STATISTICAL PROPERTIES OF CHAOTIC DYNAMICAL SYSTEMS: NON-STATIONARY CENTRAL LIMIT THEOREMS AND EXTREME VALUE THEORY

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

Licheng Zhang

May 2015

STATISTICAL PROPERTIES OF CHAOTIC DYNAMICAL SYSTEMS: NON-STATIONARY CENTRAL LIMIT THEOREMS AND EXTREME VALUE THEORY

Licheng Zhang

APPROVED:

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

Dr. Hongkun Zhang
Department of Mathematics, University of Houston

Dr. Andrew Török
Department of Mathematics, University of Houston

Dr. Gemunu Gunaratne
Department of Physics, University of Houston

Dean, College of Natural Sciences and Mathematics

Acknowledgements

My special thanks go to my thesis advisor, Professor Matthew J. Nicol, who has shown great guidance, extreme patience, generosity, humor, and insight: he continually and persuasively conveyed a spirit of adventure in regard to research and scholarship, and an excitement in regard to teaching. Without his supervision and constant help this dissertation would not have been possible.

My sincerest appreciation goes to my committee members, Dr. Hongkun Zhang, Dr. Andrew Török, and Dr. Gemunu Gunaratne, for spending precious time on reading through my thesis, suggesting improvements, future projects, and pointing out mistakes;

Many thanks go to Dr. Mark Holland and Dr. Sandro Vaienti for reading through my paper and providing valuable suggestions for my research.

I want to thank Dr. Renato Feres for inviting me to speak at the AMS Sectional Meeting AMS Special Session on Statistical Properties of Dynamical Systems at St. Louis in 2013.

I would also like to thank all my friends for being supportive all the time and encouraging me with their best wishes.

Last, but certainly not least, I would like to thank my family: my parents Jian Zhang and Yulan Chen, 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: NON-STATIONARY CENTRAL LIMIT THEOREMS AND EXTREME VALUE THEORY

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

Licheng Zhang

May 2015

Abstract

In this thesis, some statistical properties of two interesting problems are studied.

The first one is about non-stationary central limit theorems. We establish central limit theorems for a sequence of nested balls using martingale difference array technique, under certain conditions. It applies to various dynamical systems, including smooth expanding maps of the interval, Rychlik type maps, Gibbs-Markov maps, rational maps, and piecewise expanding maps in higher dimension.

And the second problem is about the Lorenz Systems. We study a family of Geometrical Lorenz Models, which have very similar properties to Lorenz Systems and are easier to study. Based on the model, we establish dynamical Borel Cantelli lemmas and convergence of rare event points processes to Poisson processes, which implies Extreme Value Theory.

Contents

Chapter 1

Introduction

The theory of dynamical systems is one of the most important part of mathematics. It has many applications, like in biomechanics, cognitive science, and human development. It studies the global orbit structure of maps (discrete-time case) and flows (continuous-time case) with emphasis on some invariant properties. While discrete dynamical systems are usually charactered iterations of maps, continuous dynamical systems are usually charactered by differential equations. For a simple dynamical system, knowing the trajectory is often sufficient; but for some complicated deterministic dynamical systems, there exhibit chaotic phenomena and unpredictable behaviors so the dynamical systems are too complicated to be understood just in terms of individual trajectories and in such chaotic systems, the future behavior of the systems are too sensitive to the initial conditions. Many problems in science and engineering are often reduced to studying the asymptotic behavior of discrete dynamical systems, like condensed-matter physics, convection-diffusion equations, and molecular dynamics, whose asymptotic property exhibit complex chaotic phenomena.

So a statistical description of the system's behavior is often the most appropriate object to study. The tools needed to study chaotic dynamical systems are analytic and measure-theoretic rather than geometric. We can study the chaotic systems from the statistical point of view, instead of viewing the long-term outcomes themselves, and chaos in the deterministic sense usually show some kind of regularity in the probability sense. We may investigate whether suitable versions of classical limit theorems from probability theory such as law of large numbers, central limit theorem, Borel-Cantelli lemma, extreme value theory and so on, hold and use this knowledge to make predictions about the system's behavior.

Mathematically, a dynamical system is a space Ω with a transformation T from Ω to itself, and we use the pair (X, T) to denote it. For a point or an initial state x in Ω, Tx represents the state at time 1 of a system, T^2x represents the state at time 2 of the system, and so on. If we consider a continuous variable for the time, we use a one-parameter family $\{T_t : t \in \mathbb{R}\}$ of maps of Ω into itself, with the property that $T_{s+t} = T_s T_t$ so that $\{T_t : t \in \mathbb{R}\}$ is a *flow*. Of course, one is more interested in the cases that when Ω is equipped with some structures and T has some restrictions. There are three major cases:

1. **Topological Dynamics**: if Ω is a topological space and T is a homeomorphism.

2. **Differentiable Dynamics:**: if Ω is a differentiable manifold and T is a diffeomorphism.

3. **Ergodic Theory**: if Ω is a measure space and T is a measure-preserving transformation.

Speaking of Ω being a measure space, we need introduce the concepts of σ-algebra and measure. Let Ω be a compact metric space.

Definition 1.0.1. *A family $\mathscr{B}$ of subsets of Ω is called a* σ**-algebra** *if the following are satisfied:*

1) $\Omega \in \mathscr{B}$*;*

2) for any $A \in \mathscr{B}$, $A^c \in \mathscr{B}$*,where* $A^c = \{x \in \Omega : x \notin A\}$*;*

3) if $A_n \in \mathscr{B}$*, for* $n = 1, 2, \ldots,$ *then* $\cup_{n=1}^{\infty} A_n \in \mathscr{B}$*.*

Definition 1.0.2. *A set function* $\mu : \mathscr{B} \to [0, \infty)$ *is called a* **measure** *on* $\mathscr{B}$ *if it satisfies:*

1) $\mu(\emptyset) = 0$*;*

2) for any sequence $\{A_n\}$ *of disjoint measurable sets,* $A_n \in \mathscr{B}$*,* $n = 1, 2, \ldots,$

$$\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

The triple $(\Omega, \mathscr{B}, \mu)$ is called a measure space and if $\mu(\Omega) = 1$, it is called a *normalized measure space* or *probability space*. If Ω is a countable union of sets of finite μ-measure, then we say Ω is a σ-finite measure space. In the rest of this thesis, we shall work with probability space.

When the map T preserves the measure μ, i.e., $\forall A \in \mathscr{B}$, $\mu(T^{-1}(A)) = \mu(A)$, we call T is a measure-preserving transformation(m.p.t.), and the corresponding dynamical system $(\Omega, \mathscr{B}, T, \mu)$ is called a measure-preserving dynamical system. For such μ, we call it a T-invariant measure, or invariant measure if we don't emphasize the transformation . Ergodic theory is the study of invariant measures in dynamical systems, which we will talk about more in section 1.1.

Remark 1.0.3. *Since measure-preserving is a very important property of dynamical systems, in the rest of the thesis, without special declaration, all the systems to be considered are measure-preserving dynamical systems, with some invariant measures.*

1.1 Ergodic Theory

Ergodic theory is a branch of mathematics that studies dynamical systems with an invariant measure and the long-term average behavior of systems. Its initial development was motivated by problems of statistical physics.

A primal concern of ergodic theory is the asymptotic behavior of a dynamical system $(\Omega, \mathscr{B}, T, \mu)$ when it is allowed to run for a long time. If $T : \Omega \to \Omega$ is a m.p.t., the n-iterate of T is defined by T^n, i.e. $T^n(x) = T \circ \cdots \circ T(x)$. The orbit $\{T^n x : n \in \mathbb{Z}\}$ of a point $x \in \Omega$ represents a single complete history of the system, from the infinite past to the infinite future. The σ-algebra $\mathscr{B}$ is thought of as the family of observable events, with the T-invariant measure μ specifying the time-independent probability of their occurrences. One powerful result regarding the recurrence of orbits is the following famous *Poincaré Recurrence Theorem* (1899):

Theorem 1.1.1. *Let T be a m.p.t on a probability space $(\Omega, \mathscr{B}, \mu)$. Let $E \in \mathscr{B}$ such that $\mu(E) > 0$. Then almost all points of E return infinitely often to E under iterations of T.*

We now introduce the concept of ergodicity:

Definition 1.1.2. *A m.p.t $T : (\Omega, \mathscr{B}, \mu) \to (\Omega, \mathscr{B}, \mu)$ is called* **ergodic** *if for any $A \in \mathscr{B}$, such that $T^{-1}A = A$, then $\mu(A) = 0$ or $\mu(A^c) = 0$.*

An equivalent definition of ergodicity is by using measurable functions: *If f is measurable and $(f \circ T)(x) = f(x)$ a.e. then f is constant a.e. $\Longleftrightarrow T$ is ergodic.*

Definition 1.1.3. *The transformation $T : \Omega \to \Omega$ is said to be* **mixing** *if for any two measurable sets $A, B \subset \Omega$,*

$$\mu(T^{-n} \cap B) \to \mu(A)\mu(B) \quad \text{as } n \to \infty$$

Remark 1.1.4. *The mixing property implies ergodicity.*

For ergodic transformations, we have a stronger version of the Poincaré Recurrence Theorem, known as Kac's Lemma. Let A be a measurable set with $\mu(A) > 0$ and we define, for $x \in A$, the first hitting time of x to the set A,

$$n_A(x) = \min\{k \geq 1 : T^k(x) \in A\}$$

Theorem 1.1.5 (Kac's Lemma, [5] Theorem 3.2.4)**.** *If μ is T-invariant and A is measurable with $\mu(A) > 0$, then*

$$\int_A n_A(x) d\mu(x) = 1.$$

In terms of the conditional measure $\mu_A(B) = \frac{\mu(A\cap B)}{\mu(A)}$, this result can be written as

$$\int_A n_A(x) d\mu_A(x) = \frac{1}{\mu(A)},$$

which says that the expected time of return to a set A is $\frac{1}{\mu(A)}$.

Now we consider an observable or a measurement made on the systems, for which we use a measurable function $f : \Omega \to \mathbb{R}$ to denote it, and $f(x), f(Tx), f(T^2x), \ldots$ may be thought of as the values of some physically interesting variable at successive moments of time, beginning in initial state x. In statistical mechanics, information theory, and other areas of application it is interesting and sometimes necessary to consider the long-term time average

$$\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i(x)$$

of a large number n of successive observations. A basic question of ergodic theory is that of the *convergence of these averages*: when does

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

exist in some sense? If it exists, the limit may be thought of as an equilibrium or central value of the variable f. The convergence had been proved in special cases earlier (e.g., Borel's Strong Law of Large Number(1909)) in the case the $f \circ T^i$ are independent and identically distributed (i.i.d.), which is a very strong assumption because in most cases they are not independent at all. But the general convergence in the mean square (L^2) sense was proved by von Neumann (1931) and the almost everywhere convergence by Birkhoff (1931). Birkhoff's Ergodic Theorem (1931) is one of the most important theorems in ergodic theory, which can be stated as the following:

Theorem 1.1.6 (Birkhoff Ergodic Theorem)**.** *Let $(\Omega, \mathscr{B}, \mu)$ be a probability space, $T : \Omega \to \Omega$ a m.p.t and $f \in L^1(\Omega, \mathscr{B}, \mu)$. Then*

(1) $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x) = f^*(x)$ exists a.e.;

(2) $f^* \circ T = f^*$;

(3) $f^* \in L^1$ and in fact $||f^*||_1 \leq ||f||_1$;

(4) If T is ergodic, then $f^* = \int f d\mu$.

In physics, $\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x)$ is called time average and $\int f d\mu$ is usually called space average while (1) and (4) together is what physicists like to call 'ergodic hypothesis'. It is desirable that the equilibrium or central value of a physical variable coincide with its weighted average over all possible states of the system. So an alternative way to define a system (or just T itself) to be ergodic, if the time average of every measurable function coincides almost everywhere with its space average.

In probability theory this property that $\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x) \to \int f d\mu$ is known as **Strong Law of Large Numbers (SLLN)**. The first question after the convergence is ensured is that how fast the time average converges to the space average. This is where decay of correlation comes into play. We define the **correlations** for two square integrable functions f and g as

$$C_{f,g}(n) = \frac{1}{||f||_2 \cdot ||g||_2} \left[\int f \cdot g \circ T^n d\mu - \int f d\mu \int g d\mu \right]$$

for any n. When $g = f$, we call it **autocorrelation**. The mixing property is related to correlations, that is, T is mixing if and only if the correlations decay, i.e. for every pair f and $g \in \mathscr{L}^2$,

$$C_{f,g}(n) \to 0 \quad \text{as } n \to \infty$$

The rate of the decay of correlations, which is also called the rate of mixing, is very crucial when we deal with particular observables. Specially, in Chapters 2, 3, 4 and 5, fast decay of correlations plays an important role in establishing several statistical properties, which will be introduced in the following sections.

1.2 Borel-Cantelli Lemmas

In probability theory and dynamical systems, the classical Borel-Cantelli lemmas are a powerful tool used to establish the almost-sure behavior of random variables. They are lemmas about sequences of events, which state that, under certain conditions, an event will occur with probability zero or with probability one.

Suppose $(\Omega, \mathscr{B}, \mathbb{P})$ is a probability space. Let $\mathbf{1}_A$ be the characteristic function of A, given A is a measurable set of Ω. Then the classical Borel-Cantelli lemmas are as follows(see [10] for proofs):

1. **First Borel-Cantelli lemma(BC1):** If $(A_n)_{n=0}^{\infty}$ is a sequence of measurable sets in Ω and $\sum_{n=0}^{\infty} \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(x \in A_n \ i.o.) = 0$ ($i.o.$ means "infinitely often")

2. **Second Borel-Cantelli lemma(BC2):** If $(A_n)_{n=0}^{\infty}$ is a sequence of independent sets in Ω and $\sum_{n=0}^{\infty} \mathbb{P}(A_n) = \infty$, then $\mathbb{P}(x \in A_n \ i.o.) = 1$

3. **Third Borel-Cantelli lemma(BC3):** If $(A_n)_{n=0}^{\infty}$ is a sequence of pairwise independent sets in Ω and $\sum_{n=0}^{\infty} \mathbb{P}(A_n) = \infty$, then for $\mathbb{P}$ a.e. $x \in \Omega$

$$\frac{S_n(x)}{E_n} \to 1$$

where $S_n(x) = \sum_{j=0}^{n-1} \mathbf{1}_{A_j}(x)$ and $E_n = \sum_{j=0}^{n-1} \mathbb{P}(A_j)$.

Note (BC3) implies (BC2), the proof of (BC3) is more complicated.

In the dynamical setting, let $T : \Omega \to \Omega$ be a measure-preserving transformation of the probability space $(\Omega, \mathscr{B}, \mu)$. Suppose that $(A_n)_{n=0}^{\infty}$ is a sequence of sets in $\mathscr{B}$ such that $\sum_{n=0}^{\infty} \mu(A_n) = \infty$. Let $E_n = \sum_{j=0}^{n-1} \mu(A_j)$ and $S_n(x) = \sum_{j=0}^{n-1} \mathbf{1}_{A_j} \circ T^j(x)$. We may consider in what conditions the dynamical version (BC3):

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

does hold. There are some sufficient conditions we can use for the dynamical version (BC3) to hold: e.g.,

1. The sets A_n are all equal to A, $\mu(A) > 0$ and μ is ergodic.

2. The sets $T^{-j}A_j$ are pairwise independent (together with $\sum_{n=0}^{\infty} \mu(A_n) = \infty$)

The first one follows from the Birkhoff ergodic theorem, since $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_A \circ T^i(x) = \int \mathbf{1}_A d\mu = \mu(A)$ and the second one by (BC3) since $\mathbf{1}_{A_j} \circ T^j(x) = \mathbf{1}_{T^{-j}A_j}(x)$. However, for a dynamical system (Ω, T, μ), $T^{-j}A_j$ are usually not pairwise independent. It is reasonable to consider some sufficiently fast decay of correlations, hoping after long enough time, the sequence can be considered almost pairwise independent. We can define the sequence (A_n) to be :

1. **a Borel Cantelli sequence(BC)** if $\mu(x : T^n x \in A_n \ i.o.) = 1$, i.e. $S_n(x)$ is unbounded.

2. **a Strong Borel Cantelli sequence(SBC)** if $\lim\limits_{n\to\infty} \frac{S_n(x)}{E_n} = 1, \ a.s.$.

Note that our interest is actually to show a sequence is a (SBC) under certain conditions.

1.3 Central Limit Theorem

In probability theory, the Central Limit Theorem (CLT) states that, given certain conditions, the arithmetic mean of a sufficiently large number of iterates of independent random variables, each with a well-defined expected value and well-defined variance, will be approximately normally distributed, regardless of the underlying distribution. That is, suppose that a sample is obtained containing a large number of observations, each observation being randomly generated in a way that does not depend on the values of the other observations, and that the arithmetic average of the observed values is computed. If this procedure is performed many times, the central limit theorem says that the computed values of the average will be distributed according to the normal distribution (commonly known as a "bell curve").

The classical central limit theorem (CLT) as follows:

Theorem 1.3.1 ([10], Theorem 3.4.1)**.** *Let* $X_1, X_2, \ldots$ *be i.i.d. random variables, with* $var(X_i) = \sigma^2 \in (0, \infty)$*. If* $S_n = X_1 + \cdots + X_n$ *then*

$$\frac{S_n - nEX_1}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$$

where $N(0,1)$ *denotes the standard normal distribution and* $\xrightarrow{d}$ *denotes convergence in distribution.*

For an ergodic measure preserving transformation T on a probability measure space $(\Omega, \mathscr{B}, \mu)$, take a measurable function(i.e., a random variable) f on Ω. Let

$$(S_n f)(x) = \sum_{i=0}^{n-1} f \circ T^i(x) = f(x) + f \circ T(x) + \cdots + f \circ T^{n-1}(x)$$

We consider asymptotic statistical distribution of $\frac{1}{n} S_n f$. The *mean* and the *variance* of f are given by

$$\mu(f) = \int_\Omega f d\mu$$

and

$$\int_\Omega (f(x) - \mu(f))^2 d\mu$$

respectively. Since T preserves the measure μ, $f \circ T^i$ has the same probability distribution for every $i \geq 0$, i.e.,

$$\mu(f \circ T^i \in E) = \mu(f \in E), \text{ for } E \in \mathscr{B}.$$

Note that the mean of $\frac{1}{n} S_n f$ is equal to $\mu(f)$ and its variance is given by

$$\int_\Omega \left(\frac{S_n f - n\mu(f)}{n} \right)^2 d\mu = \int_\Omega \left(\frac{S_n f}{n} - \mu(f) \right)^2 d\mu.$$

The Central Limit Theorem does not necessarily hold true for general transformation. From probabilistic point of view, if we assume that $f, f \circ T, f \circ T^2, \ldots$ form an independent sequence and without loss of generality, we can assume that $f \in L^2(\mu)$ is centered, i.e. $\mu(f) = \int_\Omega f d\mu = 0$, then

$$\mu(\left\{x \in \Omega \;\middle|\; \frac{S_n f}{||S_n f||_2} < u\right\}) \to \Phi(u) \text{ as } n \to \infty \tag{1.1}$$

where $||S_n f||_2 = L^2$ - norm of $S_n f$ = standard deviation of $S_n f$, and $\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{v^2}{2}} dv$. In this case $||S_n f||_2 = \sqrt{n}||f||_2$. We shall then say that $f \in L^2(\mu)$ satisfies the CLT if (1.1) holds. But independence is a strong assumption and in real situations, it is very unlikely that $\{f \circ T^i\}$ is an independent sequence, but it is stationary. Moreover, if we consider different measurable functions at each time, i.e. $\{f_i\}$, then the sequence $\{f_i \circ T^i\}$ are no longer stationary, in which case, it is more difficult for us to establish central limit theorem. In chapter 2, we will be dealing with some special kinds of measurable functions and establish self-norming central limit theorem.

1.4 Extreme Value Theory

Extreme value theory is a branch of statistics dealing with the extreme deviations from the median of probability distributions. The half-century-old Classical extreme value theory concerns the asymptotic distribution for maxima of independent, identically distributed random variables. Most recent work has been developing interest in the extension of the theory to dependent and stationary sequences. Extreme value theory is widely used in many disciplines, such as finance, earth sciences and traffic prediction. Many real-world cases have shown that the theory works well in predicting the probability distribution of, like extreme floods, equity risks, and large wildfires.

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

Let $X_0, X_1, \ldots, X_n, \ldots$ be a sequence of independent, and identically distributed (i.i.d.) random variables, and let F denote the common distribution function for X_i, so $\mathbb{P}(X_i \leq x) = F(x)$. Instead of studying the sequence itself, we consider the maximum $M_n = \max\{X_0, X_1, \ldots, X_{n-1}\}$, then it is easy to see that

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

The limiting distribution is degenerate, and provides little useful information about the asymptotic properties of the sequence $\{X_i\}$. However, there may exist scaling sequence a_n and b_n such that $a_n(M_n - b_n)$ has a nontrivial limiting distribution, that is,

$$\mathbb{P}(a_n(M_n - b_n) \leq x) = F^n(a_n^{-1}x + b_n) \xrightarrow{w} G(x) \tag{1.2}$$

where $G(x)$ is some continuous probability distribution function and $\xrightarrow{w}$ denotes weak convergence, i.e. convergence on all the possible continuous points of the limiting function G. If (1.2) holds for some sequences $\{a_n > 0\}$, $\{b_n\}$, we call that F belongs to the (i.i.d) *Domain of Attraction* (for maxima) of G and write $F \in D(G)$. So for a given distribution function F, it may turn out that there is *no* extreme value distribution function G such that $F \in D(G)$. This simply means that the maximum M_n does not have a nondegenerate limiting distribution under any *any* normalization(a common example is the Poisson distribution). And if F is in the domain of attraction of G, then we have the following:

Theorem 1.4.1 (Extremal Types Theorem, [31] Theorem 1.4.2)**.** *If for some constants $a_n > 0$, b_n, we have*

$$\mathbb{P}(a_n(M_n - b_n) \leq x) \xrightarrow{w} G(x)$$

for some nondegenerate G, then G is one of the following three extreme value type distributions:

$$\begin{aligned} &\textit{Type I} \;\; G(x) = e^{-e^{-x}}, \quad -\infty < x < \infty; \\ &\textit{Type II} \;\; G(x) = \begin{cases} 0, & x \leq 0 \\ e^{-x^{-\alpha}}, & x > 0 \end{cases} \quad \alpha > 0; \\ &\textit{Type III} \;\; G(x) = \begin{cases} e^{-(-x)^{\alpha}}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0; \end{aligned} \tag{1.3}$$

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

We have considered convergence of probabilities of the form $\mathbb{P}(a_n(M_n - b_n) \leq x)$, which may be rewritten as $\mathbb{P}(M_n \leq u_n)$ where $u_n(x) = x/a_n + b_n$. The convergence was required for every $x \in \mathbb{R}$. Also it is of our interest to consider sequences $\{u_n\}$ which may be more complicated functions than the linear one above. The following theorem is almost trivial in the i.i.d case, but nevertheless is very useful, and will be extended in important ways to apply to dependent sequences and to continuous time processes.

Theorem 1.4.2 (Extreme Value Laws, [31], Theorem 1.5.1)**.** *Let $\{X_n\}$ be an i.i.d sequence. Let $0 \leq v \leq \infty$ and suppose that $\{u_n\}$ is a sequence of real numbers such that*

$$n(1 - F(u_n)) \to v \text{ as } n \to \infty. \tag{1.4}$$

Then

$$\mathbb{P}(M_n \leq u_n) \to e^{-v} \text{ as } n \to \infty. \tag{1.5}$$

Conversely, if (1.5) *holds for some v, $0 \leq v \leq \infty$, then so does* (1.4)*.*

1.4.2 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.

We shall keep the assumption that the $\{X_n\}$ have a common distribution, so it is natural to consider *stationary* sequence, i.e. sequences such that the distributions of $(X_{j_1}, \ldots, X_{j_n})$ and $(X_{j_{1+k}}, \ldots, X_{j_{n+k}})$ are identical for any choice of $n, j_1, \ldots, j_n$, and m. We shall consider a condition, to be called $D(u_n)$, which applies only to a certain sequence of values $\{u_n\}$, as follows:

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

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

where $F_{n_1,\ldots,n_t}$ denotes the joint distribution of $X_{n_1}, \ldots, X_{n_t}$, $\alpha(n, l_n) \to 0$ as $n \to \infty$ for some sequence $l_n = o(n)$.

Condition $D(u_n)$ is like distributional mixing, and it establishes that if two large blocks are sufficiently far apart, then the joint distribution of the two blocks is approximately the product of distributions on the individual blocks. Then the Extremal Types Theorem follows easily under Condition $D(u_n)$.

