INFINITE INVARIANT MEASURES FOR NON-UNIFORMLY EXPANDING TRANSFORMATIONS OF $[0,1]$: WEAK LAW OF LARGE NUMBERS WITH ANOMALOUS SCALING.

Massimo Campanino[1]

Stefano Isola

Dipartimento di Matematica, Universitá degli Studi di Bologna,

piazza di Porta S.Donato 5, I-40127 Bologna, Italy

March 15, 1994

Abstract. We consider a class of maps of $[0,1]$ which are expanding everywhere but at a fixed point where the derivative has modulus one. Using the invariant ergodic probability measure of a suitable, everywhere expanding, induced transformation we are able to study the infinite invariant measure of the original map in some detail. Given a continuous function with compact support in $]0,1]$, we prove that its time averages satisfy a 'weak law of large numbers' with anomalous scaling $n/\log n$ and give an upper bound for the decay of correlations.

[1] Work supported by EC grant SC1-CT91-0695

1. Introduction.

We consider a smooth map f of the interval $[0,1]$ into itself with a neutral fixed point, such as those modelling Pomeau-Manneville type 1 intermittency [M.P]. When an orbit falls in the vicinity of this fixed point stays there for a time that can be arbitrarily long before reaching again the 'turbulent region'. Due to this fact, the SRB measure of this dynamical system is simply the Dirac delta measure concentrated at the indifferent fixed point.

However, even though the ordinary Cesaro average along a typical orbit would converge to the above trivial measure, the main result of this paper shows that Cesaro averages rescaled by the logarithm of the time yield a convergence in measure to a σ-finite invariant absolutely continuous measure which describes the statistical properties of intermittent events. A L^1 convergence to this measure was already proved by Collet and Ferrero in [C.F] with different techniques. On the other hand, Aaronson showed in [A] that convergence almost everywhere cannot hold, even for a single observable (see also below: the remark after the proof of Theorem 3.1), so that our result provides, in a sense, a sort of 'optimal' ergodic theorem for this class of transformations.

In the next Section, we introduce the model and define a suitable everywhere expanding induced map that was proved in [C.I] to preserve an absolutely continuous probability measure which is also ergodic and uniformly mixing (for a similar construction see also [C.G]). Using the explicit relation between this measure and the infinite invariant measure of the original map some properties of the latter are studied in detail. In particular, in Section 3, by studying the typical occupation times of subintervals belonging to a suitable partition, we prove a weak law of large number for continuous observables with compact support in $]0,1]$ and show that it cannot be sharpened to a strong one (see also [A]). An upper bound for the decay of correlations is also obtained.

2. The basic model and its induced version.

We first recall the basic setting. Let f be the map of the unit interval $[0,1]$ defined as follows:

(i) $f(0)=0,\ f(1)=1$;

(ii) f is monotone and non decreasing on $I_0=[0,\frac{1}{2}[$ and $I_1=]\frac{1}{2},1]$;

(iii) For each $I_i, i=0,1,$ f extends to a C^2 function f_i on its closure which is onto $[0,1]$.

(iv) There are two numbers $\alpha>1$ and $L>0$ such that:

$$f'_{|I_1}\geq\alpha,\ \ f'(0)=1,\ \ f'_{|]0,1/2[}\geq 1,\ \ \sup_{x\in[0,1]}|f''(x)/f'(x)|\leq L.$$

(v) $f''(0)\neq 0$ which implies $f''(0)>0$.

Let us consider the sequence of points c_k, $k\geq 0$, given by

$$c_0=1,\ \ c_k=f_0^{-1}(c_{k-1}),\quad k\geq 1.$$

This sequence generates a countable partition of $[0,1]$ into the intervals $A_k=[c_k,c_{k-1}]$, $k\geq 1$, which is a Markov partition. In particular, $f(A_k)=A_{k-1}$, $k\geq 1$.

Example. One simple example of such maps is given by the following formula:

$$f(x)=\begin{cases} x/(1-x), & \text{if } x\leq 1/2;\\ 2x-1, & \text{if } x>1/2.\end{cases}$$

In this example we find $c_k=1/(k+1)$, $k\geq 0$, and thus $|A_k|=1/k(k+1)$, $k\geq 1$, where $|I|$ denotes the length of I. More generally, one can easily realize that if $f(x)-x\sim x^\delta$ when $x\to 0$ (where, in order to satisfy the fourth assumption, $\delta\geq 2$) then $c_k\sim k^{-\beta}$, where $\beta=1/(\delta-1)$ and $a(k)\sim b(k)$ means $a(k)/b(k)\to 1$ when $k\to\infty$ (see Lemma 2.4 below). However, if the fifth assumption above has to be satisfied, then δ must be equal to 2.

We now construct a coding. Let $\Omega_{\geq}$ be the set of one-sided sequences $\omega = (\omega_0, \omega_1, \ldots)$, $\omega_i \in \{1, 2, \ldots\}$ satisfying the compatibility condition: given ω_i then either $\omega_{i-1} = \omega_i + 1$ or $\omega_{i-1} = 1$. Then, the map

$$\phi : \omega \to \phi(\omega) = x \quad \text{according to} \quad f^i(x) \in A_{\omega_i}, i \geq 1$$

is a bijection between $\Omega_{\geq}$ and the points of $[0, 1]$ which are not preimages of the origin. Moreover, ϕ conjugates the map f with the shift τ on $\Omega_{\geq}$.

For every integer $i \geq 1$ we denote by x_i the projection on the i^{th} symbol, i.e. $x_i(\omega) = \omega_i$, and define the "free" probability measure μ by

$$\mu(\omega_i) = |A_{\omega_i}|, \quad i \geq 1 \tag{2.1}$$

With slight abuse of language we shall again denote by μ the measure $\mu \circ \phi^{-1}$, i.e. the Lebesgue measure on $[0, 1]$.

We now introduce the infinite sequence τ_j, $j \geq 1$, of successive entrance times in the state 1: $\tau_1(\omega) = \inf\{i \geq 0 \ : \ x_i(\omega) = 1\}$ and, for $j \geq 2$, $\tau_j(\omega) = \inf\{i > \tau_{i-1} \ : \ x_i(\omega) = 1\}$. Furthermore, we define a sequence of integer valued random variables by

$$\sigma_j(\omega) = \tau_{j+1} - \tau_j, \quad j \geq 0 \tag{2.2}$$

with the convention that $\tau_0 = -1$.

