STATISTICAL PROPERTIES OF CHAOTIC DYNAMICAL SYSTEMS: EXTREME VALUE THEORY AND BOREL-CANTELLI LEMMAS

A Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Chinmaya Gupta

August 2010

STATISTICAL PROPERTIES OF CHAOTIC DYNAMICAL SYSTEMS: EXTREME VALUE THEORY AND BOREL-CANTELLI LEMMAS

Chinmaya Gupta

APPROVED:

Dr. Matthew Nicol, Chairman
Department of Mathematics, University of Houston

Dr. William Ott
Department of Mathematics, University of Houston

Dr. Andrei Török
Department of Mathematics, University of Houston

Dr. Gemunu Gunaratne
Department of Physics, University of Houston

Dean, College of Natural Sciences and Mathematics

Acknowledgements

No list that I may compile will include all the people who have, in some form or the other, contributed to this thesis. With apologies to all those who I may have inadvertently overlooked, I wish to thank, in particular, the following people:

My thesis advisor, Professor Matthew J. Nicol, without whose guidance, extreme patience, generosity, humor, and insight, none of what follows would be attributed to me. I feel immense gratitude for the support, academic and financial, which he provided. It is also due to him that I developed an appreciation of many areas of mathematics outside of dynamical systems, such as statistical genetics, and I thank him for introducing me to the many possibilities that lie in the applications of mathematics to problems in biology;

The members of my committee, Dr. William Ott, Dr. Andrew Török, and Dr. Gemunu Gunaratne, for painstakingly reading through my thesis, suggesting improvements, future projects, and pointing out mistakes;

My collaborators, Dr. Mark Holland, Dr. William Ott, and my advisor, Professor Matthew Nicol. I have learned much from each of them, and it is my earnest hope that in my future career, I will have an opportunity to continue these collaborations;

Dr. Ian Melbourne, for suggesting interesting problems, and Dr. Sébastein Gouëzel and Dr. Dima Dolgopyat for entertaining my queries regarding some implications of their work;

Dr. Ian Melbourne and Dr. Mike Todd for inviting me to speak at the mini-symposium on "Thermodynamic Formalism for Dynamical Systems" at Snowbird in

2009;

Dr. Ian Melbourne and Dr. Henk Bruin for hosting me at Surrey in 2009 for a workshop on dynamical systems;

Dr. Mark Pollicott and Dr. Thomas Jordan, for the “Workshop on Ergodic Theory and Iterated Function Systems” at Warwick in 2007;

Dr. Jeffery Morgan for his financial help in my job hunt, as well as Dr. Matthew O’Malley, Dr. Shanyu Ji, Dr. Mike Field, and Dr. Robert Azencott for writing me letters of recommendation;

My friends, who made it possible for me to keep my sanity, and provided a much appreciated support structure which many international students do not have, but yearn for;

And my parents, for whom no words that I may say will express completely the gratitude that I feel for their affection, advice, and involvement.

STATISTICAL PROPERTIES OF CHAOTIC DYNAMICAL SYSTEMS: EXTREME VALUE THEORY AND BOREL-CANTELLI LEMMAS

An Abstract of a Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Chinmaya Gupta

August 2010

Abstract

In this thesis, we establish extreme value (EV) theory and dynamical Borel-Cantelli lemmas for a class of deterministic chaotic dynamical systems. We establish the distributional convergence (to the three classical extreme value distributions) of the scaled sequence of partial maxima of some time series arising from an observable on systems such as the planar dispersing billiards, Lozi-like maps, and compact group skew-extensions of non-uniformly hyperbolic base maps with Hölder cocyles. We also establish Borel-Cantelli lemmas for a large class of one-dimensional non-uniformly expanding maps, and for these, we also obtain an almost sure characterization of the exceedences of the sequence of partial maxima.

Contents

1 Introduction **1**

1.1 Ergodic Theory 2

1.2 Borel-Cantelli Lemmas 9

1.3 Extreme Value Theory 11

1.3.1 Extreme value theory for i.i.d. processes 11

1.3.2 Extreme value theory for dependent processes 14

1.4 Extreme Value Theory: Applications 16

1.5 Extreme Value Theory for Dynamical Systems 18

1.5.1 Hitting time statistics and return time statistics 20

1.5.2 Extreme value theory for skew-extensions 21

1.5.3 Extreme value theory for dispersing billiards and Lozi-like maps 23

1.6 Dynamical Borel-Cantelli Lemmas 26

1.7 Young Towers 29

2 EVT for Skew Extensions **34**

2.1 Framework of the Problem 34

2.2 Preliminaries 39

2.3 Proof of Theorem 44

2.4 Applications and Examples 49

2.4.1 Uniformly and non-uniformly expanding maps of an interval modeled by Young towers 49
2.4.2 Skew product with a curve of neutral points 51
2.4.3 Some other extensions 52

3 EVT for Uniformly Hyperbolic Maps Exhibiting Singularities 54
3.1 Statement of Results 54
3.1.1 Framework of the proof 55
3.2 Condition $D_2(u_n)$ 57
3.2.1 The blocking argument. 57
3.3 Proof of $D_2(u_n)$. 58
3.4 Applications 63
3.4.1 Planar dispersing billiard maps and flows 63
3.4.2 Lozi-like maps 72

4 Some Borel-Cantelli Lemmas 82
4.1 Setting and Statements of Results 82
4.1.1 Gibbs-Markov maps 82
4.1.2 One-dimensional maps 89
4.2 Proof for the Gibbs-Markov Case 90
4.3 Proofs 95
4.3.1 Relating base dynamics to tower dynamics and preliminaries . 95
4.3.2 Proofs of Theorem 4.9 97
4.3.3 Proof of Theorem 4.12 102
4.4 Almost-sure Results in EVT for Dynamical Systems 102

Bibliography 105

Chapter 1

Introduction

Chaotic dynamical systems are those which display a sensitive dependence on initial conditions. These systems are completely deterministic in that if the current state of the system x_0 is known, all future states at time n, denoted by x_n, are known. However, even small numerical errors in the computation of x_0 can lead to very large differences in the values of x_n, so while these systems are deterministic, they are not predictable. One of the most popular examples of chaotic systems was discovered by Edward Lorenz in 1958. He observed that his computer model for the weather, a system of differential equations with twelve variables evolved very differently even when the initial conditions changed only very slightly [58]. Further work on Lorenz's observation lead to the discovery of a very simple deterministic system, a system of three ordinary differential equations [51] which displayed an extreme sensitivity to initial conditions.

This discovery by Lorenz came in the 1960's, around which time a great amount of research was being done on dynamical systems. Much of the groundwork for the study of chaotic dynamical systems was laid down by Poincaré in 1899 in his treatise [61] on celestial mechanics. Van der Pol discovered around the 1930's some experimental evidence for the existence of "deterministic chaos" [69] in the form of a complicated unstable invariant set. Cartwright and Littlewood [18], Hopf [42], Smale [64, 65], and others* then laid the foundations of the modern study of chaotic deterministic systems†.

Chaotic systems exhibit a large departure in the long-term evolution for very tiny changes in initial conditions, and hence to study the limiting behavior of these systems, one cannot study the global evolution orbit by orbit. Instead, one studies the statistical properties of these systems, as we describe in Section 1.1.

1.1 Ergodic Theory

Ergodic theory is the study of measure-preserving systems. One of the fundamental and remarkable theorems in ergodic theory is Birkhoff's Ergodic Theorem, which states that if $T : X \to X$ is a transformation preserving the probability measure μ and if $f \in L^1(\mu)$ is an "observable" ($f : X \to \mathbb{R}$), then for μ almost every $x \in X$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x) = f^*(x) \tag{1.1.1}$$

*See also page 9

†A substantially more complete, and chronologically faithful, history of the evolution of the field of chaotic dynamical systems can be found in Viana's book [70].

for some $f^* \in L^1(\mu)$ with $f^* \circ T = f^*$ and

$$\int f^* d\mu = \int f d\mu. \tag{1.1.2}$$

In particular, when the transformation T satisfies the additional property of being ergodic ($g \in L^1(\mu)$, $g \circ T = g$ then g is constant μ almost everywhere), then $f^* = \int f d\mu$.

The ergodic theorem, however, does not assume or prove ergodicity, and some of the earliest work in identifying concrete systems which display ergodicity was done by Anosov [6], Pesin [59, 11], Sinai [72], et cetera. They showed that a wide class of "chaotic" maps exhibit natural ergodic invariant measures, the maps being "chaotic" in the sense that the orbits of two points, which may be arbitrarily close to each other with respect to a metric, may evolve very differently over time. "Chaos", in this respect, refers to a huge difference in the evolution of the orbits of nearby points, and hinders the orbit-wise study of the dynamical system; the ergodicity of T, on the other hand, says that for typical observables f, the time-series generated by evaluating f along the orbit of almost every initial condition has the same mean. Ergodicity is thus the "law of large numbers" for dynamical systems. While the ergodic theorems prove convergence to the mean, natural questions arise about the deviation of these processes from the mean and about other statistical properties such as rates of mixing.

Let $(\Omega, \mathcal{B}, P)$ be a probability space with $\mathcal{B}$ the set of Borel-measurable subsets of Ω. A random variable ξ is a measurable function with domain Ω and codomain $\mathbb{R}$. A sequence of random variables $\xi_1, \ldots, \xi_n, \ldots$ is called a stochastic process. Two random variables ξ_1 and ξ_2 are said to be independent if for every pair V, W of Borel subsets of $\mathbb{R}$, $P(\xi_1 \in V, \xi_2 \in W) = P(\xi_1 \in V)P(\xi_2 \in W)$. Extending this notion of pairwise independence, we call a stochastic process (ξ_i) independent if for every $N \in \mathbb{N}$, for every sequence of sets $V_1 \in \mathcal{B}, \dots, V_N \in \mathcal{B}$ and for every set $\{j_1, \dots, j_N\} \subset \mathbb{N}$, $P(\xi_{j_i} \in V_1, \dots, \xi_{j_N} \in V_N) = \prod_{i=1}^{N} P(\xi_{j_i} \in V_i)$. A stochastic process (ξ_i) is said to be stationary if for every $N, M \in \mathbb{N}$, any ordered tuple $(j_1, \dots, j_N) \in \mathbb{N}^N$ with $j_1 < j_2 < \dots < j_N$, and any sequence of Borel measurable subsets $V_1, \dots, V_N$,

$$P(\xi_{j_1} \in V_1, \dots, \xi_{j_N} \in V_N) = P(\xi_{j_1+M} \in V_1, \dots, \xi_{j_N+M} \in V_N).$$

It is an easy observation that for an independent stochastic process, stationarity is equivalent to the following: for every $(i, j) \in \mathbb{N} \times \mathbb{N}$ and for every $v \in \mathbb{R}$, $P(\xi_i \leqslant v) = P(\xi_j \leqslant v)$.

Many statistical properties, such as the Central Limit Theorem and the Law of Iterated Logarithm, displayed by independent identically distributed stochastic processes (henceforth called i.i.d. processes), have been shown to hold for stochastic processes generated by chaotic dynamical systems[‡]. One of the better known techniques for establishing such results was proved by Young in [75], where she establishes that many dynamical systems can be modeled by "towers" and in some senses behave as Markov chains on spaces with infinitely many states (see Section 1.7). She proves that for such dynamical systems, should they mix sufficiently quickly, the process $f \circ T^n$ satisfies the Central Limit Theorem. She establishes that the completely deterministic process $f \circ T^n$ behaves, at least in the first two moments, as an

[‡]For an example of a stochastic process generated by a dynamical system, see, for instance, Section 1.5.

independent stochastic process, and the degree of independence is characterized by the rate at which the process mixes.

We now make a few definitions which will carry us through most of what follows. Let X be a compact connected $C^{1+\delta}$ manifold, and let $T : X \to X$ be a smooth ($C^{1+\delta}$) map. We will denote the tangent space at $x \in X$ by $\mathfrak{T}_x X$. DT will denote derivative of T (the matrix of partial derivatives). Definitions analogous to the ones we provide here can be made for continuous-time systems (flows) instead of discrete-time maps.

Definition 1.1 (Expanding Map). *(see [3, Example 2.1]) We say T is expanding if there exists a $\lambda > 1$ such that*

$$\|DT(x)v\| \geqslant \lambda \|v\| \quad \text{for all } x \in X \text{ and } v \in \mathfrak{T}_x X, \tag{1.1.3}$$

for some Riemannian metric $\|\cdot\|$.

We call a C^r manifold $\tilde{X}$ an extension of a C^r manifold X if $\dim \tilde{X} = \dim X$ and if $\bar{X} \subset \tilde{X}$. The manifold X is said to be extendible if there exists an extension of X. A submanifold $Y \subset X$ is said to be extendible if there exists an extension $\tilde{Y}$ of Y which is itself a submanifold of X.

Definition 1.2 (Piecewise Expanding). *(see [23, Pages 1746-1747]) Let $\{\pi_i\}$ be a finite collection of disjoint subsets of X such that each $\partial\pi_i$ is the closure of a union of finitely many disjoint, connected, codimension-one, extendible, $C^{1+\delta}$ submanifolds of π_i. We say that T is piecewise expanding if $T|_{\pi_i} : \pi_i \to X$ is an expanding map for each i.*

Definition 1.3 (Uniformly Hyperbolic)**.** *(see [13]) We say that a map T is uniformly hyperbolic if there exists a compact invariant set Λ and constants $C > 0$ and $\lambda \in (0,1)$, such that for every $x \in \Lambda$, there is a splitting of $\mathfrak{T}_x X$ into $E^s(x) \oplus E^u(x)$ satisfying*

$$DT(E^i(x)) = E^i(Tx), \quad i \in \{u, s\},$$

and for every $n \in \mathbb{N}$,

$$\|DT^n v\| \leqslant C\lambda^n \|v\| \quad \forall v \in E^s(x) \tag{1.1.4}$$

and

$$\|DT^{-n} v\| \leqslant C\lambda^n \|v\| \quad \forall v \in E^u(x). \tag{1.1.5}$$

Sometimes, the condition in equation (1.1.5) *is also written as*

$$\|DT^n v\| \geqslant C\lambda^{-n} \|v\| \quad \forall v \in E^u(x).$$

Note that in the context of the splittings of the tangent space, the expanding maps are those for which $\mathfrak{T}_x X$ splits as $E^u(x) \oplus \{0\}$ and $\Lambda = X$. We refer to $E^u(x)$ as the unstable direction at x and $E^s(x)$ as the stable direction at x.

Definition 1.4 (Non-uniformly Hyperbolic)**.** *(see [11, Section 2.2]) We say the smooth map T is non-uniformly hyperbolic if there exists an invariant set Λ, functions C, K, and constants $0 < \lambda_1 < 1$, $\lambda_2 > 1$ and $\epsilon \in (0, \min\{-\log \lambda_1, \log \lambda_2\})$ such that for every $x \in \Lambda$, there is a splitting of $\mathfrak{T}_x X$ into $E^u(x) \oplus E^s(x)$ satisfying*

$$DT(E^i(x)) = E^i(Tx), \quad i \in \{u, s\},$$

and for every $n \in \mathbb{N}$,

$$\|DT^n v\| \leqslant C(x)\lambda_1^n e^{\epsilon n}\|v\| \quad \forall x \in E^s(x), \tag{1.1.6}$$

$$\|DT^n v\| \geqslant C(x)^{-1}\lambda_2^n e^{-\epsilon n}\|v\| \quad \forall x \in E^u(x), \tag{1.1.7}$$

$\angle(v, w) \geqslant K(x)$ *for every* $v \in E^s(x), w \in E^u(x)$, $C \circ T^n(x)/C(x) \leqslant e^{\epsilon n}$ *and* $K \circ T^n(x)/K(x) \geqslant e^{-\epsilon n}$.

Definition 1.4 is a generalization of Definition 1.3 in that the expansion and contraction rates can depend on the point in question, rather than being the same for all points. The condition on the inner product allows for the stable and unstable directions to come arbitrarily close; this usually adds an order of magnitude to the complexity of the analysis for such systems. In case $E^s(x) \equiv \{0\}$, we have

Definition 1.5 (Non-uniformly Expanding). *We say that* T *is non-uniformly expanding if there exists a constant* $\lambda > 1$ *and a function* $C(x) > 0$*, such that for every* $x \in X$*,*

$$\|DT^n v\| \geqslant C(x)\lambda^n\|v\| \quad \forall v \in \mathfrak{T}_x X.$$

We can further generalize the setting of Definition 1.4 by introducing a 'central' direction which neither exhibits asymptotic expansion, nor contraction.

Definition 1.6 (Non-uniformly partially hyperbolic). *If in Definition 1.4 we replace the conditions* $0 < \lambda_1 < 1, \lambda_2 > 1$ *by* $0 < \lambda_1 < \min\{1, \lambda_2\}$*, the map is called non-uniformly partially hyperbolic with a center-unstable direction and a stable direction. Similarly, replacing* $0 < \lambda_1 < 1, \lambda_2 > 1$ *by* $\lambda_2 > \max\{1, \lambda_2\} > 0$*, we get* *a non-uniformly partially hyperbolic map with a center-stable direction and unstable direction.*

Definition 1.7 (Uniformly partially hyperbolic)**.** *If in Definition 1.6 we require that $C(x)$ be bounded above by some constant C, and $K(x)$ be bounded below by some constant K, we have a uniformly partially hyperbolic map T.*

A substantial body of literature exists which proves that completely deterministic "chaotic" maps, such as the ones satisfying Definitions 1.1, 1.2, 1.3, 1.4, 1.6, and 1.7, generate stochastic processes which behave like i.i.d. processes; see [56] for many classical limit theorems such as the Central Limit Theorem, the law of iterated logarithms, and the almost sure invariance principle (approximation by a Brownian motion) for time-one maps of hyperbolic flows; [24] for many classical limit theorems for uniformly partially hyperbolic systems with exponential decay of correlations for C^k functions; [38] for convergence to a Poisson process of the normalized returns of an orbit to a neighborhood of a point in the non-wandering set of an Axiom-A diffeomorphism; and [55], [73], [8] for large deviations for these systems.

Many books have been written on the subject of chaotic dynamical systems. [13] provides an introduction to hyperbolic dynamical systems with emphasis on symbolic systems, one-dimensional dynamics and measure-theoretic entropy. Pilyugin [60] discusses in depth the technique of "shadowing", and discusses topics in topological stability, structural stability and numerical aspects. Katok [43] provides a comprehensive overview of the modern theory of dynamical systems and its connections with the various other branches of mathematics. Anosov [7] provides a discussion of hyperbolic dynamical systems theory, as well as modifications to the hyperbolic theory which include systems such as the Lorenz attractor. The role of invariant manifold theory for dynamical systems has been studied in [39]. Operator-theoretic techniques for obtaining invariant measures and studying decay of correlations and other statistical properties has been discussed in Baladi [9] [§].

1.2 Borel-Cantelli Lemmas

Borel-Cantelli lemmas are a fundamental tool used to establish the almost-sure behavior of random variables. For example, a Borel-Cantelli lemma is used in the standard proof that Brownian motion has a version with continuous sample paths.

Suppose $(X, \mathcal{B}, \mu)$ is a probability space. For a measurable set $A \subset X$, let $\mathbf{1}_A$ denote the characteristic function of A. We abbreviate the standard terms "infinitely often" to i.o., "almost every" to a.e., and "almost surely" to a.s.. The phrases a.e. and a.s. have the same meaning and we use them interchangeably. The classical Borel-Cantelli lemmas (see for example [25, Section 4]) state that

1. if $(A_n)_{n=0}^{\infty}$ is a sequence of sets in $\mathcal{B}$ and $\sum_{n=0}^{\infty} \mu(A_n) < \infty$ then

$$\mu(\{x \in X : x \in A_n \text{ i.o.}\}) = 0;$$

2. if $(A_n)_{n=0}^{\infty}$ is a sequence of independent events in $\mathcal{B}$ and $\sum_{n=0}^{\infty} \mu(A_n) = \infty$, then

$$\frac{\sum_{i=0}^{n-1} \mathbf{1}_{A_i}}{\sum_{i=0}^{n-1} \mu(A_i)} \to 1 \quad \text{a.s..}$$

[§] A very comprehensive bibliography on dynamical systems can be found in [43].

We establish some more notation. Suppose $T : X \to X$ is a measure-preserving transformation of the probability space $(X, \mathcal{B}, \mu)$. Suppose that $(A_n)_{n=0}^{\infty}$ is a sequence of sets in $\mathcal{B}$ such that $\sum_{n=0}^{\infty} \mu(A_n) = \infty$. For $n \in \mathbb{N}$, let $E_n = \sum_{i=0}^{n-1} \mu(A_i)$ and define $S_n : X \to \mathbb{Z}^+$ by

$$S_n(x) = \sum_{i=0}^{n-1} \mathbf{1}_{A_i} \circ T^i(x).$$

Definition 1.8. *We will call a sequence (A_n) a*

1. ***Borel-Cantelli (BC) sequence*** *if $\mu(\{x \in X : T^n(x) \in A_n \ i.o.\}) = 1$;*

2. ***strong Borel-Cantelli (sBC) sequence*** *if*

$$\lim_{n\to\infty} \frac{S_n(x)}{E_n} = 1 \quad a.s.;$$

3. ***dense Borel-Cantelli (dBC) sequence*** *with respect to the measure γ if there exists $C > 0$ for which*

$$\underline{\lim}_{n\to\infty} \frac{S_n(x)}{\sum_{i=0}^{n-1} \gamma(A_i)} \geqslant C \quad a.s..$$

In the context of dynamical systems, Borel-Cantelli lemmas can be used to study the asymptotics of the frequency with which orbits visit a sequence of sets. For the constant sequence, it is easy to see that the strong Borel-Cantelli property (see 1.8) is equivalent to the conclusion of the Birkhoff Ergodic Theorem, assuming ergodicity.

To see this, we argue as follows. Suppose μ is ergodic and A is a Borel-measurable subset of X with $\mu(A) > 0$. Let $A_n := A$ be the constant sequence. By the ergodic theorem, for a.e. $x \in X$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} 1_{A_n} \circ T^j(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} 1_A \circ T^j(x) = \mu(A) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(A_n).$$

It then follows that $S_n(x)/E_n \to 1$ for a.e. $x \in X$. Conversely, if $S_n(x)/E_n \to 1$ for some $x \in X$, then for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for $n \geqslant N$

$$\mu(A)(1-\epsilon) \leqslant \frac{1}{n}\sum_{j=0}^{n-1} 1_A \circ T^j(x) \leqslant \mu(A)(1+\epsilon)$$

and so

$$\left|\frac{1}{n}\sum_{j=0}^{n-1} 1_A \circ T^j(x) - \mu(A)\right| \leqslant \mu(A)\epsilon \leqslant \epsilon.$$

Hence for such x, $\frac{1}{n}\sum_{j=0}^{n-1} 1_A \circ T^j(x) \to \mu(A)$ and $n \to \infty$.

There has been some recent interest in studying Borel-Cantelli lemmas for dynamical systems, but results in this field are still scarce¶.

1.3 Extreme Value Theory

1.3.1 Extreme value theory for i.i.d. processes

Let $\xi_1, \xi_2, \ldots, \xi_n$ be a sequence of independent, and identically distributed (i.i.d.) random variables, and let F denote the common distribution function for ξ_i. By definition, $P(\xi_i \leqslant x) = F(x)$. If one were to denote by M_n the maximum of $\xi_1, \ldots, \xi_n$, then it is easy to see that

$$P(M_n \leqslant x) = F(x)^n \to \begin{cases} 1 & x \in \{y : F(y) = 1\} \\ 0 & x \in \{y : 0 \leqslant F(y) < 1\} \end{cases}.$$

¶For more on dynamical Borel-Cantelli lemmas, see Section 1.6 on Page 26

This limiting distribution is degenerate, and provides little useful information about the asymptotic properties of the sequence ξ_i. However, there may exist scaling sequences a_n and b_n such that the random variable $\tilde{M}_n := (M_n - b_n)/a_n$ has a nontrivial limiting distribution. This idea of finding scaling sequences is not alien to statistics, and is used in, for instance, the Central Limit Theorem where the sequence $\sum_{i=1}^{n} \xi_i$ is scaled by subtracting $n\mathbb{E}[\xi_1]$ and dividing by $\sqrt{n\mathbb{E}[(\xi_0 - \mathbb{E}(\xi_0))^2]}$ to obtain the standard Normal distribution as the asymptotic limit.

The central questions in extreme value theory for i.i.d. random processes are to establish the existence of scaling sequences a_n and b_n so that

$$P(\tilde{M}_n \leqslant v) \to_d G(v) \;\forall\; v \in \mathbb{R}, \tag{1.3.8}$$

where G is a non-degenerate limiting distribution and to characterize all distributions G which can arise as such limits[‖].

A remarkable theorem, the Extremal Types Theorem (see [47]), states that the only non-degenerate G which can arise as limits in (1.3.8) are the following three parametric forms called the Extreme Value Distributions:

$$\textbf{Type I} \quad G(x) = e^{-e^{-x}}, -\infty < x < \infty \qquad ;$$

$$\textbf{Type II} \quad G(x) = \begin{cases} 0, & x \leqslant 0 \\ e^{-x^{-\alpha}}, & x > 0 \end{cases} \quad \alpha > 0;$$

$$\textbf{Type III} \quad G(x) = \begin{cases} e^{-(-x)^{\alpha}}, & x \leqslant 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0.$$

[‖] From now on, whenever $G(v)$ is a distribution as in equation (1.3.8), we will use $\to$ instead of $\to_d$ to denote convergence in distribution.

These are the only non-degenerate limits up to type, where two distributions are said to be of the same type if $G_1(v) = G_2(av+b)$ for some constants $a > 0, b \in \mathbb{R}$ and for every $v \in \mathbb{R}$.

The extremal types theorem establishes that the distributional limit of $\tilde{M}_n$, if it exists, must be either Type I, II, or III. However, to show the existence of the limiting distribution, one still needs to establish the existence of the scaling sequences a_n and b_n. For i.i.d. processes, this is relatively easy: one need only find a_n and b_n so that for each $v \in \mathbb{R}$,

$$n(1 - F(a_n v + b_n)) \to \tau(v) \quad \text{as } n \to \infty \tag{1.3.9}$$

where $0 < \tau(v) < \infty$. The limit in (1.3.9) holds if and only if $P(M_n \leqslant a_n v + b_n) \to e^{-\tau(v)}$, which happens if and only if $P(\tilde{M}_n \leqslant v) \to e^{-\tau(v)}$. Together with the extremal types theorem, this completely determines the extreme value distributions.

An easy illustration of the idea is as follows. Let ξ_i be an i.i.d. sequence sampled from the uniform distribution on $[0,1]$. For each $v \in (-\infty, 0]$, we wish to obtain $\tau(v) > 0$ such that

$$n(1 - F(a_n v + b_n)) = n(1 - a_n v - b_n) \to \tau(v).$$

Choose, for instance, $a_n = 1/n, b_n = 1$. Then $n(1 - v/n - 1) = -v$ and so $\tau(v) = -v$. The limiting distribution for $P(M_n \leqslant v/n + 1) = e^{-(-v)} = e^v$ which is a Type III distribution.

1.3.2 Extreme value theory for dependent processes

For dependent processes, establishing extreme value theory is considerably more difficult than for independent processes. The general strategy in the dependent case is to establish that the dependent process is close to an independent process in some quantifiable way, to establish extreme value theory for the independent process, and to then show that the same theory applies also to the dependent process.

One way of showing that the dependent process is close to an independent process is to check if the following conditions, $D(u_n)$ and $D'(u_n)$, hold for an appropriately chosen scaling sequence u_n.