Theorem 1.4.3 ([31], Theorem 3.3.3)**.** *Let $\{X_n\}$ be a stationary sequence and $a_n > 0$ and b_n given constants such that $\mathbb{P}(a_n(M_n - b_n) \le x)$ converges to a nondegenerate distribution function $G(x)$. Suppose that $D(u_n)$ is satisfied for $u_n = x/a_n + b_n$ for each real x. Then $G(x)$ has one of the three extreme value forms listed in* (1.3).

Extremal Types Theorem has been concerned with the *possible* forms of limiting extreme value distributions. Now we turn to the *existence* of such a limit, in that we formulate conditions under which

$$\mathbb{P}(M_n \le u_n) \to e^{-v} \text{ as } n \to \infty. \tag{1.6}$$

is equivalent to

$$n(1 - F(u_n)) = n\mathbb{P}\{X_0 > u_n\} \to v \text{ as } n \to \infty. \tag{1.7}$$

It may be seen in [31] that if (1.7) holds, then Condition $D(u_n)$ is sufficient to guarantee that $\liminf_n \mathbb{P}(M_n \leq u_n) \geq e^{-v}$. However we need a further assumption to obtain the opposite inequality for the upper limit. Here we use one of various conditions and we refer to this as $D'(u_n)$.

Condition ($D'(u_n)$): The condition $D'(u_n)$ is said to hold for the stationary sequence $\{X_n\}$ and the sequence u_n if

$$\limsup_{n\to\infty} \; n \sum_{j=2}^{[n/k]} P(X_1 > u_n, X_j > u_n) \to 0$$

as $k \to \infty$. (where [ ] denotes the integer part)

Note that under (1.7), the level u_n in (1.4.2) is such that there are on the average approximately v exceedances of u_n among $X_1, \ldots, X_n$, and thus v/k among $X_1, \ldots, X_{[n/k]}$. Condition $D'(u_n)$ bounds the probability of more than one exceedance among $X_1, \ldots, X_{[n/k]}$. This will eventually ensure that there are no multiple points in the point process of exceedances which is necessary in obtaining a simple Poisson limit for this point process. See [31] for more details.

Therefore, Theorem 1.4.2 can be generalized to stationary sequences under $D(u_n)$, $D'(u_n)$.

Theorem 1.4.4 (Extreme Value Law, [31] Theorem 3.4.1)**.** *Let $\{u_n\}$ be constants such that $D(u_n), D'(u_n)$ hold for the stationary sequence $\{X_n\}$. Let $0 \leq v < \infty$. Then (1.6) and (1.7) are equivalent, i.e. $\mathbb{P}(M_n \leq u_n) \to e^{-v}$ if and only if $n\mathbb{P}\{X_0 > u_n\} \to v$.*

1.4.3 Extreme Value Laws(EVL) in Dynamical Systems

We consider a dynamical system $(\Omega, \mathscr{B}, T, \mu)$ where T preserves an invariant measure μ. Consider the time series $X_0, X_1, X_2 \dots$ arising from such a system simply by evaluating a given random variable(r.v.) $\varphi : \Omega \to \mathbb{R} \cup \{\pm\infty\}$ along the orbits of the system, that is to say, we define

$$X_n = \varphi \circ T^n, \tag{1.8}$$

for each $n \in \mathbb{N}$. Clearly, $X_0, X_1, \dots$ defined in this way is not an independent sequence, but T-invariance of μ guarantees that this stochastic process is stationary.

Here we suppose that φ has one global maximum at $\zeta \in \Omega$ ($\varphi(\zeta) = +\infty$ is allowed), and let $u_F := \varphi(\zeta)$. By assuming that φ and μ are sufficiently regular, the event

$$U(u) := \{x \in \Omega : \ \varphi(x) > u\} = \{X_0 > u\}$$

corresponds to a topological ball centered at ζ for u sufficiently close to u_F.

Definition 1.4.5. *Consider a function $\varphi : \Omega \to \mathbb{R}$ and a point $x_0 \in \Omega$. Let d be a distance function on Ω. Now we distinguish the different types of maximum at the point x_0 for φ. We say that:*

1. *φ has a logarithmic singularity at the point x_0 if x_0 has a neighborhood where $\varphi(x) = -C \log d(x, x_0) + g(x)$ with $C > 0$, where g is bounded and has a finite limit as $x \to x_0$.*

2. *φ has a power singularity at the point x_0 if $\varphi(x) \approx Cd(x, x_0)^s$ near x_0 with $s < 0, C > 0$.*

3. *φ has a power function maximum at the point x_0, if $\varphi \approx C - C_1 d(x, x_0)^s$ near x_0 with $s > 0, C_1 > 0$, and supremum of φ on the complement of any open neighborhoods of the point x_0 is less than C.*

We are always interested in studying the extremal behavior of the stochastic process $X_0, X_1, \dots$, and it is associated with the occurrence of exceedances of high levels u. The occurrence of an exceedance at time $j \in \mathbb{N}_0$ means that the event $\{X_j > u\}$ occurs, where u is close to u_F. This is equivalent to saying that the orbit of the point x hits the ball $U(u)$ at time j, i.e. $T^j(x) \in U(u)$.

Like we did in previous section, we consider the extremal behavior of the system for which we define a new sequence of random variables $M_1, M_2, \dots$ given by

$$M_n = \max\{X_0, \dots, X_{n-1}\}.$$

Definition 1.4.6. *(Extreme Value Laws) We say that we have an EVL for M_n if there is a non-degenerate distribution function(d.f.) $G : \mathbb{R} \to [0, 1]$ with $G(0) = 0$ and, for every $\upsilon > 0$, there exists a sequence of levels $u_n = u_n(\upsilon)$, $n = 1, 2, \dots$, such that*

$$n\mu(X_0 > u_n) \to \upsilon, \ \text{as } n \to \infty, \tag{1.9}$$

and for which the following holds:

$$\mu(M_n \leq u_n) \to \bar{G}(\upsilon) \ \text{as } n \to \infty,$$

where $\bar{G} = 1 - G$.

Remark 1.4.7. *As discussed in previous section 1.4.2, for dependent stationary processes* $\{X_n\}$*, Leadbetter [31] gives two conditions called* $D(u_n)$ *and* $D'(u_n)$ *for sequence* $\{u_n\}$ *satisfying* (1.9)*, in which case EVL will hold, that is,* $n\mu(X_0 > u_n) \to \upsilon$ *is equivalent to* $\mu(M_n \leq u_n) \to e^{-\upsilon}$*, which implies* $G(\upsilon) = 1 - e^{-\upsilon}$ *is a standard exponential d.f. and that the waiting times between exceedances of* u_n *is approximately, exponentially distributed.*

In dynamical setting, there are no general techniques for proving these two conditions $D(u_n)$ *and* $D'(u_n)$*. In Collet's paper [7], he used the rate of decay of correlations of Hölder observations to establish* $D(u_n)$ *for certain one-dimensional non-uniformly expanding maps. Freitas et al[12], based on Collet's work, gave a condition* $D_2(u_n)$ *which has the full force of* $D(u_n)$ *and is relatively easier to establish in the dynamical setting by estimating the rate of decay of correlations of Hölder continuous observations or those of bounded variations. The definitions of* $D_2(u_n)$ *is given later, in order to compare* $D_2(u_n)$ *with* $D_3(u_n)$*, which will be introduced in subsection 1.4.5.*

1.4.4 Hitting Time Statistics

We next introduce Hitting Time Statistics for the dynamical system $(\Omega, \mathscr{B}, T, \mu)$. For a set $A \in \mathscr{B}$, we have defined a function $n_A(x)$ that we refer to as *first hitting time function* to A earlier in this thesis and now we recall its definition $n_A : \Omega \to \mathbb{N} \cup \{\infty\}$ where

$$n_A(x) = \min\{j \in \mathbb{N} \cup \{\infty\} : T^j(x) \in A\}.$$

That is, the first time $j \geq 1$ so that $T^j(x) \in A$. One is usually interested in the fluctuations of this function as the set A shrinks. Firstly we consider the Return Time Statistics (RTS) of this system. Let the conditional measure on A be denoted by μ_A, i.e., $\mu_A = \frac{\mu|_A}{\mu(A)}$, which means $\mu_A(B) = \frac{\mu(A \cap B)}{\mu(A)}$ for any $B \in \mathscr{B}$. By Kac's Lemma, the expected value of n_A with respect to μ_A is $\int_A n_A \, d\mu_A = 1/\mu(A)$. So it is reasonable to take a normalizing factor $1/\mu(A)$ when we study the fluctuation of n_A on A. Given a sequence of sets $\{V_n\}_{n\in\mathbb{N}}$ so that $\mu(V_n) \to 0$, the system has *Return Time Statistics* $G(t)$ for $\{V_n\}_{n\in\mathbb{N}}$ if for all $t \geq 0$, we have the following:

$$\lim_{n\to\infty} \mu_{V_n} \left(n_{V_n} \geq \frac{t}{\mu(V_n)} \right) = G(t),$$

given the limit exists.

The dynamical system (Ω, T, μ) is said to have *Return Time Statistics* $G(t)$ *to balls at* ξ if for any sequence $\{\epsilon_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+$ such that $\epsilon_n \to 0$ as $n \to \infty$ we have RTS $G(t)$ for $V_n = B_{\epsilon_n}(\xi)$.

If we consider n_A on the whole space Ω, i.e., not simply restricted to A, we are studying the Hitting Time Statistics. We will use the same normalizing factor $1/\mu(A)$ in this case. Analogously to the RTS case, given a sequence of sets $\{V_n\}_{n\in\mathbb{N}}$ so that $\mu(V_n) \to 0$, the system has *Hitting Time Statistics* $G(t)$ for $\{V_n\}_{n\in\mathbb{N}}$ if for all $t \geq 0$, we have the following:

$$\lim_{n\to\infty} \mu \left(n_{V_n} \geq \frac{t}{\mu(V_n)} \right) = G(t),$$

given the limit exists. HTS to balls at a point ξ is defined analogously to RTS to balls.

In [13], the equivalence between EVL and HTS/RTS (for balls) of stochastic processes defined by (1.8) was obtained for dynamical systems (Ω, T, μ) admitting an absolutely continuous invariant probability measure μ, i.e. if such processes have an EVL G then the system has HTS G as well for balls "centered" at ζ and vice versa.

1.4.5 Rare Events Points Processes and respective convergence

We use the definitions and concepts from [14] in this section. Let $\mathcal{S} = \{[a,b) \mid a,b \in \mathbb{R}^+\}$. Let $\mathcal{R}$ denote the ring generated by $\mathcal{S}$, i.e. $\mathcal{R} = \{J \mid \exists k \text{ such that } J = \cup_{j=1}^k I_j,\ I_j \in \mathcal{S},\ j = 1, \ldots, k\}$. For $I = [a,b) \in \mathcal{S}$ and $\alpha \in \mathbb{R}$, we define $\alpha I := [\alpha a, \alpha b)$ and $I + \alpha := [a+\alpha, b+\alpha)$. Similarly, for $J \in \mathcal{R}$ define $\alpha J := \alpha I_1 \cup \cdots \cup \alpha I_k$ and $J + \alpha := (I_1 + \alpha) \cup \cdots \cup (I_k + \alpha)$.

Definition 1.4.8. *For stationary processes $X_0, X_1, \ldots$ and sequences $(u_n)_{n\in\mathbb{N}}$ satisfying* (1.9)*, we define the* **Rare Event Point Process***(REPP) by counting the number of exceedances (or hits to $U(u_n)$) during the re-scaled time period $s_n J \in \mathcal{R}$, where $J \in \mathcal{R}$ and $s_n := 1/\mu(X_0 > u_n)$ is, according to Kac's Theorem, the expected waiting time before the occurrence of one exceedance. To be more precise, for every $J \in \mathcal{R}$, set*

$$N_n(J) := \sum_{j \in s_n J \cap \mathbb{N}_0} \mathbb{1}_{\{X_j > u_n\}}$$

Definition 1.4.9. *(Poisson process) Let $T_1, T_2, \ldots$ be an i.i.d. sequence of random variables with common exponential distribution of mean $1/\theta$. Given this sequence of* *r.v.s, for $J \in \mathcal{R}$, set*

$$N(J) = \Big|\{i \in \mathbb{N} : \sum_{j=1}^{i} T_j \in J\}\Big|$$

where $\Big| \cdot \Big|$ denotes the cardinality of a set. We say that N defined in this way is a Poisson process of intensity θ. In special case, when $J = [0, t)$, we also denote $N([0, t))$ by $N(t)$.

Remark 1.4.10. *If $\theta = 1$ then we say that N is a standard Poisson process and, for every $t > 0$, the random variable $N(t)$ has a Poisson distribution of mean t. In general, the random variable $N(J)$ has distribution*

$$\mathbb{P}(N(J) = k) = e^{-m(J)} \frac{(m(J))^k}{k!}$$

where $m(J)$ is the Lebesgue measure of J.

Definition 1.4.11. *Suppose that $(N_n)_{n \in \mathbb{N}}$ is a sequence of point process defined on $\mathcal{R}$ and N is a standard Poisson process defined on $\mathcal{R}$. We say that N_n* converges in distribution *to N if the sequence of vector random variables $\{N_n(J_1), N_n(J_2), \ldots, N_n(J_k)\}$ converges in distribution to $\{N(J_1), N(J_2), \ldots, N(J_k)\}$, for every $k \in \mathbb{N}$ and all $J_1, J_2, \ldots, J_k \in \mathcal{R}$ such that $N(\partial J_i) = 0$ a.s., for $i = 1, \ldots, k$.*

Remark 1.4.12. *The convergence of the REPP to the Poisson process is stronger than the existence of an EVL for $X_0, X_1, \ldots, X_n, \ldots$. In particular, not only we can recover the distributional limit for the maxima $\{M_n\}$ since $\{M_n \leq u_n\} = \{N_n(n/s_n) = 0\}$, also we can obtain the distributional limit of the order statistics, that is, if $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$ denote the order statistics of the first n random variables of the process, then $\{X_{n-k,n} \leq u_n\} = \{N_n(n/s_n) \leq k\}$.*

In [14, page 4], two conditions $D_3(u_n)$ and $D'(u_n)$(It is the same $D'(u_n)$ we mentioned before) are given on the dependence structure of a general stationary s-tochastic process to ensure that the REPP N_n converges in distribution to a standard Poisson process. $D_3(u_n)$ is highly related to the condition $D_2(u_n)$ and their proofs are similar, since they can be easily checked by using decay of correlations. We will prove $D_3(u_n)$ and $D'(u_n)$ for REPP. The formulation of conditions $D_3(u_n), D_2(u_n)$ are given as follows.

For every $B \in \mathcal{R}$, let

$$M(B) := \max\{X_i : i \in B \cap \mathbb{Z}\}$$

In particular, when $B = [0, n)$, we have $M(B) = M([0, n)) = M_n$. Note that $\{M(B) \leq u_n\} = \{N_n(a_n^{-1}B) = 0\}$.

Condition $(D_3(u_n))$. We say that $D_3(u_n)$ holds for the sequence $X_0, X_1, X_2, \ldots$ if for all $B \in \mathcal{R}$ and $t \in \mathbb{N}$,

$$\left|\mu(\{X_0 > u_n\} \cap \{M(B+t) \leq u_n\}) - \mu(\{X_0 > u_n\})\mu(\{M(B) \leq u_n\})\right| \leq \gamma(n, t),$$

where $\gamma(n, t)$ is nonincreasing 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 $(D_2(u_n))$. We say that $D_2(u_n)$ holds for the sequence $X_0, X_1, X_2, \ldots$ if for any integers l, t and n

$$\left|\mu(\{X_0 > u_n\} \cap \{M_{t,l} \leq u_n\}) - \mu(\{X_0 > u_n\})\mu(\{M_l \leq u_n\})\right| \leq \gamma(n, t),$$

where $M_{t,l} = \max\{X_t, X_{t+1}, \ldots, X_{t+l-1}\}$, and $\gamma(n, t)$ is nonincreasing 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 $D_3(u_n)$ is a sort of mixing requirement specially adapted to the problem of counting exceedances. While $D_2(u_n)$ is a condition on the long range dependence structure of the stochastic process $X_0, X_1, \ldots$.

Remark 1.4.13. *For systems satisfying the condition called Assumption (A), condition $D_2(u_n)(D_3(u_n))$ often follows easily if there are good enough estimates on decay of correlation for observations in a suitable Banach space.*

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

J. M. Freitas, N. Haydn, and M. Nicol [14] show that the REPP N_n converges in distribution to a standard Poisson process for functions maximized at generic points in a variety of billiard systems. They prove this by verifying that the conditions $D_3(u_n)$ and $D'(u_n)$ hold for such systems.

1.5 Frobenius-Perron Operators

Let $(\Omega, \mathscr{B}, \mu)$ be a measure space. For a given measurable transformation $T : \Omega \to \Omega$, we can define the corresponding Frobenius-Perron operator, which gives the evolution of probability density functions governed by the deterministic dynamical system. First, let me give the definition of nonsingular transformation.

Definition 1.5.1. *A measurable transformation $T : \Omega \to \Omega$ on a measure space*

$(\Omega, \mathscr{B}, \mu)$ is nonsingular if $\mu(T^{-1}(A)) = 0$ for all $A \in \mathscr{B}$ such that $\mu(A) = 0$.

Note that every measurable preserving transformation is necessarily nonsingular with respect to the invariant measure. Since absolutely continuous invariant measures, which are important measures in many applications, are more of interests of mathematicians, here we restrict the domain of the Frobenius-Perron operator to the set of all probability measures on Ω which are absolutely continuous with respect to μ. By the Radon-Nikodym theorem, such a set is equivalent to the set of all densities of $\mathscr{L}^1(\mu)$, denoted by $\mathcal{D}$. This observation leads to the following definition of Frobenius-Perron operators on functions.

For a given function $f \in \mathscr{L}^1(\mu)$, define a measure

$$\mu_f(A) = \int_{T^{-1}(A)} f d\mu, \quad \forall A \in \mathscr{B}.$$

Since T is nonsingular, $\mu(A) = 0$ implies $\mu_f(A) = 0$. Thus, Radon-Nikodym theorem implies that there exists a unique function $\hat{f} \in \mathscr{L}^1(\mu)$, denoted as Pf, such that

$$\mu_f(A) = \int_A \hat{f} d\mu, \quad \forall A \in \mathscr{B}.$$

Definition 1.5.2. *The operator $P : \mathscr{L}^1(\mu) \to \mathscr{L}^1(\mu)$ defined by*

$$\int_A Pf d\mu = \int_{T^{-1}(A)} f d\mu, \quad \forall A \in \mathscr{B}, \quad \forall f \in \mathscr{L}^1(\mu).$$

is called the Frobenius-Perron operator *associated with T.*

Remark 1.5.3. *Frobenius-Perron operator is also usually called transfer operator, because it evolves the density of the absolutely continuously invariant measure.*

With the definition, we can show that P has the following properties:

(i) P is linear, that is, for all $a, b \in \mathbb{R}$ and $f_1, f_2 \in \mathscr{L}^1$,

$$P(af_1 + bf_2) = aPf_1 + bPf_2;$$

(ii) P is a positive operator, that is, $Pf \geq 0$ if $f \geq 0$;

(iii) $\int_\Omega Pf d\mu = \int_\Omega f d\mu$; and

(iv) We can write P_T for P sometimes to emphasize the dependence of the operator P on the transformation T. We have $P_{T_1 \circ T_2} = P_{T_1} P_{T_2}$ for nonsingular transformations T_1 and T_2 from Ω into itself. In particular, $P_{T^n} = (P_T)^n$.

The importance of having Frobenius-Perron operator is shown in the following theorem.

Theorem 1.5.4 ([9], Theorem 4.2.1). *Let P be the Frobenius-Perron operator associated with a nonsingular transformation $T : \Omega \to \Omega$ and let $f \in \mathscr{L}^1$ be nonnegative. The finite measure μ_f defined by*

$$\mu_f(A) = \int_A f d\mu, \quad \forall A \in \mathscr{B}$$

is T-invariant if and only if f is a fixed point of P.

Remark 1.5.5. *Note that the measure μ is invariant under T if and only if $P1 = 1$, where 1 is the constant 1 function.*

Next, we introduce the Koopman operator which is the dual operator of the Frobenius-Perron operator.

Definition 1.5.6. *Let $(\Omega, \mathscr{B}, \mu)$ be a measure space and let $T : \Omega \to \Omega$ be a nonsingular transformation. The linear operator $U : \mathscr{L}^\infty(\Omega) \to \mathscr{L}^\infty(\Omega)$ defined by*

$$(Uf)(x) = (U_T f)(x) = f(T(x)), \quad \forall x \in \Omega, \forall f \in \mathscr{L}^\infty(\Omega)$$

is called the Koopman operator with respect to T.

Because of the nonsingularity assumption of T, the Koopman operator is well defined since $f_1(x) = f_2(x)$, $x \in \Omega$ μ-a.e. implies that $f_1(T(x)) = f_2(T(x))$, $x \in \Omega$ μ-a.e.. The following are some basic properties of the Koopman operator U:

(i) U is a positive operator;

(ii) U is a (weak) contraction on $\mathscr{L}^\infty$, i.e., $||Uf||_\infty \leq ||f||_\infty$ for all $f \in \mathscr{L}^\infty$;

(iii) $U_{T_1 \circ T_2} = U_{T_1} U_{T_2}$. In particular, $U_{T^n} = (U_T)^n$.

Proposition 1.5.7 ([9], Proposition 4.3.1)**.** *The Koopman operator U is the dual of the Frobenius-Perron operator P, that is*

$$< Pf, g >=< f, Ug >, \ \forall f \in \mathscr{L}^1, \ g \in \mathscr{L}^\infty.$$

where $< \cdot, \cdot >$ denotes the inner product.

Remark 1.5.8. *Note that Frobenius-Perron operator and Koopman operator are also both well defined on $\mathscr{L}^2(\mu)$.*

1.6 Functions of Bounded Variation

There is a key concept which plays an important role in studying Frobenious-operators that are defined on $\mathscr{L}^1$ spaces. Let me introduce the classic definition of variation for functions of one variable, which will be used for the statistical study of one-dimensional mappings.

Definition 1.6.1. *Let $f : [a, b] \to \mathbb{R}$ and the variation of f on $[a, b]$ is defined as the nonnegative number (may be ∞)*

$$\bigvee_a^b f = \sup\{\sum_{i=1}^n |f(x_i) - f(x_{i-1})| : a = x_0 < x_1 < \cdots < x_n = b\}.$$

if $\bigvee\limits_a^b f < \infty$, f is said to be of bounded variation on $[a, b]$.

For two $\mathscr{L}^1$ functions f and g, which have different values only on a set of measure zero, i.e., $f(x) = g(x)$, for x-a.e., but the definition above may give different values of $\bigvee\limits_a^b f$ and $\bigvee\limits_a^b g$. So we define the variation of a $\mathscr{L}^1$ function as follows.

Definition 1.6.2. *Let $f \in \mathscr{L}^1(a, b)$, then its variation on $[a, b]$ is defined as*

$$\bigvee_{[a,b]} f = \inf\{\bigvee_a^b g : g(x) = f(x), \forall x \in [a, b] \text{ a.e.}\}$$

if $\bigvee_{[a,b]} f < \infty$, f is said to be of bounded variation on $[a, b]$.

We now make the space of functions of bounded variation into a Banach space. Let

$$BV([a, b]) = \{f \in \mathscr{L}^1 : \bigvee_{[a,b]} f < \infty\}$$

For example, the function

$$f(x) = \begin{cases} n, & \text{if } x = \frac{1}{n}, \quad n = 1, 2, \dots \\ 0, & \text{otherwise} \end{cases}$$

has infinite variation, but $f \in BV([0,1])$ because $g = 0 = f$, $x \in [0,1]$ a.e. and $\bigvee_{[0,1]} g = 0$.