Remark. If f is affine on each interval A_k of the Markov partition, with slope $s_k = |A_{k-1}|/|A_k|$ (with $|A_0| = 1$), then the stochastic process on $\Omega_{\geq}$ given by $x_i(\omega) = \omega_i$, $i \geq 1$, is a Markov chain with conditional probabilities $p_{ij} = \mu(f_0^{-1}(A_j) \bigcap A_i)/\mu(A_i)$. As a consequence, the random variables σ_j are independent and identically distributed. Their common law is given by: $\text{Prob}(\sigma_j = k) = \mu(k)$ for any $j \geq 0$ and $k \geq 1$.

Definition 2.1. *The 'first passage' map (on the interval A_1), is the map $g : [0, 1] \to [0, 1]$ induced by f in the following way:*

$$x \to g(x) = f^{n(x)}(x) \quad \text{where} \quad n(x) = 1 + \min\{n \geq 0 \ : \ f^n(x) \in A_1 \} \tag{2.3}$$

Notice that the usual return time function $r(x)$ in the interval A_1 is then given by $r(x) = \min\{n \geq 1 \ : \ f^n(x) \in A_1 \ \} = n \circ f(x)$. Equivalently, the map g is given by

$$x \to g(x) = g_k(x) = f^k(x) \quad \text{if} \quad x \in A_k, \ \ k \geq 1. \tag{2.4}$$

and defined arbitrarily on the countable set c_k, $k \geq 0$ (for similar constructions see also [T2],[P],[P.S],[C.G]). This map is expanding and surjective, i.e.:

(1) $g'(x)_{|A_k} \geq \alpha > 1$, for any $k \geq 1$;

(2) g maps each A_k monotonically onto $[0,1]$, more precisely $\overline{g_k([c_k, c_{k-1}])} = [0,1]$, $k \geq 1$.

Moreover, one can easily realize that the following property holds:

(3) for any $\omega \in \Omega_{\geq}$, let $x = \phi(\omega)$. Then $g^j(x) \in A_{\sigma_j}$, $j \geq 1$, where the integers $\sigma_j = \sigma_j(\omega)$ are defined in (2.2).

Furthermore, let $\Sigma_{\geq}$ be the set of *all* one-sided sequences σ of the form $\sigma = (\sigma_0, \sigma_1, \ldots)$, $\sigma_j \in \{1, 2, \ldots\}$. Then, the map

$$\pi : \sigma \to \pi(\sigma) = x \quad \text{according to} \quad g^j(x) \in A_{\sigma_j}, \ \ j \geq 1 \tag{2.5}$$

is a bijection between $\Sigma_{\geq}$ and the points of $[0,1]$ which are not preimages of zero. Moreover, π conjugates the map g with the shift τ on $\Sigma_{\geq}$.

In [C.I] we have constructed an invariant ergodic probability measure ρ for the dynamical system $([0,1], g)$ as a Gibbs state on $\Sigma_{\geq}$ for the function

$$V(\sigma) = \log(\phi'_{\sigma_0}(\pi(\sigma_1\sigma_2\ldots)))$$

where $\phi_k : [0,1] \to [c_k, c_{k-1}]$ is the inverse function of g_k.

Furthermore, ρ is proved to be absolutely continuous with respect to the Lebesgue measure μ (with bounded Radon-Nikodym derivative) and to satisfy the uniform mixing property (see [R]).

We now prove the following:

Lemma 2.1. *Set $D_k = \bigcup_{l\geq k} A_k$. Then the measure ν defined for any Borel subset E of $[0,1]$ by*

$$\nu(E) = \sum_{k\geq 1} \rho(f^{-k+1}E \cap D_k) \tag{2.6}$$

is invariant under f.

Proof. Clearly $D_k = D_{k+1} \cup A_k$. Therefore,

$$\begin{aligned}\nu(f^{-1}E) &= \sum_{k\geq 1} \rho(f^{-k+1}(f^{-1}E) \cap D_{k+1}) + \sum_{k\geq 1} \rho(f^{-k+1}(f^{-1}E) \cap A_k) \\ &= \sum_{k\geq 2} \rho(f^{-k+1}E \cap D_k) + \sum_{k\geq 1} \rho(g^{-1}E \cap A_k) \\ &= \sum_{k\geq 1} \rho(f^{-k+1}E \cap D_k) = \nu(E)\end{aligned} .$$

Q.E.D.

In particular (2.6) yields

$$\nu(A_k) = \sum_{l\geq k} \rho(A_l) = \rho(D_k) \tag{2.7}$$

and the mean return time in the interval A_1 can be computed using Kac's formula and (2.7) as

$$\nu([0,1]) = \int_{A_1} r(x)d\nu(x) = \int_0^1 n(x)d\rho(x) = \sum_{k\geq 1} k\rho(A_k) \tag{2.8}$$

On the other hand, we know that $\rho(A_k) \sim C/k^2$ as $k \to \infty$ for some constant $C > 0$ (see Lemma 2.4 below). Hence the above sum diverges logarithmically.

Let us now consider the transfer operators $\mathcal{P}$ and $\mathcal{L}$ defined by:

$$\mathcal{P}h(x) = \sum_{y:f(x)=y} \frac{h(y)}{|f'(y)|} \qquad \text{and} \qquad \mathcal{L}h(x) = \sum_{y:g(x)=y} \frac{h(y)}{|g'(y)|}$$

An important property of these operators is that their fixed points must be the densities of the corresponding absolutely continuous invariant measures. Indeed, if we consider for instance the operator $\mathcal{L}$, we know from [C.I] that g is leaves invariant the measure $\rho = h\mu$ where μ denotes the Lebesgue measure and h is the unique fixed point of $\mathcal{L}$ in the space of all real Hölder continuous functions on $]0,1]$. Thus,

$$\rho(g^{-1}E) = \int_{g^{-1}E} h(x)d\mu(x) = \int_E (\mathcal{L}h)(x)d\mu(x) = \int_E h(x)d\mu(x) = \rho(E)$$

and conversely.

