STUDIA MATHEMATICA 220 (2) (2014)

Error rates in the Darling–Kac law

by

DALIA TERHESIU (Roma)

Abstract. This work provides rates of convergence in the Darling–Kac law for infinite measure preserving Pomeau–Manneville (unit interval) maps. Along the way we obtain error rates for the stable law associated with the first return map and the first return time to some suitable set inside the unit interval.

1. Introduction and main results

1.1. Darling–Kac limit laws for dynamical systems preserving an infinite measure. To understand a chaotic dynamical system, methods from probability theory are an important tool. This goes back to Birkhoff's ergodic theorem, which states that for a dynamical system $f : X \to X$ that preserves a probability measure μ, the ergodic average

$$\frac{1}{n}S_n(v) = \frac{1}{n}\sum_{k=0}^{n-1} v \circ f^k$$

converges almost everywhere (a.e.) to the space average $\int v\, d\mu$, for all integrable functions v ($v \in L^1$). In contrast, if $\mu(X) = \infty$, Birkhoff's ergodic theorem is not very informative, since in this case $n^{-1}S_n$ goes to 0 a.e., for all $v \in L^1$. Even stronger, as proved in [1], the ergodic theorem cannot be recovered by rescaling. More precisely, for any positive sequence c_n and for any $v \in L^1$, either $c_n^{-1}S_n$ goes to 0 a.e. or it goes to ∞ along subsequences. However, in certain cases there exists a positive sequence a_n such that for all $v \in L^1$, $a_n^{-1}S_n$ converges in a weaker sense, namely in distribution, to a non-trivial limit (see for instance [1, 19, 3] and the plethora of references therein). Such a limit law is referred to as the Darling–Kac (DK) theorem, and usually when this applies, one can prove the existence of other interesting limit laws, such as arc-sine laws [17, 18, 19, 22].

2010 *Mathematics Subject Classification*: 37A30, 37A50, 37E05, 60Fxx.
Key words and phrases: convergence rates, Darling–Kac law, stable limit laws, infinite ergodic theory.

DOI: 10.4064/sm220-2-1

As shown in [1, Theorems 3.6.4, 3.7.2], pointwise dual ergodicity together with regular variation of the return sequence guarantee the existence of the DK law. We recall these notions below.

Pointwise dual ergodicity provides information about the asymptotic behavior of the transfer operator $L : L^1(X) \to L^1(X)$ associated with $(X, \mathcal{A}, f, \mu)$, defined by $\int_Y Lv\, w\, d\mu = \int_Y v\, w \circ f\, d\mu$, $w \in L^\infty(Y)$. More precisely, f is *pointwise dual ergodic* if there exists a positive sequence a_n such that $a_n^{-1} \sum_{j=0}^{n-1} L^j v \to \int_X v\, d\mu$ a.e. for all $v \in L^1$. The sequence a_n is referred to as the *return sequence* for f (see [1] for a definition of a_n in terms of the weaker property of rational ergodicity). A necessary and sufficient condition for f to be pointwise dual ergodic is the existence of sets $Y \in \mathcal{A}$ with $0 < \mu(Y) < \infty$ such that for $v \in L^1$ and $a_n(Y) := \mu(Y)^{-2} \sum_{j=0}^{n-1} \mu(Y \cup f^{-j} Y)$, one has $a_n(Y)^{-1} \sum_{j=0}^{n-1} L^j v \to \int_X v\, d\mu$ uniformly on Y (see [1]). The return sequence $a_n(f)$ of f is determined up to a multiplicative constant (corresponding to an arbitrary scaling of the measure μ) and asymptotic equivalence satisfies $a_n(f) = a_n(Y)(1 + o(1))$. In the following, we will choose a suitable set Y (in accordance with the inducing method below), scale μ so that $\mu(Y) = 1$, and fix $a_n := a_n(Y)$ for this choice.

While the existence of a DK law for (X, f, μ) does not require the strong property of pointwise dual ergodicity (see [3]), it does require that the return sequence a_n is regularly varying, i.e. that $a_n = \ell(n) n^\beta$ for some slowly varying function ℓ and some index $\beta \in [0, 1]$). Regular variation is an important assumption of the Darling–Kac theory for Markov chains (see, for instance, [4]). For a pointwise dual ergodic (X, f, μ) with $a_n = \ell(n) n^\beta$ the Darling–Kac law says that for all $v \in L^1$,

$$C a_n^{-1} S_n(v) \to_d \mathcal{Y}_\beta \quad \text{as } n \to \infty.$$

where a_n is as defined above, C is a positive constant that depends only on f, and $\mathcal{Y}_\beta$ is a positive random variable distributed according to the normalized Mittag-Leffler distribution of order β, that is, $E(e^{z\mathcal{Y}_\beta}) = \sum_{p=0}^{\infty} \Gamma(1+\beta)^p z^p / \Gamma(1 + p\beta)$ for all $z \in \mathbb{C}$.

A standard way of verifying regular variation for a_n associated with a dynamical system (X, f) is by inducing with respect to the first return time to some 'good' set $Y \subset X$. To simplify notation, fix $Y \subset X$ with $\mu(Y) = 1$. Let $\varphi : Y \to \mathbb{Z}^+$ be the first return time to Y defined by $\varphi(y) = \inf\{n \geq 1 : f^n y \in Y\}$. If $\mu(\varphi > n) = \ell(n) n^{-\beta}$ for some slowly varying function ℓ and some index $\beta \in [0, 1]$ then $a_n(Y) = \ell(n) n^\beta$ for $\beta \in [0, 1)$, $a_n(Y) = n \sum_{j=1}^{n} \ell(j) j^{-1}$ for $\beta = 1$ (see [1, Section 3.8]).

1.2. A classical example. A standard example of a dynamical system with infinite measure that has the desired properties (pointwise dual ergodicity along with regular variation) is given by the family of Pomeau–Manneville intermittency maps [14]. These are interval maps with indifferent fixed points; that is, they are uniformly expanding except for an indifferent fixed point at 0. To fix notation, we recall the version considered in [10]:

$$f(x) = \begin{cases} x(1+2^\alpha x^\alpha), & 0 < x < 1/2, \\ 2x-1, & 1/2 < x < 1. \end{cases} \tag{1.1}$$

For $\alpha \geq 1$, we are in the situation of infinite ergodic theory; there exist a unique (up to scaling) σ-finite, absolutely continuous invariant measure μ. In the setting of (1.1), we let $x_0 = 1/2$ and $x_{p+1} < x_p = f(x_{p+1})$ for each $p \geq 0$, and then set $Y = [x_p, 1]$ for some arbitrary $p \geq 0$. Note that one can rescale μ such that $\mu(Y) = 1$ and recall that $\mu(\varphi = n) = O(n^{-(\beta+1)})$ with $\beta = 1/\alpha$.

The methods employed so far [1, 18, 19] to establish limit theorems for dynamical systems with infinite measure do not allow one to determine the error rate present in the convergence involved. Recent progress in this sense has been made in [11, 16], which establish sharp error rates in arcsine laws associated with systems such as (1.1). The results in [11, 16] are established by exploring a 'good' expansion of the tail distribution $\mu(\varphi > n)$. For higher order expansion of $\mu(\varphi > n)$ in the special case of (1.1), we refer to [11, 12, 16].