A norm on $BV([a,b])$ is defined as

$$||f||_{BV} = ||f||_1 + \bigvee_{[a,b]} f, \quad \text{for } f \in BV([a,b])$$

Remark 1.6.3. *Note that $BV([a,b])$ is dense in $\mathscr{L}^1([a,b])$.*

1.7 Shrinking Target

Suppose (Ω, T, μ) is an ergodic dynamical system and $B_n(p)$ is a nested sequence of balls about a point $p \in X$, with radius $r(n)$ is a decreasing sequence. The study of hitting time statistics to the sequence of nested balls is called the shrinking target problem. We can establish all the related statistical properties, mentioned above, if the sequence of balls satisfy certain conditions. There have been many papers concerning the Strong Borel Cantelli properties of the nested sequence of balls, i.e., the behavior of the almost sure limit of the normalized sum $\frac{1}{E_n} \sum_{i=1}^{n} 1_{B_i(p)} \circ T^i(x)$ where $E_n := \sum_{i=1}^{n} \mu(B_i(p))$ diverges [6, 29, 16, 18, 21, 24, 27]. Many of these references consider more general sequences of sets than nested balls. In chapter 2, we study self-norming central limit theorems for the shrinking target problem, namely the distribution limit of $\frac{1}{a_n}\sum_{i=1}^{n}[1_{B_i(p)} \circ T^i - \mu(B_i)]$ where a_n is a sequence of norming constants. One important case for applications is the case where $\mu(B_i(p)) = \frac{1}{i}$, and hence $E_n = \log n$. The central limit theorem results are stated for balls satisfying $\sum_i \mu(B_i(p)) = \infty$ and $\mu(B_i(p)) \leq \frac{C_2}{i^\gamma}$ where C_2 is a positive constants and $0 < \gamma \leq 1$. The main difficulty is to establish that the non-stationary variance has a limit in probability. Our results are limited to non-uniformly expanding systems i.e. those without a contracting direction and are based upon the Gordin [17] martingale approximation approach (see also [32]). Then in chapter 3, we show some applications of results in chapter 2 for some different expanding systems. In chapter 4, we will establish strong Borel Cantelli Lemmas and Convergence of Rare Event Point Processes (therefore, Extreme Value Laws follows) for the Lorenz maps F(see chapter 4 for details) with shrinking target property.

Chapter 2

CLT for shrinking target[1]

This chapter is going to study central limit theorems for the shrinking target problem. Moreover, it is also an attempt to study the statistics of non-stationary stochastic processes arising as observations (which perhaps change over time) on an underlying dynamical system (which may change over time), that is, we consider the processes $X_n = \phi_n \circ T_n$ where $\{\phi_n\}$ is a sequence of functions, and one special case is when $T_n = T^n$. The theory of central limit theorem for independent stochastic processes, introduced in section 1.3 does not hold here, since it is neither stationary nor independent. Conze and Raugi [8] studied similar problems for sequential expanding dynamical systems. Somewhat related results were obtained by Nándori, Szász and Varjú [34] who obtained central limit theorems in the setting in which a fixed observation $\phi : \Omega \to \mathbb{R}$ was considered on a space on which a sequence of different

[1] **This chapter contains published work from** *Haydn, N., Nicol, M., Vaienti, S., and Zhang, L. (2013). Central limit theorems for the shrinking target problem. Journal of Statistical Physics, 153(5), 864-887. Chicago.*

transformations acted $T_i : \Omega \to \Omega$, preserving a common invariant measure μ. The main difficulty in [34] was also controlling the variance, but the setting in which the underlying maps change but the observation is fixed is simpler in some respects and more difficult in others.

We achieve fairly complete results in the case in which the transfer operator with respect to the invariant measure is quasicompact in the bounded variation norm. These results are contained in Proposition 3.3.1 and Theorem 3.4.4. For systems in which the transfer operator is quasicompact in a Hölder or Lipschitz space we show that under the assumption we call (SP) (derived from a Gal-Koksma lemma as formulated by Sprindzuk [41]) or a form of short returns assumption called Assumption (C) we have a central limit theorem (Theorem 3.1.1). Assumption (C) and the (SP) property have been shown to hold for generic points in a variety of non-uniformly expanding systems [7, 20, 26].

In Section 2 we discuss the set-up, describe the martingale approach we use, prove some general results on variance and discuss the (SP) property and Assumption (C). Section 3 gives the results under the assumption of quasi-compactness in Hölder norms and also some applications. In Section 4 we give our results when we have quasi-compactness of the transfer operator in the bounded variation norm, and we give applications to piecewise expanding maps in higher dimensions. In Appendices, Appendix A.1 describes the Gal-Koksma lemma we use and Appendix A.3 shows that Assumption (C) is satisfied for generic points in many of our applications.

2.1 The setup.

We suppose that (Ω, T, μ) is an ergodic dynamical system. Let the transfer operator P be defined as in section 1.5, by proposition 1.5.7, we have

$$\int P\phi \cdot \psi \, d\mu = \int \phi \cdot U\psi d\mu = \int \phi \cdot \psi \circ T \, d\mu,$$

for all functions $\phi \in \mathscr{L}^1(\mu)$, $\psi \in \mathscr{L}^\infty(\mu)$. Suppose $\mathcal{B}_\alpha$ is a Banach space of functions embedded in $\mathscr{L}^1(\mu)$ such that for any $\phi \in \mathcal{B}_\alpha$, $\|\phi\|_1 \leq C\|\phi\|_\alpha$ where $\|.\|_\alpha$ is the Banach space norm and $\|.\|_1$ is the $\mathscr{L}^1$ norm with respect to μ. We consider P restricted to the Banach space $\mathcal{B}_\alpha$, i.e., $P|_{\mathcal{B}_\alpha} : \mathcal{B}_\alpha \to \mathcal{B}_\alpha$ and we assume that $\|P^n\phi\|_\alpha \leq C_1\theta^n\|\phi\|_\alpha$ for all $\phi \in \mathcal{B}_\alpha$ such that $\int \phi \, d\mu = 0$ and for some $0 < \theta < 1$. This implies exponential decay of correlations of the form,

$$\left| \int \phi\psi \circ T^n \, d\mu - \int \phi \, d\mu \int \psi \, d\mu \right| \leq C\theta^n \|\phi\|_\alpha \|\psi\|_1$$

for all $\phi \in \mathcal{B}_\alpha$, $\psi \in \mathscr{L}^1(\mu)$.

Remark 2.1.1. *The weaker assumption of exponential decay of correlations*

$$\left| \int \phi\psi \circ T^n \, d\mu - \int \phi \, d\mu \int \psi \, d\mu \right| \leq C\theta^n \|\phi\|_\alpha \|\psi\|_\infty$$

implies that $\|P^n\phi\|_1 \leq C\theta^n\|\phi\|_\alpha$ (by taking ψ to be sign($P^n\phi$)) and hence P contracts exponentially in the $\mathscr{L}^1$ norm. This assumption is sufficient for all our results on variance in Section 2, with the exception of the proof of the boundedness of the terms w_j, given in Lemma 2.5.1 which seems to require our stronger assumption that $\|P^n\phi\|_\alpha \leq C_1\theta^n\|\phi\|_\alpha$. These estimates on the growth of w_j are used in the proof of Theorem 3.1.1. If $\mathcal{B}_\alpha$ is the space of functions of bounded variation then the w_j terms are easily seen to be uniformly bounded under the assumption $\|P^n\phi\|_{BV} \leq C\theta^n\|\phi\|_{BV}$.

Let $p \in \Omega$ and let $B_n(p)$ be a sequence of nested balls about p such that $\mu(B_n(p)) \leq \frac{C_2}{n^\gamma}$ for constants $C_2 > 0$ and $0 < \gamma \leq 1$. Let $\mathbf{1}_{B_n(p)}$ be the characteristic function of $B_n(p)$. We will sometimes write $E[\phi]$ for the integral $\int \phi \, d\mu$ when the context is understood. Our standing assumption is that $\sum_{j=1}^{n} \mu(B_j(p)) \to \infty$ as $n \to \infty$.

$\mathbf{1}_{B_n(p)}$ may not lie in $\mathcal{B}_\alpha$ but we assume we may take an approximation $\tilde{\phi}_n^\alpha \in \mathcal{B}_\alpha$ to it such that:

(i) $||\mathbf{1}_{B_n(p)} - \tilde{\phi}_n^\alpha||_1 \leq \frac{1}{n^3}$ and;

(ii) $\|\tilde{\phi}_n^\alpha\|_\alpha \leq Cn^k$ where $C,\ k$ are independent of n;

(iii) $\tilde{\phi}_n^\alpha \geq 0$, $\tilde{\phi}_n^\alpha \geq \tilde{\phi}_{n+1}^\alpha$

We define $\phi_n^\alpha = \tilde{\phi}_n^\alpha - \int \tilde{\phi}_n^\alpha d\mu$ so that $\int \phi_n^\alpha d\mu = 0$. For ease of notation we will just use ϕ_n and $\tilde{\phi}_n$ instead of ϕ_n^α and $\tilde{\phi}_n^\alpha$.

Define $\phi_0 = 0$ and for $n \geq 1$

$$w_n = P\phi_{n-1} + P^2\phi_{n-2} + \ldots + P^n\phi_0 = \sum_{j=1}^{n} P^j \phi_{n-j}$$

so that $w_1 = P\phi_0$, $w_2 = P\phi_1 + P^2\phi_0$, $w_3 = P\phi_2 + P^2\phi_1 + P^3\phi_0$ etc... For $n \geq 1$ define

$$\psi_n = \phi_n - w_{n+1} \circ T + w_n$$

Recall our assumptions $\|\tilde{\phi}_n\|_\alpha \leq Cn^k$(so $\|\phi_n\|_\alpha \leq \tilde{C}n^k$) and $\|P^n\phi\|_\alpha \leq C_1\theta^n\|\phi\|_\alpha$ for all $\phi \in \mathcal{B}_\alpha$ such that $\int \phi d\mu = 0$; moreover we have the monotonicity property

$||\phi_{n-j}||_\alpha \leq ||\phi_n||_\alpha$, for $j < n$. These facts immediately imply that $\|w_n\|_\alpha \leq C_2\|\phi_n\|_\alpha$ (where the constant C_2 takes care of the sum of the geometric series of in θ), $\|w_n \circ T\|_\alpha \leq C_3\|\phi_n\|_\alpha$ (since $\|UP\phi\|_\alpha \leq C\|\phi\|_\alpha$ for all $\phi \in \mathcal{B}_\alpha$) and hence $\|\psi_n\|_\alpha \leq C_4\|\phi_n\|_\alpha$. Using the fact that $P(w_{n+1} \circ T) = w_{n+1}P\mathbf{1} = w_{n+1}$ we get that

$$P\psi_n = P\phi_n - P(w_{n+1} \circ T) + Pw_n = 0$$

.

Since $UP(\cdot) = E[\cdot|T^{-1}\mathcal{B}]$, $P\psi_j = 0$ implies that $E[\psi_j|T^{-1}\mathcal{B}] = 0$ and in turn $E[\psi_j \circ T^j|T^{-1-j}\mathcal{B}] = 0$ (since T preserves μ). Furthermore $\psi_j \circ T^j$ is $T^{-j}\mathcal{B}$ measurable for all $j \geq 0$.

We will use the approach of Gordin to express $\sum\limits_{j=1}^{n} \phi_j \circ T^j$ as the sum of a (non-stationary) martingale difference array and a controllable error term and then use Theorem 3.2 from Hall and Heyde [22]. The definition of martingale is the following:

Definition 2.1.2 (Martingales, see [10] chapter 5)**.** *Let $\mathscr{F}_n$ be a filtration, i.e., an increasing sequence of σ-fields. A sequence $\{X_n\}$ is said to be adapted to $\mathscr{F}_n$ if $X_n \in \mathscr{F}_n$ for all n. If $\{X_n\}$ is sequence with*

(i) $E|X_n| < \infty$,

(ii) X_n *is adapted to* $\mathscr{F}_n$,

(iii) $E(X_{n+1}|\mathscr{F}_n) = X_n$ *for all* n,

then $\{X_n\}$ is said to be a martingale(with respect to $\mathscr{F}_n$).

Theorem 2.1.3 (Theorem 3.2 [22])**.** *Let $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq k_n, n \geq 1\}$ be a zero-mean square-integrable martingale difference array with differences $X_{n,i}$ and let η^2 be an almost-sure finite random variable. Suppose that:*

(a) $\max_i |X_{n,i}| \to 0$ *in probability;*

(b) $\sum_i X_{n,i}^2 \to \eta^2$ *in probability;*

(c) $E(\max_i X_{n,i}^2)$ *is bounded in n;*

(d) the σ-fields are nested: $\mathcal{F}_{n,i} \subset \mathcal{F}_{n+1,i}$ for $1 \leq i \leq k_n$, $n > 1$.

Then $S_{n,k_n} \to Z$ (in distribution) where the random variable Z has the characteristic function $E(\exp(-\frac{1}{2}\eta^2 t^2))$.

As is common in the application of martingale theory to non-invertible dynamical systems we will have to consider the natural extension so that we have a martingale in backwards time. We outline our scheme of proof.

Let $(\tilde{\Omega}, \tilde{\sigma}, \tilde{\mu})$ be the natural extension of (Ω, T, μ). Each ψ_j lifts to to a function ψ_j^* on $\tilde{\Omega}$ in a natural way, $\psi_j^*(\ldots \omega_{-2}\omega_{-1}.\omega_0\omega_1 \ldots) := \psi_j(\omega_0)$. To simplify notation we write simply ψ_j instead of ψ_j^*.

We define scaling constants by $a_n^2 = E[(\sum_{j=1}^{n} \phi_j \circ T^j)^2]$. This sequence of constants play the role of non-stationary variance. Giving estimates on the growth and non-degeneracy of a_n in this non-stationary setting is more difficult than in the usual stationary case.

We define a triangular array $X_{n,i} = \frac{1}{a_n}\psi_{n-i} \circ \tilde{\sigma}^{-i}$, $i = 1, \ldots, n, n \in \mathbb{N}$, and put $S_{n,i} = \sum_{j=1}^{i} X_{n,j}$ for the partial sums (along rows). Then $X_{n,i}$ is $\mathcal{F}_i := \tilde{\sigma}^i \mathcal{B}_0$

measurable where $\mathcal{B}_0$ is the σ-algebra $\mathcal{B}$ lifted to Ω. Note that in Theorem 2.1.3 we take $\mathcal{F}_{n,i} := \mathcal{F}_i$ for all n and $k_n = n$. The $\mathcal{F}_i$ form an increasing sequence of σ-algebras. We obtain $E[S_{n,i+1}|\mathcal{F}_i] = S_{n,i} + E[X_{n,i+1}|\mathcal{F}_i]$ where by stationarity $E[X_{n,i+1}|\mathcal{F}_i] = E[\psi_{n-i-1}|\tilde{\sigma}\mathcal{B}_0] = 0$. Hence $E[S_{n,i+1}|\mathcal{F}_i] = S_{n,i}$ and for every $n \in \mathbb{N}$, $X_{n,i}$ is a martingale difference array with respect to $\mathcal{F}_i$.

We will then verify conditions (a), (b), (c) and (d) of Theorem 2.1.3. The hard part lies in establishing (b). This is in contrast with the stationary setting where condition (b) is usually a straightforward consequence of the ergodic theorem. Condition (b) is established in [34] by using [40, Lemma 3.3], however in next chapter, we will see that the Lipschitz norms of the observations $\tilde{\phi}_i$ are unbounded and other techniques have to be used.

Once we have established (a), (b), (c) and (d) it follows that $\lim\limits_{n\to\infty} \frac{1}{a_n} \sum\limits_{j=0}^{n-1} \psi_j \circ T^j \to N(0,1)$ in distribution. In the final step we show that $\frac{1}{a_n} \sum\limits_{j=1}^{n} [w_j \circ T^j - w_j \circ T^{j+1}] \to 0$ in $\mathscr{L}^1$ which implies that $\lim\limits_{n\to\infty} \frac{1}{a_n} \sum\limits_{j=0}^{n-1} \phi_j \circ T^j \to N(0,1)$ in distribution.

2.2 Some lemmas on variance

In this section we establish some preliminary results on the growth of the variance $E[(\sum\limits_{j=1}^{n} \phi_j \circ T^j)^2]$, i.e., the scaling constants a_n.

For further reference let us notice that $\|P^n\phi\|_\alpha \le C_3\theta^n\|\phi\|_\alpha$ and $\|\phi\|_1 \le C_3\|\phi\|_\alpha$

and that there exists a constant a such that

$$\| \sum_{j > a \log i} P^j \phi_i \|_1 \leq \frac{1}{i^3}. \tag{2.1}$$

Lemma 2.2.1.

$$\limsup_{n \to \infty} \frac{1}{E_n} E[(\sum_{i=1}^{n} \phi_i \circ T^i)^2] \geq 1$$

where $E_n = \sum_{j=1}^{n} E(\phi_j^2)$.

Proof: By exponential decay of correlations and (2.1) we get for the long term interactions:

$$\sum_{j > a \log i + i} \left| \int \phi_i \circ T^i \cdot \phi_j \circ T^j d\mu \right| \leq \frac{c_1}{i^2},$$

where we used exponential decay and our bound $\|\phi_j\|_1 \leq C_3 \|\phi_j\|_\alpha \leq c_1 j^k$, where c_1,k are independent of j. This bound is from assumption (ii). Recall $\phi_j = \tilde{\phi}_j - \int \tilde{\phi}_j$ and $\|\tilde{\phi}_j\|_1 \leq \frac{c_2}{j^\gamma}$ (for some c_2). Thus for the short term interactions we get

$$\sum_{j=i+1}^{i+a\log i} \int \phi_i \circ T^i \cdot \phi_j \circ T^j d\mu = \sum_{j=i+1}^{i+a\log i} \int \tilde{\phi}_i \circ T^i \cdot \tilde{\phi}_j \circ T^j d\mu + O(a \log i E(\tilde{\phi}_i)^2)$$

whence

$$\sum_{i=1}^{n} \sum_{j > i} E[\phi_i \circ T^i \phi_j \circ T^j] = \sum_{i=1}^{n} \sum_{j=i+1}^{i+a\log i} E[\tilde{\phi}_i \circ T^i \tilde{\phi}_j \circ T^j] + \sum_{i=1}^{n} O(a \log i E(\tilde{\phi}_i)^2).$$

Since

$$E[(\sum_{i=1}^{n} \phi_i \circ T^i)^2] = \sum_{i=1}^{n} E(\phi_i^2) + 2 \sum_{i=1}^{n} \sum_{j > i} E[\phi_i \circ T^i \cdot \phi_j \circ T^j]$$

and $\sum_{i=1}^{n} \sum_{j=i+1}^{i+a\log i} E[\tilde{\phi}_i \circ T^i \cdot \tilde{\phi}_j \circ T^j] + \sum_{i=1}^{n} a \log i E(\tilde{\phi}_i)^2 \geq 0$ the lemma is proved. $\square$

Lemma 2.2.2.

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\int \phi_i \circ T^i \cdot \phi_j \circ T^j d\mu = \sum_{i=2}^{n}\int (\phi_i w_i)\circ T^i d\mu$$

Proof: Recalling that $\phi_0 = 0$ this follows by a direct calculation and rearrangement of terms as

$$\begin{aligned}
\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\int \phi_i \circ T^i \cdot \phi_j \circ T^j d\mu &= \sum_{j=2}^{n}\sum_{i=1}^{j-1}\int \phi_i \circ T^i \cdot \phi_j \circ T^j d\mu \\
&= \sum_{j=2}^{n}\sum_{i=1}^{j-1}\int P^{j-i}\phi_i \cdot \phi_j d\mu \\
&= \sum_{j=2}^{n}\int (\sum_{i=1}^{j-1} P^{j-i}\phi_i)\phi_j d\mu \\
&= \sum_{j=2}^{n}\int w_j \phi_j d\mu \\
&= \sum_{i=2}^{n}\int (\phi_i w_i)\circ T^i d\mu.
\end{aligned}$$

The last equality is because T is measure preserving. □

The following lemma is the main result of this subsection:

Lemma 2.2.3.

$$a_n^2 = E[(\sum_{i=1}^{n}\phi_i \circ T^i)^2] = \sum_{i=1}^{n} E[\psi_i^2] - \int w_1^2 d\mu + \int w_{n+1}^2 d\mu$$

Proof: Let us first observe that factoring out yields

$$\begin{aligned}
\psi_j^2 &= \phi_j^2 + 2\phi_j(w_j - w_{j+1}\circ T) + (w_j - w_{j+1}\circ T)^2 \\
&= \phi_j^2 + 2\phi_j(w_j - w_{j+1}\circ T) + w_j^2 + w_{j+1}^2 \circ T - 2w_j w_{j+1}\circ T
\end{aligned}$$

which when integrated leads to

$$
\begin{aligned}
& \int \psi_j^2 d\mu \\
= \; & \int \phi_j^2 d\mu + 2\int \phi_j(w_j - w_{j+1}\circ T)d\mu + \int w_j^2 d\mu + \int w_{j+1}^2 d\mu - 2\int w_j w_{j+1}\circ T d\mu \\
= \; & \int \phi_j^2 d\mu + 2\int \phi_j w_j d\mu - 2\int P\phi_j w_{j+1} d\mu + \int w_j^2 d\mu + \int w_{j+1}^2 d\mu - 2\int P w_j w_{j+1} d\mu \\
= \; & \int \phi_j^2 d\mu + 2\int \phi_j w_j d\mu - 2\int P\phi_j w_{j+1} d\mu + \int w_j^2 d\mu + \int w_{j+1}^2 d\mu \\
& \qquad\qquad\qquad\qquad -2\int (w_{j+1}d\mu - P\phi_j) w_{j+1} d\mu \\
= \; & \int \phi_j^2 d\mu + 2\int \phi_j w_j d\mu + \int w_j^2 d\mu - \int w_{j+1}^2 d\mu.
\end{aligned}
$$

Since by Lemma 2.2.2

$$
a_n^2 = \sum_{i=1}^{n} E(\phi_i^2) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \int \phi_i \circ T^i \cdot \phi_j \circ T^j d\mu = \sum_{i=1}^{n}\left(E(\phi_i^2) + 2\int (\phi_i w_i)\circ T^i d\mu \right)
$$

the statement follows by substituting $\int \psi_j^2 d\mu - \int w_j^2 d\mu + \int w_{j+1}^2 d\mu$ for the terms inside the sum on the RHS and then telescoping out the expected values of w_j^2. $\square$

2.3 Property (SP)

Several authors [30, 6] have used a property derived from the Gal-Koksma theorem (see Appendix) to prove the SBC property for sequences of balls. Later we will show that in certain settings the (SP) property also implies a CLT. Property (SP) (where SP stands for Sprindzuk Property), states that $E(S_{(m,n)}^2) \leq CE(S_{(m,n)})$ for some constant C, where $f_i \geq 0$ and $S_{(m,n)} = \sum_{i=m}^{n} f_i$, it is a condition that appears in one guise or another often in proofs of the Borel Cantelli property (which is easily deduced from this).

Suppose $\{B_i\}$ are balls and let $f_i = 1_{B_i} \circ T^i$. If

$$\sum_{i=m}^{n} \sum_{j=i+1}^{n} E(f_i f_j) - E(f_i)E(f_j) \leq C \sum_{i=m}^{n} E(f_i) \qquad (SP)$$

for arbitrary integers $n > m$ then the balls are said to have the (SP) property.

2.4 Short returns and Assumption (C)

In this section we discuss a condition on short return times first considered, to our knowledge, by P. Collet [7]. We have called it Assumption (C), after Collet.

Assumption (C): We say $(B_i(p))$ satisfies Assumption (C) if there exists $\eta(p) \in (0,1)$ and $\kappa(p) > 1$ such that for all i sufficiently large

$$\mu(B_i(p) \cap T^{-r} B_i(p)) \leq \mu(B_i(p))^{1+\eta}$$

for all $r = 1, \dots, (\log i)^\kappa$.

This condition has been used to establish extreme value statistics [7, 26, 20] and dynamical Borel-Cantelli lemmas [21, 24]. Assumption (C) is a strong control on measure of points making short returns. If Assumption (C) holds for a point p then the measure of a set of points in nested balls $B_i(p)$ about p returning to $B_i(p)$ in a time interval smaller than an integer power of $-\log \mu(B_i(p))$ is smaller than $\mu(B_i(p))^\xi$ where ξ is greater than one. Note by Kac's theorem the expected return time to $\mu(B_i(p))$ is of order $\mu(B_i(p))^{-1}$, so an integer power of $-\log \mu(B_i(p))$ is indeed a very short return. This condition fails for periodic points, as a fixed fraction of the mass of a ball returns after the period. Heuristically if the first return times to $B_i(p)$ follow an exponential law (which one somehow expects for generic points) then $\lim_{i\to\infty} \frac{1}{\mu(B_i(p))}\mu\{x \in B_i(p) : \tau(x) > \frac{t}{\mu(B_i(p))}\} \to e^{-t}$ and hence $\lim_{i\to\infty} \frac{1}{\mu(B_i(p))}\mu\{x \in B_i(p) : n(x) \leq \frac{t}{\mu(B_i(p))}\} \to 1 - e^{-t} \sim t$ (for small t). Suppose now we could solve for $\frac{t}{\mu(B_i(p))} = (-\log(\mu(B_i(p)))^k$, we would then have $\mu\{x \in B_i(p) : n(x) \leq (-\log(\mu(B_i(p)))^k\} \sim (-\log(\mu(B_i(p)))^k \mu(B_i(p))^2$. Note that our assumption $\mu(B_i) \leq \frac{C_2}{i^\gamma}$ implies that $(-\log \mu(B_i))^k \geq C(\log(i))^k$ for large i. This train of thought makes Assumption (C) seem reasonable for generic points.