Lemma 2.2. *Let e be such that $\mathcal{P}e = e$, then,*

$$e = h + \sum_{i=1}^{\infty} \mathcal{P}_0^i h \tag{2.9}$$

which is identical to (2.6) provided $\nu = e\mu$.

Proof. Let $\psi_k : [0,1] \to I_k$, $k = 0,1$ be such that $\psi_k \circ f_k(x) = x$, then $\mathcal{P}$ can be decomposed as

$$\mathcal{P}h(x) = \mathcal{P}_0 h(x) + \mathcal{P}_1 h(x) = h \circ \psi_0(x) \cdot \psi_0'(x) + h \circ \psi_1(x) \cdot \psi_1'(x) \tag{2.10}$$

and moreover one can easily check that

$$\mathcal{L}h(x) = \sum_{i=1}^{\infty} h \circ \phi_i(x) \cdot \phi_i'(x) = \mathcal{P}_1 h(x) + \sum_{i=1}^{\infty} \mathcal{P}_0^i \mathcal{P}_1 h(x) \tag{2.11}$$

The statement follows immediately from (2.10) and (2.11). Q.E.D.

Remark. We point out that Lemmas 2.1 and 2.2 hold in a sligthly more general setting: assumption (v) on f is not needed and assumption (iv) can be partially relaxed in such a way to include transformations f such that $f(x) - x \sim x^\delta$ when $x \to 0$ with any $\delta > 0$ (we recall that if $\delta < 2$ the invariant measure ν is finite) and obviously uniformly expanding maps like for instance:

$$f(x) = \begin{cases} (2-a)x/(1-ax), & \text{if } x \leq 1/2; \\ 2x - 1, & \text{if } x > 1/2. \end{cases}$$

where $a = 1 - \epsilon$, $\epsilon > 0$.

We conclude this Section with the following estimates (see [T1], [C.F] for related results):

Lemma 2.3. *Under the assumptions (i)-(v) above*

$$\frac{C_1}{x} \le e(x) \le \frac{C_2}{x} \tag{2.12}$$

where C_1, C_2 are positive constants.

Proof. We know from [C.I] that h satisfies the bound $d^{-1} \le h \le d$ for some positive constant d. Then from Lemma 2.2

$$d^{-1} \sum_{k=0}^{\infty} \mathcal{P}_0^k 1(x) \le e(x) \le d \sum_{k=0}^{\infty} \mathcal{P}_0^k 1(x) \tag{2.13}$$

Furthermore,

$$\sum_{k=0}^{\infty} \mathcal{P}_0^k 1(x) = \sum_{k=0}^{\infty} (\psi_0^k)'(x)$$

where $\psi_0^k \circ f_0^k(x) = x$. We now use a reasoning similar to that of [T1].

We have $\psi_0(x) < x$ for any $x \in]0,1]$. Thus, for any positive integer N, the function

$$B_N(x) = \sum_{k=0}^{N-1} \mathcal{P}_0^k 1(x)$$

is decreasing on $]0,1]$. Therefore, for any fixed $x \in]0,1]$ we have

$$B_N(\psi_0(x)) \ge B_N(y) \ge B_N(x)$$

for any $\psi_0(x) < y < x$. On the other hand,

$$\int_{\psi_0(x)}^{x} B_N(y)dy = \sum_{k=0}^{N-1} \left[\psi_0^k(x) - \psi_0^{k+1}(x)\right] = x - \psi_0^N(x)$$

so that one immediately finds the inequalities

$$\begin{aligned} (x - \psi_0(x))B_N(x) &\le x - \psi_0^N(x) \\ (x - \psi_0(x))B_N(\psi_0(x)) &\ge x - \psi_0^N(x) \end{aligned} \tag{2.14}$$

If we now let $N \to \infty$ and note that in this limit $\psi_0^N(x) \to 0$ we obtain

$$\frac{f_0(x)}{f_0(x) - x} \le \sum_{k=0}^{\infty} \mathcal{P}_0^k 1(x) \le \frac{x}{x - \psi_0(x)} \tag{2.15}$$

Now, from the assumptions (i)-(v) on the transformation f it follows that $f_0(x)$ and $\psi_0(x)$ admit expansions in the vicinity of the origin of the form

$$f_0(x) = x + ax^2 + u(x), \qquad \psi_0(x) = x - ax^2 + v(x) \tag{2.16}$$

where $2a = f''(0) > 0$ and u, v are continuous functions defined on $[0,1]$ such that $u(x)/x^2$ and $v(x)/x^2$ tend to zero as $x \to 0$. Putting together (2.15) and (2.16) we easily find

$$\frac{1}{x}\left(\frac{1}{a + u(x)/x^2}\right) \le \sum_{k=0}^{\infty} \mathcal{P}_0^k 1(x) \le \frac{1}{x}\left(\frac{1}{a - v(x)/x^2}\right) \tag{2.17}$$

which immediately yields the result.

Notice that for the paradigmatic example: $f_0(x) = x/(1-x)$, we simply have $\psi_0(x) = (1 + 1/x)^{-1}$ and $\psi_{i0}(x) = (i + 1/x)^{-1}$, so that $\psi'_{i0}(x) = (1 + ix)^{-2}$ and (2.12) follows at once with the choice $C_1^{-1} = C_2 = d$. Q.E.D.

Lemma 2.4. *Let μ be the free probability measure over $\mathbb{N}$ defined in (2.1) and ρ the probability measure invariant under g. Then, under the assumptions (i)-(v) above,*

$$\mu(k) = |A_k| \sim \frac{1}{k^2} \tag{2.18}$$

and

$$\rho(A_k) \sim \frac{C}{k^2} \tag{2.19}$$

where C is a positive constant and $a(k) \sim b(k)$ means $a(k)/b(k) \to 1$ when $k \to \infty$.

