STUDIA MATHEMATICA, T. XCIII (1989)

Refinement of the Shannon–McMillan–Breiman Theorem for some maps of an interval

by

KRYSTYNA ZIEMIAN (Warszawa)

Abstract. The Refinement of the Shannon–McMillan–Breiman Theorem is proved for a large class of maps of an interval with nonpositive Schwarzian derivative.

Introduction. Let f be a piecewise monotone map of an interval I into itself, $\mathscr{A}$ the natural partition of I into the intervals of monotonicity of f, and μ an f-invariant probability measure on I. We have the equality $h_\mu(f, \mathscr{A}) = h_\mu(f)$, because $\mathscr{A}$ is a generator. If μ is ergodic, we have the well-known Shannon–McMillan–Breiman Theorem, which says that for μ-almost every x

$$(1) \qquad \lim_{n\to\infty} \left[-n^{-1} \log \mu(A_n(x))\right] = h_\mu(f, \mathscr{A}) = h_\mu(f),$$

where $A_n(x)$ is the atom of $\mathscr{A}_n = \bigvee_{i=0}^{n-1} f^{-i}\mathscr{A}$ containing x.

For f belonging to a large class of maps of an interval with nonpositive Schwarzian derivative (the same as in [4], [7], [8]) we prove an almost sure invariance principle for the sequence $\log \mu(A_n(x)) + nh_\mu(f)$ (the so-called Refinement of the Shannon–McMillan–Breiman Theorem). This theorem implies other limit theorems for this sequence, such as the central limit theorem and the law of iterated logarithm (except for some special cases like $f(x) = 4x(1-x)$). The law of iterated logarithm says in particular that for μ-almost every x the rate of convergence in (1) is not greater than $\sqrt{(\log\log n)/n}$.

The scheme of the proof is similar to that of [5, Th. 9.1].

Section 1 contains the assumptions on f and basic definitions.

In Section 2 we prove the almost sure invariance principle for the sequence $-\log J_n(x) + nh_\mu(f)$, where $J_n(x)$ is the jacobian of f^n with respect to the invariant measure μ. One of the assumptions of this theorem is that the asymptotic variance of this sequence does not vanish. We prove in Section 3 that this assumption is satisfied for a large class of maps (e.g. for all maps $f(x) = 4\alpha x(1-x)$ except the case $\alpha = 1$). Lemma 8, Proposition 1 and the ideas of their proofs were communicated to the author by A. Zdunik.

In Section 4 we prove that if $\alpha \in (0, \frac{1}{2})$ then $\log(J_n(x)\mu(A_n(x)))$ is of order

n^α for μ-almost every x, which yields the main result of the paper, the Refinement of the Shannon–McMillan–Breiman Theorem.

The author is very grateful to M. Denker, G. Keller, M. Misiurewicz and F. Przytycki for helpful discussions.

1. Assumptions and definitions. Let f be a map of a closed interval I into itself satisfying the following conditions [4]:

(i) There is a finite subset $A \subset I$ containing the endpoints of I such that $f|_{I\setminus A}$ is of class C^3.

(ii) $f' \neq 0$ on $I\setminus A$.

(iii) $Sf \leqslant 0$ on $I\setminus A$, where Sf is the Schwarzian derivative of f.

(iv) If $f^p(x) = x$, then $|(f^p)'(x)| > 1$.

(v) There exists a neighbourhood U of the set A such that for all $a \in A$, $n \geqslant 0$, we have $f^n(a) \in A$ or $f^m(a) \notin U$ for all $m \geqslant n$.

(vi) For all $a \in A$ there exists a neighbourhood U_a of a and constants $\alpha, \omega, \delta > 0$, $u \geqslant 0$ such that

(a) $\alpha|x-a|^u \leqslant |f'(x)| \leqslant \omega|x-a|^u$,

(b) $|f''(x)| \leqslant \delta|x-a|^{u-1}$

for all $x \in U_a$.

In view of [4, Th. 6.2] there exist an integer k and an f^k-invariant probability measure μ absolutely continuous with respect to the Lebesgue measure λ such that the system (f^k, μ) is exact, in particular weakly mixing. In the sequel we will consider this system. Obviously, f^k satisfies assumptions (i)–(vi) [4, Lemma 3.1], so we can simply assume $f = f^k$.