If $(B_i(p))$ satisfies Assumption (C) then we can say more about the behavior of the constants a_n.

Lemma 2.4.1. *Under Assumption (C) there exists a constant C_4 and some large a so that*

$$||\phi_j w_j||_1 = \int |\phi_j w_j| \, d\mu \leq C_4 \mu(B_{j-a\log j})^{1+\eta} \log j.$$

Proof: By the contraction property of the transfer operator one has as in (2.1) for a sufficiently large constant a

$$\sum_{i<j-a\log j} \int \phi_j \cdot P^{j-i}\phi_i d\mu \leq \mu(B_{j-a\log j})^2.$$

Let $\phi_j = \tilde{\phi}_j - \int \tilde{\phi}_j d\mu$ where $\tilde{\phi}_j$ is the $\mathcal{B}_\alpha$ approximation to $1_{B_j(p)}$. Hence we obtain in the $\mathscr{L}^1$-norm: (as $\tilde{\phi}_j \geq 0$)

$$\begin{aligned}
\int |\phi_j w_j| d\mu &\leq \sum_{n=1}^{a\log j} (\int \tilde{\phi}_j P^n \tilde{\phi}_{j-n} d\mu + \int \tilde{\phi}_{j-n} \int \tilde{\phi}_j d\mu + \int \tilde{\phi}_j \int P^n \tilde{\phi}_{j-n} d\mu \\
&\quad + \int \tilde{\phi}_j \int \tilde{\phi}_{j-n} d\mu) + \mathcal{O}\left((j - a\log j)\mu(B_{j-a\log j})^2\right) \\
&= \sum_{n=1}^{a\log j} \left(\int \tilde{\phi}_j P^n \tilde{\phi}_{j-n} d\mu + 3\mu(\tilde{\phi}_{j-n})\mu(\tilde{\phi}_j) \right) + \mathcal{O}(j)\mu(B_{j-a\log j})^2 \\
&= \sum_{n=1}^{a\log j} \int \tilde{\phi}_j P^n \tilde{\phi}_{j-n} d\mu + \mathcal{O}(j)\mu(B_{j-a\log j})^2,
\end{aligned}$$

Now by Assumption (C) we have

$$\int \tilde{\phi}_j \cdot P^n \tilde{\phi}_{j-n} d\mu \leq \int \tilde{\phi}_{j-n} \circ T^n \cdot \tilde{\phi}_{j-n} d\mu \leq \mu(B_{j-n} \cap T^{-n} B_{j-n}) \leq \mu(B_{j-n})^{1+\eta},$$

for $n \leq a\log j$, and thus

$$\sum_{n=1}^{a\log j} \int \tilde{\phi}_j \cdot P^n \tilde{\phi}_{j-n} d\mu \leq c_2 a \mu(B_{j-a\log j})^{1+\eta} \log j,$$

proving the lemma. □

Lemma 2.4.2. *If $(B_i(p))$ satisfies Assumption (C) then*

$$\lim_{n\to\infty} \frac{E(\sum_{i=1}^n \phi_i \circ T^i)^2}{\sum_{i=1}^n E[\phi_i^2]} = 1$$

Proof: Rearranging the sums yields by Lemma 2.2.2

$$\begin{aligned}
E(\sum_{i=1}^n \phi_i \circ T^i)^2 &= \sum_{i=1}^n E[\phi_i^2] + 2\sum_{i=1}^{n-1} \sum_{j=i+1}^n \int \phi_i \circ T^i \cdot \phi_j \circ T^j d\mu \\
&= \sum_{i=1}^n E[\phi_i^2] + 2\sum_{j=2}^n \int w_j \phi_j d\mu
\end{aligned}$$

and hence the result follows by Lemma 2.4.1 as $\eta > 1$. □

2.5 Bounds on w_j

We now assume that $\|\phi\|_\infty \leq C\|\phi\|_\alpha$ which under our assumption on the transfer operator implies that for a mean-zero function $\phi \in \mathcal{B}$, $\|P^n\phi\|_\infty \leq C\theta^n\|\phi\|_\alpha$ for some C, $0 < \theta < 1$ independently of ϕ. For example if $\|.\|_\alpha$ were the Banach space of Hölder functions of exponent α on the unit interval then $\|\phi\|_\infty \leq C\|\phi\|_\alpha$. In the BV or quasi-Hölder norm indicator functions are bounded, and the proof that w_j is uniformly bounded is straightforward in this case; we would like to stress that the next result is obtained under the general assumption that $\mu(B_n(p)) \leq \frac{C_2}{n^\gamma}$.

Lemma 2.5.1. *Assume $\|P^n\phi\|_\infty \leq C\theta^n\|\phi\|_\alpha$ then there exists a constant C_5 such that $\|w_j\|_\infty < C_5$ for all j.*

Proof: For some $a > 0$ we can achieve

$$\sum_{j=\lfloor a\log n\rfloor}^{n} ||P^j\phi_{n-j}||_\infty \leq C \sum_{j=\lfloor a\log n\rfloor}^{n} \theta^j(n-j)^k = \mathcal{O}(n^{-2})$$

and in particular $||P^j\phi_{n-j}||_\infty = \mathcal{O}(n^{-2})$ for all $j \geq \lfloor a\log n\rfloor$ and all n. As in the previous lemma let $\tilde{\phi}_j$ be the $\mathcal{B}_\alpha$ approximation for 1_{B_j} and $\phi_j = \tilde{\phi}_j - \mu(\tilde{\phi}_j)$. In view of the tail estimate it is only necessary to bound $\sum_{j=1}^{\lfloor a\log n\rfloor} P^j\phi_{n-j}$ independently of n.

(i) Bound from below: Since $\phi_j \geq -\mu(\tilde{\phi}_j) \geq -c_2\mu(B_j) \geq -\frac{c_3}{j^\gamma}$ $(c_2, c_3 > 0)$ one obtains $\sum_{j=1}^{\lfloor a\log n\rfloor} P^j\phi_{n-j} \geq \sum_{j=1}^{\lfloor a\log n\rfloor} \frac{c_3}{(n-j)^\gamma} \geq \frac{-c_4\log n}{n^\gamma}$ for some constant c_4 independent of j and n. Hence $w_n \geq -c_5$ for some $c_5 > 0$ and all n.

(ii) Bound from above: Since $1_{B_{j+1}} \leq 1_{B_j}$ one has $\tilde{\phi}_{j+1} \leq \tilde{\phi}_j$ and in particular

$\mu(\tilde{\phi}_{j+1}) \leq \mu(\tilde{\phi}_j)$. Hence $\phi_{j+1} - \phi_j \leq \mu(\tilde{\phi}_j) - \mu(\tilde{\phi}_{j+1})$ and (as $\phi_0 = 0$)

$$\begin{aligned} w_m - w_{m-1} &= \sum_{j=1}^{m-1} P^j(\phi_{m-j} - \phi_{m-1-j}) + P^m\phi_0 \\ &\leq \sum_{j=1}^{m-1} \Big(\mu(\tilde{\phi}_{m-1-j}) - \mu(\tilde{\phi}_{m-j})\Big) \\ &\leq \sum_{j=1}^{\lfloor a\log m\rfloor} \Big(\mu(\tilde{\phi}_{m-1-j}) - \mu(\tilde{\phi}_{m-j})\Big) + \mathcal{O}(m^{-2}). \end{aligned}$$

Consequently ($w_1 = P\phi_0 = 0$)

$$\begin{aligned} w_n &= \sum_{m=2}^{n}(w_m - w_{m-1}) + w_1 \\ &\leq \sum_{m=2}^{n}\left(\sum_{j=1}^{\lfloor a\log m\rfloor} \Big(\mu(\tilde{\phi}_{m-1-j}) - \mu(\tilde{\phi}_{m-j})\Big) + \mathcal{O}(m^{-2})\right) \\ &= \sum_{j=1}^{\lfloor a\log n\rfloor} \sum_{m=2\vee\lceil e^{\frac{j}{a}}\rceil}^{n} \Big(\mu(\tilde{\phi}_{m-1-j}) - \mu(\tilde{\phi}_{m-j}) + \mathcal{O}(m^{-2})\Big) \\ &= \sum_{j=1}^{\lfloor a\log n\rfloor} \Big(\mu(\tilde{\phi}_{2\vee\lceil e^{\frac{j}{a}}\rceil - j}) - \mu(\tilde{\phi}_{n-j}) + \mathcal{O}((2\vee e^{\frac{j}{a}})^{-1})\Big) \\ &\leq c_6 \end{aligned}$$

for a constant c_6 independent of n because

$$\sum_{j=1}^{\lfloor a\log n\rfloor} \mu(\tilde{\phi}_{n-j}) \leq c_7 \frac{a\log n}{n^\gamma} \to 0$$

as $n \to \infty$ and

$$\sum_{j=1}^{\lfloor a\log n\rfloor} \mu(\tilde{\phi}_{2\vee\lceil e^{\frac{j}{a}}\rceil - j}) \leq c_8 \sum_{j=1}^{\lfloor a\log n\rfloor} (e^{\frac{j}{a}})^{-\gamma} = \mathcal{O}(1)$$

for constants c_7, c_8 independent of n. Here $a \vee b = \max\{a, b\}$. □

With all these assumptions and associated results, we will see some applications in next chapter.

Chapter 3

Applications of CLT to some Banach spaces[1]

In our applications we will have the pairs $(BV(\Omega), \mathscr{L}^1(\mu))$ or $(H_\gamma(\Omega), \mathscr{L}^1(\Omega))$ where $BV(\Omega)$ is the space of function of bounded variation and $H_\gamma(\Omega)$ is the space of Hölder functions of exponent γ. For example if T is a smooth uniformly expanding map of the unit interval X then $\mathcal{B}_\alpha$ could be taken as the Banach space of functions of bounded variation $BV(\Omega)$. In this chapter we will consider Lipschitz rather than Hölder functions, as our results and proofs immediately generalize to the Hölder setting with the obvious changes.

Definition 3.0.2 (Space of Lipschitz functions)**.** *Given a metric space (M, d), where d is the metric on M. Let $Lip(M, d)$ denote the vector space of bounded functions*

[1]**This chapter contains published work from** *Haydn, N., Nicol, M., Vaienti, S., and Zhang, L. (2013). Central limit theorems for the shrinking target problem. Journal of Statistical Physics, 153(5), 864-887. Chicago.*

$f : M \to \mathbb{R}$ such that

$$||f||_d = \sup\{\frac{|f(s) - f(t)|}{d(s-t)}, s \neq t\} < \infty$$

Let $||\cdot||_\infty$ be the sup-norm. Then $Lip(M, d)$ endowed with the norm $||\cdot||_{Lip} = \max\{||\cdot||_d, ||\cdot||_\infty\}$ is a Banach space, which is called the space of Lipschitz functions.

3.1 Decay in Lipschitz versus $\mathscr{L}^1$

We consider the space of Lipschitz functions, the arguments we gave hold for Hölder norms with obvious modification. We assume that the transfer operator P, when restricted to $Lip(\Omega)$, contracts exponentially:

$$||P^n \phi||_{Lip} \leq C\theta^n ||\phi||_{Lip} \tag{3.1}$$

for all Lipschitz functions ϕ such that $\int \phi \, d\mu = 0$ where $\theta \in (0, 1)$ and C are independent of ϕ.

This implies

$$\left| \int \phi\psi \circ T^n \, d\mu - E[\phi]E[\psi] \right| \leq C\theta^n \|\phi\|_{Lip} \, \|\psi\|_{\mathscr{L}^1} \tag{3.2}$$

for the same $\theta \in (0, 1)$ and C independent of ϕ, ψ.

Theorem 3.1.1. *Assume that the transfer operator, when restricted to $Lip(\Omega)$, contracts exponentially as in (3.1) for some $\theta \in (0, 1)$.*

Suppose $\{B_i(p)\}$ be nested balls about a point p such that $\sum_i \mu(B_i) = \infty$ and $\mu(B_i(p)) \leq \frac{C_2}{i^\gamma}$ for constants $C_2 > 0$ and $0 < \gamma \leq 1$. Let $a_n^2 = E\left(\sum_{i=1}^{n}(1_{B_i} \circ T^i - \mu(B_i))\right)^2$.

(I) If the nested sequence of balls $(B_i(p))$ satisfies Assumption (C) and the SBC property then

$$\frac{1}{a_n}\sum_{i=1}^{n}\big(1_{B_i}\circ T^i-\mu(B_i)\big)\to N(0,1)$$

in distribution.

(II) If $(B_i(p))$ has the (SP) property then

$$\frac{1}{a_n}\sum_{i=1}^{n}(1_{B_i}\circ T^i-\mu(B_i))\to N(0,1).$$

Proof: We will let $\phi_j=\tilde{\phi}_j-\int\tilde{\phi}_j$,where $\tilde{\phi}_j$ be a Lipschitz approximation to 1_{B_j}, such that

(i) $\|\tilde{\phi}_j-1_{B_j}\|_1<\frac{1}{j^3}$

(ii) $\|\tilde{\phi}_j\|_{Lip}\leq Cj^k$

(iii) $\tilde{\phi}_j\geq 0$

Remark 3.1.2. *If we are taking a Hölder approximation then condition (ii) is satisfied for the balls $B_i=B(p,r_i)$ if there exists $\delta(p)>0$ and $C>0$ such that $\mu\{x:r<d(x,p)<r+\epsilon\}<C\epsilon^{\delta(p)}$. This condition is satisfied if the invariant measure μ has a density h with respect to Lebesgue measure m such that $h\in\mathscr{L}^{1+\eta}(m)$ for some $\eta>0$.*

We define $w_n=P\phi_{n-1}+P^2\phi_{n-2}+\ldots+P^n\phi_0$ and put $\psi_n=\phi_n-w_{n+1}\circ T+w_n$. Then $P\psi_n=P\phi_n-w_{n+1}+\sum_{j=2}^{n}P^j\phi_{n-j+1}=0$ which corresponds to $\int\psi_n\chi\circ T\,d\mu=\int\chi P\psi_n\,d\mu=0$ for any integrable function χ. Note that $||\phi_j||_\infty\leq||\phi_j||_{Lip},||\phi_j||_1\leq||\phi_j||_{Lip}$.

Lemma 3.1.3. *There exist constants C_6, k so that*

(i) $\|w_n\|_{Lip} \leq C_6 n^k$,

(ii) $\|w_n\|_\infty \leq C_6$,

(iii) $\|w_n\|_1 \leq C_6 \frac{\log n}{n^\gamma}$.

Proof of Lemma 3.1.3.

(i) By the contraction of the transfer operator for Lipschitz continuous functions one obtains

$$\|w_n\|_{Lip} \leq \sum_{j=1}^{n} \|P^j \phi_{n-j}\|_{Lip} \leq \sum_{j=1}^{n} C_1 \theta^j \|\phi_{n-j}\|_{Lip} \leq c_1 n^k$$

(ii) Is a consequence of Lemma 2.5.1.

(iii) For sufficiently large a we get

$$\begin{aligned}
||w_n||_1 &\leq \sum_{j=1}^{a\log n} ||P^j \phi_{n-j}||_1 + \sum_{j=a\log n+1}^{n} ||P^j \phi_{n-j}||_1 \\
&\leq \sum_{j=1}^{a\log n} ||\phi_{n-j}||_1 + \sum_{j=a\log n+1}^{n} ||P^j \phi_{n-j}||_{Lip} \\
&\leq \sum_{j=1}^{a\log n} ||\phi_{n-j}||_1 + \sum_{j=a\log n+1}^{n} C_1 \theta^j ||\phi_{n-j}||_{Lip} \\
&\leq C(a\log n)\mu(B_{n-a\log n}) + c_4 \frac{\log^2 n}{n^2} \\
&\leq c_5 \frac{\log n}{n^\gamma}
\end{aligned}$$

for some c_4, c_5 independent of n.

Now put $C_6 = \max(c_1, c_5)$.

□

As before let $(\tilde{\Omega}, \tilde{\sigma}, \tilde{\mu})$ be the natural extension of (Ω, T, μ) and put $a_n^2 = E(\sum_{j=1}^{n} \phi_j \circ T^j)^2$ for the rescaling factors where the ψ_j lift to $\tilde{\Omega}$ in a natural way. Again we put $X_{n,i} = \frac{1}{a_n}\psi_{n-i} \circ \tilde{\sigma}^{-i}$, $i = 1, \dots, n$ which are $\mathcal{F}_i = \tilde{\sigma}^i \mathcal{B}_0$ measurable where $\mathcal{B}_0$ is the σ-algebra $\mathcal{B}$ lifted to Ω. The $\mathcal{F}_i$ form an increasing sequence of σ-algebras. We put $S_{n,i} = \sum_{j=1}^{i} X_{n,j}$, $i = 1, \dots, n$ $(k_n = n)$, and obtain $E[S_{n,i+1}|\mathcal{F}_i] = S_{n,i} + E[X_{n,i+1}|\mathcal{F}_i]$ but by stationarity $E[X_{n,i+1}|\mathcal{F}_i] = E[\phi_{n-i-1}|\tilde{\sigma}\mathcal{B}] = 0$. Hence $E[S_{n,i+1}|\mathcal{F}_i] = S_{n,i}$ and $X_{n,i}$ is a martingale difference array with respect to $\mathcal{F}_i$.

Condition (d) clearly holds; instead conditions (a) and (c) simply follow since $||\phi_n||_\infty$ is bounded and a_n tends to infinity.

We now prove (I) and show that under Assumption (C), $\sum_{i=1}^{n} X_{n,i}^2 \to 1$ in probability and hence condition (b) holds.

Lemma 3.1.4.
$$\frac{1}{a_n^2}\sum_{j=1}^{n} \psi_j^2 \circ T^j \to 1$$
in probability as $n \to \infty$.

Proof. We follow an argument given by Peligrad [35]. As $\psi_j = \phi_j + w_j - w_{j+1} \circ T$ we obtain

$$\begin{aligned}
\psi_j^2 &= \phi_j^2 + 2\phi_j w_j + w_j^2 + w_{j+1}^2 \circ T - 2w_{j+1} \circ T(\phi_j + w_j) \\
&= \phi_j^2 + 2\phi_j w_j + w_j^2 + w_{j+1}^2 \circ T - 2w_{j+1} \circ T(\psi_j + w_{j+1} \circ T) \\
&= \phi_j^2 + (w_j^2 - w_{j+1}^2 \circ T) - 2\psi_j w_{j+1} \circ T + 2\phi_j w_j.
\end{aligned}$$

We want to sum over $j = 1, \ldots, n$ and normalize by $\log n$ and wish to estimate the error terms which are the last four terms on the RHS. The terms $w_j^2 - w_{j+1}^2 \circ T$ are bounded and telescope so may be neglected.

In order to estimate the third of the error terms, $\psi_j w_{j+1} \circ T$ we proceed like Peligrad ([35], page 9) using a truncation argument. Let $w_j^\epsilon = w_j 1_{\{|w_j| \le \epsilon a_n\}}$, where for simplicity of notation we have left out the dependence on n. Then

$$\int \left(\sum_{j=1}^n \psi_j \circ T^j w_{j+1}^\epsilon \circ T^{j+1} \right)^2 = \sum_{j=1}^n \int \left(\psi_j \circ T^j w_{j+1}^\epsilon \circ T^{j+1} \right)^2 \le \epsilon^2 a_n^2 \sum_{j=1}^n \int \psi_j^2$$

since the cross terms vanish (for $j > i$), as

$$\begin{aligned} \int (\psi_j w_{j+1}^\epsilon \circ T) \circ T^j (\psi_i w_{i+1}^\epsilon \circ T) \circ T^i &= \int (\psi_j w_{j+1}^\epsilon \circ T) \circ T^{j-i} (\psi_i w_{i+1}^\epsilon \circ T) \\ &= \int (\psi_j w_{j+1}^\epsilon \circ T) \circ T^{j-i-1} P(\psi_i w_{i+1}^\epsilon \circ T) \\ &= \int (\psi_j w_{j+1}^\epsilon \circ T) \circ T^{j-i-1} w_{i+1}^\epsilon P\psi_i = 0 \end{aligned}$$

as $P(\psi_i w_{i+1}^\epsilon \circ T) = w_{i+1}^\epsilon P\psi_i$.

For any $a > \epsilon$ we obtain using Tchebycheff's inequality (on the second term):

$$\begin{aligned} &\mu \left(\left| \frac{1}{a_n^2} \sum_{j=1}^n \psi_j \circ T^j w_{j+1} \circ T^{j+1} \right| > a \right) \\ \le\ & \mu \left(\max_{1 \le j \le n} \left| w_{j+1} \circ T^{j+1} \right| > \epsilon a_n \right) + \mu \left(\left| \frac{1}{a_n^2} \sum_{j=1}^n \psi_j \circ T^j w_{j+1}^\epsilon \circ T^{j+1} \right| > a \right) \\ \le\ & \mu(\max_{1 \le j \le n} |w_{j+1} \circ T^{j+1}| > \epsilon a_n) + \frac{\epsilon^2}{a^2 a_n^2} \sum_{j=1}^n \int \psi_j^2 \\ =\ & \mu(\max_{1 \le j \le n} |w_j \circ T^{j+1}| > \epsilon a_n) + c_1 \frac{\epsilon^2}{a^2}. \end{aligned}$$

In the last line we used $\sum_{j=1}^n E[\psi_j^2] \sim a_n$ by Lemma 2.2.3 and 2.4.1. By boundedness of the w_j (Lemma 2.5.1) one gets that $P(\max_{1 \le j \le n} |w_{j+1} \circ T^{j+1}| > \epsilon a_n) \to 0$ for every $\epsilon > 0$ as $n \to \infty$. Choosing $a = \epsilon^{\frac{1}{2}}$ we conclude that $\frac{1}{a_n^2} \sum_{j=1}^{n} \psi_j \circ T^j w_{j+1} \circ T^{j+1}$ converges to zero in probability as $n \to \infty$.

For the fourth error term $\frac{1}{a_n^2} 2 \sum_{j=1}^{n} (\phi_j w_j) \circ T^j$ we obtain by Lemma 2.4.2:

$$\left\| \sum_{j=1}^{n} (\phi_j w_j) \circ T^j \right\|_1 \leq \sum_{j=1}^{n} \|\phi_j w_j\|_1 \leq c_2 \sum_{j=1}^{\infty} \mu(B_{j-a\log j})^{1+\eta} \log j$$

Thus $\frac{2}{a_n^2} \sum_{j=1}^{n} (\phi_j w_j) \circ T^j \to 0$ in probability.

Since the term $\frac{1}{a_n^2} \sum_{j=1}^{n} \phi_j^2 \circ T^j$ converges to 1 almost surely by the SBC property and Lemma 2.4.2, that is,

$$\frac{\sum_{j=1}^{n} \phi_j^2 \circ T^j}{a_n^2} = \frac{\sum_{j=1}^{n} \phi_j^2 \circ T^j}{\sum_{j=1}^{n} E(\phi_j^2)} \cdot \frac{\sum_{j=1}^{n} E(\phi_j^2)}{a_n^2} \to 1$$

the proof is complete. $\square$

Lemma 3.1.4 completes the proof of part (I) of the theorem. In order to show (II) we proceed as in the proof of (I) except for the verification of condition (b). We will prove an SBC property for $\phi_j^2 + 2w_j\phi_j$. Decomposing $\phi_j = \tilde{\phi}_j - \mu(\tilde{\phi}_j)$ and defining $\tilde{w}_j = P\tilde{\phi}_{j-1} + ... + P^{[a\log j]}\tilde{\phi}_{j-[a\log j]}$ we see that $\|\phi_j^2 - \tilde{\phi}_j^2\|_1 \leq C\mu(B_j)^2$ and $\|w_j\phi_j - \tilde{w}_j\tilde{\phi}_j\|_1 \leq C\mu(B_{j-a\log j})^2 \log j$. Note that both $\tilde{w}_j$ and $\tilde{\phi}_j$ are positive functions and moreover that similarly to Lemma 3.1.3 there exists a constant $\tilde{C}_6$ such that $\|\tilde{w}_j\|_\infty \leq \tilde{C}_6$ and $\|\tilde{w}_j\|_{Lip} \leq \tilde{C}_6 j^k$. Let $\mathcal{E}_n := \sum_{j=1}^{n} E[\tilde{\phi}_j^2 + 2\tilde{w}_j\tilde{\phi}_j]$. It suffices to consider the sequence $\tilde{\phi}_j^2 + 2\tilde{w}_j\tilde{\phi}_j$. This is because $\sum_{j=1}^{n} E[\tilde{\phi}_j^2 + 2\tilde{w}_j\tilde{\phi}_j] =$

$\sum_{j=1}^{n} E[\phi_j^2 + 2w_j\phi_j] + \sum_{j=1}^{n}(C\mu(B_j)^2 + C\mu(B_{j-a\log j})^2 \log j)$ and hence μ almost surely,

$$\frac{1}{\mathcal{E}_n}\sum_{j=1}^{n}(\tilde{\phi}_j^2 \circ T^j + 2(\tilde{w}_j\tilde{\phi}_j) \circ T^j) = \frac{1}{\mathcal{E}_n}\sum_{j=1}^{n}(\phi_j^2 \circ T^j + 2(w_j\phi_j) \circ T^j).$$

We will use Proposition A.1.1 in Appendix, a form of the Gal and Koksma theorem as stated by Sprindzuk to show that $\frac{1}{\mathcal{E}_n}\sum_{j=1}^{n}\tilde{\phi}_j^2 \circ T^j + 2(\tilde{w}_j\tilde{\phi}_j) \circ T^j \to 1$ almost surely. For this we want to use Proposition A.1.1 with $f_j = \tilde{\phi}_j^2 + 2\tilde{w}_j\tilde{\phi}_j$, $g_j = \int f_j$ and h_j to be determined below. We need to estimate the terms in

$$\int \left(\sum_{i=m}^{n}(\tilde{\phi}_j^2 + 2\tilde{w}_j\tilde{\phi}_j - \int(\tilde{\phi}_j^2 + 2\tilde{w}_j\tilde{\phi}_j)d\mu)\right)^2 d\mu$$

.