Condition $D(u_n)$**:** The condition $D(u_n)$ will be said to hold if for any integers $1 \leqslant i_1 < \cdots < i_p < j_1 < \cdots < j_{p'} \leqslant n$ for which $j_1 - i_p \geqslant l$, we have

$$\left| F_{i_1,\ldots,i_p,j_1,\ldots,j_{p'}}(u_n) - F_{i_1,\ldots,i_p}(u_n) F_{j_1,\ldots,j_{p'}}(u_n) \right| \leqslant \alpha(n,l)$$

where $F_{n_1,\ldots,n_t}$ denotes the joint distribution of $\xi_{n_1} \ldots, \xi_{n_t}$, $\alpha(n, l_n) \to 0$ as $n \to \infty$, $l_n \to \infty$ and $l_n/n \to 0$.

Condition $D'(u_n)$**:** The condition $D'(u_n)$ will be said to hold for the stationary sequence ξ_i and the sequence u_n if

$$\limsup_{n\to\infty} n \sum_{j=2}^{[n/k]} P\left(\xi_1 > u_n, \xi_j > u_n\right) \to 0$$

as $k \to \infty$.

Condition $D(u_n)$ establishes that if two large blocks are sufficiently far apart, then the joint distribution of the two blocks is approximately the product of the distributions on the individual blocks. Condition $D'(u_n)$, on the other hand, is a non-clustering condition. $D'(u_n)$ states that if a large reading is observed (say that the level u_n) at some time $j < n$, then one must wait for a large time $o(n) \to \infty$ before another reading larger than or equal to u_n is observed.

If $D(u_n)$ and $D'(u_n)$ hold, then one can establish the extreme value theory for the dependent process by constructing an independent process with the same distribution function and establishing extreme value theory for the independent process. The precise strategy is given by the following

Theorem 1.9. *(see [47]) Let ξ_i be a stationary, dependent, stochastic process and suppose that there exist scaling sequences a_n and b_n such that the sequence u_n defined as*

$$u_n(v) := \frac{v}{a_n} + b_n$$

satisfies conditions $D(u_n(v))$ and $D'(u_n(v))$ for every $v \in \mathbb{R}$. Let ζ_i be an independent, identically distributed stochastic process with the same distribution function as ξ_i, that is,

$$P(\xi_1 \leqslant t) = P(\zeta_1 \leqslant t) \ \ \forall t \in \mathbb{R}.$$

Then,

$$P(\max\{\zeta_0, \zeta_1, \dots, \zeta_n\} \leqslant u_n(v)) \to F(v)$$

for some non-degenerate distribution function F, if and only if

$$P(\max\{\xi_0, \xi_1, \dots, \xi_n\} \leqslant u_n(v)) \to F(v).$$

It then follows that for a stationary dependent process ξ_i, to establish EVT, one need find sequences a_n and b_n for which equation (1.3.9) holds, and which satisfy conditions $D(u_n)$ and $D'(u_n)$.

1.4 Extreme Value Theory: Applications

Extreme value theory has traditionally been used for predicting risk. For industries such as insurance, extreme value theory is used to predict, and prepare for, the events in which large payouts are required by the company. EVT has been used as a predictive tool for studying temperature variation, flood levels in rivers, waiting times between large floods, air pollution, and ozone levels.

We will, very briefly, describe an application of EVT for studying strength of materials.[**] In the study of strength of a material subject to tension (such as a metal wire supporting a weight), it has been determined empirically that the distribution of the breaking tension depends on the length of the wire (in a regular way). Let us call F_l the distribution of the breaking tension of a piece of wire of length l. We will study the distribution F_1 of the breaking strength of a piece of unit length of this wire.

Suppose we break up a unit length of wire into n equal pieces of length $1/n$ each. We will assume that these pieces of wire are independent of each other; the distribution of the breaking strength of the piece of length $1/n$ is denoted by $F_{1/n}$. Now, the original wire does not break under weight x if none of the $1/n$ pieces break

[**]More complete details, as well as further applications and references can be found in [47].

under x, and so we must have

$$1 - F_1(x) = (1 - F_{1/n}(x))^n.$$

Under the intuitive assumption that the distribution of the unit length piece is of the same type as that of the piece of length $1/n$, we may write $F_{1/n}(x) = F_1(a_n x + b_n)$ for some sequences $a_n > 0$ and $b_n \in \mathbb{R}$. This implies that $1 - F_1(x) = (1 - F_1(a_n x + b_n))^n$, from where it follows that the distribution $G(x) := 1 - F_1(-x)$ is max stable. The extremal types theorem now implies that $G(x)$ can have one of the three classical forms, so $F_1(x)$ must be one of the following:

$$F_1(x) = \begin{cases} \textbf{Type I} : 1 - \exp(-e^x) & -\infty < x < \infty \\ \textbf{Type II} : \begin{cases} 1 - \exp(-(-x)^{-\alpha}) & x < 0 \\ 1 & x \geqslant 0 \end{cases} & \alpha > 0 \\ \textbf{Type III} : \begin{cases} 0 & x < 0 \\ 1 - \exp(-x^{\alpha}) & x \geqslant 0 \end{cases} & \alpha > 0 \end{cases}.$$

For dynamical systems, the formalism of EVT provides strong tools for studying the local behavior of orbits around some point of interest in the phase space for chaotic dynamical systems exhibiting sufficiently strong ergodic properties. In some ways, this local study builds upon the notion that chaotic systems give rise to "i.i.d."-like processes.

In the rest of this thesis, we will focus entirely on establishing EVT for time series which arise from observations on dynamical systems. We will obtain distributional convergence to the classical extreme value distributions for sequences of partial maxima for the time series we study (along the lines of [47]). In addition, we also obtain almost-sure pointwise results for the behavior of these partial maxima for some systems, as a corollary to Borel-Cantelli lemmas for those systems.

1.5 Extreme Value Theory for Dynamical Systems

In the light of Theorem 1.9, one can hope to establish extreme value theory for time series which arise from deterministic dynamical systems. This idea was first exploited by Collet [21] where he established extreme value theory for some time series arising from 1-D non-uniformly expanding maps T modeled by a Young tower (see Section 1.7) with exponentially decreasing tails for the return time function for the tower. Collet was interested in time series of the form

$$\xi_n(x) = -\log d(T^n x, x_0)$$

for some base point x_0 where d is a metric on the phase space of the dynamical system. The time series ξ_n measures how close the nth iterate of a point x comes to the base point x_0; the partial maxima of ξ_n given by

$$M_n := \max\{\xi_0, \xi_1, \ldots, \xi_n\}$$

measures how close orbits of points come to the base point x_0 in n steps. Collet established that for the systems and the time series he considered, the limiting distribution is a Type-I extreme value distribution. More precisely, he established that for a full measure set of density points x_0 and for the scaling sequence $u_n(v) = v + \log n$,

$$\lim_{n\to\infty} \mu(M_n \leqslant u_n) \to e^{-h(x_0)e^{-v}}$$

where μ is the invariant, absolutely continuous ergodic measure for T and $h = d\mu/dm$ where m is the 1-D Lebesgue measure.

Freitas and Freitas [26] showed the corresponding result for these maps when x_0 is taken to be the critical point c or critical value $f(c)$. Freitas, Freitas, and Todd [28] investigated the link between extreme value statistics and return time statistics, and showed that any multimodal map with an absolutely continuous invariant measure displays either Type I, II or III extreme value statistics. This result required no knowledge of the decay of correlations for these maps. They also proved that for these systems the excedance point process converges to a Poisson process. Dolgopyat [24, Theorem 8] has proved Poisson limit laws for the return time statistics of visits to a scaled neighborhood of a measure-theoretically generic point in uniformly partially hyperbolic systems with exponential decay of correlations for C^k functions. He also gives distributional limits for periodic orbits, but again exponential decay is required and uniform partial hyperbolicity is assumed.

One could also study extreme value theory for continuous-time dynamical systems, and we do so for some suspension flows built over measure-preserving ergodic base transformations. Suppose $(X, \mathcal{B}, T, \mu)$ is an ergodic measure-preserving system. Suppose $h : X \to \mathbb{R}^+$ is $L^1(\mu)$. We define the suspension space X^h as

$$X^h := \{(x, u) | x \in X, 0 \leqslant u \leqslant h(x)\}.$$

Define a flow $\phi_t : X^h \to X^h$ as $\phi_t(x, u) = (x, u + t)/\sim$ where $(x, h(x)) \sim (Tx, 0)$. Define $\xi_t(x, u) := -\log d(\phi_t(x, u), (x_0, u_0))$ and $\mathcal{M}_t := \sup_{0 \leqslant s \leqslant t} \xi_s$. Extreme value theory, once proved for an ergodic measure-preserving transformation T also lifts to a suspension flow over T as proved in [41]. Their theorem allows us to establish extreme value theory for the billiards flow B_t (see Chapter 3) by establishing it for the collision map T. To do this, we use the following result (for the full-strength statement, see [41, Theorem 2.6]).

Theorem 1.10. *Suppose there exist normalizing constants $a_n > 0$ and $b_n \in \mathbb{R}$ which satisfy*

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} a_n |b_{[n+\epsilon n]} - b_n| = 0 \tag{1.5.10}$$

and

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \left| 1 - \frac{a_{[n+\epsilon n]}}{a_n} \right| = 0. \tag{1.5.11}$$

Then,

$$a_N(M_N - b_N) \to F \text{ as } N \to \infty \implies a_{\lfloor T/\int h \rfloor} \left(\mathcal{M}_T - b_{\lfloor T/\int h \rfloor} \right) \to F \text{ as } T \to \infty.$$

1.5.1 Hitting time statistics and return time statistics

For a map $T : X \to X$ with invariant ergodic probability measure μ, we may define hitting and return time statistics as follows. For a set $A \subset X$, let $R_A(x)$ denote the first time $j \geqslant 1$ such that $T^j(x) \in A$. Given a sequence of sets $\{U_n\}_{n \in \mathbb{N}}$, with $\mu(U_n) \to 0$ then we say that the system has hitting time statistics (HTS) with distribution $G(t)$ for $\{U_n\}$ if for all $t \geqslant 0$

$$\lim_{n \to \infty} \mu \left(R_{U_n} \geqslant \frac{t}{\mu(U_n)} \right) = G(t). \tag{1.5.12}$$

In applications often the sequence U_n is a nested sequence of balls $B(x_0, \delta_n)$ of radius δ_n about a point x_0.

We say that the system has HTS $G(t)$ to balls at x_0 if for any sequence $\delta_n \subset \mathbb{R}^+$, with $\delta_n \to 0$ as $n \to \infty$ we have HTS $G(t)$ for $U_n = B(x_0, \delta_n)$.

Analogously we say that return time statistics (RTS) with distribution $G(t)$ holds for $\{U_n\}$ if we can replace the measure μ by the conditional measure μ_A in equation (1.5.12), where $\mu_A = \frac{\mu|A}{\mu(A)}$. RTS to balls is defined analogously to HTS to balls.

In [28] an equivalence between extreme value laws and hitting time statistics was obtained for dynamical systems (X, T, μ) admitting an absolutely continuous invariant probability measure μ. Our results in Theorem 3.5, Theorem 3.9 and Theorem 2.1 thus also establish HTS with an exponential law i.e. $G(t) = e^{-t}$, for the systems considered.

1.5.2 Extreme value theory for skew-extensions

In Chapter 2, using arguments based on Collet's and results on the rate of decay of correlations for compact group extensions of non-uniformly expanding maps by Gouëzel [35], we establish extreme value theory (or return time statistics) for non-uniformly partially hyperbolic systems. Our main result is Theorem 2.1 which gives verifiable conditions on the base transformation and a sufficient rate of decay of correlations for a Type I extreme value distribution to hold for $\Phi(p) = -\log d(p, p_0)$ for $\mu \times \lambda_Y$ a.e. $p_0 = (x_0, \theta_0) \in X \times Y$ where μ is the ergodic measure for the base map and λ_Y is the Lebesgue measure. We note that we require only a polynomial rate of decay for our results to hold. This characterizes the extreme value statistics for observations of a certain degree of regularity with maxima at such points p_0. The sufficient conditions of Theorem 2.1 are verifiable for a residual set of Hölder S^1-cocyles over certain classes of maps recorded in Corollary 2.5. The maps in this category include piecewise C^2 uniformly expanding maps and non-uniformly expanding maps with finite derivative which may be modeled by a Young tower (see 1.7) with exponential return time tails (such as logistic or unimodal maps, including the class studied by Collet). We also verify, in section 2.4.2, that Gouëzel's map satisfies the hypotheses of our theorem and hence our results also apply to this map. A key role in our verification is played by results due to Gouëzel [35] on rates of decay of correlations for S^1 extensions of non-uniformly partially hyperbolic systems. We note that our Type I law for $\Phi(p) = -\log d(p, p_0)$ also implies Type II and Type III laws for $\Phi(p) = d(p, p_0)^{-\alpha}$ and $\Phi(p) = \mathcal{C} - d(p, p_0)^{\alpha}$ (see [41, Lemma 1.3]).

Further, we verify the conditions on the base transformation for a class of intermittent-like maps, including the Liverani-Saussol-Vaienti map. Unfortunately, the rate of decay of correlations of Hölder observations on compact group extensions of such systems is not known. Nevertheless we give a sufficient decay rate to ensure Type I extreme value statistics for $-\log d(p, p_0)$ for $\mu \times \lambda_Y$ a.e. p_0. We believe it plausible that for sufficiently small $0 < \omega < 1$, where the germ of the indifferent fixed point is $x \to x + x^{1+\omega}$, this decay rate holds and will be proven to hold. We also verify all but one of the hypotheses of our theorem for the Viana map. The hypothesis that fails concerns the density of the absolutely continuous invariant measure. It is not known whether the density belongs to $\mathbb{L}^{1+\delta}(\lambda)$ for any $\delta > 0$.

1.5.3 Extreme value theory for dispersing billiards and Lozi-like maps

We establish condition $D_2(u_n)$, a version of condition $D(u_n)$ more suited to the study of dynamical systems, in Chapter 3, for the time-series of certain observations on maps modeled by Young towers (see Section 1.7) with exponential return time tails satisfying (A5) (see page 55).

Collet [21] demonstrated a technique involving maximal functions for establishing $D'(u_n)$ for one-dimensional non-uniformly expanding maps modeled by a Young tower (Section 1.7). His argument relies on the absence of a stable direction and the boundedness of the derivative and these are obstacles to generalizing his argument. The one-dimensional feature can be generalized to expanding maps in higher dimension [28]. One of our main contributions is that we extend Collet's approach to handle dynamical systems with stable foliations. We also establish condition $D^{'}(u_n)$ for planar dispersing billiard maps and flows and a class of Lozi-maps and show that from the point of view of extreme value theory they behave as i.i.d. processes. We do not provide many details about these maps here, beyond what is needed to be able to state our results and motivate our techniques, but we refer to Chapter 3 for more on these systems.

Our results on billiard flows are immediate consequences of the results in [41] which show in essence that suspension flows inherit the extreme value behavior of their base transformations.

As in Collet [21] we will consider the observation $\phi(x) = -\log d(x, x_0)$ on the metric space (X, d). Knowledge of the extreme value statistics for this observation determines the extreme value statistics of a wide class of observation maximized at the same point x_0 [41].

Let $B(x, r)$ denote the ball of radius $r > 0$ about $x \in X$. For a measure-preserving transformation $T : (X, \mu) \to (X, \mu)$ define, if it exists,

$$d(x) = \lim_{r \to 0} \frac{\log \mu(B(x, r))}{\log(r)}$$

It is known from [48] that if μ is an SRB-measure for a $C^{1+\alpha}$ diffeomorphism then the limit $d(x)$ exists and has the same value for μ almost every x [10].

We assume the existence of a scaling sequence $u_n(v)$ such that $n\mu(B(x_0, e^{-u_n(v)})) \to e^{-v}$ for μ a.e. x_0 and prove our results with respect to this sequence $u_n(v)$. Planar dispersing billiards possess an absolutely continuous invariant measure, with density $\rho(x) := \frac{d\mu}{d\lambda}(x)$, where λ is the Lebesgue measure. By the Lebesgue differentiation theorem for any $a > 0$, $n\mu B(x_0, \frac{\sqrt{a}}{\sqrt{n}}) \to \rho(x_0)a$ for μ a.e. x_0.

Henceforth, we suppress the dependence of the sequence u_n on v unless it is explicitly required.

Lozi maps have an SRB measure μ with absolutely continuous conditional measures on local unstable manifolds. In fact the conditional measure on a local unstable manifold is one-dimensional Lebesgue measure. Hence the μ measure of an annulus about a generic point x_0 of inner radius r and width ϵ is bounded by $C\epsilon^\delta$ (see Proposition 3.8) for some $\delta > 0$ and hence for μ a.e. x_0, the function $r \mapsto \mu(B(x_0, r))$ is continuous and strictly increasing. Thus we may choose a sequence u_n satisfying $n\mu(B(x_0, e^{-u_n})) \to e^{-v}$ for μ a.e. x_0 in this setting also. An immediate corollary of the existence of the dimension d is that $du_n \approx v + \log n$ (in a sense made precise in Lemma 3.4).

The relation $du_n \approx v + \log n$ does not imply that $\lim_{n\to\infty} \mu(M_n \leqslant u_n(v)) = \lim_{n\to\infty} \mu(M_n \leqslant (\log n + v)/d)$ but rather that for all $\epsilon > 0$, $\lim_{n\to\infty} \mu(M_n \leqslant (1 - \epsilon)(\log n + v)/d)) \leqslant G(v) \leqslant \lim_{n\to\infty} \mu(M_n \leqslant (1+\epsilon)(\log n + v)/d))$, where $G(v) = \lim_{n\to\infty} \mu(M_n \leqslant u_n(v))$. In the case of the Lozi map this is the best we can do. For Sinai dispersing billiards, as the invariant measure is absolutely continuous, we are able to obtain the scaling constants u_n explicitly. We prove that for μ a.e. x_0, if $\phi(x) = -\log d(x, x_0)$ then $\lim_{n\to\infty} \mu(M_n \leqslant (\log n + v + \log(\rho(x_0)))/2) = e^{-v}$ where $\rho(x_0) := \frac{d\mu}{dm}(x_0)$.

We now state the versions of $D(u_n)$ and $D'(u_n)$ which we use. If $\{X_n\}$ is a stochastic process define

$$M_{j,l} := \max\{X_j, X_{j+1}, \ldots, X_{j+l}\}.$$

We will often write $M_{0,n}$ as M_n, as this causes no confusion.

Condition $D_2(u_n)$ **[27]** We say condition $D_2(u_n)$ holds for the sequence $X_0, X_1, \ldots$, if for any integers l,t and n

$$|\mu(X_0 > u_n, M_{t,l} \leqslant u_n) - \mu(X_0 > u_n)\mu(M_l \leqslant u_n)| \leqslant \gamma(n,t)$$

where $\gamma(n,t)$ is non-increasing in t for each n and $n\gamma(n,t_n) \to 0$ as $n \to \infty$ for some sequence $t_n = o(n)$, $t_n \to \infty$.

Condition $D^{'}(u_n)$ **[47]** We say condition $D^{'}(u_n)$ holds for the sequence $X_0, X_1, \ldots$, if

$$\lim_{n\to\infty} n \sum_{j=1}^{o(n)} \mu(X_0 > u_n, X_j > u_n) = 0,$$

where $o(n) \to \infty, o(n)/n \to 0$.

We consider a class of maps of Riemannian manifolds, perhaps with singularities, modeled by a Young tower [75] (see also 1.7) with SRB measure μ and exponential return time tails. Lozi-like maps and Sinai dispersing billiards fit into this scheme. We establish $D_2(u_n)$ for the process $X_n(x) = -\log(d(x_0, T^n x))$. The proof of $D_2(u_n)$ requires only sufficiently high polynomial decay of correlations but as our applications all have exponential decay of correlations to simplify exposition we assume exponential tails. Furthermore if $D'(u_n)$ can be verified for these systems, then the process has the same extreme value statistics as its associated i.i.d. process, even for more general observations [41]. We verify $D'(u_n)$ for the systems we mentioned but we do not have a general method to establish $D'(u_n)$ for all systems modeled by a Young tower. Our method of proof for $D'(u_n)$ in these cases is an extension of the argument in Collet [21][††].

1.6 Dynamical Borel-Cantelli Lemmas

In Chapter 4 we establish dynamical Borel-Cantelli lemmas for one-dimensional ($1D$) non-uniformly expanding maps and give some applications of these results to the extreme value theory of dynamical systems.

[††]Poisson-limit laws for return-time statistics in the Axiom-A setting have been established by Hirata [38] and in the uniformly partially hyperbolic setting by Dolgopyat [24]. For recent related work on extreme value theory for deterministic dynamical systems see [26, 27, 28, 41].

Note that the finite direction of the Borel-Cantelli lemma (1) on page 9 does not require independence. In the dynamical setting suppose $T : X \to X$ is a measure-preserving transformation of the probability space $(X, \mathcal{B}, \mu)$ and (A_n) is a sequence of sets such that $\sum_n \mu(A_n) = \infty$. We are interested in the following question: does $T^n(x) \in A_n$ occur for infinitely many values of n for μ a.e. $x \in X$? and, if so, is there a quantitative estimate of the asymptotic number of entry times? For example, the sequence (A_n) may be a nested sequence of intervals, a setting which is often called the shrinking target problem. The assumption of independence of the events $T^{-n}A_n$ is seldom valid for deterministic dynamical systems; thus establishing Borel-Cantelli lemmas is a more difficult task. In Chapter 4 we establish results analogous to (1) and (2) on page 9 for certain classes of nested intervals in the setting of one-dimensional non-uniformly expanding dynamical systems: Theorems 4.9 and 4.12. To do this, we establish a more general Borel-Cantelli lemma for sequences of intervals in Gibbs-Markov systems: Theorem 4.7.

There have been some results on Borel-Cantelli lemmas for uniformly hyperbolic systems. Chernov and Kleinbock [20] establish the sBC property for certain families of cylinders in the setting of topological Markov chains and for certain classes of dynamically-defined rectangles in the setting of Anosov diffeomorphisms preserving Gibbs measures. Dolgopyat [24] has related results for sequences of balls in uniformly partially hyperbolic systems preserving a measure equivalent to Lebesgue.

More recently, Kim [45] has established the sBC property for sequences of intervals in the setting of one-dimensional piecewise-expanding maps f with $1/|f'|$ of bounded variation.

Kim uses this result to prove some sBC results for non-uniformly expanding maps with an indifferent fixed point. More precisely, he considers maps of the form

$$T_\alpha(x) = \begin{cases} x(1+2^\alpha x^\alpha) & \text{if } 0 \leqslant x < \frac{1}{2}; \\ 2x-1 & \text{if } \frac{1}{2} \leqslant x \leqslant 1. \end{cases} \tag{1.6.13}$$

If $0 < \alpha < 1$ then T_α admits an invariant probability measure μ that is absolutely continuous with respect to Lebesgue measure λ. Kim shows that if (I_n) is a sequence of intervals in $(d, 1]$ for some $d > 0$ and $\sum_n \mu(I_n) = \infty$ then I_n is an sBC sequence if (a) $I_{n+1} \subset I_n$ for all n (nested intervals) or (b) $\alpha < (3 - \sqrt{2})/2$. Kim shows that the condition $I_n \subset (d, 1]$ for some $d > 0$ is in some sense optimal (with respect to the invariant measure μ) by showing that setting $A_n = [0, n^{-1/(1-\alpha)})$ gives a sequence such that $\sum_n \mu(A_n) = \infty$ yet the sBC property does not hold; in fact, $T_\alpha^n(x) \in A_n$ for only finitely many values of n for μ a.e. $x \in [0, 1]$.

For the same class of maps T_α, Gouëzel [34] considers Lebesgue measure λ (rather than μ) and shows that if (I_n) is a sequence of intervals such that $\sum_n \lambda(I_n) = \infty$ then (I_n) is a BC sequence. Gouëzel uses renewal theory and obtains BC results but not sBC results.

In the setting of continuous-time systems, Maucourant [54] considers geodesic flows on hyperbolic manifolds of finite volume. He proves a BC result for nested balls in this context.

The above references comprise the near complete body of literature which address the Borel-Cantelli lemmas in dynamical systems; in the dynamical systems literature, therefore, Borel-Cantelli lemmas are scant.

In Chapter 4, we prove sBC results for intervals satisfying a bounded ratio condition for $1D$ Gibbs-Markov maps. We use this result to establish the dBC property for sequences of nested intervals in the setting of non-uniformly expanding $1D$ systems modeled by Young towers (Section 1.7). More precisely, our dBC results are formulated for sequences of nested intervals $I(n)$ with center x_n and length $g(n)$. Our assumption that the intervals are nested implies that $g(n)$ is a decreasing sequence. In specific situations, one often sets $g(n) = n^{-\beta}$ for some $0 \leqslant \beta \leqslant 1$. Many non-uniformly expanding $1D$ maps can be modeled by Young towers. If $(M, \mathcal{B}, \mu, T)$ is a $C^{1+\varepsilon}$ dynamical system on a compact interval M such that μ is ergodic, $\mu \ll \lambda$, and μ has a positive Lyapunov exponent, then the system can be modeled by a Young tower (personal communication by José Alves and Henk Bruin; see also [4, 5, 15]). The results of Chapter 4 therefore apply to such maps.

1.7 Young Towers

Young towers are powerful tools for coding the dynamics of a system in such a way so as to enable the computation of its statistical properties. Here we will summarize the setting and results of [75] and will focus more on the axiomatic construction of the tower than on the dynamical consequences of the construction. Let $T : X \to X$ be a $C^{1+\epsilon}$ diffeomorphism. We will denote the Lebesgue measure by λ and the restriction of λ to a submanifold S of X by λ_S. A submanifold $\gamma \subset X$ will be called an unstable disk if $d(T^{-n}x, T^{-n}y) \to 0$ exponentially fast as $n \to \infty$ for $x, y \in \gamma$. Unstable disks will be denoted by γ^u. γ will be called a stable disk if $d(T^n x, T^n y) \to 0$ as $n \to \infty$

for $x, y \in \gamma$; stable disks will be denoted by γ^s. We will say a subset $\Lambda \subset X$ has a hyperbolic product structure if there exist families of stable disks $\{\gamma^s\}$ and unstable disks $\{\gamma^u\}$ satisfying:

1. $\dim \gamma^s + \dim \gamma^u = \dim X$,
2. γ^u disks are transversal to γ^s disks with angles between them bounded away from 0,
3. each γ^u disk meets each γ^s disk in exactly one point,
4. $\Lambda = (\cup \gamma^u) \cap (\cup \gamma^s)$.

Let Γ_s and Γ_u denote the families $\{\gamma^s\}$ and $\{\gamma^u\}$. Let $\Gamma_{*,s} \subset \Gamma_s$. The families $\Gamma_{*,s}$ and Γ_u define a subset Λ_* of Λ; such a subset is called an *s-subset*. Similarly, if $\Gamma_{*,u} \subset \Gamma_u$ is a subset of the unstable family, then the subset $\Lambda_* \subset \Lambda$ generated by $\Gamma_{*,u}$ and Γ_s is called a *u-subset*.

Assume that

(P1) There exists $\Lambda \subset X$ with a hyperbolic product structure and with $\lambda_{\gamma^u}(\gamma^u \cap \Lambda) > 0$ for every γ^u an unstable leaf.