Our aim in this work is to establish error rates in the Darling–Kac law associated with systems such as (1.1). In the rest of the paper we say that (f, μ), Y and $a_n := a_n(Y)$ are defined by (1.1) in the following sense:

(i) f is the map defined by (1.1);
(ii) $Y = [x_p, 1] \subset (0, 1]$, where x_p, $p \geq 0$, is as defined in the paragraph following (1.1) (by taking p sufficiently large, we will be able to deal with observables v that are supported on a compact subset of $(0, 1]$);
(iii) the f-invariant measure μ is rescaled such that $\mu(Y) = 1$;
(iv) set $a_n(Y) = \sum_{j=0}^{n-1} \mu(Y \cap f^{-n}Y)$ (a representative of the return sequence for f).

1.3. Main results. Our main result reads as follows

THEOREM 1.1 (Error rates in the DK law associated with (1.1)). *Let (f, μ), Y and $a_n := a_n(Y)$ be as in Section* 1.2. *Suppose that the function $v : [0, 1] \to \mathbb{R}$ can be written as $v = 1_Y - \tilde{v}$, a.e. on Y, where $\int \tilde{v}\, d\mu = 0$ and $\mu_Y\big(|a_n^{-1} S_n(\tilde{v})| > g(n)\big) < g(n)$, where g is a positive decreasing function such that $g(n) = O(n^{-\beta})$. Then for any $z > 0$,*

$$|\mu_Y(a_n^{-1} S_n(v) > z) - \mathbb{P}(\mathcal{Y}_\beta > z)| = E(n),$$

where

$$E(n) = \begin{cases} O(n^{\beta-1}) & \text{if } \beta \in (1/2, 1), \\ O((\log n)^2 n^{-1/2}) & \text{if } \beta = 1/2, \\ O((\log n) n^{-\beta}) & \text{if } \beta \in (0, 1/2). \end{cases}$$

We are not aware of any result on error rates in the DK theorem associated with null recurrent Markov chains characterized by regular variation. We claim that the error rates in Theorem 1.1 are *optimal*. As we explain in what follows, the proof of Theorem 1.1 for the function 1_Y is obtained via Lemma 1.2 below, which provides *optimal* error rates for the stable law associated with the induced map f_Y and observable φ.

On the *negative* side, we acknowledge that the assumption on the zero mean function $\tilde{v}$ (and thus, v) in the statement of Theorem 1.1 is very strong. Recent work of Thomine [20] suggests that general zero mean functions $\tilde{v}$ such that $\sum_{j=0}^{\varphi-1} |\tilde{v}| \circ f^j$ belongs to $L^p(Y, \mu)$ for some $p > 2$ are not in the restrictive class of functions considered in the statement of Theorem 1.1 (see the explanatory Remark 3.1). Hence, finding a reasonably large class of functions v that yields the conclusion of Theorem 1.1 is open.

Theorem 1.1 is proved in Section 2. For a version of Theorem 1.1 for more general dynamical systems satisfying the abstract assumptions of Section 4 we refer to Lemma 5.2.

We recall that in the case of (1.1), regular variation of $\mu(\varphi > n)$ implies a stable law for the induced map $f_Y := f^\varphi$ (this follows from [2]). More precisely, let $\varphi_n = \sum_{j=0}^{n-1} \varphi \circ f_Y^j$ and assume that the sequence b_n is an asymptotic inverse of the sequence $a_n := a_n(Y) = \sum_{j=0}^{n-1} \mu(Y \cap f^{-n} Y)$ (that is, if the corresponding functions $t \mapsto a_{[t]}, t \mapsto b_{[t]}$ satisfy $a(b(n)) = n(1 + o(1))$ and $b(a(n)) = n(1 + o(1))$. Then $b_n^{-1} \varphi_n \to_d \mathcal{Z}_\beta$, where $\mathcal{Z}_\beta =_d (\mathcal{Y}_\beta)^{-1/\beta}$ and $\mathcal{Y}_\beta$ is a positive random variable distributed according to the normalized Mittag-Leffler distribution of order β (see Section 1). Hence, the real Laplace transform of $\mathcal{Z}_\beta$ is given by $E(e^{-t\mathcal{Z}_\beta}) = e^{-t^\beta}$. Alternatively, the variable $\mathcal{Z}_\beta$ can be defined in terms of its known characteristic function. For details we refer to [2]; see also Section 5 below.

Our next result provides error rates for the stable law associated with the map f_Y and observable φ. The proof is deferred to Section 5. In the present context it serves as the key result: Theorem 1.1 for the case $v = 1_Y$ can be deduced from it by standard computations (used in Proposition 2.3 and its proof).

LEMMA 1.2 (Error rates for the stable law associated with f_Y and φ). *Let (f, μ) and Y be as in Section* 1.2. *Assume $\beta \in (0, 1)$. Let φ be the first return time function to Y. Set $b_n = (n/C_0)^{1/\beta}$, where C_0 is the constant* *defined in Lemma* 2.1. *Then for any* $a > 0$,

$$|\mu_Y(b_n^{-1}\varphi_n < a) - \mathbb{P}(\mathcal{Z}_\beta < a)| = d(n),$$

where

$$d(n) = \begin{cases} O(n^{1-1/\beta}) & \textit{if } \beta \in (1/2,1), \\ O((\log n)/n) & \textit{if } \beta = 1/2, \\ O(1/n) & \textit{if } \beta \in (0,1/2). \end{cases}$$

Remark 1.3. Lemma 1.2 matches the *optimal* results on rates of convergence to a stable law of index $\beta \in (0,1)$ for sequences of independent random variables in [9]. More generally, we refer to [5, 9, 15, 21] for rates of convergence to a stable law of index $\beta \in (0,2)$ for sequences of independent random variables.

The paper is organized as follows. In Section 2, we prove Theorem 1.1 using Lemma 1.2 and some results in [11], which we recall below.

Section 5 is devoted to the proof of Lemma 1.2 in the more general setting of Section 4.

Notation. We use "big O" and $\ll$ notation interchangeably, writing $c_n = O(d_n)$ or $c_n \ll d_n$ if there is a constant $C > 0$ such that $c_n \le C d_n$ for all $n \ge 1$. We also write $\mu_Y(\,\cdot\,)$ for $\mu(x \in Y :\,\cdot\,)$.

2. Results for the function 1_Y.

Given the existence of a stable law for (f_Y, φ), it seems natural that Theorem 1.1 for the special case $v = 1_Y$ will follow from Lemma 1.2 together with the duality rule $\mu(S_m(1_Y) > n) = \mu(\varphi_n < m)$ (see Proposition 2.3 below).

