ABSTRACT

Title of Dissertation: POISSON LIMIT THEOREMS IN DYNAMICAL SYSTEMS AND ERGODIC SUMS OF NON-INTEGRABLE OBSERVABLES

Max Auer
Doctor of Philosophy, 2025

Dissertation Directed by: Professor Dmitry Dolgopyat
Department of Mathematics

This thesis investigates statistical properties of rare events in dynamical systems, with a particular focus on recurrence statistics and limit laws for ergodic sums of heavy-tailed observables. A central theme is the Poisson Limit Theorem (PLT), which describes the limiting distribution of return times to small sets, and its role in establishing asymptotics for extreme events. Concretely, this thesis addresses the following topics:

First, we construct examples of non-mixing dynamical systems that still satisfy the PLT, demonstrating that strong statistical properties can persist even in the absence of mixing. This is significant because, until now, all known systems satisfying the PLT have been mixing, with existing proofs relying on mixing assumptions.

Second, we establish strong limit laws for ergodic sums of non-integrable observables. In such cases, large fluctuations due to close visits to the singularity typically obstruct strong limit laws. To address this, we employ trimming, a well-known probabilistic technique, by systematically removing the closest visits to the singularity. We prove trimmed strong laws for irrational rotations, emphasizing once more that our results seem to be the first of their kind that do not rely on mixing.

We also present in the conclusion some natural questions and directions a future research, including hitting time statistics, similar to the PLT setting, and distributional limit theorems for ergodic sums of non-integrable observables.

POISSON LIMIT THEOREMS IN DYNAMICAL SYSTEMS AND ERGODIC SUMS OF NON-INTEGRABLE OBSERVABLES

by

Max Auer

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Advisory Committee:
Professor Professor Dmitry Dolgopyat, Chair/Advisor
Professor Adam Kanigowski, Co-Advisor
Professor Bassam Fayad
Professor Leonid Koralov
Professor David Mount

Foreword

A version of Chapter 2 has been published [M. Auer, Poisson limit theorems for systems with product structure, Discrete Contin. Dyn. Syst. 45 (2025), no. 5, 1454–1491]. I am the sole author of this work, and responsible for all its contents.

A version of Chapter 3 has been submitted for publication at J. London Math. Soc. by myself and T. Schindler, and is available as a preprint at https://arxiv.org/abs/2503.22242. I was responsible for the majority of the proofs, concept formation and manuscript composition. T. Schindler contributed to the background and manuscript edits.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Dmitry Dolgopyat, for his invaluable guidance and insightful feedback throughout the course of my doctoral studies. His mentorship has shaped both the direction of this thesis and my development as a researcher.

I am also sincerely grateful to my co-advisor, Professor Adam Kanigowski, for his support and time. Our discussions and his insights have had a significant influence on the direction of my research. Many parts of this thesis, as well as my other work, would not be the same without him.

Many thanks go to my scientific colleagues and collaborators, especially Dr. Tanja Schindler, who co-authored the publication on which Chapter 3 is based. I would also like to thank Professor Roland Zweimüller, who supported and encouraged me early on and whose mentorship during my studies in Austria had a lasting impact on my academic path.

I gratefully acknowledge the administrative and technical staff in the Department of Mathematics at the University of Maryland for their assistance and for helping create such a warm and supportive environment.

Finally, I am deeply grateful to my family and friends for their constant love, patience, and encouragement. Their belief in me has been a source of strength throughout both the challenges and joys of this journey.

None of this would have been possible without the incredible people in my life. Thank you all.

Table of contents

Chapter 1

Motivation and introduction

1.1 Motivation

Rare events play a crucial role in understanding the long-term behaviour of complex systems, with applications spanning physics, finance, climate science, and engineering. In dynamical systems, such events often correspond to extreme deviations from typical behaviour, including large fluctuations in turbulent flows, financial crashes, or sudden transitions in climate models. While individually infrequent, rare events can have an outsized impact, making their statistical properties essential to study. Mathematical techniques from probability theory and ergodic theory provide powerful tools to analyse their occurrence and structure, offering insights into both theoretical foundations and practical predictions.

A central question in this area is to describe the frequency with which events of low probability occur. This problem has been extensively studied in chaotic systems, where strong mixing properties allow for relatively straightforward probabilistic arguments. However, in many important classes of dynamical systems, particularly those where chaotic behaviour is weak and mixing is slow, traditional techniques relying on fast decay of correlations cannot be applied. Understanding the statistical properties of rare events in such settings requires new approaches.

Another key aspect of rare events in dynamical systems arises in the study of ergodic sums of heavy-tailed observables. In these cases, time averages over long trajectories can be dominated by just a few visits to regions where the observable takes exceptionally large values. This leads to large oscillations in the averages and strong dependence on the initial condition. A natural way to mitigate this issue is trimming, a technique that removes the largest observations from the sum. Trimming has been successfully used to analyse such sums and obtain meaningful limit theorems. While this approach has been well studied in certain settings, its behaviour in systems with low complexity, such as irrational rotations, remains less understood.

1.2 Background

Probably the earliest fundamental result concerning rare events is the Poincaré Recurrence Theorem. The Poincaré Recurrence Theorem states that every event of positive probability will almost certainly occur infinitely often. More precisely;

Theorem 1. *Let (X, μ, T) be a probability preserving dynamical system, and let $A \subset X$ be a measurable set with positive probability $\mu(A) > 0$. Then almost every point $x \in A$ revisits A infinitely often and it holds that*

$$\mu(x \in A \mid T^n(x) \in A \text{ only for finitely many } n) = 0.$$

Corollary 2. *If, in addition (X, μ, T) is ergodic, then every point $x \in X$ visits A infinitely often and it holds that*

$$\mu(x \in X \mid T^n(x) \in A \textit{ for infinitely many } n) = 1.$$

A natural follow-up question is with what frequency rare events occur. Therefore, for an event A of positive probability $\mu(A) > 0$, denote by φ_A the time until a point first visits A, formally

$$\varphi_A(x) = \min(n \geq 1 \mid T^n(x) \in A).$$

By the Poincaré Recurrence Theorem, it holds that $\varphi_A < \infty$ almost surely, provided that T is ergodic. A fortiori the average value of φ_A on A itself is $\frac{1}{\mu(A)}$, as described by Kac's formula

$$\int_A \varphi_A \, \mathrm{d}\mu_A = \frac{1}{\mu(A)},$$

where μ_A denotes the measure conditional on A given as $\mu_A(B) = \frac{\mu(A \cap B)}{\mu(A)}$ for measurable $B \subset A$.

The Poisson Limit Theorem describes the asymptotic behaviour of $\mu(A)\varphi_A$ as $\mu(A) \to 0$. Loosely speaking, it states that, for regular enough sets A and chaotic enough T it holds that

$$\mu(A)\varphi_A \Rightarrow \phi_{Exp} \quad \text{as } \mu(A) \to 0,$$

where ϕ_{Exp} is a standard exponential distribution, and $\Rightarrow$ denotes convergence in distribution. To avoid overly technical statements at this point, we leave precise statements to Chapter 2.

Beyond recurrence and return times, the Poisson Limit Theorem has significant applications in the study of ergodic sums of non-integrable observables $f \notin L^1$. As already described in §1.1, the ergodic sums $S_N(f) = \sum_{n=0}^{N-1} f \circ T^n$, for $N \geq 1$, are dominated by only a few large observations. This leads to large deviations in the values of $S_N(f)(x)$ as we vary x, making almost sure convergence[1], after suitable normalisation, to a constant, as in the Birkhoff Ergodic Theorem impossible. In the following, whenever there are constant $d_N > 0$ such that

$$\lim_{N \to \infty} \frac{S_N(f)(x)}{d_N} = 1 \quad \text{for almost every } x,$$

we will refer to it as a *strong law*, whereas if it holds that

$$\lim_{N \to \infty} \mu\left(\left| \frac{S_N(f)}{d_N} - 1 \right| > \epsilon \right) = 0 \quad \forall \epsilon > 0,$$

we will refer to it as a *weak law.*

This is accentuated by the stable limit theorem obtained by Kolmogorov and Gnedenko [GK68], which we state for simplicity only for the function $x^{-\beta}$, $\beta \geq 1$, on the unit interval $[0, 1]$ with the Lebesgue measure. In the setting of i.i.d.s[2] they obtain distributional convergence to a stable limit distribution, more explicitly

$$\frac{S_N(f) - N\log(N)}{N} \Rightarrow \Delta_1 \quad \text{as } N \to \infty,$$

for $\beta = 1$, and

$$\frac{S_N(f)}{N^\beta} \Rightarrow \Delta_\beta \quad \text{as } N \to \infty,$$

[1] As we will see in the next paragraph convergence in probability might still be possible.
[2] Meaning that $(f \circ T^n)_{n \geq 1}$ is an i.i.d. random process.

for $\beta > 1$, where Δ_β is a non-degenerate stable distribution. We shall sketch a proof that makes use heavily of the Poisson Limit Theorem in §4.3. Note that the above, for $\beta > 1$, already implies that there cannot even be a weak law of large numbers. In contrast, if $\beta = 1$, then a weak law does hold with $d_N = N\log(N)$.

Now we turn our attention to strong laws. As elaborated earlier, if f is non-integrable, big oscillations prohibit any strong laws for $S_N(f)$ itself. This is made precise by the following result of Aaronson ([Aar77]); let $f \geq 0$ be non-negative and non-integrable, then, for every normalising sequence $(d_N)_{N\geq 1}$, it holds that

$$\limsup_{N\to\infty} \frac{S_N(f)(x)}{d_N} = \infty \quad \text{or} \quad \liminf_{N\to\infty} \frac{S_N(f)(x)}{d_N} = 0 \quad \text{almost surely,}$$

meaning that we either underestimate or overestimate the sum infinitely often. A common method to deal with the big oscillations of $S_N(f)$ is to simply remove the largest observations. For $N \geq 1$ and $0 \leq k(N) \leq N$, define the *trimmed sum* as

$$S_N^{k(N)}(f)(x) = S_N(f)(x) - \max_{0\leq n_1<n_2<\ldots<n_{k(N)}\leq N-1} \sum_{j=1}^{k(N)} f(T^{n_j}x).$$

As it turns out, for the trimmed sum $S_N^{k(N)}(f)$ and a suitable $k(N)$, it **is** possible to obtain strong laws. We will review relevant literature for iids and systems with strong mixing properties in §3.1.2, and show strong trimmed laws for rotations in Chapter 3.

1.3 Outline of the thesis

The outline of this thesis is as follows:

In Chapter 2, we delve into the problem of asymptotic hitting times to shrinking targets, providing all relevant definitions. This Chapter consists of my paper [Aue25], in which I establish conditions under which the Poisson limit theorem (PLT) holds in systems that lack strong chaotic behaviour. This extends the applicability of such limit laws beyond the usual settings of hyperbolic or strongly mixing systems.

Chapter 3 focuses on trimmed ergodic sums for rotations. This Chapter consists of the paper [AS25], co-authored with T. Schindler, where we study trimming in the context of irrational rotations of the torus. We provide precise conditions under which trimmed strong and weak laws hold for ergodic sums of non-integrable observables in this setting.

Finally, Chapter 4 presents related open problems and potential directions for future research.

Minor changes have been made to the source papers [Aue25] and [AS25] to improve readability. Chapters 2-4 are self-contained.

Chapter 2

Poisson limit theorems for systems with product structure

In this Chapter, we obtain a Poisson Limit for return times to small sets for product systems. Only one factor is required to be hyperbolic, while the second factor is only required to satisfy polynomial deviation bounds for ergodic sums. In particular, the second factor can be elliptic or parabolic. As an application of our main result, several maps of the form Anosov map $\times$ another map are shown to satisfy a Poisson Limit Theorem at typical points, some even at all points. The methods can be extended to certain types of skew products, including T, T^{-1}-maps of high rank.

2.1 Introduction

One of the prominent limit theorems in classical probability theory is the Poisson Limit Theorem. (PLT). Due to the PLT, a variety of probabilistic models describing waiting times until unlikely events occur are well approximated by exponentially distributed variables. It has been a great discovery that many deterministic systems satisfy the same kind of limit theorems for rare events.

Limit distributions of waiting times are most classical for mixing Markov chains, where one considers returns to small cylinders, for example, see [Pit91, Theorem A]. As remarked there, this result can be immediately generalized to systems with a Markov partition, the only caveat being that the sets are still cylinders, so geometrically not the most intuitive class. Nonetheless, waiting time limits for returns to small balls can be shown in concrete settings; for example hyperbolic toral automorphisms [DGS04, Theorem 2.3] or more general hyperbolic maps [BS22, Theorem 2.8], Rychlik-maps and unimodal maps [BSTV03, Theorem 3.2 and 4.1], partially or nonuniformly hyperbolic maps [Dol04, Theorem 8] [PS16, Theorem 3.3] [CC13, Theorem 3.3], open billiard systems [BS24, Theorem 1], some intermittent interval maps [CG93, Main Theorem], and many more. It is sometimes interesting to also ask for explicit rates of convergence, this can be shown under strong mixing properties, see [HSV99, Theorem 2.1], [AS11, Theorem 7], [HV09, Theorem 8]. Interestingly enough, in [BS22, Theorem 2.8] and [BS23, Theorem 5.11] novel techniques have been used to obtain rates in certain billiard systems without strong mixing assumptions. We do not make any claims on completeness of the list of references given above, for a more complete picture see [LFF$^+$16].

Some related topics are extreme value laws [FFT11], [FFV17], spatiotemporal limits [PS20], [Zwe22] or Borel-Cantelli like Lemmas [HLSW22], [HNPV13], [DFL22].

Similar questions can be asked for flows as well, this topic has not been studied as thoroughly as the question for maps. As shown in [PY17] for suspension flows, this reduces to the study for maps. Moreover, the Poisson Limit Theorem for flows can be reduced to the Poisson Limit Theorem for time 1 map with the target being the set of points that visit $B(x, r)$ within the next unit of time.

From the list above, we see that the PLT is often associated with strong mixing properties of the system. In the present work, we construct systems that are not even weakly mixing but nevertheless satisfy the PLT (a precise definition is given beneath). The systems will have a special structure $S = T \times R$, where T is hyperbolic, but R is not.

We will develop a machinery to show the PLT for such systems. This will be used to construct systems satisfying the PLT, but otherwise exhibiting properties uncharacteristic of chaotic systems - like non weak mixing, or zero entropy[1]. This suggests that the PLT is much more common than it was believed before. In fact, discovering the most flexible conditions for the validity of the PLT is a promising direction of future research.

2.2 Preliminaries

Definition 3. *Given a probability-preserving ergodic dynamical system $(X, \mathcal{A}, \mu, T)$ and a measurable set $A \in \mathcal{A}$, we will define the first entry time to A as*

$$\varphi_A(x) = \min(n \geq 1 \mid T^n(x) \in A),$$

the restriction $\varphi_A|_A$ to the set A itself shall be referred to as the first return time to A. The first return map shall be denoted by $T_A(x) = T^{\varphi_A(x)}(x)$, and the sequence of consecutive return times by

$$\Phi_A = (\varphi_A, \varphi_A \circ T_A, \varphi_A \circ T_A^2, ...).$$

Hence, it is natural to study limits of $\mu(A)\varphi_A$ as $\mu(A) \to 0$. More explicitly let $(A_l)_{l\geq 1}$ be a sequence of rare events, that is each A_l is measurable with $\mu(A_l) \to 0$, we want to find weak limits of the form

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Rightarrow} \Phi \quad \text{as } l \to \infty,$$

or

$$\mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Rightarrow} \tilde{\Phi} \quad \text{as } l \to \infty,$$

where $\Rightarrow$ denotes convergence in distribution. In the above situation, we shall call Φ the *hitting time limit* and $\tilde{\Phi}$ the *return time limit*. An important fact is that the hitting and return time limits are intimately related, this relation was first formulated in [HLV05, Main Theorem] (albeit only for the first marginal). The analogous relation for the entire process is shown in [Zwe16, Theorem 3.1]. For exponential returns, which is what we are concerned with, the result is as follows.

Theorem 4. *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving dynamical system, and let $(A_l)_{l\geq 1}$ be a sequence of rare events. Then $\Phi \overset{d}{=} \Phi_{Exp}$ is equivalent to $\tilde{\Phi} \overset{d}{=} \Phi_{Exp}$, where Φ_{Exp} is a process of i.i.d. standard exponentially distributed random variables.*

This suggests that we should expect exponential hitting and return time limits for geometrically sensible sequences of rare events.

In the following, let X be a C^r Riemannian manifold with $\dim(X) = d$ and assume $\mu \ll m_X$ the volume on X, with continuous density, say $\frac{\mathrm{d}\mu}{\mathrm{d}m_X} = \rho$. Most of the statements can be reformulated to hold for arbitrary invariant μ, but for the sake of simplicity, we shall keep this assumption.

[1]This cannot be done with products, since $h(T \times R) = h(T) + h(R)$, where h denotes the metric entropy of a system. We extend our methods to skew-products of a certain form (Theorem 9.)

Definition 5. *(i) Let $x^* \in X$ and, for $r > 0$, denote by $B_r(x^*)$ the geodesic ball of radius r centred at x^*. We will say that T satisfies the PLT at x^* if*

$$\mu(B_r(x^*))\Phi_{B_r(x^*)} \overset{\mu}{\Rightarrow} \Phi_{Exp} \quad \text{as } r \to 0.$$

(ii) Let

$$PLT := \{x^* \in X \mid T \text{ satisfies the PLT at } x^*\}.$$

If $\mu(PLT) = 1$ we say that T satisfies the PLT almost everywhere, and if $PLT = X$ we will say that T satisfies the PLT everywhere.

If T is Lipschitz-continuous along the (finite) orbit of a periodic point, then it **does not** satisfy the PLT at that point. To see this, note first that the PLT at x^* in particular implies, via Theorem 4, that, for each $N \geq 1$,

$$\mu_{B_r(x^*)}(\varphi_{B_r(x^*)} \leq N) \to 0 \quad \text{as } r \to 0.$$

Now suppose x^* is a point with period p, and say $\rho(x^*) > 0$, and $|T^p(x) - T^p(y)| \leq C|x - y|$ near x^* then[2]

$$\mu_{B_r(x^*)}(\varphi_{B_r(x^*)} \leq p) \geq \mu_{B_r(x^*)}(B_{\frac{r}{C}}(x^*)) = \frac{1}{C^d} + o(1)$$

as $r \to 0$.

Situations where x^* is a periodic point are more delicate, and the limiting distribution is not exponential any more (due to immediate returns). For example, in [Zwe22, Theorem 3.3 and Theorem 10.1] this question was studied for expanding interval maps.

The main goal is to prove the PLT (almost) everywhere for some (skew-) product systems.

In the following we will consider return times in different systems - namely, we will have three different maps $T : X \to X$ (or $T_y : X \to X$) which is usually assumed hyperbolic, $R : Y \to Y$ which is parabolic or elliptic, and $S = T \times R : X \times Y \to X \times Y$ - in an attempt to keep notation simple we will (by slight abuse of notation) always denote the return times by φ. Which map is meant will always be clear by the specified set.

2.3 The PLT for (skew-)products

In this paper, we study the PLT for systems that can be written as a product (or skew product of a special type). Therefore, let Y be another Riemannian $C^{r'}$-manifold with $\dim(Y) = d'$, and assume $R : Y \to Y$ preserves a probability measure $\nu \ll m_Y$ with continuous density. Instead of $T : X \to X$, consider now some[3] $T : X \times Y \to X$. We will prove the PLT for certain systems of the form $S(x, y) = (T(x, y), R(y))$. The case of direct products can be recovered if

[2]Assuming $C > 1$, we have that $B_r(x^*)$ is diffeomorphic via the exponential map exp_{x^*} to a ball in $\mathbb{R}^d$. W.l.o.g. assume that $X \subset \mathbb{R}^N$

$$\begin{aligned}
\mu(B_r(x^*)) &= \int_{B_r(x^*)} \rho(x)\, \mathrm{d}m_X(x) = (\rho(x^*) + o(1))m_X(B_r(x^*)) \\
&= (\rho(x^*) + o(1)) \int_{(B_r(0))} \sqrt{|\det D_{\mathbf{u}} exp_{x^*} (D_{\mathbf{u}} exp_{x^*})^t}\, \mathrm{d}\lambda^d(\mathbf{u}) \\
&= (\rho(x^*)\sqrt{|\det D_{\mathbf{0}} exp_{x^*} (D_{\mathbf{0}} exp_{x^*})^t} + o(1))\lambda^d(B_r(0)),
\end{aligned}$$

where λ^d is the d-dimensional Lebesgue measure.

[3]$X \times Y$ is considered as a Riemannian manifold with the natural (Euclidean) product metric $d((x, y), (x', y')) = \sqrt{d(x, x')^2 + d(y, y')^2}$. Analogously, one could consider different metrics, e.g. the box metric $d((x, x'), (y', y)) = \max(d(x, y), d(x', y'))$, where many of the proofs become easier. However, we will use here the Euclidean product metric as the most natural choice.

$T(x,y)=T(x)$ is independent of y (which will be the case for most of our examples). Denote also $T_y(x)=T(x,y)$. We will assume that T_y preserves a probability measure μ (independent of y). For measurable $A\subset X$ we introduce analogously the *consecutive fiberwise return times* as

$$\begin{aligned}&\varphi_{A\times Y}(x,y)=\min(j\geq 1\mid S^j(x,y)\in A\times Y),\\&\Phi_{A\times Y}=(\varphi_{A\times Y},\varphi_{A\times Y}\circ S_{A\times Y},\varphi_{A\times Y}\circ S^2_{A\times Y},\dots),\end{aligned}$$

where $S_{A\times Y}=S^{\varphi_{A\times Y}}$ is the first return map to $A\times Y$, note that we only fix a small target in the fiber.

For our purposes, it is convenient to think of y as fixed. For $n\geq 1$ denote $T_y^n(x)=T_{R^{n-1}(y)}(T_{R^{n-2}(y)}(...(T_y(x))))$, and define

$$\begin{aligned}&\varphi_{A,y}(x)=\varphi^{(1)}_{A,y}(x)=\min(j\geq 1\mid T_y^j(x)\in A),\\&\varphi^{(n+1)}_{A,y}(x)=\min(j\geq 1\mid T_y^{\varphi^{(1)}_{A,y}(x)+\varphi^{(2)}_{A,y}(x)+\cdots+\varphi^{(n)}_{A,y}(x)+j}(x)\in A),\\&\Phi_{A,y}=(\varphi^{(1)}_{A,y},\varphi^{(2)}_{A,y},\cdots).\end{aligned}$$

Clearly the definitions coincide and $\Phi_{A,y}(x)=\Phi_{A\times Y}(x,y)$.

We will list here the main assumptions[4] we make in order to prove the PLT.

- (MEM) We will say T is *multiple exponentially mixing* there are constants $r>0$, $C>1$ and $\gamma>0$ such that, for almost all $y\in Y$,

$$\begin{aligned}&\left|\int_X\prod_{j=0}^{n-1}f_j\circ T_y^{k_j}\,\mathrm{d}\mu-\prod_{j=0}^{n-1}\int_X f_j\,\mathrm{d}\mu\right|\\&\qquad\leq Ce^{-\gamma\min_{0\leq j_1<j_2\leq n-1}|k_{j_1}-k_{j_2}|}\prod_{j=0}^{n-1}||f_j||_{C^r},\end{aligned}\tag{2.1}$$

  for $n\geq 1$, $f_0,...,f_{n-1}\in C^r$ and $0\leq k_0\leq ...\leq k_{n-1}$.

- (EE) There are $r'>0$ and $\delta<1$ such that

$$\left|\left|\sum_{n=0}^{N-1}f\circ R^n-N\int f\,\mathrm{d}\nu\right|\right|_{L^2(\nu)}\leq C||f||_{C^{r'}}N^\delta\quad\forall f\in C^{r'},\ \forall N\geq 1.\tag{2.2}$$

- (LR(y^*)) There is a $c>0$ such that, for $r>0$ and ν-a.e $y\in B_r(y^*)$, we have

$$\varphi_{B_r(y^*)}(y)\geq c|\log(r)|.\tag{2.3}$$

- (SLR(y^*)) There is a $\psi:(0,\infty)\to(0,\infty)$ with $|\log(r)|=o(\psi(r))$ as $r\to 0$ such that, for $r>0$ and ν-a.e $y\in B_r(y^*)$, we have

$$\varphi_{B_r(y^*)}(y)\geq\psi(r).\tag{2.4}$$

- (LR'(x^*)) For ν-a.e $y\in Y$ there is a $c=c_y>0$ such that, for $r>0$ and μ-a.e $x\in B_r(x^*)$, we have

$$\varphi_{B_r(x^*),y}(x)\geq c|\log(r)|.\tag{2.5}$$

[4]We often only assume a subset of these, most commonly (MEM), (EE), and (BR(x^*,y^*)). But we will always state the current assumptions.

- (NSR(x^*)) There is a $\xi : (0,\infty) \to (0,\infty)$ with $|\log(r)| = o(\xi(r))$ as $r \to 0$ such that, for ν-a.e $y \in Y$, we have

$$\mu_{B_r(x^*)}(\varphi_{B_r(x^*),y} \leq \xi(r)) \to 0 \quad \text{as } r \to 0. \tag{2.6}$$

- (BR(x^*, y^*)) One of the following is satisfied
  - (SLR(y^*)),
  - (NSR(x^*)) AND (LR(y^*)),
  - or (NSR(x^*)) AND (LR'(x^*)).

Colloquially, we will also refer to (MEM) as multiple exponential mixing, and to (EE) as the Quantitative Ergodic Theorem or effective ergodicity. Both are standard assumptions and have been studied for many classes of systems.

Conditions (LR), (LR'), (SLR), and (NSR) all are concerned with the fact that points in a small ball B cannot return to B too quickly. Sometimes in literature, the center x^* or y^* is referred to as a slowly recurrent point. For technical reasons, we need to distinguish different versions of slow recurrence, (SLR) being the strongest. The abbreviations (LR), (SLR), (NSR), and (BR) stand for 'large returns', 'strong large returns', 'no short returns', and 'big returns' respectively.

Remark 6. *(i) In the case $T(x,y) = G_{\tau(y)}(x)$, where G is a flow satisfying (a continuous version of) (MEM)*[5] *and τ is bounded, the condition (NSR(x^*)) is satisfied at almost every x^*. Indeed, it was shown in [DFL22, Lemma 4.13], albeit for maps instead of flows, that condition (NSR(x^*)) is satisfied*[6] *for G at almost every x^*. Since τ is bounded, T also has this property.*

(ii) It is shown in [BS01, Lemma 5] that, for a map of positive entropy, condition (LR) is satisfied at almost every point (In fact (2.3) *is satisfied for all $y \in B_r(y^*)$). This remains true for maps of the form $T(x,y) = G_{\tau(y)}(x)$, (in this case (LR') is satisfied) for bounded τ, where G has positive entropy.*

(iii) Considering the previous remarks, it may seem unnecessary to state condition (SLR). Note however that none of the conditions can be satisfied at periodic points, and the maps we want to use for T will have plenty of periodic points. (SLR) will be useful to show the PLT ***everywhere****, if we can choose R without periodic points.*

(iv) In most of the examples (see §2.4) we will have

$$\left\| \sum_{j=0}^{n-1} f \circ R^j - n \int f \, d\nu \right\|_{L^2(\nu)} \leq C||f||_{H^{r'}} n^{\delta} \quad \forall f \in H^{r'}, \, \forall n \geq 1.$$

Since $C^{r'} \subset H^{r'}$ and $||f||_{H^{r'}} \leq ||f||_{C^{r'}}$ for $f \in C^{r'}$, this implies condition (EE).

[5] It is in fact enough to assume exponential mixing.

[6] It is shown that, for every fixed $A, K > 0$, we have

$$\mu_{B_r(x^*)}(B_r(x^*) \cap G^{-n}B_r(x^*)) \leq |\log(r)|^{-A} \quad \forall n \leq K|\log(r)|. \tag{2.7}$$

For $A > 1$, summing over $n \in [1, K|\log(r)|]$ yields

$$\mu_{B_r(x^*)}(\varphi_{B_r(x^*),G} \leq K|\log(r)|) \to 0 \quad \text{as } r \to 0.$$

Since this is true for all $K > 0$, we can easily replace K by some $K(r) \nearrow \infty$ growing slowly enough. This is a routine argument which is left to the reader.

Theorem 7. *Assume that $S(x,y) = (T(x,y), R(y))$ satisfies conditions (MEM), (EE), and (BR(x^*,y^*)) for some $(x^*,y^*) \in X \times Y$. If*

$$d > \frac{r' + d'\delta}{1-\delta} \tag{2.8}$$

then S satisfies the PLT at (x^*,y^*).

Corollary 8. *If $T(x,y) = T(x)$ preserves a smooth measure and satisfies (MEM), and R satisfies (EE), then $S = T \times R$ satisfies the PLT almost everywhere.*

If T preserves a smooth measure, then, by [DKH24], T is Bernoulli, in particular, it has positive entropy. (NSR(x^*)) and (LR'(x^*)) are satisfied almost everywhere by Remark 6.

For some applications it will be useful to choose $T(x,y) = G_{\tau(y)}(x)$, where $\int_Y \tau \,\mathrm{d}\nu = 0$. However, in this case, T will not satisfy condition (MEM). Fortunately, we can apply similar techniques if ergodic averages of τ grow faster than logarithmically. More explicitly denote $\tau_n = \sum_{j=0}^{n-1} \tau \circ R^j$, assume there is a $\zeta : \mathbb{N} \to (0,\infty)$ with $\log(n) = o(\zeta(n))$ and a $\kappa > 0$ such that

$$\nu(|\tau_n| < \zeta(n)) \leq O(n^{-\kappa}). \tag{BA}$$

Theorem 9. *Assume that $S(x,y) = (T(x,y), R(y))$, where $T(x,y) = G_{\tau(y)}(x)$, satisfies conditions (MEM) with G instead of T. Suppose that R satisfies (EE), and τ satisfies (BA). Let $x^* \in X$, $y^* \in Y$. If there is a $\delta_2 > 0$ such that for small enough $\rho > 0$ we have*

$$\varphi_{B_\rho(y^*)} \geq \rho^{-\delta_2} \quad \text{on } B_\rho(y^*), \tag{2.9}$$

and

$$d > \frac{r' + d'\delta}{1-\delta} \quad \text{and} \quad \kappa > \frac{d'}{\delta_2}. \tag{2.10}$$

then S satisfies the PLT at (x^*,y^*).

Remark 10. *Let us remark here, that hitting times for skew product, say $S : X \times Y \to X \times Y$ with $S(x,y) = (T(x,y), R(y))$, have previously been investigated by other authors, for example, [HRY19] and [RSV14]. The main differences between [HRY19] and our results are;*

- *In [HRY19] the system is viewed from the standpoint of random dynamics, therefore the relevant target sets are of the form $B_\rho(x^*) \times Y$. In contrast, we focus on geometric balls $B_\rho(x^*,y^*)$.*
- *In [HRY19], R is a full shift. This is needed to prove a "no short return" property akin to (NSR(x^*)), which for us, is one of the assumptions. This allows different choices of R, namely, for us, R need not be hyperbolic or even mixing. This is the main novelty of our approach.*

2.4 Examples

The definitions of the maps in Examples 11, 15, and Lemma 14 are given in §2.11. For most of the examples we present, the choice of R is more interesting than the choice of T, mostly because (MEM) implies chaotic behaviour and so the PLT in that setting is not surprising. We will thus not focus too much on T for this section. We only present some examples here, there are many others one can verify using Theorem 7.

Example 11. *Let T be a map satisfying (MEM) on a manifold of sufficiently high[7] dimension then*

[7] Sufficient bounds are given in §2.11.

(i) if R is a Diophantine rotation, then $T \times R$ satisfies the PLT everywhere;

(ii) if R is the time 1 map of a horocycle flow on $\Gamma\backslash SL_2(\mathbb{R})$ where Γ is a cocompact lattice, then $T \times R$ satisfies the PLT everywhere;

(iii) if R is a skew-shift, then $T \times R$ satisfies the PLT everywhere.

Remark 12. *(i) In §2.11 we will show that the map R from example 11(i)-(iii) satisfies (EE) and (SLR(y^*)) for every y^*. The conclusion then follows from Theorem 7.*

(ii) The PLT **almost everywhere** *can be shown more readily. By Corollary 8, we just have to check (EE) for the map R, which holds for a big class of maps, examples will be given in §2.11.5.*

Example 13. *At this point, let us point out that, while T and R act on manifolds and preserve smooth measure, T and R themselves need not be smooth maps (not even continuous). For example, if $T : (0,1] \to (0,1]$ is the Gauss map (or a more general mixing, expanding interval map as in [Ryc83])*

$$T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor ,$$

then T preserves the density $\rho(x) = \frac{1}{\log 2\cdot(1+x)}$. It is standard to show that T is multiple exponentially mixing, in fact, a fortiori, the Perron-Frobenius transfer operator has a spectral gap. Now if R_α is an irrational rotation on $\mathbb{T}$, and α is of bounded type, then, as shall be demonstrated in §2.11, Theorem 7 applies to show that $T \times R_\alpha$ satisfies the PLT everywhere.

Theorem 9 can be used to construct T, T^{-1} transformations of zero entropy that satisfy the PLT. All that remains to do is to construct a τ satisfying (BA), this can be done with the construction given in [DDKN22a, Proposition 3.9].

Lemma 14. *Let $R_\alpha : \mathbb{T}^{d'} \to \mathbb{T}^{d'}$ be a Diophantine rotation, i.e*

$$|\langle k, \alpha\rangle - l| > C|k|^{-\lambda} \quad \forall k \in \mathbb{Z}^{d'}, k \neq 0, l \in \mathbb{Z}, \tag{D}$$

for some $\lambda \geq d'$. For $\frac{n}{2} < \rho < d'$ there is a $d \geq 1$ and a function $\tau \in C^\rho(\mathbb{T}^{d'}, \mathbb{R}^d)$ such that $\nu(\tau) = 0$, while

$$\nu(||\tau_n|| < \log^2(n)) = o(n^{-5}).$$

Note that in order to apply Theorem 9 we can always make d as big as we want.

Example 15. *Let $R = R_\alpha$ be a Diophantine rotation with $d' = 2$ and $\lambda = 2$, τ be the function from Lemma 14, and let G be the Weyl Chamber flow on $SL(d, \mathbb{R})/\Gamma$, where Γ is a uniform lattice. If $d > 2$ then $S(x, y) = (G_{\tau(y)}(x), R_\alpha(y))$ satisfies the PLT everywhere.*

2.5 The delayed PLT

The main step in the proof will be to show a generalised version of the PLT (for fiberwise returns), along a subsequence, this is what we will call a delayed PLT.

This 'delayed PLT' in itself is of independent interest, so let us make a more general statement.

Definition 16. *Let $(X, \mathcal{A}, \mu, T)$ be a probability-preserving ergodic dynamical system and $\alpha = (\alpha^{(n)})_{n\geq 1}$ be a sequence of natural numbers, we will refer to α as the delay sequence, and denote $\tilde{\alpha}^{(n)} = \sum_{j=1}^{n} \alpha^{(n)}$. For measurable $A \subset X$ we define the delayed consecutive return times to A*

along α as

$$
\begin{aligned}
&\varphi_{A,\alpha}(x) := \varphi^{(1)}_{A,\alpha}(x) := \min(j \geq 1 \mid T^{\tilde{\alpha}^{(j)}}(x) \in A)\\
&\varphi^{(n+1)}_{A,\alpha}(x) := \min(j \geq 1 \mid T^{\tilde{\alpha}^{(\varphi^{(1)}_{A,\alpha}(x)+\varphi^{(2)}_{A,\alpha}(x)+\cdots+\varphi^{(n)}_{A,\alpha}(x)+j)}}(x) \in A)\\
&\Phi_{A,\alpha} := (\varphi^{(1)}_{A,\alpha}, \varphi^{(2)}_{A,\alpha}, \ldots).
\end{aligned}
\tag{2.11}
$$

Like for classical return times, we will consider delayed return times for different systems. In an attempt to keep notation simple, we will not specify the underlying map, only the target set A.

The main example in this paper will be $\alpha^{(n)} = \varphi_B(R^n_B(y))$ for some $y \in Y$ and $B \subset Y$[8], however other choices are of interest, for example $\alpha^{(n)} = g(n)$, where g is a polynomial, or $\alpha^{(n)} = p_n$, where p_n is the nth prime.

Given a rare sequence $(A_l)_{l\geq 1}$, we will distinguish between two cases

1) the delay sequence is fixed in l,
2) the delay sequence is allowed to vary with l.

Definition 17. *(i) Let $x^* \in X$ and, for $r > 0$, denote by $B_r(x^*)$ the geodesic ball of radius r centred at x^*. Let $\alpha = (\alpha^{(n)})_{n\geq 1}$ be a sequence of natural numbers. We will say that T satisfies the delayed PLT along α at x^* if*

$$\mu(B_r(x^*))\Phi_{B_r(x^*),\alpha} \overset{\mu}{\Rightarrow} \Phi_{Exp} \quad \text{as } r \to 0.$$

(ii) Let $PLT(\alpha) = \{x^* \in X \mid T$ satisfies the delayed PLT along α at $x^*\}$.

If $x^* \in PLT(\alpha)$ for all sequences of natural numbers α, then we say that T satisfies the delayed PLT at x^*.

The analogous definition for varying α is

Definition 18. *(i) Let $x^* \in X$ and, for $r > 0$, denote by $B_r(x^*)$ the geodesic ball of radius r centred at x^*. Let $\alpha = ((\alpha^{(n)}_r)_{n\geq 1})_{r>0}$ be a collection of sequences of natural numbers. We will say that T satisfies the varying delayed PLT along α at x^* if*

$$\mu(B_r(x^*))\Phi_{B_r(x^*),\alpha_r} \overset{\mu}{\Rightarrow} \Phi_{Exp} \quad \text{as } r \to 0.$$

(ii) Let $PLT(\alpha) = \{x^* \in X \mid T$ satisfies the varying delayed PLT along α at $x^*\}$. If $x^* \in PLT(\alpha)$ for all sequences of natural numbers α, then we say that T satisfies the delayed PLT at x^*.

In case 1) the main result is a straightforward modification of Theorems 7 and 9.

Theorem 19. *Let α be a sequence of positive integers. Assume the conditions of Theorem 7 or 9 hold replacing (EE) in both cases by*

$$\left\| \sum_{j=0}^{n-1} f \circ R^{\tilde{\alpha}^{(j)}} - n \int f \, d\nu \right\|_{L^2(\nu)} \leq C||f||_{C^{r'}} n^{\delta} \quad \forall f \in C^{r'},\ \forall n \geq 1.$$

Then S satisfies the delayed PLT along α at (x^*, y^*).

[8] As in §2.7, see especially (2.13).

The proof is analogous to the proof of Theorem 7 resp. 9 (with Φ_{B_l} replaced by $\Phi_{B_l,\alpha}$ and Lemma 21 replaced by Lemma 39). Therefore, no detailed proof will be given [9].

In case 2) the main result is a special case of Proposition 25(III), however, it is worth stating by itself.

Theorem 20. *Suppose T satisfies conditions (MEM), (SLR(x^*)) and (NSR(x^*)), then T satisfies the varying delayed PLT at x^*.*

2.6 Cumulative return times

Most of the statements (and proofs) below are much more convenient to state in terms of cumulative return times. For a measurable set $A \subset X$ (or $B \subset Y, C \subset X \times Y$) the *sequence of cumulative return times* to A is given by[10]

$$\sigma_A^{(n)} = \sum_{j=0}^{n-1} \varphi_A \circ T_A^j, \quad n \geq 1$$
$$\Sigma_A = (\sigma_A^{(1)}, \sigma_A^{(2)}, \dots),$$

and similar notation for delayed returns.

When studying distributional convergence of Φ_A, one can equivalently study for distributional convergence of Σ_A.

Indeed, let $\iota : [0,\infty)^{\mathbb{N}} \to [0,\infty)^{\mathbb{N}}$ be the map

$$\iota(x_1, x_2, x_3, ...) = (x_1, x_1 + x_2, x_1 + x_2 + x_3, \dots).$$

Since ι is a homeomorphism, standard theory shows that

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Rightarrow} \Phi \text{ if and only if } \mu(A_l)\Sigma_{A_l} \overset{\mu}{\Rightarrow} \iota(\Phi), \tag{2.12}$$

where we use the obvious extension of ι to $[0,\infty]^{\mathbb{N}}$. Denote $\Sigma_{Exp} = \iota(\Phi_{Exp})$.

2.7 Idea of the proof

Let us first outline the strategy of proving Theorem 7.

Say we want to show the PLT for the geodesic balls (Q_l) converging to (x^*, y^*), then, for every $\epsilon > 0$, we find suitable rectangles[11] approximating Q_l in the sense that $\bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)} \subset Q_l$ and

$$(\mu \times \nu)_{Q_l} \left(Q_l \setminus \bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)} \right) < \epsilon.$$

Denote $Q'_l := \bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)}$. We will use Proposition 25 to show the PLT for Q'_l, i.e

$$(\mu \times \nu)(Q'_l)\Sigma_{Q'_l} \overset{(\mu\times\nu)}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty.$$

[9] Note however that there is no relation of delayed hitting times to delayed return times as in Theorem 4.

[10] Recall from Definition 3 that $\varphi_A(x) = \min(n \geq 1 \mid T^n(x) \in A)$ is the first entry/return time to A, and $T_A = T^{\varphi_A}$ is the first return map.

[11] Using the exponential map, this is a simple exercise in $\mathbb{R}^{d+d'}$. Here we use continuity of the densities of μ resp. ν!

Letting $\epsilon \to 0$, the next Lemma will help us conclude that

$$(\mu \times \nu)(Q_l)\Sigma_{Q_l} \overset{(\mu\times\nu)}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty.$$

The critical approximation here is the following Lemma - see [Zwe16, Theorem 4.4] for a qualitative version, and the following statement can be easily deduced from the proof - for convenience, we give a proof of this quantitative version in §2.12.

Lemma 21. *There is a metric D defined on the space of probability measures on $[0,\infty]^{\mathbb{N}}$ and modelling weak convergence of measures*[12] *such that*

$$D\left(law_{(\mu\times\nu)_Q}((\mu \times \nu)(Q)\Phi_Q), law_{(\mu\times\nu)_{Q'}}((\mu \times \nu)(Q')\Phi_{Q'}\right) \leq 4(\mu \times \nu)_Q(Q \setminus Q'),$$

for measurable $Q' \subset Q \subset X \times Y$.