Proof. We first prove that there are constants $C, D > 0$ and $\gamma > 0$ such that for k large enough,

$$C - Dc_{k-1}^{\gamma} \le \frac{\rho(A_k)}{|A_k|} \le C + Dc_{k-1}^{\gamma} \tag{2.20}$$

Indeed, let $\{x_k\}_{k\ge1}$ be the sequence defined by

$$\rho(A_k) = \int_{A_k} h(x)dx = h(x_k)|A_k| \qquad x_k \in A_k, \;\; k \ge 1$$

Obviously $c_k \leq x_k \leq c_{k-1}$. We know from [C.I] that h is γ-Hölder for some $\gamma > 0$, so that, for some $L > 0$,

$$|h(x_k) - h(x_{k+1})| \leq L|x_k - x_{k+1}|^\gamma \to 0 \quad \text{as} \quad k \to \infty$$

Hence $\{h(x_k)\}_{k\geq 1}$ is a Cauchy sequence whose limit we denote by h_0. Then we write,

$$\frac{\rho(A_k)}{|A_k|} = h_0 + (h(x_k) - h_0)$$

The result follows immediately form the inequality $|h(x_k) - h_0| \leq L x_k^\gamma \leq L c_{k-1}^\gamma$.

Now, from Lemma 2.3 we have that if k is sufficently large there exist constants $0 < M_1 \leq M_2$ such that

$$M_1 \log{(c_{k-1}/c_k)} \leq \nu(A_k) \equiv \int_{c_k}^{c_{k-1}} e(x)dx \leq M_2 \log{(c_{k-1}/c_k)} \tag{2.21}$$

On the other hand, we know from Lemma 2.1 that

$$\nu(A_k) = \rho(D_k) = \sum_{l\geq k} \rho(A_k) \tag{2.22}$$

so that

$$\nu(A_{k-1}) - \nu(A_k) = \rho(A_k) \tag{2.23}$$

¿From (2.21) and (2.23) we get

$$K_1 \log{(c_{k-2}/c_k)} \leq \rho(A_k) \leq K_2 \log{(c_{k-2}/c_k)} \tag{2.24}$$

for suitable constants $0 < K_1 \leq K_2$. Finally, (2.20), (2.23) and (2.24) entail that there exist constants $0 < R_1 \leq R_2$ such that for k large enough

$$R_1 \log{(c_{k-1}/c_k)} \leq |A_k| \leq R_2 \log{(c_{k-2}/c_k)} \tag{2.25}$$

If we now set $c_k \sim k^{-\beta}$ then (2.25) implies $\beta = 1$ (for $|A_k| = c_{k-1} - c_k$) so that (2.18) follows at once. Notice that (2.20) becomes

$$C - \frac{D}{k^\gamma} \leq \frac{\rho(A_k)}{|A_k|} \leq C + \frac{D}{k^\gamma} \tag{2.26}$$

which, together with (2.18), implies (2.19). Q.E.D.

3. Ergodic properties of the invariant measure ν.

Since the invariant probability measure $\rho = h\mu$ is ergodic (and mixing) for the dynamical system $([0,1], g)$, the two following limit properties hold true and are in a sense dual one of the other [L.M]:

i) for every density u, i.e. $u \in L^1([0,1])$ such that $u \geq 0$, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathcal{L}^k u = h \cdot \mu(u) \tag{3.1}$$

ii) for any $u \in L^1([0,1])$ we have (strong law of large numbers),

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} u(g^k(x)) = \rho(u) \quad \text{a.e.} \tag{3.2}$$

On the other side, under an hypothesis on the smoothness of f slightly stronger than (iii) above, one has the

Theorem ([C.F]) *For any real function u, bounded away from 0 and $+\infty$, $1/2$ Hölder,*

$$\lim_{n\to\infty} \frac{1}{a_n} \sum_{k=0}^{n-1} \mathcal{P}^k u = e\mu(u) \tag{3.3}$$

for the uniform convergence on compact subset of $]0,1]$; here $a_n = \sum_{k=0}^{n-1}(\mathcal{P}^k 1)(1/2)$ is of order $n/\log n$. Moreover, f is ergodic with respect to the measure $e\mu$.

On the other hand, it is easy to realize that the SRB measure of the dynamical system $([0,1], f)$ is simply δ_0, so that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} u(f^k(x)) = u(0) \quad \text{a.e.} \tag{3.4}$$

In addition, it follows from the results of Aaronson [A] that it is *impossible* to find a sequence $a_n \uparrow \infty$ such that the ergodic average $\lim_{n\to\infty} \frac{1}{a_n} \sum_{k=0}^{n-1} u(f^k(x))$

would converge almost everywhere to $\nu(u)$, even for a single $u \in L^1$. In other words, in this case we have lost the correspondence between the convergence of the (suitably rescaled) Cesaro average of the transfer operator (3.3) and a (suitably rescaled) strong law of large numbers.

Nevertheless, we shall prove the following 'weak law of large numbers':

Theorem 3.1. *For any continuous real function u compactly supported on $]0,1]$, for any $\epsilon > 0$ and for any $\tilde{\nu} \ll \nu$, such that $\tilde{\nu}([0,1]) < \infty$,*

$$\lim_{n\to\infty} \tilde{\nu}\left(\left\{x : \left|\frac{1}{a_n}\sum_{k=0}^{n-1} u(f^k(x)) - \nu(u)\right| \geq \epsilon\right\}\right) = 0 \tag{3.5}$$

where $a_n \sim cn/\log n$ and $c > 0$ is a positive constant.

We prove Theorem 3.1 through a sequence of Lemmas.

Definition 3.1. *Let $u :]0,1] \to I\!R$ be any real function. Its* induced *version $\tilde{u}$ is defined by*

$$\tilde{u}(x) = \sum_{s=0}^{n(x)-1} u(f^s x) \quad \text{where} \quad n(x) = 1 + \min\{k \geq 0 \ : \ f^k(x) \in A_1\} \tag{3.6}$$

Notice that from (2.5) it follows immediately that $n(g^k(x)) = \sigma_k$ where $\sigma = (\sigma_0, \sigma_1, \ldots) = \pi^{-1}(x)$.