(P2) There are pairwise disjoint s-subsets $\Lambda_1, \Lambda_2 \cdots \subset \Lambda$ with the properties that

- on each γ^u disk, $\lambda_{\gamma^u}((\Lambda \setminus \cup \Lambda_i) \cap \gamma^u) = 0$
- for each i, there exists an $R_i \in \mathbb{N}$ such that $T^{R_i}\Lambda_i$ is a u-subset of Λ and $T^{R_i}(\gamma^s(x)) \subset \gamma^s(T^{R_i}x)$, $T^{R_i}(\gamma^u(x)) \supset \gamma^u(T^{R_i}(x))$

- for each n there are at most finitely many i with $R_i = n$
- $\min R_i \geqslant R_0 > 1$ depending only on T.

(P3) There exists a $C > 0$ and $0 < \alpha < 1$ such that for $y \in \gamma^s(x), d(T^n x, T^n y) \leqslant C\alpha^n$ for all $n \geqslant 0$.

$T^{R_i}(\Lambda_i)$ corresponds to a "return" of an s-subset as a u-subset. We can define a notion of how long points stay together under these returns by defining a separation function s_0. We set $s_0(x,y) = n$ if the orbits of x and y stay in the same s-subsets Λ_i up to time n, and if $T^{n+1}x$ and $T^{n+1}y$ belong to different Λ_i. If two points start in the same Λ_i, then we define $s_0(T^k x, T^k y) = s_0(x,y) - k$ if $k < R_i$. Hence, points do not separate before they return. Defining $T^u := T|_{\gamma^u}$, we assume

(P4) For $y \in \gamma^u(x)$ and $0 \leqslant k \leqslant n < s_0(x,y)$, we have

(a) $d(T^n(x), T^n(y)) \leqslant C\alpha^{s_0(x,y)-n}$

(b)

$$\log \prod_{i=k}^{n} \frac{detDT^u(T^i x)}{detDT^u(T^i y)} \leqslant C\alpha^{s_0(x,y)-n}$$

(P5) **(a)** For $y \in \gamma^s(x)$,

$$\log \prod_{i=n}^{\infty} \frac{detDT^u(T^i(x))}{detDT^u(T^i(y))} \leqslant C\alpha^n \quad \forall n \geqslant 0$$

(b) For γ, γ' unstable curves, if $\Theta : \gamma \cap \Lambda \to \gamma' \cap \Lambda$ is the holonomy given by $\Theta(x) = \gamma^s(x) \cap \gamma'$, then Θ is absolutely continuous and

$$\frac{d(\Theta_*^{-1}\lambda_{\gamma'})}{d\lambda_\gamma}(x) = \prod_{i=0}^{\infty} \frac{detDT^u(T^i(x))}{detDT^u(T^i(\Theta x))}.$$

Define the Young tower by

$$\Delta = \cup_{i,l \leqslant R_i - 1}\{(x,l) : x \in \Lambda_{0,i}\}$$

where $\Lambda_{0,i} := \Lambda_i$ and the tower map $F : \Delta \to \Delta$ by

$$F(x,l) = \begin{cases} (x, l+1) & \text{if } x \in \Lambda_{0,i}, l < R_i - 1 \\ (T^{R_i}x, 0) & \text{if } x \in \Lambda_{0,i}, l = R_i - 1 \end{cases}.$$

For convenience, we will refer to $\Delta_0 := \cup_i(\Lambda_{0,i}, 0)$ as the base of the tower Δ. We define $\Delta_l = \{(x,l) : l < R(x)\}$, the lth level of the tower. Define the map $f = T^R : \Delta_0 \to \Delta_0$ i.e. $f(x) = T^{R(x)}(x)$. We may form a quotiented tower (see [75] for details) by introducing an equivalence relation for points on the same stable manifold.

There exists an invariant measure λ_0 for $f : \Delta_0 \to \Delta_0$ which has absolutely continuous conditional measures on local unstable manifolds in Δ_0, with density bounded uniformly from above and below.

The tower structure allows us to construct an invariant measure ν for F on Δ by defining for a measurable set $B \subset \Lambda_l$, $\nu(B) = \frac{\lambda_0(F^{-l}B)}{\int_{\Delta_0} R d\lambda_0}$ and extending the definition to disjoint unions of such sets in the obvious way. We define a projection $\pi : \Delta \to X$ by $\pi(x,l) = T^l(x)$. We note that $\pi \circ F = T \circ \pi$. The invariant measure μ, which is an SRB measure for $T : X \to X$, is given by $\mu = \pi_*\nu$. $W^s_{loc}(x)$ will denote the local stable manifold through x i.e there exists $\epsilon(x) > 0$ and $C > 0$, $0 < \alpha < 1$ such that $W^s_{loc} = \{y : d(x,y) < \epsilon \text{ and } d(T^n y, T^n x) < C\alpha^n \text{ for all } n \geqslant 0\}$. We use the notation W^s_{loc} rather than $W^s_\epsilon(x)$ in contexts where the length of the local stable manifold is not important. We analogously define $W^u_{loc}(x)$ and let $B(x,r)$ denote the ball of radius r centered at the point x. We lift a function $\phi : X \to \mathbb{R}$ to Δ by defining $\phi(x,l) = \phi(T^l x)$.

In Chapters 3 and 4 we will use the structure of the Young towers in an essential way to prove our results. In chapter 3, we will prove condition $D_2(u_n)$ (a version of condition $D(u_n)$, see Section 1.5.3) by using statistical properties proved in [75] for the quotiented tower. Our calculations, however, are done on the full tower. We show that the error in considering the quotiented tower, instead of the full tower, becomes arbitrarily small for T^n when n is large. In Chapter 4, we prove a result about Gibbs-Markov maps, and we then extend it to a larger class of systems modeled by a Young tower by using the fact that the base map f is a Gibbs-Markov map.

Chapter 2

EVT for Skew Extensions

2.1 Framework of the Problem

We will assume that Y is a compact, connected M-dimensional manifold equipped with metric d_Y and X is a compact N-dimensional manifold with metric d_X. Let $D = M + N$. Define the product metric on $X \times Y$ by

$$d((x_1, \theta_1), (x_2, \theta_2)) = \sqrt{d_X(x_1, x_2)^2 + d_Y(\theta_1, \theta_2)^2}. \tag{2.1.1}$$

Let λ_X denote the Lebesgue measure on X and λ_Y the Lebesgue measure on Y. The product Lebesgue measure on $X \times Y$ will then be $\lambda_X \times \lambda_Y$.

We will call a function $\Phi : X \times Y \to \mathbb{R}$ Hölder continuous of exponent ζ if there exists some constant K such that

$$|\Phi(x) - \Phi(y)| \leq K d((x_1, \theta_1), (x_2, \theta_2))^\zeta$$

for all (x_1, θ_1) and (x_2, θ_2) in $X \times Y$. We define the C^ζ norm of Φ as

$$\|\Phi\|_{C^\zeta} = \sup_{(x,\theta)\in X\times Y} |\Phi(x,\theta)| + \sup_{\substack{(x,\theta),(y,\rho)\in X\times Y \\ (x,\theta)\neq(y,\rho)}} \frac{|\Phi(x,\theta) - \Phi(y,\rho)|}{d((x,\theta),(y,\rho))^\zeta}.$$

If $f : X \to X$ is a measurable transformation and $u : X \times Y \to Y$ is a measurable function, then we may define T, the Y**-skew-extension of** f **by** u by,

$$T : X \times Y \to X \times Y$$

$$T(x,\theta) = (f(x), u(x,\theta)). \tag{2.1.2}$$

We make the following assumptions about f.

(A1) f has an ergodic invariant measure μ_X with $supp(\mu_X) = X$.

(A2) $T : X \times Y \to X \times Y$ preserves an invariant probability measure ν, with $\nu \ll \lambda_X \times \lambda_Y$, $H := d\nu/d(\lambda_X \times \lambda_Y) \in \mathbb{L}^{1+\delta}(\mu_X \times \lambda_Y)$ and $H \in \mathbb{L}^{1+\delta}(\lambda_X \times \lambda_Y)$ (locally) for some $\delta > 0$.

From this point on, in Chapter 2, we will let $\lambda := \lambda_X \times \lambda_Y$.

We will obtain the extreme value statistics of observations which are maximized at a unique point (x_0, θ_0). For the given point (x_0, θ_0) we define a function $\Phi_{(x_0,\theta_0)}$ on $X \times Y$ by

$$\Phi_{(x_0,\theta_0)}(x,\theta) = -\log d((x,\theta),(x_0,\theta_0))$$

(from here on, the dependence on (x_0, θ_0) is omitted for notational simplicity). For a given $v \in \mathbb{R}$ we define $u_n = v + \frac{1}{D}\log n$ and denote by M_n (more precisely $M_n^{(x_0,\theta_0)}$)

the random variable

$$M_n^{(x_0,\theta_0)} = \max(\Phi, \Phi\circ T, \dots, \Phi\circ T^n).$$

We will prove the following result:

Theorem 2.1. *Assume (A1) and (A2). Let δ be as in (A2) and let $\kappa > 1$ be conjugate to $1+\delta$, i.e.,*

$$\frac{1}{1+\delta} + \frac{1}{\kappa} = 1.$$

Further, assume

(A3) *that there exist constants $C_1 > 0, \beta > 0$ and an increasing function $g(n) \approx n^{D\gamma'}$ (with $0 < \gamma' < \frac{\beta}{D}$) such that if*

$$E_n^X := \left\{ x \in X \colon d_X(f^j x, x) < \frac{1}{n} \text{ for some } j \in \{1, 2 \dots g(n)\} \right\}$$

then $\mu_X(E_n^X) < \frac{C_1}{n^\beta}$

(A4) *that there exists $0 < \hat{\alpha} \leq 1$ such that, for all Hölder continuous functions Φ with Hölder exponent $\hat{\alpha}$, and $\Psi \in \mathbb{L}^\infty(\nu)$,*

$$\left| \int \Psi\circ T^j \Phi \, d\nu - \int \Phi \, d\nu \int \Psi \, d\nu \right| \leq C_2 \Theta(j) \|\Psi\|_\infty \|\Phi\|_{C^{\hat{\alpha}}} \tag{2.1.3}$$

where $\Theta(j) \leq j^{-\alpha}$ and $\alpha > \frac{\frac{1}{D}\left(1 + D\kappa\left(\frac{3}{2} - \frac{1}{\kappa}\right)\right) + \frac{3}{2}}{\min\left\{\gamma', \frac{1}{2}\right\}}$.

Then for ν a.e. (x_0, θ_0) and for every $v \in \mathbb{R}$,

$$\lim_{n\to\infty} \nu\left(M_n^{(x_0,\theta_0)} < u_n\right) = e^{-H(x_0,\theta_0)e^{-Dv}} \tag{2.1.4}$$

We will prove theorem 2.1 for an arbitrary fiber Y that is a compact connected M-dimensional manifold. Our corollaries, however, will involve the special cases $Y = S^1$ and $Y = [0,1]$. This is because (A4) of Theorem 2.1 requires a decay of correlations to hold and we only consider examples for which this decay is known to hold. Note also that we require $0 < \hat{\alpha} \leq 1$. This is because for the proof of Lemma 2.12, we need (A4) of the above theorem to hold for Lipschitz continuous functions having compact support.

We now make three definitions.

Definition 2.2 (Residual Set)**.** *A set will be called* **residual** *if its complement can be written as a countable union of nowhere dense sets.*

Definition 2.3 (Cocycle)**.** *A C^r* **cocycle** *h on an interval I into a group Y will be defined as a C^r map $h : I \to Y$.*

Definition 2.4 (Skew-Extension)**.** *If h is a cocycle, the skew-extension T will be defined as $T(x, \theta) = (f(x), \theta.h(x))$.*

We now state the corollaries to the above theorem (see Section 2.4).

Corollary 2.5. *If $Y = S^1$ and f is one of the following transformations:*

(a) a piecewise C^2 uniformly expanding map $f : I \to I$ of an interval I.

(b) a one-dimensional non-uniformly expanding map $f : I \to I$ of an interval I with bounded derivative and modeled by a Young tower with exponential decay of correlations

then for a residual set of Hölder cocycles $h : I \to S^1$, for $\mu_X \times \lambda_Y$ a.e. (x_0, θ_0) and for all $v \in \mathbb{R}$,

$$\lim_{n\to\infty} \nu\left(M_n^{(x_0,\theta_0)} < u_n\right) = e^{-H(x_0,\theta_0)e^{-Dv}}. \tag{2.1.5}$$

In the next corollary, we take the skew extension over the map $f(\omega) = 4\omega \mod 1$ with the map T_α defined in Section 2.4.2.

Corollary 2.6. *Let $T : S^1 \times [0,1] \to S^1 \times [0,1]$ be the map $T(\omega, x) = (4\omega, T_{\alpha(\omega)}(x))$ where the maps α and T_α, an intermittent type map, are as defined in Section 2.4.2. Suppose that*

$$\sup_{\omega\in[0,1]} \alpha(\omega) := \alpha_{\max} < \frac{\min\{\gamma', \frac{1}{2}\}}{\min\{\gamma', \frac{1}{2}\} + \frac{1}{D}\left(1 + \frac{D}{2}\right) + \frac{3}{2}}.$$

Then for ν a.e. (ω_0, x_0) and for each $v \in \mathbb{R}$,

$$\lim_{n\to\infty} \nu\left(M_n^{(\omega_0,x_0)} < u_n\right) = e^{-H(\omega_0,x_0)e^{-Dv}}. \tag{2.1.6}$$

There are other important classes of maps such as Y extensions of Manneville-Pommeau-type maps (for a compact connected Lie group Y, for instance, $Y = S^1$) and the Viana-type maps that satisfy most, but not all, of our hypotheses. It is not known for S^1 extensions of Manneville-Pommeau-type maps whether a sufficiently high polynomial rate of decay satisfying our hypotheses holds. Similarly, for the Viana map, all of our hypotheses are satisfied except we do not know whether the density of the invariant measure is locally $\mathbb{L}^{1+\delta}$ for some $\delta > 0$. A further discussion of these maps may be found in Section 2.4.

2.2 Preliminaries

For the rest of this chapter, we will refer to the function T^0 as the identity function and χ_A as the characteristic function for A. Upper-case greek letters, such as Φ and Ψ, will usually denote functions, while lower-case letters, such as ϕ, will usually denote scalar constants. We note that this convention is peculiar to this chapter, and we hope it causes no confusion in subsequent chapters.

This section contains the statements of some lemmas from [21] and proofs of some other lemmas which we will need for the proof of Theorem 2.1. The highlight of this section is Proposition 2.10 because it allows us to induce to the product system an important and desirable property of the base map T (see (A3)).

Lemma 2.7. *For any $k > 0$ and any $u \in \mathbb{R}$*

$$\sum_{j=1}^{k} \chi_{\{\Phi\circ T^j \geq u\}} \geq \chi_{\{M_k \geq u\}} \geq \sum_{j=1}^{k} \chi_{\{\Phi\circ T^j \geq u\}} - \sum_{l\neq j} \chi_{\{\Phi\circ T^j \geq u\}}\chi_{\{\Phi\circ T^l \geq u\}} \tag{2.2.7}$$

Lemma 2.8. *For any integers r and $k \geq 0$,*

$$0 \leq \nu(M_r < u) - \nu(M_{r+k} < u) \leq k\nu(\Phi \circ T^0 \geq u) \tag{2.2.8}$$

Lemma 2.9. *For any positive integers m, p and t,*

$$\begin{aligned} &\left|\nu(M_{m+p+t} < u) - \nu(M_m < u) + \sum_{j=1}^{p} \int \chi_{\{\Phi\circ T^0 \geq u\}}\chi_{\{M_m < u\}} \circ T^{p+t-j}\, d\nu\right| \\ &\qquad \leq 2p\sum_{j=1}^{p} \int \chi_{\{\Phi\circ T^0 \geq u\}}\chi_{\{\Phi\circ T^0 \geq u\}} \circ T^j\, d\nu + t\nu(\Phi \circ T^0 \geq u) \end{aligned} \tag{2.2.9}$$

The proofs for these lemmas can be found in [21].

Proposition 2.10. *Let μ_X be the invariant, ergodic measure with respect to the map $f : X \to X$. Suppose*

$$E_n^X := \left\{ x \in X : d(f^j x, x) < \frac{1}{n} \text{ for some } j \leq g(n) \right\}$$

satisfies $\mu_X(E_n^X) \leq \frac{C}{n^\beta}$ for some constant $C > 0$ and some $\beta > 0$. Then, under the hypotheses of Theorem 2.1, $\nu(\tilde{E}_n) \leq \frac{C}{n^\beta}$ where

$$\tilde{E}_n = \left\{ (x, \theta) \in X \times Y : d(T^j(x,\theta), (x,\theta)) < \frac{1}{n} \text{ for some } j \leq g(n) \right\}.$$

Proof. $(x, \theta) \in \tilde{E}_n$ implies $d(T^j(x,\theta),(x,\theta)) < \frac{1}{n}$ for some $j \leq g(n)$ and so

$$\sqrt{d_X(f^j x, x)^2 + d_Y(u^j(x,\theta), \theta)^2} < \frac{1}{n}$$

for such j. This forces $d_X(f^j x, x) < \frac{1}{n}$. Thus, $x \in E_n^X$ and so $\tilde{E}_n \subset E_n^X \times Y$.

Define a new measure Δ on X as $\Delta(A) := \nu(A \times Y)$. If $\lambda_X(A) = 0$ then $\mu_X(A) = 0$ and so $\mu_X \times \lambda_Y(A \times Y) = 0$ and thus $\nu(A \times Y) = 0$. Therefore, Δ is absolutely continuous with respect to the Lebesgue measure on X. Further,

$$\begin{aligned} T^{-1}(A \times Y) &= \{(x,\theta) | (fx, u(x,\theta)) \in A \times Y\} \\ &= \left\{ x \in f^{-1}A, (x,\theta) \in u^{-1}Y \right\} \\ &= \left\{ x \in (f^{-1}A \cap X), \theta \in Y \right\} \\ &= f^{-1}(A) \times Y \end{aligned}$$

and so $\nu(T^{-1}(A \times Y)) = \nu(f^{-1}(A) \times Y)$. Therefore

$$\Delta(f^{-1}A) = \nu(f^{-1}A \times Y) = \nu(T^{-1}(A \times Y)) = \nu(A \times Y) = \Delta(A).$$

To prove that Δ is ergodic for f, if $f^{-1}A = A$ then $\mu_X(A) = 0$ or 1 from which it follows that $\mu_X \times \lambda_Y(A \times Y) = 0$ or 1. Therefore by redefining H (recall that H is the density of ν) on a $\mu_X \times \lambda_Y$ measure 0 set if necessary we have

$$\nu(A \times Y) = \int_{A\times Y} H d(\mu_X \times \lambda_Y) = 0 \text{ or } 1.$$

Therefore, $\Delta(A) = 0$ or 1.

Since the measures on X are absolutely continuous with respect to Lebesgue, and hence unique, $\Delta(A) = \mu_X(A)$ from where it follows that

$$\nu(\tilde{E}_n) \le \nu(E_n^X \times Y) = \Delta(E_n^X) = \mu(E_n^X) \le \frac{C}{n^\beta}.$$

$\square$

We now prove a version of condition $D'(u_n)$.

Lemma 2.11. *Under the assumptions of Theorem 2.1, for ν a.e $(x_0, \theta_0) \in X \times Y$*

$$n \sum_{j=1}^{n^{\gamma'}} \nu(\Phi \circ T^0 > u_n, \Phi \circ T^j > u_n) \to 0 \text{ as } n \to \infty. \tag{2.2.10}$$

Proof. We begin by recalling that $H \in \mathbb{L}^{1+\delta}(\lambda) \subset \mathbb{L}^1(\lambda)$. Let

$$E_n = \{(x,\theta) : d(T^j(x,\theta),(x,\theta)) < \frac{1}{n} \text{ for some } j \le g(n)\}$$

where $g(n)$ is as in Theorem 2.1. Let $D\gamma' < \psi < \beta$ and recall that $\gamma' < \psi/D$. Define the Hardy-Littlewood Maximal function as

$$\mathfrak{L}_n(x,\theta) := \sup_{r>0} \frac{1}{\lambda(B_r(x,\theta))} \int_{B_r(x,\theta)} H\chi_{E_n} d\lambda.$$

By the Hardy-Littlewood Maximal Principle, since $H\chi_{E_n} \in \mathbb{L}^1(\lambda)$,

$$\lambda(\mathfrak{L}_n(x,\theta) > \delta) \leq \frac{C}{\delta}\|H\chi_{E_n}\|_1 \leq \frac{C}{\delta}\nu(E_n) \leq \frac{C}{\delta n^\beta}.$$

Choose γ such that $\gamma(\beta - \psi) > 1$. Replacing δ by $\frac{1}{n^{\gamma\psi}}$ and n by n^γ we get

$$\lambda\left(\mathfrak{L}_{n^\gamma} > \frac{1}{n^{\psi\gamma}}\right) \leq \frac{C}{n^{\gamma(\beta-\psi)}}.$$

Therefore we have

$$\sum_n \lambda\left(\mathfrak{L}_{n^\gamma} > \frac{1}{n^{\gamma\psi}}\right) \leq \sum_n \frac{C}{n^{\gamma(\beta-\psi)}}$$

which is summable. Hence, by the Borel-Cantelli Lemma, for λ a.e. $(x_0,\theta_0) \in X \times Y$, we have $(x_0,\theta_0) \notin \limsup\{\mathfrak{L}_{n^\gamma} > \frac{1}{n^{\gamma\psi}}\}$ and so there exists $N(x_0,\theta_0)$ such that $n \geq N(x_0,\theta_0) \implies \mathfrak{L}_{n^\gamma} \leq \frac{1}{n^{\psi\gamma}}$, i.e.,

$$\sup_{r>0} \frac{1}{\lambda(B_r(x_0,\theta_0))} \int_{B_r(x_0,\theta_0)} H\chi_{E_{n^\gamma}}\, d\lambda \leq \frac{1}{n^{\gamma\psi}}.$$

Set $r = \frac{1}{n^\gamma}$ in the above to get

$$n^{\gamma D} \int_{B_{\frac{1}{n^\gamma}}(x_0,\theta_0)} H\chi_{E_{n^\gamma}}\, d\lambda \leq \frac{1}{n^{\psi\gamma}}.$$

Therefore we have

$$\nu\left\{\left\{d((x,\theta),(x_0,\theta_0)) < \frac{1}{n^\gamma}\right\} \cap E_{n^\gamma}\right\} \leq \frac{1}{n^{\psi\gamma+\gamma D}}. \tag{2.2.11}$$

Let $\tilde{g}(n) := g(n/2)$ and $k = \left(\frac{n^{1/D}}{2e^{-v}}\right)^{\frac{1}{\gamma}}$. We see that,

$$\begin{aligned}
&\left\{(x,\theta) : d((x,\theta),(x_0,\theta_0)) \leq \frac{e^{-v}}{n^{1/D}}, d(T^j(x,\theta),(x_0,\theta_0)) \leq \frac{e^{-v}}{n^{1/D}} \text{ for some } j \leq \tilde{g}(n^{\frac{1}{D}}/e^{-v})\right\} \\
\subset \quad & \left\{(x,\theta) : d((x,\theta),(x_0,\theta_0)) \leq \frac{e^{-v}}{n^{1/D}}, d(T^j(x,\theta),(x,\theta)) < \frac{2e^{-v}}{n^{1/D}} \text{ for some } j \leq \tilde{g}\left(\frac{n^{1/D}}{e^{-v}}\right)\right\} \\
\subset \quad & \left\{(x,\theta) : d((x,\theta),(x_0,\theta_0)) < \frac{2e^{-v}}{n^{1/D}}, d(T^j(x,\theta),(x,\theta)) < \frac{2e^{-v}}{n^{1/D}} \text{ for some } j \leq \tilde{g}\left(\frac{n^{1/D}}{e^{-v}}\right)\right\} \\
\subset \quad & \left\{(x,\theta) : d((x,\theta),(x_0,\theta_0)) < \frac{1}{k^\gamma}, d(T^j(x,\theta),(x,\theta)) < \frac{1}{k^\gamma} \text{ for some } j \leq \tilde{g}(2k^\gamma)\right\} \\
\subset \quad & \left\{(x,\theta) : d((x,\theta),(x_0,\theta_0)) < \frac{1}{k^\gamma}, d(T^j(x,\theta),(x,\theta)) < \frac{1}{k^\gamma} \text{ for some } j \leq g(k^\gamma)\right\} \qquad (2.2.12)
\end{aligned}$$

so that, by (2.2.11) and (2.2.12), for any $j \leq g\left(\frac{n^{\frac{1}{D}}}{2e^{-v}}\right)$

$$\nu\{\Phi \circ T^0 > u_n, \Phi \circ T^j > u_n\} \leq \frac{(2e^{-v})^{\psi+D}}{n^{1+\frac{\psi}{D}}}.$$

Since $g(n) \approx n^{D\gamma'}$,

$$\frac{g\left(\frac{n^{1/D}}{2e^{-v}}\right)}{n^{\psi/D}} \approx n^{\gamma'-\psi/D} \to 0,$$

and so we obtain

$$n \sum_{j=1}^{g\left(\frac{n^{1/D}}{2e^{-v}}\right)} \nu\{\Phi \circ T^0 > u_n, \Phi \circ T^j > u_n\} \to 0. \tag{2.2.13}$$

□

Lemma 2.12 (Condition $D_2(u_n)$)**.** *Let $B_r(x,\theta)$ be a ball of radius r and let $\epsilon > 0$ be arbitrary. Let κ be conjugate to $1+\delta$ (i.e, $\frac{1}{1+\delta} + \frac{1}{\kappa} = 1$) and let A be any measurable set. Then, under the assumptions of Theorem 2.1, there exist constants C_1 and C_2 so that*

$$\left|\nu(B_r \cap T^{-t}(A)) - \nu(B_r)\nu(A)\right| \leq C_1 \|H\|_{1+\delta}^{\lambda,(x,\theta)} (\nu(A)+1)\, r^{\frac{D+\epsilon}{\kappa}} + \frac{C_2}{r^{1+\epsilon}t^{\alpha}} \tag{2.2.14}$$

Proof. We construct a Hölder continuous approximation to the characteristic function for B_r. Let $r' = r - r^{1+\epsilon}$. Construct Φ_B by letting it be 1 on the inside of the ball of radius r' around (x,θ) and letting it decay to 0 at a linear rate between r and r' . The Lipschitz constant of this function may be chosen to be $\frac{1}{r^{1+\epsilon}}$.

