

# ALMOST SURE ASYMPTOTIC BEHAVIOUR OF BIRKHOFF SUMS FOR INFINITE MEASURE-PRESERVING DYNAMICAL SYSTEMS

CLAUDIO BONANNO AND TANJA I. SCHINDLER

ABSTRACT. We consider a conservative ergodic measure-preserving transformation  $T$  of a  $\sigma$ -finite measure space  $(X, \mathcal{B}, \mu)$  with  $\mu(X) = \infty$ . Given an observable  $f : X \rightarrow \mathbb{R}$ , we study the almost sure asymptotic behaviour of the Birkhoff sums  $S_N f(x) := \sum_{j=1}^N (f \circ T^{j-1})(x)$ . In infinite ergodic theory it is well known that the asymptotic behaviour of  $S_N f(x)$  strongly depends on the point  $x \in X$ , and if  $f \in L^1(X, \mu)$ , then there exists no real valued sequence  $(b(N))$  such that  $\lim_{N \rightarrow \infty} S_N f(x)/b(N) = 1$  almost surely. In this paper we show that for dynamical systems with strong mixing assumptions for the induced map on a finite measure set, there exists a sequence  $(\alpha(N))$  and  $m : X \times \mathbb{N} \rightarrow \mathbb{N}$  such that for  $f \in L^1(X, \mu)$  we have  $\lim_{N \rightarrow \infty} S_{N+m(x,N)} f(x)/\alpha(N) = 1$  for  $\mu$ -a.e.  $x \in X$ . Instead in the case  $f \notin L^1(X, \mu)$  we give conditions on the induced observable such that there exists a sequence  $(G(N))$  depending on  $f$ , for which  $\lim_{N \rightarrow \infty} S_N f(x)/G(N) = 1$  holds for  $\mu$ -a.e.  $x \in X$ .

## 1. INTRODUCTION

Let  $T : X \rightarrow X$  be a conservative ergodic measure-preserving transformation of a  $\sigma$ -finite measure space  $(X, \mathcal{B}, \mu)$  and consider an observable  $f$ , that is a measurable function  $f : X \rightarrow \mathbb{R}$ . The almost sure asymptotic behaviour of the *Birkhoff sums of  $f$*

$$S_N f(x) := \sum_{j=1}^N (f \circ T^{j-1})(x)$$

is the main result of the famous Birkhoff's Ergodic Theorem. It states that if  $\mu$  is a probability measure and  $f \in L^1(X, \mu)$ , then  $S_N f(x)$  is asymptotic to  $N \int_X f d\mu$  as  $N \rightarrow \infty$  for  $\mu$ -a.e.  $x \in X$ . On the other hand, if  $f \notin L^1(X, \mu)$  the almost sure asymptotic behaviour of  $S_N f$  is not described by any sequence  $(b(N))$ . In particular, letting  $f : X \rightarrow \mathbb{R}_{\geq 0}$ , one can apply [A97, Cor. 2.3.4] to show that given a sequence  $(b(N))$  of positive numbers with  $b(N)/N \rightarrow \infty$ , for  $\mu$ -almost every (a.e.)  $x \in X$  either  $\limsup_N S_N f(x)/b(N) = \infty$  or  $\liminf_N S_N f(x)/b(N) = 0$ .

The situation is different if the measure  $\mu$  is  $\sigma$ -finite but  $\mu(X) = \infty$ . In this case if  $f \in L^1(X, \mu)$  then  $S_N f(x) = o(N)$  for  $\mu$ -a.e.  $x \in X$ , but again the almost sure asymptotic behaviour of  $S_N f$  is not described by any sequence  $(b(N))$  slower than  $N$ . It is the content of [A97, Thm. 2.4.2] that given a sequence  $(b(N))$  of positive numbers, even with  $b(N)/N \rightarrow 0$ , for  $\mu$ -a.e.  $x \in X$  either  $\liminf_N S_N f(x)/b(N) = 0$  for all non-negative  $f \in L^1(X, \mu)$ , or there exists a sequence  $(N_k)$  such that  $\lim_k S_{N_k} f(x)/b(N_k) = \infty$  for all non-negative  $f \in L^1(X, \mu)$ . Despite the impossibility to obtain a strong law of large numbers for infinite measure-preserving dynamical systems and an  $L^1$  function, there are some results about the almost sure limsup and liminf behaviour of the Birkhoff

---

2020 *Mathematics Subject Classification.* 37A40, 37A25, 60F15.

*Key words and phrases.* Infinite ergodic theory; almost sure limits for Birkhoff sums;  $\psi$ -mixing; trimmed sums.

The authors are partially supported by the PRIN Grant 2017S35EHN\_004 "Regular and stochastic behaviour in dynamical systems" of the Italian Ministry of University and Research (MUR), Italy. This research is part of the authors' activity within the UMI Group "DinAmicI" [www.dinamici.org](http://www.dinamici.org) and of the first author's activity within the Gruppo Nazionale di Fisica Matematica, INdAM, Italy. The second author acknowledges the support of the Centro di Ricerca Matematica Ennio de Giorgi and of UniCredit Bank R&D group for financial support through the 'Dynamics and Information Theory Institute' at the Scuola Normale Superiore.

sum, see for instance [AD90] and [GHPZ], the latter one as a result in the study of dynamical Borel-Cantelli lemmas.

Furthermore, an information about the almost sure asymptotic behaviour of the Birkhoff sums of summable observables is given in Hopf's Ratio Ergodic Theorem (see [A97, Thm 2.2.5]). It is shown that for  $f, g \in L^1(X, \mu)$  with  $g \geq 0$  and  $\int_X g d\mu > 0$ , the ratio  $S_N f(x)/S_N g(x)$  converges to  $\int_X f d\mu / \int_X g d\mu$  as  $N \rightarrow \infty$  for  $\mu$ -a.e.  $x \in X$ .

For completeness, we also want to add that [AKZ17, K17] give almost sure limit laws for symmetric Birkhoff sums in the setting of an infinite measure space and an integrable observable.

The almost sure asymptotic behaviour of Birkhoff sums of non-summable observables in the context of infinite ergodic theory has been studied in [I94, I00, I04, ATZ05] in terms of the ratio of two ergodic sums as in Hopf's theorem, partly restricted to indicator functions. Furthermore, in [CN17] the liminf and limsup behaviour of Birkhoff sums in this setting has been studied and in [LM18] a sufficient criterion is given for a strong law of large numbers to hold.

In this paper we restrict our attention to the case that  $\mu$  is infinite and study the almost sure asymptotic behaviour of Birkhoff sums for both summable and non-summable observables with respect to sequences. We consider dynamical systems with strong mixing assumptions, namely  $\psi$ -mixing (see Definition 2.1), for the induced map on a finite measure subset  $E \subset X$  (see Section 2 for the detailed assumptions), and prove two main results. First we prove that for  $f \in L^1(X, \mu)$  there exist sequences of positive real numbers  $(\alpha(N))$  and  $(m(N, E, x))$  such that  $S_{N+m(N, E, x)} f(x)$  is asymptotic to  $\alpha(N) \int_X f d\mu$  as  $N \rightarrow \infty$  for  $\mu$ -a.e.  $x \in X$ . That is to obtain the same asymptotic behaviour for  $\mu$ -a.e.  $x \in X$  we have to change the number of terms to consider in the Birkhoff sums (see Theorem 2.3). Then we prove that for a non-summable observable  $f : X \rightarrow \mathbb{R}_{\geq 0}$ , under suitable assumptions on its induced version  $f^E$  (see (2.12)) it is possible to find a sequence  $(G(N))$  which depends on  $f$  such that  $S_N f(x) \sim G(N)$  as  $N \rightarrow \infty$  for  $\mu$ -a.e.  $x \in X$  (see Theorem 2.7). The main difference is that in this non-summable case the properties of  $f$  play an important role.

The proofs rely on the method of *trimmed sums* (see (2.8)), that is sums from which a number of largest entries is deleted. The method to trim sums of independent random variables to prove (pointwise and distributional) limit theorems which fail to hold for the untrimmed sum is long established and most results were developed in the 80th and 90th of the last century. One generally differentiates between different strengths of trimming, according to the number  $r$  of largest entries deleted. In the light trimming case the number of deleted entries  $r$  is constant for all  $n$ , while in the intermediate trimming case the number of deleted entries  $r(n)$  depends on  $n$  and tends to infinity, but with  $r(n) = o(n)$ .

The first result using trimming in the dynamical systems context was given by Diamond and Vaaler [DV86] providing a strong law of large numbers for the continued fraction digits under the use of light trimming. More precisely, if  $(a_n(x))$  are the coefficients of the continued fraction expansion of a point  $x \in [0, 1]$ , it is proved in [DV86] that  $\lim_{n \rightarrow \infty} (\sum_{k=1}^n a_k(x) - \max_{1 \leq k \leq n} a_k(x)) / (n \log n) = \log 2$  holds for Lebesgue-a.e.  $x \in [0, 1]$ . Hence in this case it is enough to trim the sum of the coefficients  $(a_n(x))$  by deleting only the largest entry, so  $r = 1$ . These results were generalised by Aaronson and Nakada [AN03] giving strong laws of large numbers under light trimming for sufficiently fast  $\psi$ -mixing random variables, thus giving the analog to some statement for independent random variables from [KM92]. Haynes further quantified in [H14] the results by Aaronson and Nakada by giving precise error terms.

Kesseböhmer and the second author of this paper proved strong laws of large numbers under intermediate trimming using a spectral gap property of the transfer operator, see [KS19b] and [KS20a] for an application of these results to subshifts of finite type and [HM87, H93, KS19a] for these and further reaching results in the independent case. See further [KS20b] for a convergence in mean result concerning the same intermediately trimmed sums.

The paper is organised as follows. In Section 2 we describe the setting we consider and state the main results, Theorems 2.3 and 2.7, whose proofs are collected in Section 4. In Section 3 we describe two classes of dynamical systems to which our results are applicable, and discuss the assumptions on the observables for Theorem 2.7. We also study in more detail our results for the Farey map, a well-known dynamical system on the interval, also in relation to the results in [LM18] which are shown to be a particular case of ours for this system. Finally, two appendices complete the paper. In particular Appendix B contains definitions and results on slowly varying functions, an important class of functions used in the statement of Theorem 2.7.

## 2. THE SETTING AND MAIN RESULTS

Let  $T : X \rightarrow X$  be a conservative ergodic measure-preserving transformation of the measure space  $(X, \mathcal{B}, \mu)$  with  $\mu$  a  $\sigma$ -finite measure with  $\mu(X) = \infty$ . Let  $E \in \mathcal{B}$  with  $\mu(E) = 1$ , and denote by  $\varphi_E$  the *first return time*

$$\varphi_E : E \rightarrow \mathbb{N}, \quad \varphi_E(x) := \min \left\{ k \geq 1 : T^k(x) \in E \right\}.$$

The first return time is finite for  $\mu$ -a.e.  $x \in E$ , so we define  $T_E(x) := T^{\varphi_E(x)}(x)$  to be the *induced map*  $T_E : E \rightarrow E$ , which is an ergodic measure-preserving transformation of the probability space  $(E, \mathcal{B}|_E, \mu)$ . The first return time generates two families of sets, the level sets

$$(2.1) \quad A_n := \{x \in E : \varphi_E(x) = n\}$$

which are a measurable partition of  $E$ , and the super-level sets

$$(2.2) \quad A_{>n} := \{x \in E : \varphi_E(x) > n\} = \bigsqcup_{k>n} A_k.$$

We also use the notation  $A_{\geq n} = A_n \cup A_{>n}$ . Applying Kac's Theorem one has

$$(2.3) \quad \sum_{k \geq 1} k \mu(A_k) = \sum_{n \geq 0} \mu(A_{>n}) = \mu(X) = \infty.$$

A useful notion is the *longest excursion out of  $E$  beginning in the first  $N$ -steps* defined for  $\mu$ -a.e.  $x \in X$  as

$$(2.4) \quad m(N, E, x) := 1 + \max \left\{ k \geq 1 : \exists \ell \in \{1, \dots, N+1\} \text{ s.t. } T^{\ell+j}(x) \notin E, \forall j = 0, \dots, k-1 \right\}.$$

If we denote by  $R_{E,N}(x)$  the *number of visits to  $E$  up to time  $N$*  along the orbit of a point  $x$ , that is

$$(2.5) \quad R_{E,N}(x) := \sum_{k=1}^{N+1} (\mathbb{1}_E \circ T^{k-1})(x)$$

it is immediate to verify that for  $\mu$ -a.e.  $x \in E$

$$(2.6) \quad m(N, E, x) = \max \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N}(x) \right\}.$$

A notion related to  $m(N, E, x)$  is that of the *longest excursion out of  $E$  seen up to time  $N$* , defined for  $\mu$ -a.e.  $x \in E$  by

$$(2.7) \quad w(N, E, x) := 1 + \max \left\{ k \geq 1 : \exists \ell \in \{1, \dots, N-k\} \text{ s.t. } T^{\ell+j}(x) \notin E, \forall j = 0, \dots, k-1 \right\}.$$

Notice that  $w(N, E, x)$  can correspond to an excursion which does not return to  $E$  up to time  $N$ .

In this paper we are interested in studying the pointwise asymptotic behaviour of the Birkhoff sums

$$S_N f(x) := \sum_{n=1}^N (f \circ T^{n-1})(x)$$

for summable and non-summable observables  $f : X \rightarrow \mathbb{R}$ . Let us first state the definitions and the results we need in the general case of random variables. We refer to [B05] for more definitions.

**Definition 2.1.** Let  $(Y_n)$  be a sequence of random variables on a probability space  $(Y, \mathbb{P})$ , and let  $\mathcal{F}_h^k$ , for  $0 \leq h < k \leq \infty$ , be the  $\sigma$ -field generated by  $(Y_n)_{h \leq n \leq k}$ . The sequence  $(Y_n)$  is  $\psi$ -mixing if

$$\psi(n) := \sup \left\{ \left| \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(B)\mathbb{P}(C)} - 1 \right| : B \in \mathcal{F}_0^j, C \in \mathcal{F}_{j+n}^\infty, \mathbb{P}(B) > 0, \mathbb{P}(C) > 0, j \in \mathbb{N} \right\}$$

satisfies  $\psi(n) \rightarrow 0$  as  $n \rightarrow \infty$ .

We now recall the notion of *lightly trimmed sums*. Given a sequence  $(Y_n)$  of random variables on a probability space  $(Y, \mathbb{P})$  and a point  $y \in Y$ , for each  $N \in \mathbb{N}$  we choose a permutation  $\pi$  of  $\{1, 2, \dots, N\}$  such that  $Y_{\pi(1)}(y) \geq Y_{\pi(2)}(y) \geq \dots \geq Y_{\pi(N)}(y)$ . For a given  $r \in \mathbb{N}_0$ , the lightly trimmed sum of the  $(Y_n)$  is defined by

$$(2.8) \quad S_N^r(y) := \sum_{n=r+1}^N Y_{\pi(n)}(y)$$

that is the sum of the first  $N$  random variables trimmed by the largest  $r$  entries.

The next result is adapted from [AN03] as shown in Appendix A.

**Lemma 2.2.** Let  $(Y_n)$  be a sequence of non-negative, identically distributed random variables on a probability space  $(Y, \mathbb{P})$  which is  $\psi$ -mixing with coefficient  $\psi(n)$  fulfilling  $\sum_{n \geq 1} \psi(n)/n < \infty$ . Let  $F$  be the distribution function of the  $(Y_n)$ , that is  $F(y) = \mathbb{P}(Y_1 \leq y)$ , and let for some  $y_0 > 0$

$$(2.9) \quad W := \min \left\{ r \in \mathbb{N} : \int_{y_0}^\infty \left( \frac{y(1-F(y))}{\int_0^y (1-F(t)) dt} \right)^{r+1} \frac{1}{y} dy < \infty \right\},$$

where we set the min as  $\infty$  if such an  $r$  does not exist. Then there exists a sequence  $(d(N))$  such that

$$\lim_{N \rightarrow \infty} \frac{S_N^W(y)}{d(N)} = 1, \quad \mathbb{P}\text{-a.s.}$$

If we set  $a(y) := y / \int_0^y (1-F(t)) dt$ , then  $d(n)$  can be set as the inverse function of  $a(n)$ .

Furthermore, if we denote by  $M_N^r(y)$  the  $r$ -th maximum in  $\{Y_1, \dots, Y_N\}$  then

$$(2.10) \quad \lim_{N \rightarrow \infty} \frac{M_N^r(y)}{d(N)} = 0, \quad \mathbb{P}\text{-a.s.}$$

for all  $r > W$ .

We now state our main results. We first consider the Birkhoff sums of an observable  $f \in L^1(X, \mu)$ . We prove that  $S_N f$  converges almost surely to the spatial mean of the observable if we consider the average with respect to a sequence  $\alpha(n)$  growing slower than  $n$  and consider more iterates for the observable. The dependence on the point  $x$  is given by the number  $m(N, E, x)$ , as defined in (2.4), of iterates for the observable that we have to add. We remark that the sequence  $\alpha(n)$  coincides with that found in the Weak Law of Large Numbers for dynamical system which are pointwise dual ergodic and have wandering rate given by a slowly varying sequence (see [A97, Chapter 3]).