Precise information on $a_n(Y) = \sum_{j=0}^{n-1} \mu(Y \cap f^{-n}Y)$ follows from the asymptotic behavior of the transfer operator $L : L^1(\mu) \to L^1(\mu)$ associated with f. Higher order asymptotics of L^n and $\sum_{j=0}^{n-1} L^j$ have been obtained in [11, 12]. For the present purpose, we recall

Lemma 2.1 ([12, Theorem 1.5]). *Let f be defined by* (1.1) *with* $\beta \in (0,1)$. *Suppose that $v : [0,1] \to \mathbb{R}$ is Hölder or of bounded variation supported on a compact subset of* $(0,1]$. *Set* $k = \max\{j \ge 0 : (j+1)\beta - j > 0\}$. *Let* $\tau = 1$ *for* $\beta \ne 1/2$ *and* $\tau = 2$ *for* $\beta = 1/2$. *Then*

$$\sum_{j=0}^{n-1} L^j v = (C_0 n^\beta + C_1 n^{2\beta-1} + C_2 n^{3\beta-2} + \cdots + C_k n^{(k+1)\beta-k}) \int_0^1 v\,d\mu + O(\log^\tau n),$$

uniformly on compact subsets of $(0,1]$, *where* $C_0 = (c\Gamma(1-\beta)\Gamma(1+\beta))^{-1}$ *with c a positive constant depending only on f, and $C_1, C_2, \ldots$ are real constants* (*depending only on* f).

An immediate consequence of the above result is

COROLLARY 2.2. *Suppose that* (f,μ), Y *and* $a_n := a_n(Y)$ *are as in Section* 1.2. *Let* $C_0, C_1, \dots$ *and* C *be the real constants defined in Lemma* 2.1. *If* $\beta \in (0,1)$, *then*

$$a_n = (C_0 n^{\beta} + C_1 n^{2\beta-1} + C_2 n^{3\beta-2} + \cdots + C_k n^{(k+1)\beta-k}) + O(\log^{\tau} n).$$

The following result will be instrumental in the proof of Theorem 1.1.

PROPOSITION 2.3. *Assume the setting of Lemma* 1.2 *with* $\beta \in (0,1)$. *Let* $a_m := a_m(Y)$ *be as in Section* 1.2. *Then for any* $z > 0$,

$$|\mu_Y(a_m^{-1} S_m(1_Y) > z) - \mathbb{P}(\mathcal{Y}_\beta > z)| = e(m),$$

where

$$e(m) = \begin{cases} O(m^{\beta-1}) & \text{if } \beta \in (1/2,1), \\ O((\log m)^2 m^{-1/2}) & \text{if } \beta = 1/2, \\ O((\log m) m^{-\beta}) & \text{if } \beta \in (0,1/2). \end{cases}$$

Proof. By the triangle inequality,

$$|\mu_Y(a_m^{-1} S_m(1_Y) > z) - \mathbb{P}(\mathcal{Y}_\beta > z)| \le I + II \tag{2.1}$$

for

$$I = \left|\mu_Y(S_m(1_Y) > z a_m) - \mathbb{P}\left(\mathcal{Y}_\beta > \frac{[z a_m]}{C_0 m^\beta}\right)\right|,$$

$$II = \left|\mathbb{P}(\mathcal{Y}_\beta > z) - \mathbb{P}\left(\mathcal{Y}_\beta > \frac{[z a_m]}{C_0 m^\beta}\right)\right|.$$

We start with I. Let $b_m = (m/C_0)^{1/\beta}$ as in Lemma 1.2. Since

$$\mu_Y(S_m(1_Y) > z a_m) = \mu_Y(S_m(1_Y) > [z a_m]) = \mu_Y(\varphi_{[z a_m]} < m)$$

and $\mathcal{Z}_\beta =_d (\mathcal{Y}_\beta)^{-1/\beta}$, we have

$$\begin{aligned} I &= \mu_Y(S_m(1_Y) > z a_m) - \mathbb{P}\left(\mathcal{Y}_\beta > \frac{[z a_m]}{m^\beta C_0}\right) \\ &= \mu_Y\left(\frac{\varphi_{[z a_m]}}{b_{[z a_m]}} < \frac{m}{b_{[y a_m]}}\right) - \mathbb{P}\left(\mathcal{Z}_\beta < \frac{m C_0^{1/\beta}}{[z a_m]^{1/\beta}}\right) \\ &= \mu_Y\left(\frac{\varphi_{[y a_m]}}{b_{[z a_m]}} < \frac{m C_0^{1/\beta}}{[z a_m]^{1/\beta}}\right) - \mathbb{P}\left(\mathcal{Z}_\beta < \frac{m C_0^{1/\beta}}{[z a_m]^{1/\beta}}\right), \end{aligned}$$

where for the last equality we used $b_{[z a_m]} = [z a_m]^{1/\beta}/C_0^{1/\beta}$. Applying Lemma 1.2 with $n = [z a_m]$ and $a = C_0^{1/\beta}/[zc(m)]^{1/\beta}$, we obtain

$$I = \left|\mu\left(\frac{\varphi_{[y a_m]}}{b_{[z a_m]}} < \frac{C_0^{1/\beta}}{[z a_m]^{1/\beta}}\right) - \mathbb{P}\left(\mathcal{Z}_\beta < \frac{C_0^{1/\beta}}{[z a_m]^{1/\beta}}\right)\right| =: e_I(m),$$

where

$$e_I(m) = d([za_m]) = \begin{cases} O(m^{\beta-1}) & \text{if } \beta \in (1/2, 1), \\ O((\log m)m^{-1/2}) & \text{if } \beta = 1/2, \\ O(m^{-\beta}) & \text{if } \beta \in (0, 1/2). \end{cases}$$

We continue with II from (2.1). Since $\mathcal{Z}_\beta =_d (\mathcal{Y}_\beta)^{-1/\beta}$, we have

$$\begin{aligned} II &= \left| \mathbb{P}(\mathcal{Y}_\beta > z) - \mathbb{P}\left(\mathcal{Y}_\beta > \frac{[za_m]}{C_0 m^\beta} \right) \right| \\ &= \left| \mathbb{P}\left(\mathcal{Z}_\beta < \frac{1}{z^{1/\beta}} \right) - \mathbb{P}\left(\mathcal{Z}_\beta < \frac{m C_0^{1/\beta}}{[za_m]^{1/\beta}} \right) \right|. \end{aligned} \tag{2.2}$$

It is known (see for instance [13]) that for every $\epsilon > 0$ there exists $C > 0$ such that for all $a, b > 0$ with $|a - b| < \epsilon$, we have

$$|\mathbb{P}(\mathcal{Z}_\beta < a^{1/\beta}) - \mathbb{P}(\mathcal{Z}_\beta < b^{1/\beta})| \le C|a^{-1} - b^{-1}|. \tag{2.3}$$

This fact together with (2.2) implies that

$$\left| \mathbb{P}\left(\mathcal{Z}_\beta < \frac{1}{z^{1/\beta}} \right) - \mathbb{P}\left(\mathcal{Z}_\beta < \frac{m C_0^{1/\beta}}{[za_m]^{1/\beta}} \right) \right| \le C \left| z - \frac{[za_m]}{C_0 m^\beta} \right|.$$