Lemma 3.1. *For any u as in Theorem 3.1,*

$$\nu(u) = \rho(\tilde{u}) \tag{3.7}$$

Proof. From Lemma 2.2, part (i), we have

$$\nu(u) = \int_{[0,1]} u(x)e(x)d\mu(x) = \sum_{k=0}^{\infty}\int_{[0,1]} u(x)(\mathcal{P}_0^k h)(x)d\mu(x)$$
$$= \sum_{k=0}^{\infty}\int_{f_0^{-k}([0,1])} u(f_0^k x)h(x)d\mu(x)$$

On the other hand, one can easily check that

$$\sum_{k=n(x)}^{\infty}\int_{f_0^{-k}([0,1])} u(f_0^k x)h(x)d\mu(x) = \sum_{k=0}^{n(x)-1}\int_{[0,1]\setminus f_0^{-k}([0,1])} u(f_0^k x)h(x)d\mu(x) = 0$$

so that

$$\sum_{k=0}^{\infty} \int_{f_0^{-k}([0,1])} u(f_0^k x) h(x) d\mu(x) = \sum_{k=0}^{n(x)-1} \int_{[0,1]} u(f_0^k x) h(x) d\mu(x)$$
$$= \int_{[0,1]} \sum_{k=0}^{n(x)-1} u(f^k x) h(x) d\mu(x) = \rho(\tilde{u})$$

Q.E.D.

Let us now consider an orbit $\{f^k x\}_{k=0}^{n-1}$, for some $x \in]0,1]$ and denote by $N(n,x)$ the number its passages in A_1, or, in other terms, the number of symbols in its (truncated) σ-coding $(\sigma_0, \sigma_1, \ldots \sigma_{N(n,x)-1})$.

We can thus write:

$$\sum_{k=0}^{n-1} u(f^k x) = \sum_{s=0}^{N(n,x)-1} \tilde{u}(g^s x) + R(n,x,u) \tag{3.8}$$

where the remainder is given by

$$R(n,x,u) = \sum_{s=m(n,x)}^{n-1} u(f^s x) \quad \text{with} \quad m(n,x) = \sum_{k=0}^{N(n,x)-1} n(g^k x) \tag{3.9}$$

Lemma 3.2. *For any u as in Theorem 3.1,*

$$\lim_{n\to\infty} \frac{1}{N(n,x)} \sum_{s=0}^{n-1} u(f^s x) = \nu(u) \quad \text{a.e.} \tag{3.10}$$

Proof. Obvioulsy $N(n,x) \to \infty$ as $n \to \infty$ for any $x \in]0,1]$. Moreover, since u is bounded and compactly supported on $]0,1]$ then $\tilde{u}$ is bounded as well and in particular is in L^1. Therefore we may apply (3.2) so that,

$$\lim_{n\to\infty} \frac{1}{N(n,x)} \sum_{s=0}^{N(n,x)-1} \tilde{u}(g^s x) = \rho(\tilde{u}) \quad \text{a.e.} \tag{3.11}$$

Finally,

$$|R(n,x,u)| \leq \sum_{s=m(n,x)}^{n(m(n,x))-1} |u(f^s x)|$$

and the r.h.s. is uniformly bounded. This entails

$$\lim_{n\to\infty} \frac{R(n,x,u)}{N(n,x)} = 0$$

The assertion is now is an immediate consequence of (3.11) and Lemma 3.1. Q.E.D.

We now prove the key Lemma:

Lemma 3.3. *For any $\epsilon > 0$,*

$$\lim_{n\to\infty} \rho\left(\left\{x : 1-\epsilon < \frac{N(n,x)}{a_n} < 1+\epsilon\right\}\right) = 1 \tag{3.12}$$

where a_n is as in Theorem 3.1.

Proof.

We first consider the function $m(n,x)$ defined in (3.9) and observe that for $x \in]0,1]$

$$\frac{m(n,x)}{n} \to 1 \quad \text{as} \quad n \to \infty \tag{3.13}$$

in measure. This implies that the sequence a_n of Theorem 3.1 can be replaced by the sequence $a_{m(n,x)}$ without altering (3.12).

We then first consider the following inverse problem:

Let N and $x \in]0,1]$ be given and let $(\sigma_0, \ldots, \sigma_{N-1})$ be the first N symbols of the sequence $\pi^{-1}(x)$. We want to prove a 'weak law of large numbers' for the sequence

$$m(N,x) = \sum_{i=0}^{N-1} \sigma_i \tag{3.14}$$

We partition the non negative integers into some intervals $J_k = [l_k, r_k]$ for $k \geq 0$ that are defined below. Correspondingly we define the events

$$E_k^{(i)} = \{\sigma_i \in J_k^N\}, \qquad i \geq 0 \tag{3.15}$$

and the sequences

$$Z_k(N) = \sum_{i=0}^{N-1} \chi_{E_k^{(i)}} \tag{3.16}$$

Thus, we have

$$\sum_{k=0}^{\infty} l_k Z_k(N) \leq m(N) \leq \sum_{k=0}^{\infty} r_k Z_k(N) \tag{3.17}$$

Set $\varepsilon_N = \exp(-c_1(\log N)^\alpha)$ where c_1 is some positive constant. Then, the intervals $J_k = J_k(N) = [l_k, r_k]$ are defined inductively as follows:

$l_0 = 0$, $r_0 = 1$ and

$$l_k = r_{k-1} + 1, \quad r_k = [(1+\varepsilon_N)l_k] \quad \text{for} \quad k \geq 1 \tag{3.18}$$

so that $r_k \leq (1+\varepsilon_N)l_k$ for any $k \geq 1$[1].

Let α be a real constant $0 < \alpha < 1$. We then introduce two integer-valued sequences $\{g_N\}$ and $\{h_N\}$ defined by

$$\begin{aligned} g_N &= [N \exp(-c_3(\log N)^\alpha)] \\ h_N &= [N \exp(c_4(\log N)^\alpha)] \end{aligned} \tag{3.19}$$

where c_3, c_4 are positive constants to be specified later on, and divide the intervals J_k's into three categories denoted by $\mathcal{I}_1$, $\mathcal{I}_2$, $\mathcal{I}_3$:

the intervals contained in $[1, g_N]$ belong to $\mathcal{I}_1$, those intersecting $[g_N + 1, h_N]$ belong to $\mathcal{I}_2$ and $\mathcal{I}_3$ contains the remainder.