In order to verify the condition of the proposition we look at the three individual sums of terms $\int \tilde{\phi}_j \circ T^{j-i}\tilde{\phi}_i$, $\int(\tilde{\phi}_j\tilde{w}_j) \circ T^{j-i}\tilde{\phi}_i\tilde{w}_i$ and $\int(\tilde{\phi}_j\tilde{w}_j) \circ T^{j-i}\tilde{\phi}_i$ as follows:
(i) The fact that condition (SP) holds for the functions $\tilde{\phi}_j$ implies

$$\sum_{i=m}^{n}\sum_{j=i+1}^{n}\int \tilde{\phi}_j \circ T^{j-i} \cdot \tilde{\phi}_i d\mu - E[\tilde{\phi}_j]E[\tilde{\phi}_i] \leq c_1 \sum_{i=m}^{n} E[\tilde{\phi}_j]. \tag{3.3}$$

Since $E(\tilde{\phi}_j^2) - E(\tilde{\phi}_j) = \mathcal{O}(j^{-k})$ we obtain

$$\sum_{i=m}^{n}\sum_{j=i+1}^{n}\int \tilde{\phi}_j^2 \circ T^{j-i} \cdot \tilde{\phi}_i^2 d\mu - E[\tilde{\phi}_j^2]E[\tilde{\phi}_i^2] \leq c_1 \sum_{i=m}^{n} E[\tilde{\phi}_j] + \sum_{i=m}^{n}\mathcal{O}(i^{-k+1}).$$

(ii) Here we estimate the sums of the terms $\int(\tilde{\phi}_j\tilde{w}_j) \circ T^{j-i} \cdot (\tilde{\phi}_i\tilde{w}_i)d\mu - E[\tilde{\phi}_j\tilde{w}_j]E[\tilde{\phi}_i\tilde{w}_i]$.
By exponential decay of correlations one has for some constant $\beta > 0$

$$\begin{aligned}\sum_{j=i+1}^{n}&\left(\int \tilde{\phi}_j \circ T^{j-i} \cdot \tilde{\phi}_i d\mu - E[\tilde{\phi}_j]E[\tilde{\phi}_i]\right) \\ &= \sum_{j=i+1}^{i+\beta\log i}\left(\int \tilde{\phi}_j \circ T^{j-i} \cdot \tilde{\phi}_i d\mu - E[\tilde{\phi}_j]E[\tilde{\phi}_i]\right) + \mathcal{O}(i^{-k})\end{aligned}$$

and therefore

$$\sum_{j=i+1}^{i+\beta\log i}\int\tilde{\phi}_j\circ T^{j-i}\cdot\tilde{\phi}_i d\mu$$
$$\leq \sum_{j=i+1}^{i+\beta\log i}E[\tilde{\phi}_j]E[\tilde{\phi}_i]+\mathcal{O}(i^{-k})+\sum_{j=i+1}^{n}\left(\int\tilde{\phi}_j\circ T^{j-i}\cdot\tilde{\phi}_i d\mu-E[\tilde{\phi}_j]E[\tilde{\phi}_i]\right)$$

Using the uniform bound on $||\tilde{w}_j||_\infty$ this implies

$$\sum_{j=i+1}^{i+\beta\log i}\int(\tilde{\phi}_j\tilde{w}_j)\circ T^{j-i}\cdot(\tilde{\phi}_i\tilde{w}_i)d\mu$$
$$\leq \tilde{C}_6^2\left(\sum_{j=i+1}^{i+\beta\log i}E[\tilde{\phi}_j]E[\tilde{\phi}_i]+\mathcal{O}(i^{-k})+\sum_{j=i+1}^{n}\left(\int\tilde{\phi}_j\circ T^{j-i}\cdot\tilde{\phi}_i d\mu-E[\tilde{\phi}_j]E[\tilde{\phi}_i]\right)\right)$$

By decay of correlation and since $\|\tilde{w}_j\tilde{\phi}_j\|_{Lip}\leq\tilde{C}_6 j^k$ we get with a possibly larger β that

$$\begin{aligned}
&\sum_{j=i+1}^{n}\int(\tilde{\phi}_j\tilde{w}_j)\circ T^{j-i}\cdot(\tilde{\phi}_i\tilde{w}_i)d\mu-E[\tilde{\phi}_j\tilde{w}_j]E[\tilde{\phi}_i\tilde{w}_i]\\
=&\sum_{j=i+1}^{i+\beta\log i}\left(\int(\tilde{\phi}_j\tilde{w}_j)\circ T^{j-i}\cdot(\tilde{\phi}_i\tilde{w}_i)d\mu-E[\tilde{\phi}_j\tilde{w}_j)]E[\tilde{\phi}_i\tilde{w}_i]\right)+\mathcal{O}(i^{-k})\\
\leq&\tilde{C}_6^2\sum_{j=i+1}^{i+\beta\log i}E[\tilde{\phi}_j]E[\tilde{\phi}_i]+\tilde{C}_6^2\sum_{j=i+1}^{n}\left(\int\tilde{\phi}_j\circ T^{j-i}\cdot(\tilde{\phi}_i)d\mu-E[\tilde{\phi}_j]E[\tilde{\phi}_i]\right)+\mathcal{O}(i^{-k})\\
\leq&\tilde{C}_6^2\left(E[\tilde{\phi}_i]^2\beta\log i+\sum_{j=i+1}^{n}\left(\int\tilde{\phi}_j\circ T^{j-i}\cdot(\tilde{\phi}_i)d\mu-E[\tilde{\phi}_j]E[\tilde{\phi}_i]\right)\right)+\mathcal{O}(i^{-\tilde{k}})
\end{aligned}$$

for some $\tilde{k}>1$. Since $E[\tilde{\phi}_i]^2\log i\leq c_2E[\tilde{\phi}_i]\ \forall i$ and some c_2 and since for every m,n inequality (3.3) holds we now obtain

$$\sum_{i=m}^{n}\sum_{j=i+1}^{n}\left(\int(\tilde{\phi}_j\tilde{w}_j)\circ T^{j-i}\cdot(\tilde{\phi}_i\tilde{w}_i)d\mu-E[\tilde{\phi}_j\tilde{w}_j]E[\tilde{\phi}_i\tilde{w}_i]\right)\leq c_3\sum_{i=m}^{n}E[\tilde{\phi}_i]+\mathcal{O}(i^{-\tilde{k}})$$

(iii) A similar argument shows that for the 'mixed' terms

$$\sum_{i=m}^{n}\sum_{j=i+1}^{n}\int(\tilde{\phi}_j\tilde{w}_j)\circ T^{j-i}\cdot\tilde{\phi}_i^2 d\mu - E[\tilde{\phi}_j\tilde{w}_j]E[\tilde{\phi}_i^2] \le c_4\sum_{i=m}^{n}E[\tilde{\phi}_j] + \sum_{i=m}^{n}\left(\mathcal{O}(i^{-\tilde{k}})\right).$$

Combining (i), (ii) and (iii) yields for all $m < n$ and some constant c_5:

$$\int\left(\sum_{i=m}^{n}\left(\tilde{\phi}_j^2 + 2\tilde{w}_j\tilde{\phi}_j - \int(\tilde{\phi}_j^2 + 2\tilde{w}_j\tilde{\phi}_j)d\mu\right)\right)^2 \le c_5\sum_{i=m}^{n}\left(E[\tilde{\phi}_i] + \mathcal{O}(i^{-\tilde{k}})\right).$$

We choose $h_i = E[\tilde{\phi}_i] + \mathcal{O}(i^{-\tilde{k}})$, and so Proposition A.1.1 implies that

$$\frac{1}{E_n}\sum_{j=1}^{n}\left(\tilde{\phi}_j^2\circ T^j + 2(\tilde{w}_j\tilde{\phi}_j)\circ T^j\right)\to 1$$

almost surely, provided $k \ge 2$. □

3.2 Applications to dynamical systems.

Theorem 3.1.1 applies to a variety of dynamical systems including Gibbs-Markov maps [1] and rational maps [23]. For Gibbs-Markov maps it has been shown [21, Theorem 1] that nested sequences of balls $(B_i(p))$ satisfy both the Strong Borel Cantelli property and Assumption (C), so that (I) applies. For rational maps [23, Theorem 10] shows that the transfer operator contracts exponentially in the $\mathscr{L}^\infty$ norm hence if the (SP) property is also proved then (II) holds. More generally (II) shows that proving the (SP) property for systems whose associated transfer operator has exponential decay suffices to prove the SBC property and the CLT for shrinking targets.

3.3 Decay in $BV(\Omega)$ versus $\mathscr{L}^1$

It is known that summable decay of correlations in $BV(\Omega)$ versus $\mathscr{L}^1$ implies the (SP) property by work of Kim [29, Proof of Theorem 2.1] (see also Gupta et al [21, Proposition 2.6]). Hence the statement in this setting is simpler.

We assume that the restriction of P to the space $BV(\Omega)$ is exponentially contracting, i.e. $P : BV(\Omega) \to BV(\Omega)$ satisfies

$$\|P^n \phi\|_{BV} \leq C\theta^n \|\phi\|_{BV} \tag{3.4}$$

for all $\phi \in BV(\Omega)$ such that $\int \phi \, d\mu = 0$.

This implies that (Ω, T, μ) has exponential decay of correlations in BV versus $\mathscr{L}^1$, so that for some $0 < \theta < 1$,

$$\left| \int \phi \psi \circ T^n \, d\mu - \int \phi \, d\mu \int \psi \, d\mu \right| \leq C\theta^n \|\phi\|_{BV} \|\psi\|_1 \tag{3.5}$$

for all $\phi \in BV(\Omega)$, $\psi \in \mathscr{L}^1(\mu)$. In particular the measure μ is ergodic.

Proposition 3.3.1. *Assume the transfer operator P contracts exponentially as given by (3.4)*

Let $B_i := B(p, r_i)$ be nested balls of radius r_i about a point p such that $\sum_i \mu(B_i) = \infty$ and $\mu(B_i(p)) \leq \frac{C_2}{i^\gamma}$ for constants $C_2 > 0$ and $0 < \gamma \leq 1$. Let $a_n^2 = E(\sum_{j=1}^{n} (1_{B_i} \circ T^i - \mu(B_i)))^2$. Then:

$$\frac{1}{a_n} \sum_{j=1}^{n} (1_{B_i} \circ T^i - \mu(B_i)) \to N(0, 1).$$

Proof: The proof is the same as for Theorem 3.1.1 with the simplification that the (SP) property holds automatically as we have summable decay of correlations in $BV(\Omega)$ versus $\mathscr{L}^1$ (see proof of [29, Theorem 2.1]). $\square$

Remark 3.3.2. *For one-dimensional maps of the interval, Proposition 3.3.1 is basically a consequence of Conze and Raugi [8, Theorem 5.1]. Follow the proof of [8, Theorem 5.1] taking $T_k = T$ for all k, m to be the invariant measure μ and choosing $f_n = 1_{B_n}(p)$. The rates of growth are given by Lemma 2.2.1 which shows that the variance is unbounded. Lemma 2.4.2 gives a precise rate of growth in the case that Assumption (C) holds. In Proposition 3.4.2 we extend these results to piecewise expanding maps in higher dimensions.*

3.4 Applications of Proposition 3.3.1.

Proposition 3.3.1 applies to certain classes of one-dimensional maps such as piecewise expanding maps of the interval $T : \Omega \to \Omega$ with $\frac{1}{T'}$ of bounded variation and possessing an absolutely continuous invariant measure with density bounded away from zero (those maps satisfying the assumptions of [29, Theorem 2.1], see also [21]). For these systems, Assumption (C) has been shown to hold for nested balls about μ a.e. $p \in X$ [26, 20]. In the next subsection we generalize these results to piecewise expanding maps in higher dimensions.

3.4.1 Piecewise expanding maps in higher dimensions

In this section we prove the Strong Borel Cantelli property and the CLT for shrinking balls in a class of expanding maps in higher dimensions. We also show that Assumption (C) holds for μ-a.e. point.

The Banach spaces will be given by $\mathscr{L}^1$, defined with respect to the Lebesgue measure on $\mathbb{R}^n$, and a quasi-Hölder space with properties analogous to BV which we define below. A key property of the quasi-Hölder space is that characteristic functions of balls have bounded norm (as in the BV norm) which turns out to be a very useful property.

The maps are defined on compact sets $Z \in \mathbb{R}^N$. Denote by $\operatorname{dist}(\cdot, \cdot)$ the usual metric in $\mathbb{R}^N$ and for $\varepsilon > 0$ let $B_\varepsilon(x) = \{y \in \mathbb{R}^N : \operatorname{dist}(x, y) < \varepsilon\}$ be the ε-ball centered at x. Let $B_\varepsilon(A) = \{y \in \mathbb{R}^N : \operatorname{dist}(y, A) \leq \varepsilon\}$ and write Z° for the interior of Z and $\overline{Z}$ its closure.

A map $T : Z \to Z$ is said to be a multidimensional piecewise expanding map, if there exists a family of finitely many disjoint open sets $\{Z_i\}$ such that $\operatorname{Leb}(Z \setminus \bigcup_i Z_i) = 0$ and there exist open sets $\widetilde{Z}_i \supset \overline{Z_i}$ and $C^{1+\alpha}$ maps $T_i : \widetilde{Z}_i \to \mathbb{R}^N$ (for some $0 < \alpha \leq 1$) and some sufficiently small real number $\varepsilon_1 > 0$ such that for all i,

- (H1) $T_i(\widetilde{Z}_i) \supset B_{\varepsilon_1}(T(Z_i))$ and $T_i|_{Z_i} = T|_{Z_i}$;
- (H2) For $x, y \in T(Z_i)$ with $\operatorname{dist}(x, y) \leq \varepsilon_1$,

$$|\det DT_i^{-1}(x) - \det DT_i^{-1}(y)| \leq c|\det DT_i^{-1}(x)|\operatorname{dist}(x, y)^\alpha;$$

- (H3) There exists $s = s(T) < 1$ such that $\forall x, y \in T(\widetilde{Z}_i)$ with $\mathrm{dist}(x, y) \leq \varepsilon_1$, we have

$$\mathrm{dist}(T_i^{-1}x, T_i^{-1}y) \leq s\,\mathrm{dist}(x, y).$$

- (H4) Let $G(\varepsilon) := \sup_x G(x, \varepsilon, \varepsilon_1)$ where

$$G(x, \varepsilon, \varepsilon_1) := \sum_i \frac{\mathrm{Leb}(T_i^{-1} B_\varepsilon(\partial T Z_i) \cup B_{(1-s)\varepsilon_1}(x))}{\mathrm{Leb}(B_{(1-s)\varepsilon_1}(x))} \tag{3.6}$$

and assume that

$$\sup_{\delta \leq \varepsilon_1} \left(s^\alpha + 2 \sup_{\varepsilon \leq \delta} \frac{G(\varepsilon)}{\varepsilon^\alpha} \delta^\alpha \right) < 1^2 \tag{3.7}$$

We now introduce the Banach space of quasi-Hölder functions in which the spectrum of the Perron-Frobenius operator P is investigated. Given a Borel set $\Gamma \subset Z$, we define the oscillation of $\varphi \in \mathscr{L}^1(\mathrm{Leb})$ over Γ as

$$\mathrm{osc}(\varphi, \Gamma) := \operatorname*{ess\,sup}_\Gamma \varphi - \operatorname*{ess\,inf}_\Gamma \varphi.$$

The function $x \mapsto \mathrm{osc}(\varphi, B_\varepsilon(x))$ is measurable (see [28, Proposition 3.1]) For $0 < \alpha \leq 1$ and $\varepsilon_0 > 0$, we define the α-seminorm of φ as

$$|\varphi|_\alpha = \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^{\mathbb{N}}} \mathrm{osc}(\varphi, B_\varepsilon(x))\,\mathrm{dLeb}(x).$$

Let us consider the space of functions with bounded α-seminorm

$$V_\alpha = \{\varphi \in \mathscr{L}^1(\mathrm{Leb}) : |\varphi|_\alpha < \infty\},$$

and endow V_α with the norm

$$\|\cdot\|_\alpha = \|\cdot\|_1 + |\cdot|_\alpha$$

which makes it into a Banach space. We note that V_α is independent of the choice of ε_0 and that V_α is continuously injected in $\mathscr{L}^\infty(\mathrm{Leb})$. According to [37, Theorem 5.1], there exists an absolutely continuous invariant probability measure (a.c.i.p.) μ, with density bounded above, and bounded below from zero, which has exponential decay of correlations against $\mathscr{L}^1$ observables on the finitely many mixing components of V_α: in view of the next Theorem 3.4.4 we will from now restrict ourselves to one of those components, by taking a mixing iterate of T. More precisely, if the map T is as defined above and if μ is the mixing a.c.i.p., then there exist constants $C < \infty$ and $\vartheta < 1$ such that

$$\Big| \int_Z \psi \circ T^n \, h \, \mathrm{d}\mu - \int \psi \mathrm{d}\mu \int h \mathrm{d}\mu \Big| \leq C \|\psi\|_{\mathscr{L}^1} \|h\|_\alpha \vartheta^n \tag{3.8}$$

for all $\psi \in \mathscr{L}^1$ and for all $h \in V_\alpha$. Moreover $\|P^n \phi\|_\alpha \leq C\|\phi\|_\alpha$ for all $\phi \in V_\alpha$ and thus equation 3.4 holds.

We now show that characteristic functions of balls are bounded in the $\|\cdot\|_\alpha$ norm.

Lemma 3.4.1. *Let $B_i(p)$ be a nested sequence of balls about a point $p \in \Omega$, then there exists a constant $D(\alpha)$ such that*

$$\|1_{B_i}\|_\alpha \leq D(\alpha)$$

for all i.

Proof: Take any set A with a rectifiable boundary. If p is not in a 2ϵ neighborhood of the boundary of A, then the oscillation is zero, otherwise it is 1. Therefore we have $\int osc(1_A, B_\epsilon(p)) \, \mathrm{dLeb}(p) \leq c_1 \epsilon$. Then we must divide by ϵ^α. As $\alpha \leq 1$ we have the ratio bounded by $c_1 * \epsilon^{1-\alpha}$. $\square$

The boundedness of the characteristic functions in the $\|\cdot\|_\alpha$-norm allows us to proceed as in Proposition 3.3.1 (see also [8]) and to obtain the following result.

Proposition 3.4.2. *Assume a piecewise expanding map T on a compact set $Z \subset \mathbb{R}^n$ satisfies conditions (H1)–(H4) and is mixing with respect to its absolutely continuous invariant measure μ. Let $B_i := B(p, r_i)$ be balls of radius r_i about a point p such that $\sum_i \mu(B_i) = \infty$ and $\mu(B_i) \leq \frac{C_2}{i^\gamma}$ for some constants C_2 and $\gamma \in (0,1]$. Then the variance $a_n^2 := E[(\sum\limits_{j=1}^{n}(1_{B_i} \circ T^i - \frac{1}{i}))^2]$ satisfies*

$$\limsup_{n\to\infty} \frac{a_n}{\sqrt{\log n}} \geq 1$$

and

$$\frac{1}{\sqrt{a_n}} \sum_{j=1}^{n} (1_{B_i} \circ T^i - \frac{1}{i}) \to N(0,1)$$

in distribution.

Proof: The SBC property (I) is immediate from the decay of correlations, Equation 3.8 and the bound $\|1_{B_i}\|_\alpha \leq D(\alpha)$ by the proof of Proposition 3.4.1. The growth estimate follows from Lemma 2.4. □

We now make an additional assumption. Suppose that we have M domains of local injectivity for the map T; if we take the join $\mathcal{Z}^j := \bigvee_{i=0}^{j-1} T^{-i}\mathcal{Z}$, where $\mathcal{Z}$ denotes the partition, mod-0, into the closed sets $\overline{Z_i}, i = 1, \cdots, M$, then on each element $Z_l^{(j)}, l = 1, \cdots, |\mathcal{Z}^j|$, each of which is the closure of its interior, the map T^j is injective and of class $C^{1+\alpha}$ on an open neighborhood of $Z_l^{(j)}$: we call $\tilde{Z}_l^{(j)}$ such an extension. In order to prove condition (C) we require a further assumption which is also called the *finite range structure*. We assume:

- (H5) Let $\mathcal{U}^{(j)} := \{T^j Z_l^{(j)}, \forall l = 1, \cdots, |\mathcal{Z}^j|\}$, and put $\mathcal{U} = \cup_{j=1}^{\infty} \mathcal{U}^{(j)}$. Then $\mathcal{U}$ consists of only finitely many subsets of Z with positive Lebesgue measure, hence $U_m = \inf_{U \in \mathcal{U}} m(U)$ is bounded below.

Lemma 3.4.3. *Under the assumptions (H1)–(H5) Assumption (C) is satisfied.*

Proof. Denote

$$\mathscr{E}_k(\varepsilon) := \{x; \mathrm{dist}(T^k x, x) \leq \varepsilon\}.$$

By Lemma A.3.1 (see Appendix) it is enough to prove that there exists $C > 0$, $\delta > 0$ such that for all k and ε,

$$\mu(\mathscr{E}_k(\varepsilon))) < C\varepsilon^{\delta}.$$

We now fix j and consider the cylinder, say, $Z_l^{(j)}$. Let us suppose that $\{z_k\}_{k\geq 1}$ is a sequence of points in $Z_l^{(j)}$ converging to $x \in Z_l^{(j)}$, namely $\mathrm{dist}(z_k, x) \to 0$ when $k \to \infty$, and that $\mathrm{dist}(T^j(z_k), x) \to 0$ for $k \to \infty$. With abuse of definition we say that such a point x is *fixed*. If there are points in the sequence $\{z_k\}_{k\geq 1}$ which are on the boundary of $Z_l^{(j)}$, we think of T^j as its $C^{1+\alpha}$ extension on $\tilde{Z}_l^{(j)}$. We want to show that in $\tilde{Z}_l^{(j)}$ there is only one fixed point x. By contradiction, suppose y is another fixed point and $\{w_k\}_k$ a sequence converging to y and whose T^j images converge to y as well. Suppose that $\tilde{Z}_l^{(j)}$ is a convex set in such a way the segment $[x, y]$ is contained in $\tilde{Z}_l^{(j)}$ [3]. We now fix η small enough and take k big enough and such that $\mathrm{dist}(x, z_k)$, $\mathrm{dist}(x, T^j(z_k))$, $\mathrm{dist}(y, w_k)$, $\mathrm{dist}(y, T^j(w_k))$, are all smaller than η. We also put $D_{m,j} := \inf\{||DT^j(x)||\} > 1$, where the inf is taken over the points x where

[3] If not we could join x and y with a chain of segments contained each in $\tilde{Z}_l^{(j)}$: the argument will work again since the sum of the lengths of those segments is larger than the distance between x and y and this is what we need in bounding from below.

the derivative is defined. The norm is the operator norm, which is strictly larger than 1 since the map is uniformly expanding. Then we have

$$\mathrm{dist}(x,y) \geq \mathrm{dist}(T^j(z_k), T^j(w_k)) - \mathrm{dist}(x, T^j(z_k)) - \mathrm{dist}(y, T^j(w_k))$$

and by applying Taylor's formula

$$\mathrm{dist}(x,y) \geq D_{m,j}\mathrm{dist}(z_k, w_k) - 2\eta \geq D_{m,j}[\mathrm{dist}(x,y) - 2\eta] - 2\eta$$

which gives a contradiction, since $D_{m,j} > 1$, by sending η to 0. Hence x is the only fixed point.