Corollary 2.2 gives $a_m = C_0 m^\beta + O((\log m)^\tau) + O(m^{2\beta-1})$, so we get

$$\begin{aligned} \left| z - \frac{[za_m]}{C_0 m^\beta} \right| &\le \left| z - \frac{za_m}{C_0 m^\beta} \right| + \frac{1}{C_0 m^\beta} \\ &= O(m^{\beta-1}) + O\left(\frac{(\log m)^\tau}{m^\beta} \right) + \frac{1}{C_0 m^\beta} =: e_{II}(m), \end{aligned}$$

satisfying

$$e_{II}(m) = \begin{cases} O(m^{\beta-1}) & \text{if } \beta \in (1/2, 1), \\ O((\log m)^2 m^{-1/2}) & \text{if } \beta = 1/2, \\ O((\log m)m^{-\beta}) & \text{if } \beta \in (0, 1/2). \end{cases}$$

Combining the estimates, we find $e(m) = e_I(m) + e_{II}(m)$ of the required form. ∎

3. Proof of Theorem 1.1. Recall that $v : [0, 1] \to \mathbb{R}$ on Y can be written as $v = 1_Y - \tilde{v}$, where (i) $\int \tilde{v}\, d\mu = 0$ and (ii) $\mu_Y(|a_m^{-1} S_m \tilde{v}| > g(m)) < g(m)$, where g is a positive decreasing function such that $g(m) = O(m^{-\beta})$.

Note that $S_m(v) = S_m(1_Y) + S_m(\tilde{v})$ a.e. on Y. Since Proposition 2.3 gives the desired estimate for 1_Y, to conclude we need to estimate $|\mu_Y(a_m^{-1} S_m(v) > z) - \mu_Y(a_m^{-1} S_m(1_Y) > z)|$ for $z > 0$.

Remark 3.1. We note that the assumption (ii) on the function $\tilde{v}$ is very strong. Suppose that $\tilde{v} : [0, 1] \to \mathbb{R}$ is a mean zero function such that

$\sum_{j=0}^{\varphi-1} |\tilde{v}| \circ f^j$ belongs to $L^p(Y,\mu)$ for some $p>2$. As shown in the proof of [20, Theorem 4.7], the following holds a.e. on Y:

$$\Big|\sum_{j=0}^{m-1} \tilde{v} \circ f^j(x)\Big| \le C(x) m^{\beta/2+\epsilon}$$

for some $C(x)>0$, for any $\epsilon>0$ and all m sufficiently large. Assuming that $\int C(x)\,d\mu_Y < \infty$, the above inequality implies that

$$\int \Big|\sum_{j=0}^{m-1} \tilde{v} \circ f^j\Big|\, d\mu_Y \ll m^{\beta/2+\epsilon}.$$

Together with Markov's inequality, the above displayed estimate implies that given some function h and some positive constant C such that $h(m) > C$ and $h(m) = O(m^{-(\beta-\epsilon)})$, for any $\epsilon>0$, we have

$$\mu_Y\left(\left|\frac{S_m\tilde{v}}{a_m}\right| > h(m)\right) \le C^{-1} a_m^{-1} \int |S_m \tilde{v}|\, d\mu \ll a_m^{-1} m^{\beta/2+\epsilon} \ll h(m)^{1/2}.$$

The above inequality together with the argument used in the proof of Theorem 1.1 below (with $g=h$) shows that

$$\left|\mu_Y\left(\frac{S_m(v)}{a_m} > z\right) - \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z\right)\right| \ll h(m)^{1/2}.$$

This inequality together with Proposition 2.3 implies that

$$|\mu_Y(a_m^{-1} S_m(v) > z) - \mathbb{P}(\mathcal{Y}_\beta > z)| = E(m),$$

where $E(m) \ll m^{-(\beta/2-\epsilon)}$, which is a much weaker form of Theorem 1.1.

Proof of Theorem 1.1. Let g be a function as defined above. We claim that

$$\begin{aligned}&\left|\mu_Y\left(\frac{S_m(v)}{a_m} > z\right) - \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z\right)\right| \\ &\quad \le \left(\mu_Y\left(\frac{S_m(1_y)}{a_m} > z - g(m)\right) - \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z + g(m)\right)\right) + g(m)^{1/2}.\end{aligned}$$

By the triangle inequality,

$$\begin{aligned}&\left|\mu_Y\left(\frac{S_m(1_Y)}{a_m} > z - g(m)\right) - \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z + g(m)\right)\right| \\ &\quad \le \left|\mu_Y\left(\frac{S_m(1_Y)}{a_m} > z - g(m)\right) - \mathbb{P}(\mathcal{Y}_\beta > z - g(m))\right| \\ &\qquad + |\mathbb{P}(\mathcal{Y}_\beta > z - g(m))| - \mathbb{P}(\mathcal{Y}_\beta > z + g(m))| \\ &\qquad + \left|\mu_Y\left(\frac{S_m(1_Y)}{a_m} > z + g(m)\right) - \mathbb{P}(\mathcal{Y}_\beta > z + g(m))\right|.\end{aligned} \tag{3.1}$$

The first and third terms in (3.1) can be estimated using Proposition 2.3, but we should be aware that the function $e(m)$ from that proposition depends on z. Indicating this dependence as a subscript, we can estimate the two terms by $e_{z-g(m)}(m) + e_{z+g(m)}(m)$. Following the estimates of Proposition 2.3, we can see that $e_z(m)$ can be chosen to be decreasing in z, so $e_{z-g(m)}(m)+e_{z+g(m)}(m) \le e_{z/2}(m)$, which satisfies the estimate in the statement of Proposition 2.3 with $z/2$ instead of z.

Recall $\mathcal{Z}_\beta =_d (\mathcal{Y}_\beta)^{-1/\beta}$. Using (2.3), we can estimate the middle term of (3.1) as

$$|\mathbb{P}(\mathcal{Y}_\beta > z - g(m))| - \mathbb{P}(\mathcal{Y}_\beta > z + g(m))| \ll g(m).$$

Combining these estimates gives

$$\left|\mu_Y\left(\frac{S_m(v)}{a_m} > z\right) - \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z\right)\right| \ll g(m),$$

and the conclusion follows since $g(m) = O(m^{-\beta})$.