For returns to rectangles, say $A_l \times B_l$, for fixed $y \in Y$, we can ignore all the times j where $R^j(y) \notin B_l$. Hence, denoting $\alpha_l(y) = \Phi_{B_l}(y)$, we first show that for ν-a.e $y \in Y$

$$\mu(A_l)\Sigma_{A_l,\alpha_l(y)} \overset{\mu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty. \tag{2.13}$$

To show this, the idea is that, due to assumption (BR), the times[13] $\tilde{\alpha}_l^{(j)}(y)$ are sufficiently far apart to use (MEM), we apply Proposition 25. Since (2.13) is now true for ν-a.e $y \in Y$, we also have

$$\mu(A_l)\Sigma_{A_l,\Phi_{B_l}} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty, \tag{2.14}$$

where $\Sigma_{A_l,\Phi_{B_l}}(x,y) = \Sigma_{A_l,\Phi_{B_l}(y),y}(x)$.

These are not quite the returns that we wanted to consider. However, notice that, in every step, we skip exactly φ_{B_l} steps, thus we have the following relation

$$\sigma_{A_l\times B_l}(x,y) = \sum_{j=0}^{\sigma_{A_l,\Phi_{B_l}(y),y}(x)-1} \varphi_{B_l} \circ R^j_{B_l}(y). \tag{2.15}$$

We now use (EE) to control the ergodic sums of φ_{B_l} and show that

$$\mu(A_l)\nu(B_l)\Sigma_{A_l\times B_l} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty.$$

The same idea works for unions of rectangles $\bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)}$, however, in (2.13), we will have to slightly modify the definition of return times, which shall be done at the beginning of §2.8.2.

2.8 PLT along varying subsequences

2.8.1 Approximation

In our work, we often need to apply the mixing condition (MEM), and the quantitative ergodicity (EE), for indicator functions, hence we have to approximate them by functions in C^r resp. $C^{r'}$.

Definition 22. *Let M be a C^r manifold with dimension $dim(M) = d$, λ be a measure on M. Let $B_\alpha \subset M$ be measurable subsets for α in some index set. We say that $\{B_\alpha\}$ is regularly approximable in C^r if there is a constant $\mathcal{B} > 0$ such that, for each α and for $0 < \epsilon < \frac{1}{10}\lambda(B_\alpha)^{\frac{1}{d}}$, there are $\overline{h}, \underline{h} \in C^r$ with $\underline{h} \leq 1_{B_\alpha} \leq \overline{h}$ and for $h \in \{\overline{h}, \underline{h}\}$*

$$\lambda(1_{B_\alpha} \neq h) \leq \lambda(B_\alpha)^{\frac{d-1}{d}}\epsilon, \quad \textit{while} \ \ ||h||_{C^r} \leq \mathcal{B}\epsilon^{-r}. \tag{2.16}$$

We call the least such constant $\mathcal{B} > 0$ the approximant of B and denote it by $app(\{B_\alpha\})$.

[12] In the sense that $\lambda_l \Rightarrow \lambda$ if and only if $D(\lambda_l, \lambda) \to 0$.

[13] It can be recalled from Definition 16 that $\tilde{\alpha}_l^{(n)} = \sum_{j=1}^n \alpha_l^{(j)} = \sum_{j=0}^{n-1} \varphi_{B_l} \circ R^j$.

Lemma 23. *Let M, r, d and λ be as in Definition 22. Assume in addition that λ is absolutely continuous w.r.t volume with bounded density. Suppose that $\bigcup_\alpha B_\alpha$ is relatively compact, and there is an open $U \supset \overline{\bigcup_\alpha B_\alpha} \supset \bigcup_\alpha B_\alpha$ and a C^r-diffeomorphism $\iota : U \to V$ for some open set $V \subset \mathbb{R}^d$ such that each $\iota(B_\alpha)$ is a ball. Then $\{B_\alpha\}$ is regularly approximable in C^r.*

Proof. (i) We may assume that $M = \mathbb{R}^d$ and λ is Lebesgue measure, otherwise we pick up another constant, which can be absorbed into C. Furthermore, we can assume that all B_α are balls centered at the origin. In the following fix α and, dropping the α from our notation, let $B = B_t(0)$. We have $\lambda(B) = C_1 t^d$, where C_1 is the volume of the d-dimensional unit ball.

(ii) Let $\theta : \mathbb{R} \to [0,1]$ be a smooth function with $\theta(x) = 1$ if $x < 0$, and $\theta(x) = 0$ if $x > 1$. For $t > \epsilon > 0$ consider

$$\hat{\theta}(x) = \theta(\epsilon^{-1}(x - t))$$

then $\hat{\theta}$ is still smooth and $||\hat{\theta}||_{C^r} = \epsilon^{-r}||\theta||_{C^r}$. Consider $\overline{h} : \mathbb{R}^d \to [0,1]$ given by $\overline{h}(x) = \hat{\theta}(|x|)$, then

- $\overline{h}$ is smooth, away from the origin because it is the composition of smooth functions, and near the origin, it is constant 1,
- $\overline{h}(x) = 1$ if $|x| < t$ and $\overline{h}(x) = 0$ if $|x| > t + \epsilon$,
- and $||\overline{h}||_{C^r} \leq C_3\epsilon^{-r}$ where $C_3 = rC_2||\theta||_{C^r}$ and C_2 is the C^r norm of the smooth function $x \mapsto |x|$ on $\{t \leq |x| \leq t + \epsilon\}$.

Furthermore, we have

$$\begin{aligned}\lambda(\overline{h} \neq 1_B) &= \lambda(t \leq |x| \leq t + \epsilon) = C_1((t+\epsilon)^d - t^d) \\ &\leq C_1 d t^{d-1}\epsilon \leq dC_1^{\frac{1}{d}}\lambda(B)^{\frac{d-1}{d}}\epsilon,\end{aligned}$$

all the constants can be absorbed in the constant C from the claim, the constant only depends on r, d.

For $\underline{h}$ repeat the calculations with $\hat{\theta} = \theta(\epsilon^{-1}(x - (t - \epsilon)))$ instead. □

2.8.2 Proof of the PLT along varying subsequences

For our purposes it will not be enough to consider the delayed PLT for a single rare sequence $(A_l)_{l \geq 1}$, rather let $\mathcal{K} \geq 1$, and $A_l^{(1)}, ..., A_l^{(\mathcal{K})}$ be subsets of X such that $\{A_l^{(k)}\}$ is regularly approximable in C^r. Assume that there are $\omega^{(1)}, ..., \omega^{(\mathcal{K})} > 0$ and $r_l \to 0$ such that

$$\mu(A_l^{(k)}) = \omega^{(k)} r_l^d + o(r_l^d). \tag{2.17}$$

Given $\kappa_l^{(j)} \in \{1, ..., \mathcal{K}\}$, for $l, j \geq 1$, define the *cumulative return times* by

$$\begin{aligned}&\sigma^{(1)}_{\kappa_l,\alpha_l,y}(x) = \min(j \geq 1 \mid T_y^{\tilde{\alpha}_l^{(j)}}(x) \in A_l^{(\kappa_l^{(j)})}), \\ &\sigma^{(n+1)}_{\kappa_l,\alpha_l,y}(x) = \min(j \geq \tau^{(n)}_{\kappa_l,\alpha_l,y}(x) + 1 \mid T_y^{\tilde{\alpha}_l^{(j)}}(x) \in A_l^{(\kappa_l^{(j)})}), \\ &\Sigma_{\kappa_l,\alpha_l,y} = (\sigma^{(1)}_{\kappa_l,\alpha_l,y}, \sigma^{(2)}_{\kappa_l,\alpha_l,y}, ...).\end{aligned}$$

Denote the frequency with which $A_l^{(\kappa_l)} = A_l^{(k)}$ by

$$p_{l,t}^{(k)} := \frac{1}{t}\#\{j = 1, ..., t_l \mid \kappa_l^{(j)} = k\}.$$

For now suppose that there are positive constants $\theta^{(k)} > 0$ such that, for all $k = 1, ..., \mathcal{K}$, and for all $t_l \nearrow \infty$ with $t_l = O(r_l^{-d})$,

$$p_{l,t_l}^{(k)} \to \frac{\theta^{(k)}}{\sum_{j=1}^{\mathcal{K}} \theta^{(j)}} =: p^{(k)} \quad \text{as } l \to \infty. \tag{2.18}$$

Later on, when we prove PLT in product systems, we will choose κ and θ in a specific way[14] and (2.18) will be satisfied by Lemma 30.

The main estimate of mixing rates for regularly approximable sets is the following.

Lemma 24. *Suppose T satisfies (MEM), and let $y \in Y$ be such that (2.1) is satisfied. Let $m \geq 1$, $A^{(1)}, ..., A^{(k)} \subset X$ be regularly approximable in C^r, and $1 \leq n_1 < ... < n_k$, then*

$$\left| \mu\left(\bigcap_{i=1}^{k} T_y^{-n_i} A^{(i)} \right) - \prod_{i=1}^{k} \mu(A^{(i)}) \right| \leq K \max_{i=1,\dots,k} \mu(A^{(i)})^{\frac{d-1}{d} \frac{kr}{kr+1}} e^{-\frac{\gamma p}{kr+1}},$$

where $p = \min_{i=1,\dots,k-1} |n_{i+1} - n_i|$ *and the constant $K > 0$ only depends on k and app$(\{A^{(1)}, ..., A^{(k)}\})$.*

Proof. Let $C = \text{app}(\{A^{(1)}, ..., A^{(k)}\})$. By Lemma 23, for every $\epsilon > 0$, there are $h^{(i)} \in C^r$ such that $0 \leq h^{(i)} \leq 1_{A^{(i)}}$ and

$$\mu(1_{A^{(i)}} \neq h^{(i)}) \leq \mu(A^{(i)})^{\frac{d-1}{d}} \epsilon, \quad \text{while } ||h^{(i)}||_{C^r} < C\epsilon^{-r}.$$

We estimate

$$\begin{aligned}
\left| \mu\left(\bigcap_{i=1}^{k} T^{-n_i} A^{(i)} \right) - \prod_{i=1}^{k} \mu(A^{(i)}) \right| &\leq \left| \mu\left(\bigcap_{i=1}^{k} T^{-n_i} A^{(i)} \right) - \int_X \prod_{i=1}^{k} h^{(i)} \circ T^{n_i} \, \mathrm{d}\mu \right| \\
&\quad + \left| \int_X \prod_{i=1}^{k} h^{(i)} \circ T^{n_i} \, \mathrm{d}\mu - \prod_{i=1}^{k} \int_X h^{(i)} \circ T^{n_i} \, \mathrm{d}\mu \right| \\
&\quad + \left| \prod_{i=1}^{k} \int_X h^{(i)} \circ T^{n_i} \, \mathrm{d}\mu - \prod_{i=1}^{k} \mu(A^{(i)}) \right| \\
&\leq 4k \max_{i=1,\dots,k} \mu(A^{(i)})^{\frac{d-1}{d}} \epsilon + C^k e^{-\gamma p} \epsilon^{-kr}.
\end{aligned}$$

This bound is optimised for

$$\epsilon^* = \left(\frac{C^k r}{4k} \max_{i=1,\dots,k} \mu(A^{(i)})^{-\frac{d-1}{d}} e^{-\gamma p} \right)^{\frac{1}{kr+1}},$$

so

$$\begin{aligned}
\left| \mu\left(\bigcap_{i=1}^{k} T^{-n_i} A^{(i)} \right) - \prod_{i=1}^{k} \mu(A^{(i)}) \right| &\leq \hat{C} \max_{i=1,\dots,k} \mu(A^{(i)})^{\frac{d-1}{d}\left(1 - \frac{1}{kr+1}\right)} e^{-p \frac{,\gamma}{kr+1}} \\
&\quad + \bar{C} \max_{i=1,\dots,k} \mu(A^{(i)})^{\frac{kr}{kr+1} \frac{d-1}{d}} e^{-\gamma p \left(1 - \frac{kr}{kr+1}\right)} \\
&\leq K \max_{i=1,\dots,k} \mu(A^{(i)})^{\frac{d-1}{d} \frac{kr}{kr+1}} e^{-\frac{\gamma}{kr+1} p},
\end{aligned}$$

where the constants $\hat{C}, \bar{C}, K > 0$ only depend on k and on C. $\square$

[14]Say we want to prove a PLT for the system $T \times R$ and sets of the form $B_r(x^*) \times B_r(y^*)$ then we choose $\mathcal{K} = 1$ and $\theta > 0$ such that $\nu(B_r(y^*)) = \theta r^{d'} + o(r^{d'})$.

In the following Proposition, we shall demonstrate exponential limit distributions for fiberwise return times (compare (2.13)). The proof is standard and uses multiple exponential mixing and the method of moments, however, due to having multiple sets, the notations are quite cumbersome. We will present a proof for the convenience of the reader.

Proposition 25. *Suppose that T satisfies (MEM), $(A_l^{(1)}), ..., (A_l^{(\mathcal{K})})$ are sequences of rare events with $\mu(A_l^{(k)}) = \omega^{(k)} r_l^d + o(r_l^d)$, for $\mathcal{K} \geq 1$, κ_l satisfy (2.18) for some $p^{(k)} > 0$ with $\sum_{k=1}^{\mathcal{K}} p^{(k)} = 1$, and let $\alpha_l = (\alpha_l^{(n)})_{n\geq 1}$ be sequences of natural numbers. Denote $A_l = \bigcup_{k=1}^{\mathcal{K}} A_l^{(k)}$ and suppose that either*

(I) α_l grows faster than $|\log \mu(A_l)|$ in the sense that

$$|\log \mu(A_l)| = o(\min_{n\geq 2} |\alpha_l^{(n)}|), \tag{2.19}$$

(II) short returns to A_l are rare in the sense that

$$\mu_{A_l}(\varphi_{A_l} \leq a_l) \to 0 \quad \text{as } l \to \infty, \tag{2.20}$$

for some sequence $(a_l)_{l\geq 1}$ with $|\log(\mu(A_l))| = o(a_l)$, and α_l grows at least as fast $|\log \mu(A_l)|$ in the sense that

$$|\log \mu(A_l)| = O(\min_{n\geq 2} |\alpha_l^{(n)}|), \tag{2.21}$$

(III) or short returns to A_l are rare in the sense that

$$\mu_{A_l}(\varphi_{A_l} \leq a_l) \to 0 \quad \text{as } l \to \infty, \tag{2.22}$$

for some sequence $(a_l)_{l\geq 1}$ with $|\log(\mu(A_l))| = o(a_l)$, and returns are at least logarithmically large, i. e there exists $c > 0$ such that

$$\varphi_{A_l}(x) \geq c \log(\mu(A_l)) \quad \mu - a.e\ x \in A_l. \tag{2.23}$$

Then for ν-a.e $y \in Y$

$$\Omega r_l^d \Sigma_{\kappa_l, \alpha_l, y} \overset{\mu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty,$$

where $\Omega = \sum_{k=1}^{\mathcal{K}} \omega^{(k)} p^{(k)}$.

Proof. Fix y as in (MEM) and denote $T^n = T_y^n$.

(i) Taking a subsequence if necessary, we may assume that there is a $[0, \infty]$-valued process $\Sigma = (\sigma^{(1)}, \sigma^{(2)}, ...)$ such that

$$\Omega r_l^d \Sigma_{\kappa_l, \alpha_l} \overset{\mu}{\Rightarrow} \Sigma \quad \text{as } l \to \infty.$$

For $J \geq 1$ and $0 < t_1 < ... < t_J$ we show

$$\mathbb{P}\left(\sigma^{(j)} \leq t_j,\ j = 1, \ldots, J\right) = \mathbb{P}\left(\sigma_{Exp}^{(j)} \leq t_j,\ j = 1, \ldots, J\right). \tag{2.24}$$

The trick is to look at the "dual" object

$$S_{\kappa_l, \alpha_l}^{(n)} = \sum_{j=1}^{n} 1_{A_l^{\kappa_l^{(j)}}} \circ T^{\tilde{\alpha}_l^{(j)}}.$$

The important relation here is the following

$$S_{\kappa, \alpha_l}^{(n)} \geq N \iff \sigma_{\kappa, \alpha_l}^{(N)} \leq n. \tag{2.25}$$

Indeed, evaluating both sides at some $x \in X$, the left side says that there are at least N different times $1 \leq j_1 < ... < j_N \leq n$ such that $T^{\tilde{\alpha}_l^{(j)}}(x) \in A_l^{\kappa_l^{(j)}}$. The right-hand side expresses that, if $1 \leq j_1 < ... < j_N$ are the first N times that $T^{\tilde{\alpha}_l^{(j)}}(x) \in A_l^{\kappa_l^{(j)}}$, then $j_1 + \sum_{i=2}^{N}(j_i - j_{i-1}) \leq n$. Let $(P_t)_{t\geq 0}$ be a Poisson process, such that Σ_{Exp} are the cumulative waiting times of (P_t), i.e. (P_t) and Σ_{Exp} are related by

$$P_t \geq N \iff \sigma_{Exp}^{(N)} \leq t.$$

The right side of (2.24) is equal to

$$\mathbb{P}\left(\sigma_{Exp}^{(k)} \leq t_k,\ k = 1, ..., J\right) = \mathbb{P}(P_{t_k} \geq k,\ k = 1, ..., J).$$

Due to (2.25) it is enough to show

$$\left(S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_1}{\Omega r_l^d}\right\rfloor}, S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_2}{\Omega r_l^d}\right\rfloor}, ..., S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_J}{\Omega r_l^d}\right\rfloor}\right) \overset{\mu}{\Rightarrow} (P_{t_1}, ..., P_{t_J}) \quad \text{as } l \to \infty.$$

(ii) Taking a further subsequence if necessary, there are $[0,\infty]$-valued $\tilde{P}_{t_1}, ..., \tilde{P}_{t_J}$ such that

$$\left(S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_1}{\Omega r_l^d}\right\rfloor}, S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_2}{\Omega r_l^d}\right\rfloor}, ..., S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_J}{\Omega r_l^d}\right\rfloor}\right) \overset{\mu}{\Rightarrow} (\tilde{P}_{t_1}, ..., \tilde{P}_{t_J}) \quad \text{as } l \to \infty.$$

We will show that

(A) $\tilde{P}_{t_k} - \tilde{P}_{t_{k-1}}$ is Poisson distributed with intensity $t_k - t_{k-1}$ for $k = 1, ..., d$,

(B) and $(\tilde{P}_{t_1} - \tilde{P}_{t_0}, \tilde{P}_{t_2} - \tilde{P}_{t_1}, ..., \tilde{P}_{t_J} - \tilde{P}_{t_{J-1}})$ is an independent vector, where $t_0 = 0$.[15]

Clearly $P_0 = \tilde{P}_0 = 0$.

For $j = 1, \ldots, J$ denote $S_{t_j,l} = S_{A_l,\kappa_l}^{\left\lfloor \frac{t_j}{\Omega r_l^d}\right\rfloor}$. Assertions (A) and (B) will follow[16] once we show that, for all $m_1, ..., m_d \geq 1$,

$$\int_X \prod_{j=1}^{J} \binom{S_{t_j,l} - S_{t_{j-1},l}}{m_j} \mathrm{d}\mu = \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!} + o(1) \quad \text{as } l \to \infty.$$

In the rest of the proof fix $J \geq 1$, $0 = t_0 < t_1 < \cdots < t_J$ and $m_1, \ldots, m_J \geq 1$.

(iii) First, for each $j = 1, \ldots, J$, rewrite

$$S_{t_j,l} - S_{t_{j-1},l} = \sum_{i=\left\lfloor \frac{t_{j-1}}{\Omega r_l^d}\right\rfloor + 1}^{\left\lfloor \frac{t_j}{\Omega r_l^d}\right\rfloor} 1_{A_l^{(\kappa_l^{(i)})}} \circ T^{\tilde{\alpha}_l^{(i)}}.$$

So

$$\prod_{j=1}^{J} \binom{S_{t_j,l} - S_{t_{j-1},l}}{m_j} = \prod_{j=1}^{d} \sum_{\left\lfloor \frac{t_{j-1}}{\Omega r_l^d}\right\rfloor + 1 \leq k_{1,j} < ... < k_{m_j,j} \leq \left\lfloor \frac{t_j}{\Omega r_l^d}\right\rfloor} \prod_{i=1}^{m_j} \xi_{i,j} \tag{2.26}$$

[15] This is essentially Watanabe's characterisation of Poisson-processes.
[16] Here we apply the method of moments, see eg [Bil95, Theorem 30.2].

where $\xi_{i,j} = 1_{A_l^{(\kappa_l^{(k_{i,j})})}} \circ T^{\tilde{\alpha}_l^{(k_{i,j})}}$. To simplify notation we will also denote $m = (m_1 + \cdots + m_J)$, $\omega = \min_{k=1,\dots,\mathcal{K}} \omega^{(k)}$, $p_l = 2m\frac{mr+1}{\gamma}\left|\log(\omega r_l^d)\right|$,

$$\Delta_l := \{\mathbf{k} = (k_{i,j})_{\substack{j=1,\dots,J\\ i=1,\dots,m_j}} \mid \left\lfloor \frac{t_{j-1}}{\Omega r_l^d} \right\rfloor + 1 \leq k_{1,j} < \ldots < k_{m_j,j} \leq \left\lfloor \frac{t_j}{\Omega r_l^d} \right\rfloor \quad \text{for } j = 1, \dots, J\}$$

and

$$\Delta_l' := \{\mathbf{k} \in \Delta_l \mid \min_{\substack{j=1,\dots,J,\ i=1,\dots,m_j\\ j'=1,\dots,J,\ i'=1,\dots,m_{j'},\ (j,i)\neq(j',i')}} |\tilde{\alpha}_l^{(k_{i,j})} - \tilde{\alpha}_l^{(k_{i',j'})}| \leq p_l\}.$$

We will split the sum in (2.26) into two terms

$$\prod_{j=1}^{J} \binom{S_{t_j,l} - S_{t_{j-1},l}}{m_j} = M_l + R_l,$$

where

$$M_l = M_{(t_j),l,(m_j)} = \sum_{\mathbf{k}\in\Delta_l\setminus\Delta_l'} \prod_{\substack{j=1,\dots,J\\ i=1,\dots,m_j}} \xi_{i,j}, \quad R_l = R_{(t_j),l,(m_j)} = \sum_{\mathbf{k}\in\Delta_l'} \prod_{\substack{j=1,\dots,J\\ i=1,\dots,m_j}} \xi_{i,j}.$$

We will show that

$$\int_X M_l \,\mathrm{d}\mu \to \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!} \quad \text{and} \quad \int_X R_l \,\mathrm{d}\mu \to 0 \quad \text{as } l \to \infty. \tag{2.27}$$

(iv) Let us first treat M_l. For $l \geq 1$ and $\mathbf{k} \in \Delta_l$, by Lemma 24 we have

$$\begin{aligned}
&\left|\mu\left(\bigcap_{\substack{j=1,\dots,J\\ i=1,\dots,m_j}} T^{-\tilde{\alpha}_l^{(k_{i,j})}} A_l^{(\kappa_l^{(k_{i,j})})}\right) - \prod_{j,i} \mu\left(A_l^{(\kappa_l^{(k_{i,j})})}\right)\right| \\
&\leq K \max_{j,i} \mu\left(A_l^{(\kappa_l^{(k_{i,j})})}\right)^{\frac{d-1}{d}\frac{mr}{mr+1}} e^{-\frac{\gamma \min_{i,j} \alpha_l^{(k_{i,j})}}{mr+1}}.
\end{aligned}$$

For $\mathbf{k} \in \Delta_l \setminus \Delta_l'$, this yields

$$\mu\left(\bigcap_{\substack{j=1,\dots,J\\ i=1,\dots,m_j}} T^{-\tilde{\alpha}_l^{(k_{i,j})}} A_l^{(\kappa_l^{(k_{i,j})})}\right) = r_l^{md} \prod_{\substack{j=1,\dots,J\\ i=1,\dots,m_j}} \omega_l^{(\kappa_l^{k_{i,j}})} + o(r_l^{md}),$$

and the o-term does not depend on $\mathbf{k}$. Summing over $\mathbf{k} \in \Delta_l \setminus \Delta_l'$, and using (2.18), yields

$$\begin{aligned}
\int_X M_l \,\mathrm{d}\mu &= \int_X \sum_{\mathbf{k}\in\Delta_l\setminus\Delta_l'} \prod_{\substack{j=1,\dots,J\\ i=1,\dots,m_j}} 1_{A_l^{(\kappa_l^{(k_{i,j})})}} \circ T^{\tilde{\alpha}_l^{(k_{i,j})}} \,\mathrm{d}\mu \\
&= r_l^{md} \sum_{\mathbf{k}\in\Delta_l\setminus\Delta_l'} \prod_{\substack{j=1,\dots,J\\ i=1,\dots,m_j}} \omega^{(\kappa_l^{(k_{i,j})})} + o(1) \\
&\overset{*}{=} r_l^{md} \sum_{\mathbf{k}\in\Delta_l} \prod_{\substack{j=1,\dots,J\\ i=1,\dots,m_j}} \omega^{(\kappa_l^{(k_{i,j})})} + o(1) \\
&= r_l^{md} \prod_{j=1}^{J} \frac{1}{m_j!} \left(\sum_{k=1}^{\mathcal{K}} \left\lfloor \frac{t_j - t_{j-1}}{\Omega r_l^d} \right\rfloor p^{(k)}_{l,\left\lfloor \frac{t_j - t_{j-1}}{\Omega r_l^d} \right\rfloor} \omega^{(k)} \right)^{m_j} + o(1) \\
&= \Omega^{-m} \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!} \left(\sum_{k=1}^{\mathcal{K}} p^{(k)} \omega^{(k)} \right)^{m_j} + o(1) \\
&= \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!} + o(1),
\end{aligned}$$

for $*$ note that $\#\Delta_l' = O(r_l^{-md+1}|\log(r_l)|)$. This shows the first assertion of (2.27).

(v) In order to treat R_l, first note that under assumption (I) we have $R_l = 0$ for big enough l. In the following, we focus on assumptions (II) and (III). Note that

$$\int_X R_l \,\mathrm{d}\mu = \int_X 1_{R_l \neq 0} R_l \,\mathrm{d}\mu \leq \mu(\mathrm{supp}(R_l)) \|R_l\|_{L^2}$$

and

$$\mathrm{supp}(R_l) \subset \bigcup_{j=1}^{\left\lfloor \frac{t_J}{\Omega r_l^d} \right\rfloor} T^{-\tilde{\alpha}_l^{(j)}} \left(A_l \cap \{\varphi_{A_l} \leq 3p_l\} \right) =: U_l,$$

since $p_l = O(|\log(r_l)|)$, from (2.20) resp. (2.22), it follows that

$$\mu(\mathrm{supp}(R_l)) \leq \mu(U_l) = O(r_l^{-d}) \mu(A_l) o(1) = o(1).$$

Therefore, in order to show $\int_X R_l \,\mathrm{d}\mu \to 0$, it is enough to show that $(R_l)_{l\geq 1}$ is bounded in L^2. Notice that

$$R_l^2 \leq \sum_{1\leq k_1,\dots,k_{2m} \leq \left\lfloor \frac{t_J}{\Omega r_l^d} \right\rfloor} \prod_{i=1}^{m} 1_{A_l^{(\kappa_l^{(k_i)})}} \circ T^{\tilde{\alpha}_l^{(i)}} = S_{t_J,l}^{2m}.$$

We may write

$$S_{t_J,l}^{2m} = \sum_{k=1}^{2m} \left\{ {2m \atop k} \right\} \binom{S_{t_J,l}}{k} \leq C_m \sum_{k=1}^{2m} \binom{S_{t_J,l}}{k} \leq C_m \sum_{k=1}^{2m} (M_{t_J,l,k} + R_{t_J,l,k}),$$

where $\left\{ {2m \atop k} \right\}$ are the Stirling numbers of the second kind and $C_m = \max_{k=1,\dots,2m} \left\{ {2m \atop k} \right\}$. Now the previous parts of the proof show that

$$\int_X M_{t_J,l,m} \,\mathrm{d}\mu \quad \text{is bounded as } l \to \infty, \forall m \geq 1,$$

it remains to show that

$$\int_X R_{t_J,l,m}\,\mathrm{d}\mu \quad \text{is bounded as } l\to\infty, \forall m\geq 1.$$

(vi) In order to bound $\int_X R_{t_J,l,m}\,\mathrm{d}\mu$ for fixed $m\geq 1$ we first split up[17] Δ_l' into

$$\Delta_l^{(j)} := \left\{ \mathbf{k}=(k_1,...,k_m)\in\Delta_l \;\middle|\; \begin{array}{l} \exists 1\leq i_i<...<i_{m-j}\leq m \text{ such that} \\ |\tilde{\alpha}_l^{(k_{i+1})}-\tilde{\alpha}_l^{(k_i)}|\leq \tilde{p}_l \;\forall i\in\{1,...,m\}\setminus\{i_1,...,i_{m-j}\}, \\ \text{and } |\tilde{\alpha}_l^{(k_{i_r+1})}-\tilde{\alpha}_l^{(k_{i_r})}|>\tilde{p}_l \;\forall r=1,...,m-j \end{array} \right\},$$

for $j=1,...,m-1$, so that $\Delta_l'=\bigcup_{j=1}^{m-1}\Delta_l^{(j)}$. Under assumption (II), because of (2.21), there is a constant $c>0$ such that for $\rho=2m\frac{mr+1}{c\gamma}$ and big enough l we have

$$\#\Delta_l^{(j)}\leq \binom{m}{j}\binom{\left\lfloor \frac{t}{\Omega r_l^d}\right\rfloor}{m-j}\rho^j\leq m^m t^m\rho^m(\Omega^{-m}+1)r_l^{-d(m-j)}.$$

On the other hand, for $\mathbf{k}\in\Delta_l^{(j)}$, we can again use Lemma 24 to estimate

$$\int_X\prod_{s=1}^m\xi_s\,\mathrm{d}\mu\leq\int_X\prod_{s\in\{1,...,m\}\setminus\{i_1,...,i_j\}}\xi_s\leq\Omega r_l^{d(m-j)}+o(r_l^{d(m-j)}),$$

where the o-term does not depend on $\mathbf{k}$, and

$$\xi_s=1_{A_l^{\kappa_l^{(k_s)}}}\circ T^{\tilde{\alpha}_l^{(k_s)}}.$$

Summing over $\mathbf{k}\in\Delta_l'$ we obtain

$$\int_X R_{t_j,l,m}\,\mathrm{d}\mu\leq 2m^{m+1}t^m\rho^m(\Omega^{-m+1}+\Omega),\quad\forall m\geq 1,$$

and we conclude $\int_X R_l\,\mathrm{d}\mu\to 0$ under assumption (II).

(vii) Finally assume (III). For fixed $j=1,...,m-1$, $l\geq 1$ and $1\leq i_1<...<i_j\leq\left\lfloor\frac{t}{\Omega r_l^d}\right\rfloor$ with $\min_{s=1,...,j-1}|\tilde{\alpha}_l^{(i_{s+1})}-\tilde{\alpha}_l^{(i_s)}|\geq p_l$ set

$$A_{l,i_1,...,i_j}=\bigcap_{s=1}^j T^{-\tilde{\alpha}_l^{(i_s)}}A_l^{\kappa_l^{(i_s)}}.$$

By Lemma 24

$$\mu(A_{l,i_1,...,i_j})\leq(\omega r_l)^{dj}+o(r_l^{dj}),$$

[17] Here the sets of sequences should be modified, i.e

$$\Delta_l:=\{\mathbf{k}=(k_i)_{i=1,...,m}\;\Big|\;1\leq k_1<\ldots<k_m\leq\left\lfloor\frac{t_J}{\Omega r_l^d}\right\rfloor\}$$

and

$$\Delta_l':=\{\mathbf{k}\in\Delta_l\mid\min_{\substack{i=1,...,m\\ i'=1,...,m,\,i\neq i'}}|\tilde{\alpha}_l^{(k_i)}-\tilde{\alpha}_l^{(k_{i'})}|\leq p_l\}$$

and the o term doesn't depend on $(i_1, ..., i_j)$. For $x \in A_{l,i_1,...,i_j}$ consider

$$\mathfrak{K}_{l,i_1,...,i_j}(x) = \left\{ \mathbf{k} = (k_1, ..., k_m) \in \Delta_l^{(j)} \,\middle|\, \begin{array}{l} \exists r_1, ..., r_j \text{ such that } k_{r_s} = i_s \ \forall s = 1, ..., j, \\ \text{and } \prod_{s=1}^{m} \xi_s(x) = 1 \end{array} \right\}.$$

Then, since $\varphi_{A_l}(x) \geq c\log(\mu(A_l))$, we have

$$\#\mathfrak{K}_{l,i_1,...,i_j}(x) \leq \left(2\frac{p_l}{dc\log(r_l)}\right)^{m-j} =: \rho^{m-j},$$

since $p_l = constant * |\log(r_l)|$ this quantity doesn't depend on l. At the same time, we have

$$\operatorname{supp}\left(\prod_{s=1}^{m} \xi_s\right) \subset A_{l,i_1,...,i_j} \quad \text{if } \mathbf{k} \in \mathfrak{K}_{l,i_1,...,i_j}(x) \text{ for some } x.$$

Also, every $\mathbf{k} \in \tilde{\Delta}_l^{(j)}$ is in some $\mathfrak{K}_{l,i_1,...,i_j}(x)$, therefore

$$\begin{aligned}
\sum_{\mathbf{k}\in\tilde{\Delta}_l^{(j)}} \int_X \prod_{s=1}^{m} \xi_s \,\mathrm{d}\mu &\leq \sum_{\substack{1\leq i_1<...<i_j\leq\left\lfloor \frac{t}{\Omega r_l^d}\right\rfloor \\ \min_{s=1,...,j-1} |\tilde{\alpha}_l^{(i_s+1)} - \tilde{\alpha}_l^{(i_s)}|\geq p_l}} \int_{A_{l,i_1,...,i_j}} \#\mathfrak{K}_{l,i_1,...,i_j}(x)\,\mathrm{d}\mu(x) \\
&\leq \sum_{\substack{1\leq i_1<...<i_j\leq\left\lfloor \frac{t}{\Omega r_l^d}\right\rfloor \\ \min_{s=1,...,j-1} |\tilde{\alpha}_l^{(i_s+1)} - \tilde{\alpha}_l^{(i_s)}|\geq \tilde{p}_l}} \int_{A_{l,i_1,...,i_j}} \rho^j \,\mathrm{d}\mu \\
&\leq \rho^j \left(\omega r_l^{dj} + o(r_l^{dj})\right)\left(\left\lfloor \frac{t}{\Omega r_l^d}\right\rfloor\right)^j \\
&\leq \rho^j t^j \frac{\omega}{\Omega^j} + o(1).
\end{aligned}$$

Summing up over j we get

$$\int_X R_l \,\mathrm{d}\mu \leq m \max(1, \rho^m) \max(1, t^m) \frac{\omega}{\max(1, \Omega^m)} + o(1).$$

Following the argument in step (v) this shows $\int_X R_l \,\mathrm{d}\mu \to 0$, hence (2.27), in case of assumption (III). This concludes the proof. $\square$

Remark 26. *(i) Note that assumptions (I), (II), or (III) directly correspond to the three possible cases of (BR(x^*, y^*))–(SLR(y^*)), (NSR(x^*)) AND (LR(y^*)), or (NSR(x^*)) AND (LR'(x^*)) respectively.*

(ii) In all of the examples we give in section 2.4, the α_l will satisfy a condition stronger than (I). In fact, in this set-up, there is a $\delta_2 > 0$ such that

$$\min_{i\geq 2} |\alpha_l^{(i)}| \geq \mu(A_l)^{-\delta_2},$$

compare also (2.9)[18]*. If this stronger condition is satisfied instead of (I), then we do not need the full strength of exponential mixing in (MEM). Any superpolynomial rate will be enough. Details are given in the proof of Theorem 9, but we shall give a heuristic here.*

[18]The constant δ_2 given there is not exactly the same. In the notation there, we have to use $\delta_2 \frac{d'}{d}$.

When using mixing of all orders with indicators of the form $1_{B_{r_l}(x^*)}$, the error term will contain a term coming from the C^r norm in the definition of regularly approximable. In this case, using (2.16), this term will be of order Cr_l^{-drm}. To compensate, say the rate of mixing is ψ, since the gaps α_l are large we can multiply with $\psi(\min_{i\geq 2}|\alpha_l^{(i)}|)$. So we want to show

$$r_l^{-drm}\psi(r_l^{-d\delta_2}) = o(1),$$

for all $m \geq 1$, thus ψ should decay superpolynomially.

2.9 The PLT for rectangles

2.9.1 Quantitative Ergodic Theorem

In section 2.9.2, it will be convenient to use a pointwise (almost everywhere) Quantitative Ergodic Theorem instead of the L^2 bound we assume in (EE). Furthermore, using a Borel-Cantelli argument, we will show such bounds simultaneously along a sequence of functions $(f_l)_{l\geq 1}$.

Proposition 27. *Assume (EE) is satisfied, and let $(f_l)_{l\geq 1}$ be a sequence of functions in $C^{r'}$. Then, for every $\epsilon > 0, \epsilon' > 0$, there are, for ν-a.e $y \in Y$, $L_y \geq 1$ such that*

$$\left|\sum_{j=1}^{n} f_l \circ R^j(y) - n\int_Y f_l d\nu\right| \leq ||f_l||_{C^{r'}} n^{\delta+\epsilon} \quad \textit{for } \nu\textit{-a.e } y \in Y, \forall n \geq l^{\epsilon'}, l \geq L_y,$$

Proof. The proof of[19] [BS23, Theorem 3.1] shows that there is a constant $K > 0$ such that

$$\left\|\sup_{n\geq N}\frac{1}{n}\left|S_n(f) - n\int_Y f\,\mathrm{d}\nu\right|\right\|_{L^2} \leq K||f||_{C^{r'}}N^{-(1-\delta)} \quad \forall N \geq 1, f \in C^{r'}. \tag{2.28}$$

Now let $k, \delta' > 0$ to be chosen later be such that $2(1-\delta)k > 2\delta' + 1$. Using the Chebyshev inequality, from (2.28) it follows that

$$\begin{aligned}\nu\left(\sup_{n\geq N^k}\frac{1}{n}\left|S_n(f) - n\int_Y f\,\mathrm{d}\nu\right| > \frac{1}{2}||f||_{C^{r'}}N^{-\delta'}\right)&\\ \leq 4K^2N^{2\delta'-2(1-\delta)k}&\end{aligned} \tag{2.29}$$

for all $N \geq 1$ and $f \in C^{r'}$. For $l \geq 1, N = N_l = \lceil l^{\frac{2}{2(1-\delta)k-2\delta'-1}}\rceil$ denote by $B_{l,N}$ the set

$$B_{l,N} = \left\{y \in Y \mid \sup_{n\geq N^k}\frac{1}{n}\left|S_n(f_l)(y) - n\int_Y f_l\,\mathrm{d}\nu\right| > \frac{1}{2}||f_l||_{C^{r'}}N^{-\delta'}\right\}, \tag{2.30}$$

then it holds that $\nu(B_{l,N}) \leq N^{2\delta'-2(1-\delta)k}$ and

$$\begin{aligned}\sum_{l\geq 1, N\geq N_l}\nu(B_{l,N}) &\leq 4K^2\frac{1}{2\delta'-2(1-\delta)k+1}\sum_{l\geq 1}N_l^{2\delta'-2(1-\delta)k+1}\\ &\leq \frac{4K^2}{2\delta'-2(1-\delta)k+1}\sum_{l\geq 1}l^{-2} < \infty.\end{aligned}$$

[19] [BS23, Theorem 3.1] shows that (2.28) holds under the stronger assumption that $P^n \xrightarrow{L^2} Id$ polynomially fast on L^∞, where P denotes the Perron-Frobenius transfer operator. Note however that Lemma 3.3 of that paper shows that, under this assumption, (EE) holds; the rest of the proof only uses (EE), and not the stronger assumption on P.

Hence, by Borel-Cantelli, for ν-a.e $y \in Y$ there are only finitely many pairs (l, N) with $N \geq N_l$ such that $y \in B_{l,N}$, therefore, for such y, there is an $L_y \geq 1$ such that $y \notin B_{l,N}$ whenever $l \geq L_y, N \geq N_l$. For such y, l and $n \geq N_l^k$, say $N^k \leq n \leq (N+1)^k$ and $N \geq N_l$, it holds that

$$\begin{aligned}\frac{1}{n}\left|S_n(f_l)(y) - n\int_Y f \,\mathrm{d}\nu\right| &\geq \sup_{n \geq N^k} \frac{1}{n}\left|S_n(f)(y) - n\int_Y f \,\mathrm{d}\nu\right| \\ &\leq \frac{1}{2}||f_l||_{C^{r'}} N^{-\delta'} \leq ||f_l||_{C^{r'}} n^{-\frac{\delta'}{k}}.\end{aligned}$$

Choosing k so big that $\frac{1}{k} < \frac{\epsilon\epsilon'}{2}$, the claim follows by setting $\delta' = k(1-\delta-\epsilon)$ (then $2(1-\delta)k - 2\delta' - 1 = 2k\epsilon - 1 > \frac{2}{\epsilon'}$). ☐

2.9.2 Uniform estimates

Let $(B_l)_{l\geq 1}$ be a sequence of rare events[20] in Y, such that $(B_l)_{l\geq1}$ is regularly approximable in $C^{r'}$. The goal of this section will be to show that, after taking a subsequence if necessary[21]

$$\sup_{l\geq 1}\left|\frac{\nu(B_l)}{sN_l}\sum_{j=1}^{\lceil sN_l\rceil}\varphi_{B_l}^{(j)}(y) - 1\right| \to 0 \quad \text{as } s \to \infty,\ \nu\text{-a.e } y \in Y, \tag{2.31}$$

for some $N_l > 0$.

Lemma 28. *Let R satisfy (EE) and let $(B_l)_{l\geq1}$ be a sequence of rare events in Y, such that $(B_l)_{l\geq1}$ is regularly approximable in $C^{r'}$. Then, after taking a subsequence if necessary, for $\epsilon > 0$, there is a constant $K > 0$ only depending on app$(B_l)_{l\geq1}$, and, for ν-a.e $y \in Y$, $N_y > 0$ only depending on y and app$((B_l)_{l\geq1})$ such that*

$$\begin{aligned}&\left|\sum_{j=1}^{n} 1_{B_l}\circ R^j(y) - n\nu(B_l)\right| \\ &\leq K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\epsilon}{r'+1}}, \quad \forall n \geq N_y N_l, l \geq L_y\ \nu\text{-a.e } y \in Y,\end{aligned} \tag{2.32}$$

where $N_l = \nu(B_l)^{-\frac{r'+1}{1-\delta-\epsilon}}$.