Let us now take a measurable set $V \subset \tilde{Z}_l^{(j)}$ containing the fixed point $x \in \tilde{Z}_l^{(j)}$. We require that the diameter of the image $T^j(V)$ be at most ε; such an image will therefore be contained in the ball of center $T^j(x)$ and of radius ε. The Lebesgue measure of this ball will be equal to $\xi_N \varepsilon^N$, where the factor ξ_N was defined in the preceding footnote. Then we have

$$\mathrm{Leb}(B_\varepsilon(x)) = \xi_N \varepsilon^N \geq \mathrm{Leb}(T^j(V)) \geq |\det(DT^j(\kappa))|\mathrm{Leb}(V)$$

for a suitable point $\kappa \in \tilde{Z}_l^j$, where in the last inequality we used a local change of variable and the continuity of DT^j. By distortion, we could replace this point by another one, say ι such that $\mathrm{Leb}(T^j(Z_l^{(j)}) = |\det(DT^j(\iota))|\mathrm{Leb}(Z_l^{(j)})$; we call B the distortion constant satisfying $\frac{|\det(DT^j(\iota))|}{|\det(DT^j(\kappa))|} \leq B$. We therefore get

$$\mathrm{Leb}(V) \leq \frac{\xi_N \varepsilon^N B}{|\det(DT^j(\iota))|} \leq \frac{\xi_N \varepsilon^N B\ \mathrm{Leb}(Z_l^{(j)})}{U_m}$$

Since the density of the absolutely continuous invariant measure μ is bounded from above (remember it is in $\mathscr{L}^\infty(\mathrm{Leb})$), by, say, h_M, and since each $Z_l^{(j)}$ will contribute with at most one fixed point, by taking the sum over the l we will equivalently get an upper bound on the total measure of the balls including the $T^j(V)$; hence we finally get

$$\mu\{x; \mathrm{dist}(T^j x, x)\} \leq \frac{\xi_N h_M \, \varepsilon^N B}{U_m}.$$

and this bound is independent of j. □

As a consequence of Lemma 2.8 we have,

Theorem 3.4.4. *Assume a piecewise expanding map T on a compact set $Z \subset \mathbb{R}^n$ satisfies conditions (H1)–(H5) and is mixing with respect to its absolutely continuous invariant measure μ. For μ a.e. p if $B_i(p)$ are nested balls about p such that $\sum_i \mu(B_i) = \infty$ and $\mu(B_i) \leq \frac{C_2}{i^\gamma}$ for some constants $C_2 > 0$ and $\gamma \in (0, 1]$. Then*

$$a_n^2 = E[(\sum_{j=1}^{n} (1_{B_i} \circ T^i - \frac{1}{i}))^2] = \log n + \mathcal{O}(1)$$

and

$$\frac{1}{\sqrt{\log n}} \sum_{j=1}^{n} (1_{B_i} \circ T^i - \frac{1}{i}) \to N(0, 1)$$

in distribution.

3.5 Open questions

There are several natural questions remaining unanswered. In particular can the CLT for shrinking targets be proved for Anosov systems or non-uniformly hyperbolic diffeomorphisms? Chernov and Kleinbock have proved the SBC property for balls in Anosov systems [6] but the SBC property is unknown for non-uniformly hyperbolic diffeomorphisms. More generally can a limit theory be developed for the statistics of non-stationary stochastic processes arising as observations (which change in time) on deterministic dynamical systems which may also may evolve in time, such as sequential dynamical systems?

Chapter 4

BCL for Geometric Lorenz Models

In this chapter and next chapter, we study a particular chaotic system, the Lorenz system, and establish Strong Borel Cantelli lemma, Poisson laws for REPP and Extreme Value Laws for it.

The equations defining the Lorenz system were first published in the *Journal of Atmospheric Sciences*([33]) as a parametrized polynomial system of differential equations:

$$\dot{x} = \sigma(y - x)$$

$$\dot{y} = x(\rho - z) - y$$

$$\dot{z} = xy - \beta z$$

where $\sigma = 10$, $\rho = 28$, $\beta = 8/3$. The system was proposed as a simplified model for thermal fluid convection, motivated by a desire to understand weather systems.

What is interesting is that the equations are deterministic but they produce chaot-

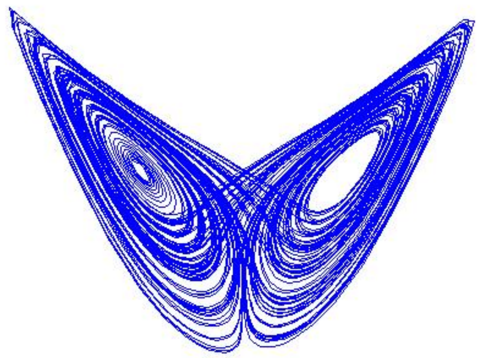

Figure 4.1: Lorenz attractor

ic behavior, with trajectories spiraling around two attractors seemingly randomly. Figure 4.1 is the Lorenz attractor, the famous "butterfly".

In order to achieve insights on this system, a very successful approach was taken by Afraimovich, Bykov and Shil'nikov[2], and Guckenheimer, Williams[19], independently: they constructed the so-called *Geometric Lorenz models*. These models are flows in three-dimension which have properties very similar to the Lorenz systems and are easier to study. One can rigorously prove the existence of an attractor that contains an equilibrium point of the flow, together with regular solutions. The original proof of the existence of a chaotic attractor was made by Warwick Tucker in the year 2000, with the help of computer (see [42, 43]).

A brief version of construction of the Geometric Lorenz model is given in Appendix A.4, and more detailed version can be found in [15, section 2.1]. As we can see in the construction, the Lorenz map F has a skew product form $F(x, y) = (T(x), G(x, y))$, where F does not preserve the two-dimensional Lebesgue measure m_2 but the one-dimensional map T preserves a measure absolutely continuous with respect to the one-dimensional Lebesgue measure m, with a Lipschitz density.

Much recent work has focused on the ergodic and statistical properties of Lorenz-like maps including rates of mixing, extreme value theory and return time statistics. S. Galatolo and M.J. Pacifico [15] proved that the Poincaré map, i.e., our Lorenz map F, associated to a Lorenz like flow has exponential decay of correlations with respect to Lipschitz observables and the hitting time statistics satisfies a logarithm law.

4.0.1 Local dimension

Let (M, d) be a metric space and assume that μ is a Borel probability measure on M. Given $x \in M$, let $B_r(x) = \{y \in M : d(x, y) \leq r\}$ be the ball centered at x with radius r. The *local dimension* of μ at $x \in M$ is defined by

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

if this limit exists. In this case, for any $\epsilon > 0$, there exists small r_0 such that for $0 < r \leq r_0$, $r^{d_\mu(x)-\epsilon} \leq \mu(B_r(x)) \leq r^{d_\mu(x)+\epsilon}$.

A result of Afraimovich and Pesin [3, Theorem 9] ensures that for the Lorenz system, the local dimension exists and is constant for μ a.e. point.

4.1 Main results

In section 4.2, to introduce the main ideas of our analysis of the Lorenz system in a simpler setting, we establish results of independent interest, namely shrinking target properties for general skew product maps which do preserve the two-dimensional Lebesgue measure. We have the following theorem:

Theorem 4.1.1. *Suppose $(\Omega, \mathscr{B}, m_2)$ is a probability space, where $\Omega = I \times I$, with $I = [-\frac{1}{2}, \frac{1}{2}]$, and m_2 is the two-dimensional Lebesgue measure. Let m denote the one-dimensional Lebesgue measure. Suppose $F : \Omega \to \Omega$ is a map in the form $F(x, y) = (T(x), G(x, y))$, where $T : I \to I$. Here F preserves the two-dimensional Lebesgue measure and T preserves the one-dimensional Lebesgue measure. In addition, T satisfies exponential decay of correlation with observables in BV norm versus L^1 norm (or L^1 v.s. L^1), i.e.*

$$|\int \phi\psi \circ T^n dm - \int \phi dm \int \psi dm| \leq C\theta^n ||\phi||_{BV} ||\psi||_1,$$

or

$$|\int \phi\psi \circ T^n dm - \int \phi dm \int \psi dm| \leq C\theta^n ||\phi||_1 ||\psi||_1,$$

and F satisfies exponential decay of correlation with observables in Lipschitz norm versus Lipschitz norm (or Lipschitz norm v.s. L^∞ norm):

$$|\int \phi\psi \circ F^n dm_2 - \int \phi dm_2 \int \psi dm_2| \leq C\alpha^n ||\phi||_{Lip} ||\psi||_{Lip}.$$

or

$$|\int \phi\psi \circ F^n dm_2 - \int \phi dm_2 \int \psi dm_2| \leq C\alpha^n ||\phi||_{Lip} ||\psi||_\infty.$$

Consider nested balls $\{B_i(p)\}$, centered at some $p \in \Omega$, with $m_2(B_i(p)) \geq \frac{C}{i^{\gamma_1}}$ for some $\gamma_1 > 0$, and $\limsup(\log i)(m_2(B_i))^{\frac{1}{2}} \leq C$. Then we have the Strong Borel Cantelli property

$$\frac{S_n(x,y)}{E_n} \to 1 \ a.s.$$

where $S_n(x,y) = \sum_{j=0}^{n-1} 1_{B_j} \circ F^j(x,y)$, $E_n = \sum_{j=0}^{n-1} m_2(B_j)$.

Then in section 4.3, we establish SBC property for the shrinking target problems for the two-dimensional Lorenz map F. When we consider shrinking target balls, they have different shapes according to different metrics. Technically, balls of different shapes are equivalent to each other, but we are able to deal with rectangles and circle balls. While rectangles consist of local unstable manifold of same length, circle balls don't. For nested rectangles, we assume they centered at a same point and have same ratio between lengths and widths. One special case is when they are squares, we prove the following theorem:

Theorem 4.1.2. *Consider a sequence of nested squares $\{A_i\}$, centered at a point p, of side length $2r(i)$ such that $\mu(A_i) \geq \frac{C_2}{i^{\gamma_1}}$, with $\gamma_1 \geq 0$. Assume that p has a local product structure for sufficiently small neighborhoods. Also we assume $(\log i)m(\pi(A_i)_\gamma)$ is bounded, where γ is any local unstable manifold of A_i and π is the projection map onto the unstable manifold. Then if $\sum_i \mu(A_i)$ diverges we have the Strong Borel Cantelli Property for the squares $\{A_i\}$ of side length $2r(i)$.*

Remark 4.1.3. *Theorem 4.1.2 holds for rectangles of bounded aspect ratio as well.*

And for circle balls, we show:

Theorem 4.1.4. *Consider a sequence of nested circle balls $\{A_i\}$, centered at a point p, of radius $r(i)$ such that $\mu(A_i) \geq \frac{C_2}{i^{\gamma_1}}$, with $\gamma_1 \geq 0$. Also assume that p has a local product structure for sufficiently small neighborhoods. Then if $\sum_i \mu(A_i)$ diverges we have the Strong Borel Property for the balls $\{A_i\}$.*

4.2 Volume Preserving Skew Products

As we mentioned above, the Lorenz map F has a skew product form $F(x, y) = (T(x), G(x, y))$. And F does not preserve the two-dimensional Lebesgue measure m_2 but the one-dimensional map T preserves a measure absolutely continuous with respect to the one-dimensional Lebesgue measure m, with a Lipschitz density. In this section, we talk about general skew product maps $\tilde{F}(x, y) = (T(x), G(x, y))$, which do preserve the two-dimensional Lebesgue measure and T does preserve the one-dimensional Lebesgue measure. We stated our Theorem 4.1.1, and now let's prove it, under the assumption that T satisfies exponential decay of correlation with observables in BV norm versus L^1 norm

$$|\int \phi\psi \circ T^n dm - \int \phi dm \int \psi dm| \leq C\theta^n ||\phi||_{BV} ||\psi||_1,$$

and F satisfies exponential decay of correlation with observables in Lipschitz norm versus Lipschitz norm:

$$|\int \phi\psi \circ F^n dm_2 - \int \phi dm_2 \int \psi dm_2| \leq C\alpha^n ||\phi||_{Lip} ||\psi||_{Lip}.$$

For other options of norms, the proof works in an analogous way with slight changes.

Proof of Theorem 4.1.1.

Let $f_k(x,y) = \mathbf{1}_{B_k} \circ \tilde{F}^k(x,y)$, $E(f_k) = m_2(B_k)$ and let $a \geq \frac{-7\gamma_1}{\log \alpha}$; we will use a later. To prove Strong Borel Cantelli Property, according to [25], it suffices to show the (SP) property (see Appendix A.2), i.e. for all $m < n$

$$\sum_{i=m}^{n} \sum_{j=i+1}^{n} (E(f_i f_j) - E(f_i)E(f_j)) \leq C \sum_{i=m}^{n} E(f_i).$$

We calculate

$$\begin{aligned}
E(f_i f_j) &= \int \mathbf{1}_{B_i} \circ \tilde{F}^i(x,y) \cdot \mathbf{1}_{B_j} \circ \tilde{F}^j(x,y) dm_2 \\
&= \int \mathbf{1}_{B_i} \cdot \mathbf{1}_{B_j} \circ \tilde{F}^{j-i}(x,y) dm_2 \\
&= m_2(B_i \cap \tilde{F}^{-(j-i)} B_j) \\
&\leq C_1 (m_2(B_i))^{\frac{1}{2}} \cdot m(\pi_X B_i \cap T^{-(j-i)} \pi_X B_j) \quad (*)
\end{aligned}$$

where $\pi_X B_i$ is the projection of the maximal horizontal section of the ball B_i onto the unstable manifold. If $\tilde{x} = (x,y) \in B_i$ and $\tilde{F}^{j-i}(\tilde{x}) \in B_j$, then their projection $x \in \pi_X B_i$ and $T^{j-i}(x) \in \pi_X B_j$ so $m_2(B_i \cap \tilde{F}^{-(j-i)} B_j) = \int m((B_i \cap \tilde{F}^{-(j-i)} B_j)_y) dm(y) = \int m(\{\tilde{x} = (x,y) | \tilde{x} \in B_i, \tilde{F}^{j-i}(\tilde{x}) \in B_j\}) dm(y) \leq \int m(x \in \pi_X B_i, T^{j-i}(x) \in \pi_X B_j) dm(y) \leq m(\pi_Y B_i) \cdot m(x \in \pi_X B_i, T^{j-i}(x) \in \pi_X B_j)$, where $\pi_Y B_i$ is the projection of the maximal vertical section of B_i onto stable manifold. And $m(\pi_X B_i) = m(\pi_Y B_i) \leq C_1 (m_2(B_i))^{\frac{1}{2}}$. Thus we continue calculating $E(f_i f_j)$:

$$\begin{aligned}
(*) &= C_1 (m_2(B_i))^{\frac{1}{2}} \cdot \int \mathbf{1}_{\pi_X B_i} \cdot \mathbf{1}_{\pi_X B_j} \circ T^{j-i}(x) dm \\
&\leq C_1 (m_2(B_i))^{\frac{1}{2}} \cdot (\int \mathbf{1}_{\pi_X B_i} \int \mathbf{1}_{\pi_X B_j} + C\theta^{j-i} ||\mathbf{1}_{\pi_X B_i}||_{BV} ||\mathbf{1}_{\pi_X B_j}||_1) \\
&\leq C_1 (m_2(B_i))^{\frac{1}{2}} \cdot \left[(m_2(B_i))^{\frac{1}{2}} \cdot (m_2(B_j))^{\frac{1}{2}} + C\theta^{j-i} (m_2(B_j))^{\frac{1}{2}} \right] \\
&\leq C_1 (m_2(B_i))^{\frac{3}{2}} + C\theta^{j-i} m_2(B_i).
\end{aligned}$$

since $\int \mathbf{1}_{\pi_X B_i} dm = m(\pi_X B_i) \leq C(m_2(B_i))^{\frac{1}{2}}$.

Recall $a \geq \frac{-7\gamma_1}{\log \alpha}$, so that

$$\begin{aligned}
& \sum_{j=i+1}^{n} (E(f_i f_j) - E(f_i)E(f_j)) \\
\leq \;& \left(\sum_{j=i+1}^{i+a\log i} + \sum_{j>i+a\log i}\right)[E(f_i f_j) - E(f_i)E(f_j)] \\
\leq \;& C_1(\log i)(m_2(B_i))^{\frac{3}{2}} + Cm_2(B_i) + \sum_{j>i+a\log i} C\alpha^{j-i}||\tilde{f}_i||_{Lip}||\tilde{f}_j||_{Lip} + O(\frac{1}{i^{7/2}}).
\end{aligned}$$

where $\tilde{f}_i$ is a Lipschitz approximation to f_i, which is constructed as following: $\tilde{f}_i = 1$ if $x \in B_i$, $\tilde{f}_i = 0$ if $d(B_i, x) > 1/i^{3\gamma_1}$, $0 \leq \tilde{f}_i \leq 1$ and $||\tilde{f}_i||_{Lip} \leq i^{3\gamma_1}$.

Then

$$\begin{aligned}
\sum_{j>i+a\log i} \alpha^{j-i}||\tilde{f}_i||_{Lip}||\tilde{f}_j||_{Lip} &\leq \sum_{j>i+a\log i} \alpha^{j-i} i^{3\gamma_1} j^{3\gamma_1} \\
&= \sum_{\beta=1}^{\infty} \alpha^{a\log i+\beta} i^{3\gamma_1}(i + a\log i + \beta)^{3\gamma_1} \\
&\leq \alpha^{a\log i} i^{3\gamma_1} C i^{3\gamma_1} \\
&\leq \frac{C}{i^{\gamma_1}} \leq Cm_2(B_i).
\end{aligned}$$

Since $(\log i)(m_2(B_i))^{\frac{1}{2}} \leq C$, the (SP) property is satisfied. □

Remark 4.2.1. *As we mention in section 2.3, (SP) property (Sprindzuk Property) is derived from Gal-Koksma theorem, which is given in the Appendix. Once we have the (SP) property, then the Strong Borel Cantelli Property is established because in the Gal-Koksma theorem, if we take* $f_k(\omega) = \mathbf{1}_{B_k} \circ \tilde{F}^k(x, y)$, $h_k = g_k = E(f_k) = m_2(B_k)$, *dividing both sides of the equation by* $\sum_{k=1}^{n} g_k$, *we will have* $\frac{S_n(x,y)}{E_n} \to 1$ *a.s..*

4.3 Lorenz System

In this section, we present our results on the statistical properties of the Lorenz system, i.e., Borel Cantelli Lemmas. And we will see convergence of REPP(which implies Extreme Value Laws) in chapter 5. Recall that the Lorenz map F does not preserve the two-dimensional Lebesgue measure m_2, but preserves an invariant measure μ which has absolutely continuous conditional measures on local unstable manifolds.

4.3.1 Borel-Cantelli Lemma

We let $A_r(p)$ denote the square of side length $2r$ centered at a point p in the two dimensional space $I \times I$. As a consequence of [4, Proposition 2.4], for μ a.e. p, there exists an $r(p) > 0$ such that for all $r < r(p)$, $A_r(p)$ has a local product structure and in particular μ a.e. $q \in A_r(p)$ has a local unstable manifold $\gamma(q) := W^u_{loc}(q)$ which extends fully across $A_r(p)$. The local stable manifolds are arbitrarily long for μ a.e. q. The set of local unstable manifolds $\Gamma = \{\gamma(q)\}$ partition $A_r(p)$ up to a set of zero μ measure i.e. $\mu(A_r(p)) = \mu(\cup_{q\in A_r(p)}\gamma(q) \cap A_r(p))$. We will drop the dependence on q and write $\Gamma = \{\gamma\}$ for simplicity.

Remark 4.3.1 (Young Tower Structure and Local Product Structure)**.** *A recent paper of Araujo, Melbourne and Varandas [4] used a Young Tower construction to establish that a broad range of geometric Lorenz flows are rapidly mixing. Along the way they showed that μ a.e. $p \in M$ has a local product structure (this follows from their Proposition 2.4). More precisely, μ a.e. p has the property that there exists a*

$r(p) > 0$ such that for all $r < r(p)$ if $A(r)$ is a square of side length $2r$, or a ball of radius r, centered at p, then μ a.e. $q \in A(r)$ has a local unstable manifold and a local stable manifold which fully crosses $A(r)$. Moreover if q_1, q_2 are in $A(r)$ then there is a unique point $z = W^u_{loc}(q_1) \cap W^s_{loc}(q_2) \in A(r)$.

Remark 4.3.2. *As we mentioned before that Theorem 4.1.2 holds for rectangles of bounded aspect ratio. We just need to change the setting above a little. Let $A_r(p)$ denote the rectangle with side length $2r$ of unstable manifold and side length $2sr$ of stable manifold, for some s, where $0 < s < \infty$. The proof will be the same as the proof of Theorem 4.1.2 later in this section.*

Before we prove Theorem 4.1.2, we introduce the notation.

Given such a point p we let A be a square based at p, with side length smaller than $2r(p)$.

Let $A_\gamma = A \cap \gamma$ for $\gamma \in \Gamma$,

$$\mu(A) = \int_I m_\gamma(A_\gamma) d\nu(\gamma),$$

where m_γ is the induced measure of μ on γ and ν is conditional measure in the decomposition of μ with respect to the partition $\{\gamma\}$.

Let π be the projection map onto the unstable manifold and note that $m_\gamma(A_\gamma) \sim m(\pi A)$. Here "$\sim$" means the equivalence of measures on the various unstable manifolds.

For the Lorenz map F, we have exponential decay in Lipschitz versus Lipschitz (see [15, Theorem 4.7])

$$|\int \phi\psi \circ F^n d\mu - \int \phi d\mu \int \psi d\mu| \leq C\alpha^n ||\phi||_{Lip}||\psi||_{Lip},$$

and for the base map T, we have exponential decay in L^1 versus BV (see [15, Propostion 2.2])

$$|\int \phi\psi \circ T^n dm - \int \phi dm \int \psi dm| \leq C\theta^n ||\phi||_{BV}||\psi||_1,$$

By taking $\phi = \psi = \mathbf{1}_{\pi A} = \mathbf{1}_{\pi A_\gamma}$, we have

$$m(\pi A_\gamma \cap T^{-n}(\pi A_\gamma)) - (m(\pi A_\gamma))^2 \leq C\theta^n ||\mathbf{1}_{\pi A_\gamma}||_{BV}||\mathbf{1}_{\pi A_\gamma}||_1.$$

That is

$$m(\pi A_\gamma \cap T^{-n}(\pi A_\gamma)) \leq (m(\pi A_\gamma))^2 + C'\theta^n m(\pi A_\gamma),$$

since $||\mathbf{1}_{\pi A_\gamma}||_{BV}$ is bounded.

Proof of Theorem 4.1.2.

It suffices to establish (SP) property. Without loss of generality, we assume $i < j$.

We notice $m(\pi(A_i)_\gamma)$ is equal for all $\gamma \in \Gamma$ since they are square balls. Thus,

$$
\begin{aligned}
& \mu(A_i \cap F^{-(j-i)}A_j) \leq \mu(A_i \cap F^{-(j-i)}A_i) \\
\sim \; & \int_I m(\tilde{x} \in (A_i)_\gamma : F^{j-i}(\tilde{x}) \in A_i) d\nu(\gamma) \\
\leq \; & \int_I m(\pi(A_i)_\gamma \cap T^{-(j-i)}(\pi(A_i)_\gamma)) d\nu(\gamma) \\
\leq \; & \int_I (m(\pi(A_i)_\gamma))^2 + C'\theta^{(j-i)} m(\pi(A_i)_\gamma) d\nu(\gamma) \\
= \; & \int_I (m(\pi(A_i)_\gamma))^2 d\nu(\gamma) + C'\theta^{(j-i)} \int_I m(\pi(A_i)_\gamma) d\nu(\gamma) \\
= \; & \int_I (m(\pi(A_i)_\gamma))^2 d\nu(\gamma) + C'\theta^{(j-i)} \mu(A_i) \\
\leq \; & Cm(\pi(A_i)_\gamma) \int_I m(\pi(A_i)_\gamma) d\nu(\gamma) + C'\theta^{(j-i)} \mu(A_i) \\
= \; & Cm(\pi(A_i)_\gamma) \mu(A_i) + C'\theta^{(j-i)} \mu(A_i),
\end{aligned}
$$

where $\sim$ is defined the same as earlier, i.e. it means the equivalence of measures on the various unstable manifolds, and we use the notation $m(\pi(A_i)_\gamma)$ for any $\gamma \in \Gamma$ since they are all equal.

Since $\mathbf{1}_{A_n}$ is not Lipschitz, we use similar technique as in chapter 2, i.e., let ϕ_n be a Lipschitz approximation of $\mathbf{1}_{A_n}$ such that

1. $||\mathbf{1}_{A_n} - \phi_n||_1 < (\mu(A_n))^3$,

2. $||\phi_n||_{Lip} < (\mu(A_n))^{-3}$.

