do you mean, ‘tipping point’? *Trends in ecology & evolution*, 31(12):902–904, 2016.
- [56] A. M. Yaglom. Certain limit theorems of the theory of branching processes. *Doklady Akademii Nauk SSSR*, 56:795—798, 1947.
- [57] J.-C. Yoccoz. Centralisateurs et conjugaison différentiable des difféomorphismes du cercle. *Astérisque*, 231:89–242, 1995.
- [58] E. Zeidler. Nonlinear functional analysis and its applications, i: Fixed points springer. *New York, NY*, 1986.
- [59] H. Zmarrou and A. J. Homburg. Bifurcations of stationary measures of random diffeomorphisms. *Ergodic Theory and Dynamical Systems*, 27(5):1651–1692, 2007.