It remains to prove the claim. We have

$$\begin{aligned}
&\mu_Y\left(\frac{S_m(v)}{a_m} > z\right) - \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z\right)\\
&\quad\le \mu_Y\left(\frac{S_m(v)}{a_m} > z \wedge \frac{S_m(\tilde v)}{a_m} \ge g(m)\right) - \mu_Y\left(\frac{S_m(1_y)}{a_m} > z\right)\\
&\qquad\qquad + \mu_Y\left(\frac{S_m(\tilde v)}{a_m} > g(m)\right)\\
&\quad\le \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z - g(m)\right) - \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z\right) + \mu_Y\left(\frac{S_m(\tilde v)}{a_m} > g(m)\right)\\
&\quad\le \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z - g(m)\right) - \mu_Y\left(\frac{S_m(1_Y)}{a_m} > z + g(m)\right)\\
&\qquad\qquad + \mu_Y\left(\frac{S_m(\tilde v)}{a_m} > g(m)\right)
\end{aligned}$$

and

$$\begin{aligned}
&\mu_Y\left(\frac{S_m(v)}{a_m} > z\right) - \mu_Y\left(\frac{S_m(1_y)}{a_m} > z\right)\\
&\quad\ge \mu_Y\left(\frac{S_m(v)}{a_m} > z \wedge \frac{S_m(\tilde v)}{a_m} \ge -g(m)\right) - \mu_Y\left(\frac{S_m(1_y)}{a_m} > z\right)\\
&\quad\ge \mu_Y\left(\frac{S_m(v)}{a_m} > z + g(m) \wedge \frac{S_m(\tilde v)}{a_m} \ge -g(m)\right) - \mu_Y\left(\frac{S_m(1_y)}{a_m} > z\right)\\
&\quad\ge -\left(\mu_Y\left(\frac{S_m(1_y)}{a_m} > z - g(m)\right) - \mu_Y\left(\frac{S_m(1_y)}{a_m} > z + g(m)\right)\right.\\
&\qquad\qquad \left. + \mu_Y\left(\frac{S_m(\tilde v)}{a_m} < -g(m)\right)\right).
\end{aligned}$$

Recall that $\mu_Y(|(S_m\tilde{v})/a_m| > g(m)) < g(m)$. This fact together with the previous two estimates implies that

$$\begin{aligned}
&\left|\mu_Y\left(\frac{S_m(v)}{a_m} > z\right) - \mu_Y\left(\frac{S_m(1_y)}{a_m} > z\right)\right| \\
&\quad \le \mu_Y\left(\frac{S_m(1_y)}{a_m} > z - g(m)\right) - \mu_Y\left(\frac{S_m(1_y)}{a_m} > z + g(m)\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad + \mu_Y\left(\left|\frac{S_m(\tilde{v})}{a_m}\right| > g(m)\right) \\
&\quad \le \left(\mu_Y\left(\frac{S_m(1_y)}{a_m} > z - g(m)\right) - \mu_Y\left(\frac{S_m(1_y)}{a_m} > z + g(m)\right)\right) + g(m),
\end{aligned}$$

which ends the proof of the claim. ∎

4. Abstract setting. Let (X, μ) be an infinite measure space, and $f : X \to X$ a conservative measure preserving map. Fix $Y \subset X$ with $\mu(Y) = 1$. Let $\varphi : Y \to \mathbb{Z}^+$ be the first return time $\varphi(y) = \inf\{n \ge 1 : f^n y \in Y\}$ and define the first return map $F = f^\varphi : Y \to Y$.

The return time function $\varphi : Y \to \mathbb{Z}^+$ satisfies $\int_Y \varphi\, d\mu = \infty$. Throughout we let $\beta \in (0,1)$ and assume

- (H) $\mu(\varphi > n) = c(n^{-\beta} + H(n))$, where $c > 0$ and $H(n) = O(n^{-q})$ for some $q > \beta$. If $q \le 1$, we assume further that $H(n) = m(n) + \tilde{m}(n)$, where m is monotone with $m(n) = O(n^{-q})$ and $\tilde{m}(n)$ is summable.

Recall that the transfer operator $R : L^1(Y) \to L^1(Y)$ for the first return map f_Y is defined via the formula $\int_Y Rv\, w\, d\mu = \int_Y v\, w \circ F\, d\mu$, $w \in L^\infty(Y)$. Let $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ and $\bar{\mathbb{D}} = \{z \in \mathbb{C} : |z| \le 1\}$. Given $z \in \bar{\mathbb{D}}$, define the perturbed operator $R(z) : L^1(Y) \to L^1(Y)$ by $R(z)v = R(z^\varphi v)$.

Also, for each $n \ge 1$, we define $R_n : L^1(Y) \to L^1(Y)$ by

$$R_n v = 1_Y R(1_{\{\varphi = n\}} v) = R(1_{\{\varphi = n\}} v).$$

It is easily verified that $R(z) = \sum_{n=1}^\infty R_n z^n$. We assume that there is a function space $\mathcal{B} \subset L^\infty(Y)$ containing the constant functions, with norm $\|\ \|$ satisfying $|v|_\infty \le \|v\|$ for $v \in \mathcal{B}$, such that

- (H1) There is a constant $C > 0$ such that $\|R_n\| \le C\mu(\varphi = n)$ for all $n \ge 1$.

It follows that $z \mapsto R(z)$ is an analytic family of bounded linear operators on $\mathcal{B}$ for $z \in \mathbb{D}$, and that this family extends continuously to $\bar{\mathbb{D}}$. Since $R(1) = R$ and $\mathcal{B}$ contains the constant functions, 1 is an eigenvalue of $R(1)$. Throughout, we assume:

- (H2) The eigenvalue 1 is simple and isolated in the spectrum of $R(1)$, and the spectrum of $R(z)$ does not contain 1 for all $z \in \bar{\mathbb{D}}$.

By (H1) and (H2), there exists $\epsilon > 0$ and a continuous family $\lambda(z)$ of simple eigenvalues of $R(z)$ for $z \in \bar{\mathbb{D}} \cap B_\epsilon(1)$ with $\lambda(1) = 1$. In what follows, we let $\lambda(\theta) := \lambda(z)$ for $z = e^{i\theta}$, $\theta \in [0, 2\pi)$.

As shown in [11, 12], the main assumptions above are enough for higher order expansion of $\lambda(z)$, $z \in \bar{\mathbb{D}} \cap B_\epsilon(1)$.

LEMMA 4.1 ([12, Lemma A.4], [11, Lemma 3.2]). *Assume* (H), (H1) *and* (H2).

If $q > 1$, set $c_H = -\Gamma(1-\beta)^{-1} \int_0^\infty H_1(x)\,dx$ where $H_1(x) = [x]^{-\beta} - x^{-\beta} + H([x])$. If $q \le 1$, set $c_H = 0$. Define $c_\beta = -i \int_0^\infty e^{i\sigma} \sigma^{-\beta}\,d\sigma$. Then as $\theta \to 0$,

$$\lambda(\theta) = 1 - cc_\beta \theta^\beta + icc_H \theta + O(\theta^{2\beta}) + D(\theta),$$

where $D(\theta) = O(\theta^q)$ if $q \ne 1$, and $D(\theta) = O\big(\theta \log \frac{1}{\theta}\big)$ if $q = 1$.