Then

$$
\begin{aligned}
& \left|\int \mathbf{1}_{A_i} \cdot \mathbf{1}_{A_j} \circ F^{j-i} d\mu - \int \mathbf{1}_{A_i} d\mu \int \mathbf{1}_{A_j} d\mu\right| \\
= & \ \Big| \int ([\mathbf{1}_{A_i} - \phi_i] + \phi_i) \cdot ([\mathbf{1}_{A_j} - \phi_j] + \phi_j) \circ F^{j-i} d\mu \\
& \quad - \int ([\mathbf{1}_{A_i} - \phi_i] + \phi_i) d\mu \int ([\mathbf{1}_{A_j} - \phi_j] + \phi_j) d\mu \Big| \\
\leq & \ \left|\int \phi_i \phi_j \circ F^{j-i} d\mu - \int \phi_i d\mu \int \phi_j d\mu\right| + \left|\int (\mathbf{1}_{A_i} - \phi_i)(\mathbf{1}_{A_j} - \phi_j) \circ F^{j-i} d\mu\right| \\
& + \left|\int \phi_i (\mathbf{1}_{A_j} - \phi_j) \circ F^{j-i} d\mu\right| + \left|\int (\mathbf{1}_{A_i} - \phi_i) \phi_j \circ F^{j-i} d\mu\right| \\
& + \left|\int (\mathbf{1}_{A_i} - \phi_i) d\mu \int (\mathbf{1}_{A_j} - \phi_j) d\mu\right| + \left|\int \phi_i d\mu \int (\mathbf{1}_{A_j} - \phi_j) d\mu\right| \\
& + \left|\int (\mathbf{1}_{A_i} - \phi_i) d\mu \int \phi_j d\mu\right| \\
\leq & \ C\alpha^{j-i} ||\phi_i||_{Lip} ||\phi_j||_{Lip} + \tilde{C}(\mu(A_i))^3.
\end{aligned}
$$

If we choose $a \geq \frac{-7\gamma_1}{\log \alpha}$, then

$$
\begin{aligned}
\sum_{j > i + a\log i} \alpha^{j-i} ||\phi_i||_{Lip} ||\phi_j||_{Lip} & \leq \sum_{j > i + a\log i} \alpha^{j-i} i^{3\gamma_1} j^{3\gamma_1} \\
& = \sum_{\beta=1}^{\infty} \alpha^{a\log i + \beta} i^{3\gamma_1} (i + a\log i + \beta)^{3\gamma_1} \\
& \leq \alpha^{a\log i} i^{3\gamma_1} C i^{3\gamma_1} \\
& \leq \frac{C}{i^{\gamma_1}} \\
& \leq C\mu(A_i).
\end{aligned}
$$

Thus, let $f_k = \mathbf{1}_{A_k} \circ F^k(x,y)$, $E(f_k) = \mu(A_k)$, so that we have

$$\begin{aligned}
& \sum_{j=i+1}^{n} (E(f_i f_j) - E(f_i)E(f_j)) \\
= & \sum_{j=i+1}^{n} \mu(A_i \cap F^{j-i} A_j) - \mu(A_i)\mu(A_j) \\
= & \sum_{j=i+1}^{i+a\log i} \big[\mu(A_i \cap F^{j-i} A_j) - \mu(A_i)\mu(A_j)\big] + \sum_{j>i+a\log i} \big[\mu(A_i \cap F^{j-i} A_j) - \mu(A_i)\mu(A_j)\big] \\
\leq & \sum_{j=i+1}^{i+a\log i} \big[m(\pi(A_i)_\gamma)\mu(A_i) + C'\theta^{j-i}\mu(A_i)\big] + C\mu(A_i) \\
\leq & \; C(\log i) m(\pi(A_i)_\gamma)\mu(A_i) + C_1\mu(A_i) + C_2\mu(A_i) \leq \tilde{C}\mu(A_i).
\end{aligned}$$

We have established the (SP) property and thus the strong Borel Cantelli lemma for $\{A_i\}$.

□

Remark 4.3.3. *For general nested rectangles, the proof will be the same with slight difference in the setting.*

For circle balls, we have Theorem 4.1.4. As we mentioned in the introduction, the proof of Theorem 4.1.4 needs the techniques from the subsection 5.1.

Proof of Theorem 4.1.4.

We follow the steps of proof of Theorem 4.1.2, and use the techniques from the proof of Theorem 5.0.4, we have:

$$\begin{aligned} & \mu(A_i \cap F^{-(j-i)} A_j) \leq \mu(A_i \cap F^{-(j-i)} A_i) \\ \sim & \int_\Gamma m(\tilde{x} \in (A_i)_\gamma : F^{j-i}(\tilde{x}) \in A_i) d\nu(\gamma) \\ \leq & \int_\Gamma m(\pi(A_i)_\gamma \cap T^{-(j-i)}(\pi(A_i)_\gamma)) d\nu(\gamma) \\ = & \int_{\Gamma_1} m(\pi(A_i)_\gamma \cap T^{-(j-i)}(\pi(A_i)_\gamma)) d\nu(\gamma) \\ & + \int_{\Gamma_2} m(\pi(A_i)_\gamma \cap T^{-(j-i)}(\pi(A_i)_\gamma)) d\nu(\gamma), \end{aligned}$$

where $\sim$ is defined the same as before and Γ_1 is the set of local unstable manifolds which has length less than r_n^3, and $\Gamma_2 = \Omega / \Gamma_1$. Thus,

$$\int_{\Gamma_1} m(\pi(A_i)_\gamma \cap T^{-(j-i)}(\pi(A_i)_\gamma)) d\nu(\gamma) \leq r_i^3 \int_{\Gamma_1} d\nu(\gamma) \leq r_i^3 \leq \mu(A_i)^{\frac{3}{d+\epsilon}},$$

where ϵ is such that $r_n^{d+\epsilon} \leq \mu(A_n) \leq r_n^{d-\epsilon}$ for n large and d is the local dimension. We let $u_i = -\log r_i$, that is, $r_i = e^{-u_i}$, where the u_i is determined by the radius r_i, not necessarily satisfying (1.9). And by (5.5)

$$\int_{\Gamma_2} m(\pi(A_i)_\gamma \cap T^{-(j-i)}(\pi(A_i)_\gamma)) d\nu(\gamma) \leq \mu(A_i) \exp(-u_i^\rho).$$

Let $f_k = \mathbf{1}_{A_k} \circ F^k(x,y)$ and $E(f_k) = \mu(A_k)$, so that

$$\begin{aligned}
& \sum_{j=i+1}^{n} (E(f_i f_j) - E(f_i)E(f_j)) \\
= & \sum_{j=i+1}^{n} \mu(A_i \cap F^{-(j-i)}A_j) - \mu(A_i)\mu(A_j) \\
= & \sum_{j=i+1}^{i+(\log i)^5} \left[\mu(A_i \cap F^{-(j-i)}A_j) - \mu(A_i)\mu(A_j)\right] \\
& + \sum_{j>i+(\log i)^5} \left[\mu(A_i \cap F^{-(j-i)}A_j) - \mu(A_i)\mu(A_j)\right] \\
\le & \sum_{j=i+1}^{i+(\log i)^5} \left[\mu(A_i)^{\frac{3}{d+\epsilon}} + \mu(A_i)\exp(-u_i^\rho)\right] + C\mu(A_i) \\
\le & (\log i)^5 \mu(A_i)^{\frac{3}{d+\epsilon}-1}\mu(A_i) + C_1\mu(A_i) + C\mu(A_i) \le \tilde{C}\mu(A_i).
\end{aligned}$$

where $0 < \rho < 1/3$ and $(\log i)^5 e^{-u_i^\rho}$ is bounded since $e^{-u_i^\rho} \le e^{-(\frac{-1}{d+\epsilon}\log\mu(A_i))^\rho} \le e^{-(\frac{\gamma_1}{d+\epsilon}\log i)^\rho}$. Therefore we have the (SP) property, and the Strong Borel Cantelli property then follows. $\square$

Chapter 5

REPP(EVL) for Geometric Lorenz Models

In this chapter, we establish Poisson Laws for REPP, which implies EVL. The main results are the following:

Theorem 5.0.4. *Consider the dynamical systems* $(\Omega, \mathscr{B}, \mu, F)$, *where* F *is the Lorenz map and* F *preserves the measure* μ, *whose decomposition on the unstable leaves is absolutely continuous with respect to the one-dimensional Lebesgue measure. Let* φ *be defined as* (5.1), $X_n = \varphi \circ F^n$. *For* μ *a.e.* x_0 *(we in particular assume* x_0 *is not periodic under* F*), we assume the sequence* $\{u_n\}$ *satisfies* $n\mu(X_0 > u_n) \to e^{-v}$. *Then REPP* N_n *given in definition 1.4.8 converges in distribution to the standard Poisson process.*

Then immediately the next corollary follows by Extreme Value Law:

Corollary 5.0.5. *Under the same setting as above, then X_n satisfies a Type I extreme value law, i.e.*

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

In section 5.2, we extend our results to the Lorenz flow f_t as well, in which case we consider continuous time stochastic process $\{X_t\}$ such that $X_t = \varphi \circ f_t$, and define the process of successive maxima $\varphi_t := \sup_{0\leq s\leq t}\{X_s\}$.

5.1 REPP and Extreme Value Laws

In this section, we establish the convergence of the rare events point processes to the standard Poisson process(thus EVL follows) for Lorenz maps by essentially showing that the two conditions $D_3(u_n)$ and $D'(u_n)$, which were introduced in section 1.4.3, are satisfied. Recall:

Condition ($D_3(u_n)$). We say that $D_3(u_n)$ holds for the sequence $X_0, X_1, X_2, \cdots$ if for all $B \in \mathcal{R}$ and $t \in \mathbb{N}$,

$$\left|\mu(\{X_0 > u_n\} \cap \{M(B+t) \leq u_n\}) - \mu(\{X_0 > u_n\})\mu(\{M(B) \leq u_n\})\right| \leq \gamma(n,t),$$

where $\gamma(n,t)$ is nonincreasing 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 ($D'(u_n)$): The condition $D'(u_n)$ is said to hold for the stationary sequence $\{X_i\}$ and the sequence $\{u_n\}$ if

$$\limsup_{n\to\infty} n \sum_{j=1}^{[n/k]} \mu(X_0 > u_n, X_j > u_n) \to 0,$$

as $k \to \infty$.

For $x_0 \in \Omega$, if we consider

$$\varphi(x) = -\log d(x, x_0), \tag{5.1}$$

where $d(\cdot,\cdot)$ is the local metric on Ω and x_0 is a generic point(we fix one point x_0). Let $X_n = \varphi \circ T^n$ and define $U_n = \{X_0 > u_n\}$, where $u_n = u_n(v)$ such that $\mu(U_n) = e^{-v}/n$ (In this case, we will see later that u_n is roughly a linear function of this v). Here u_n is an increasing sequence going to $\varphi(x_0)$(which is $+\infty$) and assume U_n corresponds to a topological ball centered at x_0 with radius e^{-u_n}. Then the corresponding processes $\{X_n\}$ will satisfy Type I extremal distribution.

To be consistent, if not specified, we will use the definitions and assumptions above in rest of this chapter.

Before we prove Theorem 5.0.4, let's prove the following two lemmas:

Lemma 5.1.1. *Suppose we have a local product structure about a point x_0 and the local dimension exists, denoted by d. Then Assumption (A) is satisfied.*

Proof. As before, the conditional measure μ_γ is equivalent to Lebesgue in the local unstable direction, and r is small, i.e., $r < 1$. Let $\epsilon = r^w$, with $w > 1$. We need to prove that the measure of the annular region $S = A_{r+\epsilon}(x_0)\backslash A_r(x_0)$ is small.

We decompose μ in a neighborhood of x_0 as follows

$$\mu(A) = \int_{\gamma\in\Gamma} m(\gamma \cap A) d\nu(\gamma)$$

where γ is the foliation into local unstable manifolds. Since we have a local product structure at x_0, these extend all the way across a sufficiently small rectangular neighborhood of x_0.

Now consider the equation of the circles $x^2 + y^2 = r^2$ and $x^2 + y^2 = (r + \epsilon)^2 = r^2 + 2r^{w+1} + r^{2w}$. The larger circle contains some local unstable manifolds which are not in the smaller circle but the greatest length of these is found by setting $y^2 = r^2$ in the second equation and solving for $\delta x \le r^{\frac{w+1}{2}}$. Their length is less than $r^{\frac{w+1}{2}}$, so that

$$\mu(S) \le \int_\Gamma (S \cap \gamma) d\nu(\gamma) < r^{\frac{w+1}{2}} < \epsilon^{\frac{w+1}{2w}} < \epsilon^{1/2}.$$

$\square$

Lemma 5.1.2. *(a) For μ a.e. x_0, for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n \ge N$*

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

where d is the local dimension.

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

$$\mu(S(n, x_0)) \le C_3 n^{-2\delta v - \delta \log n}.$$

Proof. (a) By the definition of the local dimension, for any $\epsilon > 0$, there is an N such that for all $n \geq N$, $(e^{-u_n})^{(d+\epsilon)} \leq \mu(U_n) \leq (e^{-u_n})^{(d-\epsilon)}$, and $\mu(U_n) = e^{-v}/n$. We get immediately

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

(b) According to Subsection 4.0.1 and Remark 4.3.1, Lorenz system has local dimension and local structure, and by Lemma 5.1.1, Assumption (A) is satisfied for Lorenz system. So there exists a $\delta \in (0, 1)$ such that

$$\begin{aligned}
\mu(S(n, x_0)) &\leq C(e^{-(u_n^2)})^{\delta} \\
&= Ce^{-(u_n^2)\delta} \\
&\leq C \exp\left(-\frac{\delta}{(d+\epsilon)^2}(v + \log n)^2\right) \\
&\leq C_3 n^{-2\delta' v - \delta' \log n}.
\end{aligned}$$

□

Proof of Theorem 5.0.4.

It suffices to show $D_3(u_n)$ and $D'(u_n)$. As mentioned in Remark 1.4.13, $D_3(u_n)$ is easily proven if the system satisfies Assumption (A) and good enough estimates for decay of correlations. The proof here adopts similar arguments to that in the proof of [14, Theorem 3.1]. We already know Assumption (A) is satisfied for Lorenz systems, and we have exponential decay of correlation for Lorenz map

$$|\int \phi\psi \circ F^n d\mu - \int \phi d\mu \int \psi d\mu| \leq C\alpha^n ||\phi||_{Lip} ||\psi||_{Lip}.$$

So we prove $D_3(u_n)$ in the following paragraph.

By part (b) of Lemma 5.1.2, we have $\mu(S(n,x_0)) \leq C_3 n^{-2\delta\upsilon-\delta\log n}$, where $\delta=\delta(x_0)\in(0,1)$ for n large enough and υ could be any number. Take ϕ_n to be the Lipschitz approximation of $\mathbf{1}_{U_n}=\mathbf{1}_{\{X_0>u_n\}}$ such that $\phi_n(x)=1$ if x inside $A_{e^{-u_n}-e^{-u_n^2}}(x_0)$, $\phi_n=0$ if x is outside U_n, and decays to 0 at a linear rate on $S(n,x_0)$. So we have the estimate $||\phi_n-\mathbf{1}_{\{X_0>u_n\}}||_1<\mu(S(n,x_0))$ and $||\phi_n||_{Lip}\leq e^{u_n^2}$. Also let $\psi_B^n=\mathbf{1}_{\{M(B)\leq u_n\}}$, where $B\in\mathcal{R}$ so that $B=\cup_{j=1}^{l}[a_j,b_j)$. By ([14], Lemma 3.1), we then have

$$|\int\phi_n\psi_B^n\circ F^t d\mu-\int\phi_n d\mu\int\psi_B^n d\mu|\leq O(1)(||\phi_n||_\infty\tau_1^{\lfloor t/2\rfloor}+||\phi_n||_{Lip}\alpha^{\lfloor t/2\rfloor}),$$

and

$$\begin{aligned}
& |\mu(\{X_0>u_n\}\cap\{M(B+t)\leq u_n\})-\mu(\{X_0>u_n\})\mu(\{M(B)\leq u_n\})| \\
= \ & |\int\mathbf{1}_{U_n}\cdot\psi_B^n\circ F^t d\mu-\mu(U_n)\int\psi_B^n d\mu| \\
= \ & |\int\mathbf{1}_{U_n}\cdot\psi_{B+\lfloor t/2\rfloor}^n\circ F^{t-\lfloor t/2\rfloor}d\mu-\mu(U_n)\int\psi_{B+\lfloor t/2\rfloor}^n d\mu| \\
\leq \ & |\int(\mathbf{1}_{U_n}-\phi_n)\psi_{B+\lfloor t/2\rfloor}^n\circ F^{t-\lfloor t/2\rfloor}d\mu|+|\int\phi_n\psi_{B+\lfloor t/2\rfloor}^n\circ F^{t-\lfloor t/2\rfloor}d\mu \\
& -\int\phi_n d\mu\int\psi_{B+\lfloor t/2\rfloor}^n d\mu|+|\int(\mathbf{1}_{U_n}-\phi_n)d\mu\int\psi_{B+\lfloor t/2\rfloor}^n d\mu| \\
\leq \ & 2\mu(S(n,x_0))+O(1)(||\phi_n||_\infty\tau_1^{\lfloor t/4\rfloor}+||\phi_n||_{Lip}\alpha^{\lfloor t/4\rfloor}) \\
\leq \ & O(1)(n^{-2\delta\upsilon-\delta\log n}+||\phi_n||_\infty\tau_1^{\lfloor t/4\rfloor}+||\phi_n||_{Lip}\alpha^{\lfloor t/4\rfloor}),
\end{aligned}$$

where τ_1 is from [14, Proposition 3.2]. Let $\gamma(n,t)=O(1)(n^{-2\delta\upsilon-\delta\log n}+||\phi_n||_{Lip}\theta_2^{\lfloor t/4\rfloor})$, where $\theta_2=\max\{\tau_1,\alpha\}$ and $||\phi_n||_{Lip}\leq e^{u_n^2}\leq O(1)n^{2\tilde{\delta}\upsilon+\tilde{\delta}\log n}$(by similar argument we did to get $\mu(S(n,x_0))$). Take $t=t_n=(\log n)^5$, so that $n\gamma(n,t_n)\to 0$ as $n\to\infty$. Thus $D_3(u_n)$ is established.

We now establish $D'(u_n)$. Note $F^{-j}(U_n) = \{X_j > u_n\}$, then

$$\mu(X_0 > u_n, X_j > u_n) = \mu(U_n \cap F^{-j}U_n) \leq \int_{\gamma\in\Gamma} m(\pi(U_n)_\gamma \cap T^{-j}(\pi(U_n)_\gamma))d\nu(\gamma),$$

where π is the projection map onto the unstable manifold. For each leaf $\gamma \in \Gamma$, define $B_{r_n,\gamma} = \pi(U_n)_\gamma = B(q_\gamma e^{-u_n}, \pi(x_0))$, where $r_n = e^{-u_n}$, $0 \leq q_\gamma \leq 1$.

Consider points where the local unstable manifold is less than r_n^3, so that the integral splits as follows

$$\begin{aligned}\int_{\gamma\in\Gamma} m(\pi(U_n)_\gamma \cap T^{-j}(\pi(U_n)_\gamma))d\nu(\gamma) \quad &= \quad \int_{\gamma\in\Gamma_1} m(\pi(U_n)_\gamma \cap T^{-j}(\pi(U_n)_\gamma))d\nu(\gamma)\\ &\quad + \int_{\gamma\in\Gamma_2} m(\pi(U_n)_\gamma \cap T^{-j}(\pi(U_n)_\gamma))d\nu(\gamma),\end{aligned}$$

where Γ_1 is the set of local unstable manifolds which has length less than r_n^3, and $\Gamma_2 = \Omega/\Gamma_1$. The reason for doing so is because if the local unstable manifold has a short length, the point in the projection probably has no short return.

Remark 5.1.3 (Short returns for one-dimensional Lorenz-like maps)**.** *Gupta, Holland and Nicol [20] established extreme value statistics for Lorenz-like maps. The proofs used a crucial estimate on the measure of points with short returns. In particular, they showed that for μ a.e. $p \in I$, and for all sufficiently small $r < r(p)$, if $B_r(p)$ is a ball of radius r based at p, then there are constants $C > 0$, $0 < \alpha < 1$ such that for all $1 \leq j \leq (\log r)^5$, $\mu(B_r \cap T^{-j}B_r) \leq \mu(B_r)e^{-(\log r)^\alpha}$. We will adopt similar arguments in the following part.*

For $\gamma \in \Gamma_2$, define

$$E_{k,\gamma} = \{x \in B_{r_k,\gamma} : d(T^j x, x) < \frac{1}{k^{1/3}}, \text{ for some } 1 \leq j \leq (\log k)^5\}.$$

By [20, Proposition 4.2], there exists $0 < a < 1, 0 < \tilde{\theta} < 1$ such that

$$m(E_{k,\gamma}) < \tilde{\theta}^{(\log k^{1/3})^a},$$

We only need $a < 1/2$, so we take $a = 1/3$.

Let $0 < \beta \leq \frac{1}{2}$ and let $0 < \rho < 1$ such that $\rho\beta < \beta/3$.

Define the set

$$F_{k,\gamma} := \{m(B_{q_\gamma \exp(-k^\beta),\gamma}(x) \cap E_{\exp(k^{3\beta}),\gamma}) \geq m(B_{q_\gamma \exp(-k^\beta),\gamma}(x)) \exp(-k^{\beta\rho})\}. \quad (5.2)$$

If $x \in F_{k,\gamma}$ then

$$\frac{m(B_{q_\gamma \exp(-k^\beta),\gamma}(x) \cap E_{\exp(k^{3\beta}),\gamma})}{m(B_{q_\gamma \exp(-k^\beta),\gamma}(x))} \geq \exp(-k^{\beta\rho}); \quad (5.3)$$

If we define

$$M_l(x) := \sup_{r>0} \frac{1}{m(B_r(x))} \int_{B_r(x)} \mathbf{1}_{E_l}(y) dm(y)$$

we see immediately from the definition of $M_l(x)$ and (5.3) that for every $x \in F_{k,\gamma}, M_{e^{k^{3\beta}}}(x) \geq e^{-k^{\beta\rho}}$. Hence

$$F_{k,\gamma} \subset \{M_{e^{k^{3\beta}}}(x) \geq e^{-k^{\beta\rho}}\}. \quad (5.4)$$

A theorem of Hardy and Littlewood [36, Theorem 2.19] implies that

$$m(|M_l| > c) \leq \frac{\|\mathbf{1}_{E_l}\|_1}{c};$$

$|| \cdot ||_1$ is with respect to the one-dimensional Lebesgue measure. As $m(E_{l,\gamma}) \leq \mathcal{O}(1)\tilde{\theta}^{(\log l^{1/3})^{1/3}}$ (recall $a = 1/3$),

$$m(F_{k,\gamma}) \leq \mathcal{O}(1) m(E_{\exp(k^{3\beta}),\gamma}) e^{k^{\beta\rho}} \leq \mathcal{O}(1)(e^{\alpha k^{\beta/3} + k^{\beta\rho}}),$$

where $\alpha := \log\tilde{\theta}$ and k is large enough. Since $\beta/3 > \beta\rho$, $\sum_{k>0} m(F_{k,\gamma}) < \infty$. By the Borel Cantelli lemma, $m(\limsup F_{k,\gamma}) = 0$, and hence for m almost every x there exists an N_x such that for all $k \geq N_x$, $x \notin F_{k,\gamma}$ for each γ.

Let x_0 be such a generic point, and let N_{x_0} be the corresponding index beyond which x_0 does not belong to any $F_{k,\gamma}$. Since $\lim_{k\to\infty} e^{(k+1)^\beta} e^{-k^\beta} = 1$ the fact that we restricted to a subsequence is of no consequence, and we obtain the following estimate for all n sufficiently large. If $1 \leq j \leq (\log n)^5$, then

$$m\left(B_{r_n,\gamma} \cap T^{-j} B_{r_n,\gamma}\right) \leq m(B_{r_n,\gamma}) \exp(-u_n^\rho). \tag{5.5}$$

Summing over $1 \leq j \leq (\log n)^5$ and taking limits as $n \to \infty$ we obtain:

$$\begin{aligned}
& n \sum_{1}^{(\log n)^5} \int_{\gamma\in\Gamma_2} m(\pi(U_n)_\gamma \cap T^{-j}(\pi(U_n)_\gamma)) d\nu(\gamma) \\
\leq\ & n \sum_{1}^{(\log n)^5} e^{-u_n^\rho} \int_{\gamma\in\Gamma_2} m(\pi(U_n)_\gamma) d\nu(\gamma) \\
\leq\ & n \sum_{1}^{(\log n)^5} e^{-u_n^\rho} \mu(U_n) \\
=\ & (\log n)^5 e^{-u_n^\rho} e^{-v} \to 0,
\end{aligned}$$

since u_n has estimates in part(a) of Lemma 5.1.2.