We recall that $\mathscr{A}$ denotes the natural partition of I into the intervals of monotonicity of f, $\mathscr{A}_n = \bigvee_{i=0}^{n-1} f^{-i}\mathscr{A}$, $A_n(x)$ is the atom of $\mathscr{A}_n$ which contains x, and $J_n(x)$ is the Jacobian of f^n with respect to the invariant measure μ. We will write J instead of J_1.

Let $\mathscr{A} = \{A_1, \ldots, A_k\}$. We will deal with the so-called "label" process (ξ_n) associated with the system (f, μ):

$$\xi_n(x) = i \quad \text{if } f^n(x) \in A_i,$$

which is a stationary process on the probability space $(I, \mathscr{B}, \mu)$, where $\mathscr{B}$ is the σ-field of Borel sets.

2. Almost sure invariance principle for the process $(-\log J \circ f^i + h_\mu(f))$. In this section we will show that Th. 1 of [7] holds for the function $-\log J + h_\mu(f)$ instead of a function F with bounded p-variation, $p \geqslant 1$. The function $-\log J + h_\mu(f)$ usually is not even bounded, because f can have critical points. This involves some additional difficulties. However, the main idea of the proof will be similar. We will consider the measurable function $-\log J + h_\mu(f)$ as a functional of the "label" process (ξ_n) and prove that the assumptions of [5, Th. 7.1], [1, Th. 7] are satisfied for the process $(-\log J \circ f^i + h_\mu(f))$.

LEMMA 1. *$\int |\log J|^r d\mu < +\infty$ for all $r \geqslant 1$.*

Proof. For the proof it suffices to notice that

$$(J(x))^{-1} = p(\xi_0 | \xi_1, \xi_2, \ldots)(x), \tag{2}$$

where the random variable $p(\xi_0 | \xi_1, \xi_2, \ldots)$ is defined as

$$p(\xi_0 | \xi_1, \xi_2, \ldots)(x) = \mu\{\xi_0 = i | \xi_1, \xi_2, \ldots\}(x) \quad \text{if } x \in A_i,$$

and then apply [3, Lemma 2.1]. ∎

LEMMA 2. *For every $s \in N$, $\int |\log J - E(\log J | \xi_0, \ldots, \xi_n)|^s d\mu$ tends to zero exponentially as n tends to infinity.*

Proof. By (2) and [3, Lemma 2.2] it suffices to prove the exponential convergence to zero of the sequence $\psi(n)$ defined as

$$\psi(n) = \sup_{1 \leqslant i \leqslant k} \int |\mu\{\xi_0 = i | \xi_1, \xi_2, \ldots\} - \mu\{\xi_0 = i | \xi_1, \xi_2, \ldots, \xi_n\}| d\mu.$$

Fix i, $1 \leqslant i \leqslant k$, and set

$$g_n^0 = |\mu\{\xi_0 = i | \xi_1, \xi_2, \ldots\} - \mu\{\xi_0 = i | \xi_1, \xi_2, \ldots, \xi_n\}|.$$

We have to estimate $\int g_n^0 d\mu$. First we observe that by the definition of the functions $\mu\{\xi_0 = i | \xi_1, \xi_2, \ldots\}$ and $\mu\{\xi_0 = i | \xi_1, \xi_2, \ldots, \xi_n\}$ we have

$$\operatorname{supp}\mu\{\xi_0 = i | \xi_1, \xi_2, \ldots\} = f^{-1}(f(A_i)),$$

$$\operatorname{supp}\mu\{\xi_0 = i | \xi_1, \xi_2, \ldots, \xi_n\} = f^{-1}\Big(\bigcup_{x \in A_i} A_n(f(x))\Big).$$