Next, we note that $\lambda(B_r \setminus B_r') = r^D - (r - r^{1+\epsilon})^D \leq 2^D r^{D+\epsilon}$ and so we have,

$$\begin{aligned} \|\Phi_B - \chi_{B_r}\|_1^{\nu} &= \int |\Phi_B - \chi_{B_r}|\, d\nu \leq \nu(B_r \setminus B_{r'}) = \int H\chi_{B_r \setminus B_{r'}} d\lambda \\ &\leq \|H\|_{1+\delta}^{\lambda,(x,\theta)} \|\chi_{B_r \setminus B_{r'}}\|_{\kappa}^{\lambda} \leq C_1 \|H\|_{1+\delta}^{\lambda,(x,\theta)} r^{\frac{D+\epsilon}{\kappa}}. \end{aligned} \tag{2.2.15}$$

Finally,

$$
\begin{aligned}
&\left|\int \chi_B \chi_A \circ T^t d\nu - \int \chi_B d\nu \int \chi_A d\nu\right| \\
\leq \quad & \left|\int \chi_B \chi_A \circ T^t d\nu - \Phi_B \chi_A \circ T^t d\nu\right| \\
& + \left|\int \Phi_B \chi_A \circ T^t d\nu - \int \Phi_B d\nu \int \chi_A d\nu\right| \\
& + \left|\int \Phi_B d\nu \int \chi_A d\nu - \int \chi_A d\nu \int \chi_B d\nu\right| \\
\leq \quad & \|\chi_A \circ T^t\|_\infty \|\chi_B - \Phi_B\|_1^\nu + \frac{C_2 \|\chi_A\|_\infty \|\Phi_B\|_{\hat{\alpha}}}{t^\alpha} + \nu(A)\|\chi_B - \Phi_B\|_1^\nu .
\end{aligned}
\tag{2.2.16}
$$

A substitution of estimates from equation (2.2.15) completes the proof. □

2.3 Proof of Theorem

To prove Theorem 2.1, we begin by breaking n as a product of p and q with $p = \sqrt{n}$. We note that

$$\nu(M_n < u_n) \approx \nu(M_{n+qt} < u_n)$$

where t is a monotonically increasing function chosen to satisfy $\frac{t}{p} \to 0$. The main estimate in the proof is

$$\nu(M_{n+qt} < u_n) \approx (1 - p\nu(\Phi \circ T^0 \geq u_n))^q.$$

The function t needs to be chosen so that terms of the form $n \sum_{j=1}^{p} \nu(\Phi \circ T^0 \geq u_n, \Phi \circ T^j \geq u_n)$ that appear in the error to the above approximation can be broken into sums over $1 \leq j \leq t$ and $t < j \leq p$ with t being small enough for growth of terms in the first sum to be killed by Lemma 2.11 while large enough for growth in the second sum to be killed by Lemma 2.12. We call this argument the "Blocking Argument". A similar argument is used in Section 3 to establish extreme value theory for dispersing billiards and Lozi-like maps.

Theorem 2.13. *Under the hypotheses of Theorem 2.1, for ν a.e.(x,θ) and for any $v \in \mathbb{R}$,*

$$\lim_{n\to\infty} \nu\left(M_n^{(x,\theta)} < u_n\right) = e^{-H(x,\theta)e^{-Dv}}. \tag{2.3.17}$$

Proof. Choose $(x,\theta) \notin \limsup_{n\to\infty} E_n$ such that

$$\lim_{a\to 0} \frac{1}{\lambda(B_a(x,\theta))}\nu(B_a(x,\theta)) = H(x,\theta).$$

Then from above

$$\lim_{n\to\infty} n\nu(B_{\frac{e^{-v}}{n^{1/D}}}(x,\theta)) = e^{-Dv}H(x,\theta).$$

Choose

$$\epsilon > D\kappa\left(\frac{3}{2} - \frac{1}{\kappa}\right) \tag{2.3.18}$$

and $0 < \tau < \min\{\gamma', \frac{1}{2}\}$ such that

$$\alpha > \frac{\frac{1+\epsilon}{D} + \frac{3}{2}}{\tau} > \frac{\frac{1}{D}\left(1 + D\kappa\left(\frac{3}{2} - \frac{1}{\kappa}\right)\right) + \frac{3}{2}}{\min\left\{\gamma', \frac{1}{2}\right\}} \tag{2.3.19}$$

Define $t = n^\tau$, $p = \sqrt{n}$ and $q = \sqrt{n}$. Note that, by Lemma 2.8,

$$\left|\nu(M_n < u_n) - \nu(M_{q(p+t)} < u_n)\right| \leq qt\nu(\Phi \circ T^0 \geq u_n).$$

Now, for $1 \leq l \leq q$

$$
\begin{aligned}
&\left|\nu(M_{l(p+t)} < u_n) - (1 - p\nu(\Phi \circ T^0 \geq u_n))\nu(M_{(l-1)(p+t)} < u_n)\right| \\
&\quad = \left|p\nu(\Phi \circ T^0 \geq u_n)\nu(M_{(l-1)(p+t)} < u_n) + \nu(M_{l(p+t)} < u_n) - \nu(M_{(l-1)(p+t)} < u_n)\right| \\
&\leq \left|p\nu(\Phi \circ T^0 \geq u_n)\nu(M_{(l-1)(p+t)} < u_n) - \sum_{j=1}^{p} \int \chi_{\{\Phi\circ T^j \geq u_n\}} \chi_{\left\{M_{(l-1)(p+t)} < u_n\right\}} \circ T^{p+t} \mathrm{d}\nu\right| \\
&+ \left|\nu(M_{l(p+t)} < u_n) - \nu(M_{(l-1)(p+t)} < u_n) + \sum_{j=1}^{p} \int \chi_{\{\Phi\circ T^j \geq u_n\}} \chi_{\{M_{(l-1)(p+t)} < u_n\}} \circ T^{p+t} \mathrm{d}\nu\right| \\
&= \left|p\nu(\Phi \circ T^0 \geq u_n)\nu(M_{(l-1)(p+t)} < u_n) - \sum_{j=1}^{p} \int \chi_{\{\Phi\circ T^j \geq u_n\}} \chi_{\{M_{(l-1)(p+t)} < u_n\}} \circ T^{p+t} \mathrm{d}\nu\right| \\
&+ \left|\nu(M_{lp+lt} < u_n) - \nu(M_{lp+lt-(p+t)} < u_n) + \sum_{j=1}^{p} \int \chi_{\{\Phi\circ T^j \geq u\}} \chi_{\{M_{lp+lt-(p+t)} < u_n\}} \circ T^{p+t} \mathrm{d}\nu\right|
\end{aligned}
\tag{2.3.20}
$$

By Lemma 2.2.9 we have

$$
\begin{aligned}
&\left|\nu(M_{lp+lt} < u_n) - \nu(M_{(l-1)(p+t)} < u_n) + \sum_{j=1}^{p} \int \chi_{\{\Phi\circ T^0 \geq u_n\}} \circ T^j \chi_{\{M_{(l-1)(p+t)}\}} \circ T^{p+t} \mathrm{d}\nu\right| \\
&\qquad \leq 2p \sum_{j=1}^{p} \int \chi_{\{\Phi\circ T^0 \geq u_n\}} \chi_{\{\Phi\circ T^0 \geq u_n\}} \circ T^j \mathrm{d}\nu + t\nu(\Phi \circ T^0 \geq u_n)
\end{aligned}
\tag{2.3.21}
$$

For the remaining part,

$$
\begin{aligned}
&\left|p\nu(\Phi \circ T^0 \geq u_n)\nu(M_{(l-1)(p+t)} < u_n) - \sum_{j=1}^{p} \int \chi_{\{\Phi\circ T^j \geq u_n\}} \chi_{\{M_{(l-1)(p+t)} < u_n\}} \circ T^{p+t} \mathrm{d}\nu\right| \\
&\leq \sum_{j=1}^{p} \left|\nu(\Phi \circ T^0 \geq u_n)\nu(M_{(l-1)(p+t)} < u_n) - \int \chi_{\{\Phi\circ T^j \geq u_n\}} \chi_{\{M_{(l-1)(p+t)} < u_n\}} \circ T^{p+t} \mathrm{d}\nu\right| \\
&\leq pC_1 \frac{e^{-v\frac{D+\epsilon}{\kappa}}}{n^{\frac{D+\epsilon}{D\kappa}}} + p\frac{C_2 n^{\frac{1+\epsilon}{D}}}{e^{-v(1+\epsilon)} t^{\alpha}}
\end{aligned}
\tag{2.3.22}
$$

for large n by Lemma 2.12.

Define

$$\Gamma_n := t\nu(\Phi\circ T^0\geq u_n)+2p\sum_{j=1}^{p}\int \chi_{\{\Phi\circ T^0\geq u_n\}}\chi_{\{\Phi\circ T^0\geq u_n\}}\circ T^j \mathrm{d}\nu + pC_1\frac{e^{-v\frac{D+\epsilon}{\kappa}}}{n^{\frac{D+\epsilon}{D\kappa}}}+p\frac{C_2 n^{\frac{1+\epsilon}{D}}}{e^{-v(1+\epsilon)}t^{\alpha}}$$

Therefore we have, for $1\leq l\leq q$

$$\left|\nu(M_{l(p+t)}<u_n)-(1-p\nu(\Phi\circ T^0\geq u_n))\nu(M_{(l-1)(p+t)}<u_n)\right|\leq\Gamma_n.$$

Since $n\nu(\Phi\circ T^0\geq u_n)\to e^{-Dv}H(x,\theta)$, for n large enough, $p\nu(\Phi\circ T^0\geq u_n)<1$, and so on applying the above formula inductively we get

$$\left|\nu(M_{q(p+t)}<u_n)-(1-p\nu(\Phi\circ T^0\geq u_n))^q\right|\leq q\Gamma_n+\frac{C_3\|H\|_{1+\delta}\left(1-p\nu\left(\Phi\geq u_n\right)\right)^q}{n^{\frac{1}{\kappa}}}.$$

We now show that $q\Gamma_n\to 0$ as $n\to\infty$ and this will complete the proof because

$$\left(1-\frac{pq\nu(\Phi\circ T^0\geq u_n)}{q}\right)^q\to e^{-e^{-Dv}H(x,\theta)}.$$

By Lebesgue's Differentiation Theorem, for ν a.e. (x,θ)

$$n\nu(\Phi\circ T^0\geq u_n)\to e^{-Dv}H(x,\theta)$$

and so since

$$\frac{t}{p}\to 0 \text{ as } n\to\infty$$

we have

$$\lim_{n\to\infty} qt\nu(\Phi\circ T^0\geq u_n)=0.$$

Also,

$$nC_1\frac{e^{-v\frac{D+\epsilon}{\kappa}}}{n^{\frac{D+\epsilon}{D\kappa}}}\to 0$$

because $\epsilon > \frac{3D\kappa}{2} - D$. Further,

$$n\frac{C_2 n^{\frac{1+\epsilon}{D}}}{e^{-v(1+\epsilon)}t^{\alpha}} \to 0 \text{ because } \alpha > \frac{\frac{3}{2}+\frac{1}{D}(1+\epsilon)}{\tau}$$

by equation (2.3.19).

For the remaining part

$$\begin{aligned} &qp\sum_{j=t}^{p}\nu(\{\Phi\circ T^0 \geq u_n\}\cap T^{-j}\{\Phi\circ T^0 \geq u_n\}) \\ &\leq qp^2\nu(\Phi\circ T^0 \geq u_n)^2 + qp^2 C_1\frac{e^{-v\frac{D+\epsilon}{\kappa}}}{n^{\frac{D+\epsilon}{D\kappa}}} + qp^2\frac{C_2 n^{\frac{1+\epsilon}{D}}}{e^{-v(1+\epsilon)}t^{\alpha}}. \end{aligned} \tag{2.3.23}$$

We show that the terms on the right hand side converge to 0 as $n \to \infty$. Since $qp\nu(\Phi\circ T^0 \geq u_n) \to e^{-Dv}H(x,\theta)$,

$$qp^2\nu(\Phi\circ T^0 \geq u_n)^2 \sim \frac{e^{-2Dv}H(x,\theta)^2}{q} \to 0 \text{ as } q \to \infty.$$

Next, by (2.3.18),

$$qp^2 C_1\frac{e^{-v\frac{D+\epsilon}{\kappa}}}{n^{\frac{D+\epsilon}{D\kappa}}} \sim \frac{1}{n^{\frac{3}{2}-\frac{D+\epsilon}{D\kappa}}} \to 0.$$

And, further,

$$qp^2\frac{C_2 n^{\frac{1+\epsilon}{D}}}{e^{-v(1+\epsilon)}t^{\alpha}} \sim \frac{1}{n^{\tau\alpha-\frac{3}{2}+\frac{1+\epsilon}{D}}} \to 0.$$

Also, from Lemma 2.11,

$$qp\sum_{j=1}^{t}\nu(\Phi\circ T^0 > u_n, \Phi\circ T^j > u_n) \to 0 \text{ because } t = n^{\tau} \text{ and } \tau \leq \gamma'. \tag{2.3.24}$$

This completes the proof. $\square$

2.4 Applications and Examples

We now verify the conditions of Theorem 2.1 and hence establish extreme value theory for S^1 skew extensions of piecewise C^2 uniformly expanding maps of the interval, non-uniformly expanding maps of the interval modeled by Young towers and a skew-product map with a curve of neutral points. We will also discuss briefly two other important classes of maps: extensions to the Manneville-Pommeau-type maps and the Viana-type maps. In the course of the discussion we will sketch why these maps satifsy all but one of the hypotheses of Theorem 2.1.

2.4.1 Uniformly and non-uniformly expanding maps of an interval modeled by Young towers

2.4.1.1 Piecewise C^2 uniformly expanding maps of the interval

We suppose that $f : I \to I$ is a piecewise C^2 map of an interval I onto itself in the sense that there is a finite partition $\{I_i\}$ of the interval I, f is C^2 on the interior of each I_j, $f : I_j \to I$ is onto and monotone, and $|f^{'}(x)| > 1 + \delta$ for all x lying in the interior of each I_j. It is known from [30] that such maps possess an absolutely continuous mixing invariant measure μ and there exists a C such that $\frac{1}{C} \leq \frac{d\mu}{dm} \leq C$. Let $x, y \in \{z : d(z, f^j z) < \frac{1}{n}\} \cap I_i$. We can see that, by the mean-value theorem,

$$\begin{aligned}(1+\delta)d(x,y) < (1+\delta)^j d(x,y) < |(f^j)'|d(x,y) = d(f^j x, f^j y) \\ \leq d(f^j x, x) + d(f^j y, y) + d(x,y)\end{aligned}$$

and so $d(x,y) < \frac{2/\delta}{n}$. Thus on summing over the contribution of each I_i we get an estimate of the form $m\{x : d(x, f^j x) < \frac{1}{n}\} \leq \frac{C_2}{n}$. Thus, for any $1 > \gamma' > 0$,

$$m\left\{x : d(x, f^j x) < \frac{1}{n} \text{ for some } j \in \{1, \ldots, n^{2\gamma'}\}\right\} \leq \frac{1}{n^{1-2\gamma'}}.$$

In particular, choosing $\gamma' < \frac{1}{4}$, we see that (A3) of Theorem 2.1 holds.

Such maps possess a Young tower with exponential return time tails [75], hence, as shown in [35] for a residual set of S^1 cocycles $h : I \to S^1$, the skew-extension T of the base map f has exponential decay of correlations. Thus this class of maps satisfies the conditions of Theorem 2.1.

2.4.1.2 Non-uniformly expanding maps modeled by a Young tower

Suppose $f : X \to X$ is a non-uniformly expanding map of an interval with bounded derivative, i.e., $\sup_{x \in X} |f'(x)| < C$, modeled by a Young tower with exponential return time tails. Collet [21] has shown that there exists a $\beta > 0$ for which $\mu(E_n^X) < \frac{C}{n^\beta}$ and so by Proposition 2.10 we may conclude that the system $T : X \times S^1 \to X \times S^1$ defined by $T(x,\theta) = (f(x), \theta + h(x))$ for any measurable cocycle h satisfies this property. Further, Gouëzel shows in [35] that for a residual set of Hölder cocycles, such systems satisfy the second hypothesis (A4) of Theorem 2.1 for an arbitrary α (by showing that decay is in fact exponential). Since the map T along the group S^1 is an isometry, it's density with respect to the Lebesgue measure is 1 and hence the density of the invariant measure is just the density for f. Collet [21] shows that this density lies in $\mathbb{L}^{1+\delta}$ for some δ larger than 0, and so all the hypotheses of Theorem 2.1 are satisfied.

2.4.2 Skew product with a curve of neutral points

We consider Gouëzel's map studied, for instance, in [32]. Define $f : S^1 \to S^1$ by $f(x) = 4x$ and $T_\alpha : [0,1] \to [0,1]$ as

$$T_\alpha(y) = \begin{cases} y(1 + 2^\alpha y^\alpha) & \text{if } 0 \le y \le \frac{1}{2} \\ 2y - 1 & \text{if } \frac{1}{2} < y \le 1 \end{cases} \tag{2.4.25}$$

where $\alpha : S^1 \to (0,1)$ is a C^2 map with minimum $\alpha_{\min}$ and a maximum $\alpha_{\max}$ and satisfies $0 < \alpha_{\min} < \alpha_{\max} < 1$, $\alpha_{\max} < \frac{3}{2}\alpha_{\min}$, and $\{y : \alpha(y) = \alpha_{\min}\} = \{y_0\}$ with $\alpha''(y_0) > 0$. The map $T : S^1 \times [0,1] \to S^1 \times [0,1]$ is defined as $T(x,y) = (f(x), T_{\alpha(x)}(y))$. From [32, Theorem 2.10], the density H of the map T is $\mathbb{L}^1$ with respect to the product $\mu \times \text{Leb}$ where μ is the invariant measure on S^1 for f (and is the same as the Lebesgue measure). Since f is uniformly expanding, by exactly the same argument as in section 2.4.1.1 we see that (A3) of Theorem 2.1 is satisfied. Further, from [31], if Φ is any Hölder function with exponent $\hat{\alpha}$,

$$\left| \int \Phi\Psi \circ T^n - \int \Phi \int \Psi \right| \le C n^{1-1/\alpha_{max}} \|\Phi\|_{\hat{\alpha}} \|\Psi\|_\infty \tag{2.4.26}$$

and so (A4) is also satisfied. Further, by [32, Theorem 2.10], the density H is Lipschitz on every compact subset of $S^1 \times (0,1]$. The only places in the proof of Theorem 2.1 that we require the density to be in $\mathbb{L}^{1+\delta}$ is to estimate the volume of balls, and this requirement can be replaced by the Lipschitz requirement on every compact subset. Recall, that B_r is a ball about a fixed point (x,y) of radius r, and

$$\|\Phi_B - \chi_B\|_1^\nu \le \nu(B_r \setminus B_{r'}) = \int H \chi_{B_r \setminus B_{r'}} d\lambda.$$

Fix a closed ball Γ with center (x, y). For r sufficiently small, $B_r \subset \Gamma$ and so $\|H\,|_\Gamma\,\|_\infty < \infty$. Therefore

$$\int H\chi_{B_r \setminus B_{r'}} d\lambda \leq \|H\,|_\Gamma\,\|_\infty \lambda(B_r \setminus B_{r'})$$

and so bounds of the type of Lemma 2.12 may be obtained with κ set equal to 1. Now, if we choose $\epsilon > \frac{D}{2}$, in equation (2.3.23) we have

$$qp^2 \frac{e^{-v(D+\epsilon)}}{n^{1+\frac{\epsilon}{D}}} \to 0.$$

The last term in equation (2.3.23) will converge to 0 if the function α is chosen so that $\alpha_{\max}$ satisfies

$$\alpha_{\max} < \frac{\min\{\gamma', \frac{1}{2}\}}{\min\{\gamma', \frac{1}{2}\} + \frac{1}{D}\left(1 + \frac{D}{2}\right) + \frac{3}{2}}.$$

2.4.3 Some other extensions

2.4.3.1 The Viana maps

Let f be a uniformly expanding map of the circle S^1 given by $f(\theta) = d\theta \bmod 1$ for $d \geq 16$. Suppose $b : S^1 \to S^1$ is a Morse function, that $u_\alpha(\theta, x) = a_0 + \alpha b(\theta) - x^2$ and that a_0 is chosen so that $x = 0$ is pre-periodic for $a_0 - x^2$. Let $T_\alpha(\theta, x) = (f(\theta), u_\alpha(\theta, x))$. From [2], for small enough α, there is an interval $I \subset (-2, 2)$ for which $T_\alpha(S^1 \times I) \subset \text{ int } (S^1 \times I)$.

Along the base, this map exhibits a uniformly expanding behavior, and thus, from Proposition 2.10, we can conclude that (A3) of Theorem 2.1 is satisfied. Also, it has been shown in [33] that such a system displays a decay of correlations at the rate of $O(e^{-c\sqrt{n}})$ which is faster than any polynomial. From [2], we know that the density of the absolutely continuous invariant measure lies in $\mathbb{L}^1(\lambda)$. If we knew that this density was in $\mathbb{L}^{1+\delta}(\lambda)$ for small $\delta > 0$, then all the hypotheses of Theorem 2.1 would be satisfied and in that case the limiting distribution obtained would be

$$\lim_{n\to\infty} \nu(M_n^{(x,\theta)} < u_n) = e^{-H(x,\theta)e^{-2v}}.$$

2.4.3.2 Manneville-Pommeau-type maps

We will consider the Liverani-Saussol-Vaienti map $f : [0,1] \to [0,1]$ defined as

$$f(x) = \begin{cases} x(1+2^\omega x^\omega) & x \in [0, \frac{1}{2}) \\ 2x-1 & x \in [\frac{1}{2}, 1] \end{cases}$$

Near the origin, this map is $x \mapsto x + 2^\omega x^{1+\omega}$ and the density near the origin is seen to be $h(x) \approx x^{-\omega}$ so $h \in \mathbb{L}^{\frac{1}{\omega}-\epsilon}$ for any $\epsilon > 0$. It is a result from [41] that

$$\mu_X \left\{ x : d(f^j x, x) < \frac{1}{n} \text{ for some } 0 \le j \le g(n) \right\} \le \left(\frac{g(n)}{\sqrt{n}} \right)^{1-\omega}$$

so if we choose u to be a cocycle, $g(n) = n^{\frac{1-\omega}{24}}$ and $\beta = \frac{1-\omega}{8}$, we see that for $Y = S^1$ we have $D = 2$, $\gamma' = \frac{1-\omega}{24} < \frac{\beta}{D}$ and $\mu_X(E_n^X) < \frac{C}{n^\beta}$. Further, since we have an isometry along the fiber, the density H for ν will lie in $\mathbb{L}^{\frac{1}{\omega}-\epsilon}$ and so all the hypotheses of Theorem 2.1 are met except that the rate of decay of correlations for such an extension $T = (f, u)$ is not known. If a rate satisfying (A4) can be established, we will be able to establish the extreme value law.

Chapter 3

EVT for Uniformly Hyperbolic Maps Exhibiting Singularities

3.1 Statement of Results

Let X be a Riemannian manifold with Lebesgue measure λ and let $T : X \to X$ be a (local) diffeomorphism modeled by a Young tower. The Young tower assumption implies that there exists a subset $\Lambda \subset X$ such that Λ has a hyperbolic product structure and that (P1)-(P5) of [75] hold. We refer the reader to Young's paper [75] and the book by Baladi [9] for details. A similar axiomatic construction of a tower is given by Chernov [19] which is a good reference for background on dispersing billiard maps and flows.

By taking T to be a local diffeomorphism we allow the map T or its derivative to have discontinuities or singularities.

We make the following assumption.

(A5) For μ a.e. $x_0 \in X$ there exists $\tilde{d} := \tilde{d}(x_0) > 0$ such that if $A_{r,\epsilon}(x_0) = \{y \in X : r \leq d(x_0, y) \leq r + \epsilon\}$ is a shell of inner radius r and outer radius $r + \epsilon$ about the point x_0 and if r sufficiently small, $0 < \epsilon \ll r < 1$, then $\mu(A_{r,\epsilon}(x_0)) \leq \epsilon^{\tilde{d}}$.

For systems modeled by a Young tower with exponential return time tails satisfying (A5), we will verify condition $D_2(u_n)$. Planar dispersing billiards with finite horizon and Lozi-like maps satisfy (A5) and may be modeled by a Young tower with exponential return time tails. For planar dispersing billiards with infinite horizon we will use the results of [19]. For these systems we also verify condition $D'(u_n)$. Our method of proof uses ideas from Collet [21] but the arguments need to be modified due to the stable foliation, unbounded derivative and, in the case of Lozi maps, the dissipative nature of the SRB measure.

3.1.1 Framework of the proof

Henceforth, we will fix a reference point x_0 in the support of μ and define a stochastic process X_n given by $X_n(x) = -\log d(T^n x, x_0)$. This observation determines the extreme value distribution of more general functions with unique maximum at the point x_0 [41, 28]. We are interested in the distribution of the maximum of X_n, denoted by

$$M_n = \max\{X_0, X_1, \ldots, X_n\}.$$

We will prove the condition $D_2(u_n)$ [27] for a sequence u_n for which $n\mu(B(x_0, e^{-u_n(v)})) \to e^{-v}$ for some $v \in \mathbb{R}$. We define $\kappa(n)$ to be the rate of decay of correlations of Lipschitz functions with respect to the SRB measure μ on the manifold: so that

$$|\int_X \phi\psi \circ T^n d\mu - \int_X \phi d\mu \int_X \psi d\mu| \leq \kappa(n) \|\phi\|_{Lip} \|\psi\|_{Lip}$$

for all Lipschitz ϕ, $\psi : X \to \mathbb{R}$. In fact we may use the L^∞ norm of ψ in the estimate above as ψ is defined on the quotiented tower (see [75, Section 4]) and in general a faster decay rate than $\kappa(n)$ holds. We assume in this chapter that there exists $\theta \in (0,1)$ such that $\kappa(n) \leq \theta^n$.

We define

$$B_{r,k}(x_0) = \left\{x : T^k(W^s_\eta(x)) \cap \partial B(x_0, r) \neq \emptyset\right\}$$

where $B(x_0, r)$ is the ball of radius $r > 0$ about x_0.

An immediate consequence of (A5) is the following:

Proposition 3.1. *Under (A5) there exist constants $C > 0$ and $0 < \tau_1 < 1$ such that for any r, k*

$$\mu(B_{r,k}(x_0)) \leq C\tau_1^k.$$

Proof. As a consequence of [75, (P2)], there exist $\alpha \in (0,1)$ and $C > 0$ such that $d(T^n(x), T^n(y)) \leq C\alpha^n$ for all $y \in W^s_\eta(x)$. In particular, this implies that

$$|T^k(W^s_\eta(x))| \leq C\alpha^k$$

where $|\ldots|$ denotes the length with respect to the Lebesgue measure. Therefore, $T^k(B_{r,k}(x_0))$ lies in an annulus of width $2C\alpha^k$ around the boundary of the ball of radius r centered at the point x_0. By (A5) and invariance of μ the result follows. $\square$

3.2 Condition $D_2(u_n)$

In this section, we establish condition $D_2(u_n)$ for maps modeled by a Young tower with exponential tails satisfying (A5). Our main theorem for this section is:

Theorem 3.2. *Let $T : (X, \mu) \to (X, \mu)$ be a dynamical system modeled by a Young tower with exponential tails satisfying (A5). Then the stochastic process $X_n := -\log d(T^n x, x_0)$ satisfies the condition $D_2(u_n)$, namely, for any integers j, l and n,*

$$|\mu(\{X_0 > u_n\} \cap \{M_{j,l} \leq u_n\}) - \mu(\{X_0 > u_n\})\,\mu(\{M_{0,l} \leq u_n\})| \leq \gamma(n, j) \quad (3.2.1)$$

where $\gamma(n, j)$ is non-increasing in j for each n and $n\gamma(n, t_n) \to 0$ as $n \to \infty$ for some sequence $t_n = o(n)$, $t_n \to \infty$.

We now show how $D_2(u_n)$ is used, along with a version of $D'(u_n)$ to obtain extreme laws. This uses a blocking argument of Collet [21] based on extreme value statistics (Collet attributes this approach to Galambos [29]). This argument proceeds along the lines of the proof of Theorem 2.1.