We see at once that the intervals in $\mathcal{I}_3$ can be neglected in our analysis. Indeed, let the event $F^{(3)}$ be defined by

$$F^{(3)} = \{\max_{0 \leq i \leq N-1} \sigma_i > h_N\} \tag{3.20}$$

Then, for N large enough, we can find a constant $G > 0$ such that

$$\rho\left(F^{(3)}\right) \leq G\frac{N}{h_N} \tag{3.21}$$

so that it goes to 0 for $N \to \infty$.

Now, in order to exploit the exponential mixing property of the random variables σ_i's we decompose each Z_k as

$$Z_k = Z_k^{(0)} + \ldots + Z_k^{(s-1)} \tag{3.22}$$

---

[1] Notice that a finite number of the J_k's is made out of a single point.

where $s_N = \exp(c_2(\log N)^\alpha)$ with $c_2 > 0$ to be specified later, and

$$Z_k^{(j)} = \sum_{\substack{0 \le i \le N-1 \\ i = j (\bmod s)}} \chi_{E_k^{(i)}} \tag{3.23}$$

Set moreover

$$p_k = \rho(E_k^{(0)}). \tag{3.24}$$

Now, by virtue of Lemma 2.4, we can find two positive constant b, c such that have for k large enough,

$$p_k \ge \sum_{n=l_k}^{r_k} \frac{c}{n^2} \ge b \left(\frac{1}{l_k} - \frac{1}{r_k} \right) \tag{3.25}$$

We now want to estimate the probability of the event $F_k^{(1)}$ defined for k corresponding to an interval in $\mathcal{I}_1$ by

$$F_k^{(1)} = \left\{ |Z_k - Np_k| \ge t\, s_N \sqrt{\frac{Np_k(1-p_k)}{s_N}} \right\}$$

where t will be specified later.

On the other hand, if J_k belongs to $\mathcal{I}_1$ and we choose the constants c_1, c_2, c_3 in such a way that $c_3 > c_1 + c_2$ then

$$\frac{N}{s_N} p_k \ge R \exp(\gamma (\log N)^\alpha)$$

where $\gamma = c_3 - (c_1 + c_2) > 0$ and R is a suitable positive constant. Indeed, from (3.25) we have

$$\frac{N}{s_N} p_k \ge b \frac{N}{s_N} \left(\frac{1}{l_k} - \frac{1}{r_k} \right) \ge \frac{b}{l_k} \frac{N}{s_N} \left(1 - \frac{1}{1 + \varepsilon_N} \right) \ge R \exp(\gamma (\log N)^\alpha)$$

This allows us to apply the law of large numbers to the intervals J_k's in $\mathcal{I}_1$, in the following way: for any $j \in \{1, \ldots, s_N\}$ and for any real positive β,

$$\rho\left(Z_k^{(j)} - \frac{Np_k}{s_N} \geq t\sqrt{\frac{Np_k(1-p_k)}{s_N}}\right) =$$

$$\rho\left(\exp\beta\left(Z_k^{(j)} - \frac{Np_k}{s_N}\right) \geq \exp\beta t\sqrt{\frac{Np_k(1-p_k)}{s_N}}\right) \leq$$

$$\rho\left(\prod_l e^{\beta(\chi_{E_k^{(l)}} - p_k)}\right)\exp\left(-\beta t\sqrt{\frac{Np_k(1-p_k)}{s_N}}\right) \leq$$

$$\left(\rho(e^{\beta(\chi_{E_k^{(0)}} - p_k)}) + e^{\beta}R\eta^{s_N}\right)^{N/s_N}\exp\left(-\beta t\sqrt{\frac{Np_k(1-p_k)}{s_N}}\right)$$

where the product ranges over the l's of the sublattice corresponding to $Z_k^{(j)}$ and the term involving $0 < \eta < 1$ and the constant $R > 0$ accounts for the exponential mixing property. We now make the choices

$$t = \exp(c_5(\log N)^{\alpha}), \qquad \beta = \frac{c_6 t}{\sqrt{\frac{Np_k(1-p_k)}{s_N}}}$$

and, by expanding the generating function at $\beta = 0$, we find

$$\begin{aligned}\rho\left(Z_k^{(j)} - \frac{Np_k}{s_N} \geq t\sqrt{\frac{Np_k(1-p_k)}{s_N}}\right) &\leq Be^{-bt^2/2} \\ &= B\exp\left(-b\exp(2c_5(\log N)^{\alpha})/2\right)\end{aligned} \tag{3.26}$$

for suitable positive constants B and b.

In the same way we prove that

$$\rho\left(Z_k^{(j)} - \frac{Np_k}{s_N} \leq -t\sqrt{\frac{Np_k(1-p_k)}{s_N}}\right) \leq Be^{-bt^2/2} \tag{3.27}$$

By adding over $j = 1, \ldots, s_N$, we finally obtain

$$\rho\left(F_k^{(1)}\right) \leq 2Bs_N\exp(-b\exp(2c_5(\log N)^{\alpha})/2) \tag{3.28}$$

Set

$$F^{(1)} = \bigcup_{k \in \mathcal{I}_1} F_k^{(1)}$$

where, with slight abuse of language $k \in \mathcal{I}_1$ means that k ranges over the values corresponding to $J_k \in \mathcal{I}_1$. Using (3.28) we can estimate the ρ-probability of the above event as

$$\rho\left(F^{(1)}\right) = \sum_{k \in \mathcal{I}_1} \rho\left(F_k^{(1)}\right) \leq B' g_N s_N \exp(-b \exp(2c_5 (\log N)^{\alpha})/2) \tag{3.29}$$

so that

$$\lim_{N \to \infty} \rho\left(F^{(1)}\right) = 0 \tag{3.30}$$

Now, the contribution of $\mathcal{I}_1$ to $m(N, x)$, which we denote by $m_1(N)$, is

$$\sum_{k \in \mathcal{I}_1} l_k Z_k(N) \leq m_1(N) \leq \sum_{k \in \mathcal{I}_1} r_k Z_k(N)$$

and the results obtained above immediately imply that, with probability approaching 1 as $N \to \infty$, it is of order

$$m_1(N) \sim \kappa_1 N \log g_N \sim \kappa N \log N \tag{3.31}$$

where κ_1, κ are suitable positive constants.