**Theorem 2.3.** Let  $T : X \rightarrow X$  be a conservative ergodic measure-preserving transformation of the measure space  $(X, \mathcal{B}, \mu)$  with  $\mu$  a  $\sigma$ -finite measure with  $\mu(X) = \infty$ , and let  $E \in \mathcal{B}$  with  $\mu(E) = 1$ . Let's assume that the first return time to  $E$  satisfies:

- (i) the sequence of random variables  $(\varphi_E \circ T_E^{n-1})$  with  $n \geq 1$  defined on  $(E, \mu)$  is  $\psi$ -mixing with coefficient  $\psi(n)$  fulfilling  $\sum_{n \geq 1} \psi(n)/n < \infty$ ;
- (ii) the super-level sets of  $\varphi_E$  satisfy

$$\sum_{n \geq 1} \frac{n (\mu(A_{>n}))^2}{(\sum_{j=0}^{n-1} \mu(A_{>j}))^2} < \infty.$$

Then for all  $f \in L^1(X, \mu)$  and for  $\mu$ -a.e.  $x \in X$

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \sum_{k=1}^{N+m(N,E,x)} (f \circ T^{k-1})(x) = \int_X f \, d\mu$$

where

$$(2.11) \quad \alpha(n) := \frac{n}{\sum_{j=0}^{n-1} \mu(A_{>j})}.$$

In the following we also give a slightly different formulation of Theorem 2.3 which uses  $w(N, E, x)$ . The main difference is that the number of iterates for the observable is fixed and the dependence on  $x$  is found in the renormalising sequence. Thus, this second formulation is more similar to Hopf's Ratio Ergodic Theorem in the spirit.

**Theorem 2.4.** Under the same assumptions as in Theorem 2.3, for all  $f \in L^1(X, \mu)$  and for  $\mu$ -a.e.  $x \in X$  it holds

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N - w(N, E, x))} \sum_{k=1}^N (f \circ T^{k-1})(x) = \int_X f \, d\mu$$

where  $w(N, E, x)$  is defined in (2.7) and  $\alpha(n)$  is given in (2.11).

*Remark 2.5.* The asymptotic behaviour of the sequence  $(w(N, E, x))$  has been considered in [GHPZ] for some intermittent maps of the interval with infinite measure. Following the same argument and using the relations between  $m(N, E, x)$  and  $w(N, E, x)$  one can prove that under the conditions of the two previous theorems

$$\lim_{N \rightarrow \infty} \frac{\log m(N, E, x)}{\log N} = \lim_{N \rightarrow \infty} \frac{\log w(N, E, x)}{\log N} = 1$$

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

We have shown that in the case of summable observables the assumptions only depend on the set  $E$  and on its first return time. This is in accordance with what one could expect with Hopf's Ratio Ergodic Theorem in mind. We now study the Birkhoff sums of non-negative observables  $f \notin L^1(X, \mu)$ . In this case we need assumptions also on the observable itself.

Given  $T : X \rightarrow X$  and  $E \in \mathcal{B}$  with  $\mu(E) = 1$  as above, for a measurable function  $f : X \rightarrow \mathbb{R}$ , a standard approach is to introduce the induced observable  $f^E$  defined on the full measure set of points  $x \in E$  with finite first return time by

$$(2.12) \quad f^E(x) := \sum_{k=1}^{\varphi_E(x)} (f \circ T^{k-1})(x).$$

The induced observable simply gives the contribution to the Birkhoff sums  $S_N f$  of the excursions out of  $E$  for the orbit of a point  $x$ . For  $\mu$ -a.e.  $x \in E$  let's define the time  $\tau_{E,x}(N)$  of the  $N$ -th return to  $E$ , which corresponds to the Birkhoff sum of  $\varphi_E$  for the system  $(E, T_E)$ . We have

$$(2.13) \quad \tau_{E,x}(N) := \sum_{k=1}^N (\varphi_E \circ T_E^{k-1})(x).$$

Recalling now the notion  $R_{E,N}(x)$  of the number of visits to  $E$  up to time  $N$  along the orbit of  $x$  defined in (2.5), the last visit to  $E$  for  $x$  up to time  $N$  is  $T_E^{R_{E,N}(x)-1}(x)$  and the time of this last visit is then given by  $\tau_{E,x}(R_{E,N}(x) - 1) \leq N$ . Then the following relation follows for  $\mu$ -a.e.  $x \in E$

$$(2.14) \quad S_N f(x) = \sum_{n=1}^{R_{E,N}(x)-1} (f^E \circ T_E^{n-1})(x) + \sum_{k=\tau_{E,x}(R_{E,N}(x)-1)+1}^N (f \circ T^{k-1})(x)$$

with the standard convention that a summation vanishes if the upper index is smaller than the lower index. Moreover, using that  $f \geq 0$  and  $\tau_{E,x}(R_{E,N}(x)) > N$ , being  $\tau_{E,x}(R_{E,N}(x))$  the time of the first visit to  $E$  after time  $N$ , we immediately obtain

$$\sum_{n=1}^{R_{E,N}(x)-1} (f^E \circ T_E^{n-1})(x) \leq S_N f(x) \leq \sum_{n=1}^{R_{E,N}(x)} (f^E \circ T_E^{n-1})(x)$$

for  $\mu$ -a.e.  $x \in E$ . Thus the pointwise convergence of the Birkhoff sums of  $f$  could in principle be obtained by the behaviour of the Birkhoff sums of  $f^E$  for the induced system  $(E, T_E, \mu)$ . However  $f \notin L^1(X, \mu)$  implies  $f^E \notin L^1(E, \mu)$  (see for examples [Z09, Lemma 2]), hence we cannot use Birkhoff's Theorem to obtain the pointwise convergence for the Birkhoff sums of  $f^E$ . A solution is to apply Lemma 2.2 to the sequence of random variables  $(f^E \circ T_E^{n-1})$ .

The simplest case to consider is when  $f^E$  is constant on the level sets  $A_n$ , which corresponds to  $f$  constant on the level sets of the *hitting time function*

$$(2.15) \quad h_E : X \rightarrow \mathbb{N}_0, \quad h_E(x) := \min\{k \geq 0 : T^k(x) \in E\}$$

which is defined and finite  $\mu$ -a.e. on  $X$ . Setting

$$(2.16) \quad E_n := \{x \in X : h_E(x) = n\}$$

with  $E_0 = E$ , for the level sets of  $h_E$ , if  $f(x) = f_k$  for all  $x \in E_k$  then

$$(2.17) \quad f^E(x) = \sum_{k=0}^{n-1} f_k, \quad \forall x \in A_n.$$

We recall that  $\mu(A_{>n}) = \mu(E_n)$  for all  $n \geq 0$  (see [Z09, Lemma 1]).

In the proof of the main result of this section, we use the following property for the waiting times between two consecutive returns to  $E$ .

**Lemma 2.6.** *Let us consider the function  $q : \mathbb{R}_{>0} \rightarrow \mathbb{N}$  given by*

$$q(t) := \min \left\{ j \in \mathbb{N} : \mu(A_{>j}) \leq \frac{1}{t} \right\}$$

for the super-level sets  $(A_{>n})$  defined in (2.2). Then

$$\mu \left( \left\{ x \in E : \max_{1 \leq k \leq R_{E,n}(x)} (\varphi_E \circ T_E^{k-1})(x) \geq n \xi(n) \text{ for infinite } n \right\} \right) = 0$$

where

$$\xi(n) := \frac{1}{n} q\left(\alpha(n) \log(\alpha(n)) \log^2(\log(\alpha(n)))\right)$$

and  $\alpha(n)$  is given in (2.11).

**Theorem 2.7.** *Let  $T : X \rightarrow X$  be a conservative ergodic measure-preserving transformation of the measure space  $(X, \mathcal{B}, \mu)$  with  $\mu$  a  $\sigma$ -finite measure with  $\mu(X) = \infty$ , and let  $E \in \mathcal{B}$  with  $\mu(E) = 1$ . Let  $f : X \rightarrow \mathbb{R}_{\geq 0}$  be a non-negative measurable function such that  $f \notin L^1(X, \mu)$  and let  $f^E$  be its induced version on  $E$  defined in (2.12). We assume that:*

- (i) *the sequence of random variables  $(\varphi_E \circ T_E^{n-1})$  with  $n \geq 1$  defined on  $(E, \mu)$  is  $\psi$ -mixing with coefficient  $\psi(n)$  fulfilling  $\sum_{n \geq 1} \psi(n)/n < \infty$ ;*
- (ii) *there exist  $g_1, g_2 : X \rightarrow \mathbb{R}_{\geq 0}$  constant on the level sets  $(E_n)$  of the hitting time function to  $E$  such that:*
  - (ii)-(a)  $g_1(x) \leq f(x) \leq g_2(x)$  for  $\mu$ -a.e.  $x \in X$ ;
  - (ii)-(b) *the induced functions satisfy  $g_i^E(n) := g_i^E|_{A_n} = n L_i(n)$  for  $i = 1, 2$ , where  $L_i(n)$  are normalised slowly varying function (see (B.2)), and  $L_1(n) \sim L_2(n)$  as  $n \rightarrow \infty$ ;*
  - (ii)-(c) *given the function  $\xi(n)$  defined in Lemma 2.6, for all functions  $\tilde{\xi}(n)$  with  $c \leq \tilde{\xi}(n) \leq \xi(n)$  for some constant  $c > 0$ , the functions  $L_i$  for  $i = 1, 2$  satisfy  $L_i(n \tilde{\xi}(n)) \sim L_i(n)$  as  $n \rightarrow \infty$ ;*
  - (ii)-(d) *if  $\Gamma_i(n) := \min\{k \in \mathbb{N} : g_i^E(k) > n\}$  for  $i = 1, 2$ , then*

$$(2.18) \quad \int_0^\infty \frac{y (\mu(A_{\geq \Gamma_i(y)}))^2}{(\int_0^y \mu(A_{\geq \Gamma_i(t)}) dt)^2} dy < \infty;$$

- (ii)-(e) *if  $\ell_i(n) := \sum_{k=0}^n \mu(A_{\geq \Gamma_i(k)})$ , then  $\ell_i(n)$  for  $i = 1, 2$  is super-slowly varying at infinity with rate function itself (see Definition B.5) and*

$$(2.19) \quad \ell_i(\alpha(N)) \sim L_i(N) \sum_{k=0}^N \mu(A_{>k}),$$

as  $N \rightarrow \infty$  for  $i = 1, 2$ , where  $\alpha(n)$  is defined in (2.11).

Then there exists a function  $G : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$  which is asymptotically equivalent to  $f^E$ , that is there exists a set  $\tilde{E} \subset E$  with  $\mu(\tilde{E}) = 1$  such that for every sequence  $(x_n)$  with  $x_n \in A_n \cap \tilde{E}$  for all  $n \in \mathbb{N}$  we have  $f^E(x_n) \sim G(n)$  as  $n \rightarrow \infty$ , and for  $\mu$ -a.e.  $x \in X$

$$\lim_{N \rightarrow \infty} \frac{1}{G(N)} \sum_{k=1}^N (f \circ T^{k-1})(x) = 1.$$

In Section 3.3 we discuss the assumptions of Theorem 2.7 and consider particular cases for the observable  $f$ . In particular we show that the assumptions of Theorem 2.7 get simpler for specific examples. However it turns out that assumption (ii)-(e), which is very technical and seems to be quite restrictive, cannot be dropped (see Proposition 3.8).

From a more general point of view, it would be also interesting to study if the  $\psi$ -mixing condition for the induced map on  $E$ , which is the main assumption on the dynamical systems we consider, could be relaxed. This seems to be highly unlikely in general. Aaronson and Nakada already gave in [AN03] an example of a mixing though not  $\psi$ -mixing system which does not fulfil a strong law of large numbers after trimming, even though a sufficiently fast  $\psi$ -mixing system with the same distribution functions would fulfil a trimmed strong law. Haynes gave another simple example with even stronger mixing properties which does not fulfil a lightly trimmed strong law, see [H14, Thm. 4]. However, this particular system still fulfils an intermediately trimmed strong law, see [S18]. Thus, there might be possibilities to generalise Theorems 2.3 and 2.7 in this way. In particular,

in this paper we have considered assumptions on the system  $(X, T, \mu)$  for which light trimming is enough to get the results, actually it has been enough to trim the sums by deleting only the largest entry. As shown in Section 3.1 for maps of the interval with indifferent fixed points, this is related with the order with which the derivative of the map converges to 1 in the indifferent fixed points. Using light trimming we have been able to consider only the so-called “barely infinite” situation. It is possible that we can obtain similar results for systems with a different behaviour at the indifferent fixed points by using intermediate trimming.

### 3. EXAMPLES AND APPLICATIONS

First we describe two classes of dynamical systems with infinite invariant measure. For the first - interval maps with an indifferent fixed point - we give precise conditions on the parameters such that our results hold. For the second - a multidimensional continued fraction expansion - we conjecture that the theorems can also be applied, however, we point out where the results might fail to hold. Then we discuss the assumptions and the result of Theorem 2.7 in general. Some proofs of statements in this section are collected in Section 5 to ease the reading of the paper.

**3.1. Maps of the interval with an indifferent fixed point.** This class of maps is considered the easiest example of non-uniformly hyperbolic systems, and thus the easiest model of a system with “anomalous” dynamical phenomena. Here we use the general approach introduced in [Z00].

Let  $T : [0, 1] \rightarrow [0, 1]$  be a piecewise monotonic  $C^2$  map with respect to a partition  $\mathcal{P}$  of open subintervals, that is  $\mathcal{P}$  is a finite or countable family of open subintervals  $(P_j)_{j \in J}$  such that  $m([0, 1] \setminus \cup_j P_j) = 0$ , being  $m$  the Lebesgue measure, and  $T|_{P_j}$  is  $C^2$  and strictly monotonic for all  $j \in J$ .

**Definition 3.1.** A map  $T : [0, 1] \rightarrow [0, 1]$  which is piecewise monotonic and  $C^2$  with respect to a partition  $\mathcal{P} = (P_j)_{j \in J}$  of open subintervals defines a *basic AFN-system* if it is conservative and ergodic with respect to  $m$  and if

- (A)  $T''/(T')^2$  is bounded on  $\cup_j P_j$ ;
- (F)  $T\mathcal{P} = \{T(P_j) : j \in J\}$  is a finite collection;
- (N) there exists a non-empty finite collection  $\tilde{\mathcal{P}} \subset \mathcal{P}$  such that each  $P_j \in \tilde{\mathcal{P}}$  has an *indifferent fixed point*  $x_j$ , that is

$$\lim_{x \rightarrow x_j, x \in P_j} T(x) = x_j \quad \text{and} \quad T'(x_j) := \lim_{x \rightarrow x_j, x \in P_j} T'(x) = 1,$$

which is a *one-sided regular source*, that is  $T'$  decreases on  $P_j \cap (0, x_j)$  and increases on  $P_j \cap (x_j, 1)$ . Moreover  $T$  is uniformly expanding away from the indifferent fixed points, that is for all  $\epsilon > 0$  there exists  $\rho_\epsilon > 1$  such that

$$|T'(x)| \geq \rho_\epsilon, \quad \forall x \in [0, 1] \setminus \bigcup_{P_j \in \tilde{\mathcal{P}}} P_j \cap (x_j - \epsilon, x_j + \epsilon).$$

By [Z00, Thm A] a basic AFN-system admits an absolutely continuous invariant measure  $\mu$  with a lower semicontinuous density  $h(x)$  of the form

$$h(x) = c(x)g(x), \quad \text{with} \quad g(x) = \begin{cases} \frac{x-x_j}{x-(T|_{P_j})^{-1}(x)}, & \text{for } x \in P_j \in \tilde{\mathcal{P}} \\ 1, & \text{for } x \in P_j \notin \tilde{\mathcal{P}} \end{cases}$$

and  $c(x)$  satisfying  $0 < C^{-1} \leq c(x) \leq C$  for some constant  $C > 0$ .

Examples of basic AFN-systems are the Pomeau-Manneville maps, the Liverani-Saussol-Vaianti maps and the Farey map. They all share the following structure. Let  $\mathcal{T}$  be the family of maps of  $[0, 1]$  for which there is a partition  $\mathcal{P} = \{P_0, P_1\}$  of  $[0, 1]$  into two open intervals with respect to which they are piecewise monotonic and  $C^2$ , and  $T(P_0) = T(P_1) = (0, 1)$ . Moreover the collection



$\tilde{\mathcal{P}}$  contains only the set  $P_0 = (0, \bar{x})$ , and the maps have an indifferent fixed point at  $x_0 = 0$  at which they satisfy

$$(3.1) \quad T(x) = x + Cx^{1+p} + o(x^{1+p}), \quad \text{for } x \in P_0 \text{ and } x \rightarrow 0^+$$

with  $p \geq 1$  for some constant  $C > 0$ , and are uniformly expanding away from  $x_0$ .

By [Z00, Lemma 8] there is a set  $E \subset [0, 1]$  with  $\mu(E) < \infty$  such that the induced map  $T_E$  is piecewise monotonic and  $C^2$  with respect to a partition  $\mathcal{P}_E$ , satisfies conditions (A) and (F), and is uniformly expanding (U). It is then called an AFU-system. Finally [AN05, Thm 1] gives that if  $T_E$  is weakly mixing and  $\inf_{n \geq 1, Q \in \mathcal{P}_E^n} m(T_E^n(Q)) > 0$ , then the sequence of random variables  $(\varphi_E \circ T_E^{n-1})$  with  $n \geq 1$  defined on  $(E, \mu)$  is  $\psi$ -mixing with coefficient  $\psi(n)$  fulfilling  $\psi(n) \leq K\theta^n$  for some constants  $K > 0$  and  $\theta \in (0, 1)$ . Thus assumption (i) of Theorems 2.3 and 2.7 is satisfied.

For the maps in the family  $\mathcal{T}$  one can choose  $E = P_1$  to obtain an exact AFU-system  $T_E$  which satisfies the assumptions of [AN05, Thm 1].

Moreover we need information on the super-level sets  $(A_{>n})$ .