Proof. Let $\mathcal{B} = \mathrm{app}((B_l)_{l\geq1})$, then[22] for $k > \nu(B_l)^{-\frac{1}{d'}}$ there are $\underline{h}_{k,l}, \overline{h}_{k,l} \in C^{r'}$ with $\underline{h}_{k,l} \leq 1_{B_l} \leq \overline{h}_{k,l}$ and

$$\nu(1_{B_l} \neq h_{k,l}) \leq \nu(B_l)^{\frac{d'-1}{d'}}\frac{1}{k}, \quad \text{while } ||h_{k,l}||_{C^{r'}} \leq \mathcal{B}k^{r'},$$

for $h_{k,l} \in \{\overline{h}_{k,l}, \underline{h}_{k,l}\}$. In particular

$$\nu(B_l) - \nu(B_l)^{\frac{d'-1}{d'}}\frac{1}{k} \leq \int_Y \underline{h}_{k,l}\,\mathrm{d}\nu \leq \int_Y \overline{h}_{k,l}\,\mathrm{d}\nu \leq \nu(B_l) + \nu(B_l)^{\frac{d'-1}{d'}}\frac{1}{k}.$$

By Proposition 27 there are $I_y \geq 1$ and $N_l \geq 1$ such that, for $h_{k,l} \in \{\overline{h}_{k,l}, \underline{h}_{k,l}\}$, we have

$$\begin{aligned}&\left|\sum_{j=1}^{n} h_{k,l}\circ R^j(y) - n\int_Y h_{k,l}\,\mathrm{d}\nu\right| \\ &\leq ||h_{k,l}||_{C^{r'}} n^{\delta+\epsilon} \quad \forall n \geq (l+k)^{\epsilon'}, l+k \geq I_y,\ \nu\text{-a.e } y \in Y.\end{aligned}$$

[20] Later on in §2.9.3 we will take finitely many such sequences $(B_l^{(1)})_{l\geq1}, ...(B_l^{(\mathcal{K})})_{l\geq1}$, but the same arguments apply.

[21] From this point on we will often take a subsequence of $(B_l)_{l\geq1}$ to assume that l is sufficiently small compared to $\nu(B_l)^{-1}$. In an effort to keep notation simple, we will do so without explicitly stating, accepting small imprecisions in exchange for simpler notation.

[22] By Definition 22.

Therefore, for such k, l, n,

$$\sum_{j=1}^{n} 1_{B_l} \circ R^j(y) \geq \sum_{j=1}^{n} \underline{h}_{k,l} \circ R^j(y) \geq n \int_Y \underline{h}_{k,l}\, \mathrm{d}\nu - \mathcal{B} k^r n^{\delta+\epsilon}$$
$$\geq n\nu(B_l) - n\nu(B_l)^{\frac{d'-1}{d'}} \frac{1}{k} - \mathcal{B} k^{r'} n^{\delta+\epsilon}.$$

Likewise

$$\sum_{j=1}^{n} 1_{B_l} \circ R^j(y) \leq n\nu(B_l) + n\nu(B_l)^{\frac{d'-1}{d'}} \frac{1}{k} + \mathcal{B} k^{r'} n^{\delta+\epsilon}.$$

Hence,

$$\left| \sum_{j=1}^{n} 1_{B_l} \circ R^j(y) - n\nu(B_l) \right| \leq n\nu(B_l)^{\frac{d'-1}{d'}} \frac{1}{k} + \mathcal{B} k^{r'} n^{\delta+\epsilon}. \tag{2.33}$$

Let $k_n = \left\lceil n^{\frac{1-\delta-\epsilon}{r'+1}} \nu(B_l)^{\frac{d'-1}{d'(r'+1)}} \right\rceil$ and

$$N_y = (100 I_y)^{\frac{r'+1}{1-\delta-\epsilon}}, \quad \text{and} \quad N_l = \nu(B_l)^{-\frac{r'+1}{(1-\delta-\epsilon)d'} - \frac{(d'-1)(r'+1)}{d'(1-\delta-\epsilon)}} = \nu(B_l)^{\frac{r'+1}{1-\delta-\epsilon}}.$$

We can apply (2.33) for $n \geq N_y N_l$, k_n and $l \geq 1$ (since $k_n > I_y$ and[23] $k_n^{\epsilon'} < k_n^{\frac{r'+1}{1-\delta-\epsilon}} < n$) to obtain

$$\left| \sum_{j=1}^{n} 1_{B_l} \circ R^j(y) - n\nu(B_l) \right| \leq K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\epsilon}{r'+1}},$$

where the constant $K > 0$ only depends on $\mathcal{B}$. □

Proposition 29. *Suppose R satisfies (EE) and let $(B_l)_{l\geq 1}$ be a sequence of rare events in Y, such that $(B_l)_{l\geq 1}$ is regularly approximable in $C^{r'}$. Then for every $\epsilon > 0$ there is a constant $K > 0$ and, for ν-a.e $y \in Y$, there are $S_y > 0$ such that, for $s > S_y$, we have*

$$\left| \frac{\nu(B_l)}{sM_l} \sum_{j=0}^{\lceil sM_l \rceil - 1} \varphi_{B_l} \circ R^j_{B_l}(y) - 1 \right| \leq K s^{-\frac{1-\delta-\epsilon}{r'+1}} \tag{2.34}$$

where

$$M_l = \nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\epsilon)}}.$$

Proof. Let $y \in Y$ be as in the conclusion of Lemma 28. By (2.32), for $l \geq 1$ and $n \geq N_y N_l$ we have

$$n\nu(B_l) - K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\epsilon}{r'+1}}$$
$$\leq \sum_{j=1}^{n} 1_{B_l} \circ R^j_{B_l}(y) \leq n\nu(B_l) + K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\epsilon}{r'+1}}.$$

and thus

$$\sum_{j=0}^{\lfloor n\nu(B_l) - K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\epsilon}{r'+1}} \rfloor - 1} \varphi_{B_l} \circ R^j_{B_l}(y) \leq n, \tag{2.35}$$

[23]To be completely correct, one would have to consider $k_n + l$ instead of l, but by choosing a subsequence and renumbering we can assume that l is very small compared to $\nu(B_l)^{-1}$.

as well as

$$n \leq \sum_{j=0}^{\lceil n\nu(B_l)+K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\epsilon}{r'+1}}\rceil -1} \varphi_{B_l} \circ R_{B_l}^j(y).$$

For $s > S_y$ with $S_y = (100K)^{\frac{r'+1}{1-\delta-\epsilon}} N_y$, let

$$n = \xi \nu(B_l)^{-\frac{d'+r'}{d'(1-\delta-\epsilon)}},$$

where $\xi > 0$ is such that $\xi - K\xi^{1-\frac{1-\delta-\epsilon}{r'+1}} = s$. Clearly

$$\xi \in \left(s, s + 2Ks^{1-\frac{1-\delta-\epsilon}{r'+1}}\right)$$

in particular[24] $n > N_y N_l$. Hence, using this n in the lower bound of (2.35) yields

$$\sum_{j=0}^{\lfloor s\nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\epsilon)}}\rfloor -1} \varphi_{B_l} \circ R_{B_l}^j(y) \leq n \leq (s + 2Ks^{1-\frac{1-\delta-\epsilon}{r'+1}})\nu(B_l)^{-\frac{d'+r'}{d'(1-\delta-\epsilon)}},$$

setting $M_l = \nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\epsilon)}}$ we obtain

$$\frac{\nu(B_l)}{sM_l} \sum_{j=0}^{\lfloor sM_l \rfloor -1} \varphi_{B_l} \circ R_{B_l}^j(y) \leq 1 + 2Ks^{-\frac{1-\delta-\epsilon}{r'+1}}.$$

The upper bound follows analogously by setting $\xi + K\xi^{1-\frac{1-\delta-\epsilon}{r'+1}} = s$, and the claim (making K a bit bigger) is proven. □

2.9.3 PLT scaled by returns to $\{B_l\}$

For the rest of this exposition let $\mathcal{K} \geq 1$, and $A_l^{(1)}, ..., A_l^{(\mathcal{K})}$ resp. $B_l^{(1)}, ..., B_l^{(\mathcal{K})}$ be subsets of X resp. Y regularly approximable in C^r resp. $C^{r'}$. Suppose that there are $r_l \searrow 0$, and positive constants $\omega^{(k)}, \theta^{(k)} > 0$ such that

$$\mu(A_l^{(k)}) = \omega^{(k)} r_l^d + o(r_l^d), \quad \text{and} \quad \nu(B_l^{(k)}) = \theta^{(k)} r_l^{d'} + o(r_l^{d'}) \quad \forall l \geq 1, k = 1, ..., \mathcal{K},$$

and, for each l, $B_l^{(1)}, ..., B_l^{(\mathcal{K})}$ are disjoint. Denote $B_l = \bigcup_{k=1}^{\mathcal{K}} B_l^{(k)}$, which, by disjointness of $B_l^{(1)}, ..., B_l^{(\mathcal{K})}$, is also regularly approximable in $C^{r'}$, and $\alpha_l = \Phi_{B_l}$. Consider

$$\kappa_l^{(n)}(y) = k \quad \text{if} \quad R_{B_l}^n(y) \in B_l^{(k)},$$

which is well-defined as the sets $B_l^{(k)}$ are disjoint, and

$$\begin{aligned}
\sigma_{\kappa_l,\alpha_l,y}^{(1)}(x) &= \min\left(n \geq 1 \mid T_y^{\tilde{\alpha}_l^{(n)}}(x) \in A_l^{(\kappa_l^{(n)})}\right), \\
\sigma_{\kappa_l,\alpha_l,y}^{(n+1)}(x) &= \min\left(k \geq \sigma_{\kappa_l,\alpha_l,y}^{(n)}(x) + 1 \mid T_y^{\tilde{\alpha}_l^{(k)}}(x) \in A_l^{(\kappa_l^{(k)})}\right), \\
\Sigma_{\kappa_l,\alpha_l,y} &= (\sigma_{\kappa_l,\alpha_l,y}^{(1)}, \sigma_{\kappa_l,\alpha_l,y}^{(2)}, ...).
\end{aligned}$$

[24] Recall that $N_l = \nu(B_l)^{-\frac{1+r'}{d'(1-\delta-\epsilon)}}$.

Following the steps outlined in §2.7 we will first show

$$\Omega r_l^{d'} \Sigma_{\kappa_l,\alpha_l,y} \overset{\mu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty,\ \nu - a.a\ y \in Y \tag{2.36}$$

for $\Omega = \frac{\sum_{k=1}^{\mathcal{K}} \omega^{(k)}\theta^{(k)}}{\sum_{k=1}^{\mathcal{K}} \theta^{(k)}}$. Then use the relation

$$\sigma_{\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}}(x,y) = \sum_{j=0}^{\sigma_{\kappa_l,\Phi_{B_l}(y)}(x)-1} \varphi_{B_l} \circ R_{B_l}^j(y) \tag{2.37}$$

to obtain

$$(\mu \times \nu)\left(\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}\right) \Sigma_{\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty. \tag{2.38}$$

Denote

$$p_{l,t}^{(k)}(y) := \frac{1}{t}\#\{j = 1, ..., t \mid \kappa_l^{(j)}(y) = k\} = \frac{1}{t}\sum_{j=1}^{t} 1_{B_l^{(k)}}(R_{B_l}^j(y)).$$

We first show that, for each k, ν-a.e $y \in Y$, and $t_l = O(r_l^{-d})$, we have

$$p_{l,t_l}^{(k)}(y) \to \frac{\theta^{(k)}}{\sum_{j=1}^{\mathcal{K}} \theta^{(j)}} =: p^{(k)} \quad \text{as } l \to \infty. \tag{2.39}$$

Lemma 30. *Suppose R satisfies (EE) and*

$$d > \frac{r' + d'\delta}{1-\delta}$$

Then κ_l satisfies (2.39)*.*

Proof. Let $\epsilon > 0$ be small enough that

$$d > \frac{r' + d'(\delta+\epsilon)}{1-\delta-\epsilon}.$$

(i) Using Lemma 28, and disjointness we obtain a constant $K > 0$ and $N_y \geq 1$ such that, for ν-a.e $y \in Y$, we have

$$\left|\sum_{j=1}^{n} 1_{B_l^{(k)}} \circ R^j(y) - n\nu(B_l^{(k)})\right| \leq K r_l^{\frac{r'(d'-1)}{r'+1}} n^{1-\frac{1-\delta-\epsilon}{r'+1}}, \quad \text{and}$$

$$\left|\sum_{j=1}^{n} 1_{B_l} \circ R^j(y) - n\nu(B_l)\right| \leq K r_l^{\frac{r'(d'-1)}{r'+1}} n^{1-\frac{1-\delta-\epsilon}{r'+1}}$$

for all $n \geq N_y, l \geq 1$ and $k = 1, ..., \mathcal{K}$. On the other hand[25], for $s > N_y$, Proposition 29 yields

$$\left|\frac{\nu(B_l)}{sM_l}\sum_{j=0}^{\lceil sM_l\rceil-1} \varphi_{B_l} \circ R_{B_l}^j(y) - 1\right| \leq K s^{-\frac{1-\delta-\epsilon}{r'+1}}.$$

where

$$M_l = \nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\epsilon)}}.$$

[25] Making K and N_y bigger if needed.

(ii) Rewrite

$$p_{l,n}^{(k)} = \frac{1}{n} \sum_{i=1}^{\sum_{j=1}^{n} \varphi_{B_l} \circ R_{B_l}^j(y)} 1_{B_l^{(k)}} \circ R^i.$$

Denote $t_l = s_l M_l$, since

$$M_l = \nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\epsilon)}} = O(r_l^{\frac{d'+r'}{1-\delta-\epsilon}-d'}) = o(r_l^d),$$

we necessarily have $s_l \to \infty$ as $l \to \infty$. Denote $a = \frac{1-\delta-\epsilon}{r'+1}$ and let l be big enough so that $s_l > N_y$, then we have

$$\begin{aligned}
t_l p_{l,t_l}^{(k)}(y) &= \sum_{j=1}^{\sum_{j=1}^{t_l} \varphi_{B_l} \circ R_{B_l}^j - 1} 1_{B_l^{(k)}} \circ R^j(y) \\
&\leq \sum_{j=1}^{t_l \nu(B_l)^{-1}(1+K s_l^{-a})-1} 1_{B_l^{(k)}} \circ R^j(y) \\
&\leq t_l \nu(B_l)^{-1} \nu(B_l^{(k)})(1+Ks^{-a}) + K r_l^{\frac{r'(d'-1)}{r'+1}} \left(t_l \nu(B_l)^{-1}(1+Ks^{-a})\right)^{1-a} \\
&\leq t_l \frac{\nu(B_l^{(k)})}{\nu(B_l)} (1 + o(s_l^{-a+a^2})).
\end{aligned}$$

The lower bound is similar. So

$$\left| p_{l,n}^{(k)}(y) - \frac{\nu(B_l^{(k)})}{\nu(B_l)} \right| = o(s_l^{-a+a^2}) = o(1). \tag{2.40}$$

□

2.9.4 Adding in the gaps

Now all that's left to do is to add back in the gaps. As mentioned in §2.9.3, having shown (2.36) (this is the content of Proposition 25), we will now explain how to conclude

$$(\mu \times \nu) \left(\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)} \right) \Sigma_{\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}} \overset{\mu \times \nu}{\Rightarrow} \Sigma_{Exp}, \quad \text{as } l \to \infty$$

using the relation

$$\sigma_{\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}} = \sum_{j=0}^{\sigma_{\mathcal{K}_l,\alpha_l,y}(x)-1} \varphi_{B_l} \circ R_{B_l}^j(y).$$

This is rather straightforward, given Proposition 29, and follows from a more general principle in probability theory. As this principle finds use in various places and has, to the author's knowledge, not been formulated in generality, let us state and prove a more general version than we need here.

Lemma 31. *Let $(\Omega, \mathbb{P})$ be a probability space, and $E_l : \Omega \to [0, \infty)$ non-negative real random variables, such that there are positive random variables $\mu_l : \Omega \to (0, 1)$ with*

$$\mu_l E_l \overset{\mathbb{P}}{\Rightarrow} E \quad \textit{as } l \to \infty,$$

for some non-negative random variable E with $\mathbb{P}(E=0)=0$. Then for any $M_l : \Omega \to [0,\infty)$ with

$$\mu_l M_l \to 0 \quad \text{as } l \to \infty \ \text{ pointwise } \mathbb{P}\text{-a.e}$$

we have

$$\mathbb{P}(E_l \le M_l) \to 0 \quad \text{as } l \to \infty.$$

Proof. Let $\epsilon > 0$, there is a $\delta > 0$ such that $\mathbb{P}(E \le \delta) < \epsilon$, and the distribution function of E is continuous at δ. By Jegorow's Theorem, there is a measurable $K \subset \Omega$ with $\mathbb{P}(K^c) < \epsilon$ and

$$\mu_l M_l \to 0 \quad \text{as } l \to \infty \ \text{ uniformly on } K.$$

Choose $\tilde{L} \ge 1$ so big that $M_l \le \delta \mu_l^{-1}$ on K for $l \ge \tilde{L}$. Now choose $L \ge \tilde{L}$ so big that

$$\mathbb{P}(\mu_l E_l \le \delta) \le 2\epsilon \quad \forall l \ge L,$$

it follows that

$$\mathbb{P}(E_l \le M_l) \le \mathbb{P}(\mu_l E_l \le \delta) + \mathbb{P}(K^c) \le 3\epsilon,$$

for $l \ge L$. □

Proposition 32. *Let $(\Omega, \mathbb{P})$ be a probability space, and $E_l : \Omega \to \mathbb{N}$ be positive integer valued observables. Assume there are positive real numbers $q_l \searrow 0$, and a $[0,\infty)$-valued random variable E with $\mathbb{P}(E=0)=0$ such that*

$$q_l E_l \overset{\mathbb{P}}{\Rightarrow} E \quad \text{as } l \to \infty,$$

Let $\alpha_j^{(l)} : \Omega \to [0,\infty)$ be non-negative random variables, and assume there are $M_l : \Omega \to (0,\infty)$ with $q_l M_l \to 0$ as $l \to \infty$ $\mathbb{P}$-a.e and positive random variables $b_l : \Omega \to (0,\infty)$ such that

$$\sup_{l \ge 1} \left| \frac{1}{s M_l b_l} \sum_{j=1}^{\lceil s M_l \rceil} \alpha_l^{(j)}(\omega) - 1 \right| \to 0 \quad \text{as } s \to \infty, \ \ \mathbb{P}\text{-a.e.} \ . \tag{UC}$$

Then

$$\frac{q_l}{b_l} \sum_{j=1}^{E_l} \alpha_l^{(j)} \overset{\mathbb{P}}{\Rightarrow} E \quad \text{as } l \to \infty.$$

Remark 33. *In our context we use $(\Omega, \mathbb{P}) = (X \times Y, \mu \times \nu)$, $E_l = \sigma_{A_l, \Phi_{B_l}}$, $\alpha_l^{(j)} = \varphi_{B_l} \circ R_{B_l}^j$, $q_l = \mu(A_l)$ and $b_l = \frac{1}{\nu(B_l)}$. The existence of M_l is the content of Proposition 29.*

Proof. (i) Let F be the distribution function of E and $C = \{t \mid F \text{ is continuous at } t\}$ the set of its continuities. Let $t \in C$, and $\epsilon > 0$ such that $\frac{t}{1+\epsilon}, \frac{t}{1-\epsilon} \in C$. We will show

$$\mathbb{P}\left(\frac{q_l}{b_l} \sum_{j=1}^{E_l} \alpha_l^{(j)} \le t \right) \to F(t) \quad \text{as } l \to \infty.$$

(ii) By Jegorow's Theorem, for $l \ge 1$, there is a measurable set $K \subset \Omega$ with $\mathbb{P}(K^c) < \epsilon$

$$\sup_{l \ge 1} \sup_{\omega \in K} \left| \frac{1}{s M_l b_l} \sum_{j=1}^{\lceil s M_l \rceil} \alpha_l^{(j)}(\omega) - 1 \right| \to 0 \quad \text{as } s \to \infty.$$

(iii) Choose $S > 0$ so big that

$$(1-\epsilon)b_l \leq \frac{1}{sM_l}\sum_{j=1}^{\lceil sM_l\rceil} \alpha_l^{(j)} \leq (1+\epsilon)b_l \quad \text{on } K,\ \forall s \geq S, l \geq 1,$$

and restricting to $\{E_l \geq SM_l\}$, we obtain

$$(1-\epsilon)q_l E_l \leq \frac{q_l}{b_l}\sum_{j=1}^{E_l} \alpha_l^{(j)}(\omega) \leq (1+\epsilon)q_l E_l \quad \text{on } K \cap \{E_l \geq SM_l\},\ \forall l \geq 1.$$

We get

$$\begin{aligned}&\mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \leq t\right) - \mathbb{P}(E_l \leq SM_l) - \mathbb{P}(K^c)\\ &\leq \mathbb{P}\left(K \cap \{E_l \geq SM_l\} \cap \left\{\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \leq t\right\}\right) \leq \mathbb{P}\left(q_l E_l \leq \frac{t}{1-\epsilon}\right),\end{aligned}$$

and likewise

$$\begin{aligned}\mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \leq t\right) &\geq \mathbb{P}\left(K \cap \{E_l \geq SM_l\} \cap \left\{q_l E_l \leq \frac{t}{1+\epsilon}\right\}\right)\\ &\geq \mathbb{P}\left(q_l E_l \leq \frac{t}{1+\epsilon}\right) - \mathbb{P}(E_l \leq SM_l) - \mathbb{P}(K^c).\end{aligned}$$

Taking $\limsup_{l\to\infty}$ resp. $\liminf_{l\to\infty}$ the above two equations, and using Lemma 31, we obtain

$$\begin{aligned}F\left(\frac{t}{1+\epsilon}\right) - \epsilon &\leq \liminf_{l\to\infty}\mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \leq t\right)\\ &\leq \limsup_{l\to\infty}\mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \leq t\right) \leq F\left(\frac{t}{1-\epsilon}\right) + \epsilon.\end{aligned}$$

Since $C \subset (0,\infty)$ is dense, we can let $\epsilon \searrow 0$ while $\frac{t}{1+\epsilon}, \frac{t}{1-\epsilon} \in C$, this yields

$$\mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \leq t\right) \to F(t) \quad \text{as } l \to \infty.$$

□

Remark 34.

(i) We can extend this statement to sequences in the following manner: under the assumptions of the proposition, let $E_l^{(n)} : \Omega \to \mathbb{N}$ *be such that*

$$q_l(E_l^{(1)}, E_l^{(2)}, \ldots) \overset{\mu}{\Rightarrow} (E^{(1)}, E^{(2)}, \ldots) \quad \textit{as } l \to \infty,$$

for some $E^{(n)} : \Omega \to [0, \infty)$ *with* $\mathbb{P}(E^{(n)} = 0) = 0$. *Then*

$$\frac{q_l}{b_l}\left(\sum_{j=1}^{E_l^{(n)}} \alpha_l^{(j)}, \sum_{j=1}^{E_l^{(n)}} \alpha_l^{(j)}, ...\right) \overset{\mathbb{P}}{\Rightarrow} (E^{(1)}, E^{(2)}, ...) \quad \textit{as } l \to \infty.$$

The proof of this statement is almost the same as for the proposition, therefore we won't repeat it.

(ii) The probability measure $\mathbb{P}$ *can be replaced by a sequence* $(\mathbb{P}_l)_{l\geq 1}$ *by also replacing* (UC) *with the following condition*

$$\forall \epsilon > 0 \textit{ there is are measurable sets } K_l \subset \Omega \textit{ with } \limsup_{l\to\infty} \mathbb{P}_l(K_l^c) < \epsilon \textit{ such that}$$

$$\sup_{l\geq 1} \sup_{\omega\in K_l} \left| \frac{1}{sM_l b_l} \sum_{j=1}^{\lceil sM_l \rceil} \alpha_l^{(j)}(\omega) - 1 \right| \to 0 \quad \textit{as } s \to \infty.$$

Proof of Theorem 7. (i) Let $r_l \searrow 0$, and denote by $Q_l = B_{r_l}(x^*, y^*)$ the geodesic ball of radius r_l centred at (x^*, y^*), and let $\epsilon > 0$. W.l.o.g. r_1 is small enough that the exponential map at (x^*, y^*) is a diffeomorphism from the ball of radius $2r_1$ in $\mathbb{R}^{d+d'}$ onto $B_{2r_1}(x^*, y^*)$. Let $\epsilon > 0$, it is easy to construct, for some $\mathcal{K} \geq 1$, sets $A_l^{(1)}, ..., A_l^{(\mathcal{K})}$ and $B_l^{(1)}, ..., B_l^{(\mathcal{K})}$ as in §2.8. Let $\omega^{(k)}, \theta^{(k)} > 0$ be such that[26]

$$\mu(A_l^{(k)}) = \omega^{(k)} r_l^d + o(r_l^d), \quad \text{and} \quad \nu(B_l^{(k)}) = \theta^{(k)} r_l^{d'} + o(r_l^{d'}) \quad \forall l \geq 1, k = 1, ..., \mathcal{K},$$

and set $\Omega = \frac{\sum_{k=1}^{\mathcal{K}} \omega^{(k)}\theta^{(k)}}{\sum_{k=1}^{\mathcal{K}} \theta^{(k)}}$.

Due to Lemma 23 all those sets can be chosen to be regularly approximable, such that[27] $(\mu \times \nu)_{Q_l}\left(Q_l \setminus \bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}\right) < \epsilon$ for all $l \geq 1$. Let $\Lambda > 0$ be such that $(\mu \times \nu)(Q_l) = \Lambda r_l^{d+d'} + o(r_l^{d+d'})$, then

$$\left|\Lambda - \Omega \sum_{k=1}^{\mathcal{K}} \theta^{(k)}\right| < \epsilon$$

(ii) Denote $B_l = \bigcup_{k=1}^{\mathcal{K}} B_l^{(k)}$, and, for y as in assumption (MEM), consider $\alpha_l(y) = \Phi_{B_l}(y)$ and

$$\kappa_l^{(n)}(y) = k \quad \text{if} \quad R_{B_l}^n(y) \in B_l^{(k)},$$

by disjointness $\kappa_l(y)$ is well-defined. By Lemma 30, $\kappa_l(y)$ satisfies (2.39), and $p^{(k)} = \frac{\theta^{(k)}}{\sum_{s=1}^{\mathcal{K}} \theta^{(s)}}$. We can use Proposition 25 and Remark 26(i) to obtain

$$\Omega r_l^d \Sigma_{\kappa_l(y),\alpha_l(y),y} \overset{\mu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty, \ \nu\text{-a.e } y.$$

Since the convergence holds for ν-a.e $y \in Y$, it follows that

$$\Omega r_l^d \Sigma_{\kappa_l,\alpha_l} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty,$$

where $\Sigma_{\kappa_l,\alpha_l}(x, y) = \Sigma_{\kappa_l(y),\alpha_l(y),y}(x)$.

[26] Choose x^*, y^* such that the densities of μ, ν are positive at the respective points.

[27] Here we again use the continuity of the density.

(iii) By Proposition 29, α_l satisfies (UC) with

$$b_l = \frac{1}{\nu(B_l)} = \frac{1}{\sum_{k=1}^{\mathcal{K}} \theta^{(k)}} r_l^{-d'} + o(r_l^{-d'})$$

and

$$M_l = \nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\epsilon)}}.$$

Note that, for $Q_l' = \bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}$, we have

$$\Sigma_{Q_l'} = \left(\sum_{j=0}^{\sigma^{(1)}_{\kappa_l,\alpha_l}-1} \varphi_{B_l} \circ R^j_{B_l}, \sum_{j=\sigma^{(1)}_{\kappa_l,\alpha_l}}^{\sigma^{(2)}_{\kappa_l,\alpha_l}-1} \varphi_{B_l} \circ R^j_{B_l}, \ldots \right).$$

We can apply Proposition 32 resp. Remark 34(i)

$$\Omega \left(\sum_{k=1}^{\mathcal{K}} \theta^{(k)} \right) r_l^{d+d'} \Sigma_{Q_l'} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty.$$

By disjointness

$$(\mu \times \nu)(Q_l') = \Omega \left(\sum_{k=1}^{\mathcal{K}} \theta^{(k)} \right) r_l^{d+d'} + o(r_l^{d+d'})$$

hence

$$(\mu \times \nu)(Q_l')\Sigma_{Q_l'} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty.$$

By the equivalence (2.12), we have

$$(\mu \times \nu)(Q_l')\Phi_{Q_l'} \overset{\mu\times\nu}{\Rightarrow} \Phi_{Exp} \quad \text{as } l \to \infty.$$

(iv) By Theorem 4 also

$$\mu \times \nu)(Q_l')\Phi_{Q_l'} \overset{\mu\times\nu_{Q_l'}}{\Rightarrow} \Phi_{Exp} \quad \text{as } l \to \infty.$$

At the same time, taking a subsequence if necessary, there are $[0,\infty]$-valued processes Φ and $\tilde{\Phi}$ such that

$$(\mu \times \nu)(Q_l)\Phi_{Q_l} \overset{\mu\times\nu}{\Rightarrow} \Phi \quad \text{as } l \to \infty, \quad \text{and}$$
$$(\mu \times \nu)(Q_l)\Phi_{Q_l} \overset{(\mu\times\nu)_{Q_l}}{\Rightarrow} \tilde{\Phi} \quad \text{as } l \to \infty.$$

Hence

$$D\left(\Phi_{Exp}, \tilde{\Phi}\right) \leq 7\epsilon,$$

where D is given by Lemma 21. Since this is true for every $\epsilon > 0$ we have $\tilde{\Phi} \overset{d}{=} \Phi_{Exp}$, and by Theorem 4 also $\Phi \overset{d}{=} \Phi_{Exp}$. □

2.10 The skewing time

Here we will prove Theorem 9, to do this we will verify that the map $T(x,y) = G_{\tau(y)}(x)$ satisfies superpolynomial mixing of all orders, as in Remark 26(ii).

Lemma 35. *Under the assumptions of Theorem 9, suppose that* $\sum_{l\geq 1} r_l^{\frac{1}{2}(d'-\delta_2\kappa)} < \infty$*. Then, for each* $t>0$*, there is a set* $\mathfrak{G}_t$ *with* $\nu(\mathfrak{G}_t)=1$ *such that, for* $y\in\mathfrak{G}_t$*, there are* $L_{y,t}>0$ *and sets* $\mathcal{B}_{l,y,t}\subset\{1,...,\left\lceil\frac{t}{\mu(A_l)}\right\rceil\}$ *with* $\#\mathcal{B}_{l,y,t}=o(\mu(A_l)^{-1})$ *such that*

$$\begin{aligned}&|\tau_{\tilde{\alpha}_l^{(n)}}(y)-\tau_{\tilde{\alpha}_l^{(m)}}(y)|\\ &\geq \zeta\left(\tilde{\alpha}_l^{(n)}(y)-\tilde{\alpha}_l^{(m)}(y)\right) \quad \forall l\geq L_{y,t}, 1\leq n<m\leq \left\lceil\frac{t}{\mu(A_l)}\right\rceil, n\notin\mathcal{B}_{l,y,t}\end{aligned}$$

where ζ *is the function from condition (BA) and*[28] $\alpha_l=\Phi_{B_l}$*.*

Proof. Fix $t>0$, to keep notation simple we assume $\mu(A_l)=r_l^d+o(r_l^d)$ and $\nu(B_l)=r_l^{d'}+o(r_l^{d'})$, otherwise there is an extra constant in the estimates below.

(i) We call $n\in\left\{1,...,\left\lceil\frac{t}{\mu(A_l)}\right\rceil\right\}$ a (l,y)-bad return (or simply (l,y)-bad) if there is a $m>n$ such that

$$|\tau_{\tilde{\alpha}_l^{(n)}}(y)-\tau_{\tilde{\alpha}_l^{(m)}}(y)|<\zeta(\tilde{\alpha}_l^{(n)}(y)-\tilde{\alpha}_l^{(m)}(y)),$$

denote $\mathcal{B}_{l,y}=\{n\geq 1\mid n \text{ is a } (l,y)-\text{bad return}\}$. Let $\epsilon_1>0$, we call $y\in Y$ an l-bad point if $\#\mathcal{B}_{l,y}>r_l^{-d+\epsilon_1}$.

(ii) Using Proposition 29 and Jegorov's Theorem, for $\epsilon_2>0$, we can find a measurable $G=G_{\epsilon_2}\subset Y$ with $\nu(G^c)<\epsilon_2$ and an $\tilde{L}\geq 1$ depending on G such that

$$\tilde{\alpha}_l^{\left(\left\lfloor\frac{t}{\mu(A_l}\right\rfloor\right)}\leq 2tr_l^{-(d+d')} \quad \forall y\in G, l\geq\tilde{L}. \tag{2.41}$$

(iii) For $l\geq 1$ denote $G_l=\{\tilde{\alpha}_l^{\left(\left\lfloor\frac{t}{\mu(A_l}\right\rfloor\right)}\leq 2tr_l^{-(d+d')}\}$, we have

$$\begin{aligned}\nu(y\in G_l\mid y\text{ is } l\text{-bad})&\leq\frac{\int_{G_l}\#\mathcal{B}_{l,y}\,\mathrm{d}\nu}{r_l^{-d+\epsilon_1}}\leq r_l^{d-\epsilon_1}\sum_{n=1}^{\left\lceil\frac{t}{\mu(A_l)}\right\rceil}\nu(y\in G_l\mid n\text{ is }(l,y)\text{-bad})\\
&\leq r_l^{d-\epsilon_1}\sum_{j=1}^{\lceil 2tr_l^{-(d+d')}\rceil}\nu\left(\exists i\geq 1\mid|\tau_{j+\tilde{\alpha}_l^{(i)}}-\tau_j|<\zeta\left(\tilde{\alpha}_l^{(i)}\right)\right)\\
&\leq r_l^{d-\epsilon_1}\sum_{j=1}^{\lceil 2tr_l^{-(d+d')}\rceil}\nu\left(R^{-j}\left(\exists i\geq 1\mid|\tau_{\tilde{\alpha}_l^{(i)}}|<\zeta\left(\tilde{\alpha}_l^{(i)}\right)\right)\right)\\
&\leq r_l^{d-\epsilon_1}\sum_{j=1}^{\lceil 2tr_l^{-(d+d')}\rceil}\nu(\exists i\geq r_l^{-\delta_2}\mid|\tau_i|<\zeta(i))\\
&\leq 2Ktr_l^{d-\epsilon_1-d-d'+\delta_2\kappa}\leq 2Ktr_l^{\delta_2\kappa-d'-\epsilon_1},\end{aligned}$$

for some constant $K>0$, for small enough ϵ_1 this is summable. An application of the Borel-Cantelli Lemma yields that for almost every $y\in Y$; for big enough l, either $y\notin G_l$ or

$$|\tau_{\tilde{\alpha}_l^{(n)}}(y)-\tau_{\tilde{\alpha}_l^{(m)}}(y)|\geq\zeta(\tilde{\alpha}_l^{(n)}(y)-\tilde{\alpha}_l^{(m)}(y)) \quad \forall 1\leq n<m\leq\left\lceil\frac{t}{\mu(A_l)}\right\rceil, n\notin\mathcal{B}_{l,y,}.$$

At the same time, by (2.41), we have $G_l\nearrow Y$. Thus ν-a.e $y\in Y$ is in G_l for big enough l, and the conclusion follows. $\square$

[28] It can be recalled from Definition 16 that $\tilde{\alpha}_l^{(n)}=\sum_{j=1}^n\alpha_l^{(j)}=\sum_{j=0}^{n-1}\varphi_{B_l}\circ R^j$.

Proof of Theorem 9. In order to keep notation simple we will only show the PLT for regularly approximable rectangles, this can be easily extended to geodesic balls, by following the same arguments as in the proof of Theorem 7.

(i) For ν-a.e $y \in Y$ and $0 = t_0 < t_1 < ... < t_J$ choose $L_y = L_{y,t_1+...+t_J}$ and $\mathcal{B}_{l,y} = \mathcal{B}_{l,y,t_1+...+t_J}$ as in Lemma 35. For such a y and $l \geq L_y$ (in the following we suppress y from the notation) consider

$$S_{t_j,l} = \sum_{i=1}^{\left\lceil \frac{t_j}{\mu(A_l)} \right\rceil} 1_{A_l} \circ T^{\tilde{\alpha}_l^{(i)}} = S'_{t_j,l} + S''_{t_j,l},$$

where

$$S'_{t_j,l} = \sum_{\substack{i=1,...,\left\lceil \frac{t_j}{\mu(A_l)} \right\rceil \\ i \notin \mathcal{B}_l}} 1_{A_l} \circ T^{\tilde{\alpha}_l^{(j)}}.$$

As in Proposition 25, the first goal is to show

$$\left(S_{t_1,l} - S_{t_0,l}, ..., S_{t_J,l} - S_{t_{J-1},l}\right) \stackrel{\mu}{\Rightarrow} \left(P_{t_1-t_0}, ..., P_{t_J-t_{J-1}}\right) \quad \text{as } l \to \infty,$$

where (P_t) is a standard Poisson process. Since $||S''_{t_j,l}||_{L^1} \to 0$ for all $j = 1, ..., J$, it is equivalent to show

$$\left(S'_{t_1,l} - S'_{t_0,l}, ..., S'_{t_J,l} - S'_{t_{J-1},l}\right) \stackrel{\mu}{\Rightarrow} \left(P_{t_1-t_0}, ..., P_{t_J-t_{J-1}}\right) \quad \text{as } l \to \infty.$$

For $m_1, ..., m_J \geq 1$ it will be enough to show

$$\int_X \prod_{j=1}^{J} \binom{S'_{t_j,l} - S'_{t_{j-1},l}}{m_j} \mathrm{d}\mu = \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!}. \tag{2.42}$$

We have

$$\prod_{j=1}^{J} \binom{S'_{t_j,l} - S'_{t_{j-1},l}}{m_j} = \sum_{\substack{\left\lceil \frac{t_{j-1}}{\mu(A_l)} \right\rceil + 1 \leq k_{j,1} < ... < k_{j,m_j} \leq \left\lceil \frac{t_j}{\mu(A_l)} \right\rceil \\ k_{j,i} \notin \mathcal{B}_l \text{ for } j=1,...,J,\ i=1,...,m_j}} \prod_{i,j} 1_{A_l} \circ T^{\tilde{\alpha}_l^{(k_{j,i})}}. \tag{2.43}$$

(ii) Due to assumption (MEM) for G, and Lemma 35, we have

$$\left| \int_X \prod_{j=1}^{m} f_j \circ T^{\tilde{\alpha}_l^{(n_j)}} \mathrm{d}\mu - \prod_{j=1}^{m} \int_X f_j \,\mathrm{d}\mu \right| \leq C_y \psi(\min_{j \neq j'} |\tilde{\alpha}_l^{(n_j)} - \tilde{\alpha}_l^{(n_{j'})}|) \prod_{j=1}^{m} ||f_j||_{C^r},$$

where $\psi(x) = e^{-\gamma \zeta(x)}$ and ζ is as in assumption (BA), for $f_1, ..., f_m \in C^r$ and $1 \leq n_1 \leq ... \leq n_m \leq \left\lceil \frac{t1+...+t_J}{r_l^d} \right\rceil$ with $n_i \notin \mathcal{B}_l$. Due to (2.9) and assumption (BA) we have

$$\psi(\min_{j \neq j'} |\tilde{\alpha}_l^{(n_j)} - \tilde{\alpha}_l^{(n_{j'})}|) = O(r_l^{w_l}),$$

for some $w_l > 0$ with $w_l \to \infty$ as $l \to \infty$. Approximating 1_{A_l} by functions in C^r it is straightforward[29] to show that

$$\left| \int_X \prod_{i,j} 1_{A_l} \circ T^{\tilde{\alpha}_l^{(k_{j,i})}} \mathrm{d}\mu - \mu(A_l)^{m_1+...+m_J} \right| = o(\mu(A_l)^{m_1+...+m_J}),$$

for $k_{j,i}$ as in (2.43). The sum in (2.43) has

$$\mu(A_l)^{-(m_1+...+m_J)} \prod_{j=1}^{J} \frac{(t_j - t_{j-1})_j^m}{m_j!} + o(\mu(A_l)^{-(m_1+...+m_J)})$$

many terms, so (2.42) follows. $\square$

[29]The calculation is analogous to Lemma 24.

2.11 Examples

Here, we verify conditions (EE) and (BR) for the examples listed in §2.4.

2.11.1 Diophantine rotations

Let $\alpha \in ((0,1)\setminus\mathbb{Q})^{d'}$ satisfy a Diophantine condition, i.e. there are $C>0$ and $n\geq 1$ such that

$$|\langle k,\alpha\rangle - l| > C|k|^{-n} \quad \forall k\in\mathbb{Z}^{d'}, k\neq 0, l\in\mathbb{Z}, \tag{D}$$

and $R=R_\alpha : x\mapsto x+\alpha \ (mod\ 1)$, for $x\in\mathbb{T}^{d'}$, the rotation by α. Almost all α satisfy (D) for some $n>d'$ (this is a consequence of a higher dimensional version of Khinchin's Theorem, see e.g. [BRV16]). If $d'=n=1$, then we say α is of *bounded type.*

Note that (D) implies that there is a constant $C'>0$ such that

$$\left|1-e^{2\pi i\langle k,\alpha\rangle}\right| \geq C'|k|^{-n} \quad \forall k\in\mathbb{Z}^{d'}\setminus\{0\}.$$

Property $(SLR(y^*))$ follows directly from (D). Due to the self-symmetry of R_α it is enough to consider returns of $x=0$ to a rectangle $(-r,r)^{d'}$, but

$$R_\alpha^m(0)\in(-r,r)^{d'} \iff |m\alpha-k|<r \quad \text{for some } k\in\mathbb{Z}^{d'}.$$

Then, by (D), $Cm^{-n}<r$, equivalently $m>(C^{-1}r)^{-\frac{1}{n}}$. Hence, (2.4) is satisfied with $\psi(r)=(C^{-1}r)^{-\frac{1}{n}}$.

To show effective equidistribution (EE), we solve the homological equation. Let $f\in H^n(\mathbb{T}^{d'})$ with $\int f=0$. Then

$$f(x)=\sum_{k\in\mathbb{Z}^{d'}} a_k e^{2\pi i\langle k,x\rangle},$$

where $\sum_{k\in\mathbb{Z}^{d'}}|a_k|^2\sum_{j_1+\ldots+j_{d'}=n}\prod_{i=1}^{d'}|k_i|^{2j_i}<\infty$ and $a_0=0$. To solve $f=g-g\circ R_\alpha$ we write

$$g(x)=\sum_{k\in\mathbb{Z}^{d'}} b_k e^{2\pi i\langle k,x\rangle}.$$

By comparing coefficients, this is satisfied for $b_k=\frac{a_k}{1-e^{2\pi i\langle k,\alpha\rangle}}$ for $k\neq 0$ and $b_0=0$. We have

$$\sum_{k\in\mathbb{Z}^{d'}}|b_k|^2 \leq (C')^2\sum_{k\in\mathbb{Z}^{d'}}|a_k|^2|k|^{2n} \leq (C')^2\sum_{k\in\mathbb{Z}}|a_k|^2\sum_{j_1+\ldots+j_{d'}=n}\prod_{i=1}^{d'}|k_i|^{2j_i}<\infty.$$

In particular $||g||_{L^2}\leq C'||f||_{H^n}$.