Let us now consider the intervals in $\mathcal{I}_2$.

For $J_k \in \mathcal{I}_2$ we want to estimate the ρ-probability of the event

$$F_k^{(2)} = \{Z_k \geq tNp_k\} \tag{3.32}$$

where this time we keep t constant, say $t = 4$. We perform the same decomposition of Z_k as in (3.21) and estimate for any $j = 1, \ldots, s$,

$$\rho\left(Z_k^{(j)} \geq t\frac{Np_k}{s_N}\right)$$

By proceeding as above with the choice β =constant, say $\beta = 1$, we obtain, for N large enough,

$$\rho\left(Z_k^{(j)} \geq t\frac{Np_k}{s_N}\right) \leq \left(\left(1 + p_k(e^{\beta} - 1) + e^{\beta}\eta^{s_N}\right) e^{-\beta t p_k}\right)^{N/s_N} \leq De^{(-dN/s_N)}$$

for some constants $D, d > 0$.

By adding over $j = 1, \ldots, s$ we find

$$\rho(F_k^{(2)}) \leq D s_N e^{(-dN/s_N)} \tag{3.33}$$

and finally, setting

$$F^{(2)} = \bigcup_{k \in \mathcal{I}_2} F_k^{(2)}$$

we find

$$\rho\left(F^{(2)}\right) = \sum_{k \in \mathcal{I}_2} \rho\left(F_k^{(2)}\right) \leq D'(h_N - g_N) s_N e^{(-dN/s_N)}$$

so that, again,

$$\lim_{N \to \infty} \rho\left(F^{(2)}\right) = 0$$

This implies that the contribution to $m(N, x)$ of I_2, which satisfies

$$m_2(N) \leq \sum_{k \in \mathcal{I}_2} r_k Z_k(N)$$

is, with large probability, of order strictly smaller than $N \log N$ and therefore can be neglected in the limit $N \to \infty$. Indeed, using (3.32) and (3.33), it can be estimated as follows

$$m_2(N) \leq r_1 N(\log h_N - \log g_N) \leq r_2 N(\log N)^\alpha$$

for a suitable choice of the constants r_1, r_2. To summarize, we have proved so far that there is a constant $\kappa > 0$ such that, with probability approaching 1 as $N \to \infty$, $m(N, x) \sim \kappa N \log N$. Otherwise stated, for any $\epsilon > 0$

$$\rho\left(\left\{x : 1 - \epsilon < \frac{m(N, x)}{\kappa N \log N} < 1 + \epsilon\right\}\right) \to 1 \quad \text{as} \quad N \to \infty \tag{3.34}$$

This concludes the first part of the proof.

We now proceed to reverse the above result. Let M be fixed and let us denote by $\overline{N}_1(M)$ the maximum integer for which the inequality

$$(1 + \epsilon)\kappa \overline{N}_1 \log \overline{N}_1 \leq M$$

is satisfied. Similarly let $\overline{N}_2(M)$ be the minimum integer satisfying

$$(1-\epsilon)\kappa\overline{N}_2\log\overline{N}_2 \geq M$$

¿From (3.34) we have

$$\lim_{M\to\infty}\rho\left(\{\, x : m(\overline{N}_1, x) \leq M \leq m(\overline{N}_2, x)\,\}\right) = 1 \tag{3.35}$$

Finally, define

$$N_1(M, x) = \max\{N \,:\, m(N, x) \leq M\}$$

$$N_2(M, x) = \min\{N \,:\, m(N, x) \geq M\}$$

so that

$$N_1(M, x) \leq N(m, x) \leq N_2(M, x) \tag{3.36}$$

Then, from (3.34), (3.35) and (3.36) we have immediately that,

$$\lim_{m\to\infty}\rho\left(\left\{\, x : (1-\epsilon)\frac{cm}{\log m} \leq N(m, x) \leq (1+\epsilon)\frac{cm}{\log m}\,\right\}\right) = 1 \tag{3.37}$$

with $c = 1/\kappa$. Lemma 3.3 now follows by putting together (3.37) and (3.13). Q.E.D.

Proof of Theorem 3.1. Theorem 3.1 is now an immediate consequence of Lemmas 3.1, 3.2, 3.3. Q.E.D.

Remark. For the particular situation considered in this paper one can give a simple argument showing that the above result cannot be sharpened (for a more general result see [A]).

Consider for instance eq. (3.34). It tells us that if we set

$$m_N = \sum_{i=1}^{N}\sigma_i \qquad \text{and} \qquad b_N = \kappa N\log N$$

then

$$\lim_{N\to\infty}\frac{m_N}{b_N} = 1 \qquad \text{in probability} \tag{3.38}$$

One may then wonder if the above limit property holds in a strong sense, i.e. if

$$\rho\left(\lim_{N\to\infty}\frac{m_N}{b_N} = 1\right) = 1 \tag{3.39}$$

We now show that it does not and in fact

$$\rho\left(\lim_{N\to\infty}\frac{m_N}{b_N}=1\right)=0 \tag{3.40}$$

Indeed, from Lemma 2.4 we have that for a large enough one can choose a suitable constant c such that for any $N\geq 0$

$$\rho\left(\sigma_N>a\right)\geq\frac{c}{a} \tag{3.41}$$

Hence, for any constant $\ell>1$ and N sufficently large

$$\rho\left(\sigma_N>\ell\, b_N\right)\geq\frac{c}{\ell\, b_N}=\frac{c}{\ell\kappa N\log N} \tag{3.42}$$

and thus

$$\sum_{N=1}^{\infty}\rho\left(\sigma_N>\ell\, b_N\right)=\infty \tag{3.43}$$

Therefore, from the extension of the Borel-Cantelli lemma to non-disjoint events (see e.g. [B]) it follows that

$$\rho\left(\frac{\sigma_N}{b_N}>\ell\ \text{ infinitely often}\right)=1$$

so that

$$\rho\left(\frac{m_N}{b_N}>\ell\ \text{ infinitely often}\right)=1 \tag{3.44}$$

and consequently

$$\rho\left(\overline{\lim_{N\to\infty}}\,\frac{m_N}{b_N}=\infty\right)=1 \tag{3.45}$$

which implies (3.40). We can actually say more. Since (3.38) implies the convergence a.e. on a subsequence, (3.40) is valid for *every* sequence of constants b_N.