**Definition 3.2.** We say that a basic AFN-system  $T$  admits *regularly varying nice expansions* if for all  $P_j \in \tilde{\mathcal{P}}$  there are  $p_j \geq 1$  and  $L_j$  slowly varying functions at infinity (see Appendix B) such that

$$T(x) \sim x + |x - x_j|^{1+p_j} (L_j(1/|x - x_j|))^{-1}, \quad \text{for } x \in P_j \text{ and } x \rightarrow x_j.$$

Using then computations similar to those in [T83, Lemma 2], one can show that a basic AFN-system with regularly varying nice expansions satisfies assumption (ii) of Theorem 2.3 if  $p := \max_{P_j \in \tilde{\mathcal{P}}} p_j = 1$  and  $L(y) := \max_{P_j \in \tilde{\mathcal{P}}} |L_j(y)|$  satisfies

$$(3.2) \quad \int_1^\infty \left( \frac{L(y)}{\int_1^y \frac{L(t)}{t} dt} \right)^2 \frac{1}{y} dy < \infty.$$

Indeed if  $p = 1$  the super-level sets satisfy  $\mu(A_{>n} \cap P_j) \sim |L_j(n)|/n$  as  $n \rightarrow \infty$  for all  $P_j \in \tilde{\mathcal{P}}$ , so that by (3.2) assumption (ii) is satisfied.

Maps in  $\mathcal{T}$  have  $L_0(y) \equiv C$ , so that the previous condition is easily satisfied.

**3.2. Fibred systems related to multi-dimensional continued fraction expansions.** The class of systems in dimension greater than one for which the  $\psi$ -mixing property has been studied includes the so-called *fibred systems* on compact and connected subsets  $E$  of  $\mathbb{R}^d$ .

The notion of fibred system was introduced in [S95] in relation to the theory of real number expansions. We refer to [S95] and [NN03] for the detailed definitions.

Given a compact and connected set  $E \subset \mathbb{R}^d$  and a measurable map  $V : E \rightarrow E$ , the main features of a fibred system  $(E, V)$  we need to recall are:

- (i) the existence of a finite or countable measurable partition  $\mathcal{C} = \{C_j\}_{j \in J}$  of  $E$  into open sets such that  $V|_{C_j}$  is injective and differentiable for all  $j \in J$ ;
- (ii) given the iterated partition  $\mathcal{C}^n = \bigvee_{k=0}^{n-1} V^{-k}\mathcal{C}$ , the existence of a sequence  $(\sigma(n))_{n \geq 0}$  with  $\sigma(n) \rightarrow 0$  as  $n \rightarrow \infty$  and such that

$$\sup_{(j_1, \dots, j_n) \in J^n} \text{diam } C_{j_1, \dots, j_n} \leq \sigma(n).$$

for  $C_{j_1, \dots, j_n} \in \mathcal{C}^n$ .

In addition we assume that

- (iii) there exists a finite number of measurable subsets  $U_1, \dots, U_N$  of  $E$  such that for any cylinder  $C_{j_1, \dots, j_n}$  of positive measure, there exists  $U_i$  with  $1 \leq i \leq N$  such that  $V^n(C_{j_1, \dots, j_n}) = U_i$  up to measure-zero sets, and the partition generated by  $U_1, \dots, U_N$  covers  $E$  up to sets of Lebesgue measure zero.

It is proved in [NN03, Thm 2], that a fibred system  $(E, V)$  is  $\psi$ -mixing with respect to the iterated partition  $\mathcal{C}^n$ , and that there exist constants  $K > 0$  and  $\theta \in (0, 1)$  such that

$$(3.3) \quad \psi(n) \leq K \left( \theta^{\sqrt{n}} + \sigma(\sqrt{n}) \right).$$

The estimate in (3.3) is the simplified version of the result in [NN03] valid under assumption (iii).

Examples of fibred systems are obtained in the study of multi-dimensional continued fractions algorithms. Here we consider the two-dimensional system studied in [BDM21], which is related to the two-dimensional continued fractions algorithm introduced in [G01]. Let

$$\Delta := \{(x, y) \in \mathbb{R}^2 : 1 \geq x \geq y \geq 0\}$$

and define the partition  $\{\Gamma_0, \Gamma_1\}$  of  $\Delta$  by

$$\Gamma_0 := \{(x, y) \in \mathbb{R}^2 : 1 \geq x \geq y > 1 - x\},$$

and

$$\Gamma_1 := \{(x, y) \in \mathbb{R}^2 : 1 - y \geq x \geq y \geq 0\}.$$

We consider the map  $S : \Delta \rightarrow \Delta$  defined as

$$S(x, y) := \begin{cases} \left( \frac{y}{x}, \frac{1-x}{x} \right) & \text{if } (x, y) \in \Gamma_0 \\ \left( \frac{x}{1-y}, \frac{y}{1-y} \right) & \text{if } (x, y) \in \Gamma_1 \end{cases}.$$

In [BDM21] it is proved that  $S$  admits an absolutely continuous invariant measure  $\mu$  with density  $h(x, y) = \frac{1}{xy}$ , so that  $\mu(\Delta) = \infty$ , and that the system  $(\Delta, S, \mu)$  is ergodic and conservative.

The next result shows that there exists a set  $E \subset \Delta$  for which assumption (i) of Theorem 2.3 holds. The proof of the proposition is in Section 5.

**Proposition 3.3.** *Let  $E \subset \Delta$  be the triangle with vertices  $Q_1 = (\frac{1}{2}, \frac{1}{2})$ ,  $Q_2 = (\frac{2}{3}, \frac{1}{3})$  and  $Q_3 = (1, 1)$ , with the sides  $Q_1Q_2$  and  $Q_2Q_3$  not included. Then  $\mu(E) < \infty$ , the induced map  $S_E$  on  $E$  is  $\psi$ -mixing with respect to the level sets  $(A_n)$  of the return time function  $\varphi_E$ , and  $\sum_{n \geq 1} \psi(n)/n < \infty$ .*

Finally, assumption (ii) of Theorem 2.3 is indirectly studied in [BDM21, Appendix B] but not proved to hold. Its validity implies the wandering rate  $w_n(E) := \sum_{k=0}^n \mu(A_{>n})$  of  $E$  to be slowly varying as discussed in [AN03], and this property of  $w_n(E)$  is not known. However strong evidence suggests that assumption (ii) of Theorem 2.3 is satisfied by the map  $S$ .

**3.3. Discussions on Theorem 2.7.** As previously remarked, in the case of non-summable observables the pointwise asymptotic behaviour of the Birkhoff sums depends on some properties of the observable itself. We have been able to show that for  $f \geq 0$  it holds  $S_N f(x) \sim G(N)$  for  $\mu$ -a.e.  $x \in X$ , where  $G(n)$  is the asymptotic behaviour of the induced observable  $f^E$  on the level sets  $(A_n)$ , if  $f$  can be bounded by functions  $g_1$  and  $g_2$  with induced functions  $g_i^E$  of the form  $nL_i(n)$  for suitable normalised slowly varying functions (see Appendix B for definitions and details on slowly varying functions). The assumptions on the functions  $L_i(n)$  concern their asymptotic behaviour, as in assumptions (ii)-(b) and (c), and their link with the first return time function to the set  $E$ , as in assumptions (ii)-(d) and (e).

First we discuss assumptions (ii)-(b) and (c), showing sufficient conditions for  $f$  and for a slowly varying function to satisfy them. Then we discuss assumptions (ii)-(d) and (e) in a specific example, and consider the necessity of (ii)-(e). Further we consider the existing literature on this problem.

**Proposition 3.4.** *Let  $f : X \rightarrow \mathbb{R}_{\geq 0}$  be constant on the level sets  $(E_n)$  of the hitting time function (2.15) and assume that the induced function  $f^E$  satisfies  $f^E|_{A_n} = nL(n)$  for all  $n \geq 0$ , with  $L$  slowly varying. Then there exist  $g_1, g_2 : X \rightarrow \mathbb{R}_{\geq 0}$  satisfying (ii)-(a) and (b) of Theorem 2.7.*

*Proof.* Apply Lemma B.3 to  $f^E$  to find  $g_1^E$  and  $g_2^E$ . Then  $g_1$  and  $g_2$  are obtained using relation (2.17) between a function which is constant on the level sets  $(E_n)$  and its induced function.  $\square$

We can thus reduce our attention to the case of a function  $f$  with induced function which is asymptotically equivalent to a function  $G(n) = nL(n)$  with  $L$  normalised slowly varying. Hence using normalised slowly varying functions does not reduce generality.

Let us now consider (ii)-(c). First notice that if  $\xi(n)$  is bounded then assumption (ii)-(c) is easily obtained by the properties of slowly varying functions. Moreover, if  $h(n)$  is a function for which  $\xi(n) \leq h(n)$  and assumption (ii)-(c) is satisfied for  $h$ , then it is immediately satisfied also for  $\xi$ .

Fixed a slowly varying function  $h$ , in Lemma B.4 we give a sufficient condition on slowly varying functions  $L$  to satisfy  $L(n\tilde{h}(n)) \sim L(n)$  for all  $\tilde{h}$  such that  $c \leq \tilde{h}(n) \leq h(n)$ . Notice that in Lemma B.4 we assume that  $h$  is bounded below from 1. If  $h$  does not satisfy this assumption, simply apply the lemma to  $k(n) := \min\{1 + h_0, h(n)\}$  for some  $h_0 > 0$ , which is still slowly varying.

To study assumption (ii)-(c) using Lemma B.4 for the function  $\xi$  it is thus sufficient to show that it is slowly varying or that there exists a slowly varying function  $h$  such that  $\xi(n) \leq h(n)$ . It follows from computations in Proposition 3.6 that  $\xi(n) \sim \log^2(\log n)$ , and therefore is slowly varying, for the class  $\mathcal{T}$  of maps of the interval considered in Section 3.1. For the map studied in Section 3.2 instead, it is shown in [BDM21] that  $\mu(A_{>n}) = O(\frac{\log^2 n}{n})$  and  $\alpha(n) = O(n/\log^2 n)$ , hence in Lemma 2.6 one finds  $q(t) = O(t \log^2 t)$ , and

$$\xi(n) = O\left(\log n \log^2(\log n)\right).$$

Thus we estimate  $\xi$  with a slowly varying function and Lemma B.4 gives a sufficient condition on the slowly varying function  $L$  to satisfy assumption (ii)-(c). Hence if the measure of the super-level sets  $(A_{>n})$  is bounded by a regular varying function with index -1, and the sequence  $\alpha(n)$  defined in (2.11) is bounded by a regular varying function with index 1, then  $\xi$  is bounded by a slowly varying function. In Remark 4.6 we show that it is the case if (2.18) holds.

Finally we discuss the relation between the observable  $f$  and the first return time function to  $E$ . In particular we show that assumptions (ii)-(d) and (e) simplify when the two limiting functions  $g_1$  and  $g_2$  have induced functions which increase asymptotically linearly. The proof of the following proposition is in Section 5.

**Proposition 3.5.** *Let  $T$ ,  $(X, \mathcal{B}, \mu)$  and  $E$  be defined as in Theorem 2.7. Let the sequence of random variables  $(\varphi_E \circ T_E^{n-1})$  with  $n \geq 1$  satisfy assumption (i) of Theorem 2.7. Let  $f : X \rightarrow \mathbb{R}_{\geq 0}$  be a non-negative measurable function such that  $f \notin L^1(X, \mu)$  and let  $f^E$  be its induced version on  $E$  defined in (2.12). We assume that:*

- (c1) *there exist  $g_1, g_2 : X \rightarrow \mathbb{R}_{>0}$  constant on the level sets  $(E_n)$  of the hitting time function to  $E$  such that  $g_1(x) \leq f(x) \leq g_2(x)$  for  $\mu$ -a.e.  $x \in X$ , and the induced functions satisfy  $g_i^E(n) \sim n$  for  $i = 1, 2$ ;*
- (c2) *the super-level sets of  $\varphi_E$  satisfy*

$$\sum_{n \geq 1} \frac{n(\mu(A_{>n}))^2}{(\sum_{j=0}^{n-1} \mu(A_{>j}))^2} < \infty;$$

- (c3) *if  $\ell(n) := \sum_{k=0}^n \mu(A_{\geq k})$ , then  $\ell(N/\ell(N)) \sim \ell(N)$  as  $N \rightarrow \infty$ .*

*Then for  $\mu$ -a.e.  $x \in X$  it holds*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N (f \circ T^{k-1})(x) = 1.$$

**Corollary 3.6.** *Let us consider a map of the interval  $T \in \mathcal{T}$  with only one indifferent fixed point at 0 satisfying (3.1) with  $p = 1$  and  $C = 1$ , and let  $\mu$  be the infinite absolutely continuous  $T$ -invariant measure. If  $f : X \rightarrow \mathbb{R}_{\geq 0}$  is constant on the level sets  $(E_n)$  of the hitting time function (2.15) and the induced function  $f^E$  satisfies  $f^E|_{A_n} \sim n$  as  $n \rightarrow \infty$  then*

$$(3.4) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N (f \circ T^{n-1})(x) = 1, \quad \text{for } \mu\text{-a.e. } x \in [0, 1].$$

*Proof.* This is a special case of Proposition 3.5. Assumption (i) of Theorem 2.7 holds as discussed in Section 3.1, and (c1) follows by definition of the observable  $f$ . Finally for  $T \in \mathcal{T}$  with  $p = 1$  and  $C = 1$  it holds  $\mu(A_{>n}) \sim \frac{1}{n}$ . Thus  $\ell(n) \sim \log n$  and (c2)-(c3) hold true.  $\square$

In order to not only considering the trivial case of  $G(n) \sim n$ , we give a more involved example generalising Corollary 3.6.

*Example 3.7.* Let  $(X, \mathcal{B}, \mu)$  and  $E \in \mathcal{B}$  be a system fulfilling (i) of Theorem 2.7 with  $\mu(A_{>n}) \sim \log^b(\log n)/n$  as  $n \rightarrow \infty$  for  $b \in \mathbb{R}$ . Furthermore we assume  $f : X \rightarrow \mathbb{R}_{\geq 0}$  to be constant on the level sets  $(E_n)$  of the hitting time function (2.15) and let the induced function  $f^E$  satisfy  $f^E|_{A_n} \sim n \log^c(\log n)$  as  $n \rightarrow \infty$  with  $c \in \mathbb{R}$ , then

$$(3.5) \quad \lim_{N \rightarrow \infty} \frac{1}{N \log^c(\log N)} \sum_{n=1}^N (f \circ T^{n-1})(x) = 1, \quad \text{for } \mu\text{-a.e. } x \in X.$$

We will use Theorem 2.7 to verify this statement. Obviously, conditions (i), (ii)-(a), and (ii)-(b) are already fulfilled by assumption. To verify condition (ii)-(c) we notice that  $q(n) \sim n \log^b(\log n)$  (using [BGT87, Cor. 2.3.4]). Furthermore  $\alpha(n) \sim n/(\log n \log^b(\log n))$  implies  $\xi(n) \sim n \log^2(\log n)$  and  $\log^c(\log(n \xi(n))) \sim \log^c(\log n)$  gives condition (ii)-(c).

To verify condition (ii)-(d) we notice that  $\Gamma(n) \sim n/\log^c(\log n)$  (also by applying [BGT87, Cor. 2.3.4]) and so  $\mu(A_{>\Gamma(n)}) \sim \log^{b+c}(\log n)/n$ . Thus

$$\int_0^\infty \frac{y (\mu(A_{\geq \Gamma(y)}))^2}{\left(\int_0^y \mu(A_{\geq \Gamma(t)}) dt\right)^2} dy \sim \int_0^\infty \frac{\log^{2(b+c)}(\log y)}{y \left(\int_0^y \log^{b+c}(\log t)/t dt\right)^2} dy \sim \int_0^\infty \frac{1}{y \log^2 y} dy < \infty.$$

Finally, we are left to verify condition (ii)-(e). We have  $\ell(n) \sim \log n \log^{b+c}(\log n)$  which is super-slowly varying with rate function itself and  $\ell(\alpha(n)) \sim \ell(n)$ . On the other hand

$$L(n) \sum_{k=1}^n \mu(A_{>k}) \sim \left(\log^c(\log n)\right) \left(\log n \log^b(\log n)\right) \sim \ell(n)$$

giving all properties of Theorem 2.7 and (3.5) holds.

In the last examples we have seen that not only the trivial case  $G(n) \sim n$  and  $\mu(A_{>n}) \sim \frac{1}{n}$  is applicable on Theorem 2.7. However, as we will see in the next proposition, condition (ii)-(e) can be quite restrictive and, as we will see, we will not be able to drop it.