Thus, for every[30] $h\in H^n(\mathbb{T}^{d'})$ we have

$$\left|\left|\sum_{j=1}^{J} h\circ R_\alpha^j - J\int_{\mathbb{T}^{d'}} h\,\mathrm{d}\lambda^{d'}\right|\right|_{L^2} \leq C''||h||_{H^n},$$

where $\lambda^{d'}$ is the d'-dimensional Lebesgue measure on $\mathbb{T}^{d'}$. Due to Remark 6(iii), condition (EE) is satisfied with $r=n$ and $\delta=0$.

We can apply Theorem 7 with

$$d>n.$$

[30] If $\int h\neq 0$ consider $f=h-\int h$.

2.11.2 Horocycle flows

Consider the classical horocycle flow h_t on compact homogeneous space $\Gamma\backslash PSL(2,\mathbb{R})$ generated by

$$h_t = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

For fixed $t > 0$ we will consider the time t map $R = h_t$.

Condition $(SLR(y^*))$ follows from the relation $h_{e^{2s}t} = g_s \circ h_t \circ g_{-s}$, where g_s is the geodesic flow

$$g_s = \begin{bmatrix} e^s & 0 \\ 0 & e^{-s} \end{bmatrix}.$$

Indeed, we will show that there is a constant $c > 0$ such that for small enough $r > 0$, $0 < |s| < cr^{-1}$, and $y, y^* \in \Gamma/PSL(2,\mathbb{R})$ with $d(y, y^*) < \frac{c}{2} r^{\frac{1}{2}}$ we have

$$d(h_s y, y^*) \geq \frac{c}{2} r^{\frac{1}{2}}.$$

By the triangle inequality, it is enough to show $d(h_s y, y) \geq cr^{\frac{1}{2}}$. By compactness choose $c = \inf_x d(h_1 x, x) > 0$ (by Hedlund's Theorem there are no periodic orbits). Let $t = \frac{\log(|s|)}{2}$, then

$$d(h_s y, y) = d(g_t h_{sgn(s)} g_{-t} y, g_t g_{-t} y) \geq e^{-|t|} d(h_{sgn(s)} g_{-t} y, g_{-t} y) \geq cr^{\frac{1}{2}}.$$

For small enough r, g_t contracts distances at most by a factor of $e^{-|t|}$. Renaming $r = \frac{c}{2} r^{\frac{1}{2}}$ we obtain (2.4) with $\psi(r) = 2c^3 r^{-2}$.

In order to show effective ergodicity (EE), we combine [FFT16, Corollary 2.8] and [FF03, Theorem 1.5] to conclude that there is a constant $C > 0$ with

$$\left\| \sum_{j=0}^{n-1} f(R^j(y)) - n \int f \, \mathrm{d}\nu \right\| \leq C ||f||_{W^{15}} N^{\frac{5}{6}+\epsilon} \quad \forall f \in W^s, y \in \Gamma/PSL(2,\mathbb{R}), N \geq 1, \tag{2.44}$$

for all $\epsilon > 0$.

Indeed, for $s > 3$, [FF03, Theorem 1.5] yields

$$\begin{aligned} \left| \int_0^T f(h_t(y)) \, \mathrm{d}t - T \int_{\Gamma/PSL(2,\mathbb{R}} \varphi \, \mathrm{d}\nu \right| &\leq C(s) ||f||_{W^s} T^{\frac{1}{2}} \log(T) \\ &\forall f \in W^s, y \in \Gamma/PSL(2,\mathbb{R}), T > 0, \end{aligned} \tag{2.45}$$

for some constant $C(s) > 0$.

A consideration involving twisted integrals as in [FFT16, Corollary 2.8] yields, for $s > 14$,

$$\begin{aligned} \left| \sum_{n=0}^{N-1} f(h_n(y)) - \int_0^N f(h_t(y)) \, \mathrm{d}t \right| &\leq C'(s) ||f||_{W^s} N^{\frac{5}{6}} \log^{\frac{1}{2}}(N) \\ &\forall f \in W^s, y \in \Gamma/PSL(2,\mathbb{R}), N \geq 1, \end{aligned} \tag{2.46}$$

for some constant $C'(s) > 0$. Now (2.45) and (2.46) together imply (2.44).

Now, setting $s = 15$ in (2.46), Theorem 7 applies with

$$d > 6(15 + \frac{5}{3}) = 100.$$

2.11.3 Skew shifts

Let $\alpha \in (0,1) \setminus \mathbb{Q}$ satisfy the Diophantine condition (D) for some $n \geq 2$ and $R : \mathbb{T}^2 \to \mathbb{T}^2$ be given by

$$R(x,y) = (x+\alpha, y+x).$$

Since R has a Diophantine rotation as a factor $(SLR(y^*))$ is satisfied by §2.11.1.

For $k = (k_1, k_2) \in \mathbb{Z}^2$ denote $e_k(x) = e^{2\pi i \langle k,x\rangle}$. Note that

$$\langle e_k, e_{k'} \circ R^j \rangle_{L^2(\mathbb{T}^2)} = \delta^{(k_1,k_2)}_{(k_1'+jk_2', k_2')}.$$

For $f \in H^2(\mathbb{T})$ we can write $f = \sum_{k\in\mathbb{Z}^2} a_k e_k$. If $a_{(k_1,0)} \equiv 0$ (in particular $\int f = 0$) then

$$\sum_{j\geq 1} \left|\langle f, f\circ R^j\rangle_{L^2}\right| = \sum_{j\geq 1} \left| \sum_{k\in\mathbb{Z}^2} a_{(k_1,k_2)} \overline{a_{(k_1+jk_2,k_2)}} \right| \leq \left(\sum_{k\in\mathbb{Z}^2} |a_{(k_1,k_2)}| \right)^2 \leq C||f||^2_{H^2},$$

where $C > 0$ does not depend on f. From this we obtain

$$\begin{aligned} \left|\left| \sum_{j=1}^J f\circ R^j \right|\right|^2_{L^2(\mathbb{T}^2)} &= \sum_{j=1}^J \sum_{j'=1}^J \langle f\circ R^{j'}, f\circ R^j\rangle_{L^2(\mathbb{T}^2)} \\ &\leq \sum_{j=0}^{J-1} (J-j)\langle f, f\circ R^j\rangle_{L^2(\mathbb{T}^2)} \leq CJ||f||^2_{H^2(\mathbb{T}^2)}. \end{aligned}$$

For general $f \in H^n(\mathbb{T}^2)$ with $\int f = 0$, again write $f = \sum_{k\in\mathbb{Z}^2} a_k e_k$ and set

$$f_1 = \sum_{k\in\mathbb{Z}^2, k_2\neq 0} a_k e_k \quad \text{and} \quad f_2 = \sum_{k\in\mathbb{Z}^2, k_2=0} a_k e_k.$$

Applying the above, and the analysis for Diophantine rotations, we find

$$\left|\left| \sum_{j=1}^J f\circ R^j \right|\right|^2_{L^2(\mathbb{T}^2)} \leq (CJ + C')||f||^2_{H^n(\mathbb{T}^2)}.$$

Thus, condition (EE) is satisfied with $\delta = \frac{1}{2}$.

So we can apply Theorem 7 with

$$d > 2(n+1).$$

2.11.4 Example 15

Recall the definition of the Weyl Chamber flow on $\Gamma\backslash SL(d,\mathbb{R})$. Let $d \geq 3$, and Γ be a uniform lattice. Denote by D_+ the subgroup of diagonal elements of $SL(d,\mathbb{R})$ with positive entries. It is easy to see that D_+ is isomorphic to $\mathbb{R}^{d-1}$. D_+ acts on $\Gamma\backslash Sl(d,\mathbb{R})$ by right translation, giving us a $\mathbb{R}^{d-1}$-action. By [BEG20, Theorem 1.1] the action G satisfies (a $\mathbb{R}^{d-1}$ version of) (MEM).

The Diophantine rotation R_α satisfies (EE) and (2.9) by §2.11.1. Hence, we can apply Theorem 9.

2.11.5 Other systems satisfying (EE)

From Example 11 it might seem like (EE) is a very special property and only a few systems satisfy this. The opposite is true, in fact, most classical systems have this property.

To convince ourselves of this, let us give some more examples and point out the mechanisms.

Definition 36. *The system (Y, R, ν) is called mixing of order α if, for each $f, g \in C^{r'}$ with $\int_Y f\, d\nu = \int_Y g\, d\nu = 0$, we have*

$$\left| \int_Y f \circ R^n \cdot g\, d\nu \right| < ||f||_{C^{r'}} ||g||_{C^{r'}} \alpha(n) \quad \forall n \geq 1. \tag{2.47}$$

We say that (Y, R, ν) is polynomially mixing if it is mixing with rate $\alpha(n) = O(n^{-\epsilon})$ for some $\epsilon > 0$.

Lemma 37. *Polynomial mixing implies (EE). More precisely, if (Y, R, ν) is mixing of order $\alpha(n) = O(n^{-\epsilon})$, for some $\epsilon > 0$, then, for all $\epsilon' > 0$, it satisfies property (EE) with*

$$\delta = \begin{cases} \frac{2-\epsilon}{2} & \text{if } \epsilon < 1 \\ \frac{1}{2} + \epsilon' & \text{if } \epsilon = 1 \\ \frac{1}{2} & \text{if } \epsilon > 1. \end{cases} \tag{2.48}$$

Proof. For $f \in C^{r'}$ with $\int_Y f\, \mathrm{d}\nu = 0$ we have, for $N \geq 1$,

$$\begin{aligned} \left\| \sum_{n=0}^{N-1} f \circ R^n \right\|_{L^2}^2 &\leq \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} \left| \int_Y f \circ R^{n_1} f \circ R^{n-2}\, \mathrm{d}\nu \right| \\ &\leq 2N \sum_{n=0}^{N-1} \left| \int_Y f \circ R^n \cdot f\, \mathrm{d}\nu \right| \leq K ||f||_{C^{r'}}^2 N^{2\delta} \end{aligned}$$

for some $K > 0$. □

Remark 38. *In fact, the proof above remains valid if (2.47) holds for all $n \leq N$ except for a subset of $\{1, ..., N\}$ of size $N^{1-\bar{\epsilon}}$ for some $\bar{\epsilon} > 0$. We call such systems* polynomially weakly mixing.[31]

Many classical systems exhibit polynomial (or faster) mixing we list just a few examples referring to [DDKN22b, Section 8] for a more comprehensive list

- mixing piecewise expanding interval maps [BSTV03, Theorem 3.1] as well as expanding interval maps with critical points and singularities [LM13, Theorem 1.5],
- uniformly hyperbolic systems [Liv95, Theorem 3.9],
- some quadratic maps [You92, Theorem 3],
- noncompact translations on finite volume homogeneous spaces of semisimple Lie groups without compact factors [KM96, §2.4.3],
- time change of horocycle flow [FU12, Theorem 3].

For parabolic and elliptic systems, one can often use a harmonic analytic argument akin to (but more involved than) §2.11.1 or 2.11.3. Other concrete examples include

[31] In fact, a slight modification of the proof shows that if R_1 is polynomially mixing and R_2 satisfies (EE) then $R_1 \times R_2$ satisfies (EE).

- nilflows [FF06, Theorem 1.1],
- almost every interval exchange transformation [AFS23, Theorem 7.1],
- time 1 map of certain smooth surface flows, this follows from a work in progress by the author, where polynomial weak mixing is shown.

2.12 Robustness of return times

Lastly, we mention the proof for the delayed PLT. All of the above proofs can be done using $\Phi_{B_l,\alpha}$ instead of Φ_{B_l}, this shows (with the notation from the proof of Theorem 7)

$$(\mu\times\nu)\left(\bigcup_{k=1}^{K} A_l^{(k)}\times B_l^{(k)}\right)\Phi_{\bigcup_{k=1}^{K} A_l^{(k)}\times B_l^{(k)},\alpha}\overset{\mu}{\Rightarrow}\Phi_{Exp}.$$

To conclude, we only need a version of the approximation Lemma 21 for delayed return times.

Let (M, d_M) be a compact metric space, let $(\tilde{\vartheta}_n)_{n\geq 1}$ be a sequence of Lipschitz functions on M dense in $C(M)$, and denote $\vartheta_n=\frac{\tilde{\vartheta}_n}{||\tilde{\vartheta}_n||_{Lip}}$. The metric

$$D_M(\lambda,\lambda')=\sum_{n\geq 1}2^{-n}\left|\int_M\vartheta_n\,\mathrm{d}\lambda'-\int_M\vartheta_n\,\mathrm{d}\lambda\right|,$$

for probability measures λ and λ', models distributional convergence[32].

Lemma 39. *Let (X,μ,T) be a probability-preserving dynamical system, α be a sequence of natural numbers, $(A_l)_{l\geq 1}$ be a sequence of rare events, and $\Phi=(\phi^{(1)},\phi^{(2)},...)$ be a random process in $[0,\infty)$. Assume that, for each $\delta>0$, there is a sequence of rare events $(A_l^{(\delta)})_{l\geq 1}$ with $A_l^{(\delta)}\subset A_l$ and $\mu_{A_l}(A_l\setminus A_l^{(\delta)})<\delta$ such that*

$$\mu(A_l^{(\delta)})\Phi_{A_l^{(\delta)},\alpha}\overset{\mu}{\Rightarrow}\Phi\quad;\textit{as } l\to\infty,\forall\delta>0.$$

Then

$$\mu(A_l)\Phi_{A_l,\alpha}\overset{\mu}{\Rightarrow}\Phi\quad\textit{as } l\to\infty.$$

Proof. Taking a subsequence if necessary, we may assume that there is a $[0,\infty]$-valued random process Φ' with

$$\mu(A_l)\Phi_{A_l,\alpha}\overset{\mu}{\Rightarrow}\Phi'\quad\text{as } l\to\infty.$$

For $s,t\in[0,\infty]$ denote $d_{[0,\infty]}(s,t)=|e^{-s}-e^{-t}|$, where by convention $e^{-\infty}=0$, then $([0,\infty],d_{[0,\infty]})$ is a compact metric space. Also, the infinite product $([0,\infty]^{\mathbb{N}},d_{[0,\infty]^{\mathbb{N}}})$ is a compact metric space with $diam([0,\infty]^{\mathbb{N}})=1$, where

$$d_{[0,\infty]^{\mathbb{N}}}((s_j),(t_j))=\sum_{j\geq 1}2^{-j}d_{[0,\infty]}(s_j,t_j).$$

We claim that for every $\epsilon>0$ there exist $\delta_0>0$ and an $L\geq 1$ such that

$$D_{[0,\infty]^{\mathbb{N}}}\Big(law_\mu(\mu(A_l^{(\delta)})\Phi_{A_l^{(\delta)},\alpha}),law_\mu(\mu(A_l)\Phi_{A_l,\alpha})\Big)<5\epsilon\quad\forall l\geq L.\tag{2.49}$$

Then taking $l\to\infty$ shows $D_{[0,\infty]^{\mathbb{N}}}(\Phi,\Phi')<5\epsilon$ and the conclusion follows by $\epsilon\to 0$.

[32] In the sense that $\lambda_n\Rightarrow\lambda$ if and only if $D_M(\lambda,\lambda_n)\to 0$.

Let $1 > \epsilon > 0$. First, note that[33]

$$D_{[0,\infty]^{\mathbb{N}}}\left(law_\mu(\mu(A_l^{(\delta)})\Phi_{A_l^{(\delta)},\alpha}), law_\mu(\mu(A_l)\Phi_{A_l^{(\delta)},\alpha})\right) < \delta,$$

so it is enough to show that there exist $\epsilon > \delta_0 > 0$ and an $L \geq 1$ such that

$$D_{[0,\infty]^{\mathbb{N}}}\left(law_\mu(\mu(A_l)\Phi_{A_l^{(\delta)},\alpha}), law_\mu(\mu(A_l)\Phi_{A_l,\alpha})\right) < 4\epsilon \quad \forall l \geq L. \tag{2.50}$$

Denote $\Phi_{A_l,\alpha} = (\varphi^{(1)}_{A_l,\alpha}, \varphi^{(2)}_{A_l,\alpha}, ...)$ and $\Phi_{A_l^{(\delta)},\alpha} = (\varphi^{(1)}_{A_l^{(\delta)},\alpha}, \varphi^{(2)}_{A_l^{(\delta)},\alpha}, ...)$. Now choose $J \geq 1$ so big that $\sum_{j\geq J} 2^{-j} < \epsilon$, and $T > 0$ such that

$$\mathbb{P}\left(\sum_{j=1}^{J} \phi^{(j)} > T\right) < \epsilon.$$

For $\delta = \min\left(\frac{\epsilon}{2}, \frac{\epsilon}{2T}\right)$ choose $L \geq 1$ so big that

$$\mu\left(\sum_{j=1}^{J} \mu(A_l^{(\delta)})\varphi^{(j)}_{A_l^{(\delta)},\alpha} > T\right) < 2\epsilon \quad \forall l \geq L.$$

Since $\sum_{j=1}^{J} \varphi^{(j)}_{A_l^{(\delta)},\alpha} > \sum_{j=1}^{J} \varphi^{(j)}_{A_l,\alpha}$ and $\mu(A_l^{(\delta)}) > (1-\delta)\mu(A_l) > \frac{1}{2}\mu(A_l)$, in particular

$$\mu\left(\sum_{j=1}^{J} \mu(A_l)\varphi^{(j)}_{A_l,\alpha} > 2T\right) < 2\epsilon \quad \forall l \geq L.$$

Now, for $j = 1, ..., J$, we have

$$\begin{aligned}
&\mu\left(\varphi^{(j)}_{A_l,\alpha} \neq \varphi^{(j)}_{A_l^{(\delta)},\alpha}\right) \leq \mu\left(\varphi^{(J)}_{A_l,\alpha} \neq \varphi^{(J)}_{A_l^{(\delta)},\alpha}\right) \\
&\leq \mu\left(\bigcup_{i=1}^{\left\lfloor \frac{2T}{\mu(A_l)} \right\rfloor} T^{-\tilde{\alpha}^{(i)}}(A_l \setminus A_l^{(\delta)})\right) + \mu\left(\sum_{j=1}^{J} \mu(A_l)\varphi^{(j)}_{A_l,\alpha} > 2T\right) \\
&\leq \frac{2T}{\mu(A_l)}\mu(A_l)\delta + 2\epsilon \leq 3\epsilon.
\end{aligned}$$

Thus

$$\begin{aligned}
&D_{[0,\infty]^{\mathbb{N}}}\left(law_\mu(\mu(A_l)\Phi_{A_l^{(\delta)},\alpha}), law_\mu(\mu(A_l)\Phi_{A_l,\alpha})\right) \\
&= \sum_{j\geq 1} 2^{-j} \int_X d_{[0,\infty]}(\mu(A_l)\varphi^{(j)}_{A_l,\alpha}, \mu(A_l)\varphi^{(j)}_{A_l^{(\delta)},\alpha})\, \mathrm{d}\mu \leq 3\epsilon \sum_{j=1}^{J} 2^{-j} + \epsilon \leq 4\epsilon
\end{aligned}$$

proving (2.50). □

To conclude, we give a

[33]For $k \geq 1$ and $s, t \in [0,\infty]$ we have $d_{[0,\infty]}(ks, kt) \geq d_{[0,\infty]}(s,t) \leq |s-t|$.

Proof of Lemma 21. Denote $\lambda = \mu \times \nu$ and $M = [0, \infty]$, with $D_{M^{\mathbb{N}}}$ as in the proof of Lemma 39 we have

$$\begin{aligned}
&D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_{Q'}}(\lambda(Q')\Phi_{Q'})\\
&\qquad \leq \lambda_Q(Q')D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_Q}(\lambda(Q')\Phi_{Q'})\\
&\qquad\quad + \lambda_Q(Q \setminus Q')D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_{Q'}}(\lambda(Q')\Phi_{Q'})\\
&\qquad \leq D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_Q}(\lambda(Q')\Phi_{Q'}) + \lambda_Q(Q \setminus Q').
\end{aligned}$$

Furthermore

$$\begin{aligned}
&D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_Q}(\lambda(Q')\Phi_{Q'})\\
&\qquad \leq \sum_{j\geq 0} 2^{-j-1} \int_Q d_M(\lambda(Q)\varphi_Q \circ S_Q^j, \lambda(Q')\varphi_{Q'} \circ S_{Q'}^j)\,\mathrm{d}\lambda_Q\\
&\qquad \leq \sum_{j\geq 0} 2^{-j-1} \int_Q d_M(\lambda(Q)\varphi_Q \circ S_Q^j, \lambda(Q)\varphi_{Q'} \circ S_{Q'}^j)\,\mathrm{d}\lambda_Q + \lambda_Q(Q \setminus Q').
\end{aligned}$$

For each $j \geq 0$ we have

$$\lambda_Q(\varphi_Q \circ S_Q^j \neq \varphi_{Q'} \circ S_{Q'}^j) \leq \lambda_Q\left(\bigcup_{i=0}^{j} S_Q^i(Q \setminus Q')\right) \leq (j+1)\lambda_Q(Q \setminus Q').$$

The claim follows since $\sum_{j\geq 1} j2^{-j} = 2$. $\square$

Chapter 3

Trimmed ergodic sums for non-integrable functions with power singularities over irrational rotations

Studying Birkhoff sums of non-integrable functions involves the challenge of large observations depending on the sampled orbit, which prevents pointwise limit theorems. To address this issue, the largest observations are removed, this process is commonly known as trimming. While this method is well studied for independent identically distributed sequences and systems with strong mixing behaviour, this paper examines regular dynamical systems with polynomial deviations of ergodic averages. Focusing on irrational rotations of $\mathbb{T}$, we establish trimmed weak and strong laws for the functions $\frac{1}{x}$ and $\frac{1}{x^\beta}$ with $\beta > 1$, providing explicit conditions on the rotation angle.

3.1 Introduction

3.1.1 Motivation

Let (X, T, μ) be a probability preserving ergodic dynamical system. The much celebrated Birkhoff Ergodic Theorem states that, given an integrable function $f \in L^1$, the average of observations over a long time approximates the integral, more formally

$$\lim_{N\to\infty} \frac{S_N(f)}{N} = \int f \,\mathrm{d}\mu \quad \text{almost surely,}$$

where $S_N(f) = \sum_{n=0}^{N-1} f \circ T^n$. In this paper, we will investigate the question

$$\text{How do we describe asymptotics of } S_N(f) \text{ if } \int |f| \,\mathrm{d}\mu = \infty?$$

For simplicity, we shall restrict ourselves to functions that are non-negative and finite almost everywhere, denote the collection of these functions by

$$\mathcal{F} = \{f : X \to [0, \infty] \text{ measurable} \mid f \notin L^1(\mu), \text{ but } \mu(f = \infty) = 0\}.$$

Often, in the situation when X is a metric space, we will also require that $f \in \mathcal{F}$ is continuous[1].

[1] In this case we endow $[0, \infty]$ with the metric $d(s,t) = |e^{-s} - e^{-t}|$, for $s, t \in [0, \infty]$, making it a compact metric space. Note that f is continuous if and only if, for every $s \in [0, \infty)$, the sets $\{f < s\}$ and $\{f > s\}$ are open.

The most naive way to approach this problem would be to normalize by a factor different from N, say $d_N > 0$. However, Aaronson's "anti-convergence" result ([Aar77]) asserts that this is impossible: for every normalising sequence $(d_N)_{N\geq 1}$ it holds that

$$\limsup_{N\to\infty} \frac{S_N(f)}{d_N} = \infty \quad \text{or} \quad \liminf_{N\to\infty} \frac{S_N(f)}{d_N} = 0 \quad \text{almost surely,}$$

meaning that we either underestimate or overestimate the sum infinitely often. This non-convergence is caused by large observations, meaning n such that $f(T^n x)$ is big, which may or may not occur depending on x. To circumvent this problem, a common method is to exclude the largest observations altogether. This method is referred to as trimming.

More formally, for $N \geq 1$ and $0 \leq k(N) \leq N$, define the *trimmed sum* as

$$S_N^{k(N)}(f)(x) = \sum_{n=0}^{N-1} f(T^n x) - \max_{0\leq n_1 < n_2 < \ldots < n_{k(N)} \leq N-1} \sum_{j=1}^{k(N)} f(T^{n_j} x).$$

We will also refer to $k(N)$ as the *trimming sequence*. For simplicity set $S_N^M(f) = 0$ if $M > N$.

The central problem we study is the following;

Problem A. *Do there exist positive $k(N) = o(N)$ and $d_N > 0$ such that*

$$\lim_{N\to\infty} \frac{S_N^{k(N)}(f)(x)}{d_N} = 1 \quad \text{for } \mu\text{-a.e. } x \in X? \tag{3.1}$$

If (3.1) holds, we refer to it as a trimmed strong law.

Remark 40. *Note that $S_N^N(f) = 0$ by definition. Furthermore, heuristically speaking, if $c \in (0,1)$, f is continuous and $k(N) \sim cN$ (i.e. $\lim_{N\to\infty} k(N)/(cN) = 1$) then*

$$S_N^{k(N)}(f) \sim S_N(\min(K, f)) \text{ almost surely,}$$

where $K > 0$ is such that $\mu(f > K) = c$.

This explains why we choose $k(N) = o(N)$ in Problem A; if $k(N) \sim cN$ then we lose information on f, making $S_N^{k(N)}(f)$ uninteresting.

3.1.2 Background

Ideally, one would hope for[2] $k(N) = constant$, this is referred to as *light trimming*. On the other hand, situations where $k(N) \to \infty$, but $k(N) = o(N)$ as $N \to \infty$ are referred to as *intermediate trimming*.

Trimming is well-studied in the set-up of identically distributed random variables[3] (iids). The literature is vast, for brevity's sake, we will only mention situations where explicitly strong laws (or results having immediate consequences for strong laws) are studied. Trimming for iids has been used in other contexts, for example, weak laws, laws of iterated logarithm, CLTs, and more.

Mori ([Mor76, Mor77]) developed general criteria for lightly trimmed strong laws to hold, his results were later generalized by Kesten and Maller ([KM92, KM95, Mal84]).

[2] By the above result by Aaronson, if $\int |f| \, d\mu = \infty$, which is the case that we are concerned with, then strong laws with $k(N) = 0$ are not possible.

[3] Meaning that $f \circ T^n = X_n$ is an independent and identically distributed random process.

From Mori's works ([Mor76] gives necessary and sufficient conditions for certain lightly trimmed laws to hold) it already becomes apparent that light trimming is not always sufficient. A fortiori Kesten ([Kes93]) showed later that even a lightly trimmed weak law already necessitates an untrimmed weak law (which might not hold).

A special case, that is of interest, is when[4] X has regularly varying tails with index $\alpha \in (0,1)$. This means that there is a slowly varying function[5] such that $\mathbb{P}(X > t) = t^{-\alpha} L(t)$. Note that in this situation[6] there is no untrimmed weak law, and hence, by Kesten's result, also no lightly trimmed weak (or strong) law. In this case, a law of the iterated logarithm under trimming was obtained by Haeusler and Mason ([HM87]). Their result was shown to be optimal by Haeusler ([Hae93]). From those results, a strong law under intermediate trimming can be deduced.

In the context of dynamical systems, so far results have been shown only for systems exhibiting strong mixing properties. Some known results are:

- If T is the Gauss map and $f(x) = \frac{1}{x}$ (in the case S_N is essentially the sum of coefficients in the CFE[7] of x) then it was proved[8] by Diamond and Vaaler ([DV86]) that $\frac{S_N^1(f)}{N \log(N)} \to \frac{1}{\log(2)}$.
- Even though not formulated in terms of dynamical systems, Aaronson and Nakada ([AN03]) showed lightly trimmed strong laws (extending Mori's results) for ψ-mixing random variables [9] with speed $\sum_{n=1}^{\infty} \frac{\psi(n)}{n} < \infty$. For Example, Gibbs-Markov maps, together with a function that is measurable for the partition, are exponentially ψ-mixing (this follows from [AD01]). Their results apply (among others) if $\mathbb{P}(X > t) = \frac{1}{t}$ for $t > 0$. In terms of functions, it applies to $f = \frac{1}{x}$ (with Lebesgue measure).
- Haynes ([Hay14]) later showed a quantitative version of the above. Under the same assumptions he showed a strong law for

  $$S_N(f)(x) - \delta(N,x) \max_{n=0,\ldots,N-1} f(T^n x) \quad \text{for some} \quad \delta(N,x) \in \{0,1\},$$

  where $(f(T^n x))$ is a ψ-mixing process with a certain speed, and provided explicit error terms. We remark that, in contrast to Aaronson and Nakada's result, the number of trimmed terms is allowed to depend on x.
- Kesseböhmer and Schindler ([KS19]) studied intermediate trimming for maps satisfying a spectral gap condition. This method applies to some expanding interval maps (more general than Gibbs-Markov). Their results apply to functions like $f(x) = \frac{1}{x^\beta}$ with $\beta > 1$ (with Lebesgue measure).
- Schindler recently proved ([Sch18]) a strong law for the doubling map and the function $f = \frac{1}{x}$. In this situation[10] there is no lightly trimmed strong law as already noted by Haynes ([Hay14]). Instead, an intermediately trimmed strong law strong law is obtained.

[4] Here X is a random variable having the same distribution as f.

[5] A function $L : [0,\infty) \to \mathbb{R}$ is called *slowly varying* if, for every $c > 0$, it holds that $\lim_{x\to\infty} \frac{L(cx)}{L(x)} = 1$.

[6] As can be deduced from [Fel57, VII.7 Theorem 2].

[7] The continued fraction expansion (CFE) shall be defined later in §3.3.

[8] They study a situation where the trimming sequence $k(N)$ is allowed to depend on x. More explicitly, they show a strong law for

$$S_N(f)(x) - \theta(N,x) \max_{n=0,\ldots,N-1} f(T^n x) \quad \text{for some} \quad \theta(N,x) \in [0,1].$$

However a statement about S_N^1 can be easily deduced from their proof.

[9] For a precise definition of various mixing properties, see [Bra05].

[10] This highlights the influence of periodic points; the function $\frac{1}{x}$ has its singularity at the fixed point 0.

3.1.3 General results

Our first result answers Problem A. We show that it is always possible to find a good trimming sequence $k(N) = o(N)$. The following result seems to be known to the experts, for the convenience of the reader we shall provide a proof in §3.4.

Theorem 41. *Let (X, T, μ) be a probability preserving ergodic system. Then, for $f \in \mathcal{F}$ there exist $k(N) \in \mathbb{N}$ with $k(N) = o(N)$ and $d_N > 0$ such that*

$$\lim_{N\to\infty} \frac{S_N^{k(N)}(f)(x)}{d_N} = 1 \quad \text{for } \mu\text{-a.e. } x \in X. \tag{3.2}$$

Furthermore, if (X, T, μ) is uniquely ergodic (and μ is regular) and f is continuous, then $k(N)$ and d_N can be chosen such that convergence in (3.2) *holds uniformly*[11] *at all $x \in X$.*

Note that Theorem 41 does only guarantee the existence of some trimming sequence $k(N) = o(N)$, there is no control whatsoever about how fast $\frac{k(N)}{N}$ decays. In fact, as shall be made more precise later in Remark 50, for the function $\frac{1}{x}$ over a Liouvillean rotation, $\frac{k(N)}{N}$ might decay arbitrarily slowly.

Heuristically speaking it is most desirable to choose $k(N)$ as small as we can; the smaller $k(N)$ is, the more information about f is retained after trimming $k(N)$ terms. Therefore, Problem A should be reformulated as;

Problem B. *What is the "smallest" $k(N)$ such that there are d_N for which* (3.2) *holds? Furthermore, give an explicit formula for d_N.*

This seems to suggest that if a trimmed strong law holds, then it will continue to hold if we trim more terms. However, this is not true, on the contrary;

Theorem 42. *Let (X, T, μ) be an aperiodic*[12] *probability preserving ergodic system. Then there is a function $f \in \mathcal{F}$ and a trimming sequence $k(N) = o(N)$ such that there are $d_N > 0$ so that* (3.2) *holds, but there is another trimming sequence $k'(N) = o(N)$ with $k'(N) \geq k(N)$ such that for all normalising $d'_N > 0$ there is an $\epsilon > 0$ such that*

$$\limsup_{N\to\infty} \mu\left(\left|\frac{S_N^{k'(N)}(f)}{d'_N} - 1\right| > \epsilon\right) > 0.$$

In the following, we will attempt to answer Problem B in two settings:

- If T has polynomial deviations of ergodic averages (as defined below). This will be done in Section 3.2.
- If T is an irrational rotation. This will be done in Section 3.3.

3.2 Trimming for functions with power singularities

For this section, let $X = M$ be a compact C^r-manifold, for some $r \geq 1$, with $\dim(M) = d$, and T be a transformation preserving a smooth probability measure μ with density $\frac{\mathrm{d}\mu}{\mathrm{dvol}} = \rho$ on M. We shall investigate (trimmed) ergodic sums of non-L^1 functions with power singularities.

[11]For definiteness, whenever we talk about uniform convergence we replace f by $\hat{f}$ with $\hat{f} = f$ on $\{f < \infty\}$ and $\hat{f} = 0$ otherwise.

[12]For a periodic ergodic system, i.e. a rotation on finitely many elements, the statement of the Theorem is trivial because every finite function is integrable.

Definition 43. *Let $x^* \in M$, and $f : M \setminus \{x^*\} \to [0, \infty)$ be a non-negative function on M. We say that f has a power singularity at x^* if $f \in C^r(M \setminus \{x^*\})$ and there are $c, \beta > 0$ such that*

$$\lim_{x \to x^*} f(x) d(x, x^*)^\beta = c,$$

where $d(\cdot, \cdot)$ denotes the geodesic distance. In this case we will say that $c = Res_{x^*}(f)$ is the residue of f at x^*, and $\beta = Ord_{x^*}(f)$ is the order of the singularity.

Note that, in this case, assuming $\rho(x^*) > 0$, f is non-integrable if and only if $\mathrm{Ord}_{x^*}(f) \geq d$. We will further distinguish the cases $\mathrm{Ord}_{x^*}(f) = d$ and $\mathrm{Ord}_{x^*}(f) > d$. Integrating we obtain

- for $\beta > d$ it holds that

$$\int_{\epsilon < |\mathbf{x}|_2 < 1} |\mathbf{x}|_2^{-\beta} \, \mathrm{d}\mathbf{x} = \frac{1}{\beta - d} B_d \epsilon^{d-\beta} + O(1) \quad \text{as } \epsilon \to 0,$$

- whereas

$$\int_{\epsilon < |\mathbf{x}|_2 < 1} |\mathbf{x}|_2^{-d} \, \mathrm{d}\mathbf{x} = B_d |\log(\epsilon)| \quad \forall \epsilon \in (0, 1),$$

where B_d is the volume of the unit ball in d dimensions. So let $f \in C^r(m \setminus \{x^*\})$ be a function with only one pole at x^*, where $\rho(x^*) > 0$, it holds that[13]

- if $\mathrm{Ord}_{x^*}(f) = \beta > d$ then

$$\int_{M \setminus B_\epsilon(x^*)} f \, \mathrm{d}\mu = \frac{1}{\beta - d} \mathrm{Res}_{x^*}(f) B_d \rho(x^*) \epsilon^{d-\beta} + O(\epsilon^{d-\beta+1}) \quad \text{as } \epsilon \to \infty, \tag{3.3}$$

- and, if $\mathrm{Ord}_{x^*}(f) = d$, then

$$\int_{M \setminus B_\epsilon(x^*)} f \, \mathrm{d}\mu = \mathrm{Res}_{x^*}(f) B_d \rho(x^*) |\log(\epsilon)| + O(\epsilon |\log(\epsilon)|) \quad \forall \epsilon > 0.$$

Using the polynomial deviations as assumed will lead to cumbersome formulas. We introduce the following simplifying notation.

For $A, B \in \mathbb{R}$ we write

$$A \ll B \quad \iff \quad \exists C > 0 \text{ such that } A \leq CB,$$

the constant $C > 0$. We write $A \ll_{f,g,h} B$ if the implicit constant is allowed to depend on the quantities f, g, h. We write $A \asymp B$ if $A \ll B$ and $B \ll A$.

To obtain a trimming, we will assume *polynomial deviations of ergodic averages* for smooth observables either pointwise or in L^2. More explicitly, we will assume either of the following two

(A) there is a[14] $r \geq 1$ and a $\delta > 0$ such that

$$||S_N(g) - N \int g \, \mathrm{d}\mu||_{L^2} \ll N^{1-\delta} ||g||_{C^r} \quad \forall N \geq 1, g \in C^r, \tag{3.4}$$

[13] The extra error term comes from considering the density ρ. For $\epsilon > 0$ sufficiently small, it holds that $|\mu(B_\epsilon(x^*)) - B_d \rho(x^*) \epsilon^d| \leq \Omega \epsilon^{d+1}$, for some constant $\Omega > 0$ depending on x^*.

[14] Here the C^r norm of a function can be defined by covering M by finitely many charts and taking the maximum of all local C^r norms.

(B) or there is a $r \geq 1$ and a $\delta > 0$ such that

$$|S_N(g)(x) - N \int g \, \mathrm{d}\mu| \ll N^{1-\delta} ||g||_{C^r} \quad \forall N \geq 1, g \in C^r, x \in M. \tag{3.5}$$

Remark 44. *Since M is compact, we can define the C^r norm of a function by covering M with finitely many charts and taking the maximum of the local C^r norms in these charts.*

In both cases, we will show a trimming result for $k(N) = \lceil N^{1-\epsilon} \rceil$ with $\epsilon > 0$. In case (B), the convergence is uniform.

Theorem 45. *Assume (A) holds. Let $x^* \in M$ be a point on M with $\rho(x^*) > 0$ and let $f \geq 0$, $f \in C^r(M \setminus \{x^*\}$ be a non-negative function with a power singularity only at x^* of order $Ord_{x^*}(f) = \beta \geq d$. If $k(N) = \lceil N^\alpha \rceil$ for some*

$$1 > \alpha > \alpha_0 = \max \left(\frac{r + d - d\delta}{r + d}, \frac{2rd + d - r - d\delta}{2rd + d - r}, \frac{r + 1}{r + d + 1} \right)$$

then

- *if $\beta = d$ then*

$$\frac{S_N^{k(N)}(f)(x)}{N \log(N)} \to \frac{1 - \alpha}{d} Res_{x^*}(f) B_d \rho(x^*) \quad \textit{as } n \to \infty, \textit{ for } \mu - a.e\ x \in M, \tag{3.6}$$

- *and if $\beta > d$ then*

$$\frac{S_N^{k(N)}(f)(x)}{N^{\frac{\beta}{d}} k(N)^{1-\frac{\beta}{d}}} \to \frac{1}{\beta - d} Res_{x^*}(f) B_d^{\frac{\beta}{d}} \rho(x^*)^{\frac{\beta}{d}} \quad \textit{as } n \to \infty, \textit{ for } \mu - a.e\ x \in M. \tag{3.7}$$

If (B) holds, then the convergence in (3.6) *resp.* (3.7) *is uniform in x.*

Remark 46. *In Theorem 45, and the Theorems below, an explicit bound on the error term can be easily computed by following the estimates in the proof carefully. In an attempt to keep this exposition simple, we will only state the qualitative results without error terms.*

Remark 47. *For most systems the conclusion of Theorem 45 is far from optimal. As can be seen from the table in §3.3.4, for systems with strong enough mixing properties we expect a strong law to hold with any trimming sequence growing faster than* $\log(\log(N))$*. Moreover, for almost all rotations, as trimming by* $\log^{1+\epsilon}(N)$*, for any $\epsilon > 0$, is sufficient.*

However, it is possible to construct rotations that satisfy assumption (B), where there is no better trimming than polynomial. More precisely, in light of Theorem 56 and Lemma 77, for a rotation by an irrational angle α, with continued fraction approximants $\frac{p_n}{q_n}$, satisfying $q_n^2 < q_{n+1} < q_n^3$, and the functions $\frac{1}{x}$ or $\frac{1}{x^\beta}$ with $\beta > 1$, even weak laws can only hold after trimming polynomially many terms.

The significance of Theorem 45 is that it allows us, given a function $f \in C^r(M \setminus \{x^*\})$ with one pole at some $x^* \in M$, to effectively determine both the order and the residue of the pole by sampling along an orbit of a suitable dynamical system.

Having these results, one might ask what happens if there are finitely many poles, say f has poles at $x_1, ..., x_p$ with $\mathrm{Ord}_{x_i}(f) = \beta_i \geq d$ for $i = 1, ..., p$. One might expect to obtain an asymptotic expansion of the form

$$S_N^{k(N)}(f)(x) = \sum_{\mathrm{Ord}_{x_i}(f)>d} \frac{1}{\beta_i - d}\mathrm{Res}_{x_i}(f) B_d^{\frac{\beta_i}{d}} \rho(x_i)^{\frac{\beta_i}{d}} N^{\frac{\beta_i}{d}} k(N)^{1-\frac{\beta_i}{d}}$$
$$+ \sum_{\mathrm{Ord}_{x_i}(f)=d} N\log(N) d(1-\alpha)\mathrm{Res}_{x_i}(f) B_d \rho(x_i) + O(N),$$

with $k(N)$ as in Theorem 45, using similar techniques. However, this is not true, unless all of the orders are the same.

Following the proof carefully, we see that in Lemma 60 an error term appears that is of the order $o(N^{\frac{\beta}{d}} k(N)^{1-\frac{\beta}{d}})$, but not $o(N^{\frac{\beta'}{d}} k(N)^{1-\frac{\beta'}{d}})$ for $\beta' < \beta$, where $\mathrm{Res}_{x^*}(f) = \beta = \max_{i=1,\dots,p} \beta_i > d$. This error term comes from the contribution of the $f(T^j(x))$ close but not too close to[15] x^*, and can, using our methods, not be avoided.

On the other hand, if all of the orders are the same, say $\beta_1 = \dots = \beta_p = \beta$, and $\sum_{i=1}^p \mathrm{Res}_{x_i}(f) \neq 0$, then the corresponding limit theorems do hold. And we have[16]

- if $\beta = d$ then
$$\frac{S_N^{k(N)}(f)(x)}{N\log(N)} \to d(1-\alpha)\sum_{i=1}^p \mathrm{Res}_{x_i}(f) B_d \rho(x_i) \quad \text{as } n\to\infty, \text{ for } \mu - \text{a.e } x \in M,$$

- and if $\beta > d$ then
$$\frac{S_N^{k(N)}(f)(x)}{N^{\frac{\beta}{d}} k(N)^{1-\frac{\beta}{d}}} \to \frac{1}{\beta - d}\sum_{i=1}^p \mathrm{Res}_{x_i}(f) B_d^{\frac{\beta}{d}} \rho(x_i)^{\frac{\beta}{d}} \quad \text{as } n\to\infty, \text{ for } \mu - \text{a.e } x \in M.$$

3.3 Trimming for rotations

3.3.1 Background

For irrational $\alpha \in (0,1)$ we consider the rotation $R(x) = R_\alpha(x) = x + \alpha \pmod 1$, $x \in [0,1)$ (we identify $\mathbb{T}$ with $[0,1)$ by fixing 0). First, we shall review some basic facts.