Proof. The case $q > 1$ is contained in the proof of [11, Lemma 3.2]. For $q < 1$, the argument for the exact term $1 - cc_\beta\theta^\beta$ in the expression of $\lambda(\theta)$ is again contained in the proof of [11, Lemma 3.2]. The estimate for $D(\theta)$ follows by the argument used in the proof of [12, Lemma A.4] (in estimating $D(z)$ there, with $z = e^{-u+i\theta}$ in the case $0 < u < \theta$). ∎

LEMMA 4.2 ([12, Theorem 4.1]). *Assume* (H1) *and* (H2). *Suppose that* (H) *holds with $q = 2\beta$. Let $k = \max\{j \ge 0 : (j+1)\beta - j > 0\}$. Let $L : L^1(X) \to L^1(X)$ be the transfer operator for f. Then for all $v \in \mathcal{B}$, there exist positive constants $C_0, \dots, C_k$ (depending only on f) such that*

$$\sum_{j=0}^{n-1} 1_Y L^j v = (C_0 n^\beta + C_1 n^{2\beta-1} + C_2 n^{3\beta-2} + \dots + C_k n^{(k+1)\beta-k}) \int_Y v\,d\mu + E_n v,$$

where $|E_n v|_\infty \le C(\log^\tau n)|v|_\infty$, C constant, and $\tau = 1$ for $\beta \ne 1/2$, $\tau = 2$ for $\beta = 1/2$.

The exact expression of the constants $C_0, \dots, C_k$ is provided in [12] and for later use we recall that $C_0 = (c\Gamma(1-\beta)\Gamma(1+\beta))^{-1}$.

5. Results for the abstract setting. In this section we provide a more general version of Lemma 1.2 and formulate a version of Theorem 1.1 for systems that satisfy (H), (H1) and (H2).

Throughout this section we use the notation

$$a_n(Y) := \sum_{j=1}^{n} \mu(Y \cup f^{-j} Y).$$

We assume that (H) holds. Lemma 4.2 gives

$$a_n(Y) = C_0 n^\beta + C_1 n^{2\beta-1} + C_2 n^{3\beta-2} + \dots + C_k n^{(k+1)\beta-k} + O(\log^\tau n).$$

Recall that $C_0 = (c\Gamma(1-\beta)\Gamma(1+\beta))^{-1}$ and set $b_n = (n/C_0)^{1/\beta}$. Define $c_\beta = -i\int_0^\infty e^{i\sigma}\sigma^{-\beta}\,d\sigma$ and $C_\beta = c_\beta(\Gamma(1-\beta)\Gamma(1+\beta))^{-1}$.

In what follows, we let $\mathcal{Z}_\beta$ be a positive random variable with characteristic function $E(e^{i\theta\mathcal{Z}_\beta}) = e^{-C_\beta\theta^\beta}$. With these specified, we state

LEMMA 5.1. *For any $a > 0$,*

$$|\mu_Y(b_n^{-1}\varphi_n < a) - \mathbb{P}(\mathcal{Z}_\beta < a)| = d(n),$$

where

$$d(n) = \begin{cases} O(n^{1-1/\beta}) & \text{if } \beta \in (1/2,1),\ q > 1, \\ O(n^{1-1/\beta}(\log n) + n^{-1}) & \text{if } \beta \in (0,1),\ q = 1, \\ O(n^{1-q/\beta}) & \text{if } \beta \in (0,1),\ q < 1. \end{cases}$$

Proof of Lemma 1.2. As shown in [11, 12], the map f defined by (1.1) satisfies (H1), (H2). Moreover, if $\beta \in (0,1)$, then (H) holds with $q = 2\beta$ and $Y = [x_p, 1]$, $p \geq 0$, where x_p is as specified in the paragraph following (1.1) (see [12, Proposition B1]). The conclusion follows immediately from Lemma 5.1. ∎

Lemma 5.1 allows us to establish a version of Theorem 1.1 for more general dynamical systems:

LEMMA 5.2. *Assume that* (H) *holds. Suppose that the function $v : X \to \mathbb{R}$ can be written as $v = 1_Y - \tilde{v}$, a.e. on Y, where $\int \tilde{v}\,d\mu = 0$ and $\mu_Y(|a_n^{-1}S_n(\tilde{v})| > g(n)) < g(n)$, where g is a positive decreasing function such that $g(n) = O(n^{-\beta})$. Then there exists a positive constant C (depending only on f) such that for any $z > C$,*

$$|\mu_Y(a_n(Y)^{-1}S_n^p(v) > z) - \mathbb{P}(\mathcal{Y}_\beta > z)| = E(n),$$

where $E(n) = O(n^{\beta-1})$ if $\beta \in (1/2,1)$, $E(n) = O((\log n)^2 n^{-1/2})$ if $\beta = 1/2$ and, $E(n) = O((\log n)n^{-\beta})$ if $\beta \in (0,1/2)$.

Proof. The result follows by the argument used in the proof of Theorem 1.1 together with Lemma 5.1. ∎

The remainder of this section is devoted to the proof of Lemma 5.1. Below, we collect some instrumental results.

Recall $b_n = (c\Gamma(1-\beta)\Gamma(1+\beta))^{1/\beta}n^{1/\beta}$, $c_\beta = -i\int_0^\infty e^{i\sigma}\sigma^{-\beta}\,d\sigma$ and $C_\beta = c_\beta(\Gamma(1-\beta)\Gamma(1+\beta)^{-1}$.

PROPOSITION 5.3. *Let c and c_H be the real constants defined in* (H) *and Lemma* 4.1*, respectively. Assume $\beta \in (0,1)$. Set $e_\beta = cc_H(c\Gamma(1-\beta) \cdot \Gamma(1+\beta))^{-1/\beta}$. Choose $\epsilon > 0$ such that $\lambda(\theta)$ is well defined for $\theta \in (0,\epsilon)$. In particular, this ensures that $\theta < \epsilon b_n$ for all n large enough. Then*

$$\lambda\left(\frac{\theta}{b_n}\right)^n = e^{-C_\beta\theta^\beta}(1 - ie_\beta n^{1-1/\beta}\theta + E(\theta/b_n)),$$

where $E(\theta/b_n)$ *satisfies the following for all* n *sufficiently large and all* $\theta < \epsilon b_n$*:*

$$E(\theta/b_n) \ll \begin{cases} n^{-1}\theta^{2\beta} + n^{1-q/\beta}\theta^q & \text{if } q \neq 1, \\ n^{-1}\theta^{2\beta} + n^{1-1/\beta}\theta \log(n/\theta) & \text{if } q = 1. \end{cases}$$

Proof. The conclusion follows from Lemma 4.1 and standard computations. We provide the argument for completeness.

Note that for all n sufficiently large and all $\theta < \epsilon b_n$,

$$n \log[\lambda(\theta/b_n)] = -n(1 - \lambda(\theta/b_n)) + O(n|(1 - \lambda(\theta/b_n))^2|).$$

Lemma 4.1 and straightforward calculations imply that

$$1 - \lambda(\theta/b_n) = C_\beta n^{-1}\theta^\beta - ie_\beta n^{-1/\beta}\theta + D(\theta/b_n),$$

where

$$D(\theta/b_n) \ll \begin{cases} n^{-2}\theta^{2\beta} + n^{-q/\beta}\theta^q & \text{if } q \neq 1, \\ n^{-2}\theta^{2\beta} + n^{-1/\beta}\theta \log(n/\theta) & \text{if } q = 1. \end{cases}$$

Thus, we can write

$$\lambda(\theta/b_n)^n = e^{-C_\beta\theta^\beta} \exp\big(-ie_\beta n^{1-1/\beta}\theta + nD(\theta/b_n) + D_1(\theta/b_n)\big), \tag{5.1}$$

where $|D_1(\theta/b_n)| \ll n|(1 - \lambda(\theta/b_n))^2|$.