3.2.1 The blocking argument.

We will divide successive observations $\{X_0, \dots, X_{n-1}\}$ of length n into q blocks of length $p + t$. The gap t will be large enough that successive p blocks are approximately independent but small enough so that $\mu(M_n \leq u_n)$ is approximately equal to $\mu(M_{qp} \leq u_n)$. For the purposes of our applications, which have exponential decay of correlations, we may take $p \approx \sqrt{n}$ and $t = t_n = (\log(n))^5$ but the method is quite flexible. Using approximate independence of p blocks we show $\mu(M_n > u_n) \approx \mu(M_p > u_n)^q$ and $\mu(M_n \leq u_n) \approx 1 - \mu(M_p \geq u_n)^q$. More precisely Collet, using general set inclusions and probabilistic arguments shows

$$|\mu(M_n \leq u_n) - (1 - p\mu(X_0 > u_n))^q| \leq q\Gamma_n$$

where

$$\Gamma_n = p\gamma(n,t) + t\mu(X_0 > u_n) + 2p\sum_{j=1}^{p-1} \mu(\{X_0 > u_n\} \cap \{X_j > u_n\}).$$

By assumption

$$\lim_{n\to\infty} n\mu(X_0 > u_n) = e^{-v}$$

so

$$\lim_{n\to\infty} \mu(M_n \leq u_n) = e^{-e^{-v}}$$

provided $q\Gamma_n \to 0$. The term $pq\gamma(n,t_n) \to 0$ from $D_2(u_n)$ while $qt\mu(X_0 > u_n) \to 0$ as $n\mu(X_0 > u_n) \to e^{-v}$ and $t = o(n)$. Finally we need to check $n\sum_{j=1}^{p-1}\mu(\{X_0 > u_n\} \cap \{X_j > u_n\}) \to 0$. This is a version of $D'(u_n)$ as in applications p is prescribed as a function of n (for example $p = \sqrt{n}$). In our applications we will give more details in our proofs on the interplay of non-recurrence and decay of correlations needed to ensure $n\sum_{j=1}^{p-1}\mu(\{X_0 > u_n\} \cap \{X_j > u_n\}) \to 0$.

3.3 Proof of $D_2(u_n)$.

We now turn to the proof of Theorem 3.2. The constant τ_1 below is from (A5).

Lemma 3.3. *Suppose $\Phi : M \to \mathbb{R}$ is Lipschitz and $\Psi_{a,b}$ is the indicator function*

$$\Psi_{a,b} := 1_{\{X_a \le u_n, X_{a+1} \le u_n, \ldots, X_{a+b} \le u_n\}}$$

There exists $\tau_2 > 0$, $0 < \tau_1 < \tau_2 < 1$, such that for all $j \ge 0$

$$\left| \int \Phi \Psi_{0,l} \circ T^j \, d\mu - \int \Phi \, d\mu \int \Psi_{0,l} \, d\mu \right| \le \mathcal{O}(1) \left(\|\Phi\|_\infty \tau_2^{\lfloor j/2 \rfloor} + \|\Phi\|_{Lip} \theta^{\lfloor j/2 \rfloor} \right) \quad (3.3.2)$$

Proof. Define the function $\tilde{\Phi} : \Delta \to \mathbb{R}$ by $\tilde{\Phi}(x,r) = \Phi(T^r(x))$ and the function $\tilde{\Psi}_{a,b}(x,r) = \Psi_{a,b}(T^r(x))$. We choose a reference unstable manifold $\tilde{\gamma}^u \subset \Delta_0$ and by the hyperbolic product structure each local stable manifold $W^s_\eta(x)$ will intersect $\tilde{\gamma}^u$ in a unique point $\hat{x}$. Here x denotes a point in the base of the tower Δ_0 and we therefore have $x \in W^s_\eta(\hat{x})$.

We define the function $\overline{\Psi}_{a,b}(x,r) := \tilde{\Psi}_{a,b}(\hat{x}, r)$. We note that $\overline{\Psi}_{a,b}$ is constant along stable manifolds in Δ and the set of points where $\overline{\Psi}_{a,b} \ne \tilde{\Psi}_{a,b}$ is, by definition, the set of (x,r) which project to points $T^r(x)$ for which there exist x_1, x_2 on the same local stable manifold as $T^r(x)$ for which

$$x_1 \in \{X_a \le u_n, \ldots, X_{a+b} \le u_n\}$$

but

$$x_2 \notin \{X_a \le u_n, \ldots, X_{a+b} \le u_n\}$$

This set is contained inside $\cup_{k=a}^{a+b} B_{u_n,k}$. If we let $a = \lfloor j/2 \rfloor$ and $b = l$ then by Proposition 3.1 we have

$$\nu \left\{ \tilde{\Psi}_{\lfloor j/2 \rfloor, l} \ne \overline{\Psi}_{\lfloor j/2 \rfloor, l} \right\} \le \sum_{k=\lfloor j/2 \rfloor}^{l+\lfloor j/2 \rfloor} \mu(B_{u_n,k}) \le \mathcal{O}(1) \tau_2^{\lfloor j/2 \rfloor},$$

$0 < \tau_1 < \tau_2 < 1$.

By the decay of correlations as proved in [75] under the assumption of exponential tails, we have

$$\left|\int \tilde{\Phi}\overline{\Psi}_{\lfloor j/2\rfloor,l} \circ F^{j-\lfloor j/2\rfloor} d\nu - \int \tilde{\Phi} d\nu \int \overline{\Psi}_{\lfloor j/2,l\rfloor} d\nu\right| \leq \mathcal{O}(1)\|\Phi\|_{\mathrm{Lip}}\|\Psi\|_{\infty}\theta^{\lfloor j/2\rfloor}.$$

Therefore,

$$\begin{aligned}
&\left|\int \Phi\Psi_{\lfloor j/2\rfloor} \circ T^{j-\lfloor j/2\rfloor} d\mu - \int \Phi d\nu \int \Psi_{\lfloor j/2\rfloor,l} d\mu\right| \\
&= \left|\int \tilde{\Phi}\tilde{\Psi}_{\lfloor j/2\rfloor,l} \circ F^{j-\lfloor j/2\rfloor} d\nu - \int \tilde{\Phi} d\nu \int \tilde{\Psi}_{\lfloor j/2,l\rfloor} d\nu\right| \\
&\leq \left|\int \tilde{\Phi}\left(\tilde{\Psi}_{\lfloor j/2\rfloor,l} - \overline{\Psi}_{\lfloor j/2\rfloor,l}\right) \circ F^{j-\lfloor j/2\rfloor} d\nu\right| + \mathcal{O}(1)\|\Phi\|_{\mathrm{Lip}}\theta^{\lfloor j/2\rfloor} \\
&\quad + \left|\int \tilde{\Phi} d\nu \int \left(\overline{\Psi}_{\lfloor j/2\rfloor,l} - \tilde{\Psi}_{\lfloor j/2\rfloor,l}\right) \circ F^{j-\lfloor j/2\rfloor} d\nu\right| \\
&\leq \mathcal{O}(1)\left(\|\Phi\|_{\infty}\nu\left\{\overline{\Psi}_{\lfloor j/2\rfloor,l} \neq \tilde{\Psi}_{\lfloor j/2\rfloor,l}\right\} + \|\Phi\|_{\mathrm{Lip}}\theta^{\lfloor j/2\rfloor}\right) \\
&\leq \mathcal{O}(1)\left(\|\Phi\|_{\infty}\tau_2^{\lfloor j/2\rfloor} + \|\Phi\|_{\mathrm{Lip}}\theta^{\lfloor j/2\rfloor}\right).
\end{aligned} \tag{3.3.3}$$

We complete the proof by observing that $\int \Psi_{0,l} d\mu = \int \Psi_{\lfloor j/2\rfloor,l} d\mu$ by the μ invariance of T and that $\Psi_{\lfloor j/2\rfloor,l} \circ T^{j-\lfloor j/2\rfloor} = \Psi_{j,l} = \Psi_{0,l} \circ T^j$.

□

To prove condition $D_2(u_n)$, we will approximate the characteristic function of the set $\{X_0 > u_n\}$ by a suitable Lipschitz function. This approximation will decrease sharply to zero near the boundary of the set $\{X_0 > u_n\}$. The bound in Lemma 3.3 involves the Lipschitz norm, therefore, we need to be able to bound the increase in this norm. To this end, we prove our next lemma.

Lemma 3.4. *1. For μ a.e. x_0 for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n \geq N$*

$$\frac{1}{d+\epsilon}(v + \log n) \leq u_n(v) \leq \frac{1}{d-\epsilon}(v + \log n)$$

2. Denote by $S(n, x_0) := A_{(e^{-u_n} - e^{-u_n^2}, e^{-u_n^2})}(x_0)$ the annulus formed by the region between balls of radius e^{-u_n} and $e^{-u_n} - e^{-u_n^2}$ about x_0. There exists a $\delta(x_0) \in (0,1)$ such that for n large enough

$$\mu(S(n, x_0)) \leq \mathcal{O}(1)(n^{-2\delta v - \delta \log n}).$$

Proof. (1) By the definition of d, for any $\epsilon > 0$ there exists an N_1 such that for all $n \geq N_1$, $(e^{-u_n})^{(d+\epsilon)} \leq \mu(B(x, e^{-u_n})) \leq (e^{-u_n})^{(d-\epsilon)}$. Since we have assumed $\lim_{n\to\infty} n\mu(B(x, e^{-u_n})) \to e^{-v}$, we must have $\limsup n(e^{-u_n})^{d+\epsilon} \leq e^{-v}$. Since $e^{-v} > 0$, this implies given $\eta > 0$ there exists N_2 such that $n(e^{-u_n})^{d+\epsilon} \leq (1+\eta)e^{-v}$ for all $n \geq N_2$.

For the other direction, since $\liminf n(e^{-u_n})^{d-\epsilon} \geq e^{-v}$ there exists N_3 such that for all $n \geq N_3$, $n(e^{-u_n})^{d-\epsilon} \geq (1-\eta)e^{-v}$. Since η was arbitrary the result follows.

(2) The proof follows from part *(1)* and (A5). There exists a $\delta \in (0,1)$ such that

$$\mu(S(n, x_0)) \leq \mathcal{O}(1)|S(n, x_0)|^{\delta}$$

where $|\cdot|$ denotes the width of the annulus. From part (1), $|S(x_0, n)|^{\delta} \leq e^{-u_n^2 \delta} \leq \exp(-\delta(1/(d+\epsilon)^2 v^2 + (\log n)^2 + 2v \log n)) \leq \mathcal{O}(1)(n^{-2\delta' v - \delta' \log n})$ for some $\delta' > 0$.

∎

We note that if the map T preserves an absolutely continuous measure, as in the case of dispersing billiards, then this estimate can be obtained trivially. We are now ready to prove Theorem 3.2.

Proof of Theorem 3.2. We approximate the indicator function $1_{\{X_0>u_n\}}$ by a Lipschitz continuous function Φ as follows. The set $\{X_0 > u_n\}$ corresponds to a ball of radius e^{-u_n} centered at the point x_0. We define Φ to be 1 inside a ball centered at x_0 of radius $e^{-u_n} - e^{-u_n^2}$ and decaying to 0 at a linear rate on $S(n, x_0)$ so that on the boundary of $\{X_0 > u_n\}$, Φ vanishes. The Lipschitz norm of Φ is seen to be bounded by $\exp(u_n^2)$. Since

$$\begin{aligned}
&\left|\int 1_{\{X_0>u_n\}}\Psi_{\lfloor j/2\rfloor,l}\circ T^{j-\lfloor j/2\rfloor}d\mu - \mu(X_0>u_n)\int \Psi_{\lfloor j/2\rfloor,l}d\mu\right| \\
\le &\left|\int \left(1_{\{X_0>u_n\}}-\Phi\right)\Psi_{\lfloor j/2\rfloor,l}d\mu\right| + \mathcal{O}(1)\left(\|\Phi\|_\infty j^2\tau_2^{\lfloor j/4\rfloor} + \|\Phi\|_{\text{Lip}}\theta^{\lfloor j/2\rfloor}\right) \\
&+\left|\int \left(1_{\{X_0>u_n\}}-\Phi\right)d\mu\int \Psi_{\lfloor j/2\rfloor,l}d\mu\right|, \qquad (3.3.4)
\end{aligned}$$

and because $\|1_{\{X_0>u_n\}} - \Phi\|_1 \le \mu(S(n, x_0))$, we have

$$|\mu(\{X_0 > u_n\}| \cap \{M_{j,l} \le u_n\}) - \mu(\{x_0 > u_n\})\mu(\{M_{0,l} \le u_n\}) \le \gamma(n, j)$$

where

$$\gamma(n,j) = \mathcal{O}(1)\left(n^{-2\delta' v-\delta'\log n} + n^{2v+\log n}\theta_1^{\lfloor j/2\rfloor}\right)$$

where $\theta_1 = \max\{\tau_2, \theta\}$. Let $j = t_n = (\log n)^5$. Then $n\gamma(n, t_n) \to 0$ as $n \to \infty$. Note that we had considerable freedom of choice of t_n. Anticipating our applications, we choose $t_n = (\log n)^5$.

□

3.4 Applications

We now prove condition $D'(u_n)$ for some concrete examples. We consider Lozi maps and Sinai dispersing billiards. These are (almost) hyperbolic systems that admit invariant cone fields, but the derivative map DT is discontinuous or singular. We discuss these in further detail below.

3.4.1 Planar dispersing billiard maps and flows

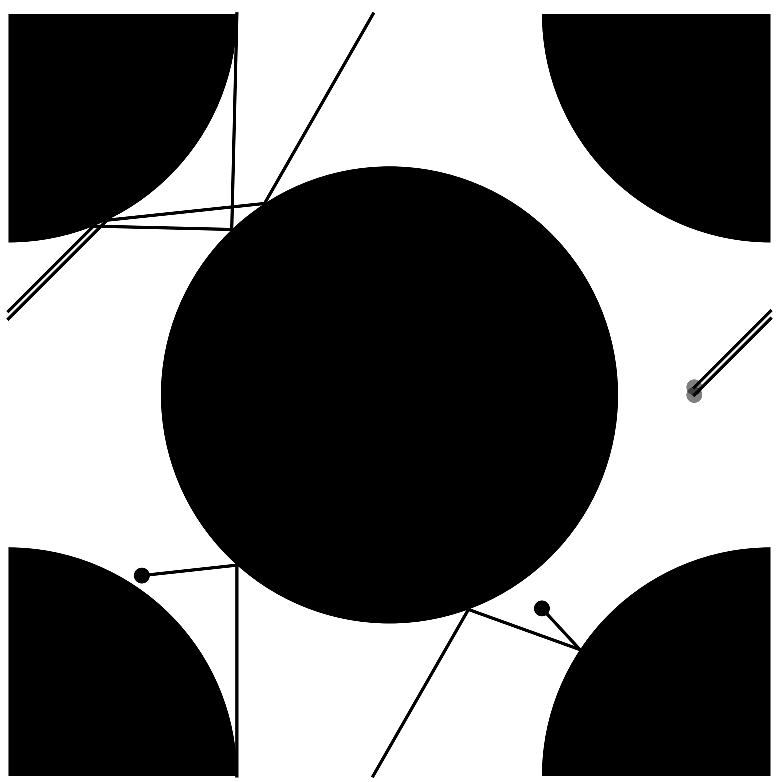

Figure 3.4.1 *Planar dispersing billiards with convex circular scatterers, also known as Lorenz gas. The picture shown here is on $\mathbb{T}^2$, the two-dimensional torus. The light gray dots (in the center right of the image) shows the initial position of two trajectories, while the dark dots show their position at time $t = 1$.*

Let $\Gamma = \{\Gamma_i, i = 1 : k\}$ be a family of pairwise disjoint, simply connected C^3 curves with strictly positive curvature on the two-dimensional torus $\mathbb{T}^2$.

The billiard flow B_t is the dynamical system generated by the motion of a point particle in $Q = \mathbb{T}^2/(\cup_{i=1}^k($ interior$\Gamma_i)$ with constant unit velocity inside Q and with elastic reflections at $\partial Q = \cup_{i=1}^k \Gamma_i$, where elastic means "angle of incidence equals angle of reflection" (see figure 3.4.1).

If each Γ_i is a circle then this system is called a periodic Lorentz gas, a well-studied model in physics. The billiard flow is Hamiltonian and preserves a probability measure (which is Liouville measure) $\tilde{\mu}$ given by $d\tilde{\mu} = C_Q dq\ dt$ where C_Q is a normalizing constant and $q \in Q$,$t \in \mathbb{R}$ are Euclidean coordinates.

From the flow B_t, we may construct the "collision map" $T : \partial Q \to \partial Q$ by considering only the positions and angles of collisions. Let r be a one-dimensional coordinatization of Γ corresponding to length and let $n(r)$ be the outward normal to Γ at the point r. For each $r \in \Gamma$ we consider the tangent space at r consisting of unit vectors v such that $(n(r), v) \geq 0$. This gives us the possible angles a particle can take after it collides with the boundary. We identify each such unit vector v with an angle $\theta \in [-\pi/2, \pi/2]$, which is the angle the outgoing trajectory makes with the unit normal. The boundary X is then parametrized by $X := \partial Q = \Gamma \times [-\pi/2, \pi/2]$ so that X consists of the points (r, θ). $T : X \to X$ is the Poincaré map that gives the position and angle $T(r, \theta) = (r_1, \theta_1)$ after a point (r, θ) flows under B_t and collides again with X, according to the rule "angle of incidence equals angle of reflection". Thus if a trajectory starting at (r, θ) flies for time $h(r, \theta)$ before it collides with ∂Q again, we may set $T(r, \theta) = B_{h(r,\theta)}(r, \theta)$. The billiard map preserves a measure

$d\mu = c_X \cos\theta dr d\theta$ equivalent to two-dimensional Lebesgue measure $dm = dr d\theta$ with density $\rho(x)$ where $x = (r, \theta)$.

Under the assumption of finite horizon, namely, that the time of flight $h(r, \theta)$ is bounded above, Young [75] proved that the billiard map has exponential decay of correlations for Hölder observations. This settled a long-standing question about the rate of decay of correlations in such systems. Chernov [19] extended this result to planar dispersing billiards with infinite horizon where $h(x, r) < \infty$ for all but finitely many points (r, θ) but is not essentially bounded. Chernov also proved exponential decay for dispersing billiards with corner points (a class of billiards we do not discuss in this chapter). A good reference for background results for this section are the papers [17, 16, 75, 19]. We first establish extreme value statistics for billiard maps and then, in the next section, deduce corresponding limit laws for billiard flows using the results of Holland et al [41].

(A5) is satisfied by planar dispersing billiards with finite and infinite horizon as the invariant measure is equivalent to Lebesgue. This is proved in [17, Appendix 2] where it is shown that $\tilde{d}$ may be taken as 1 in the case of finite horizon and $4/5$ in the case of infinite horizon. The proof of $D_2(u_n)$ is immediate in the case of dispersing billiard maps with finite horizon, as they are modeled by a Young tower in [75], have exponential decay of correlations. Chernov [19, Section 5] (see also [17, Section 5]) constructs a Young tower for billiards with infinite horizon to prove exponential decay of correlations so that condition $D_2(u_n)$ is satisfied by this class of billiard map as well. Hence we need only prove condition $D^{'}(u_n)$.

It is known (see [19, Lemma 7.1] for finite horizon and [19, Section 8] for infinite horizon) that dispersing billiard maps expand in the unstable direction in the Euclidean metric $|.| = \sqrt{(dr)^2 + (d\phi)^2}$, in that $|DT_u^n v| \geq C\varrho^n |v|$ for some constants C, $\varrho > 1$ which is independent of v. In fact $|L_n| \geq C\varrho^n |L_0|$ where L_0 is a segment of unstable manifold (once again in the Euclidean metric) and L_n is $T^n L_0$.

We choose N_0 so that $\rho := C\varrho^{N_0} > 1$ and then T^{N_0} (or DT^{N_0}) expands unstable manifolds (tangent vectors to unstable manifolds) uniformly in the Euclidean metric.

It is common to use the p-metric in proving ergodic properties of billiards. Young uses this semi-metric in [75]. Recall that for any curve γ, the p-norm of a tangent vector to γ is given as $|v|_p = \cos\phi(r)|dr|$ where γ is parametrized in the (r, ϕ) plane as $(r, \phi(r))$. The Euclidean metric in the (r, ϕ) plane is given by $ds^2 = dr^2 + d\phi^2$; this implies that $|v|_p \leq \cos\phi(r)ds \leq ds = |v|$. We will use $l_p(C)$ to denote the length of a curve in the p-metric and $l(C)$ to denote length in the Euclidean metric. If γ is a local unstable manifold or local stable manifold then $C_1 l(\gamma)_p \leq l(\gamma) \leq C_2\sqrt{l_p(\gamma)}$.

For planar dispersing billiards there exists an invariant measure μ (which is equivalent to 2-dimensional Lebesgue measure) and through μ a.e. point x there exists a local stable manifold $W^s_{loc}(x)$ and a local unstable manifold $W^u_{loc}(x)$. The SRB measure μ has absolutely continuous (with respect to Lebesgue measure) conditional measures μ_x on each $W^u_{loc}(x)$. The expansion by DT is unbounded however in the p-metric at $\cos\theta = 0$ and this may lead to quite different expansion rates at different points on $W^u_{loc}(x)$. To overcome this effect and obtain uniform estimates on the densities of conditional SRB measure it is common to definite homogeneous local unstable and local stable manifolds. This is the approach adopted in [17, 16, 19, 75].

Fix a large k_0 and define

$$I_k = \{(r,\theta) : \frac{\pi}{2} - k^{-2} < \theta < \frac{\pi}{2} - (k+1)^{-2}\}$$

$$I_{-k} = \{(r,\theta) : -\frac{\pi}{2} + (k+1)^{-2} < \theta < -\frac{\pi}{2} + k^{-2}\}$$

and

$$I_0 = \{(r,\theta) : -\frac{\pi}{2} + k_0^{-2} < \theta < \frac{\pi}{2} - k_0^{-2}\}.$$

In our setting we call a local unstable (stable) manifold $W^u_{loc}(x)$ ($W^s_{loc}(x)$) homogeneous if for all $n \geq 0$ $T^n W^u_{loc}(x)$ ($T^{-n} W^s_{loc}(x)$) does not intersect any of the line segments in $I_{k_0} \cup \cup_k (I_k \cup I_{-k})$. Homogeneous $W^u_{loc}(x)$ have almost constant conditional SRB densities $\frac{d\mu_x}{d\lambda_x}$ in the sense that there exists $C > 0$ such that $\frac{1}{C} \leq \frac{d\mu_x(z_1)}{d\lambda_x} / \frac{d\mu_x(z_2)}{d\lambda_x} \leq C$ for all $z_1,\ z_2 \in W^u_{loc}(x)$ (see [19, Section 2] and the remarks following Theorem 3.1).

From this point on, all the local unstable (stable) manifolds that we consider will be homogeneous. Bunimovich et al [17, Appendix 2, Equation A2.1] give quantitative estimates on the length of homogeneous $W^u_{loc}(x)$. They show that there exists $C,\ \tau > 0$ such that $\mu\{x : l(W^s_{loc}(x)) < \epsilon \text{ or } l(W^u_{loc}(x)) < \epsilon\} \leq C\epsilon^\tau$ where $l(C)$ denotes 1-dimensional Lebesgue measure or length of a rectifiable curve C. In our setting τ could be taken to be $\frac{2}{9}$, its exact value will play no role, but for simplicity, in the forthcoming estimates we assume $0 < \tau < \frac{1}{2}$.

The natural measure μ has absolutely continuous conditional measures μ_x on local unstable manifolds $W^u_{loc}(x)$ which have almost uniform densities with respect to Lebesgue measure on $W^u_{loc}(x)$ by [19, Equation 2.4].

We prove,

Theorem 3.5. *Let $T : X \to X$ be a planar dispersing billiard map. Then for μ a.e. x_0 the stochastic process defined by $X_n(x) = -\log(d(x_0, T^n x))$ satisfies a Type I extreme value law in the sense that* $\lim_{n\to\infty} \mu(M_n \leq (v + \log n + \log(\rho(x_0)))/2) = e^{-e^{-v}}$.

Proof. Let $A_{\sqrt{\epsilon}} = \{x : |W^u_{loc}(x)| > \sqrt{\epsilon}\}$ then $\mu(A^c_{\sqrt{\epsilon}}) < C\epsilon^{\tau/2}$. Let $x \in A_{\sqrt{\epsilon}}$ and consider $W^u_{loc}(x)$. Since $|T^{-k}W^u_{loc}(x)| < \lambda^{-1}|W^u_{loc}(x)|$ for $k > N_0$ the optimal way for points $T^{-k}(y)$ in $T^{-k}W^u_{loc}(x)$ to be close to their preimages $y \in W^u_{loc}(x)$ is for $T^{-k}W^u_{loc}(x)$ to overlay $W^u_{loc}(x)$, in which case it has a fixed point and it is easy to see that

$$l\{y \in W^u_{loc}(x) : d(y, T^{-k}y) < \epsilon\} \leq l\{y \in \mathbb{R} : d(y, \frac{y}{\lambda}) < \epsilon\} \leq (1 - \lambda^{-1})^{-1}\epsilon.$$

Accordingly $l\{y \in W^u_{loc}(x) : d(y, T^{-k}y) < \epsilon\} \leq C\sqrt{\epsilon}l\{y \in W^u_{loc}(x)\}$. Recalling that the density of the conditional SRB-measure λ_x is bounded above and below with respect to one-dimensional Lebesgue measure we obtain $\mu_x(A_{\sqrt{\epsilon}} \cap \{y \in W^u_{loc}(x) : d(y, T^{-k}y) < \epsilon\}) < C\sqrt{\epsilon}$. Integrating over all unstable manifolds in $A_{\sqrt{\epsilon}}$ (throwing away the set $\mu(A^c_{\sqrt{\epsilon}})$ we have $\mu\{x : d(T^{-k}x, x) < \epsilon\} < C\epsilon^{\tau/2}$. Since μ is T-invariant, and because T is invertible, $\mu\{x : d(T^k x, x) < \epsilon\} < C\epsilon^{\tau/2}$ for $k > N_0$. Hence for any iterate T^k, $k > N_0$

$$\mathcal{E}_k(\epsilon) := \mu\{x : d(T^k x, x) < \epsilon\} < C\epsilon^{\tau/2}$$

Recall that the scaling constant $u_n(v)$ is chosen so that $n\mu(B(x_0, e^{-u_n(v)}) \to e^{-v}$. For hyperbolic billiards we take $u_n(v) = \frac{1}{2}(v + \log n + \log(\rho(x_0)))$ and shrinking balls of radius roughly $\frac{1}{\sqrt{n}}$ about points. This leads to the use of $\frac{1}{\sqrt{k}}$ in the next definition. Define

$$E_k := \{x : d(T^j x, x) \leq \frac{2}{\sqrt{k}} \text{ for some } 1 \leq j \leq (\log k)^5\}$$

We have shown that for any $\delta > 0$, for all sufficiently large k, $\mu(E_k) \leq k^{-\tau/4+\delta}$. For simplicity we take $\mu(E_k) \leq k^{-\sigma}$ where $\sigma < \tau/4 - \delta$ and omit the constant e^{-v} in the following equations.