There is a unique[17] representation

$$\alpha = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\dots}}}, \tag{3.8}$$

with $a_j \in \mathbb{N}$ for $j \geq 1$, called the *continued fraction expansion* (CFE), we will sometimes also denote (3.8) by $\alpha = [a_1, a_2, \dots]$. In this case $a_1, a_2, \dots$ shall be referred to as the CFE coefficients.

If $\alpha = [a_1, a_2, \dots]$, then α is well approximated by finite iterates of the expansion, i.e. by the rational numbers

$$\frac{p_n}{q_n} = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \dots + \cfrac{1}{a_{n-2} + \cfrac{1}{a_{n-1}}}}}, \tag{3.9}$$

[15] With the notation of Lemma 60, near the boundary of $B_{\epsilon(N)}(x^*)$.

[16] Which can be proven using the same techniques as in Theorem 45.

[17] For rational α the CFE also exists, but is not unique. There are always two representations, namely $\alpha = [a_1, a_2, \dots, a_n]$ and $\alpha = [a_1, a_2, \dots, a_n - 1, 1]$.

with the convention $p_1 = 0$, $q_1 = 1$. The rational numbers $\frac{p_n}{q_n}$ shall be referred to as the CFE approximants. They are the best rational approximations to α in the sense that

$$\left|\alpha - \frac{p_n}{q_n}\right| < \left|\alpha - \frac{p}{q}\right| \quad \forall q \le q_n, p \in \mathbb{N}, (p,q) \neq (p_n, q_n),$$

furthermore, we have

$$\frac{1}{q_{n+1}(q_{n+1}+q_n)} < \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_{n+1}^2}.$$

From (3.9), it is easy to see that p_n and q_n satisfy the relation

$$q_{n+1} = a_n q_n + q_{n-1} \text{ and } p_{n+1} = a_n p_n + p_{n-1} \quad n \ge 2.$$

These basic facts can be found in any book on Diophantine approximations.

For $N \ge 1$ and n such that $N \in [q_n, q_{n+1} - 1]$ we can uniquely expand N as

$$N = \sum_{j=1}^{n} b_j q_j, \tag{3.10}$$

with $b_j \in \mathbb{N}$ and $b_j \le a_j \le \frac{q_{j+1}}{q_j}$, the expansion in (3.10) is unique if we require that, for each $J = 1, ..., n-1$, we have[18] $\sum_{j=1}^{J} b_j q_j < q_{J+1}$. This is called the Ostrowski expansion, note that $b_n = \left\lfloor \frac{N}{q_n} \right\rfloor$. In an attempt to keep the notation simple we will, in the following, always implicitly assume that $N \in [q_n, q_{n+1}-1]$ and (3.10) holds, thus suppressing the dependence of n and b_j on N.

We will often make use of the Denjoy-Koksma inequality; for $f \in BV$ and $n \ge 1$ it holds that

$$\left|S_{q_n}(f)(x) - q_n \int_0^1 f \, \mathrm{d}\lambda\right| \le \mathrm{Var}(f) \quad \forall x \in [0,1),$$

where λ denotes the Lebesgue measure on $[0,1)$ and $\mathrm{Var}(f)$ denotes the total variation of f. In light of the Ostrowski expansion (3.10), it follows that

$$\left|S_N(f)(x) - N \int_0^1 f \, \mathrm{d}\lambda\right| \le \sum_{j=1}^{n} b_j \mathrm{Var}(f) \quad \forall x \in [0,1).$$

For simplicity's sake, in this section, we will focus on the functions

$$f(x) = x^{-\beta}, \ \beta \ge 1.$$

Ergodic sums of $x^{-\beta}$ - or the "zero average" version $x^{-\beta} - (1-x)^{-\beta}$ - have already been studied without trimming, among others in [DF15], [SU08]. However, there the rotation number α is also randomised and one obtains a distributional limit, to a non-constant distribution, rather than convergence to a constant. For example [SU08, Theorem 2] shows; For $f(x) = \frac{1}{x} - \frac{1}{1-x}$ there is a distribution $\mathcal{D}$ on $\mathbb{R}$ such that

$$\lambda^2\left((\alpha, x) \;\middle|\; a \le \frac{1}{N} S_N(f)(x) \le b\right) \to \mathcal{D}[a,b] \quad \text{as } N \to \infty \quad \forall a, b \in \mathbb{R},$$

where λ^2 denotes the Lebesgue measure on $\mathbb{T}^2$. In [DF15] an analogous result is obtained for $f(x) = x^{-\beta}$ with $\beta > 1$. By linearity and symmetry their result can be generalised to

[18]Equivalently $b_1 \neq a_1$ and there is no $j \in [2, n]$ with $b_j = a_j$ and $b_{j-1} \ge 1$.

$f(x) = c_1 x^{-\beta} + c_2(1-x)^{-\beta}$ for any $c_1, c_2 \in \mathbb{R}$. Usually, the case $c_1 = -c_2$ is referred to as the *symmetric* case, whereas $c_1 \neq -c_2$ is called the *asymmetric* case.

This highlights, that the strength of our results is that α can be fixed. The price we have to pay is that trimming might be necessary.

The following theorems of this section are stated for the functions $\frac{1}{x}$ resp. $x^{-\beta}$. More generally, they apply also to $\frac{c_1}{x} + \frac{c_2}{1-x}$ resp. $c_1 x^{-\beta} + c_2(1-x)^{-\beta}$ with $c_1 \neq -c_2$, with only minor changes in the proof which we leave to the interested reader. However, note that, in contrast to [DF15] and [SU08], we only discuss the asymmetric case.

3.3.2 The function $\frac{1}{x}$

First let us consider $f(x) = \frac{1}{x}$. As it turns out, for almost all α and the function f, a weak law of large numbers holds, even without trimming. On the other hand, for a G_δ-dense set of α, even the weak law does not hold without trimming.

The crucial condition here is a well-known Diophantine condition on α.

Definition 48. *The number α is said to be of Roth type if, for all $\epsilon > 0$ and large enough n, it holds that $q_{n+1} < q_n^{1+\epsilon}$. Equivalently*

$$\lim_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} = 1.$$

Theorem 49. *There are $d_N > 0$ such that*

$$\lambda\left(\left|\frac{S_N(f)}{d_N} - 1\right| > \epsilon\right) \to 0 \quad \text{as } N \to \infty \quad \forall \epsilon > 0,$$

if and only if α is of Roth type. Furthermore, in this situation

$$d_N = N\log(N).$$

Remark 50. *Contrary to the situation in Theorem 49, for α not being of Roth type, we may need a trimming sequence arbitrarily close to N to obtain a trimmed weak law. More precisely, given any sequence $l(N) = o(N)$, there is a G_δ dense set of α such that, whenever $k(N) \leq l(N)$ and $d_N > 0$, there is an $\epsilon > 0$ such that*

$$\limsup_{N\to\infty} \lambda\left(\left|\frac{S_N^{k(N)}(f)}{d_N} - 1\right| > \epsilon\right) > 0.$$

Remark 51. *A trimmed weak law generally implies an untrimmed weak law under condition* (3.11). *More precisely, let (X, T, μ) be a probability-preserving dynamical system, and suppose there exist $k(N) = o(N)$ and $d_N > 0$ such that*

$$\mu\left(\left|\frac{S_N^{k(N)}(f)}{d_N} - 1\right| > \epsilon\right) \to 0 \quad \text{as } N \to \infty, \quad \forall \epsilon > 0.$$

If, in addition,

$$\mu\left(f > c\frac{d_N}{k(N)}\right) = o\left(\frac{1}{N}\right) \quad \forall c > 0, \tag{3.11}$$

then the untrimmed weak law follows:[19]

$$\mu\left(\left|\frac{S_N(f)}{d_N}-1\right|>\epsilon\right)\to 0 \quad \textit{as } N\to\infty, \quad \forall\epsilon>0.$$

Next, we investigate strong laws. As above for iids, or the Gauss map, for almost all α it suffices to trim by $k(N)=1$.

Theorem 52. *For almost all α it holds that*

$$\lim_{N\to\infty}\frac{S_N^1(f)(x)}{N\log(N)}=1 \quad \textit{for a.e. } x\in[0,1). \tag{3.12}$$

Furthermore, if α is of bounded type, then this convergence is uniform in x, i.e.

$$\lim_{N\to\infty}\frac{S_N^1(f)(x)}{N\log(N)}=1 \quad \textit{uniformly in } x\in[0,1).$$

The convergence in (3.12) only holds for almost every x, in fact;

Theorem 53. *For almost all α it holds that*

$$\limsup_{N\to\infty}\frac{S_N^1(f)(\alpha)}{N\log(N)}>1\geq\liminf_{N\to\infty}\frac{S_N^1(f)(\alpha)}{N\log(N)}.$$

Remark 54. *The set of all x such that*

$$\limsup_{N\to\infty}\frac{S_N^1(f)(x)}{N\log(N)}>1\geq\liminf_{N\to\infty}\frac{S_N^1(f)(x)}{N\log(N)}$$

is invariant under R, as will be shown in the proof of Theorem 55. Therefore it is also true that

$$\limsup_{N\to\infty}\frac{S_N^1(f)(0)}{N\log(N)}>1\geq\liminf_{N\to\infty}\frac{S_N^1(f)(0)}{N\log(N)}.$$

Lastly, there are[20] α, such that the untrimmed weak law holds[21], but not the trimmed strong law for any $k(N)$ being constant.

Theorem 55. *There is an α, such that*

$$\lambda\left(\left|\frac{S_N(f)}{N\log(N)}-1\right|>\epsilon\right)\to 0 \quad \textit{as } N\to\infty \quad \forall\epsilon>0,$$

but, for every $K\geq 1$ and almost every x, it holds that

$$\limsup_{N\to\infty}\frac{S_N^K(f)(x)}{N\log(N)}>1\geq\liminf_{N\to\infty}\frac{S_N^K(f)(x)}{N\log(N)}. \tag{3.13}$$

[19] For any $\epsilon>0$, we estimate

$$\begin{aligned}\mu\left(|S_N(f)-d_N|>\epsilon d_N\right)&\leq\mu\left(\left|S_N^{k(N)}(f)-d_N\right|>\frac{\epsilon d_N}{2}\right)+\mu\left(\exists 0\leq n\leq N-1 \mid f\circ T^n>\frac{\epsilon d_N}{2k(N)}\right)\\&\leq\mu\left(\left|S_N^{k(N)}(f)-d_N\right|>\frac{\epsilon d_N}{2}\right)+N\mu\left(f>\frac{\epsilon d_N}{2k(N)}\right)\to 0,\end{aligned}$$

as $N\to\infty$.

[20] As before, the set of such α will be G_δ-dense. We will not show this fact, as it can easily be verified and is not essential to this work, leaving it instead to the reader.

[21] Or equivalently α is of Roth type.

3.3.3 The function $x^{-\beta}$ for $\beta > 1$

Let $\beta > 1$ and $f(x) = x^{-\beta}$. We will prove sufficient and necessary conditions for $k(N)$ such that trimmed weak or strong laws hold. The general conditions are slightly complicated to state, but, in the case of monotone trimming sequences, reduce to a much simpler one. Here we will only state the condition for monotone trimming sequences, leaving the more general statements for the proof section.

Theorem 56. *Let $k(N)$ be monotone with $k(N) = o(N)$ and $d_N > 0$, then the following are equivalent*

(I) the trimmed strong law holds uniformly, i.e.

$$\lim_{N\to\infty} \frac{S_N^{k(N)}(f)(x)}{d_N} = 1 \quad \textit{uniformly in } x \in [0,1), \tag{3.14}$$

(II) the trimmed weak law holds, i.e.

$$\lambda\left(\left|\frac{S_N^{k(N)}(f)}{d_N} - 1\right| > \epsilon\right) \to 0 \textit{ as } N \to \infty \quad \forall \epsilon > 0, \tag{3.15}$$

(III)

$$\frac{k(N)}{\max(b_n, \max_{j\le n-1} a_j)} \to \infty \textit{ as } N \to \infty. \tag{3.16}$$

Furthermore, in this situation, $d_N = \frac{1}{\beta-1} N^\beta k(N)^{1-\beta}$.

3.3.4 Comparison of results to literature

As our results seem to be the first that do not require strong mixing properties of the system, it is of special interest to compare our results, to the known ones listed in §3.1.2. We shall compare trimming results for the non-integrable functions $\frac{1}{x}$ and $x^{-\beta}$ with $\beta > 1$. All of the results below were discussed in §3.1.2 or earlier in this section. In each cell of the table, we write sufficient conditions for a trimming sequence $k(N) = o(N)$ such that there are $d_N > 0$ with

$$\lim_{N\to\infty} \frac{S_N^{k(N)}(f)(x)}{d_N} = 1 \quad \text{for } \mu\text{-a.e. } x \in X.$$

To simplify notation, from now on we introduce the following; For x sufficiently large such that the below is defined denote

$$\log_1(x) = \log(x) \;\text{ and }\; \log_k(x) = \log(\log_{k-1}(x)), \;\; k \ge 2.$$

	$\frac{1}{x}$	$x^{-\beta}$, $\beta > 1$
iid	$k(N) = 1$	$\frac{k(N)}{\log_2(N))} \to \infty$
Gauss map	$k(N) = 1$	$\frac{k(N)}{\log_2(N))} \to \infty$
doubling map	$k(N) = \lceil \kappa \log_3(N) \rceil$, $\kappa > \frac{1}{\log(2)}$	$\frac{k(N)}{\log_2(N))} \to \infty$
almost every rotation	$k(N) = 1$	$\frac{k(N)}{\log^{1+\epsilon}(N)} \to \infty$

In all of the cases we have $d_N = c_1 N \log(N)$ for the function $\frac{1}{x}$ and $d_N = \frac{c_\beta}{\beta-1} N^\beta k(N)^{-\beta+1}$ for $x^{-\beta}$, where $c_1 = c_\beta = 1$ in all cases except for the Gauss map - in case of the Gauss map we have $c_1 = \frac{1}{\log(2)}$ and $c_\beta = \frac{1}{(\log(2))^\beta}$[22].

[22]The phenomenon that we get a different constant in case of the Gauss map is due to the fact the Gauss map is invariant with respect to the Gauss measure which is equivalent but not equal to the Lebesgue measure - all other ergodic transformations are invariant with respect to the Lebesgue measure.

A similar comparison can be made for trimmed weak laws;

- For iids, the Gauss map, and the doubling map, together with the function $\frac{1}{x}$ weak laws hold even without trimming. This was shown in [Fel57, Theorem 2 VII.], [Khi35] and [Sch18] respectively, but can be also deduced directly from the trimmed strong laws in the table above.[23]
- For iids, the Gauss map, and the doubling map, and the function $x^{-\beta}$ trimmed weak laws hold for any intermediate trimming sequence $k(N) \to \infty$ with $k(N) = o(N)$. This was shown in [KS20].
- For rotations of Roth type and the function $\frac{1}{x}$ the weak law also holds without trimming, as shown in Theorem 49. However, as demonstrated in Remark 50, for a generic rotation number α, we might require a trimming sequence arbitrarily close to N to obtain a trimmed weak law.
- For rotations and the function $x^{-\beta}$, Theorem 56 states exact conditions on $k(N)$ so that a trimmed weak law holds. Also here generically we will need a sequence arbitrarily close to N. Only for α of bounded type we can choose a trimming sequence that approaches infinity arbitrarily slowly.

This suggests, that trimming is well applicable even beyond the mixing case. Finding more flexible conditions and techniques to obtain effective trimming results without using strong mixing properties is a promising direction for future research.

3.4 Proof of general trimmed laws

Proof of Theorem 41. (i) For every $l \geq 1$ denote $f_l = \min(f, l)$, and $a_l = \int_X f_l \, \mathrm{d}\mu$. Using Jegorow's Theorem, there are sets B_l with $\mu(B_l) < 2^{-l}$ and $1 \leq N_1 \leq N_2 \leq \ldots$, with $N_l \geq l$ such that

$$|S_n(1_{\{f \geq l\}})(x) - n\mu(f \geq l)| < 2^{-l} n\mu(f \geq l) \tag{3.17}$$

and

$$|S_n(f_l)(x) - na_l| < 2^{-l} na_l \tag{3.18}$$

for $x \notin B_l$ and $n \geq N_l$. Let $k(n)$ be defined as

$$k(n) = \lceil (1 + 2^{-l}) n\mu(f \geq l) \rceil, \quad n \in [N_l, N_{l+1} - 1].$$

Since $\mu(f \geq l) \to 0$ as $l \to \infty$, it is clear that $k(n) = o(n)$. By the Borel-Cantelli lemma, almost every x is only in finitely many B_l. For such an x and l big enough such that $x \notin B_j$ for $j \geq l$, using (3.17) and (3.18), for $n \geq N_l$ it holds that

$$\begin{aligned} |S_n^{k(n)}(f)(x) - na_l| &\leq |S_n^{k(n)}(f)(x) - S_n(f_l)(x)| + |S_n(f_l)(x) - na_l| \\ &< 2^{-l+1} l n\mu(f \geq l) + l + 2^{-l} na_l. \end{aligned}$$

Since $a_l \nearrow \infty$ as $l \to \infty$, and $n \geq l$, the right-hand side is $o(na_l)$, hence

$$\lim_{n \to \infty} \frac{S_n^{k(n)}(f)(x)}{d_n} = 1 \quad \text{for } \mu\text{-a.e. } x \in X,$$

[23] This follows from Remark 51 since in all cases, for all $c > 0$,

$$\mu\left(f > c\frac{d_N}{k(N)}\right) \ll \frac{k(N)}{N \log(N)} = o\left(\frac{1}{N}\right).$$

where $d_n = na_l$ for $n \in [N_l, N_{l+1} - 1]$.

(ii) Now suppose (X, T, μ) is uniquely ergodic (and μ is regular) and f is continuous. Since the sets $\{f = s\}$, for $s \geq 0$, are disjoint, only countably many can have positive measure. Therefore, let $s_l \nearrow \infty$ be a sequence with $\mu(f = s_l) = 0$, let $f_l = \min(f, s_l)$ and $a_l = \int_X f_l \, \mathrm{d}\mu$, note that f_l is continuous. By regularity of μ, there is an open set $O_l \supset \{f \geq s_l\}$ such that $\mu(O_l \setminus \{f \geq s_l\}) < 2^{-l}\mu(f \geq s_l)$, and by Urysohn's Lemma there is a continuous function χ_l^+ that is 1 on $\{f \geq s_l\}$ and 0 outside O_l. Likewise, there is a closed set $C_l \subset \{f > s_l\}$, with $\mu(\{f > s_l\} \setminus C_l) < 2^{-l}\mu(f \geq s_l)$, and a continuous function χ_l^- that is 1 on C_l and 0 outside $\{f > s_l\}$. There are $1 \leq N_1 \leq N_2 \leq ...$, with $N_l \geq s_l$, such that

$$\begin{aligned}
\left| S_n(\chi_l^+)(x) - n \int_X \chi_l^+ \, \mathrm{d}\mu \right| &< 2^{-l} n\mu(f \geq s_l) \text{ and} \\
\left| S_n(\chi_l^-)(x) - n \int_X \chi_l^- \, \mathrm{d}\mu \right| &< 2^{-l} n\mu(f \geq s_l)
\end{aligned}$$

as well as

$$|S_n(f_l)(x) - na_l| < 2^{-l} na_l,$$

for $x \in X$ and $n \geq N_l$. For

$$k(n) = \lceil (1 + 2^{-l+1}) n\mu(f \geq s_l) \rceil, \quad n \in [N_l, N_{l+1} - 1],$$

analogously to the above, it follows that

$$|S_n^{k(n)}(f)(x) - na_l| < 2^{-l+2} s_l n\mu(f \geq s_l) + s_l + 2^{-l} na_l.$$

Since $a_l \geq s_l \mu(f \geq s_l)$, $a_l \to \infty$ and $n \geq N_l \geq s_l$, the right-hand side is $o(na_l)$ as $l \to \infty$. The claim follows with $d_n = na_l$ for $n \in [N_l, N_{l+1} - 1]$. $\square$

To prove Theorem 42, we will construct a function f using Rokhlin Towers. Here we use the following version of Rokhlin's Tower theorem ([Aar97, Theorem 1.5.9]).

Lemma 57. *Let (X, T, μ) be an aperiodic probability preserving ergodic system. Then, for $N \geq 1$ and $\epsilon > 0$, there is a measurable set A such that $\{T^{-j}(A)\}_{j=0}^{N-1}$ are disjoint and $\mu\left(X \setminus \bigcup_{j=0}^{N-1} T^{-j}(A)\right) < \epsilon$.*

Proof of Theorem 42. In order to prove Theorem 42 it will be enough to construct a function $f \in \mathcal{F}$, a sequence $(N_l)_{l \geq 1}$, and normalising sequences $D_l, D_l' > 0$ such that

$$\lim_{l \to \infty} \frac{S_{N_l}(f)(x)}{D_l} = 1 \quad \text{for } \mu\text{-a.e. } x \in X, \tag{3.19}$$

but

$$\mu\left(S_{N_l}^1(f) \geq D_l'\right) \geq \frac{1}{3}, \quad \text{while} \quad \mu\left(S_{N_l}^1(f) \leq \frac{D_l'}{10}\right) \geq \frac{1}{3}. \tag{3.20}$$

To conclude the proof, let $\tilde{k}(N) = o(N)$ and $\tilde{d}_N > 0$ be such that

$$\lim_{N \to \infty} \frac{S_N^{\tilde{k(N)}}(f)(x)}{\tilde{d}_N} = 1 \quad \text{for } \mu\text{-a.e. } x \in X,$$

the existence of such $\tilde{k}(N)$, $\tilde{d}_N$ is guaranteed by Theorem 41. Now the conclusion follows for

$$k(N) = \begin{cases} 0 & \text{if } N = N_l \text{ for some } l \geq 1, \\ \tilde{k}(N) & \text{otherwise,} \end{cases}$$

$$k'(N) = \begin{cases} 1 & \text{if } N = N_l \text{ for some } l \geq 1, \\ \tilde{k}(N) & \text{otherwise,} \end{cases}$$

$$d_N = \begin{cases} D_l & \text{if } N = N_l \text{ for some } l \geq 1, \\ \tilde{d}_N & \text{otherwise,} \end{cases}$$

and

$$d'_N = \begin{cases} D'_l & \text{if } N = N_l \text{ for some } l \geq 1, \\ \tilde{d}_N & \text{otherwise.} \end{cases}$$

Let $N_l = 2^{2^{2^l}}$ and $\epsilon_l = 2^{-N_l}$, the Rokhlin Tower theorem, Lemma 57, yields a measurable set A_l such that $\{T^{-j}(A_l)\}_{j=0}^{N_l-1}$ are disjoint and $\mu(R_l) < \epsilon_l$, where $R_l = X \setminus \bigcup_{j=0}^{N_l-1} T^{-j}(A_l)$. Note that $\frac{1-2^{-N_l}}{N_l} \leq \mu(A_l) \leq \frac{1}{N_l}$. Let $B_l \subset T^{-1}(A_l)$ be a measurable set with $\mu(B_l) = \frac{\mu(A_l)}{2}$ and define a function $g_l : X \to [0, \infty)$ by

$$g_l(x) = \begin{cases} 2^{N_l} & \text{if } x \in A_l, \\ 2^{N_l - l} & \text{if } x \in B_l, \\ 0 & \text{otherwise.} \end{cases}$$

Let $f = \sum_{l=1}^{\infty} g_l$. Since each g_l is non-zero only on a set of measure at most $\frac{3}{2N_l}$, the function f is almost everywhere defined by a finite sum. Furthermore, f is non-integrable since $\int g_l \,\mathrm{d}\mu \geq \frac{1}{2}$ for each $l \geq 1$.

For each $l \geq 1$, every point $x \notin R_l$ will visit A_l within N_l steps, therefore $S_{N_l}(f)(x) \geq 2^{N_l}$. On the other hand, if x does not visit $A_k \cup B_k$ within[24] N_l steps for any $k > l$, then, for big enough l, it holds

$$S_{N_l}(f)(x) \leq (1 + 2^{-l})2^{N_l} + N_l \sum_{j=1}^{l-1} ||g_j||_{L^\infty} \leq (1 + 2^{-l})2^{N_l} + N_l \sum_{j=1}^{l-1} 2^{N_j} \leq (1 + 2^{-l+1})2^{N_l}.$$

Therefore, for big enough l,

$$\mu \left(\left| \frac{S_{N_l}(f)}{2^{N_l}} - 1 \right| > 2^{-l+1} \right) \leq \mu \left(R_l \cup \tilde{R}_l \right) \leq 2^{-N_l} + 2N_l \sum_{k=l+1}^{\infty} \frac{1}{N_k} \leq \frac{1}{N_l},$$

where $\tilde{R}_l = \bigcup_{k=l+1}^{\infty} \bigcup_{j=0}^{N_l-1} T^{-j}(A_k \cup B_k)$. By Borel-Cantelli (3.19) holds with $D_l = 2^{N_l}$.

On the other hand, every point $x \notin A_l \cup R_l$ will visit $T^{-1}(A_l)$ within N_l steps, of those points half will visit B_l and the other half will visit[25] $T^{-1}(A_l) \setminus B_l$. In addition, all of these points will visit A_l exactly once, that is the largest value that we trim, unless $A_k \cup B_k$ is visited for some $k > l$, then the largest value will be one of the terms contributed by g_k. Therefore, for big enough l, we have

$$\mu \left(S^1_{N_l}(f) \geq 2^{N_l - l} \right) \geq \mu \left(\bigcup_{j=0}^{N_l-2} T^{-j}(B_l) \right) \geq \frac{(1 - 2^{-N_l})(N_l - 1)}{2N_l} \geq \frac{1}{3}.$$

If x does not visit B_l within N_l steps, and avoids all $A_k \cup B_k$ for $k > l$, then all the non-zero contribution can only come from visiting A_l, which however is the value that will be trimmed,

[24] Here we count x itself as the first step.

[25] And hence they will avoid B_l altogether since it would take at least N_l steps to visit $T^{-1}(A_l)$ again.

and visits to $A_k \cup B_k$ for $k < l$. Therefore, for $x \in \bigcup_{j=0}^{N_l-2} T^{-j}(T^{-1}(A_l) \setminus B_l) \setminus \tilde{R}_l$ and big enough l it holds that

$$S^1_{N_l}(f)(x) \leq N_l \sum_{k=1}^{l-1} ||g_k||_{L^\infty} \leq N_l \sum_{k=1}^{l-1} 2^{N_k} \leq \frac{1}{10} 2^{N_l - l}.$$

The set in question has measure

$$\mu\left(\bigcup_{j=0}^{N_l-2} T^{-j}(T^{-1}(A_l) \setminus B_l) \setminus \tilde{R}_l\right) \geq \frac{(1-2^{-N_l})(N_l-1)}{2N_l} - 2N_l \sum_{k=l+1}^{\infty} \frac{1}{N_k} \geq \frac{1}{3}.$$

It follows that (3.20) holds with $D'_l = 2^{N_l - l}$. □

3.5 Proof of polynomial trimming

In this section, we shall provide a proof of Theorem 45. First, we shall prove uniform convergence assuming (polynomial) unique ergodicity, i.e. assumption (B). Throughout this Chapter, we shall use the notation introduced in §3.2.

In order to make use of the deviation of ergodic averages, we will approximate indicators of small balls by smooth functions.

Lemma 58. *For $\epsilon, \hat{\epsilon} > 0$, with $\epsilon + \hat{\epsilon}$ small enough, there is a C^r-function $g_{\epsilon,\hat{\epsilon}} : M \to [0,1]$ with $g_{\epsilon,\hat{\epsilon}} \equiv 1$ on $B_\epsilon(x^*)$ and $g_{\epsilon,\hat{\epsilon}} \equiv 0$ on $B_{\epsilon+\hat{\epsilon}}(x^*)^c$, furthermore it holds that $||g_{\epsilon,\hat{\epsilon}}||_{C^r} \ll \hat{\epsilon}^{-r}$.*

Proof. Let $\chi : [0,1] \to [0,1]$ be a smooth function with $\chi(0) = 1, \chi(1) = 0$ and

$$\chi^{(j)}(0) = \chi^{(j)}(1) = 0 \quad \forall j \geq 1.$$

Suppose $\epsilon + \hat{\epsilon}$ is so small that the exponential map $\exp_{x^*} : B_{\epsilon+\hat{\epsilon}}(0) \to B_{\epsilon+\hat{\epsilon}}(x^*)$ is a well-defined diffeomorphism. Then for

$$g_{\epsilon,\hat{\epsilon}}(x) = \begin{cases} 1 & \text{if } d(x,x^*) < \epsilon, \\ \chi\left(\hat{\epsilon}^{-1}(|\exp_{x^*}^{-1}(x)|_2 - \epsilon)\right) & \text{if } \epsilon \leq d(x,x^*) < \epsilon + \hat{\epsilon}, \\ 0 & \text{if } d(x,x^*) \geq \epsilon + \hat{\epsilon}, \end{cases}$$

all the desired properties are easy to check. □

There are a constant $\Omega > 0$ and an $\epsilon^* > 0$ such that

$$(1-\Omega\epsilon) B_d \rho(x^*) \epsilon^d \leq \mu(B_\epsilon(x^*)) \leq (1+\Omega\epsilon) B_d \rho(x^*) \epsilon^d \quad \forall \epsilon < \frac{\epsilon^*}{2}.$$

Recall that by assumption there is a $\delta > 0$ such that

$$|S_n(g)(x) - n \int g \, \mathrm{d}\mu| \ll n^{1-\delta} ||g||_{C^r} \quad \forall n \geq 1, g \in C^r, x \in M. \tag{3.21}$$

Let us first show bounds on sums of indicators.

Lemma 59. *For $\epsilon > 0$ small enough and $n > \epsilon^{-\frac{r+d}{\delta}}$ it holds that*

$$|S_n(1_{B_\epsilon(x^*)})(x) - n B_d \rho(x^*) \epsilon^d| \ll n\epsilon^{d+1} + \epsilon^{-\frac{r(d-1)}{r+1}} n^{1-\frac{\delta}{r+1}} \quad \forall x \in M, \tag{3.22}$$

the implicit constant is independent of ϵ and n.

Proof. For $\epsilon > \hat{\epsilon} > 0$ to be chosen later let $g = g_{\epsilon,\hat{\epsilon}}$ be the function from Lemma 58. From (3.21) we obtain

$$|S_n(g)(x) - n \int g \,\mathrm{d}\mu| \ll n^{1-\delta}\hat{\epsilon}^{-r}.$$

Since $(1-\Omega\epsilon)\rho(x^*)B_d\epsilon^d \le \int g \,\mathrm{d}\mu \le (1+2\Omega\epsilon)\rho(x^*)B_d(\epsilon+\hat{\epsilon})^d$ and $g \ge 1_{B_\epsilon(x^*)}$, it follows that

$$S_n(1_{B_\epsilon(x^*)})(x) - nB_d\rho(x^*)\epsilon^d \le S_n(g)(x) - B_d\rho(x^*)\epsilon^d \ll n\epsilon^{d+1} + n\epsilon^{d-1}\hat{\epsilon} + n^{1-\delta}\hat{\epsilon}^{-r},$$

and setting[26] $\hat{\epsilon} = (n^\delta\epsilon^{d-1})^{-\frac{1}{r+1}}$ yields

$$S_n(1_{B_\epsilon(x^*)})(x) - nB_d\rho(x^*)\epsilon^d \ll n\epsilon^{d+1} + \epsilon^{-\frac{r(d-1)}{r+1}}n^{1-\frac{\delta}{r+1}}. \tag{3.23}$$

The same arguments for $g = g_{\epsilon-\hat{\epsilon},\hat{\epsilon}}$ instead of $g_{\epsilon,\hat{\epsilon}}$ yields

$$nB_d\rho(x^*)\epsilon^d - S_n(1_{B_\epsilon(x^*)})(x) \ll n\epsilon^{d+1} + \epsilon^{-\frac{r(d-1)}{r+1}}n^{1-\frac{\delta}{r+1}}. \tag{3.24}$$

The claim follows from (3.23) and (3.24). □

Let $\epsilon(n) = \left(\frac{k(n)}{\rho(x^*)B_d n}\right)^{\frac{1}{d}}$ and denote

$$f_n(x) = f(x)(1 - 1_{B_{\epsilon(n)}(x^*)}(x)) = \begin{cases} 0 & \text{if } d(x,x^*) < \epsilon(n), \\ f(x) & \text{otherwise} \end{cases}.$$

Considering Lemma 59, the idea is that $S_n^{k(n)}(f) \approx S_n(f_n)$ as will be made clear in the next Lemma.

Lemma 60. *It holds that*

$$|S_n^{k(n)}(f)(x) - S_n(f_n)(x)| = o(n^{\frac{\beta}{d}}k(n)^{1-\frac{\beta}{d}}) \textit{ as } n \to \infty, \textit{ uniformly in } x.$$

Proof. Let $C_d > 1$ be the implicit constant in (3.22), and

$$\epsilon'(n) = \min\left(\left(\frac{k(n)}{10\rho(x^*)B_d n}\right)^{\frac{1}{d}}, \left(\frac{k(n)}{10C_d n}\right)^{\frac{1}{d}}\right).$$

For n big enough[27] $n > \epsilon'(n)^{-\frac{r+d}{\delta}}$ and we can apply Lemma 59 to obtain

$$S_n(1_{B_{\epsilon'(n)}(x^*)})(x) \le \frac{k(n)}{10} + \frac{k(n)}{10} + k(n)^{-\frac{r(d-1)}{d(r+1)}}n^{1-\frac{\delta}{r+1}+\frac{r(d-1)}{d(r+1)}},$$

and since $n^{1-\frac{\delta}{r+1}+\frac{r(d-1)}{d(r+1)}} = o\left((n^{\alpha\left(1-\frac{\delta}{r+1}+\frac{r(d-1)}{d(r+1)}\right)}\right) = o\left(k(n)^{1+\frac{r(d-1)}{d(r+1)}}\right)$ we can choose n so big that

$$S_n(1_{B_{\epsilon'(n)}(x^*)})(x) \le \frac{k(n)}{2}. \tag{3.25}$$

Number the points $\{x, Tx, ..., T^{n-1}x\}$ as $x_1, ..., x_n$ in such a way that $f(x_1) \ge f(x_2) \ge ... \ge f(x_n)$, then

$$|S_n^{k(n)}(f)(x) - S_n(f_n)(x)| \le \sum_{\substack{j>k(n) \\ x_j \in B_{\epsilon(n)}(x^*)}} f(x_j) + \sum_{\substack{j \le k(n) \\ x_j \notin B_{\epsilon(n)}(x^*)}} f(x_j). \tag{3.26}$$

[26] The assumption $n > \epsilon^{-\frac{r+d}{\delta}}$ is to ensure that $\hat{\epsilon} < \epsilon$.

[27] Because $\epsilon'(n)^{-\frac{r+d}{\delta}} = O(n^{\frac{(1-\alpha)(r+d)}{\delta}}) = o(n)$.

By Lemma 59,

$$|S_n(1_{B_{\epsilon(n)}(x^*)})(x) - k(n)| = \phi(n)k(n),$$

for some function $\phi : \mathbb{N} \to [0,1]$ with $\phi(n) \to 0$ as $n \to \infty$, so

$$\begin{aligned} x_j \in B_{\epsilon(n)}(x^*) &\implies j \leq k(n) + \phi(n)k(n) \text{ and} \\ x_j \notin B_{\epsilon(n)}(x^*) &\implies j \geq k(n) - \phi(n)k(n). \end{aligned} \tag{3.27}$$

Due to (3.25), all the x_j on the right-hand side of (3.26) lie outside $B_{\epsilon'(n)}(x^*)$, therefore there is a constant $c > 0$, independent of n and x, such that $f(x_j) \leq c\epsilon'(n)^{-\beta}$ for all such points. This fact together with (3.26) and (3.27) implies that

$$|S_n^{k(n)}(f)(x) - S_n(f_n)(x)| \leq 2c\phi(n)k(n)(\epsilon'(n))^{-\beta} = o(n^{\frac{\beta}{d}}k(n)^{1-\frac{\beta}{d}}).$$

□

Lemma 61. *For $\beta > d$ we have*

$$S_n(f_n)(x) = \frac{1}{\beta - d} Res_{x^*}(f) B_d \rho(x^*) n\epsilon(n)^{d-\beta}(1 + o(1)) \quad \textit{uniformly in } x.$$

If $\beta = d$ then

$$S_n^{k(n)}(f_n)(x) = Res_{x^*}(f) B_d \rho(x^*) n |\log(\epsilon(n))| (1 + o(1)) \textit{ as } n \to \infty, \textit{ uniformly in } x.$$

Proof. We focus on the case $\beta > d$, the case $\beta = d$ can be proven analogously. Let $\hat{\epsilon}(n) > 0$ with $\hat{\epsilon}(n) = o(\epsilon(n))$ to be chosen later, let $g = g_{\epsilon(n),\hat{\epsilon}(n)}$ be the function from Lemma 58 and denote $\hat{f}_n = (1-g)f$. Then $\hat{f}_n \leq f_n$ and $||\hat{f}_n||_{C^r} \ll \epsilon(n)^{-r-\beta}\hat{\epsilon}(n)^{-r}$, furthermore[28]

$$\int \hat{f}_n \, \mathrm{d}\mu = \frac{1}{\beta - d} \mathrm{Res}_{x^*}(f) B_d \rho(x^*) \epsilon(n)^{d-\beta}(1 + o(1)).$$

It follows from (3.21) that

$$\begin{aligned} S_n(f_n)(x) \leq S_n(\hat{f}_n)(x) &\leq \\ &\frac{1}{\beta - d} \mathrm{Res}_{x^*}(f) B_d \rho(x^*) n\epsilon(n)^{d-\beta}(1 + o(1)) + O(\epsilon(n)^{-r-\beta}\hat{\epsilon}(n)^{-r}). \end{aligned}$$

Choosing $\hat{\epsilon}(n)$ such that[29] $n^{-\frac{1}{r}}\epsilon(n)^{-\frac{r+d}{r}} = o(\hat{\epsilon}(n))$, we obtain

$$S_n(f_n)(x) \leq \frac{1}{\beta - d} \mathrm{Res}_{x^*}(f) B_d \rho(x^*) n\epsilon(n)^{d-\beta}(1 + o(1)).$$

The same arguments with $g = g_{\epsilon(n)-\hat{\epsilon}(n),\hat{\epsilon}(n)}$ instead of $g_{\epsilon(n),\hat{\epsilon}(n)}$ yield

$$S_n(f_n)(x) \geq \frac{1}{\beta - d} n \mathrm{Res}_{x^*}(f) B_d \rho(x^*) \epsilon(n)^{d-\beta}(1 + o(1)).$$

□

[28] Since $\int_{||\mathbf{x}|| \in [\epsilon,1]} ||\mathbf{x}||^{-\beta} \, \mathrm{d}\mathbf{x} = \frac{1}{\beta - d} B_d \epsilon^{d-\beta}(1 - o(1))$.

[29] This is possible since $n^{-\frac{1}{r}}\epsilon(n)^{-\frac{r+d}{r}} = O\left(n^{-\frac{1}{r} + \frac{(1-\alpha)(r+d)}{rd}}\right) = o\left(n^{-\frac{(1-\alpha)}{d}}\right) = o(\epsilon(n))$.

Proof of Theorem 45 under assumption (B). First assume $\beta > d$. Lemma 61 yields

$$S_n^{k(n)}(f_n)(x) = \frac{1}{\beta - d}\mathrm{Res}_{x^*}(f)\rho(x^*)^{\frac{\beta}{d}} B_d^{\frac{\beta}{d}} n^{\frac{\beta}{d}} k(n)^{1-\frac{\beta}{d}}(1 + o(1)) \text{ as } n \to \infty, \text{ uniformly in } x,$$

while Lemma 60 ascertains that

$$|S_n^{k(n)}(f)(x) - S_n(f_n)(x)| = o(n^{\frac{\beta}{d}} k(n)^{1-\frac{\beta}{d}}) \text{ as } n \to \infty, \text{ uniformly in } x.$$

If $\beta = d$ similar arguments apply. □

Under assumption (A), the conclusion can be proven by almost the same arguments, utilizing the pointwise estimates from [Aue25, Proposition 9.1].

Proposition 62. *Assume that (A) holds, and let $(g_l)_{l\geq 1}$ be a sequence of C^r functions. Then, for every $\epsilon > 0, \epsilon' > 0$, there are $L_x \geq 1$ such that*

$$\left| S_n(g_l)(x) - n \int g_l d\mu \right| \leq ||g_l||_{C^r} n^{\delta+\epsilon} \quad \text{for } \mu\text{-a.e } x \in M, \forall n \geq l^{\epsilon'}, l \geq L_x.$$

Proof of Theorem 45, general case. We shall present here the argument for $\beta > d$, the argument for $\beta = d$ goes analogously. For $\epsilon > 0$, and choosing[30] $(g_l)_{l\geq 1}$ in a clever way, Proposition 62 yields for μ-a.e $x \in M$

$$\left| S_n(g_l)(x) - n \int g_l d\mu \right| \leq ||g_l||_{C^r} n^{\delta+\epsilon} \quad \forall n \geq l^{\epsilon'}, \quad l \geq L_x.$$

For such x, the proof of case (B) then shows

$$S_n^{k(n)}(f_n)(x) = \frac{1}{\beta - d} c\rho(x^*)^{\frac{\beta}{d}} B_d^{\frac{\beta}{d}-1} n^{\frac{\beta}{d}} k(n)^{1-\frac{\beta}{d}}(1 + o(1)),$$

whenever $k(n) = \lceil n^\alpha \rceil$ for some

$$1 > \alpha > \alpha_0 = \max\left(\frac{r + d - d(\delta + \epsilon)}{r + d}, \frac{2rd + d - r - d(\delta + \epsilon)}{2rd + d - r}, \frac{r + 1}{r + d + 1}\right)$$

The proof is concluded by letting $\epsilon \to 0$. □

3.6 Proof of trimmed laws for rotations

3.6.1 Background

Recall that, for $N \geq 1$ and n such that $N \in [q_n, q_{n+1} - 1]$, there are $0 \leq b_j < \frac{q_{j+1}}{q_j}$, for $j = 0, ..., n$, such that

$$N = \sum_{j=0}^{n} b_j q_j,$$

this is called the Ostrowski expansion.