Using the expansion of $1 - \lambda(\theta/b_n)$, we find that for all n sufficiently large and all $\theta < \epsilon b_n$,

$$D_1(\theta/b_n) \ll \begin{cases} n^{-1}\theta^{2\beta} + n^{-q/\beta}\theta^{q+\beta} & \text{if } q \neq 1, \\ n^{-1}\theta^{2\beta} + n^{-1/\beta}\log(n/\theta)\theta^{\beta+1} + n^{-1}\theta^{2\beta} & \text{if } q = 1. \end{cases}$$

Hence, $|D_1(\theta/b_n)| \ll n^{-1}\theta^{2\beta}$ for all $q > \beta$.

Clearly, $|D_1(\theta/b_n)| \ll n|D(\theta/b_n)|$ as $n \to \infty$. Define

$$D_2(\theta/b_n) = nD(\theta/b_n) + D_1(\theta/b_n).$$

Note that

$$D_2(\theta/b_n) \ll \begin{cases} n^{-1}\theta^{2\beta} + n^{1-q/\beta}\theta^q & \text{if } q \neq 1, \\ n^{-1}\theta^{2\beta} + n^{1-1/\beta}\theta \log(n/\theta) & \text{if } q = 1. \end{cases}$$

This together with (5.1) yields

$$\lambda(\theta/b_n)^n = e^{-C_\beta\theta^\beta}\big(1 - ie_\beta n^{1-1/\beta}\theta + D_2(\theta/b_n) + D_3(\theta/b_n)\big),$$

where $|D_3(\theta/b_n)| \ll n^{2(1-1/\beta)}\theta^2$. To conclude, put $E(\theta/b_n) = D_2(\theta/b_n) + D_3(\theta/b_n)$. ■

A useful consequence of the above result is

Corollary 5.4. *Choose* $\epsilon > 0$ *such that* $\lambda(\theta)$ *is well defined for* $\theta \in (0, \epsilon)$*. Then*

$$\int_0^{\epsilon b_n} \theta^{-1}|\lambda(\theta/b_n) - e^{-C_\beta\theta^\beta}|\, d\theta = d'(n),$$

where

$$d'(n) = \begin{cases} O(n^{1-1/\beta}) & \text{if } \beta \in (1/2, 1),\ q > 1, \\ O(n^{1-1/\beta}(\log n) + n^{-1}) & \text{if } \beta \in (0,1),\ q = 1, \\ O(n^{1-q/\beta}) & \text{if } \beta \in (0,1),\ q < 1. \end{cases}$$

Proof. Define $d_\beta = \mathrm{Re}(C_\beta)$. By Proposition 5.3 with $\beta > 1/2$ and $q > 1$,

$$\int_0^{\epsilon b_n} \theta^{-1} |\lambda(\theta/b_n) - e^{-C_\beta \theta^\beta}|\, d\theta \ll n^{1-1/\beta} \int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta}\, d\theta + n^{-1} \int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta} (\theta^{2\beta-1} + \theta^{q-1})\, d\theta.$$

Clearly, for any $p > \beta - 1$ and all $n \geq 1$,

$$\int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta} \theta^p\, d\theta = \int_0^1 e^{-d_\beta \theta^\beta} \theta^p\, d\theta + \frac{1}{\beta} \int_1^{(\epsilon b_n)^{1/\beta}} e^{-d_\beta \sigma} \sigma^{(p+1)/\beta - 1}\, d\sigma \ll \int_1^\infty e^{-\sigma} \sigma^{(p+1)/\beta-1}\, d\sigma = \text{constant}.$$

Hence, $\int_0^{\epsilon b_n} \theta^{-1} |\lambda(\theta/b_n) - e^{-c_\beta \theta^\beta}|\, d\theta \ll n^{1-1/\beta}$, as desired. The estimate for the case $q < 1$, $\beta \in (0,1)$ follows by a similar argument.

It remains to consider the case $q = 1$, $\beta \in (0,1)$. By Proposition 5.3,

$$\begin{aligned} \int_0^{\epsilon b_n} \theta^{-1} |\lambda(\theta/b_n) - e^{-C_\beta \theta^\beta}|\, d\theta &\ll n^{-1} \int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta} \theta^{2\beta-1}\, d\theta + n^{1-1/\beta} \int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta} \log(n/\theta)\, d\theta \\ &\ll n^{-1} + n^{1-1/\beta} \int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta} \log(n/\theta)\, d\theta. \end{aligned}$$

By Potter's bounds (see [4]), for any $\delta > 0$,

$$\begin{aligned} \int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta} \log(n/\theta)\, d\theta &= \log n \int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta} \log(n/\theta) (\log n)^{-1}\, d\theta \\ &\ll \log n \int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta} (\theta^{-\delta} + \theta^\delta)\, d\theta. \end{aligned}$$

Hence, $n^{1-1/\beta} \int_0^{\epsilon b_n} e^{-d_\beta \theta^\beta} \log(n/\theta)\, d\theta \ll n^{1-1/\beta} \log n$, providing the required estimate. ∎

Proof of Lemma 5.1. By the smoothness inequality for characteristic functions (see, for instance, [8]), for any $\epsilon > 0$,

$$|\mu_Y(b_n^{-1}\varphi_n < a) - \mathbb{P}(\mathcal{Z}_\beta < a)| \le \int_0^{\epsilon b_n} \theta^{-1}|E(e^{i\theta b_n^{-1}\varphi_n}) - E(e^{i\theta\mathcal{Z}_\beta})|\,d\theta + O(b_n^{-1}).$$

Let $d(n)$ be defined as in the statement of Lemma 5.1. Clearly, for all $\beta \in (0,1)$, $b_n^{-1} \ll n^{-1/\beta} \ll d(n)$. Hence, the result will follow once we show that $\int_0^{\epsilon b_n} \theta^{-1}|E(e^{i\theta b_n^{-1}\varphi_n}) - E(e^{i\theta\mathcal{Z}_\beta})|\,d\theta \ll d(n)$.

Choose $\epsilon > 0$ such that $\lambda(z)$ is well defined for $z \in \bar{\mathbb{D}} \cap B_\epsilon(1)$. Let $P(z) : \mathcal{B} \to \mathcal{B}$ denote the family of spectral projections associated with $\lambda(z)$ with $P(1) = P$. Hence, $P(v)(y) \equiv \int_Y v\,d\mu$.