Define the Hardy-Littlewood maximal function $\mathfrak{L}_l$ for $\phi(x) = 1_{E_l}(x)\rho(x)$ where $\rho(x) = \frac{d\mu}{dm}(x)$, so that

$$\mathfrak{L}_l(x) := \sup_{a>0} \frac{1}{\lambda(B_a(x))} \int_{B_a(x)} 1_{E_l}(y)\rho(y)dm(y)$$

A theorem of Hardy and Littlewood [53, Theorem 2.19] implies that

$$\lambda(|M_l| > C) \leq \frac{\|1_{E_l}\rho\|_1}{C}$$

where $\|\cdot\|_1$ is the L^1 norm with respect to m. Let

$$F_k := \{x : \mu(B_{k^{-\gamma/2}}(x) \cap E_{k^{\gamma/2}}) \geq (k^{-\gamma\beta/2})k^{-\gamma/2}\}$$

Then $F_k \subset \{M_{k^{\gamma/2}} > k^{-\gamma\beta/2}\}$ and hence

$$\lambda(F_k) \leq \mu(E_{k^{\gamma/2}})k^{\gamma\beta/2} \leq Ck^{-\gamma\sigma}k^{\gamma\beta/2}$$

If we take $0 < \beta < \sigma$ and $\gamma > \sigma/2$ then for some $\delta > 0$, $k^{-\gamma\sigma}k^{\gamma\beta/2} < k^{-1-\delta}$ and hence

$$\sum_k \lambda(F_k) < \infty$$

Thus for λ a.e. (hence μ a.e.) $x_0 \in X$ there exists $N(x_0)$ such that $x_0 \notin F_k$ for all $k > N(x_0)$. Thus along the subsequence $n_k = k^{\gamma/2}$, $\mu(X_0 > u_{n_k}, X_0 \circ T^j >$

$u_{n_k}) \leq n_k^{-1-\delta}$ for $k > N(x_0)$, $j \leq (\log n_k)^5$. This is sufficient to obtain an estimate for all u_n. From the estimate for n_k, we may deduce that if $k^{\gamma/2} \leq n \leq (k+1)^{\gamma/2}$, then $\mu(X_0 > u_n, X_0 \circ T^j > u_n) \leq \mu(X_0 \geq u_n, X_0 \circ T^j \geq u_n) \leq n_k^{-1-\delta}$ for $j \leq \min\{(\log(n_k))^5, (\log(n))^5\}$. But if n is large enough (note that this forces k to be large), then

$$k^{\gamma/2} \leq n \leq (k+1)^{\gamma/2} \implies 0 \leq \log n - \gamma/2 \log k \leq \frac{\gamma}{2} \log \frac{k+1}{k} \to 0.$$

Hence $\log n \approx \log n_k$, so we may in fact take $j \leq (\log n)^5$.

We now control the iterates $1 \leq j \leq N_0$. If x_0 is not periodic then

$$s(x_0) := \min_{1 \leq i < j \leq N_0} d(T^i x_0, T^j x_0) > 0$$

and hence for large enough n, for all $1 \leq j \leq N_0$, $\mu(X_0 > u_n, X_0 \circ T^j > u_n) = 0$.

Recalling that u_n was chosen so that $n\mu(B_{e^{-u_n}}(x)) \to e^{-v}$, we get, for any $1 \leq j \leq (\log n)^5$,

$$\mu(X_0 > u_n, X_0 \circ T^j > u_n) \leq 2n^{-1-\delta}$$

Hence

$$\lim_{n\to\infty} n \sum_{j=1}^{(\log n)^5} \mu(X_0 > u_n, X_0 \circ T^j > u_n) = 0$$

We now use exponential decay of correlations to show

$$\lim_{n\to\infty} n \sum_{(\log n)^5}^{p=\sqrt{n}} \mu(X_0 > u_n, X_0 \circ T^j > u_n) = 0$$

We let 1_{u_n} denote the indicator function of the set $\{X_0 > u_n\}$. We approximate the indicator function 1_{u_n} by a Lipschitz function ϕ_n which is 1 on a neighborhood of x_0 of radius $e^{-u_n} - e^{-u_n^2}$ and linearly decays to 0 outside it, so that on the boundary of the ball of radius e^{-u_n}, $\phi_n = 0$. ϕ_n has Lipschitz norm bounded by $e^{(u_n)^2} \approx e^{(v+\log n)^2/4}$.

Note that

$$
\begin{aligned}
|\int 1_{u_n}\ 1_{u_n}\circ T^j\ d\mu - (\int 1_{u_n}\ d\mu)^2| \ &\leq\ |\int \phi_n\ \phi_n\circ T^j\ d\mu - (\int \phi_n\ d\mu)^2| \\
&+\ |(\int \phi_n\ d\mu)^2 - (\int 1_{u_n}\ d\mu)^2| \\
&+\ |\int 1_{u_n}\ 1_{u_n}\circ T^j\ d\mu - \int \phi_n\ \phi_n\circ T^j\ d\mu|.
\end{aligned}
$$

If $(\log n)^5 \leq j \leq p = \sqrt{n}$ then by decay of correlations $|\int \phi_n(x)\phi_n\circ T^j(x) - (\int \phi_n)^2 d\mu| \leq Ce^{2u_n^2}\theta^j \leq Ce^{-2\log n} = \frac{C}{n^2}$, if n is sufficiently large. Furthermore if n is large $|(\int \phi_n d\mu)^2 - (\int 1_{u_n}\ d\mu)^2| < Cn^{-2(\delta v - \delta\log n)} < Cn^{-2}$.

Finally $|\int \phi_n\phi_n\circ T^j\ d\mu - 1_{u_n}1_{u_n}\circ T^j\ d\mu| \leq \mu(\phi_n(x) \neq 1_{X_0>u_n}) + \mu(\phi_n\circ T^j(x) \neq 1_{X_0\circ T^j(x)}) \leq \frac{C}{n^2}$ since the supports of ϕ_n and $\phi_n\circ T^j$ are contained inside the supports of 1_{u_n} and $1_{u_n}\circ T^j$.

Hence

$$
n\lim_{n\to\infty}\sum_{j=(\log n)^5}^{p=\sqrt{n}} \mu(X_0 > u_n, X_0\circ T^j > u_n) = 0.
$$

This concludes the proof of Theorem 3.5.

□

3.4.1.1 Billiard flows

In section 3.4.1, we established extreme value theory for the map T obtained by restricting attention to successive collisions for the billiards flow B_t. The full flow B_t

can be viewed, given the collision map T, as a suspension flow over T, with time of flight given by the function h. This viewpoint lets us exploit the results of Holland et al. [41], on extreme value theory for suspension flows as a consequence of the theory for the base maps.

Suppose $B_t : Q \to Q$ is the billiard flow preserving the ergodic invariant natural measure $\tilde{\mu}$ and $\bar{h}$ is the average first return time of the billiard flow from the boundary to the boundary with respect to μ i.e. $\bar{h} = C_X \int_X h(r, \phi) \cos \phi d\phi dr$ where $h(r, \phi)$ is the time of flow till the point $(r, \phi) \in \partial Q$ hits the boundary again ∂Q. As a consequence of [41, Theorem 2.10], we have the following corollary,

Corollary 3.6. *For $\tilde{\mu}$ a.e. $p_0 \in Q$, if $\phi(p) = -\log d(p, p_0)$ and $M_t := \max_{s\le t}\{\phi(B_s(p))\}$ then*

$$\lim_{t\to\infty} \tilde{\mu}(M_t \le v + \log(t/\bar{h}) + \log \rho(x_0)/2) = e^{-e^{-v}}.$$

3.4.2 Lozi-like maps

The Lozi mapping T is a homeomorphism of $\mathbb{R}^2$ given by

$$(x, y) \to (1 + y - a|x|, bx)$$

where a and b are parameters. It has been studied as a model of chaotic dynamics intermediate in complexity (or difficulty) between Axiom A diffeomorphisms and Henon diffeomorphisms [57, 22, 74]. The derivative is discontinuous on the y-axis and this leads to arbitrarily short smooth local unstable manifolds. Misuiurewicz [57] proved that there exists an open set G of parameters such that if $(a, b) \in G$ the map T is hyperbolic. If $(a, b) \in G$, then $T_{a,b}$ has invariant stable and unstable directions (where the derivative is defined) and the angle between them is bounded below by $\pi/5$. We will restrict our attention to maps with parameters in the set G.

These maps admit a strict cone, and the tangent derivatives, where defined, satisfy uniform expansion estimates [22] in that there exists $\rho > 1$ such that $|DT^n v| \geq \rho^n v$ for all $v \in E^u$ (the unstable direction) and correspondingly for E^s (the stable direction). $T_{a,b}$ has an invariant ergodic probability measure μ [22] which is absolutely continuous with respect to the one-dimensional Lebesgue measure along local unstable curves. In fact the conditional invariant measure on local unstable manifolds is simply one-dimensional Lebesgue measure [22]. Young [74] established similar results for a broader class of maps, 'generalized' Lozi maps which are piecewise C^2 mappings of the plane. But one reason for restricting to maps $T_{a,b}$, $(a, b) \in G$ is that for such maps Collet and Levy have also shown that for μ almost every point on the attractor the Hausdorff dimension of μ exists and is constant [22]. We need this fact for verifying Lemma 3.4, which is an essential ingredient in the verification of condition $D_2(u_n)$ for these maps.

The existence of a dimension d implies that for almost every x in the attractor, the dimension constant $d(x)$ in the definition of u_n is the same. We will use a sequence of scaling constants $u_n(x_0, v)$ defined for a generic point x_0 by the requirement that $n\mu(B(x_0, e^{-u_n(x_0,v)})) \to e^{-v}$.

In later work [75, Section 7] Young constructs SRB measures via a tower construction for a broader class of piecewise C^2 uniformly hyperbolic maps of the plane. The

Lozi map $T_{a,b}$, with $(a,b) \in G$, b sufficiently small, may be modeled by a Young tower with exponential tails [75]. Hence the Lozi maps we consider satisfy exponential decay of correlations for Hölder continuous observations.

We now summarize the ergodic properties of the Lozi maps that we will use. T has an invariant SRB-measure μ and μ a.e. point x has a local stable manifold $W^s_{loc}(x)$ and local unstable manifold $W^u_{loc}(x)$. In [22, Proposition IV.1] it is shown that the conditional measures of μ on the local unstable manifolds are the corresponding 1-d Lebesgue measures. Furthermore μ a.e. point x has a quadrilateral $\beta(x)$ with a local product structure, in the sense that $y \in \beta(x)$ implies there exists a unique $z \in \beta(x)$ such that $z = W^u_{loc}(y) \cap W^s_{loc}(x)$ and a unique $z' \in \beta(x)$ such that $z' = W^s_{loc}(y) \cap W^u_{loc}(x)$ [22, Section 4]. Suppose that $W^u_{loc}(x)$ and $W^u_{loc}(x')$ are local unstable manifolds. Then the holonomy $h : W^u_{loc}(x) \to W^u_{loc}(x')$ is defined on the set $D(h) := \{x \in W^u_{loc}(x) : W^s_{loc}(x) \cap W^s_{loc}(x') \neq \emptyset\}$. The holonomy between local unstable manifolds satisfies the following quantitative estimates,

Proposition 3.7. *[22, Proposition II.4] Given $W^u_{loc}(x)$ and $W^u_{loc}(x')$ there is a constant L such that for any Borel subset $A \subset W^u_{loc}(x) \cap D(h)$,*

$$(1 - L(d(W^u_{loc}(x), W^u_{loc}(x'))^{1/3})l(A) \leq l(h(A)) \leq (1 + L(d(W^u_{loc}(x), W^u_{loc}(x'))^{1/3})l(A)$$

Note that the local unstable manifolds lie in a strict cone and the conditional invariant measure on local unstable manifolds is one-dimensional Lebesgue measure. Suppose that $x_0 \in X$, $A_{r,\epsilon}(x_0)$ is an annulus with center x_0 and λ_x is conditional measure on $W^u_{loc}(x)$ with x in the quadrilateral $\beta(x_0)$.

Proposition 3.8. *(A5) is satisfied for the Lozi maps, if $(a,b) \in G$.*

Proof. Around almost every point x_0, there exists a quadrilateral $\beta(x_0)$ which has a hyperbolic product structure. Furthermore, the conditional measure on each local unstable manifold is the 1-d Lebesgue measure. If any unstable leaf completely stretches across the annulus of width ϵ around a ball of inner radius r, then it intersects the annulus in a curve of length at most $C\sqrt{\epsilon}$ for some constant C which may depend only on x_0 but not on ϵ or r. Since the measure is supported on the unstable leafs, the assertion holds.

□

Thus we need only show condition $D^{'}(u_n)$. We will establish $D^{'}(u_n)$ in this section and prove the following theorem.

Theorem 3.9. *Let $T_{a,b} : X \to X$ be a Lozi map with $(a, b) \in G$ with b sufficiently small. Then for μ a.e. x_0 the stochastic process defined by $X_n(x) = -\log(d(x_0, T^n x))$ satisfies a Type I extreme value law in the sense that $\lim_{n\to\infty} \mu(M_n \leq u_n(x_0, v)) = e^{-e^{-v}}$*

Remark 3.10. *We do not know the precise scaling constants $u_n(x_0, v)$, but for all $\epsilon > 0$, $\lim_{n\to\infty} \mu(M_n \leq (1 - \epsilon)(\log n + v)/d)) \leq e^{-e^{-v}} \leq \lim_{n\to\infty} \mu(M_n \leq (1 + \epsilon)(\log n + v)/d))$ which provides an estimate of the correct sequence u_n.*

Proof. We need only establish $D^{'}(u_n)$. We will denote the length of a rectifiable curve C by $l(C)$ in the usual Euclidean metric. As tangent vectors to local unstable manifolds lie in a strict cone the projected length onto either the horizontal or vertical axis of a connected component $\mathcal{C}$ of $T^k W^u_{loc}(x)$ is bounded below by $\kappa l(\mathcal{C})$ for some

$\kappa > 0$. This constant will be absorbed into our C's below, so that the expansion of a local unstable manifold under T^j may be used to estimate the measure of points which satisfy $d(x, T^j x) < \epsilon$. The projection of $W^u_{loc}(x)$ onto the horizontal axis expands uniformly for all $j > N_0$ for some N_0, but as in the case of billiards this does not affect our argument if x_0 is not a periodic point. For simplicity of exposition we assume $N_0 = 1$.

One would think that as the derivative is bounded and there is uniform expansion in the unstable direction, which lies within a cone, the proof of $D^{'}(u_n)$ would be immediate but the presence of discontinuities for the derivative complicates the picture. If $W^u_{loc}(x)$ is a local unstable manifold then $T^n(W^u_{loc}(x))$ is either a line segment or a connected broken line segment. Here is what could possibly go wrong in the latter case. Suppose that the map T (restricted to local unstable manifolds), expands uniformly and $|T^{'}(x)| > \rho > 1$. Let L be a segment of unstable manifold and consider $T^n L$. It expands but may encounter the set of discontinuities/singularities S. Suppose $T^n L$ is partitioned into X smooth components β_i with corresponding pre-image intervals $\alpha_i \subset L$ so that $T^n \alpha_i = \beta_i$. Suppose the map T^n folds back on itself many times and places each β_i atop α_i such that the left endpoint x_i of α_i lies very close to the left endpoint $T^n x_i$ of β_i. If $|\beta_i| < \epsilon$ then each point in α_i lies within ϵ of its image under T^n. We have to show this cannot happen. We use the structure of a Young tower to do this.

We first show there exists $\sigma > 0$ such that that for a generic point z (generic here means a set of points of full measure) if ϵ is sufficiently small then $\lambda_z(y \in W^u_{loc}(z) : d(y, T^j y) < \epsilon) < \epsilon^\sigma l(W^u_{loc}(z))$ where λ_z is conditional measure on $W^u_{loc}(z)$.

Assume that $z = \pi(x, r)$ for some $(x, r) \in \Delta$ i.e. for some i, $x \in \Lambda_i \subset \Lambda_0$, $r < R(x) = R_i$ we have $T^r x = z$. We may assume without loss of generality that $W^u_{loc}(z) \supset T^r(W^u_{loc}(x) \cap \Lambda_i)$, otherwise we could refine the partition on the tower by defining a new return time on the base $R_k(y) = R(y) + \ldots + R(f^k y)$. By refining in this way we could also require $(T^R)'(x) > 2$. This is equivalent to considering a tower with return time partition $\bigvee_0^k P_0 := \bigvee_{j=0}^{k-1} f^{-j} P_0$ on the base, where P_0 is the original partition into sets $\{\Lambda_i\}$. For large enough k, $W^u_{loc}(z) \supset T^r(W^u_{loc}(x) \cap \Lambda_i)$. Note that the new tower will also have exponential return time tails. We identify $W^u_\rho(z) := T^r(W^u_{loc}(x) \cap \Lambda_i) \cap W^u_{loc}(z) \subset X$ with $W^u_{loc}(x, r)$ on the tower. The portion of local unstable manifold $W^u_\rho(z)$ may not be symmetrical about z but this will not affect our argument.

There exists $\tau > 0$ such that if $\epsilon > 0$ is sufficiently small then except for a set of $\Lambda_i' s$ of λ_0 measure less than ϵ^τ, $|W^u_\eta(y) \cap \Lambda_i| > \sqrt{\epsilon}$ for all $y := (y, 0) \in \Lambda_i$. This observation uses exponential decay of the return time. To see this suppose that Λ_n has return time R_n. Let γ be the length of Λ_0 in the unstable direction. Since $|T'(x)| < K$ is bounded we have $K^{R_n}|\Lambda_n| \approx \gamma$ and hence $|\Lambda_n| \geq \gamma e^{-R_n \log K}$. So if $R_n < \frac{-\log \epsilon}{2 \log K}$ then $|\Lambda_n| > \gamma\sqrt{\epsilon}$. Since we have exponential return time tails, $m(x \in \Lambda_0 : R(x) > T) \leq C\theta^T$ for some $0 < \theta < 1$. Hence $\lambda_0(\cup \Lambda_n \subset \Lambda_0 : |\Lambda_n| < \sqrt{\epsilon}) < \epsilon^\tau$ for some $\tau \approx \frac{-\log \theta}{2 \log K}$. Choose $0 < \sigma < 1$ so that $\epsilon^\sigma > \epsilon^\tau + \sqrt{\epsilon}$ for sufficiently small $\epsilon > 0$.

Now $T^j(W^u_\rho(z))$ expands uniformly for $j = 1$ to $R_i - r$ then makes a full crossing of the base Λ_0. By full crossing we mean that on the quotiented tower $T^j(W^u_\rho(z)) = \tilde{\Lambda_0}$, where $\tilde{\Lambda_0}$ is Λ_0 quotiented along stable manifolds. By the same argument as in the case of billiards, by uniform expansion, for sufficiently small $\epsilon > 0$ for each $j < R_i - r$, $l\{x \in W^u_\rho(z) : d(x, T^j x) < \epsilon\} \leq \tilde{C}\epsilon < C\sqrt{\epsilon} l(W^u_\rho(z))$ where l is one-dimensional Lebesgue measure. For $j = R_i - r$, $T^j(W^u_\rho(z))$ has made a full crossing and this partitions $W^u_\rho(z)$ into components C_k such that for each C_k, $T^{R_i - r} C_k$ crosses Λ_k. By bounded distortion, (except for a set of measure less than $\epsilon^\tau l(W^u_\rho(z))$ corresponding to those C_k such that $T^{R_i-r}C_k$ crosses Λ_k and $|\Lambda_k| < \sqrt{\epsilon}$) each C_k satisfies $l(T^{R_i-r}C_k) > \sqrt{\epsilon}$. If $l(T^{R_i-r}C_k) > \sqrt{\epsilon}$ then $l(y \in C_k : d(y, T^{R_i-r}y) < \epsilon) \leq \sqrt{\epsilon} l(C_k)$. This proves $l(y \in W^u_\rho(z) : d(T^{R_i-r}y, y) < \epsilon) \leq \epsilon^\sigma l(W^u_\rho(z))$ as $\epsilon^\sigma > \sqrt{\epsilon} + \epsilon^\tau$.

Each set $T^{R_i-r}C_k$ expands uniformly under T^s until $s = R_k$, so for $s < R_k$, $l(y \in C_k : d(y, T^{s+R_i-r}y) < \epsilon) \leq \sqrt{\epsilon} l(C_k)$. When $s = R_k$, $T^s T^{R_i-r} C_k$ has made a full crossing and C_k is partitioned into sets A^k_j such that $T^s T^{R_i-r} A^k_j = \Lambda_j \cap T^s T^{R_i r} C_k$. By bounded distortion except for a set of A^k_j's of measure less than $\epsilon^\tau l(C_k)$, each A^k_j satisfies $l(T^s T^{R_i-l} A^k_j) > \sqrt{\epsilon}$ in which case $l(y \in A^k_j : d(y, T^{s+R_i-r}y) < \epsilon) \leq \sqrt{\epsilon} l(A^k_j)$. Thus $l(y \in C_k : d(y, T^{s+R_i-r}y) < \epsilon) \leq \epsilon^\sigma l(C_k)$ for $0 \leq s \leq R_k$.

Given t we induce a partition of $W^u_\rho(z)$ by writing, for each $x \in W^u_\rho(z)$, $t = R(x) - r + R(fx) + \ldots + R(f^{n(x)}(x)) + k(x)$, so that $F^t(x)$ has made precisely n returns to the base and moved k levels up the tower. This defines a partition of $W^u_\rho(z)$ into intervals I^j_t such that points in I^j_t have not been separated on the tower for $n-1$ returns to the base ($n(x)$ a random variable), then made a full crossing and moved up to level $k(x)$. By bounded distortion and the same argument as in the case of the components C_k, for each I^j_r, $l(x \in I^j_r : d(x, T^r x) < \epsilon) \leq \epsilon^\sigma l(I^j_r)$. This proves that for all j, $l(y \in W^u_\rho(z) : d(y, T^j y) < \epsilon) < \epsilon^\sigma l(W^u_\rho(z))$.

Now let $x \in X$ be a generic point so that for a (sufficiently small) local unstable manifold $W^u_\eta(x)$ we have $l(y \in W^u_\eta(x) : d(y, T^j y) < \epsilon) < \epsilon^\sigma l(W^u_\eta(x))$. We consider $W^u_\eta(x)$ as a measure space equipped with one-dimensional conditional Lebesgue measure λ_x.

Let

$$E_k := \{y \in W^u_\eta(x) : d(T^j y, y) \leq \frac{3}{\sqrt{k}} \text{ for some } 1 \leq j \leq (\log k)^5\}$$

We have shown that for any $\delta > 0$, for all sufficiently large k, $\lambda_x(E_k) \leq k^{-\sigma+\delta}$. For simplicity, we will take $\lambda_x(E_k) \leq k^{-\sigma/2}$.

Define the Hardy-Littlewood maximal function $\mathfrak{L}_l$ for $\phi(y) = 1_{E_l}(y)$ so that

$$\mathfrak{L}_l(p) := \sup_{a>0} \frac{1}{2a} \int_{I_a(p)} 1_{E_l}(y) d\lambda_x(y)$$

where $I_a(x) = \{y \in W^u_\eta(x) : d(y,p) \leq a\}$. By the Hardy and Littlewood Theorem [53, Theorem 2.19] , for any $C > 0$,

$$\lambda_x(|\mathfrak{L}_l| > C) \leq \frac{\|1_{E_l}\|_1}{C}$$

where $\|, \|_1$ is the L^1 norm with respect to λ_x. Let, for some β and γ,

$$F_k := \{z \in W^u_\eta(x) : \lambda_x(I_{k^{-\gamma/2}}(z) \cap E_{k^{\gamma/2}}) \geq (k^{-\gamma\beta/2})k^{-\gamma/2}\}$$

Then $F_k \subset \{\mathfrak{L}_{k^{\gamma/2}} > k^{-\gamma\beta/2}\}$ and hence

$$\lambda_x(F_k) \leq \lambda_x(E_{k^{\gamma/2}})k^{\gamma\beta/2} \leq Ck^{-\gamma\sigma/4}k^{\gamma\beta/2}$$

If we take $0 < \beta < \sigma/4$ and $\gamma > 8/\sigma$ then for some $\delta > 0$ $\lambda_x(F_k) < k^{-1-\delta}$. This implies that

$$\sum_k \lambda_x(F_k) < \infty$$

and hence by the Borel-Cantelli lemma for λ_x a.e. x_0 there exists $N(x_0)$ such that $x_0 \notin F_k$ for all $k > N(x_0)$.

Since x was arbitrary and the invariant measure is carried on unstable manifolds this implies that for μ a.e. x_0 there exists an $N(x_0)$ such that $\lambda_{x_0}\{y \in W^u_{k^{-\gamma/2}}(x_0) : d(T^j y, x_0) < \frac{3}{\sqrt{k^{\gamma/2}}}$ for any $j = 1, \dots, (\log k^{\gamma/2})^5 \} < (k^{\gamma/2})^{-1-\delta}$ for all $k \geq N(x_0)$, or, equivalently, $l(\{y \in W^u_{k^{-\gamma/2}}(x_0) : d(T^j y, x_0) < 3/\sqrt{k^{\gamma/2}}$ for any $j = 1, \dots, (\log k^{\gamma/2})^5\}) < (k^{\gamma/2})^{-1-\delta} k^{-\gamma/2}$. As in the case of billiards, since

$$\lim_{k\to\infty} \left(\frac{k+1}{k}\right)^{\gamma/2} = 1,$$

we obtain the same estimate for all k sufficiently large, not just along the subsequence $k^{\gamma/2}$.

If $\max\{d(y, x_0), d(z, x_0)\} < \frac{1}{\sqrt{k}}$ and $z \in W^s_\eta(y)$ then $d(T^j z, x_0) < \frac{1}{\sqrt{k}}$ implies that $d(T^j y, x_0) < \frac{2}{\sqrt{k}}$ since $d(T^j y, x_0) \leq d(T^j y, T^j z) + d(T^j z, x_0)$. Thus $d(T^j y, x_0) > \frac{2}{\sqrt{k}}$ for all $j = 1, \dots, (\log k)^5$ implies that $d(T^j z, x_0) > \frac{1}{\sqrt{k}}$ for all $j = 1, \dots, (\log k)^5$ for all $z \in W^s_{\frac{1}{\sqrt{k}}}(y)$.

Since the holonomy map satisfies the quantitative estimates of Proposition 3.7 on each unstable manifold $W^u_\eta(x)$ in a neighborhood of x_0 of diameter $\frac{1}{\sqrt{k}}$ for sufficiently large k, $l\{y \in W^u_{1/\sqrt{k}}(x) : d(T^j y, y) < \frac{1}{\sqrt{k}}$ for any $j = 1, \dots, (\log k)^5 \} < \sqrt{k}^{-1-\delta}(1 + \sqrt{k}^{-1/3})k^{-1/2} < \sqrt{k}^{-1-\delta'} k^{-1/2}$ for some $\delta' > 0$. Thus the *fractional* conditional measure on each unstable manifold in a $\frac{1}{\sqrt{k}}$ neighborhood of x_0 of points $y \in W^u_\eta(x)$ such that $d(T^j y, y) < \frac{1}{k}$ for any $j = 1, \dots, (\log k)^5$ is bounded by $(\sqrt{k}^{-1-\delta}(1+\sqrt{k}^{-1/3})) < \sqrt{k}^{-\delta'-1}$.

Recalling that u_n was chosen so that $n\mu(B_{e^{-u_n}}(x)) \to e^{-v}$, we obtain that for any

$1 \le j \le (\log n)^5$,

$$\mu(X_0 > u_n, X_0 \circ T^j > u_n) \le Cn^{-1-\delta'/2}$$

Hence $n \sum_{j=1}^{(\log n)^5} \mu(X_0 > u_n, X \circ T^j > u_n)) \to 0$.

The argument that exponential decay of correlations implies that $n \sum_{(\log n)^5}^{p} \mu(X_0 > u_n, X_0 \circ T^j > u_n) \to 0$ is the same as that for billiards. This concludes the proof of Theorem 3.9.