In the following we will, by slight abuse of notation, identify $[0, 1)$ with $\left[-\frac{1}{2}, \frac{1}{2}\right)$ via the identification $\iota : [0, 1) \to \left[-\frac{1}{2}, \frac{1}{2}\right)$

$$\iota(x) = \begin{cases} x & \text{if } x < \frac{1}{2}, \\ x - 1 & \text{otherwise.} \end{cases}$$

[30] Here g_l are all the functions used in the proof of case (B); either smooth functions approximating $1_{B_{\epsilon(n)}}$, or smooth functions approximating f_n.

Whenever we write $x < 0 < y$ for some points $x, y \in [0,1)$, we implicitly refer to this identification.

Denote

$$\delta_n = |\alpha q_{n-1}|. \tag{3.28}$$

Then

$$\delta_n \in \left(\frac{1}{2q_n}, \frac{1}{q_n}\right).$$

Furthermore, as is easily deduced by the approximation properties of α, δ_n is given by a simple recursion.

Lemma 63. *For all $n \geq 1$ it holds that $a_n\delta_{n+1} + \delta_{n+2} = \delta_n$.*

Proof. Assume $\alpha - \frac{p_n}{q_n} > 0$, the other case can be proven similarly. Since $\alpha - \frac{p_{n-1}}{q_{n-1}} < 0$ and $\alpha - \frac{p_{n+1}}{q_{n+1}} < 0$, we have

$$R^{q_{n-1}}(0) < R^{q_{n-1}+q_n}(0) < ... < R^{q_{n-1}+a_n q_n}(0) = R^{q_{n+1}}(0) < 0.$$

The claim follows since $d(R^{q_{n-1}}(0), 0) = \delta_n$, $d(R^{q_{n+1}}(0), 0) = \delta_{n+2}$ and

$$d(R^{q_{n-1}+(i-1)q_n}(0), R^{q_{n-1}+iq_n}(0)) = \delta_{n+1} \quad \forall i = 1, ..., a_n.$$

$\square$

For $N \geq 1$ and $x \in [0,1)$ denote by $j_1^N(x), ..., j_N^N(x)$ the clockwise ordering of the points $\{x, ..., R^{N-1}(x)\}$ around the circle, for convenience leaving out 0 so that

$$0 < R^{j_1^N(x)}(x) < ... < R^{j_N^N(x)}(x) \leq 1,$$

where we identify $\mathbb{T}$ with $(0,1]$, and

$$f(R^{j_1^N(x)}(x)) > ... > f(R^{j_N^N(x)}(x)),$$

where we recall that, by the convention in Theorem 41, we set $f(0) = 0$. Sometimes we will also denote $R^{j_1^N(x)}(x) = x_{\min}^N$.

It holds that

$$f(R^{j_1^N(x)}(x)) > ... > f(R^{j_N^N(x)}(x)),$$

where $f(x) = x^{-\beta}$ for some $\beta \geq 1$. Hence, for $N \geq 1$ and $k \leq N$, it holds that

$$S_N^k(f)(x) = \sum_{l=k+1}^{N} f(R^{j_l^N(x)}(x)).$$

From this, we can immediately deduce some helpful bounds. Consider the case when $N \in [q_n, q_{n+1} - 1]$, $\alpha - \frac{p_n}{q_n} > 0$ and $j_1^N(x) = j_1^{q_n}(x)$. Then the points $\{x, ..., R^{N-1}(x)\}$ are grouped into clusters of size b_n or $b_n + 1$. If in addition $R^{j_1^N(x)}(x) > \frac{\epsilon}{q_n}$, for some $\epsilon > 0$, then, for each $k \geq 0$, we can bound

$$S_N^k(f)(x) \leq 2^\beta q_n^\beta (b_n + 1) \sum_{j=\left\lfloor \frac{k}{b_n} \right\rfloor}^{q_n - 1} (\epsilon + j)^{-\beta}.$$

This estimate is obtained by noticing that the leftmost point of each group[31] is at distance at least $\delta_n \geq \frac{1}{2q_n}$. A similar estimate can be made, assuming $k \geq b_n + 1$ instead of $j_1^N(x) = j_1^{q_n}(x)$,

[31] These are the points $\{x, ..., R^{q_n-1}(x)\}$ numbered clockwise, starting with $R^{j_1^{q_n}(x)}(x)$.

but the point of the above is that this calculation is independent of k. Many of our proofs will rely on this or a similar estimate, therefore it will be crucial to first understand better when $j_1^N(x) = j_1^{q_n}(x)$.

Lemma 64. *If $\alpha - \frac{p_n}{q_n} > 0$ and*[32] $R^{j_{q_n}^{q_n}(x)}(x) \leq -b_n\delta_{n+1}$*, then $j_1^N(x) = j_1^{q_n}(x)$.*

Proof. It holds that

$$R^{j_{q_n}^{q_n}(x)}(x) < R^{j_{q_n}^{q_n}(x)+q_n}(x) < ... < R^{j_{q_n}^{q_n}(x)+b_nq_n}(x) < R^{j_1^{q_n}(x)}(x),$$

and no other point of $\{x, ..., R^{N-1}(x)\}$ is between $R^{j_{q_n}^{q_n}(x)}(x)$ and $R^{j_1^{q_n}(x)}(x)$. Furthermore,

$$R^{j_{q_n}^{q_n}(x)+b_nq_n}(x) \leq R^{j_{q_n}^{q_n}(x)}(x) + b_n\delta_{n+1} \leq 0,$$

and the conclusion is immediate. □

3.6.2 The function $\frac{1}{x}$

For this section, denote $f(x) = \frac{1}{x}$.

Before starting the section with some technical results, we first give an overview of the main propositions in this section. First, using the Denjoy-Koksma inequality, we find suitable $d_N > 0$ and $\epsilon_n > 0$ such that

$$|S_N^{b_n+1}(f)(x) - d_N| < \epsilon_N \quad \forall x \in [0, 1),$$

this is shown in Lemma 66 and 70. If α is of Roth type, then we can show that $d_N \sim N\log(N)$ and $\epsilon_N = o(N\log(N))$, this will be shown in Proposition 74 and Lemma 71 respectively. This leads to the intermediate result

$$\lim_{N\to\infty} \frac{S_N^{b_n+1}(f)(x)}{N\log(N)} = 1 \quad \text{uniformly in } x \in [0, 1), \tag{3.29}$$

for α of Roth type, as stated in Proposition 75.

In order to prove Theorems 49 and 52, we will in both cases compare $S_N^{b_n+1}(f)$ to $S_N^1(f)$ and then make use of the convergence (3.29). In the case of Theorem 49 we note that, as in §3.3.4, a lightly trimmed weak law for the function $\frac{1}{x}$ already implies an untrimmed weak law.

In the last part of this section we will prove Theorems 53 and 55 making use of the previous results to prove the lim inf case and construct some sets for which the lim sup case holds.

Lemma 65. *For $n \geq 1$ and $x \in [0, 1)$ there is at most one $k^* \in \{1, ..., q_n\}$ such that*

$$R^{k^*}(x) \in \left[0, \frac{1}{q_n + q_{n-1}}\right).$$

Proof. For $k, k' \in \{1, ..., q_n\}$ with $k \neq k'$ and $y \in [0, 1)$ we have

$$d(R^k(y), R^{k'}(y)) \geq d(k\alpha, k'\alpha) \geq d(|k-k'|\alpha, 0) \geq d(q_{n-1}\alpha, 0) = \delta_n \geq \frac{1}{q_n + q_{n-1}}.$$

Hence there is at most one such point in the interval $\left[0, \frac{1}{q_n+q_{n-1}}\right)$. □

Denote

$$f_n(x) = \begin{cases} f(x) & \text{if } x \in \left[\frac{1}{q_n+q_{n-1}}, 1\right), \\ 0 & \text{if } x \in \left[0, \frac{1}{q_n+q_{n-1}}\right). \end{cases}$$

[32]This is to be understood in $[-\frac{1}{2}, \frac{1}{2})$.

Lemma 66. *For $N \geq 1$ we have*

$$S_N^{b_n+1}(f)(x) \leq S_N(f_n)(x) \leq S_N^{b_n+1}(f)(x) + 3N \quad \forall x \in [0,1).$$

Proof. By Lemma 65 there are at most b_n+1 points of the form $R^j(x)$ with $0 \leq j \leq N-1$ that are in $\left[0, \frac{1}{q_n+q_{n-1}}\right)$. Let $k' \leq b_n+1$ be such that $R^{j_{k'}^N(x)}(x) \in \left[0, \frac{1}{q_n+q_{n-1}}\right)$ but $R^{j_{k'+1}^N(x)}(x) \notin \left[0, \frac{1}{q_n+q_{n-1}}\right)$ then[33]

$$S_N^{b_n+1}(f)(x) = \sum_{i=b_n+2}^{N} f\big(R^{j_i^N(x)}(x)\big) \leq \sum_{i=k'+1}^{N} f\big(R^{j_i^N(x)}(x)\big) = S_N(f_n)(x).$$

On the other hand

$$\begin{aligned} S_N(f_n)(x) &= \sum_{i=k'+1}^{N} f\big(R^{j_i^N(x)}(x)\big) \leq \sum_{i=b_n+2}^{N} f\big(R^{j_i^N(x)}(x)\big) + (b_n+1-k')f((q_n+q_{n-1})^{-1}) \\ &\leq \sum_{i=b_n+2}^{N} f\big(R^{j_i^N(x)}(x)\big) + (b_n+1)(q_n+q_{n-1}) \leq S_N^{b_n+1}(f)(x) + 3N. \end{aligned}$$

□

Using the Denjoy-Koksma inequality, we can immediately deduce a useful estimate for $S_{q_n}^1(f)$.

Proposition 67. *For $x \in [0,1)$ and $n \geq 1$ it holds that*

$$\left| S_{q_n}^1(f)(x) - q_n \log(q_n) \right| \leq 7q_n.$$

Proof. By Denjoy-Koksma inequality we have

$$\left| S_{q_n}(f_n)(x) - q_n \int_0^1 f_n(y)\,\mathrm{d}y \right| \leq \mathrm{Var}(f_n),$$

and the claim follows from Lemma 66. □

We will use a more general form of this computation (see [BK21, Lemma 4.4]).

Lemma 68. *For $n \geq 1$ it holds that*

$$2q_{n+1} \geq \sum_{j=1}^{n} a_j q_j. \tag{3.30}$$

Proof. Recall that $q_{n+1} = a_n q_n + q_{n-1}$. For $n = 1$, (3.30) becomes $2q_2 \geq a_1 q_1$, which, considering that $q_1 = 1$ and $q_2 = a_1$, clearly holds. Now let $n \geq 2$ and assume (3.30) holds for all $k < n$, then

$$\sum_{j=1}^{n} a_j q_j \leq a_n q_n + a_{n-1} q_{n-1} + 2q_{n-1} \leq (a_n+1)q_n + 2q_{n-1} \leq 2q_{n+1}.$$

□

[33]Besides the points $R^{j_1^N(x)}(x), \ldots, R^{j_{k'}^N(x)}(x)$ there might additionally be one orbit point in $\left[0, \frac{1}{q_n+q_{n-1}}\right)$, namely in the case $R^{j_N^N(x)}(x) = 0$. However, even in this case, the calculation below remains true, since $f(0) = f_n(0) = 0$.

Lemma 69. *For each $n \geq 1, i \in [0, n-1]$ we have*

$$|S_{q_{n-i}}(f_n)(x) - q_{n-i}\log(q_{n-i})| \leq \min\left(2q_n, \frac{1}{x_{\min}^{q_{n-i}}}\right) + 3q_{n-i} \quad \forall x \in [0,1).$$

Proof. Denote $h = 1_{[\frac{1}{2q_{n-i}},1)} f_n$. Since any two points in $\{x, ..., R^{q_{n-i}-1}(x)\}$ are at distance at least $\delta_{n-i} = |\alpha q_{n-i-1}| \geq \frac{1}{2q_{n-i}}$, we have

$$\#\left(\{x, ..., R^{q_{n-i}-1}(x)\} \cap \left(0, \frac{1}{2q_{n-i}}\right)\right) \leq 1,$$

hence

$$|S_{q_{n-i}}(f_n)(x) - S_{q_{n-i}}(h)(x)| \leq f_n(x_{\min}^{q_{n-i}}) \leq \min\left(2q_n, \frac{1}{x_{\min}^{q_{n-i}}}\right).$$

Applying Denjoy-Koksma to h yields

$$|S_{q_{n-i}}(h)(x) - q_{n-i}\log(q_{n-i})| \leq \mathrm{Var}(h) + \log(2)q_{n-i} \leq 3q_{n-i},$$

and the claim follows. $\square$

Lemma 70. *If $N \in [q_n, q_{n+1})$ then*

$$\left|S_N(f_n)(x) - \sum_{j=1}^{n} b_j q_j \log(q_j)\right| \leq 16N + 2\sum_{1\leq j\leq n-1, b_j\neq 0} a_j q_j \log(b_j).$$

Proof. For $j = 1, ..., n-1$ and $s = 0, ..., b_j - 1$ denote $x(j,s) = R^{\sum_{i=j+1}^{n} b_i q_i + s q_j}(x)$, and $x(n,s) = R^{s q_n}(x)$ for $s = 0, ..., b_n - 1$. Then

$$S_N(f_n)(x) = \sum_{j=1}^{n}\sum_{s=0}^{b_j-1} S_{q_j}(f_n)(x(j,s)),$$

and by Lemma 69 we have

$$\begin{aligned}\left|S_N(f_n)(x) - \sum_{j=1}^{n} b_j q_j \log(q_j)\right| &\leq 3N + \sum_{j=1}^{n}\sum_{s=0}^{b_j-1}\min\left(2q_n, \frac{1}{(x(j,s))_{\min}^{q_j}}\right)\\ &\leq 3N + 2b_n q_n + \sum_{j=1}^{n-1}\sum_{s=0}^{b_j-1}\min\left(2q_n, \frac{1}{(x(j,s))_{\min}^{q_j}}\right)\\ &\leq 5N + \sum_{j=1}^{n-1}\sum_{s=0}^{b_j-1}\min\left(2q_n, \frac{1}{(x(j,s))_{\min}^{q_j}}\right). \end{aligned} \tag{3.31}$$

For $j = 1, ..., n$ and $s = 0, ..., b_j - 1$ denote $\tilde{x}(j,s) = (x(j,s))_{\min}^{q_j}$. For $j = 1, ..., n$ with $b_j \neq 0$ it holds that

$$\{R^{\sum_{l=j+1}^{n} b_l q_l}(x), R^{\sum_{l=j+1}^{n} b_l q_l + 1}(x), ..., R^{\sum_{l=j}^{n} b_l q_l - 1}(x)\} = \bigcup_{s=0}^{b_j-1}\{x(j,s), ..., R^{q_j-1}(x(j,s))\}$$

and the union is disjoint.

The set on the left is a finite orbit of length $b_j q_j < q_{j+1}$, therefore all of the points in the set on the right are at distance at least $\frac{1}{2q_{j+1}}$, in particular it holds that

$$|\tilde{x}(j,s) - \tilde{x}(j,s')| \geq \frac{1}{2q_{j+1}}, \tag{3.32}$$

for any $s, s' \in \{0, \ldots, b_j - 1\}$, $s \neq s'$.

Similarly, for $k = 1, ..., n-1$, we have

$$\{R^{\sum_{j=k+1}^{n} b_j q_j}(x), R^{\sum_{j=k+1}^{n} b_j q_j + 1}, ..., R^{N-1}(x)\} = \bigcup_{j=1}^{k} \bigcup_{s=0}^{b_j - 1} \{x(j,s), ..., R^{q_j - 1}(x(j,s))\}$$

and the union is disjoint.

The set on the left is a finite orbit of length $\sum_{j=1}^{k} b_j q_j < q_{k+1}$ and therefore, by Lemma 65, contains at most one point in $\left[0, \frac{1}{2q_{k+1}}\right)$. Letting (j_k^*, s_k^*) with $1 \leq j_k^* \leq k$ and $0 \leq s_k^* \leq b_{j_k^*} - 1$ be the index with $\tilde{x}(j_k^*, s_k^*) = \min_{j=1,...,k,\ s=0,...,b_j-1} \tilde{x}(j,s)$, it follows that[34]

$$\tilde{x}(j,s) \geq \frac{1}{2q_{k+1}} \quad j = 1, ..., k,\ s = 0, ..., b_j - 1,\ (j,s) \neq (j_k^*, s_k^*). \tag{3.33}$$

Let

$$\Delta^* = \{(j_1^*, s_1^*), ..., (j_{n-1}^*, s_{n-1}^*)\} \text{ and}$$
$$\Delta = \{(j,s) \mid 1 \leq j \leq n-1,\ 0 \leq s \leq b_j - 1,\ (j,s) \notin \Delta^*\}.$$

We will split up the sum in (3.31) as

$$\begin{aligned} \sum_{j=1}^{n-1} \sum_{s=0}^{b_j - 1} \min\left(2q_n, \frac{1}{(x(j,s))_{\min}^{q_j}}\right) &= \sum_{(j,s)\in\Delta} \min\left(2q_n, \frac{1}{(x(j,s))_{\min}^{q_j}}\right) \\ &+ \sum_{(j,s)\in\Delta^*} \min\left(2q_n, \frac{1}{(x(j,s))_{\min}^{q_j}}\right). \end{aligned} \tag{3.34}$$

For the first sum, notice that (3.33) implies

$$\tilde{x}(j,s) \geq \frac{1}{2q_{j+1}} \quad \forall (j,s) \in \Delta. \tag{3.35}$$

[34] It is possible that also $\tilde{x}(j_k^*, s_k^*) \geq \frac{1}{2q_{k+1}}$, we do not make any claims of these points.

Using (3.32), (3.35) and Lemma 68, we obtain

$$\begin{aligned}
&\sum_{(j,s)\in\Delta} \min\left(2q_n, \frac{1}{(x(j,s))_{\min}^{q_j}}\right) \\
&\quad\le \sum_{j=1}^{n-1}\sum_{s=0}^{b_j-1} \frac{1}{\frac{1}{2q_{j+1}}+\frac{s}{2q_{j+1}}} \\
&\quad\le 2\sum_{1\le j\le n-1, b_j\neq 0} q_{j+1}\left(\log(b_j)+1\right) \\
&\quad\le 2\sum_{1\le j\le n-1, b_j\neq 0} a_jq_j\left(\log(b_j)+1\right) + 2\sum_{1\le j\le n-1, b_j\neq 0} q_{j-1}\left(\log(b_j)+1\right) \\
&\quad\le 2\sum_{1\le j\le n-1, b_j\neq 0} a_jq_j\log(b_j) + 3\sum_{j=1}^{n-1} a_jq_j \\
&\quad\le 2\sum_{1\le j\le n-1, b_j\neq 0} a_jq_j\log(b_j) + 3N.
\end{aligned} \tag{3.36}$$

Note that the indices (j_k^*, s_k^*) may coincide for different k, in fact if $j_k^* = k-i$ for some $i \ge 1$, then $(j_k^*, s_k^*) = (j_{k-l}^*, s_{k-l}^*)$ for $l = 1, ..., i$. In other words, by (3.33), for each $1 \le k \le n-2$, it holds that

$$\tilde{x}(j_k^*, s_k^*) < \frac{1}{2q_{k+2}} \implies (j_k^*, s_k^*) = (j_{k+1}^*, s_{k+1}^*).$$

Using Lemma 68, it follows that

$$\sum_{(j,s)\in\Delta^*} \min\left(2q_n, \frac{1}{\tilde{x}(j,s)}\right) \le 2q_n + 2\sum_{k=1}^{n-2} q_{k+2} \le 4q_n + 2\sum_{k=1}^{n-1} q_k \le 8N. \tag{3.37}$$

From (3.31), (3.34), (3.36) and (3.37) it follows that

$$\left| S_N(f_n)(x) - \sum_{j=1}^{n} b_jq_j\log(q_j)\right| \le 16N + 2\sum_{1\le j\le n-1, b_j\neq 0} a_jq_j\log(b_j).$$

□

Recall that the first goal will be to show that

$$\lim_{N\to\infty} \frac{S_N^{b_n+1}(f)(x)}{N\log(N)} = 1 \quad \text{uniformly in } x \in [0,1),$$

if α is of Roth type. Considering Lemmas 66 and 70, this will follow, once we show

$$\sum_{1\le j\le n-1, b_j\neq 0} a_jq_j\log(b_j) = o(N\log(N)), \tag{3.38}$$

and

$$\lim_{N\to\infty} \frac{N\log(N)}{\sum_{j=1}^{n} b_jq_j\log(q_j)} = 1. \tag{3.39}$$

First we will show (3.38).

Lemma 71. *If* $\lim_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} = 1$, *then*

$$\sum_{j=1}^{n-1} a_j q_j \log(a_j) = o(q_n \log(q_n)).$$

Proof. Let $\epsilon > 0$, $n_0 \geq 1$ be so big that $q_{m+1} < q_m^{1+\epsilon}$ for all $m \geq n_0$ and thus in particular $a_m < q_m^{\epsilon}$. Moreover, let $n_1 \geq n_0 + 1$ be so big that $q_{n_1} > \frac{q_{n_0}}{\epsilon}$. Then, using Lemma 68, for $n \geq n_1$ it holds that

$$\begin{aligned}
\sum_{j=1}^{n-1} a_j q_j \log(a_j) &= \sum_{j=1}^{n_0-1} a_j q_j \log(a_j) + \sum_{j=n_0}^{n-1} a_j q_j \log(a_j) \\
&\leq \log(q_n) \sum_{j=1}^{n_0-1} a_j q_j + \epsilon \log(q_n) \sum_{j=n_0}^{n-1} a_j q_j \\
&\leq 2\log(q_n)(q_{n_0} + \epsilon q_n) \\
&\leq 2\epsilon q_n \log(q_n).
\end{aligned}$$

□

This immediately implies (3.38). The next lemmas will be a preparation for Proposition 74 which will give (3.39).

Lemma 72. *It holds that*

$$\sum_{j=1}^{n} b_j q_j \log(q_j) \leq N\log(N) \leq \sum_{j=1}^{n} b_j q_j \log(q_j) + N \log\left(\sum_{j=1}^{n} b_j\right).$$

Proof. Since $N = \sum_{j=1}^{n} b_j q_j$ it is clear that $\sum_{j=1}^{n} b_j q_j \log(q_j) \leq N\log(N)$. Now note that the function $z \mapsto z\log(z)$ is convex on $(0,\infty)$, hence

$$\sum_{j=1}^{n} \frac{b_j}{\sum_{i=1}^{n} b_i} q_j \log(q_j) \geq \left(\sum_{j=1}^{n} \frac{b_j}{\sum_{i=1}^{n} b_i} q_j\right) \log\left(\sum_{j=1}^{n} \frac{b_j}{\sum_{i=1}^{n} b_i} q_j\right) \geq \frac{N}{\sum_{i=1}^{n} b_i} \log\left(\frac{N}{\sum_{i=1}^{n} b_i}\right).$$

Multiplying this equation by $\sum_{i=1}^{n} b_i$ yields

$$\sum_{j=1}^{n} b_j q_j \log(q_j) \geq N\log(N) - N\log\left(\sum_{i=1}^{n} b_i\right).$$

□

Lemma 73. *If* $\lim_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} = 1$, *then*

$$\log\left(\sum_{j=1}^{n} b_j\right) = o(\log(N)).$$

Proof. Let $\epsilon \in (0,1)$ and $M' \geq 1$ be so big that $q_{m+1} < q_m^{1+\epsilon}$, which implies $b_m < q_m^{\epsilon}$, for $m > M'$. Let $M > M'$ be so big that $\sum_{j=1}^{M'} a_j < q_M^{\epsilon}$ and[35] $q_m > m^{\frac{1}{\epsilon}}$ for $m \geq M$. Then, for

[35]If $\alpha = \frac{1+\sqrt{5}}{2}$ then $q_n = \left\lfloor \frac{\alpha^n}{\sqrt{5}} \right\rceil$. For the golden ratio all the CFE coefficients are 1, so for general α we have $q_n \geq \left\lfloor \frac{\alpha^n}{\sqrt{5}} \right\rceil$.

$N \geq q_M$ it holds that

$$\log\left(\sum_{j=1}^{n} b_j\right) \leq \log\left(\sum_{j=M'}^{n} q_j^{\epsilon}\right) + \log(q_M^{\epsilon})$$
$$\leq \log(q_n^{\epsilon}) + \log(n) + \log(q_M^{\epsilon}) \leq 4\epsilon \log(N).$$

□

Proposition 74. *It holds that*

$$\lim_{N\to\infty} \frac{N\log(N)}{\sum_{j=1}^{n} b_j q_j \log(q_j)} = 1 \tag{3.40}$$

if and only if $\lim_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} = 1$.

Proof. (i) If $\lim_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} = 1$, then (3.40) holds by Lemmas 72 and 73.

(ii) On the other hand, if $\liminf_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} > 1$ then there is an $\epsilon > 0$ and a subsequence $(n_l)_{l\geq 1}$ such that $q_{n_l+1} > q_{n_l}^{1+\epsilon}$, which implies $a_{n_l} > q_{n_l}^{\epsilon/2}$, for l sufficiently large. For $N_l = a_{n_l} q_{n_l} \in [q_{n_l}, q_{n_l+1} - 1]$ and l sufficiently large it holds that[36]

$$N_l \log(N_l) > \left(1 + \frac{\epsilon}{2}\right) a_{n_l} q_{n_l} \log(q_{n_l}),$$

contradicting (3.40). □

Proposition 75. *If* $\lim_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} = 1$, *then it holds that*

$$\lim_{N\to\infty} \frac{S_N^{b_n+1}(f)(x)}{N\log(N)} = 1 \quad \textit{uniformly in } x \in [0,1).$$

Proof. This is an immediate consequence of Lemmas 66 and 70, (3.38) and Proposition 74. □

Now we will show that, assuming $\lim_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} = 1$, weak laws can be obtained even without trimming. As an intermediate step, we first prove weak laws under light trimming.

Lemma 76. *If* $\lim_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} = 1$ *then*

$$\lambda\left(\left|\frac{S_N^1(f)}{N\log(N)} - 1\right| > \epsilon\right) \to 0 \quad \textit{as } n \to \infty \quad \forall \epsilon > 0. \tag{3.41}$$

Proof. Let $\epsilon \in \left(0, \frac{1}{100}\right)$ and M_0 be big enough such that $\frac{\log(q_{n+1})}{\log(q_n)} < 1 + \frac{\epsilon^2}{10}$ for all $n \geq M_0$. For $n \geq M_0$ denote $\epsilon_n = 10\frac{1}{\epsilon}\left(\frac{\log(q_{n+1})}{\log(q_n)} - 1\right) < \epsilon$, and let

$$B_n = \left\{ x \in [0,1) \;\middle|\; d(x + j\alpha, \mathbb{Z}) \leq \frac{\epsilon_n}{q_n} \text{ for some } j \in \{0, ..., q_n - 1\} \right\}$$
$$= \bigcup_{j=0}^{q_n-1} \left[\left(-j\alpha - \frac{\epsilon_n}{q_n}\right) \mod 1, \left(-j\alpha + \frac{\epsilon_n}{q_n}\right) \mod 1\right],$$

[36] In this case $a_{n_l} = b_{n_l}$.

note that $\lambda(B_n) \le 2\epsilon_n$. Note furthermore that $a_n \le \frac{q_{n+1}}{q_n} = q_n^{-1+\frac{\log(q_{n+1})}{\log(q_n)}}$, taking logarithms we obtain

$$\log(a_n) \le \frac{\epsilon\epsilon_n}{10}\log(q_n) \le \frac{\epsilon\epsilon_n}{2}\log(N). \tag{3.42}$$

We will show that for big enough n it holds that

$$S_N^1(f)(x) - S_N^{b_n+1}(f)(x) \le \epsilon N\log(N) + N \quad \forall N \in [q_n, q_{n+1}-1],\ x \notin B_n. \tag{3.43}$$

Considering Proposition 75, (3.41) follows from (3.43). Note that, if $a_n < 10$, or equivalently $q_{n+1} < 10q_n + q_{n-1}$, then (3.43) holds trivially, even without any assumption on x, because

$$S_N^1(f)(x) - S_N^{b_n+1}(f)(x) \le \sum_{i=2}^{b_n} f(R^{j_i^N(x)}(x)) \le 10q_{n+1} \le 110N.$$

So for the rest of the proof we can focus on the case $a_n \ge 10$.

We will only show (3.43), for n with $\alpha - \frac{p_n}{q_n} > 0$. The conclusion for all other n follows by considering $1-\alpha$ instead[37] .

We first claim that, for $x \notin B_n$ and big enough n, whenever $q_n \le N < q_n + \epsilon_n q_{n+1}$, it holds that

$$j_1^{q_n}(x) = j_1^N(x). \tag{3.44}$$

Indeed, since $b_n \le 1 + \frac{\epsilon_n q_{n+1}}{q_n}$, keeping in mind that $\epsilon_n < \frac{1}{100}$, and identifying $\mathbb{T}$ with $\left[-\frac{1}{2}, \frac{1}{2}\right)$, it holds that

$$R^{j_{q_n}^{q_n}(x)}(x) \le -\delta_n + \frac{\epsilon_n}{q_n} \le -\frac{1}{2q_n} + \frac{1}{100q_n} \le -\frac{1}{q_{n+1}} - \frac{1}{100q_n} \le -b_n\delta_{n+1},$$

where we recall that $\delta_n = d(q_{n-1}\alpha, \mathbb{Z}) \in \left(\frac{1}{q_n+q_{n-1}}, \frac{1}{q_n}\right)$. Now (3.44) follows by applying Lemma 64. Therefore, in this case, $x_{\min}^N = x_{\min}^{q_n} \ge \frac{\epsilon_n}{q_n}$ and using (3.42) we obtain

$$S_N^1(f)(x) - S_N^{b_n+1}(f)(x) \le b_n\frac{q_n}{\epsilon_n} \le \frac{N}{\epsilon_n} \le \epsilon N\log(N).$$

On the other hand, if $N \ge q_n + \epsilon_n q_{n+1}$, then it is not necessarily true that $j_1^{q_n}(x) = j_1^N(x)$. Nevertheless, using (3.42), we can estimate

$$\begin{aligned} S_N^1(f)(x) - S_N^{b_n+1}(f)(x) &= \sum_{i=2}^{b_n+1} f(R^{j_i^N(x)}(x)) \le \sum_{i=1}^{b_n}\frac{q_{n+1}}{i} \\ &\le q_{n+1}(1+\log(b_n)) \le \frac{1}{\epsilon_n}N(1+\log(b_n)) \le \epsilon N\log(N) + N, \end{aligned}$$

and we have shown (3.43) in both cases. □

[37]Let n be such that $\alpha - \frac{p_n}{q_n} < 0$ and denote $\tilde{\alpha} = 1-\alpha$. Then the CFE approximants of $\tilde{\alpha}$ are given by $\frac{\tilde{p}_{\tilde{n}}}{\tilde{q}_{\tilde{n}}} = \frac{q_n - p_n}{q_n}$ (the index $\tilde{n}$ is either $n-1$ or $n+1$, depending on whether $\alpha < \frac{1}{2}$) and we have $\tilde{\alpha} - \frac{\tilde{p}_{\tilde{n}}}{\tilde{q}_{\tilde{n}}} > 0$. Denote by $\tilde{S}_N$ the ergodic sum w.r.t. the rotation $\tilde{R}(x) = x + \tilde{\alpha}$, and note that for $N \in [q_n, q_{n+1}-1] = [\tilde{q}_{\tilde{n}}, \tilde{q}_{\tilde{n}+1}-1]$ it holds that $S_N^k = \tilde{S}_N^k \circ R^N$ for all $k \le N$. Since $\tilde{\alpha} - \frac{\tilde{p}_{\tilde{n}}}{\tilde{q}_{\tilde{n}}} > 0$, (3.43) yields

$$\tilde{S}_N^1(f)(x) - \tilde{S}_N^{\tilde{b}_{\tilde{n}}+1}(f)(x) \le \epsilon N\log(N) \quad \forall N \in [\tilde{q}_{\tilde{n}}, \tilde{q}_{\tilde{n}+1}-1], x \notin \tilde{B}_{\tilde{n}},$$

where $\tilde{b}_{\tilde{n}}$ and $\tilde{B}_{\tilde{n}}$ are defined in the same way as b_n and B_n, only using $\tilde{q}_{\tilde{n}}$ resp. $\tilde{q}_{\tilde{n}+1}$ instead of q_n resp. q_{n+1}. Since $\tilde{q}_{\tilde{n}} = q_n$ and $\tilde{q}_{\tilde{n}+1} = q_{n+1}$ it holds that $\tilde{b}_{\tilde{n}} = b_n$ and $\tilde{B}_{\tilde{n}} = B_n$, it follows that

$$S_N^1(f)(x) - S_N^{b_n+1}(f)(x) \le \epsilon N\log(N) \quad \forall N \in [q_n, q_{n+1}-1], x \notin R^N(B_n).$$

We will use an argument similar to this one several times in the sequel.

Lemma 77. *Suppose that, for some $\gamma \in (0,1]$ and $n \in \mathbb{N}$, it holds that $q_{n+1} > q_n^{1+\gamma}$ and $q_n^\gamma > \gamma^{-1}1000$. Then there are sets $A, B \subset [0,1)$ with $\lambda(A) = \lambda(B) = \frac{\gamma}{1000}$ such that, whenever $k \leq q_{n+1} q_n^{-\left(1+\frac{\gamma}{2}\right)}$,*

$$S_N^k(f)(x) > \frac{\gamma}{4} q_{n+1}\log(q_n), \quad \textit{while} \quad S_N^k(f)(y) < \frac{\gamma}{100} q_{n+1}\log(q_n) \quad \forall x \in A, y \in B,$$

where $N = \lceil \frac{\gamma q_{n+1}}{250} \rceil$.

Proof. We only treat the case for $\alpha - \frac{p_n}{q_n} > 0$, the other case follows by considering $1-\alpha$ instead. Let

$$A = \left\{ x \in [0,1) \;\middle|\; R^{j_{q_n}^{q_n}(x)}(x) \in \left(-\frac{\gamma}{1000 q_n}, 0\right) \right\} = \bigcup_{j=0}^{q_n-1} \left(-\frac{\gamma}{1000 q_n}, 0\right) - j\alpha,$$

and

$$\begin{aligned} B &= \left\{ x \in [0,1) \;\middle|\; R^{j_{q_n}^{q_n}(x)}(x) \in \left(-\left(\frac{1}{4} + \frac{\gamma}{1000}\right)\frac{1}{q_n}, -\frac{1}{4q_n}\right) \right\} \\ &= \bigcup_{j=0}^{q_n-1} \left(-\left(\frac{1}{4} + \frac{\gamma}{1000}\right)\frac{1}{q_n}, -\frac{1}{4q_n}\right) - j\alpha. \end{aligned}$$

Clearly $\lambda(A) = \lambda(B) = \frac{\gamma}{1000}$.

For $x \in B$ we have

$$R^{j_{q_n}^{q_n}(x)}(x) \leq -\frac{1}{4q_n} \leq -\left\lceil \frac{\gamma q_{n+1}}{250 q_n} \right\rceil \frac{1}{q_{n+1}} \leq -b_n \delta_{n+1},$$

and Lemma 64 yields $j_1^{q_n}(x) = j_1^N(x)$. Moreover, since different points in $\{x, ..., R^{q_n-1}(x)\}$ are at distance at least $\frac{1}{2q_n}$, it holds that $R^{j_1^{q_n}(x)}(x) \geq \frac{1}{8q_n}$, therefore $S_N^k(f)(x) = S_N^k(f')(x)$, where

$$f'(z) = \begin{cases} 0 & \text{if } z \in [0, \frac{1}{8q_n}), \\ f(z) & \text{otherwise.} \end{cases}$$

Using the Denjoy-Koksma inequality, we obtain

$$S_N^k(f)(x) \leq S_{(b_n+1)q_n}(f')(x) \leq 2(b_n+1)q_n\log(q_n) + 8(b_n+1)q_n \leq \frac{\gamma}{100} q_{n+1}\log(q_n).$$

On the other hand, since $\alpha - \frac{p_n}{q_n} > 0$, applying R^{q_n} will move any point to the right by a distance of $\delta_{n+1} = |q_n\alpha| > \frac{1}{2q_{n+1}}$. Therefore, if $x \in A$ then it holds that

$$R^{j_{q_n}^{q_n}(x) + \lceil \frac{b_n}{2} \rceil q_n}(x) \geq -\frac{\gamma}{1000 q_n} + \left\lceil \frac{b_n}{2} \right\rceil \delta_{n+1} \geq -\frac{\gamma}{1000 q_n} + \frac{\gamma q_{n+1}}{500 q_n}\frac{1}{2q_{n+1}} \geq 0,$$

where we identify the circle with $\left[-\frac{1}{2}, \frac{1}{2}\right)$. Since now the points

$$R^{j_{q_n}^{q_n}(x)}(x), R^{j_{q_n}^{q_n}(x)+q_n}(x), ..., R^{j_{q_n}^{q_n}(x)+\lceil \frac{b_n}{2} \rceil q_n}(x)$$

are at distance $\delta_{n+1} < \frac{1}{q_{n+1}}$ and cross from negative to positive, it follows that there is an $s^* \in \left\{1, ..., \lceil \frac{b_n}{2} \rceil\right\}$ such that $R^{j_{q_n}^{q_n}(x)+s^* q_n}(x) \in \left(0, \frac{1}{q_{n+1}}\right)$. There could be two such points in the interval, in that case choose the left one[38] so that $j_{q_n}^{q_n}(x) + s^* q_n = j_1^N(x)$. The points

$$\{R^{j_{q_n}^{q_n}(x)+(s^*+k)q_n}(x), ..., R^{j_{q_n}^{q_n}(x)+(b_n-1)q_n}(x)\}$$

[38]For the calculation below it is not important which we choose.

have k other points to their left, therefore $S_N^k(f)(x)$ can be estimated from below by summing f only over those points. More precisely

$$
\begin{aligned}
S_N^k(f)(x) &> \sum_{j=s^*+k}^{b_n-1} f(R^{j_{q_n}^{q_n}(x)+jq_n}(x)) > \sum_{j=k}^{\lfloor\frac{b_n-1}{2}\rfloor} f\left(\frac{j+1}{q_{n+1}}\right) \\
&= q_{n+1} \sum_{j=k}^{\lfloor\frac{b_n-1}{2}\rfloor} \frac{1}{j+1} \geq q_{n+1}\left(\log\left(\left\lfloor\frac{b_n-1}{2}\right\rfloor\right) - \log(k)\right) \\
&\geq \frac{1}{2} q_{n+1}\left(\log\left(\frac{q_{n+1}}{2q_n}\right) - \log\left(q_{n+1} q_n^{-1-\frac{\gamma}{2}}\right)\right) \geq \frac{\gamma}{4} q_{n+1}\log(q_n).
\end{aligned}
$$

□

Now we are in a position to prove Theorem 49.

Proof of Theorem 49. First assume that $\lim_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} = 1$, then, by Lemma 76

$$
\lambda\left(\left|\frac{S_N^1(f)(x)}{N\log(N)} - 1\right| > \epsilon\right) \to 0 \quad \text{as } n \to \infty \quad \forall \epsilon > 0. \tag{3.45}
$$

For $N \geq 1$ let $M_N = S_N(f) - S_N^1(f) = \max_{i=0,\dots,N-1} f \circ R^i$, then, for $\epsilon > 0$, we have

$$
\begin{aligned}
\lambda\left(\left|\frac{S_N(f)(x)}{N\log(N)} - 1\right| > \epsilon\right) &\leq \lambda\left(M_N(x) > \frac{\epsilon}{2} N\log(N)\right) + \lambda\left(\left|\frac{S_N^1(f)(x)}{N\log(N)} - 1\right| > \frac{\epsilon}{2}\right) \\
&\leq \lambda\left(x \;\middle|\; \exists i \in \{0,\dots,N-1\} \text{ s.t. } R^i(x) \in \left(0, \left(\frac{\epsilon}{2}N\log(N)\right)^{-1}\right)\right) \\
&\quad + \lambda\left(\left|\frac{S_N^1(f)(x)}{N\log(N)} - 1\right| > \frac{\epsilon}{2}\right) \\
&\leq \frac{2}{\epsilon\log(N)} + \lambda\left(\left|\frac{S_N^1(f)(x)}{N\log(N)} - 1\right| > \frac{\epsilon}{2}\right).
\end{aligned}
$$

In light of (3.45), it follows that

$$
\lambda\left(\left|\frac{S_N(f)(x)}{N\log(N)} - 1\right| > \epsilon\right) \to 0 \quad \text{as } n \to \infty \quad \forall \epsilon > 0.
$$

On the other hand, now suppose $\limsup_{n\to\infty} \frac{\log(q_{n+1})}{\log(q_n)} > 1$. Taking a subsequence $(n_l)_{l\geq 1}$ we may assume that there is a $\gamma \in (0,1)$ such that $q_{n_l+1} > q_{n_l}^{1+\gamma}$ and $q_{n_l}^{\gamma} > \gamma^{-1}1000$. For a contradiction assume that there are $d_N > 0$ such that

$$
\lambda\left(\left|\frac{S_N(f)(x)}{d_N} - 1\right| > \epsilon\right) \to 0 \quad \text{as } N \to \infty \quad \forall \epsilon > 0.
$$

For every $l \geq 1$ let $N_l = \left\lceil \frac{\gamma q_{n_l+1}}{250} \right\rceil$, by Lemma 77 there are sets A_l, B_l with $\lambda(A_l) = \lambda(B_l) = \frac{\gamma}{1000}$ such that

$$
S_{N_l}(f)(x) > \frac{\gamma}{4} q_{n_l+1}\log(q_{n_l}) \quad \forall x \in A_l, \tag{3.46}
$$

while

$$
S_{N_l}(f)(y) < \frac{\gamma}{100} q_{n_l+1}\log(q_{n_l}) \quad \forall y \in B_l. \tag{3.47}
$$

Due to (3.46), we must have $d_{N_l} \geq \frac{\gamma}{8} q_{n_l+1}\log(q_{n_l})$, on the other hand (3.47) implies $d_{N_l} \leq \frac{\gamma}{50} q_{n_l+1}\log(q_{n_l})$. This is a contradiction. □

Proof of Remark 50. Let $l(N) = o(N)$ be a sequence of natural numbers and $k(N) \leq l(N)$. For $\epsilon > 0$ denote

$$u(\epsilon) = \min(N \geq 1 \mid l(m) < \epsilon m, \ \forall m \geq N).$$

Let α be such that,

$$q_{n+1} > q_n^2 u\left(\frac{1}{q_n^2}\right) \quad \text{for all } n \text{ sufficiently large,}$$

noting that there is a G_δ dense set of α satisfying this condition. Moreover, this choice implies that

$$k(N) \leq l(N) \leq q_{n+1} q_n^{-\frac{3}{2}},$$

whenever $N = \lceil \frac{q_{n+1}}{250} \rceil$ and $\sqrt{q_n} > 250$. The claim follows from Lemma 77 with $\gamma = 1$. $\square$

We will now focus on the claims about strong convergence, i.e. Theorems 52, 53, and 55.