**Proposition 3.8.** *Let  $T$ ,  $(X, \mathcal{B}, \mu)$  and  $E$  be defined as in Theorem 2.7. Let the sequence of random variables  $(\varphi_E \circ T_E^{n-1})$  with  $n \geq 1$  satisfy assumption (i) of Theorem 2.7. Let  $f : X \rightarrow \mathbb{R}_{\geq 0}$  be a non-negative measurable function such that  $f \notin L^1(X, \mu)$  and let  $f^E$  be its induced version on  $E$  defined in (2.12). We assume that:*

- (w1) *there exist  $g_1, g_2 : X \rightarrow \mathbb{R}_{\geq 0}$  constant on the level sets  $(E_n)$  of the hitting time function to  $E$  such that  $g_1(x) \leq f(x) \leq g_2(x)$  for  $\mu$ -a.e.  $x \in X$ , and the induced functions satisfy  $g_i^E(n) \sim n \log n$  for  $i = 1, 2$ ;*
- (w2) *the super-level sets of  $\varphi_E$  satisfy  $\mu(A_{>n}) \sim \frac{1}{n}$ .*

Then there exist  $u_1 < u_2$  such that for  $\mu$ -a.e.  $x \in X$  it holds

$$\limsup_{N \rightarrow \infty} \frac{1}{N \log N} \sum_{k=1}^N (f \circ T^{k-1})(x) > u_2 \quad \text{and} \quad \liminf_{N \rightarrow \infty} \frac{1}{N \log N} \sum_{k=1}^N (f \circ T^{k-1})(x) < u_1.$$

We will first assure ourselves that Proposition 3.8 fulfils all assumptions of Theorem 2.7 with  $G(n) \sim n \log n$  except for (ii)-(e). It is clear that assumptions (ii)-(a) and (b) are satisfied by (w1). To verify (ii)-(c), we note that (w2) implies  $\alpha(n) \sim n/\log n$  from which we obtain  $\xi(n) \sim \log^2(\log n)$  and  $\log(n \xi(n)) \sim \log n$  verifying (ii)-(c). To verify (ii)-(d) we first notice that  $\Gamma(n) \sim n/\log n$  is an asymptotic inverse of  $G(n)$  which we either get by direct calculation or by using e.g. [BGL18, Cor. 2.3.4]. This implies that  $\mu(A_{>\Gamma(n)}) \sim \log n/n$  and by an easy calculation we obtain that (2.18) is fulfilled. The proof of Proposition 3.8 is in Section 5.

Finally, we discuss the existing literature. The pointwise convergence of Birkhoff averages for non-summable observables and infinite measure-preserving dynamical systems has been recently discussed in [LM18]. The authors of [LM18] have proved that (3.4) holds under strong assumptions on the observable  $f$  and very weak assumptions on the system. Basically they considered  $L^\infty$  observables for which it is possible to control in a very strong way the contributions to the Birkhoff sums from the excursions out of a finite measure set  $E$ , the set on which we have considered the induced map. In this paper we have less stringent assumptions on the observables but have considered more particular dynamical systems, those for which an induced map exists which is  $\psi$ -mixing. We finish this section by discussing the relations of our results with those in [LM18] for the Farey map.

*Remark 3.9.* Let us consider the particular case of the Farey map  $F : [0, 1] \rightarrow [0, 1]$  which belongs to the class  $\mathcal{T}$  considered in Section 3.1 and is defined as

$$F(x) = \begin{cases} \frac{x}{1-x}, & \text{if } x \in [0, \frac{1}{2}], \\ \frac{1-x}{x}, & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

It is well known that the absolutely continuous invariant measure is  $d\mu(x) = \frac{1}{x \log 2} dx$ . If  $E = (\frac{1}{2}, 1)$ , for which  $\mu(E) = 1$ , it holds  $E_n = (\frac{1}{n+1}, \frac{1}{n})$ . Thus  $\mu(E_n) = \mu(A_{>n}) = \log_2(1 + \frac{1}{n}) \sim \frac{1}{n \log 2}$ .

A particular class of observables studied in the context of infinite mixing theory is that of *global observables*. For maps in  $\mathcal{T}$ , we recall that a global observable is a function  $f \in L^\infty([0, 1], \mu)$  such that the limit

$$\lim_{a \rightarrow 0^+} \frac{1}{\mu(a, 1)} \int_a^1 f d\mu$$

exists and is finite (see [BGL18, BL21] for more details). We first show that observables  $f \in L^\infty(X, \mu)$  which are constant on the level sets  $(E_n)$  with  $f_n := f|_{E_n}$  and for which the induced function satisfies  $f^E|_{A_n} \sim n$  as  $n \rightarrow \infty$ , are global observables. The definition of global observable for the Farey map and for observables constant on the sets  $E_n$  reduces to show that the limit

$$(3.6) \quad \lim_{n \rightarrow \infty} \frac{1}{\mu\left(\frac{1}{k+1}, 1\right)} \int_{\frac{1}{k+1}}^1 f(x) d\mu = \lim_{n \rightarrow \infty} \frac{1}{\sum_{k=1}^n \log_2(1 + \frac{1}{k})} \sum_{k=1}^n f_k \log_2\left(1 + \frac{1}{k}\right)$$

exists and is finite. To prove this, it is enough to recall that  $\sum_{k=1}^n \log_2(1 + \frac{1}{k}) \sim \log_2 n$  and to apply Abel's summation formula to get

$$\sum_{k=1}^n f_k \log_2\left(1 + \frac{1}{k}\right) \sim \left(\sum_{k=1}^n f_k\right) \log_2\left(1 + \frac{1}{n}\right) + \frac{1}{\log 2} \int_1^n \left(\sum_{k=1}^t f_k\right) \frac{1}{t^2 + t} dt \sim$$

$$\sim f^E|_{A_n} \log_2 \left(1 + \frac{1}{n}\right) + \int_1^n \frac{f^E|_{A_t}}{\log 2} \frac{1}{t^2 + t} dt \sim n \log_2 \left(1 + \frac{1}{n}\right) + \frac{1}{\log 2} \int_1^n \frac{t}{t^2 + t} dt \sim \log_2(n+1).$$

On the other hand let's consider a global observable  $f : X \rightarrow \mathbb{R}_{\geq 0}$  which is constant on the level sets  $(E_n)$  with  $f_n := f|_{E_n}$ . By rescaling we can assume that the limit in (3.6) exists and is equal to 1. In general it is not possible to conclude about the asymptotic behaviour of  $f^E|_{A_n}$ , however some sufficient conditions can be obtained by the following argument (see [M06]). With  $\ell_n := \sum_{k=1}^n \log_2(1 + \frac{1}{k})$  let

$$\tau_n := \frac{1}{\ell_n} \sum_{k=1}^n f_k \log_2 \left(1 + \frac{1}{k}\right).$$

Then  $f_1 = \ell_1 \tau_1$  and

$$f_n = \frac{1}{\log_2 \left(1 + \frac{1}{n}\right)} (\ell_n \tau_n - \ell_{n-1} \tau_{n-1}), \quad \text{for } n \geq 2,$$

from which

$$f^E|_{A_n} = \sum_{k=1}^n f_k = \frac{\ell_n \tau_n}{\log_2 \left(1 + \frac{1}{n}\right)} - \sum_{k=1}^{n-1} \ell_k \tau_k \left( \frac{1}{\log_2 \left(1 + \frac{1}{k+1}\right)} - \frac{1}{\log_2 \left(1 + \frac{1}{k}\right)} \right).$$

Since  $\ell_n \sim \log_2 n$ , it follows for example that if  $\tau_n \sim 1 + o((\log_2 n)^{-1})$  then  $f^E|_{A_n} \sim n$ . Hence we can check if Proposition 3.6 is applicable to a given global observable.

Let us now compare the conditions on  $f$  in Proposition 3.6 with the statement in [LM18, Theorem 2.5]. In our setting and with proper norming, the assumptions in [LM18, Theorem 2.5] can be stated as follows:  $f \in L^\infty(X, \mu)$  and for all  $\epsilon > 0$  there exist  $N, K \in \mathbb{N}$  such that for all  $x \in \bigcup_{k \geq K} E_k$  we have that

$$(3.7) \quad \left| \frac{\sum_{j=1}^N (f \circ T^{j-1})(x)}{N} - 1 \right| < \epsilon.$$

First we show that if the assumptions for an observable  $f$  in [LM18, Theorem 2.5] are satisfied, then we obtain a good asymptotic behaviour for the induced observable  $f^E$ , so that we are basically in a good situation to apply Theorem 2.7.

**Lemma 3.10.** *Let  $f : [0, 1] \rightarrow \mathbb{R}_{\geq 0}$  be in  $L^\infty(X, \mu)$  and such that for all  $\epsilon > 0$  there exist  $N, K \in \mathbb{N}$  such that for all  $x \in \bigcup_{k \geq K} E_k$  we have that (3.7) holds. Then  $f^E|_{A_k} \sim k$  for  $k \rightarrow \infty$ .*

*Proof.* Let  $\epsilon > 0$  be given and consider  $N, K$  fixed as in (3.7). We set  $f_N(x) := \sum_{j=1}^N (f \circ T^{j-1})(x)$ . For  $x \in A_R$  with  $R > K$ , writing  $R = K + rN + s$  with  $r \in \mathbb{N}_0$  and  $0 \leq s \leq N - 1$ , we get

$$f^E(x) = \sum_{j=1}^{R+1} (f \circ T^{j-1})(x) = \sum_{i=1}^r f_N(T^{N(i-1)}(x)) + \sum_{j=rN+1}^{R+1} (f \circ T^{j-1})(x).$$

Since  $T^{N(i-1)}(x) \in \bigcup_{k \geq K} E_k$  for all  $i = 1, \dots, r$ , we apply (3.7), and the fact that  $f \in L^\infty(X, \mu)$  and  $f \geq 0$ , to obtain

$$rN(1 - \epsilon) \leq f^E(x) \leq rN(1 + \epsilon) + (N + K) \|f\|_\infty, \quad \forall x \in A_R.$$

Since  $R \sim rN \rightarrow \infty$  as  $R \rightarrow \infty$ , the lemma is proved.  $\square$

Finally we give two examples to which Theorem 2.7 is applicable, but which fail to fulfil the conditions in [LM18, Theorem 2.5]. The first is an example with  $f \notin L^\infty(X, \mu)$ , the second one with  $f \in L^\infty(X, \mu)$  but not satisfying (3.7).

*Example 3.11.* Let us consider a strictly monotonic sequence of natural numbers  $\gamma_n$  defined by  $\gamma_1 = 4$  and  $\gamma_{n+1} = \gamma_n + \left\lfloor \gamma_n^{1/2} \right\rfloor$ . Then we set  $f$  to be constant on the intervals  $E_k$  with

$$f(x) = \begin{cases} \gamma_n - \gamma_{n-1}, & \text{if } x \in E_{\gamma_n} \text{ for some } n \in \mathbb{N}; \\ 0, & \text{else.} \end{cases}$$

Since  $\gamma_n - \gamma_{n-1}$  tends to infinity,  $f \notin L^\infty(X, \mu)$ . On the other hand the induced function satisfies  $f^E|_{A_k} = \gamma_n$  for  $\gamma_n \leq k < \gamma_{n+1}$ , so that for all  $k \in \mathbb{N}$  and all  $x \in A_k$  we have  $k + 1 - \sqrt{k+1} \leq f^E(x) \leq k$ , implying that  $f^E|_{A_k} \sim k$  as  $k \rightarrow \infty$ .

*Example 3.12.* First we define a strictly monotonic sequence of even numbers  $\kappa_n$  by setting  $\kappa_1 = 4$  and  $\kappa_{n+1} = \kappa_n + 2 \left\lfloor \kappa_n^{1/2} / 2 \right\rfloor$ . Furthermore, we set

$$U := \{\kappa_n, \kappa_n + 1, \dots, \kappa_n + 1/2(\kappa_{n+1} - \kappa_n), n \in \mathbb{N}\}.$$

We define

$$f(x) = \begin{cases} 2, & \text{if } x \in E_k \text{ and } k \in U; \\ 0, & \text{else.} \end{cases}$$

Then  $0 \leq f(x) \leq 2$  and thus  $f \in L^\infty(X, \mu)$ . Furthermore, if  $x \in A_k$  then  $k - 2\sqrt{k} \leq f^E(x) \leq k + 2\sqrt{k}$  implying that  $f^E|_{A_k} \sim k$  as  $k \rightarrow \infty$ .

On the other hand, for each  $K \in \mathbb{N}$  there is a positive measure set  $A \subset \cup_{k \geq K} E_k$  such that for all  $x \in A$  we have that  $f(x) = \dots = (f \circ T^K)(x) = 0$  implying that (3.7) cannot be fulfilled.

#### 4. PROOFS OF THE MAIN RESULTS

**4.1. Proof of Theorem 2.3.** We prove the theorem for  $\mu$ -a.e.  $x \in E$ . The result for  $\mu$ -a.e.  $x \in X$  follows since  $T$  is conservative and ergodic. Let  $f \in L^1(X, \mu)$  and let  $R_{E, N, m}(x) := R_{E, N+m(N, E, x)}(x)$  denote the number of visits to  $E$  in the first  $N + m(N, E, x)$  steps of the orbit of  $x$ . Thus for each  $N$  we add a number of steps depending on  $x$  and being non-decreasing in  $N$ .

**Lemma 4.1.** *For  $\mu$ -a.e.  $x \in E$  it holds*

$$\lim_{N \rightarrow \infty} \frac{1}{R_{E, N, m}(x)} \sum_{k=1}^{N+m(N, E, x)} (f \circ T^{k-1})(x) = \int_X f \, d\mu.$$

*Proof.* We apply Hopf's Ratio Ergodic Theorem to  $f(x)$  and  $\mathbb{1}_E(x)$  and obtain for  $\mu$ -a.e.  $x \in E$

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{\sum_{k=1}^{N+m(N, E, x)} (f \circ T^{k-1})(x)}{R_{E, N, m}(x)} = \\ & = \lim_{N \rightarrow \infty} \frac{\sum_{k=1}^{N+m(N, E, x)} (f \circ T^{k-1})(x)}{(\mathbb{1}_E \circ T^{N+m(N, E, x)})(x) + \sum_{k=1}^{N+m(N, E, x)} (\mathbb{1}_E \circ T^{k-1})(x)} = \\ & = \frac{\int_X f \, d\mu}{\int_X \mathbb{1}_E \, d\mu} = \int_X f \, d\mu \end{aligned}$$

where we have used that  $\sum_{k=1}^{N+m(N, E, x)} (\mathbb{1}_E \circ T^{k-1})(x)$  is divergent. □

Let us now recall for  $\mu$ -a.e.  $x \in E$  the definition of the time  $\tau_{E,x}(N)$  of the  $N$ -th return to  $E$  given in (2.13). Then as in (2.8), we consider its trimmed version

$$(4.1) \quad \tau_{E,x}^1(N) := \tau_{E,x}(N) - \max_{1 \leq k \leq N} (\varphi_E \circ T_E^{k-1})(x)$$

where we are trimming  $\tau_{E,x}$  by deleting the largest entry.

**Lemma 4.2.** *For  $\mu$ -a.e.  $x \in E$  it eventually holds*

$$N < \tau_{E,x}^1(R_{E,N,m}(x)) \leq N + M_{N,m}^2(x)$$

where  $M_{N,m}^2(x)$  denotes the 2-nd maximum in  $\{(\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N,m}(x)\}$  which in case of equality might coincide with the maximum.

*Proof.* We first give some remarks on  $\tau_{E,x}$  and  $R_{E,N}(x)$ . Given  $N \geq 1$ ,  $R_{E,N}(x)$  is the number of visits to  $E$  up to time  $N$ , hence  $R_{E,N}(x) \leq N + 1$  and  $T_E^{R_{E,N}(x)-1}(x)$  is the last visit to  $E$  for  $x$  up to time  $N$ . It follows that

$$\tau_{E,x}(R_{E,N}(x)) = \sum_{k=1}^{R_{E,N}(x)} (\varphi_E \circ T_E^{k-1})(x) > N$$

as the  $(R_{E,N}(x))$ -th return to  $E$ , equivalently the  $(R_{E,N}(x) + 1)$ -th visit to  $E$ , happens after time  $N$ . Moreover

$$\sum_{k=1}^{R_{E,N}(x)} (\varphi_E \circ T_E^{k-1})(x) = \varphi_E \circ T_E^{R_{E,N}(x)-1}(x) + \tau_{E,x}(R_{E,N}(x) - 1) \leq \varphi_E \circ T_E^{R_{E,N}(x)-1}(x) + N$$

as the  $(R_{E,N}(x) - 1)$ -th return to  $E$ , equivalently the  $(R_{E,N}(x))$ -th visit to  $E$ , happens at time less than or equal to  $N$ .

Using now  $R_{E,N,m}(x)$ , we have proved

$$(4.2) \quad N + m(N, E, x) < \tau_{E,x}(R_{E,N,m}(x)) \leq N + m(N, E, x) + \varphi_E \circ T_E^{R_{E,N,m}(x)-1}(x).$$

We now show that

$$(4.3) \quad \max \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N,m}(x) \right\} = \max \left\{ m(N, E, x), \varphi_E \circ T_E^{R_{E,N,m}(x)-1}(x) \right\}.$$

From (2.6),  $m(N, E, x)$  maximises the return times  $(\varphi_E \circ T_E^{k-1})(x)$  for  $k = 1, \dots, R_{E,N}(x)$ , that is up to time  $\tau_{E,x}(R_{E,N}(x))$ . Moreover, repeating the same argument as above, the return times  $(\varphi_E \circ T_E^{k-1})(x)$  with  $k = R_{E,N}(x) + 1, \dots, R_{E,N,m}(x) - 1$  are concerned with excursions outside  $E$  for the orbit of  $x$  with time in  $(N, N + m(N, E, x)]$ . Then

$$m(N, E, x) = \max \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N}(x) - 1 \right\}$$

and (4.3) is proved.

Then, if

$$\max \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N,m}(x) \right\} = m(N, E, x)$$

it immediately follows from (4.2) that

$$N < \tau_{E,x}(R_{E,N,m}(x)) - \max_{1 \leq k \leq R_{E,N,m}(x)} (\varphi_E \circ T_E^{k-1})(x) = \tau_{E,x}^1(R_{E,N,m}(x)),$$

otherwise, we obtain

$$\tau_{E,x}^1(R_{E,N,m}(x)) = \tau_{E,x}(R_{E,N,m}(x) - 1) > N$$



since  $x$  necessarily visits  $E$  at some time  $k$  in  $(N, N + m(N, E, x)]$ . In both cases we have thus obtained the first inequality to prove.