$\square$

Chapter 4

Some Borel-Cantelli Lemmas

4.1 Setting and Statements of Results

4.1.1 Gibbs-Markov maps

We first describe $1D$ Gibbs-Markov maps (see Definition 4.2) and then show that for such maps, sequences of intervals satisfying a bounded ratio criterion have the sBC property. The base map of a Young tower (see section 1.7) is a Gibbs-Markov system and our result for such systems, Theorem 4.7, will play a crucial role in the proof of Theorems 4.9 and 4.12.

Let $(X, \mathcal{B}, m)$ be a Lebesgue probability space. Let $\mathcal{P}$ be a countable measurable partition of X such that $m(\alpha) > 0$ for all $\alpha \in \mathcal{P}$.

Definition 4.1. *A measure-preserving map $T : X \to X$ is said to be a* ***Markov***

map *if the following are satisfied.*

1. *($\mathcal{P}$ generates $\mathcal{B}$) We have* $\sigma(\{T^{-i}(\alpha) : \alpha \in \mathcal{P},\ i \in \mathbb{Z}^{+}\}) = \mathcal{B} \pmod m$, *where* $\sigma(\cdot)$ *denotes the* σ*-algebra generated by its argument.*

2. *(Markov property) For all* $\alpha, \beta \in \mathcal{P}$, *if* $m(T(\alpha) \cap \beta) > 0$ *then* $\beta \subset T(\alpha) \pmod m$.

3. *(local invertibility) For all* $\alpha \in \mathcal{P}$, $T|\alpha$ *is invertible.*

For $n \in \mathbb{N}$, let $\mathcal{P}_n$ be the refinement of $\mathcal{P}$ defined by

$$\mathcal{P}_n = \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) = \left\{ \bigcap_{i=0}^{n-1} T^{-i}(\alpha_i) : \alpha_i \in \mathcal{P} \text{ for } 0 \leqslant i \leqslant n-1 \right\}.$$

Define

$$J_{T^k} = \frac{d(m \circ T^k)}{dm}.$$

Definition 4.2. *The quintet* $(X, \mathcal{B}, m, T, \mathcal{P})$ *is said to be a* ***Gibbs-Markov system*** *if* T *is a Markov map and the following properties also hold.*

(H1) *(full branches) For all* $\alpha \in \mathcal{P}$, $T(\alpha) = X \pmod m$.

(H2) *(uniform expansion) There exists* $K_1 > 0$ *and* $\gamma_1 \in (0,1)$ *such that* $m(\alpha) \leqslant K_1\gamma_1^n$ *for all* $n \in \mathbb{N}$ *and* $\alpha \in \mathcal{P}_n$.

(H3) *(distortion control) There exists* $K_2 > 0$ *and* $\gamma_2 \in (0,1)$ *such that for all* $n \in \mathbb{N}$ *and* $\alpha \in \mathcal{P}_n$, *we have*

$$\left| \log \left(\frac{J_{T^n}(x)}{J_{T^n}(y)} \right) \right| \leqslant K_2 \gamma_2^n \tag{4.1.1}$$

for all $x, y \in \alpha$.

A consequence of $(H3)$ is that for every $\omega \in \mathcal{P}_n$ and for every $\mathcal{B}$ measurable set A , $m(A \cap \omega)/m(\omega) \leq \mathcal{O}(1)m(T^n A)/m(X)$. We will use this observation in the proof of Theorem 4.9.

Remark 4.3. *Some authors weaken (H1) in the definition of Gibbs-Markov systems by requiring that $m(T(\alpha)) > K > 0$ for some K independent of α.*

Definition 4.4. *The Gibbs-Markov system $(X, \mathcal{B}, m, T, \mathcal{P})$ is said to be a $1D$ Gibbs-Markov system if X is a compact interval and $\mathcal{P}$ is a partition of X into subintervals.*

Now let X be a compact interval. A map $T : X \to X$ is said to be piecewise-differentiable if there exists a coutable partition $\mathcal{P}$ of X into intervals with disjoint interiors such that for all $I \in \mathcal{P}$, T is differentiable on the interior of I. A piecewise-differentiable map $T : X \to X$ is said to be uniformly expanding if there exists $K > 1$ such that $|T'(x)| \geqslant K$ for all x at which $T'(x)$ exists. We similarly define piecewise-C^k maps for $k \geqslant 2$.

For certain piecewise-differentiable uniformly expanding maps, Kim [45, Theorem 2.1] establishes the sBC property for sequences of intervals. His result can be more usefully stated as

Proposition 4.5 ([45])**.** *Suppose T is a piecewise-differentiable uniformly expanding map of the compact interval X and suppose that T admits a unique absolutely continuous invariant probability measure μ with density bounded away from 0. Assume that there exists a summable sequence $(\kappa(n))_{n=1}^{\infty}$ and $C > 0$ such that for all $f \in L^1(\mu)$ and $\psi \in \mathrm{BV}(X)$, we have*

$$\left| \int_X (f \circ T^n)(\psi)\, d\mu - \left(\int_X f\, d\mu \right) \left(\int_X \psi\, d\mu \right) \right| \leqslant C\kappa(n) \|f\|_1 \|\psi\|_{\mathrm{BV}}. \tag{4.1.2}$$

If (A_n) is a sequence of intervals in X and $\sum_{n=0}^{\infty} \mu(A_n) = \infty$ then (A_n) is an sBC sequence.

The proof is the same as that of [45, Theorem 2.1]. As a corollary we have the following sBC result for certain one-dimensional Gibbs-Markov systems.

Lemma 4.6. *Suppose that $(X, \mathcal{B}, m, T, \mathcal{P})$ is a piecewise-C^2 1D Gibbs-Markov system for which there exists $L > 0$ such that*

$$\sup_{\alpha \in \mathcal{P}} \sup_{x \in \overline{\alpha}} \frac{|T''(x)|}{T'(x)^2} \leqslant L < \infty. \tag{4.1.3}$$

Let $(A_n)_{n=0}^{\infty}$ be a sequence of intervals in X. If $\sum_{n=0}^{\infty} m(A_n) = \infty$, then (A_n) is an sBC sequence.

Proof of Lemma 4.6. Condition (4.1.3) is sometimes called the Adler property. It enables us to show that $g := 1/T'$ is of bounded variation. Rychlik [63] has shown that for piecewise-differentiable uniformly expanding maps with g of bounded variation, correlations decay exponentially; that is, (4.1.2) holds with $\kappa(n)$ decaying exponentially (see also [9, 40]). Kim uses the result of Rychlik to establish the sBC property for sequences of intervals in the setting of piecewise-differentiable uniformly expanding maps with g of bounded variation, although his proof is valid if $(\kappa(n))$ is summable.

To see that g is of bounded variation, let $x, y \in \alpha \in \mathcal{P}$. Using (4.1.3) and (H3), we have

$$\begin{aligned}
|g(x)-g(y)| &= \left|\frac{T'(y)-T'(x)}{T'(x)T'(y)}\right| \\
&\leqslant \int_x^y \frac{|T''(s)|}{|T'(x)T'(y)|}\,ds \\
&= \int_x^y \left(\frac{|T''(s)|}{T'(s)^2}\right)\left(\frac{T'(s)^2}{|T'(x)T'(y)|}\right) ds \\
&\leqslant K\,|x-y|\,.
\end{aligned}$$

Using the distortion estimate (H3) again, for every $\alpha \in \mathcal{P}$ and for every $x \in \alpha$, we have

$$e^{-K_2}\left(\frac{\lambda(X)}{\lambda(\alpha)}\right) \leqslant |T'(x)| \leqslant e^{K_2}\left(\frac{\lambda(X)}{\lambda(\alpha)}\right)$$

where λ denotes Lebesgue measure on $\mathbb{R}$. Consequently, if $x \in \alpha \in \mathcal{P}$ and $y \in \beta \in \mathcal{P}$, then

$$|g(x)-g(y)| \leqslant \frac{1}{|T'(x)|} + \frac{1}{|T'(y)|} \leqslant K(\lambda(\alpha)+\lambda(\beta)).$$

We will now let $x_{n_i+1}, \ldots, x_{n_{i+1}} \in \alpha_{i+1}, 0 \leq i < \infty, \sup\{n_i\} = N$ be any finite partition of X. If an α_{i+1} contains none of the points x_i, we let $n_{i+1} = n_i$. Let $\mathfrak{A}$ denote the set of intervals α_i which contain at least one of the points $x_1, \ldots, x_N$. The elements of $\mathfrak{A}$ can be numbered as $\alpha^1, \ldots \alpha^t$ for some t, where $t \leq N$. We can calculate the variation of g as

$$\begin{aligned}
|g(x_1)-g(x_N)| &\leq \sum_{i=0}^{\infty} K|x_{n_{i+1}} - x_{n_i+1}| + \sum_{j=1}^{t-1} K(\lambda(\alpha^j)+\lambda(\alpha^{j+1})) \\
&\leq K\lambda(X) + K\lambda(X) + K\lambda(X) := K_N < \infty.
\end{aligned}$$

Since $\sup_N K_N = 3K\lambda(X)$, g is of bounded variation if the Adler condition holds. $\square$

We now state a result for $1D$ Gibbs-Markov systems without the Adler condition but with a bounded ratio restriction on the sequence of intervals. The proof of Theorem 4.7 is given in Section 4.2.

Theorem 4.7. *Let X be a compact interval and let $\mathcal{P}$ be a countable partition of X into subintervals. Suppose that $(X, \mathcal{B}, m, T, \mathcal{P})$ is a Gibbs-Markov system. Let $(A_n)_{n=0}^{\infty}$ be a sequence of intervals in X for which there exists $C > 0$ such that $m(A_j) \leqslant Cm(A_i)$ for all $j \geqslant i \geqslant 0$. If $\sum_{n=0}^{\infty} m(A_n) = \infty$, then (A_n) is an sBC sequence.*

We have already described Young towers for maps of an arbitrary compact manifold X. In the 1-D setting, we make the following additional assumptions:

(A6) Δ_0, the base of the Young tower Δ, is an interval, as are the partitions Λ_i of Δ_0.

While a general one-dimensional non-uniformly expanding map may not be modeled by a Young tower in which the base is an interval and in which the partitions of the base are also intervals, (A6) is true for all common examples in the mathematical literature, in the sense that all the partition elements can be enclosed within disjoint intervals. Therefore, the setting of our results is not a restrictive setting.

Definition 4.8. *We say that $(M, \mathcal{B}, \mu, T)$ is a $1D$ system modeled by a Young tower if it satisfies the setting described in Section 1.7 and satisfies (A6).*

Theorem 4.9. *Let $(M, \mathcal{B}, \mu, T)$ be a $1D$ system modeled by a Young tower Δ. Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a decreasing function and let $(I(n))_{n=0}^{\infty}$ be a nested sequence of closed* *intervals in M such that $I(n)$ has length $g(n)$ for all $n \in \mathbb{Z}^+$. If $\sum_{n=0}^{\infty} g(n) = \infty$ and*

$$J_\infty := \bigcap_{n=0}^{\infty} I(n) \subset \pi(\Delta),$$

then the following hold.

1. *If J_∞ is an interval, then $(I(n))$ is an sBC sequence.*
2. *There exists a set $\Gamma \subset M$, $\mu(\Gamma) = 1$, such that if $p \in \Gamma$ and $p = J_\infty$ then $(I(n))$ is a dBC sequence satistying*

$$\underline{\lim}_{n\to\infty} \frac{S_n(x)}{E_n} \geqslant 1 \quad a.s.$$

Remark 4.10. *If J_∞ is an interval then the conclusion of Theorem 4.9 follows immediately from the Birkhoff ergodic theorem.*

Remark 4.11. *For any $p \in M$ and for every k and l, either $\Lambda_{k,l} \cap \pi^{-1}(p) = \emptyset$ or $\Lambda_{k,l} \cap \pi^{-1}(p)$ consists of a single point $\hat{p}_{k,l}$. The set Γ consists of points $p \in M$ such that $\hat{p}_{k,l} \in \mathrm{int}(\Lambda_{k,l})$ for all k and l for which $\hat{p}_{k,l}$ is defined and*

$$\lim_{r\to 0} \frac{\mu(I(p,r))}{2r} = \frac{d\mu}{d\lambda}(p) > 0,$$

where $I(p,r)$ is the ball (interval) centered at p of radius r.

The following dBC result for nested intervals centered at a point $p \in M$ does not require that $p \in \Gamma$.

Theorem 4.12. *Suppose that $(M, \mathcal{B}, \mu, T)$ is a $1D$ system modeled by a Young tower Δ. Let $p \in M$ and let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a decreasing function. For $n \in \mathbb{Z}^+$, let $I(n)$ denote the closed interval centered at p of length $g(n)$. If $\sum_{n=0}^{\infty} g(n) = \infty$ and $p \in \pi(\Delta)$, then $(I(n))$ is a dBC sequence with respect to Lebesgue measure λ.*

As mentioned earlier, if $(M, \mathcal{B}, \mu, T)$ is a $C^{1+\varepsilon}$ dynamical system on a compact interval M such that μ is ergodic, $\mu \ll \lambda$, and μ has a positive Lyapunov exponent, then the system can be modeled by a Young tower (personal communication by José Alves and Henk Bruin; see also [4, 5, 15]). This implies the following corollary.

Corollary 4.13. *Suppose that* $(M, \mathcal{B}, \mu, T)$ *is a* $C^{1+\varepsilon}$ *dynamical system on a compact interval* M *such that the invariant probability measure* μ *is ergodic and absolutely continuous with respect to Lebesgue measure. If* $(M, \mathcal{B}, \mu, T)$ *has a positive Lyapunov exponent, then Theorems 4.9 and 4.12 apply to* $(M, \mathcal{B}, \mu, T)$.

Remark 4.14. *In the context of Theorem 4.9,* $(I(n))$ *may not be a dBC sequence if* $J_\infty \not\subset \pi(\Delta)$. *Kim [45] constructs explicit examples of this phenomenon in the context of the Liverani-Saussol-Vaienti maps defined by* (1.6.13) *(see [49]). Theorem 4.9 applies if* $0 < \alpha < 1$. *For every* $0 < \alpha < 1$, *Kim proves that the sequence* $A(n) = [0, n^{-1/(1-\alpha)})$ *satisfies* $\sum_{n=0}^{\infty} \mu(A(n)) = \infty$ *but for* μ *a.e.* $x \in M$, $T_\alpha^n(x) \in A(n)$ *for only finitely many values of* n. *The intervals* $A(n)$ *satisfy*

$$\bigcap_{n=0}^{\infty} A(n) = \{0\}$$

and $0 \notin \pi(\Delta)$ *for the Young tower modeling* $([0,1], \mu, T_\alpha)$ *(for an explicit construction of this tower see e.g. [12]).*

4.1.2 One-dimensional maps

Theorem 4.9 applies to many classes of one-dimensional maps. Here we give a partial list.

1. Pomeau-Manneville intermittent-type maps (such as Liverani-Saussol-Vaienti maps) [49, 62]. See [34, 45] for related results.

2. Certain classes of multimodal maps; see Bruin *et al.* [14].

3. A class of non-uniformly expanding circle maps; see Young [76, Section 6]. Let $T : \mathbb{S}^1 \to \mathbb{S}^1$ be a map of degree $d > 1$ such that T is C^1 on $\mathbb{S}^1$ and C^2 on $\mathbb{S}^1 \setminus \{0\}$, $T' > 1$ on $\mathbb{S}^1 \setminus \{0\}$, $T(0) = 0$, $T'(0) = 1$, and for some $0 < \alpha < 1$, we have $-xT''(x) \sim |x|^\alpha$ for $x \neq 0$.

4.2 Proof for the Gibbs-Markov Case

We prove Theorem 4.7 by using a sufficient condition for the sBC property given in [46, 66]. This sufficient condition has also been used by Chernov and Kleinbock [20] and by Kim [45].

Proposition 4.15. *Let $(X, \mathcal{B}, m)$ be a probability space and let $(B_n)_{n=0}^\infty$ be a sequence of measurable subsets of X such that $\sum_{n=0}^\infty m(B_n) = \infty$. If there exists a constant $C > 0$ such that for all $N \geqslant M \geqslant 0$*

$$\sum_{i,j=M}^{N} m(B_i \cap B_j) - m(B_i)m(B_j) \leqslant C \sum_{i=M}^{N} m(B_i), \tag{4.2.4}$$

then for every $\varepsilon > 0$, we have

$$\sum_{i=0}^{N-1} \mathbf{1}_{B_i}(x) = \sum_{i=0}^{N-1} m(B_i) + \mathcal{O}\left(\left(\sum_{i=0}^{N-1} m(B_i) \right)^{\frac{1}{2}} \log^{\frac{3}{2}+\varepsilon} \left(\sum_{i=0}^{N-1} m(B_i) \right) \right)$$

for m-a.e. $x \in X$.

Remark 4.16. *The implied constant in the error estimate*

$$\mathcal{O}\left(\left(\sum_{i=0}^{N-1} m(B_i)\right)^{\frac{1}{2}} \log^{\frac{3}{2}+\varepsilon}\left(\sum_{i=0}^{N-1} m(B_i)\right)\right)$$

is not uniform but a function of x.

Our proof uses the fact that Gibbs-Markov maps are exponentially continued fraction mixing [1, page 164] in the sense that there exists $\tau \in (0,1)$ and a constant $K_3 > 0$ such that

$$|m(\alpha \cap T^{-(n+k)}(\beta)) - m(\alpha)m(\beta)| \leqslant K_3 \tau^n m(\alpha)m(\beta)$$

for all measurable $\beta \in \mathcal{B}$, $\alpha \in \mathcal{P}_k$.

Proof of Theorem 4.7. Throughout this proof C will be used to denote a constant, whose value may change from line to line. Set $B_i = T^{-i}(A_i)$ in (4.2.4). Notice that if $j \geqslant i$, then $T^{-i}(A_i) \cap T^{-j}(A_j) = T^{-i}(A_i \cap T^{-(j-i)}(A_j))$. Since T preserves m, we have $m(T^{-i}(A_i \cap T^{-(j-i)}(A_j))) = m(A_i \cap T^{-(j-i)}(A_j))$. We therefore estimate

$$\begin{aligned}\sum_{i,j=M}^{N} m(B_i \cap B_j) - m(B_i)m(B_j) &= 2\sum_{i=M}^{N}\sum_{j=i+1}^{N} m(A_i \cap T^{-(j-i)}(A_j)) - m(A_i)m(A_j) \\ &\quad + \sum_{i=M}^{N} m(A_i) - (m(A_i))^2.\end{aligned} \tag{4.2.5}$$

For the diagonal terms, we have the straightforward estimate

$$\sum_{i=M}^{N} m(A_i) - (m(A_i))^2 \leqslant \sum_{i=M}^{N} m(A_i). \tag{4.2.6}$$

Now assume that $j > i$. We estimate $m(A_i \cap T^{-(j-i)}(A_j)) - m(A_i)m(A_j)$.

Let $\mathcal{V}_{i,j} = \left\{\alpha \in \mathcal{P}_{\lceil (j-i)/2 \rceil} : \alpha \cap A_i \neq \emptyset\right\}$. Let $N(i)$ be the largest integer such that A_i intersects at most 2 partition elements of $\mathcal{P}_{\lceil (j-i)/2 \rceil}$ for $j - i < N(i)$. If $N(i) > j - i > 1$ and $j - i$ is even, we have the estimate

$$\begin{aligned}
m(A_i \cap T^{-(j-i)} A_j) - m(A_i)m(A_j) &\leqslant m(A_i \cap T^{-(j-i)} A_j) \\
&= m(A_i \cap T^{-(j-i)/2}(T^{-(j-i)/2} A_j)) \\
&\leqslant 2C\gamma_1^{(j-i)/2} m(T^{-(j-i)/2} A_j) \\
&= 2C\gamma_1^{(j-i)/2} m(A_j) \\
&\leqslant C\gamma_1^{(j-i)/2} m(A_i).
\end{aligned}$$

This holds because

1. using (H1)–(H3), for each $\alpha \in \mathcal{V}_{i,j}$ we have

$$m(\alpha \cap T^{-(j-i)/2}(T^{-(j-i)/2} A_j)) \leqslant Cm(\alpha)m(T^{-(j-i)/2} A_j)$$

   and $m(\alpha) \leqslant K_1 \gamma_1^{(j-i)/2}$,

2. $\#\mathcal{V}_{i,j} \leqslant 2$ since $j - i < N(i)$, and

3. $m(A_j) \leqslant Cm(A_i)$ by assumption.

If $j - i$ is odd we estimate

$$\begin{aligned}
m(A_i \cap T^{-(j-i)} A_j) - m(A_i)m(A_j) &\leq m(A_i \cap T^{-(j-i)} A_j) \\
&= m(A_i \cap T^{-(j-i+1)/2}(T^{-(j-i-1)/2} A_j))
\end{aligned}$$

and proceed as before. In particular, if $j - i = 1$ we use the simple estimate $m(A_i \cap$

$T^{-(j-i)}A_j) \leq m(A_j)$. Thus

$$\sum_{j=i+1}^{i+N(i)-1} m(A_i \cap T^{-(j-i)}A_j) - m(A_i)m(A_j) \leq Cm(A_i)$$

We now consider $j - i \geq N(i)$. We can no longer assume that A_i intersects at most 2 elements of $\mathcal{P}_{\lceil (j-i)/2 \rceil}$. In this case the collection $\mathcal{V}_{i,j}$ induces a partition of A_i. Define

$$\mathcal{V}^1_{i,j} = \{\alpha \in \mathcal{V}_{i,j} : \alpha \subset (\inf(A_i), \sup(A_i))\}.$$

Since $\mathcal{P}$ consists of subintervals, $\mathcal{V}^2_{i,j} := \mathcal{V}_{i,j} \setminus \mathcal{V}^1_{i,j}$ contains at most 2 elements. For $k = 1, 2$ define

$$Q^k_{i,j} = \bigcup \{\alpha : \alpha \in \mathcal{V}^k_{i,j}\}.$$

We have

$$\begin{aligned}
&\sum_{j=i+N(i)}^{N} m(A_i \cap T^{-(j-i)}A_j) - m(A_i)m(A_j) \\
&\quad = \sum_{j=i+N(i)}^{N} m(A_i \cap Q^1_{i,j} \cap T^{-(j-i)}A_j) - m(A_i)m(A_j) \\
&\qquad + \sum_{j=i+N(i)}^{N} m(A_i \cap Q^2_{i,j} \cap T^{-(j-i)}A_j) - m(A_i)m(A_j) \\
&\quad \leqslant \sum_{j=i+N(i)}^{N} m(A_i \cap Q^1_{i,j} \cap T^{-(j-i)}A_j) - m(A_i \cap Q^1_{i,j})m(A_j) && (4.2.7a) \\
&\qquad + \sum_{j=i+N(i)}^{N} m(A_i \cap Q^2_{i,j} \cap T^{-(j-i)}A_j) - m(A_i \cap Q^2_{i,j})m(A_j) && (4.2.7b)
\end{aligned}$$

We estimate $(4.2.7a)$ first. For this we will use the exponential continued fraction mixing estimate of Aaronson [1, page 164]. There exists $C > 0$ and $\tau \in (0, 1)$, both independent of i and j, such that for $\alpha \in \mathcal{V}^1_{i,j}$, we have

$$\left| m(\alpha \cap T^{-(j-i)}(A_j)) - m(\alpha)m(A_j) \right| \leqslant C\tau^{(j-i)/2} m(\alpha)m(A_j).$$

Thus

$$\begin{aligned}\sum_{j=i+N(i)}^{N} \sum_{\alpha \in \mathcal{V}^1_{i,j}} m(\alpha \cap T^{-(j-i)} A_j) - m(\alpha)m(A_j) &\leqslant \sum_{j=i+N(i)}^{N} \sum_{\alpha \in \mathcal{V}^1_{i,j}} \tau^{(j-i)/2} m(\alpha)m(A_j) \\ &\leqslant \sum_{j=i+N(i)}^{N} \tau^{(j-i)/2} m(A_i)m(A_j)\end{aligned}$$

For (4.2.7b) we note that if $\beta \in \mathcal{V}^2_{i,j}$ then $m(\beta) \leqslant K_1 \gamma_1^{\lceil (j-i)/2 \rceil}$. Consider the partition of β induced by $\mathcal{P}_{j-i}$.

For $\omega \in \mathcal{P}_{j-i}$ such that $\omega \subset \beta$, $T^{j-i}(\omega) = X$ by (H1), so the distortion estimate (H3) and uniform expansion (H2) give

$$m(\beta \cap T^{-(j-i)}(A_j)) = \sum_{\substack{\omega \in \mathcal{P}_{j-i} \\ \omega \subset \beta}} m(\omega \cap T^{-(j-i)}(A_j)) \leqslant C\gamma_1^{(j-i)/2} m(A_j)$$

and hence

$$\sum_{j=i+N(i)}^{N} m(A_i \cap Q^2_{i,j} \cap T^{-(j-i)} A_j) - m(A_i \cap Q^2_{i,j})m(A_j) \leqslant \sum_{j=i+N(i)}^{N} 2C\gamma_1^{(j-i)/2} m(A_j).$$

This concludes the proof as $m(A_j) \leqslant Cm(A_i)$. $\square$

4.3 Proofs

4.3.1 Relating base dynamics to tower dynamics and preliminaries

For $x \in \Lambda_0$ and $n \in \mathbb{N}$, we define

$$R_n(x) = \sum_{i=0}^{n-1} R(f^i(x)).$$

As a consequence of the Birkhoff ergodic theorem, we have

Lemma 4.17. *Assume the setting of Theorem 4.9 and define*

$$\langle R \rangle = \int_{\Lambda_0} R \, dm.$$

Then for m a.e. $x \in \Lambda_0$, we have

$$\lim_{n\to\infty} \frac{R_n(x)}{n} = \langle R \rangle. \tag{4.3.8}$$

We now prove some elementary lemmas which will be useful in the proofs of Theorems 4.9 and 4.12.