Proof of Theorem 52. Let $\kappa = \frac{1}{8}$ and α be such that $q_{n+1} < q_n \log(q_n) \log^2(\log(q_n))$ and

$$\sum_{n \geq 1, q_{n+1} \geq q_n \log^{1-\frac{\kappa}{2}}(q_n)} \frac{1}{\log^{1-\kappa}(q_n)} < \infty, \tag{3.48}$$

by[39] [FK16], this condition is satisfied for almost all α. Denote

$$\mathcal{N} = \{n \geq 1 \mid q_{n+1} < q_n \log^{1-\frac{\kappa}{2}}(q_n)\}.$$

Clearly, α fulfils the Roth type condition and thus, Proposition 75 shows that

$$\lim_{N \to \infty} \frac{S_N^{b_n+1}(f)(x)}{N \log(N)} = 1 \quad \text{uniformly in } x \in [0, 1).$$

The claim (3.12) follows once we show that, for almost every x, it holds that

$$S_N^1(f)(x) - S_N^{b_n+1}(f)(x) = o(N \log(N)). \tag{3.49}$$

We estimate the left side in two different ways;

- since any two distinct points in $\{x, ..., R^{N-1}(x)\}$ are at distance at least $\frac{1}{2q_{n+1}}$, we have

$$S_N^1(f)(x) - S_N^{b_n+1}(f)(x) \leq \sum_{j=1}^{b_n} \frac{2q_{n+1}}{j} \leq 10 q_{n+1} \max(1, \log(b_n)), \tag{3.50}$$

- on the other hand, denoting $x_{\min}^N = R^{j_1^N(x)}(x)$, we can trivially estimate

$$S_N^1(f)(x) - S_N^{b_n+1}(f)(x) \leq b_n \frac{1}{x_{\min}^N}. \tag{3.51}$$

Since $b_n < \log(q_n) \log^2(\log(q_n))$, we have;

[39]In [FK16] it is shown that the stronger condition

$$\sum_{n \geq 1, q_{n+1} \geq q_n \log^{1-\kappa}(q_n)} \frac{1}{\log^{1-\kappa}(q_n)} < \infty,$$

holds for almost all α.

(A) if $n \in \mathcal{N}$ sufficiently large, then by (3.50) it holds that

$$\begin{aligned} S_N^1(f)(x) - S_N^{b_n+1}(f)(x) &\leq 10 q_{n+1} \max(1, \log(b_n)) \\ &\leq 20 q_n \log^{1-\frac{\kappa}{2}}(q_n) \log_2(q_n) = o(N \log(N)), \end{aligned}$$

(B) if $N > q_n \log^{\frac{\kappa}{2}}(q_n)$, then by (3.50) for n sufficiently large it holds that

$$\begin{aligned} S_N^1(f)(x) - S_N^{b_n+1}(f)(x) &\leq 10 q_{n+1} \max(1, \log(b_n)) \\ &\leq 20 q_n \log(q_n) \log^3(\log(q_n)) \leq 100 N \log^{1-\frac{\kappa}{4}}(N) = o(N \log(N)), \end{aligned}$$

(C) if $x_{\min}^N \geq \frac{1}{q_n \log^{1-\kappa}(q_n)}$, then by (3.51) for n sufficiently large it holds that

$$S_N^1(f)(x) - S_N^{b_n+1}(f)(x) \leq q_n \log^{1-\kappa}(q_n) \log^2(\log(q_n)) = o(N \log(N)).$$

In the next steps we will show that, indeed, for almost all x and all but finitely many $n \notin \mathcal{N}$, if $N \geq q_n \log^{\frac{\kappa}{2}}(q_n)$ then we have $x_{\min}^N \geq \frac{1}{q_n \log^{1-\kappa}(q_n)}$ and we have covered all cases.

For each $n \notin \mathcal{N}$ with $\alpha - \frac{p_n}{q_n} > 0$ let $\epsilon_n = \log^{-(1-\kappa)}(q_n)$ and

$$B_n = \left\{ x \in [0,1) \;\middle|\; d(x + j\alpha, \mathbb{Z}) \leq \frac{\epsilon_n}{q_n} \;\text{ for some } j \in \{0, ..., q_n - 1\} \right\} = \bigcup_{j=0}^{q_n-1} \left[\frac{-\epsilon_n}{q_n}, \frac{\epsilon_n}{q_n} \right] - j\alpha.$$

For $n \notin \mathcal{N}$ with $\alpha - \frac{p_n}{q_n} > 0$, $x \notin B_n$, and $N \leq q_n \log^{\frac{\kappa}{2}}(q_n)$ we have

$$R^{j_{q_n}^{q_n}(x)}(x) < -q_n^{-1} \log^{-(1-\kappa)}(q_n) \leq -b_n q_{n+1}^{-1},$$

and by Lemma 64 it follows that $j_1^{q_n}(x) = j_1^N(x)$. In this case

$$x_{\min}^N \geq q_n^{-1} \log^{-(1-\kappa)}(q_n).$$

Similarly, for $n \notin \mathcal{N}$, $\alpha - \frac{p_n}{q_n} < 0$ and

$$B_n' = \bigcup_{j=0}^{q_n-1} \left[0, \frac{2\epsilon_n}{q_n} \right] - j\alpha,$$

it also holds that

$$x_{\min}^N \geq q_n^{-1} \log^{-(1-\kappa)}(q_n) \quad \forall x \notin B_n'.$$

Indeed, if $x \notin B_n'$, $\alpha - \frac{p_n}{q_n} < 0$, $n \notin \mathcal{N}$ and $N \leq q_n \log^{\frac{\kappa}{2}}(q_n)$, then

$$j_1^N(x) = \begin{cases} j_1^{q_n}(x) + b_n q_n & \text{if } j_1^{q_n}(x) + b_n q_n \leq N, \\ j_1^{q_n}(x) + (b_n - 1) q_n & \text{otherwise,} \end{cases}$$

therefore, under the just stated conditions, it holds that

$$x_{\min}^N \geq 2 q_n^{-1} \log^{-(1-\kappa)}(q_n) - b_n q_{n+1}^{-1} \geq q_n^{-1} \log^{-(1-\kappa)}(q_n).$$

Since $\lambda(B_n \cup B_n') = 3 \log^{-(1-\kappa)}(q_n)$, the assumption (3.48) shows that $\sum_{n \notin \mathcal{N}} \lambda(B_n \cup B_n') < \infty$, and by Borel-Cantelli almost every x is only in finitely many B_n or B_n' with $n \notin \mathcal{N}$. By (A)-(C) it follows that (3.49) holds for almost every x.

For the "furthermore" part of the statement, note that if α is of bounded type then $\mathcal{N} = \mathbb{N} \setminus K$ for a finite set $K \subset \mathbb{N}$. Hence, case (C), the only case where we don't have a uniform statement, does not have to be considered. $\square$

Note that, if α is not of bounded type, the proof uses a Borel-Cantelli argument, hence convergence typically is not uniform. For example we shall show that, for almost all α,

$$\frac{S_N^1(f)(0)}{N\log(N)}$$

does not converge.

Lemma 78. *If* $q_{n+1} \in (q_n \log(q_n)\log_3(q_n)\,,\, q_n \log^2(q_n))$, $q_n > 10^6$, $\alpha - \frac{p_n}{q_n} > 0$ *and* $x_{\min}^{q_n} < \frac{1}{q_n \log(q_n)\log_3(q_n)}$, *then it holds that*

$$S_N^1(f)(x) - S_N^{b_n+1}(f)(x) \geq \frac{1}{2} N\log(N),$$

where $N = \lceil \frac{q_{n+1}\log_5(q_n)}{\log(q_n)} \rceil$.

Proof. Denote $\epsilon = \frac{1}{q_n \log(q_n)\log_3(q_n)}$, for n sufficiently large, it holds that

$$\begin{aligned}
S_N^1(f)(x) - S_N^{b_n+1}(f)(x) &\geq \sum_{i=1}^{b_n} \frac{1}{\epsilon + \frac{i}{q_{n+1}}} \geq \sum_{i=1+\lceil \epsilon q_{n+1}\rceil}^{b_n+\lceil \epsilon q_{n+1}\rceil} \frac{q_{n+1}}{i} \\
&\geq \frac{1}{2} q_{n+1} \log\left(\frac{b_n}{\epsilon q_{n+1}}\right) \geq \frac{1}{2} N \frac{\log(q_n)}{\log_5(q_n)} \log_4(q_n) \geq \frac{1}{2} N\log(N).
\end{aligned}$$

□

Remark 79. *The "*$\log_3$*" in the previous lemma might seem a bit odd, and in fact, the lemma works also if we replace* $\log_3$ *by any function* ψ *with* $\psi(n) \nearrow \infty$. *The reason for our choice will be apparent from the proofs of Theorems 53 and 55.*

Lemma 80. *For almost every* α *there are infinitely many* $n \geq 1$ *such that*

$$q_{n+1} > q_n \log(q_n)\log_3(q_n) \qquad \text{and } \alpha - \frac{p_n}{q_n} > 0.$$

Proof. To avoid confusion, in this proof we shall write $q_n(\alpha)$ resp. $a_n(\alpha)$ instead of q_n resp. a_n. We have $\sum_{n\geq 1} \frac{1}{n\log(n)\log_3(n)} = \infty$ and hence by Khinchine's Theorem, for almost every α, there are infinitely many n with

$$q_{n+1}(\alpha) > q_n(\alpha)\log(q_n(\alpha))\log_3(q_n(\alpha)).$$

Equivalently,[40]

$$a_n(\alpha) > \log(q_n(\alpha))\log_3(q_n(\alpha)).$$

Expanding α as

$$\alpha = \cfrac{1}{a_1(\alpha) + \cfrac{1}{a_2(\alpha) + \cfrac{1}{\dots}}}, \tag{3.52}$$

[40] This is not quite equivalent since $q_{n+1} = a_n q_n + q_{n-1}$, but almost. Using Khinchine also yields $q_{n+1}(\alpha) > 2q_n(\alpha)\log(q_n(\alpha))\log_3(q_n(\alpha))$ for almost all α, which implies $a_n(\alpha) > \log(q_n(\alpha))\log_3(q_n(\alpha))$.

we can easily see that $\alpha - \frac{p_n(\alpha)}{q_n(\alpha)} > 0$ if and only if n is even. Denote

$$A_{even} = \left\{ \alpha \in (0,1) \setminus \mathbb{Q} \;\middle|\; \begin{array}{l} \text{there are infinitely many even } n \text{ with} \\ a_n(\alpha) > \log(q_n(\alpha)) \log_3(q_n(\alpha)) \end{array} \right\},$$

$$A_{odd} = \left\{ \alpha \in (0,1) \setminus \mathbb{Q} \;\middle|\; \begin{array}{l} \text{there are infinitely many odd } n \text{ with} \\ a_n(\alpha) > \log(q_n(\alpha)) \log_3(q_n(\alpha)) \end{array} \right\}.$$

Let $T(x) = \lfloor \frac{1}{x} \rfloor$ be the Gauss-map. From (3.52) it follows that

$$a_n(T(\alpha)) = a_{n+1}(\alpha) \text{ and therefore } q_n(T(\alpha)) \leq q_{n+1}(\alpha),$$

hence $A_{odd} \subset T^{-1}(A_{even})$ and $A_{even} \subset T^{-1}(A_{odd})$. It follows that $A_{even} \subset T^{-2}(A_{even})$. Since T preserves the Gauss-measure μ, which is equivalent to Lebesgue, we have $A_{even} = T^{-2}(A_{even})$ (mod μ). And since T^2 is ergodic[41] $\mu(A_{even}) \in \{0,1\}$. From

$$A_{even} \subset T^{-1}(A_{odd}) \subset T^{-2}(A_{even}) = A_{even} \text{ (mod } \mu),$$

it follows that $A_{even} = T^{-1}(A_{odd})$ (mod μ), hence $\mu(A_{odd}) = \mu(A_{even}) \in \{0,1\}$, and by equivalence to Lebesgue also $\lambda(A_{odd}) = \lambda(A_{even}) \in \{0,1\}$. At the same time, by Khinchine it holds that $\lambda(A_{even} \cup A_{odd}) = 1$, it follows that $\lambda(A_{odd}) = \lambda(A_{even}) = 1$ as desired. □

Proof of Theorem 53. Proposition 67 ascertains that $|S^1_{q_n}(f)(x) - q_n \log(q_n)| \leq 7q_n$ and the "$\liminf$" part of the claim holds.

By Lemma 80, for almost all α, there are infinitely many n with $q_{n+1} > q_n \log(q_n) \log_3(q_n)$ and $\alpha - \frac{p_n}{q_n} > 0$. For all such n and $x = \alpha$ we have

$$x^{q_n}_{\min} = R^{q_n}(0) < \frac{1}{q_{n+1}} < \frac{1}{q_n \log(q_n) \log_3(q_n)},$$

and by Lemma 78 there is an $N_n \in [q_n, q_{n+1} - 1]$ such that

$$S^1_{N_n}(f)(x) - S^{b_n+1}_{N_n}(f)(x) \geq \frac{1}{2} N_n \log(N_n). \tag{3.53}$$

The claim follows from (3.53) together with Proposition 75. □

Theorem 55 can be proven using the same ideas. Here we will need a version of the second Borel-Cantelli Lemma, using only weak pairwise independence. The following has been proven in [Lam63], see also [Pet02] or [Cha08].

Lemma 81 ([Lam63]). *Let $(\Omega, \mathbb{P})$ be a probability space and $(A_n)_{n\geq 1}$ be a sequence of events with*

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$$

and there is a $K > 0$ such that

$$\mathbb{P}(A_i \cap A_j) \leq K \mathbb{P}(A_i) \mathbb{P}(A_j) \quad \forall i < j.$$

Then it holds that

$$\mathbb{P}(\omega \in \Omega \mid \omega \in A_n \textit{ for infinitely many } n) > 0.$$

[41] A fortiori T is mixing.

Proof of Theorem 55. Let α be such that $q_{n+1} \in (q_n \log(q_n)\log_3(q_n), q_n \log^2(q_n))$ for all n. The claim about the weak law follows from Theorem 49. We will now show that the strong law does not hold. For simplicity, we only consider n for which $\alpha - \frac{p_n}{q_n} > 0$, the other case follows by considering $1-\alpha$ instead.

(i) For the "lim inf" part of the claim, note that by Proposition 67 we always have $|S^1_{q_n}(f)(x) - q_n\log(q_n)| \leq 7q_n$. Furthermore, since the points $\{x, ..., R^{q_n-1}(x)\}$ are at distance at least $\delta_n > \frac{1}{2q_n}$, it holds that $R^{j^{q_n}_{i+1}(x)}(x) \geq \frac{i}{2q_n}$. It follows that

$$|S^K_{q_n}(f)(x) - q_n \log(q_n)| \leq (7 + 2\log(K))q_n \quad \text{if } q_n > K \geq 1.$$

(ii) Let $N_n = \lceil \frac{q_{n+1}\log_5(q_n)}{\log(q_n)} \rceil$ and G be the set

$$G = \Big\{ x \mid \limsup_{n\to\infty} \frac{S^K_{N_n}(f)(x)}{N_n \log(N_n)} > 1 \Big\}.$$

Note that [42]

$$|S^K_{N_n}(f)(x) - S^K_{N_n}(f)(R(x))| = o(N_n \log(N_n)) \quad \text{uniformly in } x,$$

therefore $G = R(G)$ is invariant. By ergodicity, the claim (3.13) will follow once we show $\lambda(G) > 0$.

(iii) Let $\epsilon_n = \frac{1}{q_n \log(q_n)\log_3(q_n)}$ and

$$A_n = \{x \mid R^{j^{q_n}_1(x)}(x) \in (0, \epsilon_n)\} = \bigcup_{l=0}^{q_n-1} (0, \epsilon_n) - l\alpha,$$

and this union is disjoint.[43] Note that $\frac{b_n}{\epsilon_n q_{n+1}} \geq \frac{\log_3(q_n)\log_5(q_n)}{2}$. Then for $x \in A_n$

$$\begin{aligned} S^K_{N_n}(f)(x) - S^{b_n+1}_{N_n}(f)(x) &\geq \sum_{i=K}^{b_n} \frac{1}{\epsilon_n + \frac{i}{q_{n+1}}} \geq \sum_{i=K+\lceil \epsilon_n q_{n+1}\rceil}^{b_n + \lceil \epsilon_n q_{n+1}\rceil} \frac{q_{n+1}}{i} \\ &\geq (1000K)^{-1} q_{n+1} \log\left(\frac{b_n}{\epsilon_n q_{n+1}}\right) \geq \kappa N_n \log(N_n), \end{aligned}$$

for some $\kappa > 0$ and n sufficiently large.

By Proposition 75 we have

$$S^{b_n+1}_N(f)(x) = N\log(N)(1 + o(1)) \quad \text{uniformly in } x \in [0,1),$$

and it follows that

$$G \supset \{x \mid x \in A_n \text{ for infinitely many } n\}.$$

It remains to show

$$\lambda(x \mid x \in A_n \text{ for infinitely many } n) > 0.$$

(iv) We will verify the conditions of Lemma 81 for A_n, i.e. we will show

$$\sum_{n=1}^{\infty} \lambda(A_n) = \infty \tag{3.54}$$

[42] Since, for big enough n, the distance between any two points in $\{R(x), ..., R^N(x)\}$ is at least $\frac{1}{2q_{n+1}}$, at most one such point can be in $[0, \frac{\log_5(q_n)}{N_n \log(N_n)}) \subset [0, \frac{1}{2q_{n+1}})$.

[43] Since $\epsilon_n < \frac{1}{2q_n} < \min_{j_1,j_2\in\{0,...,q_n-1\}, j_1\neq j_2} |j_1\alpha - j_2\alpha|$.

and

$$\lambda(A_i \cap A_j) \le 2\lambda(A_i)\lambda(A_j) \quad \forall i < j. \tag{3.55}$$

Lemma 81 will then imply $\lambda(x \mid x \in A_n$ for infinitely many $n) > 0$ and the proof will be complete.

(v) First we verify (3.54). Indeed, the assumption $q_{n+1} < q_n \log^2(q_n)$ implies[44] that $q_n < 16^n(n!)^4$, in particular $\log(q_n) < 100n\log(n)$ and $\log_3(q_n) < 10\log_2(n)$. It follows that

$$\sum_{n=1}^{\infty} \lambda(A_n) = \sum_{n=1}^{\infty} \epsilon_n q_n = \sum_{n=1}^{\infty} \frac{1}{\log(q_n)\log_3(q_n)} > \frac{1}{1000}\sum_{n=1}^{\infty} \frac{1}{n\log(n)\log_2(n)} = \infty.$$

Now we verify (3.55). Let $i < j$ and, for $l_1 = 0, ..., q_i - 1$ and $l_2 = 0, ..., q_j - 1$, denote

$$I_{l_1} = (0, \epsilon_i) - l_1\alpha \quad \text{and} \quad \tilde{I}_{l_2} = (0, \epsilon_j) - l_2\alpha,$$

additionally

$$\mathcal{N}_{l_1} = \#\{l \in \{0, ..., q_j - 1\} \mid \tilde{I}_l \cap I_{l_1} \neq \emptyset\}.$$

Since the $\tilde{I}_l$ are disjoint we have

$$\mathcal{N}_{l_1} = \#\{l \in \{0, ..., q_j - 1\} \mid -l\alpha \in I_{l_1}\} = \#\{l \in \{0, ..., q_j - 1\} \mid l\alpha \in -I_{l_1}\} = S_{q_j}(1_{-I_{l_1}})(0).$$

$$\lambda(A_i \cap A_j) \le S_{q_j}(1_{A_i})(0)\epsilon_j \le (q_j q_i \epsilon_i + 2)\epsilon_j \le 2\lambda(A_i)\lambda(A_j),$$

showing (3.55) and completing the proof. □

3.6.3 The function $x^{-\beta}$, $\beta > 1$

In this section, we shall provide a proof of Theorem 56. Let $\beta > 1$ and $f(x) = x^{-\beta}$. Clearly, the strong law implies the weak law, so (I) $\implies$ (II). The main propositions in the section are the following: We show that the strong law holds (uniformly) assuming (3.16), i.e. (III) $\implies$ (I), see Proposition 84. Then we introduce an a priori weaker condition than (3.16) - condition (D) given in Definition 85. We show in Proposition 92 that (II) $\implies$ condition (D) and Lemma 86 shows that condition (D) and (3.16) are equivalent under the condition that $k(N)$ is monotone implying (II) $\implies$ (III).

We start with some preparations to prove Proposition 84. To simplify matters, we will introduce an auxiliary sequence $K(N) = o(N)$. Instead of studying the full sum $S_N(f)$ it will be convenient to study separately[45] the sums $S_{N'}(f)$ and $S_{N''}(f)$ with

$$N' = \sum_{j=n-K(N)}^{n} b_j q_j, \qquad \text{and} \qquad N'' = N - N' = \sum_{j=1}^{n-K(N)-1} b_j q_j \le q_{n-K(N)}.$$

By carefully choosing $K(N)$, the second sum $S_{N''}(f)$ is dominated by $S_{N'}(f)$, hence it suffices to study the latter.

Let $\epsilon_\beta > 0$ be so that $(1+\epsilon_\beta)\frac{\beta-1}{\beta} < 1$, and

$$K(N) = \min\left(\kappa \ge 1 \;\middle|\; q_{n-\kappa} \le Nk(N)^{(1+\epsilon_\beta)\frac{1-\beta}{\beta}}\right).$$

[44]This can be shown by induction. Recall that, by convention $q_1 = 1$. Now assume $q_n < 16^n(n!)^4$ for some n, then it holds that

$$q_{n+1} < 16^n(n!)^4\log^2(16^n(n!)^4) \le 16^n(n!)^4(\log(16)n + 4n\log(n))^2 \le 16^{n+1}((n+1)!)^4.$$

[45]It holds that $S_N(f) = S_{N'}(f) + S_{N''}(f) \circ R^{N''}$.

With this choice it holds that

$$\begin{aligned}\sup_{x\in[0,1)} |S^1_{N''}(f)(x)| \le \sup_{x\in[0,1)} |S^1_{q_{n-K(N)}}(f)(x)| &\le \sum_{i=1}^{q_{n-K(N)}} q^{\beta}_{n-K(N)} i^{-\beta} \\ &= O\big(q^{\beta}_{n-K(N)}\big) = O\big(N^{\beta} k(N)^{(1+\epsilon_\beta)(1-\beta)}\big) = o\big(N^{\beta} k(N)^{1-\beta}\big).\end{aligned}$$

Since

$$S^{k(N)}_{N'}(f) \le S^{k(N)}_{N}(f) \le S^{k(N)-1}_{N'}(f) + S^1_{N''}(f)\circ R^{N'}, \tag{3.56}$$

the strong law (3.14), with $d_N = \frac{1}{\beta-1} N^{\beta} k(N)^{1-\beta}$, will follow once we show[46]

$$\lim_{N\to\infty} \frac{S^{k(N)}_{N'}(f)(x)}{N^{\beta} k(N)^{-\beta+1}} = \frac{1}{\beta-1} \quad \text{uniformly in } x \in [0,1). \tag{3.57}$$

First, we claim that

$$\frac{k(N)}{\sum_{n-K(N)\le j\le n} b_j} \to \infty. \tag{3.58}$$

Note that, if[47] $K(N) = 1$, then this follows trivially from (3.16). Therefore we can focus on the case $K(N) > 1$.

By (3.16), the condition (3.58) is equivalent to the a priori weaker condition

$$\frac{k(N)}{\sum_{n-K(N)+1\le j\le n} b_j} \to \infty.$$

Note that, for every $1 \le J \le n-1$, N satisfies

$$N \ge b_n q_n > b_n a_{n-1} q_{n-1} > \ldots > b_n \prod_{j=n-J}^{n-1} a_j q_{n-J}.$$

Therefore, by the definition of $K(N)$ (and since $K(N) > 1$), we have

$$q_{n-K(N)+1} > N k(N)^{(1+\epsilon_\beta)\frac{1-\beta}{\beta}} > b_n \prod_{j=n-K(N)+1}^{n-1} a_j q_{n-K(N)+1} k(N)^{(1+\epsilon_\beta)\frac{1-\beta}{\beta}}.$$

Moreover, for sufficiently large n and K, it holds that

$$q_n > \rho^{K-2} q_{n-K}, \quad \text{where } \rho = \frac{1+\sqrt{5}}{2} \text{ (the golden ratio)}.$$

It follows that $K(N) = o(k(n))$, and consequently,

$$\sum_{n-K(N)+1\le j\le n} b_j \le b_n \prod_{j=n-K(N)+1}^{n-1} a_j + K(N) \le k(N)^{(1+\epsilon_\beta)\frac{\beta-1}{\beta}} + K(N) = o(k(N))$$

[46] Applying the same arguments with $k(N)-1$ instead of $k(N)$ also yields

$$\lim_{N\to\infty} \frac{S^{k(N)-1}_{N'}(f)(x)}{\frac{1}{\beta-1} N^{\beta} k(N)^{-\beta+1}} = 1 \quad \text{uniformly in } x \in [0,1).$$

[47] To be slightly more precise; if $K(N) = 1$ along a subsequence, then (3.58) follows trivially from (3.16) along this subsequence.

and hence (3.58) holds.

Let $z(N,x) = R^{j^{N'}_{k(N)+1}(x)}(x)$, recall that this means that $z(N,x)$ is the $k(N)+1$st point in $\{x,...,R^{N'-1}(x)\}$ that we encounter moving from left to right from 0 (not counting 0 itself, if $0 \in \{x,...,R^{N'-1}(x)\}$), then we have

$$S^{k(N)}_{N'}(f)(x) = f(z(N,x)) + \sum_{y\in\{x,...,R^{N'-1}(x)\}, y>z(N,x)} f(y). \tag{3.59}$$

We first try to locate $z(N,x)$. To that end first rewrite

$$z(N,x) = \min(\zeta \in \{x,...,R^{N'-1}(x)\} \mid S_{N'}(1_{(0,\zeta]})(x) \geq k(N)+1).$$

For $\zeta \in (0,1)$ we have

$$|S_{N'}(1_{(0,\zeta]})(x) - N'\zeta| \leq 2\sum_{j=n-K(N)}^{n} b_j.$$

Since

$$S_{N'}(1_{(0,\zeta]})(x) \in [N'\zeta - 2\sum_{j=n-K(N)}^{n} b_j \,,\, N'\zeta + 2\sum_{j=n-K(N)}^{n} b_j],$$

it follows that

$$z(N,x) \in \left[\frac{k(N)+1-2\sum_{j=n-K(N)}^{n} b_j}{N'}, \frac{k(N)+1+2\sum_{j=n-K(N)}^{n} b_j}{N'}\right] =: I_N.$$

Denote

$$f_N(x) = \begin{cases} f(x) & \text{if } x \in \left[\frac{k(N)+1}{N'}, 1\right), \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 82. *For $N \in [q_n, q_{n+1}-1]$ sufficiently large*

$$|S^{k(N)}_{N'}(f)(x) - S_{N'}(f_N)(x)| = o(N^\beta k(N)^{1-\beta}) \quad \textit{uniformly in } x \in [0,1).$$

Proof. For fixed $x \in [0,1)$ let $\tilde{f}_N$ be the function defined by

$$\tilde{f}_N(y) = \begin{cases} f(y) & \text{if } y \in [z(N,x), 1)\,, \\ 0 & \text{otherwise,} \end{cases}$$

then, by (3.59), we have $S^{k(N)}_{N'}(f)(x) = S_{N'}(\tilde{f}_N)(x)$. Furthermore, for N sufficiently large

$$\begin{aligned} |S_{N'}(\tilde{f}_N)(x) - S_{N'}(f_N)(x)| &\leq 2^\beta N^\beta k(N)^{-\beta} \#\{0 \leq j \leq N'-1 \mid R^j(x) \in I_N\} \\ &\leq 2^\beta N^\beta k(N)^{-\beta} S_{N'}(1_{I_N})(x) \\ &\leq 1000^\beta N^\beta k(N)^{-\beta} \sum_{j=n-K(N)}^{n} b_j, \end{aligned}$$

and the claim follows since $\sum_{j=n-K(N)}^{n} b_j = o(k(N))$. □

Lemma 83. *It holds that*

$$\left|S^{k(N)}_{N'}(f)(x) - \frac{1}{\beta-1}N^\beta k(N)^{-\beta+1}\right| = o(N^\beta k(N)^{-\beta+1}) \quad \textit{uniformly in } x \in [0,1). \tag{3.60}$$

Proof. The Denjoy-Koksma inequality yields

$$|S_{N'}(f_N)(x) - N'\int_0^1 f_N(y)\,\mathrm{d}y| \leq N^\beta k(N)^{-\beta} \sum_{j=n-K(N)}^{n} b_j.$$

Since $\int_0^1 f_N(y)\,\mathrm{d}y = \frac{1}{\beta-1}(N^{\beta-1}k(N)^{-\beta+1}-1)$ and $\frac{N}{N'} \to 1$, the claim follows via Lemma (82). $\square$

Altogether we obtain the following.

Proposition 84. *If* (3.16) *holds, then the strong law holds uniformly, i.e.*

$$\lim_{N\to\infty} \frac{S_N^{k(N)}(f)(x)}{d_N} = \frac{1}{\beta-1} \quad \textit{uniformly in } x \in [0,1).$$

Proof. Considering (3.56), it is enough to show (3.57), which follows from (3.60). $\square$

To complete the proof of Theorem 56, it remains to show that if the weak law[48] holds and $k(N)$ is monotone, then $k(N)$ satisfies condition (3.16).

To this end, we show that if the weak law (3.15) is satisfied, then $k(N)$ satisfies the following (a priori weaker) condition.

Definition 85. $k(N)$ *is said to satisfy condition (D), if, for any subsequence* $(N_l)_{l\geq 1}$ *satisfying* $\limsup_{l\to\infty} \frac{N_l}{q_{n+1}} < 1$, *it holds that*

$$\frac{k(N_l)}{b_n} \to \infty. \tag{3.61}$$

Lemma 86. *Let* $k(N)$ *be a monotone sequence satisfying condition (D), then* (3.16) *holds.*

Since clearly (3.16) is the stronger condition, the lemma shows that if $k(N)$ is assumed to be monotone, then (D) is equivalent to (3.16).

Proof. If α is of bounded type, then both (3.16) and (3.61) are equivalent to $k(N) \to \infty$. So assume now that α is not of bounded type.

Let $\mathcal{N} := \{n \mid a_n > 2\}$ and consider the sequence

$$(N_l)_{l\geq 1} = (q_n, 2q_n, ..., \left\lceil \frac{a_n}{2} \right\rceil q_n \mid n \in \mathcal{N}),$$

clearly $\limsup_{l\to\infty} \frac{N_l}{q_{n+1}} \leq \frac{2}{3}$ and, per assumption, (3.61) holds for the sequence $(N_l)_{l\geq 1}$.

In order to show (3.16), let $C > 0$, by (3.61) there is an L_0 so big that $\frac{k(N_l)}{b_n} > 2C$ for $l \geq L_0$. Furthermore, let n_0 be such that $N_{L_0} \in [q_{n_0}, q_{n_0+1} - 1]$. Since α is not of bounded type, there is an $n_1 > n_0$ such that[49] $a_{n_1} > \max_{j\leq n_0} a_j$. Now (3.16) will follow once we show that, for[50] $N > q_{n_1+1}$, it holds that

$$\frac{k(N)}{\max(b_n, \max_{j\leq n-1} a_j)} > C.$$

[48] A priori weak or strong laws could also be possible for a different normalisation $d_N > 0$ with

$$\lim_{N\to\infty} \frac{(\beta-1)d_N}{N^\beta k(N)^{1-\beta}} \neq 1,$$

so we have to take this into account in the proof.

[49] Implicitly $n_1 \in \mathcal{N}$.

[50] Now N is not necessarily a member of the sequence $(N_l)_{l\geq 1}$ any more.

Since $n \geq n_1 + 1$ and $n_1 \in \mathcal{N}$, the above is equivalent to

$$\frac{k(N)}{\max(b_n, \max_{j \leq n-1, j \in \mathcal{N}} a_j)} > C.$$

For $j \in [n_0 + 1, n - 1] \cap \mathcal{N}$ let $l_j \geq 1$ be such that $N_{l_j} = \lceil \frac{a_n}{2} \rceil q_n$, and, if $n \in \mathcal{N}$, let $l_n \geq 1$ be such that $N_{l_n} = \min(b_n, \lceil \frac{a_n}{2} \rceil)$. Clearly ,

$$\min_{j \in [n_0+1, n]} l_j \geq L_0 \quad \text{ and } \quad N \geq \max_{j \in [n_0+1, n]} N_{l_j}.$$

Since k is monotone

- $k(N) \geq k(N_{l_{n_1}}) > C a_{n_1} > C \max_{j \leq n_0} a_j$,
- for $j \in [n_0 + 1, n - 1]$ it holds that $k(N) \geq k(N_{l_j}) > C a_j$ since $N \geq \lceil \frac{a_{n_1}}{2} \rceil q_{n_1}$,
- $k(N) \geq k(N_{l_n}) > C b_n$ since $N \geq b_n q_n$.

□

Assume now that there are $d_N > 0$ such that

$$\lambda \left(\left| \frac{S_N^{k(N)}(f)}{d_N} - 1 \right| > \epsilon \right) \to 0 \quad \text{as } N \to \infty, \quad \forall \epsilon > 0.$$

We will show that $k(N)$ satisfies condition (D).

To this end, we will show that if $k(N)$ does not satisfy (D), then $S_N^{k(N)}(f)$ has "oscillations" of order $N^\beta k(N)^{1-\beta}$ (Lemma 91). To conclude, we have to ensure additionally that, along suitable subsequences $(N_l)_{l \geq 1}$, it holds that[51] $d_{N_l} = O(N_l^\beta k(N_l)^{1-\beta})$.

Lemma 87. *There is a constant $C > 0$ only depending on β such that, for $\epsilon \in (0, \frac{1}{100})$, $N \in [q_n, (1-\epsilon)q_{n+1}]$, $k \geq 0$ and n big enough, we have*

$$\lambda \left(S_N^k(f) \leq C \epsilon^{-\beta} \min \left(N q_n^{\beta-1}, N^\beta k^{1-\beta} \right) \right) \geq \frac{\epsilon}{40}.$$

Proof. Assume $\alpha - \frac{p_n}{q_n} > 0$, the other case can be proven analogously. Let $A \subset [0, 1)$ be the set

$$A = \left\{ x \in [0,1) \;\middle|\; R^{j_1^{q_n}(x)}(x) \in \left[\frac{\epsilon}{20} \delta_n, \frac{\epsilon}{10} \delta_n \right] \right\} = \bigcup_{j=0}^{q_n - 1} \left[\frac{\epsilon}{20} \delta_n, \frac{\epsilon}{10} \delta_n \right] - j\alpha,$$

where δ_n is given by (3.28), clearly $\lambda(A) \geq \frac{\epsilon}{40}$. We have $b_n \leq (1-\epsilon) \frac{q_{n+1}}{q_n}$. Hence, by Lemma 63, for $x \in A$ it holds that

$$R^{j_{q_n}^{q_n}(x)}(x) \leq -(1 - \frac{\epsilon}{10}) \delta_n \leq -(1 - \frac{\epsilon}{10}) a_n \delta_{n+1} \leq -b_n \frac{1}{q_{n+1}},$$

and Lemma 64 ascertains that $j_1^N(x) = j_1^{q_n}(x)$. By summing separately over the first cluster of points, we obtain

$$\begin{aligned} S_N^k(f)(x) &\leq \sum_{j=k}^{b_n+1} \left(\frac{\epsilon}{40 q_n} + \frac{j}{2 q_{n+1}} \right)^{-\beta} + b_n \sum_{i=\max\left(1, \left\lfloor \frac{k}{b_n} \right\rfloor\right)}^{q_n - 1} \left(\frac{i}{2 q_n} \right)^{-\beta} \\ &\leq 100^\beta \epsilon^{-\beta} \chi_{k \leq b_n + 1} b_n q_n^\beta + \frac{2^\beta}{\beta - 1} b_n q_n^\beta \min(1, b_n^{\beta-1} k^{1-\beta}) \\ &\leq C \epsilon^{-\beta} \min(N q_n^{\beta-1}, N^\beta k^{1-\beta}), \end{aligned}$$

[51] This is not quite what we show, but morally speaking.

for some constant $C > 0$, where

$$\chi_{k \leq b_n+1} = \begin{cases} 1 & \text{if } k \leq b_n + 1 \\ 0 & \text{otherwise.} \end{cases}$$

$\square$

Lemma 88. *If the weak law of large numbers as in* (3.15) *holds and* $(N_l)_{l\geq 1}$ *is a subsequence fulfilling* $\limsup_{l\to\infty} \frac{N_l}{q_{n+1}} < 1$*, then*

$$d_{N_l} = O(\min(N_l q_n^{\beta-1}, N_l^{\beta} k(N_l)^{1-\beta})). \tag{3.62}$$

Proof. Let $\epsilon \in (0, 100^{-1})$ and L be such that $\frac{N_l}{q_{n+1}} < 1 - \epsilon$ for $l > L$. For such l, Lemma 87 yields

$$\lambda(S_{N_l}^k \leq C\epsilon^{-\beta} \min(N_l q_n^{\beta-1}, N_l^{\beta} k(N_l)^{1-\beta})) \geq \frac{\epsilon}{40}.$$

This clearly implies

$$d_{N_l} = O(\min(N_l q_n^{\beta-1}, N_l^{\beta} k(N_l)^{1-\beta})).$$

$\square$

Remark 89. *A posteriori, if* $k(N)$ *satisfies condition (D), then*

$$N_l^{\beta} k(N_l)^{1-\beta} = o(N_l q_n^{\beta-1}),$$

and (3.62) *becomes simply*

$$d_{N_l} = O(N_l^{\beta} k(N_l)^{1-\beta}).$$

Lemma 90. *For* $\epsilon \in (0,1)$ *and* $w \geq 1$ *it holds that*

$$\lfloor (1-\epsilon)w \rfloor - 1 \leq (1 - \frac{\epsilon}{2}) \lfloor w \rfloor.$$

Proof. If $w \leq \frac{2}{\epsilon}$, then

$$\frac{\lfloor (1-\epsilon)w \rfloor - 1}{\lfloor w \rfloor} \leq 1 - \frac{1}{w} \leq 1 - \frac{\epsilon}{2}.$$

On the other hand, if $w > \frac{2}{\epsilon}$

$$\lfloor (1-\epsilon)w \rfloor = \lfloor w - \epsilon w \rfloor \leq \lfloor w \rfloor - \lfloor \epsilon w \rfloor,$$

and

$$\lfloor \epsilon w \rfloor \geq \epsilon w - 1 \geq (1 - \frac{1}{2})\epsilon w \geq \frac{\epsilon}{2} \lfloor w \rfloor.$$

It follows that

$$\lfloor (1-\epsilon)w \rfloor \leq (1 - \frac{\epsilon}{2}) \lfloor w \rfloor.$$

$\square$

Lemma 91. *There is a constant* $c > 0$ *such that, for* $\epsilon \in (0, \frac{1}{100})$ *and* $N \in [q_n, (1-\epsilon)q_{n+1}]$*, there are sets* A, B *with* $\lambda(A) = \lambda(B) \geq \frac{\epsilon^2}{1000}$ *such that*

$$S_N(f)(x) - S_N(f)(y) > c\epsilon N q_n^{\beta-1} \quad \forall x \in A, y \in B. \tag{3.63}$$

Furthermore, if $k = \hat{k}\frac{N}{q_n} \in \mathbb{N}$ *then*

$$S_N^k(f)(x) - S_N^k(f)(y) > c\epsilon \min\left(N q_n^{\beta-1}, \hat{k}^{-2} N^{\beta} k^{1-\beta}\right) \quad \forall x \in A, y \in B. \tag{3.64}$$

Proof. We focus on (3.64), (3.63) can be shown using almost the same arguments. In the following we assume $\alpha - \frac{p_n}{q_n} > 0$, the other case follows by considering $1-\alpha$ instead.

Let

$$\begin{aligned} A' &= \{x \in (0,1] \mid R^{j_1^{q_n}(x)}(x) \in (0, \frac{\epsilon}{10}\delta_n) \text{ and } j_1^{q_n}(x) \in \{0, ..., q_n - q_{n-1} - 1\}\} \\ &= \bigcup_{j=0}^{q_n - q_{n-1}-1} (0, \frac{\epsilon}{10}\delta_n) - j\alpha. \end{aligned}$$

Since $\delta_n \geq \frac{1}{2q_n}$, it holds that[52] $\lambda(A') \geq \frac{\epsilon^2}{40}$. For $x \in A'$ we have

$$R^{j_{q_n}^{q_n}(x)}(x) < R^{j_{q_n}^{q_n}(x)+q_n}(x) < ... < R^{j_{q_n}^{q_n}(x)+(b_n-1)q_n}(x) < R^{j_1^{q_n}(x)}(x),$$

and no other points[53] of $\{x, ..., R^{N-1}(x)\}$ are between $R^{j_{q_n}^{q_n}(x)}(x)$ and $R^{j_1^{q_n}(x)}(x)$. Using Lemma 90, we have

$$b_n - 1 \leq \left\lfloor \frac{(1-\epsilon)q_{n+1}}{q_n} \right\rfloor - 1 \leq \left(1 - \frac{\epsilon}{2}\right) \left\lfloor \frac{q_{n+1}}{q_n} \right\rfloor = \left(1 - \frac{\epsilon}{2}\right) a_n,$$

in addition Lemma 63 yields $a_n \delta_{n+1} \geq \delta_n$ and therefore

$$R^{j_{q_n}^{q_n}(x)+(b_n-1)q_n}(x) \leq -\left(1 - \frac{\epsilon}{10}\right)\delta_n + (b_n - 1)\delta_{n+1} \leq -\left(1 - \frac{\epsilon}{10}\right)\delta_n + \left(1 - \frac{\epsilon}{2}\right)a_n\delta_{n+1} \leq -\frac{\epsilon}{4}\delta_n.$$

It follows that $j_1^N(x) = j_1^{q_n}(x)$ and $j_i^N(x) = j_i^N(x + \frac{\epsilon}{4}\delta_n)$ for all $i = 1, ..., N$. Writing

$$S_N^k(f)(x) - S_N^k(f)(x + \frac{\epsilon}{4}\delta_n) \geq \sum_{l=\lceil \hat{k} \rceil + 1}^{q_n} \sum_{i=0}^{b_n - 1} \left(f\left(R^{j_l^{q_n}(x)+iq_n}(x)\right) - f\left(R^{j_l^{q_n}(x)+iq_n}(x) + \frac{\epsilon}{4}\delta_n\right)\right),$$

we notice that all the summands are positive, hence we can bound below by taking only the terms where $l = \lceil \hat{k} \rceil + 1$. Furthermore, all of the points $R^{j^{q_n}_{\lceil \hat{k} \rceil+1}(x)+iq_n}(x)$ are to the left of $R^{j^{q_n}_{\lceil \hat{k} \rceil+2}(x)}(x)$, therefore we again bound from below by replacing them all with this point. Therefore

$$\begin{aligned} S_N^k(f)(x) - S_N^k(f)(x + \frac{\epsilon}{4}\delta_n) &\geq b_n \left(f\left(R^{j^{q_n}_{\lceil \hat{k} \rceil+2}(x)}(x)\right) - f\left(R^{j^{q_n}_{\lceil \hat{k} \rceil+2}(x)}(x) + \frac{\epsilon}{4}\delta_n\right)\right) \\ &\geq b_n \left(\left(\frac{\hat{k}+3}{q_n}\right)^{-\beta} - \left(\frac{\hat{k}+3}{q_n} + \frac{\epsilon}{8q_n}\right)^{-\beta} \right) \\ &\geq c\epsilon \min(1, \hat{k}^{-\beta-1}) b_n q_n^\beta \geq c\epsilon \min\left(N q_n^{\beta-1}, \hat{k}^{-2} N^\beta k^{1-\beta}\right), \end{aligned} \tag{3.65}$$

for a constant $c > 0$.