By (H2), we can write $R(z) = \lambda(z)P(z) + Q(z)$, where $Q(z)$ is an operator on $\mathcal{B}$ whose spectrum is contained in a disk of radius strictly less than 1. Hence, for all $n \ge 1$ and for all $z \in \bar{\mathbb{D}} \cap B_\epsilon(1)$, $\|Q(z)^n\|$ decays exponentially fast in n. Thus, $\|R(z)^n - \lambda(z)^n P(z)\| \ll \tau^n$ for some $\tau \in (0,1)$. Also, (H1) together with $\mu(\varphi > n) \ll n^{-\beta}$ implies that $\|P(\theta) - P\| \ll \theta^\beta$ (see, for instance, [11, Proposition 2.7]). Therefore there exists $\tau \in (0,1)$ such that $\|R(\theta)^n - \lambda(\theta)^n P\| \le \|\lambda(\theta)^n G(\theta)\| + \tau^n$, where $\|G(\theta)\| \ll \theta^\beta$. This together with $b_n \ll n^{-1/\beta}$ implies that for all $\theta \in (0, \epsilon b_n)$ and n sufficiently large,

$$\begin{aligned} E(e^{i\theta b_n^{-1}\varphi_n}) &= \int_Y e^{i\theta\varphi_n/b_n}\,d\mu = \int_Y R^n(e^{i\theta\varphi_n/b_n})\,d\mu \\ &= \lambda(\theta/b_n)^n + F(\theta/b_n), \end{aligned} \tag{5.2}$$

where

$$|F(\theta/b_n)| \ll \Big|\lambda(\theta/b_n)^n \int_Y G(\theta/b_n)\Big|\,d\mu \ll n^{-1}\theta^\beta|\lambda(\theta/b_n)^n|.$$

By Proposition 5.3, $\lambda(\theta/b_n)^n = e^{-C_\beta\theta^\beta}1 + E(n)$, where $E(n) \to 0$ as $n \to \infty$. Hence, $|F(\theta/b_n)| \ll n^{-1}\theta^\beta e^{-d_\beta\theta^\beta}$ with $d_\beta = \mathrm{Re}(c_\beta)$.

Recall that for $\beta \in (0,1)$, $\mathcal{Z}_\beta$ is a random variable with characteristic function $E(e^{i\theta\mathcal{Z}_\beta}) = e^{-C_\beta\theta^\beta}$. Equation (5.2) together with the fact that $\|F(\theta/b_n)\| \ll n^{-1}\theta^\beta e^{-d_\beta\theta^\beta}$ implies that

$$\begin{aligned} \int_0^{\epsilon b_n} \theta^{-1}|E(e^{i\theta b_n^{-1}\varphi_n}) - E(e^{i\theta\mathcal{Z}_\beta})|\,d\theta \ll & \int_0^{\epsilon b_n} \theta^{-1}|\lambda(\theta)^n - E(e^{i\theta\mathcal{Z}_\beta})|\,d\theta \\ & + n^{-1}\int_0^{\epsilon b_n} e^{-d_\beta\theta^\beta}\theta^{\beta-1}\,d\theta. \end{aligned}$$

By Corollary 5.4, we find $\int_0^{\epsilon b_n} \theta^{-1}|\lambda(\theta)^n - E(e^{i\theta\mathcal{Z}_\beta})|\,d\theta \ll d'(n)$. Clearly, $n^{-1}\int_0^{\epsilon b_n} e^{-d_\beta\theta^\beta}\theta^{\beta-1}\,d\theta \ll n^{-1}\int_0^n e^{-\sigma}\,d\sigma \ll n^{-1}$. To conclude, put $d(n) = d'(n) + 1/n$. ∎

Acknowledgments. This research started at Surrey University, UK where I held a research fellow position supported by EPSRC Grant EP/F031807/1 and continued at Tor Vergata University, where I hold a research fellow position supported by MALADY Grant, ERC AdG 246953.

I wish to thank Henk Bruin, Ian Melbourne and Roland Zweimüller for very useful conversations. I am grateful to a referee for pointing out several imprecisions and mistakes in previous versions of the manuscript.

References

[1] J. Aaronson, *An Introduction to Infinite Ergodic Theory*, Math. Surveys Monogr. 50, Amer. Math. Soc., 1997.

[2] J. Aaronson and M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps*, Stoch. Dynam. 1 (2001), 193–237.

[3] J. Aaronson and R. Zweimüller, *Limit theory for some positive stationary processes*, preprint, 2011.

[4] N. H. Bingham, C. M. Goldie and J. L. Teugels, *Regular Variation*, Encyclopedia Math. Appl. 27, Cambridge Univ. Press, Cambridge, 1987.

[5] H. Cramer, *On asymptotic expansion for sums of independent random variables with a limiting stable law distribution*, Sankhyā 25 (1963), 13–24.

[6] D. A. Darling and M. Kac, *On occupation times for Markoff processes*, Trans. Amer. Math. Soc. 84 (1957), 444–458.

[7] W. Feller, *Fluctuation theory of recurrent events*, Trans. Amer. Math. Soc. 67 (1949), 98–119.

[8] W. Feller, *An Introduction to Probability Theory and its Applications, II*, Wiley, New York, 1966.

[9] P. Hall, *Two sided bounds on the rate of convergence to a stable law*, Z. Wahrsch. Verw. Gebiete 57 (1981), 349–364.

[10] C. Liverani, B. Saussol and S. Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory Dynam. Systems 19 (1999), 671–685.

[11] I. Melbourne and D. Terhesiu, *Operator renewal theory and mixing rates for dynamical systems with infinite measure*, Invent. Math. 189 (2012), 61–110.

[12] I. Melbourne and D. Terhesiu, *First and higher order uniform ergodic theorems for dynamical systems with infinite measure*, Israel J. Math. 194 (2013), 793–830.

[13] J. P. Nolan, *Numerical calculations of stable densities and distribution functions*, Comm. Statist. Stoch. Models 13 (1997), 759–774.

[14] Y. Pomeau and P. Manneville, *Intermittent transition to turbulence in dissipative dynamical systems*, Comm. Math. Phys. 74 (1980), 189–197.

[15] K. Satyiabaldina, *Absolute estimates of the rate of convergence to stable laws*, Probab. Theory Appl. 18 (1972), 726–728.

[16] D. Terhesiu, *Improved mixing rates for infinite measure preserving transformations*, Ergodic Theory Dynam. Systems, to appear.

[17] M. Thaler, *Transformations on* $[0,1]$ *with infinite invariant measures*, Israel J. Math. 46 (1983), 67–96.

[18] M. Thaler, *The Dynkin–Lamperti arc-sine laws for measure preserving transformations*, Trans. Amer. Math. Soc. 350 (1998), 4593–4607.

[19] M. Thaler and R. Zweimüller, *Distributional limit theorems in infinite ergodic theory*, Probab. Theory Related Fields 135 (2006), 15–52.

[20] D. Thomine, *A generalized central limit theorem in infinite ergodic theory*, Probab. Theory Related Fields, online February 2013.

[21] V. Zolotarev, *Some new inequalities in probability connected with Lévy's metric*, Soviet Math. Dokl. 11 (1970), 231–234.

[22] R. Zweimüller, *Infinite measure preserving transformations with compact first regeneration*, J. Anal. Math. 103 (2007), 93–131.

Dalia Terhesiu
Dipartimento di Matematica
Università di Roma (Tor Vergata)
Via della Ricerca Scientifica
00133 Roma, Italy
E-mail: daliaterhesiu@gmail.com

Received July 13, 2012
Revised version August 12, 2013 (7566)