In view of [7, Ths. 2, 3] there exist $\gamma_1 \in (0, 1)$, $c_1 > 0$ such that for all $y \in I$ and $n \in N$

$$\mu(A_n(y)) \leqslant c_1 \gamma_1^n. \tag{3}$$

Therefore

$$\mu\Big(f^{-1}\Big(\bigcup_{x \in A_i} A_n(f(x))\Big) \setminus f^{-1}(f(A_i))\Big) = \mu\Big(\bigcup_{x \in A_i} A_n(f(x)) \setminus f(A_i)\Big) \leqslant 2c_1\gamma_1^n. \tag{4}$$

g_n^0 takes values from $[0, 1]$, which together with (4) gives

$$\int g_n^0 d\mu \leqslant 2c_1\gamma_1^n + \int_{f^{-1}(f(A_i))} g_n^0 d\mu. \tag{5}$$

It remains to estimate $\int_{f^{-1}(f(A_i))} g_n^0\,d\mu$. We define the function g_n by

$$g_n = |p(\xi_0|\xi_1, \xi_2, \ldots) - p(\xi_0|\xi_1, \xi_2, \ldots, \xi_n)|,$$

where $p(\xi_0|\xi_1, \xi_2, \ldots)$ is as in the proof of Lemma 1 and $p(\xi_0|\xi_1, \xi_2, \ldots, \xi_n)(x) = \mu\{\xi_0 = j|\xi_1, \xi_2, \ldots, \xi_n\}(x)$ if $x \in A_j$.

Let for $r \in N$, $0 \leq l \leq r-1$,

$$A_{l,r} = \{x \in A_i\colon\ l/r \leq g_n(x) < (l+1)/r\},$$
$$A_{l,r}^0 = \{x \in f^{-1}(f(A_i))\colon\ l/r \leq g_n^0(x) < (l+1)/r\}.$$

For every $r \in N$ define simple functions $g_{n,r}, g_{n,r}^0$ by

$$g_{n,r} = \sum_{l=0}^{r-1} \frac{l}{r}\chi_{A_{l,r}}, \qquad g_{n,r}^0 = \sum_{l=0}^{r-1} \frac{l}{r}\chi_{A_{l,r}^0}. \tag{6}$$

We have

$$\int_{f^{-1}(f(A_i))} g_n^0\,d\mu \leq \frac{1}{r} + \int_{f^{-1}(f(A_i))} g_{n,r}^0\,d\mu = \frac{1}{r} + \sum_{l=0}^{r-1} \frac{l}{r}\mu(A_{l,r}^0). \tag{7}$$

We will estimate $\sum_{l=0}^{r-1} (l/r)\mu(A_{l,r}^0)$. We claim that

$$A_{l,r}^0 = f^{-1}(f(A_{l,r})). \tag{8}$$

Indeed, if $y \in A_{l,r}^0$, then $f(y) \in f(A_i)$, which means that there exists $x \in A_i$ such that $f(y) = f(x)$. We have $g_n^0(x) = g_n^0(y)$, because g_n^0 is constant on every set $f^{-1}(z)$, $z \in I$. The functions g_n^0 and g_n are equal on A_i, so we have also $g_n(x) = g_n^0(x) = g_n^0(y)$, which means that $g_n(x)$ satisfies the same inequalities as $g_n^0(y)$. Hence $x \in A_{l,r}$ and $y \in f^{-1}(f(A_{l,r}))$.

Now take $y \in f^{-1}(f(A_{l,r}))$. There exists $x \in A_{l,r}$ such that $f(y) = f(x)$. By the same argument as above $g_n^0(y) = g_n^0(x) = g_n(x)$, so $g_n^0(y)$ satisfies the same inequalities as $g_n(x)$. Hence $y \in A_{l,r}^0$, which completes the proof of (8).

It follows from [7, Th. 3] that there exists $p > 1$ such that for every $G \in \mathscr{B}$, $\mu(G)$ does not exceed $(\lambda(G))^{1/p}$ multiplied by a positive constant L. Using this fact, (6), (8) and the concavity of the function $x^{1/p}$ we obtain