Moreover from (4.2) and (4.3) we obtain

$$\tau_{E,x}(R_{E,N,m}(x)) - \max_{1 \leq k \leq R_{E,N,m}(x)} (\varphi_E \circ T_E^{k-1})(x) \leq N + \min \left\{ m(N, E, x), \varphi_E \circ T_E^{R_{E,N,m}(x)-1}(x) \right\}$$

and by (4.3)

$$\min \left\{ m(N, E, x), \varphi_E \circ T_E^{R_{E,N,m}(x)-1}(x) \right\} \leq M_{N,m}^2(x).$$

The proof is finished.  $\square$

**Lemma 4.3.** *For  $\mu$ -a.e.  $x \in E$  we have*

$$\lim_{N \rightarrow \infty} \frac{\tau_{E,x}^1(N)}{d(N)} = 1$$

where  $d(n)$  is the asymptotic inverse function of the sequence  $(\alpha(n))$  defined in (2.11).

*Proof.* By their definition (4.1), the terms  $\tau_{E,x}^1(N)$  are the trimmed Birkhoff sums of  $\varphi_E$  for the system  $(E, T_E)$ . Thus, it is enough to show that Lemma 2.2 can be applied to the sequence of random variables  $(\varphi_E \circ T_E^{n-1})$ .

By assumption (i), the sequence is  $\psi$ -mixing with mixing coefficient  $\psi(n)$  fulfilling  $\sum_{n \geq 1} \psi(n)/n < \infty$ . Moreover, the distribution function of  $(\varphi_E \circ T_E^{n-1})$  is given by

$$F(y) = \mu(\varphi_E \leq y) = \sum_{k=1}^{\lfloor y \rfloor} \mu(A_k) = 1 - \mu(A_{> \lfloor y \rfloor}),$$

by (2.1) and (2.2). Thus, in (2.9) with  $y_0 = 1$  we find

$$\begin{aligned} \int_1^\infty \left( \frac{y(1-F(y))}{\int_0^y (1-F(t)) dt} \right)^{r+1} \frac{1}{y} dy &= \sum_{n \geq 1} \int_n^{n+1} \left( \frac{y \mu(A_{>n})}{\int_0^y \mu(A_{>[t]}) dt} \right)^{r+1} \frac{1}{y} dy \leq \\ &\leq \sum_{n \geq 1} \frac{(n+1)^r (\mu(A_{>n}))^{r+1}}{(\int_0^n \mu(A_{>[t]}) dt)^{r+1}} = \sum_{n \geq 1} \frac{(n+1)^r (\mu(A_{>n}))^{r+1}}{(\sum_{j=0}^{n-1} \mu(A_{>j}))^{r+1}} \end{aligned}$$

and by assumption (ii) we find  $W \leq 1$ . Analogously,

$$\int_1^\infty \left( \frac{y(1-F(y))}{\int_0^y (1-F(t)) dt} \right)^{r+1} \frac{1}{y} dy \geq \sum_{n \geq 1} \frac{n^r (\mu(A_{>n+1}))^{r+1}}{(\sum_{j=0}^n \mu(A_{>j}))^{r+1}}$$

and (2.3) imply  $W > 0$ , so that  $W = 1$ . Thus, the proof is finished by applying Lemma 2.2 and recalling that the Birkhoff sums of  $(\varphi_E \circ T_E^{n-1})$  trimmed by the largest entry are  $\tau_{E,x}^1(N)$ .  $\square$

**Lemma 4.4.** *For  $\mu$ -a.e.  $x \in E$  it holds*

$$\lim_{N \rightarrow \infty} \frac{R_{E,N,m}(x)}{\alpha(N)} = 1.$$

*Proof.* Since  $R_{E,N,m}(x)$  is diverging, we obtain from Lemma 4.3 that

$$\lim_{N \rightarrow \infty} \frac{\tau_{E,x}^1(R_{E,N,m}(x))}{d(R_{E,N,m}(x))} = 1$$

for  $\mu$ -a.e.  $x \in E$ . Moreover, from Lemma 4.2 we obtain

$$\frac{N}{d(R_{E,N,m}(x))} \leq \frac{\tau_{E,x}^1(R_{E,N,m}(x))}{d(R_{E,N,m}(x))} \leq \frac{N + M_{N,m}^2(x)}{d(R_{E,N,m}(x))}$$

where we recall that  $M_{N,m}^2(x)$  denotes the 2-nd maximum in  $\{(\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N,m}(x)\}$ . Thus, by applying Lemma 2.2 to  $(\varphi_E \circ T_E^{n-1})$  with  $W = 1$  as shown above, we use (2.10) with  $r = 2$  to obtain

$$\lim_{N \rightarrow \infty} \frac{N}{d(R_{E,N,m}(x))} = \lim_{N \rightarrow \infty} \frac{N + M_{N,m}^2(x)}{d(R_{E,N,m}(x))} = 1$$

for  $\mu$ -a.e.  $x \in E$ . The result now follows by using that  $\alpha(n)$  is the asymptotic inverse of  $d(n)$ .  $\square$

Theorem 2.3 follows by applying Lemma 4.1 and Lemma 4.4.  $\square$

**4.2. Proof of Theorem 2.4.** As before we prove the theorem for  $\mu$ -a.e.  $x \in E$ . The first step is the analog of Lemma 4.1. Using Hopf's Ratio Ergodic Theorem we have for  $\mu$ -a.e.  $x \in E$

$$(4.4) \quad \lim_{N \rightarrow \infty} \frac{\sum_{k=1}^N (f \circ T^{k-1})(x)}{R_{E,N}(x)} = \lim_{N \rightarrow \infty} \frac{\sum_{k=1}^N (f \circ T^{k-1})(x)}{(\mathbb{1}_E \circ T^N)(x) + \sum_{k=1}^N (\mathbb{1}_E \circ T^{k-1})(x)} = \int_X f \, d\mu.$$

Then we consider the sequence  $\tau_{E,x}$  defined in (2.13) and the trimmed version  $\tau_{E,x}^1$ . In Lemma 4.2 we have proved

$$(4.5) \quad N < \tau_{E,x}(R_{E,N}(x)) \leq N + \varphi_E \circ T_E^{R_{E,N}(x)-1}(x)$$

for  $\mu$ -a.e.  $x \in E$ . For the trimmed sums

$$\tau_{E,x}^1(R_{E,N}(x)) = \tau_{E,x}(R_{E,N}(x)) - \max_{1 \leq k \leq R_{E,N}(x)} (\varphi_E \circ T_E^{k-1})(x)$$

we now show that

$$(4.6) \quad N - w(N, E, x) \leq \tau_{E,x}^1(R_{E,N}(x)) \leq N - w(N, E, x) + M_N^2(x)$$

where  $w(N, E, x)$  is defined in (2.7) and  $M_N^2(x)$  denotes the 2-nd maximum in  $\{(\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N}(x)\}$ .

We remark that analogously to (2.6) we have for  $\mu$ -a.e.  $x \in E$

$$(4.7) \quad w(N, E, x) = \max \left\{ \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N}(x) - 1 \right\}, N - \tau_{E,x}(R_{E,N}(x) - 1) \right\}$$

since  $\tau_{E,x}(R_{E,N}(x) - 1)$  is the time of the  $(R_{E,N}(x) - 1)$ -th return to  $E$ , the last before step  $N + 1$ . We thus consider two cases. If

$$\max \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N}(x) \right\} = \max \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N}(x) - 1 \right\},$$

by (4.7) it follows

$$w(N, E, x) = \max \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N}(x) - 1 \right\}$$

since  $\varphi_E \circ T_E^{R_{E,N}(x)-1}(x) > N - \tau_{E,x}(R_{E,N}(x) - 1)$ . So (4.6) follows from (4.5) by subtracting  $w(N, E, x)$  from all terms and using  $\varphi_E \circ T_E^{R_{E,N}(x)-1}(x) \leq M_N^2(x)$ . If on the contrary

$$\varphi_E \circ T_E^{R_{E,N}(x)-1}(x) = \max \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N}(x) \right\},$$

then

$$\tau_{E,x}^1(R_{E,N}(x)) = \tau_{E,x}(R_{E,N}(x)) - \varphi_E \circ T_E^{R_{E,N}(x)-1}(x) = \tau_{E,x}(R_{E,N}(x) - 1).$$

Now from (4.7) we may have

$$w(N, E, x) = N - \tau_{E,x}(R_{E,N}(x) - 1) = N - \tau_{E,x}^1(R_{E,N}(x))$$

in which case (4.6) is immediate. Otherwise we may have

$$w(N, E, x) = \max \left\{ (\varphi_E \circ T_E^{k-1})(x) : k = 1, \dots, R_{E,N}(x) - 1 \right\}$$

in which case

$$N - \tau_{E,x}^1(R_{E,N}(x)) \leq w(N, E, x) = M_N^2(x)$$

so that

$$N - w(N, E, x) \leq \tau_{E,x}^1(R_{E,N}(x)) \leq N = N - w(N, E, x) + M_N^2(x)$$

where we have used (4.5) in the second inequality, and again we have proved (4.6).

Finally, we use Lemma 4.3 to write as in the proof of Lemma 4.4

$$\lim_{N \rightarrow \infty} \frac{\tau_{E,x}^1(R_{E,N}(x))}{d(R_{E,N}(x))} = 1$$

for  $\mu$ -a.e.  $x \in E$ . Moreover, from (4.6) we obtain

$$\frac{N - w(N, E, x)}{d(R_{E,N}(x))} \leq \frac{\tau_{E,x}^1(R_{E,N}(x))}{d(R_{E,N}(x))} \leq \frac{N - w(N, E, x) + M_N^2(x)}{d(R_{E,N}(x))}.$$

Thus, by applying Lemma 2.2 to  $(\varphi_E \circ T_E^{n-1})$  with  $W = 1$  as above, we use (2.10) with  $r = 2$  to obtain

$$\lim_{N \rightarrow \infty} \frac{N - w(N, E, x)}{d(R_{E,N}(x))} = 1$$

for  $\mu$ -a.e.  $x \in E$ . Using now that  $\alpha(n)$  is the asymptotic inverse of  $d(n)$ , we obtain

$$(4.8) \quad \lim_{N \rightarrow \infty} \frac{R_{E,N}(x)}{\alpha(N - w(N, E, x))} = 1$$

for  $\mu$ -a.e.  $x \in E$ . Theorem 2.4 follows from (4.4) and (4.8). □

**4.3. Proof of Lemma 2.6.** We first prove that

$$(4.9) \quad \mu \left( \left\{ x \in E : \max_{1 \leq i \leq n} (\varphi_E \circ T_E^{i-1})(x) > q(n \log n \log^2(\log n)) \text{ for infinite } n \right\} \right) = 0.$$

We note that for all  $2^n \leq k < 2^{n+1}$

$$(4.10) \quad \left\{ x \in E : \max_{1 \leq i \leq k} (\varphi_E \circ T_E^{i-1})(x) > q(k \log k \log^2(\log k)) \right\} \subset \\ \subset \left\{ x \in E : \max_{1 \leq i \leq 2^{n+1}} (\varphi_E \circ T_E^{i-1})(x) > q(2^n \log 2^n \log^2(\log 2^n)) \right\}.$$

Furthermore, we have by the definition of  $q$  that

$$\begin{aligned}
(4.11) \quad & \mu \left( \left\{ x \in E : \max_{1 \leq i \leq 2^{n+1}} (\varphi_E \circ T_E^{i-1})(x) > q (2^n \log 2^n \log^2(\log 2^n)) \right\} \right) \leq \\
& \leq \sum_{j=1}^{2^{n+1}} \mu \left( \left\{ x \in E : (\varphi_E \circ T_E^{j-1})(x) > q (2^n \log 2^n \log^2(\log 2^n)) \right\} \right) = \\
& = \sum_{j=1}^{2^{n+1}} \frac{1}{2^n \log 2^n \log^2(\log 2^n)} = \frac{2}{\log 2^n \log^2(\log 2^n)}.
\end{aligned}$$

Since

$$\sum_{n=1}^{\infty} \frac{2}{\log 2^n \log^2(\log 2^n)} < \infty,$$

we obtain by the first Borel-Cantelli lemma and a combination of (4.10) and (4.11) that (4.9) holds. From (4.8) we obtain that for  $\mu$ -a.e.  $x \in E$

$$\limsup_{n \rightarrow \infty} \frac{R_{E,n}(x)}{\alpha(n)} \leq 1.$$

Hence,

$$\mu \left( \left\{ x \in E : \max_{1 \leq i \leq R_{E,n}(x)} (\varphi_E \circ T_E^{i-1})(x) > q (\alpha(n) \log(\alpha(n)) \log^2(\log(\alpha(n)))) \text{ for infinite } n \right\} \right) = 0$$

and with the definition of  $\xi$  the statement of the lemma follows.  $\square$

**4.4. Proof of Theorem 2.7.** For simplicity let us first assume that  $f$  is constant on the level sets  $E_n$  and we can take  $g_1 = g_2 = f \geq 0$  in the assumptions. Hence there exists a non-decreasing  $G : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$  with  $f^E(x) = G(n)$  for all  $x \in A_n$ , such that  $G(n) = n L(n)$  with  $L$  a normalised slowly varying function and that satisfies (ii)-(c), (d) and (e) with  $\Gamma_i(n) = \Gamma(n) := \min\{k \in \mathbb{N} : G(k) > n\}$ . In this case the sequence of random variables  $(f^E \circ T_E^{n-1})$  is  $\psi$ -mixing since it generates the same  $\sigma$ -field as the sequence  $(\varphi_E \circ T_E^{n-1})$ . Thus it satisfies the first assumption of Lemma 2.2.

Moreover the distribution function of the random variables  $(f^E \circ T_E^{n-1})$  is given by

$$F(y) = \mu(f^E \leq y) = 1 - \mu(A_{\geq \Gamma(y)})$$

so that (2.18) implies  $W \leq 1$  in (2.9). Setting  $G(0) = 0$  we have by definition  $\Gamma(y) = k$  for all  $y \in [G(k-1), G(k))$  for all  $k \geq 1$ . Hence

$$\int_0^{\infty} \mu(A_{\geq \Gamma(y)}) dy = \sum_{k=1}^{\infty} (G(k) - G(k-1)) \mu(A_{> k-1}) = \sum_{k=1}^{\infty} f_{k-1} \mu(A_{> k-1}) = \int_0^{\infty} f(x) d\mu = \infty$$

where we have used (2.17) with  $f_n = f|_{E_n}$  for all  $n \geq 0$ , with  $E_n$  the level sets (2.16) of the hitting time function to  $E$ . It follows that  $W > 0$ , so  $W = 1$ .

Applying Lemma 2.2 to  $(f^E \circ T_E^{n-1})$  it follows that letting  $d(n)$  be the inverse function of

$$(4.12) \quad a(y) := \frac{y}{\int_0^y \mu(A_{\geq \Gamma(t)}) dt}$$

we have for  $N \rightarrow \infty$

$$(4.13) \quad \sum_{n=1}^{R_{E,N}(x)-1} (f^E \circ T_E^{n-1})(x) \sim d(R_{E,N}(x) - 1) + \max_{1 \leq k \leq R_{E,N}(x)-1} (f^E \circ T_E^{k-1})(x)$$

for  $\mu$ -a.e.  $x \in E$ , since  $R_{E,N}(x)$  is diverging.

Looking at (2.14) we need now to consider the last term on the right hand side. We recall that if  $\tau_{E,x}(R_{E,N}(x) - 1) = N$  this term does not appear. If  $\tau_{E,x}(R_{E,N}(x) - 1) < N$  let us first write

$$\sum_{k=\tau_{E,x}(R_{E,N}(x)-1)+1}^N (f \circ T^{k-1})(x) = \sum_{k=\tau_{E,x}(R_{E,N}(x)-1)+1}^{\tau_{E,x}(R_{E,N}(x))} (f \circ T^{k-1})(x) - \sum_{k=N+1}^{\tau_{E,x}(R_{E,N}(x))} (f \circ T^{k-1})(x)$$

where  $\tau_{E,x}(R_{E,N}(x)) \geq N + 1$  by definition. Then by definition (2.12) and by (2.17)

$$\begin{aligned} (f^E \circ T_E^{R_{E,N}(x)-1})(x) &= \sum_{k=\tau_{E,x}(R_{E,N}(x)-1)+1}^{\tau_{E,x}(R_{E,N}(x))} (f \circ T^{k-1})(x) = \\ &= \sum_{j=0}^{(\varphi_E \circ T_E^{R_{E,N}(x)-1})(x)-1} f_j = G\left(\left(\varphi_E \circ T_E^{R_{E,N}(x)-1}\right)(x)\right) \end{aligned}$$

and by (2.17)

$$\sum_{k=N+1}^{\tau_{E,x}(R_{E,N}(x))} (f \circ T^{k-1})(x) = \sum_{j=1}^{\tau_{E,x}(R_{E,N}(x))-N} f_j = G\left(\tau_{E,x}(R_{E,N}(x)) - N + 1\right) - f_0.$$

Since  $(f^E \circ T_E^{k-1})(x) = G((\varphi_E \circ T_E^{k-1})(x))$ , using (4.13) in (2.14), we finally have

$$\begin{aligned} (4.14) \quad S_N f(x) &\sim d(R_{E,N}(x) - 1) + \max_{1 \leq k \leq R_{E,N}(x)-1} G\left((\varphi_E \circ T_E^{k-1})(x)\right) + \\ &+ G\left(\left(\varphi_E \circ T_E^{R_{E,N}(x)-1}\right)(x)\right) - G\left(\tau_{E,x}(R_{E,N}(x)) - N + 1\right) \end{aligned}$$

where we have discarded the constant  $f_0$ .