Lemma 4.18. *Suppose $g : \mathbb{R}^+ \to \mathbb{R}^+$ is decreasing and $\sum_{i=0}^{\infty} g(i) = \infty$.*

(A) *For all $a > 0$ we have*

$$\frac{\int_0^{(1+a)n} g(t)\, dt}{\int_0^n g(t)\, dt} \leqslant 1 + a$$

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

(B)

$$\lim_{n\to\infty} \frac{\int_0^n g(t)\,dt}{\sum_{i=0}^{n-1} g(i)} = 1.$$

Proof of Lemma 4.18. For part (A), observe that since g is decreasing,

$$\begin{aligned}
\int_0^{(1+a)n} g(t)\,dt &= \int_0^n g(t)\,dt + \int_n^{(1+a)n} g(t)\,dt \\
&\leqslant \int_0^n g(t)\,dt + (an)g(n) \\
&\leqslant \int_0^n g(t)\,dt + a\,(g(1) + g(2) + \cdots + g(n)) \\
&\leqslant \int_0^n g(t)\,dt + a\left(\int_0^1 g(t)\,dt + \cdots + \int_{n-1}^n g(t)\,dt\right) \\
&\leqslant (1+a)\int_0^n g(t)\,dt.
\end{aligned}$$

For part (B), the bound

$$\sum_{i=0}^{n-1} g(i) \geqslant \int_0^n g(t)\,dt \geqslant \sum_{i=1}^{n} g(i)$$

implies the result since $\sum_{i=0}^{\infty} g(i) = \infty$. □

Lemma 4.19. *Let (a_n) and (b_n) be sequences in $\mathbb{R}^+$ such that*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L$$

for some L. If $\sum_{i=0}^{\infty} a_i = \infty$, then

$$\lim_{n\to\infty} \frac{\sum_{i=0}^{n-1} a_i}{\sum_{i=0}^{n-1} b_i} = L.$$

Proof. $a_n = Lb_n + c_n$ where $\lim_{n\to\infty} c_n/b_n = 0$. Therefore,

$$\frac{\sum_{i=0}^{\infty} a_i}{\sum_{i=0}^{\infty} b_i} = L + \frac{\sum_{i=0}^{\infty} c_i}{\sum_{i=0}^{\infty} b_i}.$$

If $\sum_{n=0}^{\infty} b_n < \infty$, then $\sum_{n=0}^{\infty} c_n \leq \sum_{n=0}^{\infty} b_n < \infty$ and so $\sum_{n=0}^{\infty} a_n < \infty$. So, $\sum b_n$ diverges. Since $|c_n|/b_n \to 0$, there exists an N_ϵ such that $|c_n| \leq \epsilon b_n$ for all $n \geq N_\epsilon$, from where it follows that for any $M > N_\epsilon$, $\sum_{n\geq N_\epsilon}^{M} |c_n| / \sum_{n\geq N_\epsilon}^{M} b_n \leq \epsilon$. Hence, for any $N > N_\epsilon$

$$\begin{aligned} \left| \frac{\sum_{i=0}^{N} a_i}{\sum_{i=0}^{N} b_i} - L \right| &\leq \frac{\sum_{i=0}^{N} |c_i|}{\sum_{i=0}^{N} b_i} \\ &\leq \frac{\sum_{i<N_\epsilon} |c_i|}{\sum_{i=0}^{N} b_i} + \epsilon. \end{aligned}$$

Choose N so large that $\sum_{i<N_\epsilon} |c_i| / \sum_{i=0}^{N} b_i < \epsilon$. □

4.3.2 Proofs of Theorem 4.9

Assume that

$$\bigcap_{n=0}^{\infty} I(n) = \{p\}\,.$$

The case that $\bigcap_{n=0}^{\infty} I(n)$ is an interval follows from the Birkhoff ergodic theorem. We will consider partition elements $\Lambda_{k,l}$ of the tower such that $\pi^{-1}(p) \cap \Lambda_{k,l} \neq \emptyset$. Since $\pi^*\nu = \mu$, these are the only partition elements we need to consider to determine whether $F^n(x,0) \in \pi^{-1}(I(n))$ infinitely often (and hence whether $T^n(x) \in I(n)$ infinitely often). For all k and l such that $\pi^{-1}(p) \cap \Lambda_{k,l} \neq \emptyset$, let $\hat{p}_{k,l}$ denote the point of intersection. We assume that $\hat{p}_{k,l} \in \mathrm{int}(\Lambda_{k,l})$ for all k and l for which $\hat{p}_{k,l}$ exists.

We first consider the sequence $\Lambda_{k,l} \cap \pi^{-1}(I(n))$ for a fixed partition element $\Lambda_{k,l}$. For $n \in \mathbb{Z}^+$ define

$$A'_n = \Lambda_{k,l} \cap \pi^{-1}\left(I(n\langle R\rangle)\right), \qquad G'_n = F^{-l}(A'_n)$$

and for $n \in \mathbb{N}$ let

$$\alpha(n) = \sum_{j=0}^{n-1} m(G'_j).$$

For $x \in \Lambda_0$ and $n \in \mathbb{N}$, define

$$\hat{S}(n,x) = \#\left\{ j < n : F^j(x,0) \in \Lambda_{k,l} \cap \pi^{-1}(I(j)) \right\}.$$

Step 1. We relate the recurrence properties of F to those of $f := T^R$. We claim that

$$\lim_{n\to\infty} \frac{\hat{S}(n,x)}{\alpha(\lfloor n\langle R\rangle^{-1} \rfloor)} = 1.$$

for m a.e. $x \in \Lambda_0$. As the proof of the claim proceeds, we will place finitely many restrictions on x. Each of these restrictions will be satisfied by a set of full measure. First, assume that x satisfies (4.3.8). Ergodicity relates the clock associated with the tower map F to the clock associated with the return map f. For small $\varepsilon \in \mathbb{R}$, define the sets

$$A_{n,\varepsilon} = \Lambda_{k,l} \cap \pi^{-1}(I(n(\langle R\rangle + \varepsilon))), \qquad G_{n,\varepsilon} = F^{-l}(A_{n,\varepsilon})$$

and the sums

$$\mathcal{S}_\varepsilon(n,x) = \sum_{j=0}^{n-1} \mathbf{1}_{G_{j,\varepsilon}} \circ f^j(x), \qquad \mathcal{E}_\varepsilon(n) = \sum_{j=0}^{n-1} m(G_{j,\varepsilon}).$$

For sequences (u_n) and (v_n) in $\mathbb{R}$, we write $u_n \approx v_n$ if

$$\lim_{n\to\infty} \frac{u_n}{v_n} = 1.$$

Since $\hat{p}_{k,l} \in \operatorname{int}(\Lambda_{k,l})$, we have $\mathcal{E}_\varepsilon(n) \to \infty$ as $n \to \infty$. Theorem 4.7 gives the sBC property for the base transformation: for m a.e. $x \in \Lambda_0$, we have

$$\lim_{n\to\infty} \frac{\mathcal{S}_\varepsilon(n,x)}{\mathcal{E}_\varepsilon(n)} = 1. \tag{4.3.9}$$

We now examine $\hat{S}(n,x)$. Define $q(n,x)$ by

$$R_{q(n,x)}(x) \leqslant n < R_{q(n,x)+1}(x).$$

Observe that $\hat{S}(n,x) - \hat{S}(R_{q(n,x)}, x) \in \{0,1\}$. This is so because the levels of the Young tower are pairwise disjoint and so the orbit of $(x,0)$ must enter $\Lambda_{k,l}$ in order to increment $\hat{S}$ and this can happen at most once from time $R_{q(n,x)}$ to time n. Thus it suffices to examine $\hat{S}(R_{q(n,x)}, x)$. Using (4.3.8), for $\varepsilon > 0$ small we obtain

$$\begin{aligned} &\left(\frac{\mathcal{S}_\varepsilon(q(n,x),x) + \psi(x,\varepsilon)}{\mathcal{E}_\varepsilon(q(n,x))}\right)\left(\frac{\mathcal{E}_\varepsilon(q(n,x))}{\alpha(q(n,x))}\right) \\ \leqslant\ & \frac{\hat{S}(R_{q(n,x)},x)}{\alpha(q(n,x))} \\ \leqslant\ & \left(\frac{\mathcal{S}_{-\varepsilon}(q(n,x),x) + \zeta(x,\varepsilon)}{\mathcal{E}_{-\varepsilon}(q(n,x))}\right)\left(\frac{\mathcal{E}_{-\varepsilon}(q(n,x))}{\alpha(q(n,x))}\right). \end{aligned}$$

where ψ and ζ are independent of n. Using (4.3.9) and Lemma 4.18, this implies

$$\lim_{n\to\infty} \frac{\hat{S}(n,x)}{\alpha(q(n,x))} = 1$$

and therefore another application of Lemma 4.18 gives

$$\lim_{n\to\infty} \frac{\hat{S}(n,x)}{\alpha(\lfloor n\langle R\rangle^{-1}\rfloor)} = 1$$

since $q(n,x) \approx n\langle R\rangle^{-1}$.

Step 2. We claim that

$$\alpha(\lfloor n\langle R\rangle^{-1}\rfloor) \approx \sum_{j=0}^{n-1} \nu(\hat{I}_{k,l}(j))$$

where

$$\hat{I}_{k,l}(j) = \Lambda_{k,l} \cap \pi^{-1}(I(j)).$$

This follows from a change of variable argument. Let $\tilde{p}_{k,l} \in \Lambda_k$ satisfy $T^l(\tilde{p}_{k,l}) = p$. Defining $\rho = \frac{dm}{d\lambda}$, we have

$$\begin{aligned}
\alpha(\lfloor n\langle R\rangle^{-1}\rfloor) = \sum_{j=0}^{\lfloor n\langle R\rangle^{-1}\rfloor-1} m(G'_j) &\approx \frac{\rho(p)}{|DT^l(\tilde{p}_{k,l})|} \sum_{j=0}^{\lfloor n\langle R\rangle^{-1}\rfloor-1} g(j\langle R\rangle) \\
&\approx \frac{\rho(p)}{|DT^l(\tilde{p}_{k,l})|} \int_0^{n\langle R\rangle^{-1}} g(t\langle R\rangle)\,dt \\
&= \frac{\rho(p)}{\langle R\rangle\,|DT^l(\tilde{p}_{k,l})|} \int_0^n g(u)\,du \\
&\approx \frac{\rho(p)}{\langle R\rangle\,|DT^l(\tilde{p}_{k,l})|} \sum_{j=0}^{n-1} g(j) \\
&\approx \sum_{j=0}^{n-1} \nu(\hat{I}_{k,l}(j)).
\end{aligned}$$

Steps (1) and (2) imply that

$$\hat{S}(n,x) \approx \sum_{j=0}^{n-1} \nu(\hat{I}_{k,l}(j)).$$

Step 3. We now study the sequence of preimages $(\pi^{-1}(I(n))_{n=0}^{\infty}$ on the whole tower Δ. By definition,

$$\frac{d\,\mu}{d\,\lambda}(p) = \sum_{\hat{p}_{l,k}\in\pi^{-1}(p)} \frac{d\,\nu}{d\,\lambda}(\hat{p}_{l,k}).$$

Consequently, for every $\delta > 0$, there exists $N(\delta)$ such that the truncated tower $\Delta_{N(\delta)} := \{\Lambda_{k,l} : k \leqslant N(\delta) \text{ and } l < \min\{N(\delta), L_k\}\}$ satisfies

$$\mu(I(n) \cap \pi(\Delta_{N(\delta)})) \geqslant (1-\delta)\mu(I(n))$$

for all n sufficiently large. Now fix $\delta > 0$. Repeat steps (1) and (2) for every $\hat{p}_{k,l} \in \pi^{-1}(p)$. For m a.e. $x \in \Lambda_0$, we have

$$\#\left\{0 \leqslant j < n : F^j(x,0) \in \hat{I}_{k,l}(j)\right\} \approx \sum_{j=0}^{n-1} \nu(\hat{I}_{k,l}(j)) \tag{4.3.10}$$

for every k and l such that $\Delta_{N(\delta)} \cap \pi^{-1}(p) \neq \emptyset$. For $x \in \Lambda_0$, define

$$U(n,x) = \#\left\{0 \leqslant j < n : F^j(x,0) \in \bigcup_{\hat{I}_{k,l}(j) \subset \Delta_{N(\delta)}} \hat{I}_{k,l}(j)\right\}.$$

Estimate (4.3.10) and the fact that the levels of the tower are pairwise disjoint imply the existence of $\kappa(\delta)$ satisfying $1 - \delta \leqslant \kappa(\delta) \leqslant 1$ for which m a.e. $x \in \Lambda_0$ satisfies

$$U(n,x) \approx \kappa(\delta) \sum_{j=0}^{n-1} \mu(I(j))$$

for every $\delta > 0$. Define

$$V(n,x) = \#\left\{0 \leqslant j < n : F^j(x,0) \in \bigcup_{\hat{I}_{k,l}(j) \in \Delta} \hat{I}_{k,l}(j)\right\}.$$

Since $\delta > 0$ is arbitrary, we conclude that for m a.e. $x \in \Lambda_0$, we have

$$\underline{\lim}_{n\to\infty} \frac{V(n,x)}{\sum_{j=0}^{n-1} \mu(I(j))} \geqslant 1$$

and therefore

$$\underline{\lim}_{n\to\infty} \frac{S_n(x)}{\sum_{j=0}^{n-1} \mu(I(j))} \geqslant 1 \tag{4.3.11}$$

where

$$S_n(x) = \#\left\{0 \leqslant j < n : T^j(x) \in I(j)\right\}.$$

The arguments in steps (1)–(3) extend to ν a.e. element of every level of the tower Δ. Relating the dynamics on the tower to those on M, since $\pi \circ F = T \circ \pi$ and $\pi_* \nu = \mu$, we conclude that (4.3.11) holds for μ a.e. $x \in M$. This concludes the proof of Theorem 4.9.

4.3.3 Proof of Theorem 4.12

Since $p \in \pi(\Delta)$ we may consider a single preimage $\hat{p}_{k,l} \in \pi^{-1}(p)$ and note that $J(n) = \pi^{-1}(I(n)) \cap \Lambda_{k,l}$ has the property that $\sum_n \nu(J_n) = \infty$. We then use the same argument to the one given in steps (1) and (2) of the proof of Theorem 4.9 to obtain the dBC property.

4.4 Almost-sure Results in EVT for Dynamical Systems

Including the results in Chapters 2 and 3, there has been much recent work on the extreme value theory of deterministic dynamical systems [21, 24, 27, 27, 28, 36, 37, 38, 41]. If $\phi : X \to \mathbb{R}$ is an integrable observable on a dynamical system (X, μ, T), we define $\phi_j(x) = \phi \circ T^j(x)$ and in turn define (M_n), the sequence of successive maxima, by

$$M_n(x) = \max_{0 \leqslant j \leqslant n-1} \phi_j(x).$$

Much recent research on the extreme value theory of non-uniformly hyperbolic dynamical systems [21, 27, 28, 37] has concentrated on the study of distributional limits of the sequence (M_n). There, the goal is to find scaling constants $a_n > 0$ and $b_n \in \mathbb{R}$ and a non-degenerate distribution $G(x)$ such that

$$\lim_{n\to\infty} \mu(a_n(M_n - b_n) \leqslant x) = G(x).$$

Many examples of such distributional convergences are given in Chapter 2 and Chapter 3.

The Borel-Cantelli results allow a description of the almost-sure behavior of the sequence of successive maxima $M_n(x)$, rather than just a distributional description. This allows an estimation of almost-sure upper bounds. In a similar way the law of the iterated logarithm gives an almost-sure upper bound for the rate of growth of scaled Birkhoff sums

$$b_n(x) = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \phi(T^j(x),$$

namely

$$\overline{\lim}_{n\to\infty} \frac{b_n(x)}{(n\log(\log(n)))^{1/2}} < C$$

almost surely, in contrast to the central limit theorem, which is a distributional result.

Extreme value theory is related to entrance times to nested balls by the observation that if $\phi(x) = -\log(d(x, x_0))$, $M_n(x) < \log(n) + v$ if and only if $d(T^i(x), x_0) > \frac{e^{-v}}{n}$ for $0 \leqslant i \leqslant n-1$. In the context of Theorem 4.9, for μ a.e. center x_0 we have

$\mu(M_n > v + \log(n)\text{i.o.}) = 1$ since

$$E_n = \sum_{i=0}^{n-1} \mu\left(B\left(x_0, \frac{e^{-v}}{i}\right)\right)$$

diverges. By contrast, if $\delta > 1$ then

$$\mu(M_n(x) > v + \log(n) + \delta \log(\log(n))\text{i.o.}) = 0$$

by the classical Borel-Cantelli lemma since

$$\sum_{i=0}^{n-1} \mu\left(B\left(x_0, \frac{e^{-v}}{i(\log(i))^{1+\delta}}\right)\right)$$

converges. Thus the sequence $u_n = \log(n) + \delta \log(\log(n))$ is an almost-sure upper bound for M_n for any $\delta > 1$. The consideration of the function $\phi(x) = -\log(d(x, x_0))$ is not restrictive; other functions with unique maxima can be considered in this framework and almost-sure upper bounds u_n for the sequence M_n can be derived, though the sequence u_n will depend upon the form of the function near the maximal point.

Bibliography

[1] Jon Aaronson. *An introduction to infinite ergodic theory*, volume 50 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.

[2] J Alves and M Viana. Statistical stability for robust classes of maps with non-uniform expansion. *Ergodic Theory and Dynamical Systems*, 22:1–32, 2002.

[3] José F. Alves. Stochastic behavior of asymptotically expanding maps. *Discrete Contin. Dyn. Syst.*, Added Volume:14–21, 2001.

[4] José F. Alves, Stefano Luzzatto, and Vilton Pinheiro. Lyapunov exponents and rates of mixing for one-dimensional maps. *Ergodic Theory and Dynamical Systems*, 24(3):637–657, 2004.

[5] José F. Alves, Stefano Luzzatto, and Vilton Pinheiro. Markov structures and decay of correlations for non-uniformly expanding dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 22(6):817–839, 2005.

[6] D.V. Anosov. *Geodesic flows on closed Riemann manifold with negative curvature*, volume 90 of Trundy ordena Lenina Mathematicheskogo Instituta imeni V.A. Steklova. American Mathematical Society, 1969.

[7] D.V. Anosov. *Dynamical systems: Dynamical systems with hyperbolic behavior*, volume 9 of *Dynamical Systems*. Springer, 1995.

[8] V. Araújo and M.J. Pacifico. Large deviations for non-uniformly expanding maps. *Journal of Statistical Physics*, 125:411–453, 2006.

[9] Viviane Baladi. *Positive transfer operators and decay of correlations*, volume 16 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[10] L Barreira, Y Pesin, and J Schmeling. Dimension and product structure of hyperbolic measures. *Annals of Mathematics-Second Series*, 149(3):755–784, 1999.

[11] L. Barreira and Y.B. Pesin. *Lyapunov exponents and smooth ergodic theory*, volume 23 of *University Lecture Series*. AMS Bookstore, 2002.

[12] S. Brandon. Sub-actions for young towers. *Discrete and Continuous Dynamical Systems*, 22(3):541–556, 2008.

[13] M. Brin and G. Stuck. *Introduction to dynamical systems*. Cambridge University Press, 2002.

[14] Henk Bruin, Stefano Luzzatto, and Sebastian Van Strien. Decay of correlations in one-dimensional dynamics. *Ann. Sci. École Norm. Sup. (4)*, 36(4):621–646, 2003.

[15] Henk Bruin and Mike Todd. Return time statistics of invariant measures for interval maps with positive Lyapunov exponent. *Stoch. Dyn.*, 9(1):81–100, 2009.

[16] L.A. Bunimovich, N. Chernov, and Sinai Ya.G. Markov partitions for two-dimensional hyperbolic biliards. *Uspekhi Mat. Nauk*, 45(3(273)):97–134, 1990.

[17] L.A. Bunimovich, N. Chernov, and Sinai Ya.G. Statistical properties of two-dimensional hyperbolic billiards. *Uspekhi Mat. Nauk*, 46(4(280)):43–92, 1991.

[18] M.L. Cartwright and J.E. Littlewood. On nonlinear differential equations of the second order, i: the equation $\ddot{y} + k(1 - y^2)\dot{y} + y = b\lambda k \cos(\lambda t + a), k$ large. *J. London Math. Soc.*, 20:180–189, 1945.

[19] N. Chernov. Decay of correlations and dispersing billiards. *J. Statist. Phys.*, 94(3-4):513–556, 1999.

[20] N. Chernov and D. Kleinbock. Dynamical Borel-Cantelli lemmas for Gibbs measures. *Israel J. Math.*, 122:1–27, 2001.

[21] P Collet. Statistics of closest return for some non-uniformly hyperbolic systems. *Ergodic Theory and Dynamical Systems*, 21(02):401–420, 2001.

[22] P Collet and Y Levy. Ergodic properties of the Lozi mappings. *Communications in Mathematical Physics*, 93(4):461–481, 1984.

[23] W.J. Cowieson. Stochastic stability for piecewise expanding maps in $\mathbb{R}^d$. *Nonlinearity*, 13:1745–1760, 2000.

[24] D Dolgopyat. Limit theorems for partially hyperbolic systems. *Transactions-American Mathematical Society*, 356(4):1637–1689, 2004.

[25] Richard Durret. *Probability: theory and examples.* Duxbury Press, Belmont, CA, second edition, 1996.

[26] AC Moreira Freitas and JM Freitas. Extreme values for Benedicks–Carleson quadratic maps. *Ergodic Theory and Dynamical Systems*, 28(04):1117–1133, 2008.

[27] ACM Freitas and JM Freitas. On the link between dependence and independence in extreme value theory for dynamical systems. *Statistics and Probability Letters*, 78(9):1088–1093, 2008.

[28] ACM Freitas, JM Freitas, and M Todd. Hitting time statistics and extreme value theory. *Probability Theory and Related Fields*, 147(3-4):675–710, 2009.

[29] Janos Galambos. *The asymptotic theory of extreme order statistics.* Robert E. Krieger Publishing Co. Inc., Melbourne, FL, second edition, 1987.

[30] P. Góra and A. Boyarsky. Absolutely continuous invariant measures for piecewise expanding C^2 transformation in $\mathbf{R}^N$. *Israel J. Math.*, 67(3):272–286, 1989.

[31] S Gouëzel. *Personal Communication.*

[32] S Gouëzel. Statistical properties of a skew product with a curve of neutral points. *Ergodic Theory and Dynamical Systems*, 27(01):123–151, 2007.

[33] Sébastien Gouëzel. Decay of correlations for nonuniformly expanding systems. *Bull. Soc. Math. France*, 134(1):1–31, 2006.

[34] Sébastien Gouëzel. A Borel-Cantelli lemma for intermittent interval maps. *Nonlinearity*, 20(6):1491–1497, 2007.

[35] Sébastien Gouezel. Local limit theorem for nonuniformly partially hyperbolic skew-products and Farey sequences. *Duke Math. J.*, 147(2):193–284, Apr 2009.

[36] Chinmaya Gupta. Extreme value distributions for some classes of non-uniformly partially hyperbolic dynamical systems. *Ergodic Theory and Dynamical Systems*, 30:757–771, 2010.

[37] Chinmaya Gupta, Mark Holland, and Matthew Nicol. Extreme value theory for hyperbolic billiards, Lozi-like maps, and Lorenz-like maps. preprint http://www.math.uh.edu/∼ccgupta/, 2009.

[38] Masaki Hirata. Poisson law for Axiom A diffeomorphisms. *Ergodic Theory Dynam. Systems*, 13(3):533–556, 1993.

[39] M.W. Hirsch, C.C. Pugh, and M. Shub. *Invariant manifolds*, volume 583 of *Lecture Notes in Mathematics*. Springer-Verlag, 1977.

[40] Franz Hofbauer and Gerhard Keller. Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z.*, 180(1):119–140, 1982.

[41] M Holland, M Nicol, and A Török. Extreme value theory for non-uniformly expanding dynamical systems. *math.uh.edu*.

[42] E. Hopf. Abzweigung einer periodischen l osung von einer station aren l osung eines differential-systems. *Berichte Math.-Phys. Kl. S achs. Akad. Wiss. Leipzig. Math.-Nat. Kl.*, 94:1–22, 1942.

[43] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of mathematics and its applications*. Cambridge University Press, 1996.

[44] Anatole Katok, Jean-Marie Strelcyn, F. Ledrappier, and F. Przytycki. *Invariant manifolds, entropy and billiards; smooth maps with singularities*, volume 1222 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986.

[45] Dong Han Kim. The dynamical Borel-Cantelli lemma for interval maps. *Discrete Contin. Dyn. Syst.*, 17(4):891–900, 2007.

[46] D. Y. Kleinbock and G. A. Margulis. Logarithm laws for flows on homogeneous spaces. *Invent. Math.*, 138(3):451–494, 1999.

[47] M. R. Leadbetter, Georg Lindgren, and Holger Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York, 1983.

[48] François Ledrappier. Dimension of invariant measures. In *Proceedings of the conference on ergodic theory and related topics, II (Georgenthal, 1986)*, volume 94 of *Teubner-Texte Math.*, pages 116–124, Leipzig, 1987. Teubner.

[49] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems*, 19(3):671–685, 1999.

[50] Carlangelo Liverani and Maciej P. Wojtkowski. Ergodicity in Hamiltonian systems. In *Dynamics reported*, volume 4 of *Dynam. Report. Expositions Dynam. Systems (N.S.)*, pages 130–202. Springer, Berlin, 1995.

[51] EN Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20:130–141, 1963.

[52] S Luzzatto, I Melbourne, and F Paccaut. The Lorenz attractor is mixing. *Communications in Mathematical Physics*, Jan 2005.

[53] Pertti Mattila. *Geometry of sets and measures in Euclidean spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.

[54] François Maucourant. Dynamical Borel-Cantelli lemma for hyperbolic spaces. *Israel J. Math.*, 152:143–155, 2006.

[55] I Melbourne and M Nicol. Large deviations for nonuniformly hyperbolic systems. *Transactions-American Mathematical Society*, 360(12):6661–6676, 2008.

[56] I. Melbourne and A. Torok. Central limit theorems and invariance principles for time-one maps of hyperbolic flows. *Communications in Mathematical Physics*, 229(1):57–71, 2002.

[57] Michał Misiurewicz. Strange attractors for the Lozi mappings. In *Nonlinear dynamics (Internat. Conf., New York, 1979)*, volume 357 of *Ann. New York Acad. Sci.*, pages 348–358. New York Acad. Sci., New York, 1980.

[58] T Palmer. Edward Norton Lorenz. *Physics Today*, 61(9):81, 2008.

[59] Y.B. Pesin. *Lectures on partial hyperbolicity and stable ergodicity*. European Mathematical Society, 2004.

[60] S. Pilyugin. *Shadowing in dynamical systems*, volume 1706 of *Lecture Notes in Mathematics*. Springer, 1999.

[61] H.J. Poincaré. *Les méthodes nouvelles de la mécanique céleste*, volume 1-3. Gauthiers-Villars, 1892, 1893, 1899.

[62] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Commun. Math. Phys.*, 74:189–197, 1980.

[63] Marek Rychlik. Bounded variation and invariant measures. *Studia Math.*, 76(1):69–80, 1983.

[64] S. Smale. Diffeomorphisms with many periodic points. In *Differential and combinatorial topology: A symposium in honor of Marston Morse*, pages 63–70. Princeton University Press, Princeton, NJ, 1965.

[65] S. Smale. Differential dynamical systems. *Bull. Soc. Math. Soc.*, 73:747–817, 1967.

[66] Vladimir G. Sprindžuk. *Metric theory of Diophantine approximations.* V. H. Winston & Sons, Washington, D.C., 1979. Translated from the Russian and edited by Richard A. Silverman, With a foreword by Donald J. Newman, Scripta Series in Mathematics.

[67] W Tucker. The Lorenz attractor exists. *Comptes Rendus de l'Academie des Sciences Series I Mathematics*, 328(12):1197–1202, 1999.

[68] Warwick Tucker. A rigorous ODE solver and Smale's 14th problem. *Found. Comput. Math.*, 2(1):53–117, 2002.

[69] B. van der Pol and B. van der Mark. Frequency demultiplication. *Nature*, 120:363–364, 1927.

[70] Marcelo Viana. Stochastic dynamics of deterministic systems. Lecture Notes, XXI Bras. Math. Colloq. IMPA, Rio de Janeiro, 1997.

[71] R Williams. The structure of Lorenz attractors. *Publications Mathematiques de l'IHES*, Jan 1979.

[72] Sinai Ya.G. *Introduction to ergodic theory*, volume 18 of *Mathematical Notes.* Princeton University Press, 1977.

[73] L.-S. Young and L. Ruy-Bellet. Large deviations in non-uniformly hyperbolic dynamical systems. *Ergodic Theory and Dynamical Systems*, 28(02):587–612, 2008.

[74] Lai-Sang Young. Bowen-ruelle measuers for certain piecewise hyperbolic maps. *Transactions-American Mathematical Society*, 287(1):41–48, Jan 1985.

[75] Lai-Sang Young. Statistical properties of dynamical systems with some hyperbolicity. *The Annals of Mathematics Second Series*, 147(3):585–650, May 1998.

[76] Lai-Sang Young. Recurrence times and rates of mixing. *Israel J. Math.*, 110:153–188, 1999.