And for Γ_1,

$$n \sum_{1}^{(\log n)^5} \int_{\gamma\in\Gamma_1} m(\pi(U_n)_\gamma \cap T^{-j}(\pi(U_n)_\gamma)) d\nu(\gamma) \leq n(\log n)^5 e^{-3u_n} \to 0.$$

Consequently we have

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

Finally, similarly to the argument in the case of Planar Dispersing Billiard Maps in [20, section 4.1.3], we use exponential decay of correlations to show

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

$\square$

5.2 Lorenz Flow

Let M be the Riemannian manifold, associated with Lorenz flows, endowed with a metric d_M, and $f_t : M \to M$ the Lorenz C^1-flow. $\Omega \subset M$ is a transverse cross-section of the flow which is a C^1-submanifold with boundary, as we stated in previous sections. We know $F : \Omega^* \to \Omega$ preserves a probability measure μ, where $\Omega = [-1/2, 1/2] \times [-1/2, 1/2]$ and $\Omega^* = ([-1/2, 1/2]\backslash\{0\}) \times [-1/2, 1/2]$. Let $h : \Omega \to \mathbb{R}_+$ be the first return time of the flow to Ω, and $h \in L^1(\mu)$. Consider the suspension space

$$\Omega^h = \{(p, u) \in \Omega \times \mathbb{R} \mid 0 \le u \le h(p)\ \}/\sim, \quad \text{where} \ \ (p, h(p)) \sim (F(p), 0).$$

We model the flow $f_t : M \to M$ in the standard way by the suspension flow $\tilde{f}_t : \Omega^h \to \Omega^h$, $\tilde{f}_t(p, u) = (p, u + t)/\sim$. Denote the metric on Ω by d_Ω, and we define a metric d_{Ω^h} on Ω^h by

$$d_{\Omega^h}((p, u), (q, v)) = \sqrt{d_\Omega(p, q)^2 + |u - v|^2}.$$

Then we can introduce a projection map $\pi_M : \Omega^h \to M$, $(p, t) \mapsto f_t(p)$, which is a local C^1-diffeomorphism. μ is an invariant ergodic probability measure for the first return map, i.e., our Lorenz map, $F : \Omega^* \to \Omega$. This induces (in the standard way) an invariant measure μ^h, on the suspension Ω^h, which is given by $d\mu \times dm/\bar{h}$ and $\bar{h} = \int_\Omega h d\mu$. Then μ^h determines a f_t-invariant measure μ_M on M by $\mu_M(A) = \mu^h(\pi_M^{-1} A)$ for measurable sets A.

Consider a measurable observation $\varphi : \Omega^h \to \mathbb{R}$ such that $\varphi(x) = -\log d_{\Omega^h}(x, x_0)$ where x_0 is any point in Ω^h, then φ has a logarithmic singularity at x_0. Define $\Phi : \Omega \to \mathbb{R}$ by

$$\Phi(p) := \max\{\varphi(\tilde{f}_s(p, 0)) \mid 0 \le s < h(p)\}.$$

Denote

$$\varphi_t(p) := \max\{\varphi(\tilde{f}_s(p, 0)) | 0 \le s < t\};$$

$$\Phi_n(p) := \max\{\Phi(F^k(p)) | 0 \le k < n\}.$$

Then we have our main theorem in the flow case:

Theorem 5.2.1. *Assume that F is the Lorenz map and f_t is the corresponding Lorenz flow. Assume the levels $\{u_n\}$ satisfy*

$$n\mu(\Phi_0 > u_n) \to e^{-v},$$

which gives some normalizing constants $a_n > 0$ and b_n such that

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

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

Then Φ_n satisfies a Type I extreme value law:

$$a_n(\Phi_n - b_n) \xrightarrow{d} e^{-e^{-v}},$$

and φ_t also satisfies a Type I extreme value law,

$$a_{\lfloor t/\bar{h} \rfloor}(\varphi_t - b_{\lfloor t/\bar{h} \rfloor}) \xrightarrow{d} e^{-e^{-v}}.$$

Proof. It is a consequence of [26, Theorem 2.6] and Corollary 5.0.5. In [26, Sublemma 4.18], they take care of the issue that h is not bounded. □

Remark 5.2.2. *The sequence $\{u_n\}$ in Theorem 5.2.1 is a linear scaling of v, based on the argument earlier in this paper. For u_n to be a nonlinear function of v, the situations could be very complicated, and we barely know anything about the limiting type when u_n is nonlinear function of v.*

Appendix A

A.1 Gal-Koksma Theorem.

We recall the following result of Gal and Koksma as formulated by W. Schmidt [38, 39] and stated by Sprindzuk [41]:

Proposition A.1.1. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and let $f_k(\omega)$, $(k = 1, 2, \dots)$ be a sequence of non-negative μ measurable functions and g_k, h_k be sequences of real numbers such that $0 \le g_k \le h_k \le 1$, $(k = 1, 2, \dots,)$. Suppose there exists $C > 0$ such that*

$$\int \left(\sum_{m<k\le n} (f_k(\omega) - g_k) \right)^2 d\mu \le C \sum_{m<k\le n} h_k \qquad (*)$$

for arbitrary integers $m < n$. Then for any $\epsilon > 0$

$$\sum_{1\le k\le n} f_k(\omega) = \sum_{1\le k\le n} g_k + O(\Theta^{1/2}(n) \log^{3/2+\epsilon} \Theta(n))$$

for μ a.e. $\omega \in \Omega$, where $\Theta(n) = \sum_{1\le k\le n} h_k$.

A.2 (SP) Property

This is a concept extracted from Gal-Koksma Theorem (Appendices, section A.1). Let $(\Omega, \mathcal{B}, \mu, F)$ be a dynamical system and $\{B_n\}$ be a sequence of nested balls. Define $f_k = \mathbf{1}_{B_k} \circ F^k$, then $E(f_k) = \mu(B_k)$. We say (SP) property is satisfied, if for all $n > m$, we have

$$\sum_{i=m}^{n} \sum_{j=i+1}^{n} (E(f_i f_j) - E(f_i)E(f_j)) \leq C \sum_{i=m}^{n} E(f_i),$$

for some constant C.

A.3 Assumption (C) for expanding systems

In this appendix we show that if the invariant measure μ has a density $\rho(x)$ with respect to Lebesgue measure m then Assumption (C) is valid. Recall we define

$$\mathscr{E}_k(\epsilon) := \{x : d(T^k x, x) \leq \epsilon\}$$

Lemma A.3.1. *Let $B_i(x)$ denote a decreasing sequence of balls with center x and suppose $\mu(B_i(x)) \leq \frac{C_2}{i^\gamma}$ for some constants C_1, C_2 and $0 < \gamma \leq 1$. Suppose μ has a density ρ with respect to Lebesgue measure m with support X and there exists $C_7 > 0$, $\delta > 0$ such that for all k, ϵ,*

$$\mu(\mathscr{E}_k(\epsilon)) < C_7 \epsilon^\delta$$

Then for μ a.e. $p \in X$ there exists $\eta(p) \in (0, 1)$ and $\kappa > 1$ such that for all i sufficiently large

$$\mu(B_i(p) \cap T^{-r} B_i(p)) \leq \mu(B_i(p))^{1+\eta}$$

for all $r = 1, \dots, \log^\kappa i$.

Proof. Let $\rho(x) = \frac{d\mu}{dm}(x)$ be the density of μ with respect to Lebesgue measure m.

Let $\sigma \geq 1$ and $\gamma > \sigma$. We choose ϵ_k so that for all x a ball of radius ϵ_k about x, denoted $B(x, \epsilon_k)$, satisfies $c_1/k^\sigma \leq m(B(x, \epsilon_k)) \leq c_2/k^\sigma$, so $\epsilon_k \simeq k^{-\sigma/D}$ where D is the dimension of X and $c_1, c_2 > 0$.

Let $\kappa > 1$ and define $A_k := \{x : d(T^j x, x) \leq \epsilon_k \text{ for some } 1 \leq j \leq \log^\kappa k\}$. Evidently $A_k \subset \bigcup_{j=1}^{\log^\kappa k} \mathscr{E}_j$. By the estimate on $\mathscr{E}_k(\epsilon)$ for all large k, $\mu(A_k) \leq c_3 \epsilon_k^\tau$ where $\tau < \delta$. Let

$$F_k := \{x : \mu(B(x, \epsilon_k) \cap A_k) \geq 1/k^\gamma\}$$

and define the Hardy-Littlewood maximal function M_k for $\phi(x) = 1_{A_k}(x)\rho(x)$ by

$$M_k(x) := \sup_{a>0} \frac{1}{m(B_a(x))} \int_{B_a(x)} 1_{A_k}(y)\rho(y)\, dm(y).$$

If $x \in F_k$ then $M_k > c_2^{-1} k^{\sigma-\gamma}$.

A theorem of Hardy and Littlewood ([11] Theorem 3.17) states that

$$m(|M_k| > C) \leq c_4 \frac{\|1_{A_k}\rho\|_1}{C}$$

for some constant c_4, where $\|\cdot\|_1$ is the $\mathscr{L}^1$ norm with respect to m. Hence

$$m(F_k) \leq m(M_k > c_2^{-1} k^{\sigma-\gamma}) \leq c_4 \mu(A_k) c_2 k^{\gamma-\sigma} \leq k^{\gamma-\sigma(1+\tilde{\tau})}$$

where $0 < \tilde{\tau} < \tau/D$. We need to alter τ/D to $\tilde{\tau}$ to take into account the fact that a ball of radius ϵ has measure roughly ϵ^D.

Choosing σ large enough that $\sigma\tilde{\tau} > 1$ and then taking $\sigma < \gamma < \sigma - 1 + \sigma\tilde{\tau}$ the series $\sum_k m(F_k)$ converges.

So for m a.e. x_0 there exists an $N(x_0)$ such that $x_0 \notin F_k$ for all $k > N(x_0)$. Hence for $k > N(x_0)$, $\mu(B(x, \epsilon_k) \cap A_k) \leq 1/k^\gamma$, thus $\mu(B(x, \epsilon_k) \cap A_k) \leq m(B(x, \epsilon_k))^{1+\eta}$ for some $\eta > 0$ (recall $m(B(x, \epsilon_k)) \simeq \frac{1}{k^\sigma}$ and $\gamma > \sigma$).

Furthermore by the Lebesgue differentiation theorem for m a.e. x

$$\lim_{\epsilon \to 0} \frac{1}{m(B_\epsilon(x))} \int_{B_\epsilon(x)} \rho(y) dm = \rho(x)$$

and for μ a.e. x , $\rho(x) > 0$ as X is the support of μ. Hence for m a.e. x_0 there exists an $\tilde{N}(x_0)$ and $\tilde{\eta} > 0$ such that for all $k > \tilde{N}(x_0)$ we have $\mu(B(x, \epsilon_k) \cap A_k) \leq \mu(B(x, \epsilon_k))^{1+\tilde{\eta}}$.

As κ was arbitrary by interpolating between the sequence ϵ_k we have that for μ a.e. $x \in X$ there exists $\eta' > 0$, $\kappa' > 1$ such that

$$\mu(B_i(x) \cap T^{-r} B_i(x)) \leq \mu(B_i(x))^{1+\eta'}$$

for $1 \leq r \leq \log^{\kappa'} i$. This is Assumption (C). $\square$

A.4 Geometric Lorenz Model

Consider a linear system in $[-1, 1]^3$:

$$(\dot{x}, \dot{y}, \dot{z}) = (\lambda_1 x, \lambda_2 y, \lambda_3 z),$$

with λ_1, λ_2, λ_3 satisfying

$$0 < \frac{\lambda_1}{2} \leq -\lambda_3 < \lambda_1 < -\lambda_2.$$

For any initial point $(a, b, c) \in \mathbb{R}^3$ near the equilibrium $(0, 0, 0)$, the trajectories are given by

$$\tilde{L}_t(a, b, c) = (ae^{\lambda_1 t}, be^{\lambda_2 t}, ce^{\lambda_3 t}),$$

where $\tilde{L}_t$ denotes the linear flow.

Consider $\Omega = \{(x, y, 1) : |x| \leq \frac{1}{2},\ |y| \leq \frac{1}{2}\} = \Omega^- \cup \Omega^o \cup \Omega^+$, where

$$\Omega^- = \{(x, y, 1) \in \Omega : x < 0\},$$

$$\Omega^+ = \{(x, y, 1) \in \Omega : x > 0\},$$

$$\Omega^o = \{(x, y, 1) \in \Omega : x = 0\}.$$

Ω is a transverse section to the linear flow $\tilde{L}_t$, and since $\lambda_3 < 0$, every trajectory, that would cross Ω, will cross in the direction of the negative z axis. Let $\Omega^* = \Omega^- \cup \Omega^+$ and let $\tilde{\Omega} = \{(x, y, z) : |x| = 1\} = \tilde{\Omega}^- \cup \tilde{\Omega}^+$ with $\tilde{\Omega}^{\pm} = \{(x, y, z) : x = \pm 1\}$. For each $(a, b, 1) \in \Omega^*$, the time t such that $\tilde{L}_t(a, b, 1) \in \tilde{\Omega}$ is given by

$$|ae^{\lambda_1 t}| = 1 \implies t(a) = -\frac{1}{\lambda_1} \log |a|,$$

the time only depends on the first component of the point in Ω^* and $t(a) \to \infty$ as $a \to 0$. Thus, we can express the point in $\tilde{\Omega}$ mapped from point $(a, b, 1) \in \Omega^*$ explicitly:

$$\tilde{L}_{t(a)}(a, b, 1) = (sgn(a), be^{\lambda_2 t(a)}, e^{\lambda_3 t(a)}) = (sgn(a), b|a|^{-\frac{\lambda_2}{\lambda_1}}, |a|^{-\frac{\lambda_3}{\lambda_1}}),$$

where $sgn(a) = a/|a|$ for $a \neq 0$. In this way, we just defined a map $L : \Omega^* \to \tilde{\Omega}^{\pm}$ by

$$L(x, y, 1) = (sgn(x), y|x|^{\beta}, |x|^{\alpha}),$$

where $\beta = -\frac{\lambda_2}{\lambda_1}$, $\alpha = -\frac{\lambda_3}{\lambda_1}$ satisfying $\frac{1}{2} < \alpha < 1 < \beta$, since $0 < \frac{\lambda_1}{2} \leq -\lambda_3 < \lambda_1 < -\lambda_2$.

Then we should let the sets $L(\Omega^*)$ return to the cross section Ω through a flow defined by a suitable composition of a rotation $R_\pm$, an expansion $E_{\pm\theta}$ and a translation $\mathbb{T}_\pm$. More precisely, for $(x, y, z) \in \tilde{\Omega}^\pm$,

$$R_\pm(x, y, z) = \begin{pmatrix} 0 & 0 & \pm 1 \\ 0 & 1 & 0 \\ \pm 1 & 0 & 0 \end{pmatrix},$$

and for $(x, y, z) \in \Omega$,

$$E_{\pm\theta}(x, y, z) = \begin{pmatrix} \theta & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

θ and $\mathbb{T}_\pm$ shall be chosen to satisfy certain conditions.

So the Poincaré first return map, i.e. our Lorenz map, $F : \Omega^* \to \Omega$, is defined as

$$F(x, y) = \begin{cases} \mathbb{T}_+ \circ E_{+\theta} \circ R_+ \circ L(x, y, 1) & \text{for } x > 0 \\ \mathbb{T}_- \circ E_{-\theta} \circ R_- \circ L(x, y, 1) & \text{for } x < 0 \end{cases}$$

Combining the effect of the rotation with expansion and translation, F must have the form:

$$F(x, y) = (T(x), G(x, y)),$$

where $T : I \backslash \{0\} \to I$ and $G : (I \backslash \{0\}) \times I \to I$, where $I = [-\frac{1}{2}, \frac{1}{2}]$. Here T is given by

$$T(x) = \begin{cases} f_1(x^\alpha), & x < 0 \\ f_0(x^\alpha), & x > 0 \end{cases}$$

with $f_i(x) = (-1)^i\theta \cdot x + b_i, i \in \{0, 1\}$ such that $\theta \cdot (\frac{1}{2})^\alpha < 1$ and $\theta \cdot \alpha \cdot 2^{1-\alpha} > 1$. T is the quotient map of F, usually referred to as the Lorenz like map, see Figure A.1. It has the following properties:

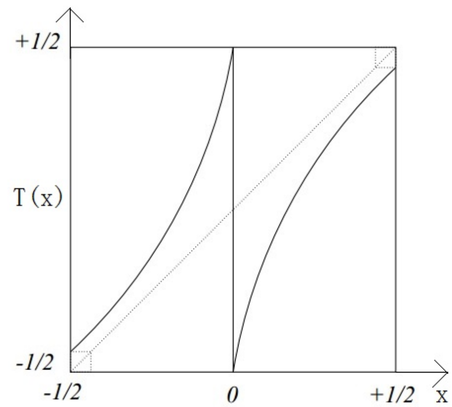

Figure A.1: Lorenz like map T

1. T is discontinuous at $x = 0$, and the lateral limits $T(0^{\pm})$ do exist, $T(0^{\pm}) = \mp\frac{1}{2}$.
2. T is C^2 on $I\backslash\{0\}$ and $T'(x) > 1$ for all $x \in I\backslash\{0\}$.
3. $\lim_{x\to 0} T'(x) = +\infty$.

And G is given by

$$G(x, y) = \begin{cases} g_1(x^\alpha, y \cdot x^\beta), & x < 0 \\ g_0(x^\alpha, y \cdot x^\beta), & x > 0 \end{cases}$$

where $g_1|I^- \times I \to I$ and $g_0|I^+ \times I \to I$ are suitable affine maps. Here $I^- = (-1/2, 0)$, $I^+ = (1/2, 0)$.

Bibliography

[1] Jon Aaronson and Manfred Denker. Local limit theorems for partial sums of stationary sequences generated by gibbs–markov maps. *Stochastics and Dynamics*, 1(02):193–237, 2001.

[2] V. S. Afraĭmovich, V. V. Bykov, and L. P. Shil′nikov. On attracting structurally unstable limit sets of Lorenz attractor type. *Trudy Moskov. Mat. Obshch.*, 44:150–212, 1982.

[3] V. S. Afraĭmovich and Ya. B. Pesin. Dimension of Lorenz type attractors. In *Mathematical physics reviews, Vol. 6*, volume 6 of *Soviet Sci. Rev. Sect. C Math. Phys. Rev.*, pages 169–241. Harwood Academic Publ., Chur, 1987.

[4] V Araujo, I Melbourne, and P Varandas. Rapid mixing for the lorenz attractor and statistical limit laws for their time-1 maps. *arXiv preprint arXiv:1311.5017*, 2013.

[5] Abraham Boyarsky et al. *Laws of chaos.* Springer, 1997.

[6] Nikolai Chernov and Dmitry Kleinbock. Dynamical borel-cantelli lemmas for gibbs measures. *Israel Journal of Mathematics*, 122(1):1–27, 2001.

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

[8] Jean-Pierre Conze, Albert Raugi, et al. Limit theorems for sequential expanding dynamical systems on [0, 1]. *Contemporary Mathematics*, 430:89–122, 2007.

[9] Jiu Ding and Aihui Zhou. *Statistical properties of deterministic systems.* Springer Science & Business Media, 2010.

[10] Rick Durrett. *Probability: theory and examples.* Cambridge university press, 2010.

[11] GB Folland. Real analysis. 1999.

[12] Ana Cristina Moreira Freitas and Jorge Milhazes Freitas. On the link between dependence and independence in extreme value theory for dynamical systems. *Statistics & Probability Letters*, 78(9):1088–1093, 2008.

[13] Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, and Mike Todd. Hitting time statistics and extreme value theory. *Probab. Theory Related Fields*, 147(3-4):675–710, 2010.

[14] Jorge Milhazes Freitas, Nicolai Haydn, and Matthew Nicol. Convergence of rare events point processes to the poisson for billiards. *arXiv preprint arXiv:1311.2649*, 2013.

[15] S. Galatolo and Maria José Pacifico. Lorenz-like flows: exponential decay of correlations for the Poincaré map, logarithm law, quantitative recurrence. *Ergodic Theory Dynam. Systems*, 30(6):1703–1737, 2010.

[16] Stefano Galatolo, Jérôme Rousseau, and Benoît Saussol. Skew products, quantitative recurrence, shrinking targets and decay of correlations. *arXiv preprint arXiv:1109.1912*, 2011.

[17] Mikhail I Gordin. Central limit theorem for stationary processes. *Doklady Akademii Nauk SSSR*, 188(4):739, 1969.

[18] Sébastien Gouëzel. A borel–cantelli lemma for intermittent interval maps. *Nonlinearity*, 20(6):1491, 2007.

[19] John Guckenheimer and R. F. Williams. Structural stability of Lorenz attractors. *Inst. Hautes Études Sci. Publ. Math.*, (50):59–72, 1979.

[20] Chinmaya Gupta, Mark Holland, and Matthew Nicol. Extreme value theory and return time statistics for dispersing billiard maps and flows, Lozi maps and Lorenz-like maps. *Ergodic Theory Dynam. Systems*, 31(5):1363–1390, 2011.

[21] Chinmaya Gupta, Matthew Nicol, and William Ott. A borel–cantelli lemma for nonuniformly expanding dynamical systems. *Nonlinearity*, 23(8):1991, 2010.

[22] Peter Hall and Christopher C Heyde. *Martingale limit theory and its application*. Academic press New York, 1980.

[23] Nicolai Haydn. Convergence of the transfer operator for rational maps. *Ergodic Theory and Dynamical Systems*, 19(03):657–669, 1999.

[24] Nicolai Haydn, Matthew Nicol, Tomas Persson, and Sandro Vaienti. A note on borel–cantelli lemmas for non-uniformly hyperbolic dynamical systems. *Ergodic Theory and Dynamical Systems*, 33(02):475–498, 2013.

[25] Nicolai Haydn, Matthew Nicol, Sandro Vaienti, and Licheng Zhang. Central Limit Theorems for the Shrinking Target Problem. *J. Stat. Phys.*, 153(5):864–887, 2013.

[26] Mark Holland, Matthew Nicol, and Andrei Török. Extreme value theory for non-uniformly expanding dynamical systems. *Transactions of the American Mathematical Society*, 364(2):661–688, 2012.

[27] Johannes Jaerisch, Marc Kesseböhmer, and Bernd O Stratmann. A fr\'{e} chet law and an erd\" os-philipp law for maximal cuspidal windings. *arXiv preprint arXiv:1109.3583*, 2011.

[28] Gerhard Keller. Generalized bounded variation and applications to piecewise monotonic transformations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 69(3):461–478, 1985.

[29] Dong Han Kim. The dynamical borel-cantelli lemma for interval maps. *Discrete and Continuous Dynamical Systems*, 17(4):891, 2007.

[30] DY Kleinbock and Gregory A Margulis. Logarithm laws for flows on homogeneous spaces. *Inventiones mathematicae*, 138(3):451–494, 1999.

[31] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, 1983.

[32] Carlangelo Liverani. Central limit theorem for deterministic systems. *Pitman Research Notes in Mathematics Series*, pages 56–75, 1996.

[33] Edward N Lorenz. Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141, 1963.

[34] Péter Nándori, Domokos Szász, and Tamás Varjú. A central limit theorem for time-dependent dynamical systems. *Journal of Statistical Physics*, 146(6):1213–1220, 2012.

[35] Magda Peligrad. Central limit theorem for triangular arrays of non-homogeneous markov chains. *Probability Theory and Related Fields*, 154(3-4):409–428, 2012.

[36] Walter Rudin. *Real and complex analysis*. Tata McGraw-Hill Education, 1987.

[37] Benoît Saussol. Absolutely continuous invariant measures for multidimensional expanding maps. *Israel Journal of Mathematics*, 116(1):223–248, 2000.

[38] Wolfgang Schmidt. A metrical theorem in diophantine approximation. *Canad. J. Math.*, 12:619–631, 1960.

[39] Wolfgang M. Schmidt. Metrical theorems on fractional parts of sequences. *Trans. Amer. Math. Soc.*, 110:493–518, 1964.

[40] Sunder Sethuraman and SRS Varadhan. A martingale proof of dobrushin's theorem for non-homogeneous markov chains. *Electron. J. Probab*, 10(36):1221–1235, 2005.

[41] Vladimir G. Sprind*v*zuk. *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.

[42] Warwick Tucker. The Lorenz attractor exists. *C. R. Acad. Sci. Paris Sér. I Math.*, 328(12):1197–1202, 1999.

[43] Warwick Tucker. A rigorous ODE solver and Smale's 14th problem. *Found. Comput. Math.*, 2(1):53–117, 2002.