Let us fix  $x \in E$ , a constant  $\eta \in (0, \frac{1}{2})^1$ , and let  $(U_j)$  be the sequence of the return times to  $E$ , that is  $U_j = \tau_{E,x}(j)$  for  $j \geq 1$  and  $R_{E,U_j}(x) - 1 = j$ . Then we consider the following subintervals of  $\mathbb{N}$  depending on  $x$ :

- the intervals corresponding to excursions outside  $E$  which do not achieve a new record for the return time

$$(4.15) \quad \mathcal{I}_j := \left\{ N \in (U_j, U_{j+1}) : (\varphi_E \circ T_E^j)(x) \leq \max_{1 \leq k \leq j} (\varphi_E \circ T_E^{k-1})(x) \right\};$$

- the intervals corresponding to excursions outside  $E$  which achieve a new record for the return time

$$(4.16) \quad \mathcal{J}_j := \left\{ N \in (U_j, U_{j+1}) : (\varphi_E \circ T_E^j)(x) > \max_{1 \leq k \leq j} (\varphi_E \circ T_E^{k-1})(x) \right\};$$

we also need to consider the subintervals for which  $N$  is asymptotically not too far from  $U_{j+1}$  with respect to  $U_j$  up to  $\eta$  in the following sense

$$(4.17) \quad \mathcal{J}_j^\eta := \{N \in \mathcal{J}_j : U_{j+1} - N \leq \eta(N - U_j)\}.$$

<sup>1</sup>The bound  $\eta < \frac{1}{2}$  simplifies a step of the proof but is not necessary.

The sequence  $(U_j)$  and the intervals  $\mathcal{I}_j$  and  $\mathcal{J}_j$  depend on  $x$ , we have dropped this dependence in the notation for simplicity.

We now show that for  $\mu$ -a.e.  $x \in E$  we have

$$(4.18) \quad \lim_{N \rightarrow \infty} \frac{1}{G(N)} \sum_{k=1}^N (f \circ T^{k-1})(x) = 1.$$

The proof follows from the following lemmas.

**Lemma 4.5.** *For  $\mu$ -a.e.  $x \in X$  we have*

$$(4.19) \quad d(R_{E,N}(x) - 1) \sim d(R_{E,N}(x)) \sim G(N - w(N, E, x)).$$

*Proof.* Here we use the notions of slowly and regularly varying functions from Appendix B. By a remark in [AN03], condition (2.18) implies that the sequence  $(a(n))$  defined in (4.12) is a regularly varying function with exponent 1, and the same is true for its inverse  $d(n)$  (see for example [BGT87, Prop. 1.5.14]). For such a function we have that  $d(b_n) \sim d(b_n - 1)$  for all diverging sequences of positive numbers  $(b_n)$ . This shows that for  $\mu$ -a.e.  $x \in X$  we have  $d(R_{E,N}(x) - 1) \sim d(R_{E,N}(x))$ .

Moreover, from (4.8) and again from  $d$  being regularly varying with index 1, we have  $d(R_{E,N}(x)) \sim d(\alpha(N - w(N, E, x)))$ , where  $\alpha(n)$  is given in (2.11). Using now the theory of the *de Bruijn conjugate* of a slowly varying function (see [BGT87, Thm. 1.5.13]), letting  $a(n) = n/\ell(n)$  with

$$\ell(n) := \sum_{k=0}^n \mu(A_{\geq \Gamma(k)}),$$

by assumption (ii)-(e) we know that  $\ell$  is super-slowly varying at infinity with rate function itself (see Definition B.5), and by [BGT87, Cor. 2.3.4] it follows that its de Bruijn conjugate  $\ell^\#$  satisfies  $\ell^\#(n) \sim 1/\ell(n)$ . Hence

$$d(n) \sim n \ell(n)$$

and by (2.19)

$$d(\alpha(n)) \sim \alpha(n) \ell(\alpha(n)) \sim G(n)$$

as  $n \rightarrow \infty$ . It follows that

$$d(\alpha(N - w(N, E, x))) \sim G(N - w(N, E, x))$$

and the lemma is proved.  $\square$

*Remark 4.6.* Here we show that under condition (2.18) also the sequence  $(\alpha(n))$  defined in (2.11) is a regularly varying function with exponent 1. Indeed by the remark in [AN03], condition (2.18) implies that  $\gamma(t) = \sum_{k=1}^{\infty} \mu(A_k) \min\{kL(k), t\}$  is normalised slowly varying. (Indeed, the remark only states slow variation, but proves actually normalised slow variation.) Since  $L(t)$  is normalised slowly varying and  $nL(n)$  is non-decreasing, we can show that

$$\gamma(tL(t)) = \sum_{k=1}^{\infty} \mu(A_k) \min\{kL(k), tL(t)\} = \sum_{k=1}^{\lfloor t \rfloor} \mu(A_{>k}) \left( kL(k) - (k-1)L(k-1) \right)$$

is normalised slowly varying as well. Referring to (B.2), we may write  $\gamma(t) = \exp(\int_{\kappa}^t (\eta(x)/x) dx)$  with  $\eta(x)$  tending to zero and  $L(t) = \exp(\int_{\kappa}^t (\varepsilon(x)/x) dx)$  with  $\varepsilon(x)$  tending to zero, by extending the domain of definition of  $L$  by linear interpolation. Then we have for  $t \notin \mathbb{N}$

$$(4.20) \quad \frac{d}{dt} \gamma(tL(t)) = \frac{\gamma(tL(t)) \eta(tL(t)) (1 + \varepsilon(t))}{t}$$

and since  $\eta(tL(t)) (1 + \varepsilon(t))$  tends to zero,  $\gamma(tL(t))$  is normalised slowly varying (see eq. (1.3.4) and subsequent comments in [BGT87]).

In the next steps we will compare  $\gamma(tL(t))$  with the wandering rate  $w_n(E) := \sum_{k=1}^{n-1} \mu(A_{>k})$ . Letting  $\tilde{L}(k) := 1/(kL(k) - (k-1)L(k-1))$ , we have

$$\begin{aligned} w_n(E) &= \sum_{k=1}^{n-1} \mu(A_{>k}) (kL(k) - (k-1)L(k-1)) \frac{1}{kL(k) - (k-1)L(k-1)} = \\ &= \sum_{k=1}^{n-1} \left( \gamma(kL(k)) - \gamma((k-1)L(k-1)) \right) \tilde{L}(k) = \sum_{k=1}^{n-1} \int_{k-1}^k \tilde{L}(k) \left( \frac{d}{dt} \gamma(tL(t)) \right) dt. \end{aligned}$$

Since  $\tilde{L}(k)$  is asymptotic to  $1/L(k)$  it is slowly varying, then by (4.20) and [BGT87, Prop. 1.5.9a] also  $w_n(E)$  is slowly varying. Thus,  $\alpha(n)$  is a regularly varying function with exponent 1.

**Lemma 4.7.** *The limit in (4.18) holds for the sequence  $(U_j)$ .*

*Proof.* For the sequence  $(U_j)$ , by (4.13) and (4.19) we first write

$$S_{U_j} f(x) = \sum_{n=1}^j (f^E \circ T_E^{n-1})(x) \sim G(U_j - w(U_j, E, x)) + \max_{1 \leq k \leq j} G((\varphi_E \circ T_E^{k-1})(x)).$$

By (4.7)

$$w(U_j, E, x) = \max_{1 \leq k \leq j} (\varphi_E \circ T_E^{k-1})(x)$$

and since  $G(n)$  is non-decreasing it follows

$$S_{U_j} f(x) \sim G(U_j - w(U_j, E, x)) + G(w(U_j, E, x)) \sim G(U_j)$$

by Lemma B.1. □

**Lemma 4.8.** *The limit in (4.18) holds for a sequence  $(N_j)$  such that  $N_j \in \mathcal{I}_j$  for all  $j$ .*

*Proof.* By (4.15) and (4.7), for all  $j$  it holds

$$w(N_j, E, x) = \max_{1 \leq k \leq R_{E, N_j}(x)-1} (\varphi_E \circ T_E^{k-1})(x).$$

Hence by (4.14) and (4.19) we have, since  $G(n)$  is non-decreasing,

$$\begin{aligned} S_{N_j} f(x) &\sim G(N_j - w(N_j, E, x)) + G(w(N_j, E, x)) + \\ &+ G\left(\left(\varphi_E \circ T_E^{R_{E, N_j}(x)-1}\right)(x)\right) - G\left(\tau_{E, x}(R_{E, N_j}(x)) - N + 1\right). \end{aligned}$$

Moreover, since  $(f^E \circ T_E^{n-1})$  satisfies the assumptions of Lemma 2.2 with  $W = 1$ , we apply (2.10) to obtain

$$G\left(\left(\varphi_E \circ T_E^{R_{E, N_j}(x)-1}\right)(x)\right) = o(d(R_{E, N_j}(x))),$$

because by (4.15)

$$G\left(\left(\varphi_E \circ T_E^{R_{E, N_j}(x)-1}\right)(x)\right) \leq G\left(\max_{1 \leq k \leq R_{E, N_j}(x)-1} (\varphi_E \circ T_E^{k-1})(x)\right)$$

so that  $G\left(\left(\varphi_E \circ T_E^{R_{E, N_j}(x)-1}\right)(x)\right)$  is the 2-nd maximum in

$$\left\{ G\left(\varphi_E \circ T_E^{k-1}(x)\right) : k = 1, \dots, R_{E, N_j}(x) \right\}.$$

Since  $G(n)$  is non-negative, by using (4.19) and Lemma B.1 the statement is proved. □

**Lemma 4.9.** *The limit in (4.18) holds for a subsequence  $(N_j)$  such that  $N_j \in \mathcal{J}_j^\eta$  for all  $j$ .*

*Proof.* Looking at (4.14), we first consider the last terms on the right-hand side. By the definition of  $(U_j)$  and (2.13)

$$(4.21) \quad \begin{aligned} \left(\varphi_E \circ T_E^{R_{E,N_j}(x)-1}\right)(x) &= \left(\varphi_E \circ T_E^{R_{E,U_j}(x)-1}\right)(x) = (\varphi_E \circ T_E^j)(x) = U_{j+1} - U_j, \\ \tau_{E,x}(R_{E,N_j}(x)) - N_j &= \tau_{E,x}(R_{E,U_j}(x)) - N_j = \tau_{E,x}(j+1) - N_j = U_{j+1} - N_j \end{aligned}$$

and by (4.16), the sequence  $U_{j+1} - U_j$  is diverging. Moreover writing

$$(\varphi_E \circ T_E^j)(x) = \left(U_{j+1} - N_j + 1\right) + \left(N_j - U_j - 1\right)$$

and using that  $N_j \in \mathcal{J}_j^\eta$ , letting  $a_j := U_{j+1} - N_j + 1$  and  $b_j := N_j - U_j - 1$ , we have that  $b_j$  is diverging by (4.17). Then by Lemma B.1 we have

$$(4.22) \quad G\left(\left(\varphi_E \circ T_E^{R_{E,N_j}(x)-1}\right)(x)\right) \sim G\left(U_{j+1} - N_j + 1\right) + G\left(N_j - U_j - 1\right).$$

Moreover, we use (4.22) and the fact that  $G(n) = nL(n)$  with  $L$  slowly varying to prove that

$$(4.23) \quad G\left(\left(\varphi_E \circ T_E^{R_{E,N_j}(x)-1}\right)(x)\right) - G\left(U_{j+1} - N_j + 1\right) \sim G\left(N_j - U_j - 1\right).$$

In fact, by (4.22) for all  $\epsilon > 0$  there exists  $\bar{j} \in \mathbb{N}$  such that for all  $j \geq \bar{j}$  we have

$$1 - \epsilon \leq \frac{G\left(\left(\varphi_E \circ T_E^{R_{E,N_j}(x)-1}\right)(x)\right)}{G\left(U_{j+1} - N_j + 1\right) + G\left(N_j - U_j - 1\right)} \leq 1 + \epsilon.$$

Since  $N_j - U_j \geq 1$  and  $\eta \in (0, \frac{1}{2})$ , for  $N_j \in \mathcal{J}_j^\eta$  we have

$$U_{j+1} - N_j + 1 \leq \frac{1}{2}(N_j - U_j) + 1 \leq N_j - U_j - 1,$$

hence  $G(U_{j+1} - N_j + 1) \leq G(N_j - U_j - 1)$  because  $G$  is non-decreasing. It follows that for  $j \geq \bar{j}$

$$1 - 2\epsilon \leq \frac{G\left(\left(\varphi_E \circ T_E^{R_{E,N_j}(x)-1}\right)(x)\right) - G\left(U_{j+1} - N_j + 1\right)}{G\left(N_j - U_j - 1\right)} \leq 1 + 2\epsilon.$$

By the arbitrariness of  $\epsilon$  we have proved (4.23).

Using (4.13) and Lemma 4.7 for the first two terms on the right-hand side of (4.14), and (4.23) for the last difference of (4.14), we have

$$S_{N_j} f(x) \sim G(U_j) + G(N_j - U_j - 1).$$

Finally Lemma B.1 and the fact that  $N_j - U_j$  is diverging imply  $S_{N_j} f(x) \sim G(N_j)$  and the lemma is proved.  $\square$

**Lemma 4.10.** *The limit in (4.18) holds for a subsequence  $(N_j)$  such that  $N_j \in \mathcal{J}_j \setminus \mathcal{J}_j^\eta$  for all  $j$ .*

*Proof.* As in the proof of Lemma 4.9 we need to study the last two terms on the right-hand side of (4.14), for which we recall (4.21). For the first two terms, Lemma 4.7 shows that they are asymptotic to  $G(U_j)$ .

To estimate the third and fourth term of (4.14) we aim to apply Lemma B.2, for which we use that  $U_{j+1} - U_j$  tends to infinity. Moreover we now need to consider two cases. Let us introduce the sets

$$\mathcal{L}_j^0 := \left\{ N \in \mathcal{J}_j : \lim_{j \rightarrow \infty} \frac{U_{j+1} - U_j}{U_j} = 0 \right\}$$



and

$$\mathcal{L}_j := \left\{ N \in \mathcal{J}_j : \exists \ell \in (0, \infty] \text{ s.t. } \lim_{j \rightarrow \infty} \frac{U_{j+1} - U_j}{U_j} = \ell \right\}.$$

It is clear that for all sequences  $(N_j)$  there exists a subsequence in  $\mathcal{L}_j^0$  or in  $\mathcal{L}_j$ . Thus we restrict ourselves to the cases  $N_j \in \mathcal{L}_j^0$  for all  $j$ , or  $N_j \in \mathcal{L}_j$  for all  $j$ .

First let  $N_j \in (\mathcal{J}_j \setminus \mathcal{J}_j^\eta) \cap \mathcal{L}_j^0$  for all  $j$ . Then  $U_j \sim U_{j+1}$  and  $N_j \in (U_j, U_{j+1})$  imply  $G(U_j) \sim G(U_{j+1}) \sim G(N_j)$  and Lemma 4.7 and the fact that  $S_{U_j}f(x) < S_{N_j}f(x) < S_{U_{j+1}}f(x)$  imply that the limit (4.18) follows for the sequences in  $\mathcal{L}_j^0$ .

Let now  $N_j \in (\mathcal{J}_j \setminus \mathcal{J}_j^\eta) \cap \mathcal{L}_j$  for all  $j$  and let  $\ell \in (0, \infty)$  be the limit of  $(U_{j+1} - U_j)/U_j$ . Notice that in this case we also have  $\lim_j (U_{j+1} - N_j) = \infty$ . A similar argument works for the case  $\ell = \infty$ . Letting  $a_j := U_{j+1} - U_j$  and  $b_j := U_{j+1} - N_j + 1$ , we apply Lemma B.2 to write

$$(4.24) \quad G(U_{j+1} - U_j) - G(U_{j+1} - N_j + 1) \sim (N_j - U_j - 1) L(U_{j+1} - U_j).$$

To study the term  $L(U_{j+1} - U_j)$ , first by Lemma 2.6 for  $\mu$ -a.e.  $x \in E$  we have for  $j$  large enough

$$(4.25) \quad U_{j+1} - U_j = \max_{1 \leq k \leq R_{E, U_j}(x)} (\varphi_E \circ T_E^{k-1})(x) \leq U_j \xi(U_j).$$

Then we want to use assumption (ii)-(c) to obtain the asymptotic behaviour of  $L(U_{j+1} - U_j)$ .

Since  $N_j \in (\mathcal{J}_j \setminus \mathcal{J}_j^\eta) \cap \mathcal{L}_j$ , for all  $\epsilon > 0$  there exists  $j$  big enough such that  $U_{j+1} - U_j \geq (\ell - \epsilon)U_j$ . From (4.25) it follows that  $\tilde{\xi}(U_j) := (U_{j+1} - U_j)/U_j$  satisfies the conditions of assumption (ii)-(c), hence  $L(U_{j+1} - U_j) \sim L(U_j)$ . From (4.24) and Lemma 4.7 we have

$$(4.26) \quad S_{N_j}f(x) \sim G(U_j) + (N_j - U_j - 1) L(U_j).$$

At this point, if

$$\liminf_{j \rightarrow \infty} \frac{N_j - U_j}{U_j} = 0,$$

for all  $\epsilon > 0$  and for  $j$  big enough we have, up to the choice of a subsequence,  $0 < N_j - U_j < \epsilon U_j$ , and thus

$$G(U_j) + (N_j - U_j - 1) L(U_j) \leq G(U_j) + \epsilon U_j L(U_j) \leq (1 + \epsilon) G(N_j),$$

and in the other direction

$$G(U_j) \geq G\left(\frac{1}{1 + \epsilon} N_j\right) \gtrsim \frac{1}{1 + \epsilon} G(N_j).$$

Hence by the arbitrariness of  $\epsilon$ , we find  $S_{N_j}f(x) \sim G(N_j)$  in (4.26).