A further consequence of our construction is an upper bound on the decay of correlations.

Theorem 3.2. *For any pair of bounded continuous functions u, v compactly supported on $]0,1]$ and such that $\nu(u)=\nu(v)=0$, there is a constant $R>0$ such that*

$$\left|\int_{[0,1]}u(f^t x)\, v(x)\, d\nu(x)\right|\leq\frac{R}{t} \tag{3.46}$$

Proof. By virtue of Lemma 3.1,

$$\int_{[0,1]} u(f^t x)\, v(x)\, d\nu(x) = \int_{[0,1]} \sum_{l=0}^{n(x)-1} u(f^{l+t}(x))\, v(f^l(x))\, d\rho(x) \tag{3.47}$$

Now, for $x \in]0,1]$ we write

$$t = \sum_{s=0}^{\tau(t,x)-1} n(g^s x) + r(t,x) \tag{3.48}$$

where $r(t,x) < n(g^{\tau(t,x)}(x))$.

Then,

$$\int_{[0,1]} u(f^t x)\, v(x)\, d\nu(x) = \int_{[0,1]} \sum_{l=0}^{n(x)-1} u(g^{\tau(t,x)}(f^{l+r(t,x)}(x)))\, v(f^l(x))\, d\rho(x) \tag{3.49}$$

Furthermore, one can easily check that the r.h.s. of (3.49) can be rewritten as

$$\int_{[0,1]} \sum_{l=0}^{n(x)-1} u(g^{\tau(t,x)}(f^{l+r(t,x)}(x)))\, v(f^l(x))\, d\rho(x) =$$
$$\sum_{k=0}^{\infty} \int_{D_{k+1}} u(g^{\tau(t,x)}(f^{k+r(t,x)}(x)))\, v(f^k(x))\, d\rho(x) =$$
$$\sum_{k=0}^{\infty} \int_{[0,1]} u(g^{\tau(t,x)}(f^{r(t,x)}(x)))\, v(x)\, d(f^k_{0*}\rho)(x)$$

where the sets D_k's are as in Lemma 2.1 and $\psi_*\rho$ denotes the measure obtained by pushing forward ρ with ψ, i.e. $\psi_*\rho(A) = \rho(\psi^{-1}(A))$. On the other hand, since u, v are compactly supported in $]0,1]$, the above sum is actually finite. Hence, we can find a constant $M > 0$ (depending of u and v) such that

$$\left| \int_{[0,1]} u(f^t x)\, v(x)\, d\nu(x) \right| \leq M \sup_{0 \leq k \leq M} \left| \int_{[0,1]} u(g^{\tau(t,x)}(f^{r(t,x)}(x)))\, v(x)\, d(f^k_{0*}\rho)(x) \right| \tag{3.50}$$

After a short reflection one easily realizes that an upper bound to the integral in the r.h.s. of (3.50) can be found by estimating the $f^k_{0*}\rho$-probability of those 'bad' points $x \in]0,1]$ such that $\tau(t,x)$ grows at the smallest possible rate (in particular smaller than the most probable rate $ct/\log t$ prescribed by Lemma 3.3). Now, the worst situation is attained when $\tau(t,x)$ is just a constant, in particular equal to 0. On the other hand it is easy to see that the above probability is maximal for $k=0$. Indeed, we have

$$\begin{aligned} & f_{0*}^k \rho\{x : \tau(t,x) = 0\} = f_{0*}^k \rho\{x : n(x) > t\} = \rho\{x : n(f_0^k x) > t\} = \\ & \rho\{x : n(x) > t + k\} \leq \rho\{x : n(x) > t\} = \rho\{x : \tau(t,x) = 0\} \end{aligned} \tag{3.51}$$

and, using Lemma 2.4, we find for t large enough,

$$\rho\{x : \tau(t,x) = 0\} = \rho\{x : n(x) > t\} \leq \frac{C}{t} \tag{3.52}$$

Moreover, it is not difficult to check that $\rho\{x : \tau(t,x) = m\}$ is, for any $m \geq 1$, of order strictly smaller than (3.52).

Finally, using (3.50), (3.51), (3.52) and the boundedness of u, v, one immediately obtains (3.47). Q.E.D.

References.

[A] Aaronson J., *The asymptotic distributional behaviour of transformations preserving infinite measures*, Journal d'analyse mathématique, **39**, 203-234 (1981)

[B] Billingsley P., *Convergence of Probability Measures*, Wiley, New York, 1968

[C.F] Collet P. and Ferrero P., *Some limit ratio theorem related to a real endomorphism with a neutral fixed point*, Ann. Inst. H. Poincaré **52**, 283 (1990)

[C.G] Collet P. and Galves A., *Statistics of close visit to the indifferent fixed point of an interval map*, Jour. Stat. Phys. **72**, 459 (1993)

[C.I] Campanino M. and Isola S., *Statistical properties of long return times in type I intermittency*, (1993) Forum Math., to appear.

[L.M] Lasota A. and Mackey M.C., *Probabilistic properties of deterministic systems*, Cambridge University Press, 1985

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

[P] Pianigiani G., *First return map and invariant measures*, Isr. Jour. of Math. **35**, 32-48 (1980)

[P.S] Prellberg T. and Slawny J., *Maps of intervals with indifferent fixed points: thermodynamic formalism and phase transitions*, Jour. of Stat. Phys. **66**, 503-514 (1992)

[R] Ruelle D., *Thermodynamic Formalism*, Addison-Wesley Publ. Co., 1978

[T1] Thaler M., *Estimates of the invariant densities of endomorphisms with indifferent fixed points*, Isr. Jour. of Math. **37**, 303-313 (1980)

[T2] Thaler M., *Transformations in* $[0,1]$ *with infinite invariant measures*, Isr. Jour. of Math. **46**, 67-96 (1983)