We distinguish two cases:

[52] Strictly speaking this is only true if $(1 - \frac{\epsilon}{2})q_n > q_{n-1}$. If $(1 - \frac{\epsilon}{2})q_n \leq q_{n-1}$, then a similar proof with the set

$$A' = \{x \in (0,1] \mid R^{j_1^{q_n}(x)}(x) \in (0, \frac{\epsilon}{20}\delta_{n-1})\} = \bigcup_{j=0}^{q_n-1} (0, \frac{\epsilon}{20}\delta_{n-1}) - j\alpha$$

leads to the same conclusion.

[53] The only candidates for points between $R^{j_{q_n}^{q_n}(x)}(x)$ and $R^{j_1^{q_n}(x)}(x)$ are

$$R^{j_{q_n}^{q_n}(x)}(x) < R^{j_{q_n}^{q_n}(x)+q_n}(x) < ... < R^{j_{q_n}^{q_n}(x)+b_nq_n}(x) < R^{j_1^{q_n}(x)}(x).$$

However the assumption $j_1^{q_n}(x) \leq q_n - q_{n-1} - 1$, or equivalently $j_{q_n}^{q_n} \geq q_{n-1}$, ensures that the point $R^{j_{q_n}^{q_n}(x)+b_nq_n}(x)$ is not present in $\{x, ..., R^{N-1}(x)\}$.

(i) If there is an s such that $\lambda(x \in A' \mid S_N^k(f)(x) = s) \geq \frac{1}{3}\lambda(A')$, we set

$$A = \{x \in A' \mid S_N^k(f)(x) = s\} \text{ and } B = A + \frac{\epsilon}{4}\delta_n.$$

Then, for $x \in A, y \in B$, it holds that

$$S_N^k(f)(x) - S_N^k(f)(y) = S_N^k(f)(y - \frac{\epsilon}{4}\delta_n) - S_N^k(f)(y) \overset{(3.65)}{>} c\epsilon \min\left(Nq_n^{\beta-1}, \hat{k}^{-2}N^\beta k^{1-\beta}\right),$$

where in the first equality we used the fact that $S_N^k(f)$ is constant on A.

(ii) Otherwise, if there is no such s, let

$$\mathcal{S} = \left\{ s > 0 \;\middle|\; \lambda(x \in A' \mid S_N^k(f)(x) \leq s) \geq \frac{1}{3}\lambda(A') \right\} \text{ and } s_0 = \inf \mathcal{S}.$$

Then, for[54]

$$A = \{x \in A' \mid S_N^k(f)(x) > s_0\} \text{ and } B = \{x \in A' \mid S_N^k(f)(x) \leq s_0\} + \frac{\epsilon}{4}\delta_n,$$

it holds that, for $x \in A, y \in B$, we have

$$S_N^k(f)(x) - S_N^k(f)(y) > S_N^k(f)(y - \frac{\epsilon}{4}\delta_n) - S_N^k(f)(y) \overset{(3.65)}{>} c\epsilon \min\left(Nq_n^{\beta-1}, \hat{k}^{-2}N^\beta k^{1-\beta}\right),$$

where we used that $S_N^k(f)(x) > s_0$ by definition of A, and $S_N^k(f)(y - \frac{\epsilon}{4}\delta_n) \leq s_0$ by definition of B.

Furthermore, by continuity of λ, we have

$$\lambda(B) = \lambda\left(\bigcap_{s \in \mathcal{S}} \{x \in A' \mid S_N^k(f)(x) \leq s\}\right) \geq \frac{1}{3}\lambda(A').$$

In order to estimate $\lambda(A)$, we distinguish two possibilities

- if $s_0 \notin \mathcal{S}$, then by definition it holds that

$$\lambda(A) = \lambda(x \in A' \mid S_N^k(f)(x) > s_0) \geq \frac{2}{3}\lambda(A'),$$

- if $s_0 \in \mathcal{S}$, then

$$\begin{aligned}\lambda(A) &= \lambda(x \in A' \mid S_N^k(f)(x) > s_0) \\ &= \lambda\Big(\bigcup_{s \notin \mathcal{S}} \{x \in A' \mid S_N^k(f)(x) > s\}\Big) - \lambda\left(x \in A' \mid S_N^k(f)(x) = s_0\right) \geq \frac{1}{3}\lambda(A').\end{aligned}$$

□

Proposition 92. *If there are some $d_N > 0$ such that*

$$\lambda\left(\left|\frac{S_N^{k(N)}(f)}{d_N} - 1\right| > \epsilon\right) \to 0 \quad \text{as } N \to \infty, \quad \forall \epsilon > 0,$$

then $k(N)$ fulfils Property D.

Proof. We assume weak convergence as in (3.15) holds and let $(N_l)_{l\geq 1}$ be a subsequence with $\lim_{l\to\infty} \frac{N_l}{q_{n+1}} < 1$, say $\lim_{l\to\infty} \frac{N_l}{q_{n+1}} = 1 - \epsilon$ for some small $\epsilon > 0$. By Lemma 88, we have

$$d_{N_l} = O(\min(N_l q_n^{\beta-1}, N_l^\beta k(N_l)^{1-\beta})). \tag{3.66}$$

We claim that necessarily $\frac{k(N_l)}{b_n} \to \infty$, equivalently $\frac{k(N_l)q_n}{N_l} \to \infty$. If not, then, by Lemma[55] 91,

[54] Clearly $\mathcal{S} \neq \emptyset$ because $S_N^k(f)(x) < 100^\beta \epsilon^{-\beta} N q_n^\beta$ whenever $x_{\min}^N > \frac{\epsilon}{10}\delta_n$.

[55] A priori it might also be possible that $k(N_l) = 0$, in this case, use (3.63). Otherwise, we use (3.64).

there are sets A_{N_l}, B_{N_l} with $\lambda(A_{N_l}) = \lambda(B_{N_l}) \geq \frac{\epsilon^2}{1000}$ with

$$S_{N_l}^{k(N_l)}(f)(x) - S_{N_l}^{k(N_l)}(f)(y) > \tilde{c}\min\left(N_l q_{n_l}^{\beta-1}, N_l^{\beta} k(N_l)^{1-\beta}\right), \quad \forall x \in A_{N_l}, y \in B_{N_l},$$

for a constant[56] $\tilde{c} > 0$ which depends on ϵ but not on l. In light of (3.66), this contradicts (3.15). □

Finally, we are in a position to give the full proof of Theorem 56.

Proof of Theorem 56. Clearly, strong convergence implies weak convergence, therefore (I) $\Longrightarrow$ (II). Additionally, if $k(N)$ satisfies condition (3.16), then Proposition 84 shows that the strong law holds uniformly, so (III) $\Longrightarrow$ (I). Moreover, by Proposition 92 we have that (II) $\Longrightarrow$ condition (D) and according to Lemma 86, if $k(N)$ is monotone, condition (D) is equivalent to (3.16). □

[56] The dependence on ϵ and "$\hat{k}$" as in Lemma 91 can be absorbed into $\tilde{c}$. This is doesn't cause any problems since $\hat{k} = \frac{k(N_l)q_n}{N_l}$ is bounded by assumption.

Chapter 4

Concluding remarks and future directions

To conclude this thesis, let us remark on some natural questions that arise from this work.

4.1 Dimensionality in the PLT

One limitation in our construction of non-mixing systems satisfying the PLT (Chapter 2) is the requirement that the fiber, which is exponentially mixing of all orders, must have a very high dimension. A natural question is whether this assumption can be relaxed.

One possible approach is to refine the partitioning scheme used in the proof so that the estimates needed for recurrence statistics remain valid even in lower-dimensional settings. This would require modifying the geometric arrangement of target sets in a way that ensures the necessary separation conditions hold while still preserving sufficient mixing in the fiber. However, adapting the arguments of Section §2.8.2 to this new setting presents technical challenges, particularly in maintaining control over the number of targets introduced at each scale.

Addressing this question could lead to a more flexible framework for establishing the PLT in weakly chaotic systems, potentially broadening its applicability to lower-dimensional settings.

A preliminary outline of how one might modify the proof is as follows:

- Note that we only the used estimates on the dimension in the calculations of §2.9 in order to show $\mu(A_l^{(j)}) = o(\nu(B_l^{(j)})^\iota)$ for some power $\iota > 0$ that was determined by the two systems T and R. Since $\mu(A_l^{(j)}) \asymp r_l^d$ and $\nu(B_l^{(j)}) \asymp r_l^{d'}$, we just had to choose $d > \iota d'$ in order to satisfy the requirement.
- Alternatively we could simply partition the ball Q_l into more rectangles, say so that each $A_l^{(j)}$ has a diameter of order r_l^a with $a > \max(1, \frac{\iota d'}{d})$, while $B_l^{(j)}$ has a diameter of order r_l. The total number of rectangles $A_l^{(j)} \times B_l^{(j)}$ required is of order $O(r_l^{1-a})$.
- With this choice it now holds that $\mu(A_l^{(j)}) = o(\nu(B_l^{(j)})^\iota)$, regardless of what d and d' are, and all the arguments of §2.9 can be applied.
- On the other hand the PLT in the fiber, obtained in §2.8.2 can no longer be applied. In addition to just shrinking the targets $A_l^{(j)}$, in each step we increase the number of targets as well. So some additional work is required to fit the arguments into this setting.

4.2 Hitting time statistics for slowly chaotic systems

Recall that the goal of Chapter 2 was to find limit distributions of

$$\mu(B_r(x^*))\varphi_{B_r(x^*)} \tag{4.1}$$

as $r \to \infty$, where $B_r(x^*)$ is a ball of radius r around x^*, for all or almost all x^*, and $\varphi_{B_r(x^*)}$ denotes the first hitting time of $B_r(x^*)$. Further recall that the distribution of (4.1) w.r.t. μ is referred to as the *hitting time statistic*, while the distribution w.r.t. the conditional measure $\mu_{B_r(x^*)}$ is referred to as the *return time statistic*.

For sufficiently chaotic systems, see also the list in §2.1, hitting time statistics for geometric balls are expected to be asymptotically

- exponentially distributed if x^* is aperiodic,
- scaled exponentially distributed if x^* is periodic. Note that in this case the limit of the hitting time statistics is only a sub-probability distribution[1].

To contrast this, a wide range of distributions are a priori possible as limits of hitting or return time statistics. In fact, in [Lac02] it is shown that any positive distribution with expectation ≤ 1 can appear as a limit point of return time statistics. Possible limits for hitting time statistics are described by [KL05], this description can also be more directly derived from the relation of hitting and return time statistics in [HLV05].

Below we shall discuss results on hitting time limits for some maps and flows with only slow chaotic behaviour. For flows the question about hitting times has to be modified as mentioned in §2.1. Note that in all the cases below, the limit is not exponential.

4.2.1 Irrational Rotations

For an irrational rotation $R = R_\alpha$ on $[0, 1)$, where $\alpha \in [0, 1] \setminus \mathbb{Q}$, hitting time statistics have been studied for specific sequences of shrinking intervals. Given a point $x^* \in [0, 1)$ and $n \geq 1$, let J_n be the interval with endpoints $R^{q_n}(x^*)$ and $R^{q_{n+1}}(x^*)$ that contains x^*, then [CdF96] investigates the normalized hitting times $\mu(J_n)\varphi_{J_n}$. The results in [CdF96] show that, along a subsequence $\{n_i\}$, the normalized hitting times exhibit the following behaviour:

- If $q_{n_i+1}/q_{n_i} \to \infty$, then $\mu(J_{n_i})\varphi_{J_{n_i}}$ converges to a uniform distribution.
- Otherwise, the limiting distribution can be a piecewise constant function, whose precise form depends on the arithmetic properties of the continued fraction approximants of α. A full classification of these distributions is given in [CdF96].

A natural question is whether these results extend to more general sequences of intervals. Unlike mixing systems, where hitting time statistics often exhibit universal behaviour, the case of irrational rotations appears highly sensitive to the arithmetic properties of the interval endpoints. If one considers a general sequence of intervals $I_n = [a_n, b_n]$, it remains unclear what types of limit distributions might arise, or even whether a meaningful classification is

[1] For example if T is the doubling map and $x^* = 0$, then [Zwe22, Theorem 10.1] describes the return time statistics as

$$\mu_{B_r(x*)}\left(B_r(x^*))\varphi_{B_r(x^*)} \leq t\right) \to \frac{1}{2} + \frac{1}{2}\left(1 - e^{-\frac{t}{2}}\right) \quad \forall t > 0, \ \textit{as } r \to 0, \tag{4.2}$$

The constant $\frac{1}{2}$ in (4.2) comes from the fact that half of $B_r(x^*)$ returns to itself immediately. Using the relation given in [HLV05, Main Theorem], the hitting time statistics satisfy

$$\mu\left(B_r(x^*))\varphi_{B_r(x^*)} \leq t\right) \to \frac{1}{2}\left(1 - e^{-\frac{t}{2}}\right) \quad \forall t > 0, \ \textit{as } r \to 0.$$

possible. The dependence on the Diophantine properties of a_n and b_n suggests that a deeper number-theoretic analysis is required to make further progress.

4.2.2 Integrable flows

For $m, d \in \mathbb{N}$ with $d \geq 2$, let $\mathcal{U} \subset \mathbb{R}^m$ be a bounded domain, and let $f : \mathcal{U} \to \mathbb{R}^d$ be a smooth function. We consider the *integrable flow*

$$\phi^t : \mathbb{T}^d \times \mathcal{U} \to \mathbb{T}^d \times \mathcal{U}, \quad (\theta, J) \mapsto (\theta + tf(J), J).$$

The target sets we consider, while avoiding overly technical definitions, are sets D_r that, in each leaf, are intervals transverse to the flow direction. A typical example is given by

$$D_r = \{(g(J) + sf^{\perp}(J), J) \mid -r < s < r\},$$

where $g : \mathcal{U} \to \mathbb{R}^d$ is a smooth function, and $f^{\perp} : \mathcal{U} \to \mathbb{R}^d$ is a unit-length vector field satisfying $f \cdot f^{\perp} = 0$.

In [DMS17], it is shown that for generic g, there exists a limiting distribution F such that

$$2r\varphi_{D_r} \Rightarrow F \quad \text{as } r \to 0. \tag{4.3}$$

Moreover, [DMS17] provides an explicit description of F.

We remark that intervals transverse to the flow direction are the natural class of sets to study. For example if we wanted to investigate hitting times of more geometric objects, such as balls in $\mathbb{T}^d \times \mathcal{U}$, a possible approach would be to realize the restricted dynamics as a suspension flow and apply the techniques of [PY17] to analyse the hitting time statistics.

4.2.3 Horocycle flow

Denote by h_t and g_s the horocyle and geodesic flow introduced in §2.11.2. Suppose $SL(2, \mathbb{R})/\Gamma$ has one cusp and let C be a closed stable horocycle of length 1 around the cusp, we will investigate the sets

$$D_R = \bigcup_{t \geq R} g_t(C).$$

In [KMK24] ($\Gamma = SL(2, \mathbb{Z})$) and [MP24] (general Γ) it is shown that there exists a limiting distribution F such that

$$\mu(D_R)\varphi_{D_R} \Rightarrow F \quad \text{as } R \to \infty. \tag{4.4}$$

Moreover, [MP24] provides an explicit description of F.

4.3 Distributional limit theorems for trimmed ergodic sums via the PLT

For ergodic sums of non-integrable random variables, distributional limit theorems are more classical than strong limit theorems, as in Chapter 3, where trimmed strong laws are discussed. More explicitly, do there exist constants a_N and b_N such that

$$\frac{S_N(f) - b_N}{a_N}, \tag{4.5}$$

or in the trimmed case

$$\frac{S_N^{k(N)}(f) - b_N}{a_N}, \tag{4.6}$$

for some trimming sequence $k(N)$, that converges to a non-degenerate distribution as $N \to \infty$?

In the case of iids[2], this question is well studied. We present the results in a simplified setting where $f = x^{-\beta}$ for $\beta \geq 1$:

- In the untrimmed case, [GK68] showed that (4.5) converges to a stable distribution as $N \to \infty$ for
$$a_N = N^\beta \text{ and } b_N = \begin{cases} N\log(N) & \text{if } \beta = 1, \\ 0 & \text{if } \beta > 1. \end{cases}$$
- The lightly trimmed case, where $k(N) = K$ is constant, has been studied[3] by several authors, including [Dar52], [Teu81], and [AB60]. They show convergence of (4.6) to a non-degenerate distribution for $a_N = N^\beta$. The limiting distribution and the constants b_N are computed in [Hal78]. Notably, the limiting distribution is not a stable distribution.
- The case of intermediate trimming, where $k(N) \to \infty$ but $k(N) = o(N)$ as $N \to \infty$, has been studied by [CoHM86]. They showed that there exist constants a_N and b_N such that (4.6) converges to a normal distribution.

Since classical techniques heavily rely on independence assumptions and characteristic function computations, they do not extend easily to deterministic settings. Instead, we sketch a proof using the Poisson Limit Theorem (PLT), which captures the behaviour of extreme events in a more flexible manner. This not only provides insight into the underlying mechanisms of convergence but also extends naturally to non-iid systems. The following sketch is work in progress by the author. In order to avoid technical difficulties, we encourage the reader to first consider the iid case; however, we note that this proof also appears to be applicable in the non-iid setting. Furthermore, for simplicity, let $k(N) = K$ be constant; this corresponds to the lightly trimmed or untrimmed case.

(i) Let $m, R \geq 1$ and
$$f_N = f 1_{\left[0,\left(\frac{R}{N}\right)^{\frac{1}{\beta}}\right]} \quad \text{and} \quad \bar{f}_N = f - f_N.$$

A straightforward calculation shows that $\mathrm{Var}(S_N(\bar{f}_N)) = O(N^{1+\frac{2\beta-1}{\beta}} R^{-\frac{2\beta-1}{\beta}}) = O(N^{2\beta}R^{-1})$. Considering that the goal is to show
$$\frac{S_N^{k(N)}(f) - b_N}{N^\beta} \Rightarrow L,$$
for some limit distribution L, it becomes apparent that, for large R, the oscillations of $\bar{f}_N$ are negligible. Therefore we will focus our efforts on f_N.

(ii) For $j = 0, ..., m-1$ denote
$$A_N^j = \left\{ \left(\frac{Nm}{Rj}\right)^\beta \geq f > \left(\frac{Nm}{R(j+1)}\right)^\beta \right\} = \left[\frac{Rj}{Nm}, \frac{R(j+1)}{Nm}\right],$$
where by convention $\frac{1}{0} = \infty$. For $N \geq 1$ let
$$\mathcal{X}_N(x) = \{T^j(x)\}_{j=0}^{N-1} \setminus \{\text{The } K \text{ closest points to } 0\}$$

[2] By this, we mean that $(f \circ T^j)$ forms a sequence of iid random variables. For the sake of being explicit, one might imagine that $T^j(x)$ is replaced by a sequence of iid uniformly distributed points on $[0, 1]$.

[3] To be more precise, the references below study the joint distribution of $S_N(f)$ and $\max_{0 \leq j_1 < \ldots < j_K \leq N-1} \sum_{i=1}^K f(T^{j_i}(x))$, from which statements about $S_N^K(f)$ can be deduced.

and

$$\mathcal{N}_N^j = \#\left(\mathcal{X}_N \cap A_N^j\right).$$

Then, for points x that do not visit A_N^0, we obtain the bound

$$\sum_{j=1}^{m-1}\left(\frac{Nm}{R(j+1)}\right)^{\beta}\mathcal{N}_N^j \le S_N^K(f_N) \le \sum_{j=1}^{m-1}\left(\frac{Nm}{Rj}\right)^{\beta}\mathcal{N}_N^j.$$

(iii) Let $\Lambda = \{\xi_j\}$ be a Poisson point process[4] on $[0,\infty)$, with the points ordered so that $\xi_1 < \xi_2 < ...$, and let $\Lambda^K = \{\xi_j\}_{j=K+1}^{\infty}$. By a natural extension of the PLT, it holds that

$$\left(\mathcal{N}_N^j\right)_{j\ge 1} \Rightarrow \left(\#\left(\Lambda^K \cap \left[\frac{Rj}{m}, \frac{R(j+1)}{m}\right]\right)\right)_{j\ge 1}.$$

This assertion follows directly from the simpler statement

$$\left(S_N\left(1_{\left[\frac{Rj}{Nm}, \frac{R(j+1)}{Nm}\right]}\right)\right)_{j\ge 1} \Rightarrow \left(\#\left(\Lambda^K \cap \left[\frac{Rj}{m}, \frac{R(j+1)}{m}\right]\right)\right)_{j\ge 1},$$

because

$$\sum_{j=0}^{J}\mathcal{N}_N^j = \min\left(S_N\left(1_{\left[0, \frac{R(J+1)}{Nm}\right]}\right) - K, 0\right), \quad \text{and}$$

$$\#\left(\Lambda^K \cap \left[0, \frac{R(J+1)}{m}\right]\right) = \min\left(\#\left(\Lambda \cap \left[0, \frac{R(J+1)}{m}\right]\right) - K, 0\right).$$

The PLT as introduced in Chapter 2 asserts that, for each $J \ge 1$,

$$S_N\left(1_{\left[0, \frac{R(J+1)}{Nm}\right]}\right) \Rightarrow \#\left(\Lambda \cap \left[0, \frac{R(J+1)}{m}\right]\right), \tag{4.7}$$

so it remains to show that the $S_N\left(1_{\left[\frac{Rj}{Nm}, \frac{R(j+1)}{Nm}\right]}\right)$ are asymptotically independent.

(iv) Letting $N \to \infty$ and then $m \to \infty$, we see that

$$\frac{S_N^K(f_N)}{N^{\beta}} \Rightarrow \sum_{j\ge K, \xi_j \le R}\frac{1}{\xi_j^{\beta}}.$$

Finally, as $R \to \infty$, it is straightforward to verify that

$$\sum_{j\ge K, \xi_j\le R}\frac{1}{\xi_j} - \log(R) \quad \text{and} \quad \sum_{j\ge K, \xi_j \le R}\frac{1}{\xi_j^{\beta}} \text{ for } \beta > 1,$$

converge in distribution.

A similar proof appears to hold in the case of intermediate trimming, where we must instead analyze

$$\sum_{j\ge k(N), \xi_j\le R}\frac{1}{\xi_j^{\beta}},$$

which, with a suitable choice of R, under correct normalization, has a normal limit.

These heuristics should also extend to cases where the limit in (4.7) is not Poisson but follows a different distribution. Inspired by §4.2.3, one could naturally consider a scenario where the function f exhibits a singularity at the cusp of $SL(2,\mathbb{R})/\Gamma$. It would be interesting to explore the limiting distributions of $S_N(f)$ and $S_N^{k(N)}(f)$ that emerge from the hitting time statistics in [MP24].

[4]For a sequence (E_j) of standard exponentially distributed random variables it holds that $\xi_J = \sum_{j=1}^{J} E_j$.

4.4 Trimmed distributional limit theorems for rotations

A particular case of the questions raised in the previous section is distributional convergence of $S_N(f)$ or $S_N^{k(N)}(f)$ after suitable normalisation for irrational rotation. However, especially considering §4.2.1, an approach as in 4.3 via the PLT seems a bit unnatural.

The untrimmed version of this question has already been studied in [SU08] and [DF15], however there the rotation number α is randomised. Their proof relies on the mixing properties of the geodesic flow.

Let us emphasize that the stochasticity has to come entirely from visits close to the singularity, everything far away contributes almost nothing. More explicitly; for some small $\epsilon > 0$ let χ be a smooth function approximating $f \cdot 1_{(\epsilon,1)}$, where $f(X) = x^{-\beta}$, then using Denjoy-Koksma, for almost all α, it holds that

$$S_N(\chi) \ll N\epsilon^{1-\beta} + \log(N)\epsilon^{-\beta} \quad \text{if } \beta > 1,$$

this bound is optimised for $\epsilon = \frac{\log(N)}{N}$. Considering [DF15] this looks promising, since $S_N(f)$ should be normalised by N^β.

Now one has to study the contribution of $f - \chi$. In the mixing case, we were able to rely on the PLT, which is no longer valid for rotations. This presents a significant challenge, as standard probabilistic tools break down in the absence of mixing, leaving only heavy computational approaches. An interesting open question is whether a more conceptual method, possibly making use of renormalisation via the Gauss-map, could be developed to establish a meaningful limit theorem in this setting.

4.5 Trimming for other functions

A natural extension of the results in Chapter 3 is to investigate whether similar trimmed limit theorems hold for a broader class of observables, possibly with more than one power singularity like

$$\frac{\cos(2\pi x)}{\cos(11\pi x) + \sin(12\pi x)}.$$

Let us first remark that the techniques and results of §3.3 extend without difficulty to general functions with one power singularity of the form

$$f(x) = c_1 x^{-\beta} + c_2(1-x)^{-\beta} + \tilde{f},$$

for real numbers $c_1 \neq -c_2$, a bounded function $\tilde{f}$ and $\beta > 1$. The contribution of $S_N(\tilde{f})$ is of order N and therefore dominated by $S_N(f - \tilde{f})$. Furthermore as already mentioned in §3.2, if there are multiple singularities of different orders, then the one with the largest order will dominate the rest. The situation becomes more complicated if there are multiple singularities of the same order.

To illustrate a potential problem, let us consider the function $f(x) = f_1(x) + f_2(x) = |x|^{-\beta} + |x - x_0|^{-\beta}$ with $\beta > 1$ then;

- On the one hand, intuitively the asymptotics of $S_N(f)$ should depend on the Diophantine properties of the singularities, in this case of x_0. In the extreme case if $x_0 = \alpha$ then $S_N(f) \approx 2S_N(f_1)$. However, now consider the case where $x_0 \in \bigcup_{j=0}^{q_n-1} \left(-\frac{1}{8q_n}, -\frac{1}{16q_n}\right) + j\alpha$ for infinitely many n say for a sequence (n_l), this concerns a positive measure set of x_0 since the sets have measure $\frac{1}{16}$. Now, if the $\left\lfloor \frac{q_{n+1}}{2} \right\rfloor$ orbit of some point x comes $\frac{1}{q_{n+1}}$ close to 0 then it only comes at most $\frac{1}{16q_n} - \frac{1}{q_{n+1}}$ close to x_0, and $S_N(f_1)(x)$ dominates $S_N(f_2)(x)$.

- On the other hand, by the arguments at the end of §3.2, if the trimming sequence $k(N)$ is big enough and α is Diophantine, then the asymptotics of $S_N^{k(N)}(f)$ do not depend on x_0 and it holds that

$$\lim_{N\to\infty} \frac{S_N(f)(x)}{N^\beta k(N)^{1-\beta}} = \frac{2}{\beta - 1} \quad \text{uniformly in } x.$$

This suggests that the dependence on x_0 becomes less the more we trim, and is rather subtle.

Bibliography

[Aar77] J. Aaronson, *On the ergodic theory of non-integrable functions and infinite measure spaces*, Israel J. Math. **27** (1977), no. 2, 163–173.

[Aar97] J. Aaronson, *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, Providence, RI, 1997.

[AB60] D. Z. Arov and A. A. Bobrov, *The extreme members of a sample and their role in the sum of independent variables*, Teor. Verojatnost. i Primenen. **5** (1960), 415–435.

[AD01] J. Aaronson and M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stoch. Dyn. **1** (2001), no. 2, 193–237.

[AFS23] A. Avila, G. Forni, and P. Safaee, *Quantitative weak mixing for interval exchange transformations*, Geom. Funct. Anal. **33** (2023), no. 1, 1–56.

[AN03] J. Aaronson and H. Nakada, *Trimmed sums for non-negative, mixing stationary processes*, Stochastic Process. Appl. **104** (2003), no. 2, 173–192.

[AS11] M. Abadi and B. Saussol, *Hitting and returning to rare events for all alpha-mixing processes*, Stochastic Process. Appl. **121** (2011), no. 2, 314–323.

[AS25] M. Auer and T. I. Schindler, *Trimmed ergodic sums for non-integrable functions with power singularities over irrational rotations*, 2025, arXiv:2503.22242, manuscript submitted for publication at J. Lond. Math. Soc.

[Aue25] M. Auer, *Poisson limit theorems for systems with product structure*, Discrete Contin. Dyn. Syst. **45** (2025), no. 5, 1454–1491.

[BEG20] M. Björklund, M. Einsiedler, and A. Gorodnik, *Quantitative multiple mixing*, J. Eur. Math. Soc. (JEMS) **22** (2020), no. 5, 1475–1529.

[Bil95] P. Billingsley, *Probability and measure*, John Wiley & Sons, 1995.

[BK21] P. Berk and A. Kanigowski, *Spectral disjointness of rescalings of some surface flows*, J. Lond. Math. Soc. (2) **103** (2021), no. 3, 901–942.

[Bra05] R. C. Bradley, *Basic properties of strong mixing conditions. A survey and some open questions*, Probab. Surv. **2** (2005), 107–144, Update of, and a supplement to, the 1986 original.

[BRV16] V. Beresnevich, F. Ramírez, and S. Velani, *Metric Diophantine approximation: aspects of recent work*, Dynamics and analytic number theory, London Math. Soc. Lecture Note Ser., vol. 437, Cambridge Univ. Press, Cambridge, 2016, pp. 1–95.

[BS01] L. Barreira and B. Saussol, *Hausdorff dimension of measures via Poincaré recurrence*, Comm. Math. Phys. **219** (2001), no. 2, 443–463.

[BS22] L. A. Bunimovich and Y. Su, *Poisson approximations and convergence rates for hyperbolic dynamical systems*, Comm. Math. Phys. **390** (2022), no. 1, 113–168.

[BS23] ———, *Maximal large deviations and slow recurrences in weakly chaotic systems*, Adv. Math. **432** (2023), 58.

[BS24] ———, *Back to boundaries in billiards*, Comm. Math. Phys. **405** (2024), no. 6, 74.

[BSTV03] H. Bruin, B. Saussol, S. Troubetzkoy, and S. Vaienti, *Return time statistics via inducing*, Ergodic Theory Dynam. Systems **23** (2003), no. 4, 991–1013.

[CC13] J.-R. Chazottes and P. Collet, *Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **33** (2013), no. 1, 49–80.

[CdF96] Z. Coelho and E. de Faria, *Limit laws of entrance times for homeomorphisms of the circle*, Israel J. Math. **93** (1996), 93–112.

[CG93] P. Collet and A. Galves, *Statistics of close visits to the indifferent fixed point of an interval map*, J. Statist. Phys. **72** (1993), no. 3-4, 459–478.

[Cha08] T. K. Chandra, *The Borel-Cantelli lemma under dependence conditions*, Statist. Probab. Lett. **78** (2008), no. 4, 390–395.

[CoHM86] S. Csörgő, L. Horváth, and D. M. Mason, *What portion of the sample makes a partial sum asymptotically stable or normal?*, Probab. Theory Relat. Fields **72** (1986), no. 1, 1–16.

[Dar52] D. A. Darling, *The influence of the maximum term in the addition of independent random variables*, Trans. Amer. Math. Soc. **73** (1952), 95–107.

[DDKN22a] D. Dolgopyat, C. Dong, A. Kanigowski, and P. Nándori, *Flexibility of statistical properties for smooth systems satisfying the central limit theorem*, Invent. Math. **230** (2022), no. 1, 31–120.

[DDKN22b] ———, *Mixing properties of generalized T, T^{-1} transformations*, Israel J. Math. **247** (2022), no. 1, 21–73.

[DF15] D. Dolgopyat and B. Fayad, *Limit theorems for toral translations*, Hyperbolic dynamics, fluctuations and large deviations, Proc. Sympos. Pure Math., vol. 89, Amer. Math. Soc., Providence, RI, 2015, pp. 227–277.

[DFL22] D. Dolgopyat, B. Fayad, and S. Liu, *Multiple Borel-Cantelli lemma in dynamics and multilog law for recurrence*, J. Mod. Dyn. **18** (2022), 209–289.

[DGS04] M. Denker, M. Gordin, and A. Sharova, *A Poisson limit theorem for toral automorphisms*, Illinois J. Math. **48** (2004), no. 1, 1–20.

[DKH24] D. Dolgopyat, A. Kanigowski, and F. Rodriguez Hertz, *Exponential mixing implies Bernoulli*, Annals of Mathematics **199** (2024), no. 3, 1225 – 1292.

[DMS17] C. P. Dettmann, J. Marklof, and A. Strömbergsson, *Universal hitting time statistics for integrable flows*, J. Stat. Phys. **166** (2017), no. 3-4, 714–749.

[Dol04] D. Dolgopyat, *Limit theorems for partially hyperbolic systems*, Trans. Amer. Math. Soc. **356** (2004), no. 4, 1637–1689.

[DV86] H. G. Diamond and J. D. Vaaler, *Estimates for partial sums of continued fraction partial quotients*, Pacific J. Math. **122** (1986), no. 1, 73–82.

[Fel57] W. Feller, *An introduction to probability theory and its applications, volume 2*, An Introduction to Probability Theory and Its Applications, no. v. 1-2, Wiley, 1957.

[FF03] L. Flaminio and G. Forni, *Invariant distributions and time averages for horocycle flows*, Duke Math. J. **119** (2003), no. 3, 465–526.

[FF06] ______, *Equidistribution of nilflows and applications to theta sums*, Ergodic Theory Dynam. Systems **26** (2006), no. 2, 409–433.

[FFT11] A. C. M. Freitas, J. M. Freitas, and M. Todd, *Extreme value laws in dynamical systems for non-smooth observations*, J. Stat. Phys. **142** (2011), no. 1, 108–126.

[FFT16] L. Flaminio, G. Forni, and J. Tanis, *Effective equidistribution of twisted horocycle flows and horocycle maps*, Geom. Funct. Anal. **26** (2016), no. 5, 1359–1448.

[FFV17] A. C. M. Freitas, J. M. Freitas, and S. Vaienti, *Extreme value laws for non stationary processes generated by sequential and random dynamical systems*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 3, 1341–1370.

[FK16] B. Fayad and A. Kanigowski, *Multiple mixing for a class of conservative surface flows*, Invent. Math. **203** (2016), no. 2, 555–614.

[FU12] G. Forni and C. Ulcigrai, *Time-changes of horocycle flows*, J. Mod. Dyn. **6** (2012), no. 2, 251–273.

[GK68] B.V. Gnedenko and A.N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley Mathematical Series, Addison-Wesley, 1968.

[Hae93] E. Haeusler, *A nonstandard law of the iterated logarithm for trimmed sums*, Ann. Probab. **21** (1993), no. 2, 831–860.

[Hal78] P. Hall, *On the extreme terms of a sample from the domain of attraction of a stable law*, J. London Math. Soc. (2) **18** (1978), no. 1, 181–191.

[Hay14] A. Haynes, *Quantitative ergodic theorems for weakly integrable functions*, Ergodic Theory Dynam. Systems **34** (2014), no. 2, 534–542.

[HLSW22] M. Hussain, B. Li, D. Simmons, and B. Wang, *Dynamical Borel-Cantelli lemma for recurrence theory*, Ergodic Theory Dynam. Systems **42** (2022), no. 6, 1994–2008.

[HLV05] N. Haydn, Y. Lacroix, and S. Vaienti, *Hitting and return times in ergodic dynamical systems*, Ann. Probab. **33** (2005), no. 5, 2043–2050.

[HM87] E. Haeusler and D. M. Mason, *Laws of the iterated logarithm for sums of the middle portion of the sample*, Math. Proc. Cambridge Philos. Soc. **101** (1987), no. 2, 301–312.

[HNPV13] N. Haydn, M. Nicol, T. Persson, and S. Vaienti, *A note on Borel-Cantelli lemmas for non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **33** (2013), no. 2, 475–498.

[HRY19] N. Haydn, J. Rousseau, and F. Yang, *Exponential law for random maps on compact manifolds*, 2019.

[HSV99] M. Hirata, B. Saussol, and S. Vaienti, *Statistics of return times: a general framework and new applications*, Comm. Math. Phys. **206** (1999), no. 1, 33–55.

[HV09] N. Haydn and S. Vaienti, *The compound Poisson distribution and return times in dynamical systems*, Probab. Theory Related Fields **144** (2009), no. 3-4, 517–542.

[Kes93] H. Kesten, *Convergence in distribution of lightly trimmed and untrimmed sums are equivalent*, Math. Proc. Cambridge Philos. Soc. **113** (1993), no. 3, 615–638.

[Khi35] A. Khintchine, *Metrische Kettenbruchprobleme*, Compositio Math. **1** (1935), 361–382.

[KL05] M. Kupsa and Y. Lacroix, *Asymptotics for hitting times*, Ann. Probab. **33** (2005), no. 2, 610–619.

[KM92] H. Kesten and R. A. Maller, *Ratios of trimmed sums and order statistics*, Ann. Probab. **20** (1992), no. 4, 1805–1842.

[KM95] ______, *The effect of trimming on the strong law of large numbers*, Proc. London Math. Soc. (3) **71** (1995), no. 2, 441–480.

[KM96] D. Y. Kleinbock and G. A. Margulis, *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, Sinaĭ's Moscow Seminar on Dynamical Systems, Amer. Math. Soc. Transl. Ser. 2, vol. 171, Amer. Math. Soc., Providence, RI, 1996, pp. 141–172.

[KMK24] M. Kirsebom and K. Mallahi-Karai, *On an extreme value law for the unipotent flow on* $\mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$, 2024, arXiv:2209.07283.

[KS19] M. Kesseböhmer and T. Schindler, *Strong laws of large numbers for intermediately trimmed Birkhoff sums of observables with infinite mean*, Stochastic Process. Appl. **129** (2019), no. 10, 4163–4207.

[KS20] M. Kesseböhmer and T. I. Schindler, *Mean convergence for intermediately trimmed Birkhoff sums of observables with regularly varying tails*, Nonlinearity **33** (2020), no. 10, 5543–5566. MR 4151417

[Lac02] Y. Lacroix, *Possible limit laws for entrance times of an ergodic aperiodic dynamical system*, Israel J. Math. **132** (2002), 253–263.

[Lam63] J. Lamperti, *Wiener's test and Markov chains*, J. Math. Anal. Appl. **6** (1963), 58–66.

[LFF+16] V. Lucarini, D. Faranda, A. C. M. Freitas, J. M. Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd, and S. Vaienti, *Extremes and recurrence in dynamical systems*, Pure and Applied Mathematics (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2016.

[Liv95] C. Liverani, *Decay of correlations*, Annals of Mathematics **142** (1995), no. 2, 239–301.

[LM13] S. Luzzatto and I. Melbourne, *Statistical properties and decay of correlations for interval maps with critical points and singularities*, Comm. Math. Phys. **320** (2013), no. 1, 21–35.

[Mal84] R. A. Maller, *Relative stability of trimmed sums*, Z. Wahrsch. Verw. Gebiete **66** (1984), no. 1, 61–80.

[Mor76] T. Mori, *The strong law of large numbers when extreme terms are excluded from sums*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **36** (1976), no. 3, 189–194.

[Mor77] ______, *Stability for sums of i.i.d. random variables when extreme terms are excluded*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **40** (1977), no. 2, 159–167.

[MP24] J. Marklof and M. Pollicott, *Extreme events for horocycle flows*, 2024, preprint, arXiv:2408.01781.

[Pet02] V. V. Petrov, *A note on the Borel-Cantelli lemma*, Statist. Probab. Lett. **58** (2002), no. 3, 283–286.

[Pit91] B. Pitskel, *Poisson limit law for Markov chains*, Ergodic Theory Dynam. Systems **11** (1991), no. 3, 501–513.

[PS16] F. Pène and B. Saussol, *Poisson law for some non-uniformly hyperbolic dynamical systems with polynomial rate of mixing*, Ergodic Theory Dynam. Systems **36** (2016), no. 8, 2602–2626.

[PS20] ———, *Spatio-temporal Poisson processes for visits to small sets*, Israel J. Math. **240** (2020), no. 2, 625–665.

[PY17] M. J. Pacifico and F. Yang, *Hitting times distribution and extreme value laws for semi-flows*, Discrete Contin. Dyn. Syst. **37** (2017), no. 11, 5861–5881.

[RSV14] J. Rousseau, B. Saussol, and P. Varandas, *Exponential law for random subshifts of finite type*, Stochastic Process. Appl. **124** (2014), no. 10, 3260–3276.

[Ryc83] M. Rychlik, *Bounded variation and invariant measures*, Studia Mathematica **76** (1983), no. 1, 69–80 (eng).

[Sch18] T. Schindler, *Trimmed sums for observables on the doubling map*, 2018, arXiv:1810.03223.

[SU08] Y. G. Sinai and C. Ulcigrai, *A limit theorem for Birkhoff sums of non-integrable functions over rotations*, Geometric and probabilistic structures in dynamics, Contemp. Math., vol. 469, Amer. Math. Soc., Providence, RI, 2008, pp. 317–340.

[Teu81] J. L. Teugels, *Limit theorems on order statistics*, Ann. Probab. **9** (1981), no. 5, 868–880. MR 628879

[You92] L.-S. Young, *Decay of correlations for certain quadratic maps*, Comm. Math. Phys. **146** (1992), no. 1, 123–138.

[Zwe16] R. Zweimüller, *The general asymptotic return-time process*, Israel J. Math. **212** (2016), no. 1, 1–36.

[Zwe22] ———, *Hitting times and positions in rare events*, Ann. H. Lebesgue **5** (2022), 1361–1415.