Instead, if

$$\liminf_{j \rightarrow \infty} \frac{N_j - U_j}{U_j} = \bar{\ell} > 0,$$

then for all  $\epsilon > 0$  and for  $j$  big enough we have  $N_j - U_j \geq (\bar{\ell} - \epsilon)U_j$ . Hence letting  $\tilde{\xi}(U_j) := (N_j - U_j)/U_j$ , we can apply assumption (ii)-(c) since

$$\bar{\ell} - \epsilon \leq \tilde{\xi}(U_j) \leq \frac{U_{j+1} - U_j}{U_j} \leq \xi(U_j).$$

Thus it follows that  $L(N_j - U_j) \sim L(U_j)$ , and in (4.26) we find

$$S_{N_j}f(x) \sim G(U_j) + (N_j - U_j - 1) L(N_j - U_j) \sim G(U_j) + G(N_j - U_j) \sim G(N_j)$$

by applying Lemma B.1 and the fact that  $(U_j)$  and  $(N_j - U_j)$  are diverging sequences.

This concludes the proof of the lemma.  $\square$

From Lemmas 4.7 - 4.10 the limit in (4.18) holds and the theorem is proved for functions  $f$  constant on  $E_n$ .

In the general case, for a function  $f : X \rightarrow \mathbb{R}_{\geq 0}$  satisfying assumption (ii) we obtain that

$$(4.27) \quad N L_1(N) \sim S_N g_1(x) \leq S_N f(x) \leq S_N g_2(x) \sim N L_2(N)$$

for  $\mu$ -a.e.  $x$ , since we can apply Lemmas 4.7 - 4.10 to  $g_1, g_2$ . Moreover, for  $\mu$ -a.e.  $x \in A_n$  we have  $g_1^E(n) \leq f^E(x) \leq g_2^E(n)$  for all  $n \in \mathbb{N}$ . Since  $g_1^E(n) \sim g_2^E(n)$  by assumption (ii)-(b), the function  $G(n)$  which is asymptotically equivalent to  $f^E$  is well-defined and satisfies  $G(n) \sim g_i^E(n) = n L_i(n)$  for  $i = 1, 2$ . Thus by (4.27) we obtain  $S_N f(x) \sim G(N)$  for  $\mu$ -a.e.  $x$  and the theorem is proved.  $\square$

## 5. PROOFS OF THE STATEMENTS FROM SECTION 3

**Proof of Proposition 3.3.** In the proof of [BDM21, Theorem 3.1] it is shown that  $(E, S_E)$  is a fibred system with respect to the level sets of the return time function  $\varphi_E$  and that assumption (iii) is satisfied, hence it is  $\psi$ -mixing. It remains to show that  $\sum_{n \geq 1} \psi(n)/n < \infty$ .

Using (3.3) we need to study the sequence  $\sigma(k)$ . Moreover it is proved in [BDM21] that there exists a constant  $C > 0$  such that  $\sigma(k) \leq C d(k)$ , where  $d(k)$  is defined by

$$d(k) = \sum_{j=0}^2 \left( \left| \frac{f_{k+j+2}}{f_{k+6}} - \frac{f_{k+j+1}}{f_{k+5}} \right| + \left| \frac{f_{k+j+2}}{f_{k+6}} - \frac{f_{k+j}}{f_{k+4}} \right| \right)$$

being  $(f_k)$  the sequence recursively defined as

$$f_0 = 0, f_1 = 1, f_2 = 0, \quad f_{k+3} = f_{k+2} + f_k, \quad \forall k \geq 0.$$

Then given the distinct roots  $\lambda, \mu, \bar{\mu}$  of the polynomial  $p(t) = t^3 - t^2 - 1$ , with  $\lambda > 1$  and  $|\mu|^2 = |\bar{\mu}|^2 = \lambda^{-1} < 1$ , there exist constants  $c_1, c_2, c_3 \in \mathbb{C}$  such that

$$f_k = c_1 \lambda^k + c_2 \mu^k + c_3 \bar{\mu}^k, \quad \forall k \geq 0.$$

It follows that  $d(k) = O(\lambda^{-k})$ , so that  $\sigma(k) = O(\lambda^{-k})$ , and finally  $\sum_{n \geq 1} \psi(n)/n < \infty$ .  $\square$

**Proof of Proposition 3.5.** We show that the assumptions of Theorem 2.7 are satisfied with  $G(N) \sim N$  as  $N \rightarrow \infty$ . It is clear that assumptions (ii)-(a), (b) and (c) are satisfied by (c1). In addition, since  $\Gamma_i(k) \sim k$  we can argue as in the proof of Lemma 4.3 to show that (c2) implies (ii)-(d).

Thus it remains to show that (ii)-(e) is satisfied. First, in this case we have  $\alpha(n) \sim n/\ell(n)$ , hence (2.19) becomes  $\ell(N/\ell(N)) \sim \ell(N)$ . Finally, using that  $\ell(n)$  is an increasing sequence which diverges, for all  $\delta \in [0, 1]$  we have the inequalities

$$\ell(n \ell^\delta(n)) \geq \ell(n)$$

and

$$\ell\left(\frac{n \ell^\delta(n)}{\ell(n \ell^\delta(n))}\right) \leq \ell\left(\frac{n \ell(n)}{\ell(n)}\right) = \ell(n).$$

Thus by (c3)

$$\ell(n) \leq \ell(n \ell^\delta(n)) \sim \ell\left(\frac{n \ell^\delta(n)}{\ell(n \ell^\delta(n))}\right) \leq \ell(n)$$

and we have shown that

$$(5.1) \quad \lim_{N \rightarrow \infty} \frac{\ell(n \ell^\delta(n))}{\ell(n)} = 1$$

for all  $\delta \in [0, 1]$ . This can be easily made uniform in  $\delta$ , proving that  $\ell$  is super-slowly varying at infinity with rate function itself (see Definition B.5). However we need this property of  $\ell$  only in the proof of Lemma 4.5 to apply [BGT87, Cor. 2.3.4], for which we only need (5.1) with  $\delta = 1$ .

Thus all the assumptions of Theorem 2.7 are satisfied with  $G(N) \sim g_i^E(N) \sim N$ .  $\square$

**Proof of Proposition 3.8.** All assumptions of Theorem 2.7 except for (ii)-(e) are satisfied, hence we can make use of all the results leading to the proof of Theorem 2.7 except for Lemma 4.5.

We will only show the divergence result of this proposition for the subsequence  $(U_j)$  fulfilling  $U_j = \tau_{E,x}(j)$  for  $j \geq 1$ . By (2.14) and (4.13) we have for this subsequence

$$(5.2) \quad \sum_{n=1}^{U_j} (f \circ T^{n-1})(x) \sim d(R_{E,U_j}(x) - 1) + \max_{1 \leq k \leq R_{E,U_j}(x)-1} (f^E \circ T_E^{k-1})(x).$$

To estimate the first summand we obtain by the same argumentation as in Lemma 4.5 that  $d(R_{E,U_j}(x)) \sim d(\alpha(U_j - w(U_j, E, x)))$ . We may use again  $\mu(A_{\geq \Gamma_n}) \sim \log n/n$  with which we calculate  $a(n) \sim 2n/\log^2(n)$  with  $a$  as in (4.12). Hence, we obtain for its asymptotic inverse function  $d(n) \sim n \log^2(n)/2$ . Using then  $\alpha(n) \sim n/\log n$  from above and (4.8) implies

$$(5.3) \quad d(R_{E,U_j}(x) - 1) \sim \frac{1}{2} (U_j - w(U_j, E, x)) \log(U_j - w(U_j, E, x)).$$

On the other hand, for the second summand of (5.2) we obtain from

$$(f^E \circ T_E^{k-1})(x) \sim (\varphi_E \circ T_E^{k-1})(x) \log(\varphi_E \circ T_E^{k-1})(x),$$

with  $\sim$  in the meaning that  $(f^E \circ T_E^{k-1})(x)$  large, and (4.7) that

$$(5.4) \quad \max_{1 \leq k \leq R_{E,U_j}(x)-1} (f^E \circ T_E^{k-1})(x) \sim w(U_j, E, x) \log w(U_j, E, x).$$

From Theorem 2.4 and the fact that  $\alpha$  is regularly varying with index 1 we can conclude that there exist  $u'_1 < u'_2$  such that for  $\mu$ -a.e.  $x \in X$  we have

$$\liminf_j w(U_j, E, x)/U_j < u'_1 < u'_2 < \limsup_j w(U_j, E, x)/U_j.$$

Indeed if the limit of  $w(U_j, E, x)/U_j$  existed, the result of Theorem 2.4 would contradict [A97, Thm. 2.4.2]. Hence, (5.2), (5.3) and (5.4) together imply the statement of Proposition 3.8.  $\square$

## APPENDIX A. PROOF OF LEMMA 2.2

Lemma 2.2 is an equivalent formulation to [AN03, Theorem 1.1]. Instead of the condition in (2.9) this theorem asks for the following conditions to hold:  $J_r := \sum_{n=1}^{\infty} \epsilon(n)^r/n < \infty$  with  $\epsilon(n) = n (\log^+ L)'(n)$  and  $L(t) = \mathbb{E}(\min\{t, Y_1\})$ . We have that

$$L(t) = \int_0^t x dF(x) + t(1 - F(t)) = tF(t) - \int_0^t F(x) dx + t(1 - F(t)) = \int_0^t (1 - F(x)) dx.$$

Furthermore,

$$(\log^+ L)'(t) = \frac{L'(t)}{L(t)} = \frac{1 - F(t)}{\int_0^t (1 - F(x)) dx} \quad \text{and} \quad \epsilon(t) = \frac{t(1 - F(t))}{\int_0^t (1 - F(x)) dx}.$$

Thus, the condition  $J_r := \sum_{n=1}^{\infty} \epsilon(n)^r/n < \infty$  and (2.9) are equivalent.

Moreover, in [AN03] the norming sequence  $(d(n))$  is set to be the inverse of  $a(t) = t/L(t) = t/\int_0^t (1 - F(x)) dx$ . Since by the remark before Theorem 1.1 in [AN03],  $L$  is slowly varying (see

Appendix B),  $\lim_{t \rightarrow \infty} t/L(t) = \infty$ . Moreover, one can easily verify that  $a'(t) > 0$  on an interval  $[K, \infty)$  and thus one can consider  $a(t)$  as an invertible function.

Equation (2.10) follows from Theorem 1.1.(ii) and its following remark in [AN03].  $\square$

## APPENDIX B. SLOWLY VARYING FUNCTIONS

In this section we collect the results on slowly varying functions used in the proof of Theorem 2.7. We recall that a function  $L : \mathbb{R}_{>0} \rightarrow \mathbb{R}$  is called *regularly varying (at infinity) with index  $\gamma \in \mathbb{R}$*  if for all  $r > 0$  it holds

$$\lim_{x \rightarrow \infty} \frac{L(rx)}{L(x)} = r^\gamma.$$

If  $\gamma = 0$ ,  $L$  is called a *slowly varying (at infinity)*. We refer to [BGT87] for more details. By Karamata's representation theorem (see for example [BGT87, Thm. 1.3.1]) we have that each slowly varying function  $L$  can be written as

$$(B.1) \quad L(x) = c(x) \exp \left( \int_\kappa^x \frac{\eta(t)}{t} dt \right),$$

where  $c(x)$  tends to a constant  $C$  as  $x \rightarrow \infty$ ,  $\kappa \geq 0$  and  $\eta(x)$  tends to zero as  $x \rightarrow \infty$ , and a function  $L$  is called *normalized slowly varying* if it can be written as

$$(B.2) \quad L(x) = C \exp \left( \int_\kappa^x \frac{\eta(t)}{t} dt \right).$$

It is immediately clear that for each slowly varying function  $L$  there exists a normalised slowly varying function  $\tilde{L}$  such that  $L(x) \sim \tilde{L}(x)$ .

In the proof of our results we need *Potter's bound*, see for example [BGT87, Thm. 1.5.6]. If  $L$  is a slowly varying function at infinity, then for all constants  $\delta > 0$  and  $A > 1$  there exists  $C = C(\delta, A)$  such that for all  $x, y \geq C$  we have

$$(B.3) \quad \frac{L(x)}{L(y)} \leq A \max \left\{ \left( \frac{x}{y} \right)^\delta, \left( \frac{y}{x} \right)^\delta \right\}.$$

**Lemma B.1.** *If  $(a_n)$  and  $(b_n)$  are two sequences of non-negative reals, at least one of them diverging, and  $L$  is non-negative and slowly varying such that  $(nL(n))$  is non-decreasing, then*

$$a_n L(a_n) + b_n L(b_n) \sim (a_n + b_n) L(a_n + b_n).$$

*Proof.* We begin writing

$$a_n L(a_n) + b_n L(b_n) = (a_n + b_n) L(a_n) - b_n (L(a_n) - L(b_n)).$$

Without restriction of generality we assume that  $a_n > b_n$  for all  $n$ . (Otherwise we could just define two new sequences  $c_n := \max\{a_n, b_n\}$  and  $d_n := \min\{a_n, b_n\}$  and continue with these sequences.) By Potter's bound (B.3) we have for all  $\epsilon > 0$  that there exists  $N \in \mathbb{N}$  such that for  $n \geq N$

$$L(a_n + b_n) \leq (1 + \epsilon) L(a_n) \left( \frac{a_n + b_n}{a_n} \right)^\epsilon \leq (1 + \epsilon) 2^\epsilon L(a_n)$$

and

$$L(a_n + b_n) \geq (1 - \epsilon) L(a_n) \left( \frac{a_n}{a_n + b_n} \right)^\epsilon \geq (1 - \epsilon) 2^{-\epsilon} L(a_n),$$

implying  $L(a_n + b_n) \sim L(a_n)$  and thus  $(a_n + b_n) L(a_n) \sim (a_n + b_n) L(a_n + b_n)$ .

Next, we have a closer look at  $b_n (L(a_n) - L(b_n))$  distinguishing two cases. For a given  $\epsilon > 0$  we first look at the subsequence  $n_k$  fulfilling  $b_{n_k} \leq \epsilon a_{n_k}$  for which we have

$$b_{n_k} (L(a_{n_k}) - L(b_{n_k})) < \epsilon (a_{n_k} + b_{n_k}) L(a_{n_k}) \sim \epsilon (a_{n_k} + b_{n_k}) L(a_{n_k} + b_{n_k}).$$

On the other hand,  $nL(n)$  being non-decreasing and non-negative we have

$$\begin{aligned} b_{n_k} (L(a_{n_k}) - L(b_{n_k})) &\geq -b_{n_k} L(b_{n_k}) \geq -\epsilon (a_{n_k} + b_{n_k}) L(\epsilon (a_{n_k} + b_{n_k})) \\ &\sim -\epsilon (a_{n_k} + b_{n_k}) L(a_{n_k} + b_{n_k}). \end{aligned}$$

Next, we look at the subsequence fulfilling  $b_{n_k} > \epsilon a_{n_k}$ . Applying again (B.3), we have that for  $k$  big enough

$$L(a_{n_k} + b_{n_k}) \leq (1 + \epsilon) L(b_{n_k}) \left( \frac{a_{n_k} + b_{n_k}}{b_{n_k}} \right)^\epsilon \leq \frac{(1 + \epsilon)^{1+\epsilon}}{\epsilon^\epsilon} L(b_{n_k})$$

and

$$L(a_{n_k} + b_{n_k}) \geq \frac{1}{1 + \epsilon} L(b_{n_k}) \left( \frac{b_{n_k}}{a_{n_k} + b_{n_k}} \right)^\epsilon \geq \frac{\epsilon^\epsilon}{(1 + \epsilon)^{1+\epsilon}} L(b_{n_k}).$$

Since  $L(a_{n_k} + b_{n_k}) \sim L(a_{n_k})$  we also have that for  $k$  big enough

$$\frac{1}{1 + \epsilon} \leq \frac{L(a_{n_k} + b_{n_k})}{L(a_{n_k})} \leq 1 + \epsilon,$$

therefore there is a constant  $c > 0$  such that choosing  $k$  big enough we have

$$\begin{aligned} b_{n_k} (L(a_{n_k}) - L(b_{n_k})) &\leq (a_{n_k} + b_{n_k}) \left( (1 + \epsilon)L(a_{n_k} + b_{n_k}) - \frac{\epsilon^\epsilon}{(1 + \epsilon)^{1+\epsilon}} L(a_{n_k} + b_{n_k}) \right) \\ &\leq c\epsilon^{1/2} (a_{n_k} + b_{n_k}) L(a_{n_k} + b_{n_k}) \end{aligned}$$

for  $\epsilon \in (0, 1)$ , and similarly

$$\begin{aligned} b_{n_k} (L(a_{n_k}) - L(b_{n_k})) &\geq b_{n_k} \left( \frac{1}{1 + \epsilon} L(a_{n_k} + b_{n_k}) - \frac{(1 + \epsilon)^{1+\epsilon}}{\epsilon^\epsilon} L(a_{n_k} + b_{n_k}) \right) \\ &\geq -c\epsilon^{1/2} (a_{n_k} + b_{n_k}) L(a_{n_k} + b_{n_k}) \end{aligned}$$

for  $\epsilon \in (0, 1)$ .

Since  $\epsilon$  was chosen arbitrarily, we have that

$$b_{n_k} (L(a_{n_k}) - L(b_{n_k})) = o((a_{n_k} + b_{n_k}) L(a_{n_k} + b_{n_k})),$$

and the lemma is proved.  $\square$

**Lemma B.2.** *Let  $(a_n)$  and  $(b_n)$  be two sequences of non-negative reals with  $a_n > b_n$  and  $(a_n)$  tending to infinity. Further let  $L$  be a non-negative normalized slowly varying function. Then*

$$(B.4) \quad a_n L(a_n) - b_n L(b_n) \sim (a_n - b_n) L(a_n).$$

*Proof.* First we look at a subsequence  $n_j$  for which  $b_{n_j}$  is bounded, in this case (B.4) immediately holds. Hence, we may assume without loss of generality that also  $b_n$  tends to infinity. We have

$$\begin{aligned} a_n L(a_n) - b_n L(b_n) &= (a_n - b_n) L(a_n) + b_n (L(a_n) - L(b_n)) \\ &= b_n \left( \frac{a_n}{b_n} - 1 \right) L(a_n) + b_n \left( 1 - \frac{L(b_n)}{L(a_n)} \right) L(a_n). \end{aligned}$$

Hence, we have to show that

$$1 - \frac{L(b_n)}{L(a_n)} = o\left(\frac{a_n}{b_n} - 1\right).$$

To do so, using (B.2) we note that

$$\frac{L(a_n)}{L(b_n)} = \exp\left(\int_{b_n}^{a_n} \frac{\eta(t)}{t} dt\right) \leq \exp\left(\sup_{y \in [b_n, a_n]} \eta(y) \int_{b_n}^{a_n} \frac{1}{t} dt\right)$$

and similarly

$$\frac{L(a_n)}{L(b_n)} = \exp\left(\int_{b_n}^{a_n} \frac{\eta(t)}{t} dt\right) \geq \exp\left(\inf_{y \in [b_n, a_n]} \eta(y) \int_{b_n}^{a_n} \frac{1}{t} dt\right).$$

Since  $(b_n)$  tends to infinity and  $\eta$  tends to zero we can conclude that for all  $\epsilon > 0$  there exists  $N \in \mathbb{N}$  such that for all  $n \geq N$  we have

$$\left(\frac{a_n}{b_n}\right)^{-\epsilon} \leq \frac{L(a_n)}{L(b_n)} \leq \left(\frac{a_n}{b_n}\right)^\epsilon.$$

Hence, for all  $\epsilon > 0$  there exists  $N \in \mathbb{N}$  such that for all  $n \geq N$

$$\frac{\left|1 - \frac{L(b_n)}{L(a_n)}\right|}{\left|\frac{a_n}{b_n} - 1\right|} \leq \frac{\left(\frac{a_n}{b_n}\right)^\epsilon - 1}{\frac{a_n}{b_n} - 1}.$$

Since the function  $g: (1, \infty) \rightarrow \mathbb{R}$  given by  $(x^\epsilon - 1)/(x - 1)$  is monotonically decreasing and tends to  $\epsilon$  for  $x \rightarrow 1^+$ , it follows that for all  $\epsilon > 0$  there exists  $N \in \mathbb{N}$  such that for all  $n \geq N$

$$\left|1 - \frac{L(b_n)}{L(a_n)}\right| \leq \epsilon \left|\frac{a_n}{b_n} - 1\right|.$$

Since  $\epsilon$  was arbitrary, the statement is proven.  $\square$

**Lemma B.3.** *Let  $L$  be a slowly varying function at infinity. Then there exist two normalised slowly varying functions  $L^-$  and  $L^+$  such that  $L^-(n) \leq L(n) \leq L^+(n)$  for all  $n \in \mathbb{N}$  and  $L^-(n) \sim L(n) \sim L^+(n)$  as  $n \rightarrow \infty$ .*

*Proof.* Let  $L$  be represented as in (B.1) for fixed  $c(x)$ ,  $\kappa$  and  $\eta(x)$ . Our approach is to find two normalised slowly varying functions  $c^-, c^+$  fulfilling  $c^- \leq c \leq c^+$  and  $c^-(n) \sim c(n) \sim c^+(n)$ . Then we may set

$$L^\pm(n) = c^\pm(n) \exp\left(\int_\kappa^n \eta(t)/t dt\right).$$

Since a product of two normalised slowly varying functions is still normalised slowly varying, the functions  $L^\pm$  fulfil all the required properties. We will in the following only give the construction of  $c^-$  as the construction of  $c^+$  follows analogously.

We notice that a function  $\ell$  is normalised slowly varying if it fulfils  $\ell'(x) = o(\ell(x)/x)$  almost everywhere. This follows immediately by taking the derivative of (B.2), but see also [BGT87, p. 15]. We construct a function  $c^-$  on a line  $[\gamma_1, \infty)$  which is continuous and piecewise differentiable.

In order to define  $c^-$  let  $\Gamma := \{n \in \mathbb{N} : c(n) < c(k) \text{ for all } k > n\}$  and let  $(\gamma_n)$  be the ordered sequence of elements in  $\Gamma$ , that is we have  $\gamma_1 < \gamma_2 < \dots$  and  $c(\gamma_1) < c(\gamma_2) < \dots$ . The set  $\Gamma$  contains infinitely many elements unless  $\min\{c(n), C\} = C$  for sufficiently large  $n$  which is a trivial case in which we might simply set  $c^-(n) = \min\{c(n), C\}$ . In this case  $(c^-)'(x) = 0$  for  $x$  sufficiently large and thus  $(c^-)'(x) \leq (x \log x)^{-1}$  holds immediately.

As we are only interested in the limit behaviour, it is sufficient to define  $c^-$  at  $\mathbb{R}_{\geq \gamma_1}$  and we set  $c^-(\gamma_i) = c(\gamma_i)$ , for all  $i \in \mathbb{N}$ . For  $x \in (\gamma_i, \gamma_{i+1})$  we set

$$c^-(x) := \max\left\{(\log x)^{1/2} + c(\gamma_i) - (\log \gamma_i)^{1/2}, c(\gamma_{i+1})\right\}.$$

The function  $c^-$  is then either piecewise constant and on those parts its derivative equals zero, or has derivative equal to  $1/2(\log x)^{-1/2} x^{-1} = o(x^{-1})$ . This is sufficient since in our case,  $c^-$  is bounded by  $C$ . Thus, at those points at which  $c^-$  is differentiable, it fulfils  $(c^-)'(x) = o(c^-(x)/x)$  and there are only countably many points on which it is not differentiable. On the other hand, since  $(\log x)^{1/2}$  tends to infinity,  $c^-$  tends to  $C$ .  $\square$

**Lemma B.4.** *Let  $h$  be a slowly varying function with  $h(t) \geq 1 + h_0$  for all  $t$  and for some  $h_0 > 0$ , and let  $L$  be a slowly varying function written as in (B.1) with  $\eta(t) = o(1/\log \xi(t))$  as  $t \rightarrow \infty$ . Then for all functions  $\tilde{h}$  which satisfy  $c \leq \tilde{h}(n) \leq h(n)$  for some constant  $c > 0$  and all  $n$ , it holds  $L(n\tilde{h}(n)) \sim L(n)$  as  $n \rightarrow \infty$ .*

The proof of the lemma follows by applying results in the theory of functions with regular variations for which we first need to introduce some definitions.

**Definition B.5.** Given a function  $h : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ , a function  $\ell : \mathbb{R}_{>0} \rightarrow \mathbb{R}$  is called *super-slowly varying at infinity with rate function  $h$*  if

$$\lim_{x \rightarrow \infty} \frac{\ell(x h(x)^\delta)}{\ell(x)} = 1 \text{ uniformly in } 0 \leq \delta \leq 1.$$

**Definition B.6.** A function  $\phi : \mathbb{R}_{>0} \rightarrow \mathbb{R}$  is called *self-controlled* if there exist positive constants  $\gamma, \Gamma, T$  such that  $\gamma\phi(t) \leq \phi(t + \delta\phi(t)) \leq \Gamma\phi(t)$ , for all  $t \geq T$  and  $\delta \in [0, 1]$ .

**Lemma B.7** ([BGT87, Thm. 3.12.5]). *Let  $h : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$  be such that  $\phi(t) := \log(h(e^t))$  is self-controlled and there exists  $T > 0$  such that  $1/\phi$  is locally integrable over  $[T, \infty)$ . Then  $L : \mathbb{R}_{>0} \rightarrow \mathbb{R}$  is super-slowly varying with rate function  $h$  if and only if  $L$  can be written as in (B.1) with  $\eta(t) = o(1/\log h(t))$  as  $t \rightarrow \infty$ .*

*Proof of Lemma B.4.* Let us use Lemma B.7. Since  $h$  is a slowly varying function and  $h(t) \geq 1 + h_0$  for all  $t$ , the function  $\phi(t) := \log(h(e^t))$  is such that  $1/\phi$  is locally integrable. Moreover we notice that  $\phi(t + \delta\phi(t)) = \log(h(e^t h(e^t)^\delta))$  and by using Potter's bound (B.3) for  $h$  we obtain for all  $\epsilon > 0$  that there exists  $T > 0$  such that for  $t > T$  we have

$$\frac{h(e^t h(e^t)^\delta)}{h(e^t)} \leq (1 + \epsilon) \max \left\{ h(e^t)^{\delta\epsilon}, h(e^t)^{-\delta\epsilon} \right\}.$$

Since  $h(t) \geq 1 + h_0$ , we have  $\phi(t) \geq \log(1 + h_0)$ , hence for all  $\epsilon \in (0, h_0)$

$$\phi(t + \delta\phi(t)) \leq \log(1 + \epsilon) + (1 + \delta\epsilon)\phi(t) \leq (2 + \delta\epsilon)\phi(t).$$

Analogously, we obtain  $\phi(t + \delta\phi(t)) \geq ((1 - \delta\epsilon) + \log(1 - \epsilon)/\log(1 + h_0))\phi(t)$ . If we choose  $\epsilon$  sufficiently small, then the factor before  $\phi(t)$  is bounded away from zero which implies that  $\phi$  is self-controlled.

Then by Lemma B.7, we obtain that if  $L$  is a slowly varying function written as in (B.1) with  $\eta(t) = o(1/\log h(t))$  as  $t \rightarrow \infty$ , it is super-slowly varying with rate function  $h$ . Let now  $\tilde{h}$  be a function satisfying  $c \leq \tilde{h}(n) \leq h(n)$  for all  $n$  and for a constant  $c > 0$ . Up to dividing by  $c$  and using that  $\eta(t) = o(1/\log(h(t)/c))$  as  $t \rightarrow \infty$ , we can set  $c = 1$ . Then we can write  $\tilde{h}(n) = h(n)^{\delta(n)}$  where  $\delta(n) \in [0, 1]$  for all  $n$ . By Definition B.5 it follows

$$\left| \frac{L(n\tilde{h}(n))}{L(n)} - 1 \right| = \left| \frac{L(n h(n)^{\delta(n)})}{L(n)} - 1 \right| \leq \sup_{\delta \in [0, 1]} \left| \frac{L(n h(n)^\delta)}{L(n)} - 1 \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

and the lemma is proved.  $\square$

## REFERENCES

- [A97] J. Aaronson, "An introduction to infinite ergodic theory", Mathematical Surveys and Monographs **50**, American Mathematical Society, Providence, RI, 1997.
- [AD90] J. Aaronson, M. Denker, *Upper bounds for ergodic sums of infinite measure preserving transformations*, Trans. Amer. Math. Soc. **319** (1990), no. 1, 101–138.
- [AKZ17] J. Aaronson, Z. Kosloff, B. Weiss, *Symmetric Birkhoff sums in infinite ergodic theory*, Ergodic Theory Dynam. Systems, **37** (2017), no. 8, 2394–2416.

- [AN03] J. Aaronson, H. Nakada, *Trimmed sums for non-negative, mixing stationary processes*, Stochastic Process. Appl. **104** (2003), no. 2, 173–192.
- [AN05] J. Aaronson, H. Nakada, *On the mixing coefficients of piecewise monotonic maps*, Israel J. Math. **148** (2005), 1–10.
- [ATZ05] J. Aaronson, M. Thaler, R. Zweimüller, *Occupation times of sets of infinite measure for ergodic transformations*, Ergodic Theory Dynam. Systems, **25** (2005), no. 4, 959–976.
- [BGT87] N. H. Bingham, C. M. Goldie, J. L. Teugels, “Regular variation”, Encyclopedia of Mathematics and its Applications **27**, Cambridge University Press, Cambridge, 1987.
- [BDM21] C. Bonanno, A. Del Vigna, S. Munday, *A slow triangle map with a segment of indifferent fixed points and a complete tree of rational pairs*, Monatsh. Math. **194** (2021), no. 1, 1–40.
- [BGL18] C. Bonanno, P. Giulietti, M. Lenci, *Infinite mixing for one-dimensional maps with an indifferent fixed point*, Nonlinearity **31** (2018), no. 11, 5180–5213.
- [BL21] C. Bonanno, M. Lenci, *Pomeau-Manneville maps are global-local mixing*, Discrete Contin. Dyn. Syst. **41** (2021), no. 3, 1051–1069.
- [B05] R. C. Bradley, *Basic properties of strong mixing conditions. A survey and some open questions*, Probab. Surv. **2** (2005), 107–144.
- [CN17] M. Carney, M. Nicol, *Dynamical Borel-Cantelli lemmas and rates of growth of Birkhoff sums of non-integrable observables on chaotic dynamical systems*, Nonlinearity, **30** (2017), no. 7, 2854–2870.
- [DV86] H. G. Diamond, J. D. Vaaler, *Estimates for partial sums of continued fraction partial quotients*, Pacific J. Math. **122** (1986), no. 1, 73–82.
- [GHPZ] S. Galatolo, M. Holland, T. Persson, Y. Zhang, *Anomalous time-scaling of extreme events in infinite systems and Birkhoff sums of infinite observables*, Discrete Contin. Dyn. Syst. **41** (2021), no. 4, 1799–1841.
- [G01] T. Garrity, *On periodic sequences for algebraic numbers*, J. Number Theory **88** (2001), no. 1, 86–103.
- [H93] E. Haeusler, *A nonstandard law of the iterated logarithm for trimmed sums*, Ann. Probab. **21** (1993), no. 2, 831–860.
- [HM87] E. Haeusler, 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.
- [H14] A. Haynes, *Quantitative ergodic theorems for weakly integrable functions*, Ergodic Theory Dynam. Systems **34** (2014), no. 2, 534–542.
- [I94] T. Inoue, *Ergodic theorems for piecewise affine Markov maps with indifferent fixed points*, Hiroshima Math. J. **24** (1994), no. 3, 447–471.
- [I00] T. Inoue, *Sojourn times in small neighborhoods of indifferent fixed points of one-dimensional dynamical systems*, Ergodic Theory Dynam. Systems **20** (2000), no. 1, 241–257.
- [I04] T. Inoue, *Ergodic sums of non-integrable functions under one-dimensional dynamical systems with indifferent fixed points*, Ergodic Theory Dynam. Systems **24** (2004), no. 2, 525–545.
- [K17] Z. Kosloff, *A universal divergence rate for symmetric Birkhoff sums in infinite ergodic theory*, Trans. Amer. Math. Soc. **369** (2017), 6373–6388.
- [KS19a] M. Kesseböhmer, T. I. Schindler, *Strong laws of large numbers for intermediately trimmed sums of i.i.d. random variables with infinite mean*, J. Theoret. Probab. **32** (2019), no. 2, 702–720.
- [KS19b] M. Kesseböhmer, T. I. 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. Corrigendum in Stochastic Process. Appl. **130** (2020), no. 11, 7019.
- [KS20a] M. Kesseböhmer, T. I. Schindler, *Intermediately trimmed strong laws for Birkhoff sums on subshifts of finite type*, Dyn. Syst. **35** (2020), no. 2, 275–305.
- [KS20b] M. Kesseböhmer, T. I. Schindler, *Mean convergence for intermediately trimmed Birkhoff sums of observables with regularly varying tails*, Nonlinearity **33** (2020), no. 10, 5543–5566.
- [KM92] H. Kesten, R. A. Maller, *Ratios of trimmed sums and order statistics*, Ann. Probab. **20** (1992), no. 4, 1805–1842.
- [LM18] M. Lenci, S. Munday, *Pointwise convergence of Birkhoff averages for global observables*, Chaos **28** (2018), no. 8, 083111, 16 pp.
- [M06] F. Móricz, *On the harmonic averages of numerical sequences*, Arch. Math. (Basel) **86** (2006), no. 4, 375–384.
- [NN03] H. Nakada, R. Natsui, *On the metrical theory of continued fraction mixing fibred systems and its application to Jacobi-Perron algorithm*, Monatsh. Math. **138** (2003), no. 4, 267–288.
- [S18] T. I. Schindler, *Trimmed sums for observables on the doubling map*, arXiv:1810.03223 [math.DS]
- [S95] F. Schweiger, “Ergodic theory of fibred systems and metric number theory”, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1995.



- [T83] M. Thaler, *Transformations on  $[0, 1]$  with infinite invariant measures*, Israel J. Math. **46** (1983), no. 1-2, 67–96.
- [Z00] R. Zweimüller, *Ergodic properties of infinite measure-preserving interval maps with indifferent fixed points*, Ergodic Theory Dynam. Systems **20** (2000), no. 5, 1519–1549.
- [Z09] R. Zweimüller, “Surrey notes on infinite ergodic theory”,  
<http://mat.univie.ac.at/%7Ezweimueller/MyPub/SurreyNotes.pdf>

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI PISA, LARGO BRUNO PONTECORVO 5, 56127 PISA, ITALY  
*Email address:* `claudio.bonanno@unipi.it`

CENTRO DI RICERCA ENNIO DE GIORGI, SCUOLA NORMALE SUPERIORE, PIAZZA DEI CAVALIERI 3, 56126 PISA, ITALY  
*Email address:* `tanja.schindler@sns.it`