$$\begin{aligned}\sum_{l=0}^{r-1} \frac{l}{r}\mu(A_{l,r}^0) &= \sum_{l=0}^{r-1} \frac{l}{r}\mu\big(f^{-1}(f(A_{l,r}))\big) = \sum_{l=0}^{r-1} \frac{l}{r}\mu\big(f(A_{l,r})\big)\\ &\leq L\sum_{l=0}^{r-1} \frac{l}{r}\big(\lambda(f(A_{l,r}))\big)^{1/p} \leq L'\sum_{l=0}^{r-1} \frac{l}{r}\big(\lambda(A_{l,r})\big)^{1/p}\\ &= L'\frac{r-1}{2}\left[\sum_{l=0}^{r-1} \frac{2l}{r(r-1)}\big(\lambda(A_{l,r})\big)^{1/p}\right]\\ &\leq L'\frac{r-1}{2}\left[\sum_{l=0}^{r-1} \frac{2l}{r(r-1)}\lambda(A_{l,r})\right]^{1/p}\\ &= L'\frac{r-1}{2}\left(\frac{2}{r-1}\right)^{1/p}\left[\sum_{l=0}^{r-1} \frac{l}{r}\lambda(A_{l,r})\right]^{1/p}\\ &\leq L' r^{1-1/p}\big(\textstyle\int g_{n,r}\,d\lambda\big)^{1/p} \leq L' r\big(\textstyle\int g_n\,d\lambda\big)^{1/p},\end{aligned} \tag{9}$$

where $L' = L(\sup_I |f'|)^{1/p}$.

We will prove that there exist $c > 0$, $\gamma \in (0, 1)$ such that

$$\int g_n\,d\lambda \leq c\gamma^n. \tag{10}$$

First we observe that

$$p(\xi_0|\xi_1, \xi_2, \ldots, \xi_n)(x) = \frac{\mu(A_{n+1}(x))}{\mu(A_n(f(x)))},$$
$$p(\xi_0|\xi_1, \xi_2, \ldots)(x) = (J(x))^{-1}.$$

Hence in view of [8, Th. 2] there exist $c_0 > 0$, $\gamma_0 \in (0, 1)$ such that

$$\int g_n\,d\mu \leq c_0\gamma_0^n. \tag{11}$$

Let $K = \operatorname{supp}\mu$. It is proved in [4, Cor. (5.5) and Th. (6.2) (b)] that K is a finite union of intervals. Let m be the number of these intervals, and put $G_n = \bigcup_{x \in K} A_{n+1}(x)$. In view of (3) we have

$$\mu(G_n \setminus K) \leq mc_1\gamma_1^{n+1}.$$

By the definition of g_n, $G_n = \operatorname{supp} g_n$ and

$$\int g_n\,d\lambda = \int_{G_n} g_n\,d\lambda \leq \int_K g_n\,d\lambda + mc_1\gamma_1^{n+1}. \tag{12}$$

The density of μ is bounded from below on K by a positive constant $d > 0$, so $\int_K g_n\,d\lambda \leq d^{-1}\int g_n\,d\mu$, which together with (12) and (11) completes the proof of (10).

Recall that r was fixed arbitrarily, so we can assume $r \geq 1/\alpha^n$, where $\gamma^{1/p} < \alpha < 1$. Now the desired estimate of $\int g_n^0\,d\mu$ (and $\psi(n)$) follows from (5), (7), (9) and (10). ∎

LEMMA 3. *The process (ξ_n) satisfies the strong mixing condition with mixing coefficient $\alpha(n)$ converging exponentially to zero.*

Proof. This follows from [7, Th. 4]. ∎

We will now consider the process $(-\log J \circ f^i + h_\mu(f))$. By Rokhlin's formula

$$\int(-\log J + h_\mu(f))\,d\mu = 0. \tag{13}$$

Define

$$S(n) = \sum_{i=0}^{n-1} (-\log J \circ f^i + h_\mu(f)),$$

$$\sigma^2 = \int (-\log J + h_\mu(f))^2 \, d\mu + 2 \sum_{i=1}^{\infty} \int (-\log J + h_\mu(f))(-\log J \circ f^i + h_\mu(f)) \, d\mu.$$

LEMMA 4. *The correlation coefficients*

$$\int (-\log J + h_\mu(f))(-\log J \circ f^i + h_\mu(f)) \, d\mu, \quad i = 1, 2, \ldots,$$

converge to zero exponentially.

Proof. The standard proof follows from (13), Lemmas 2, 3 and [5, Lemma 7.2.1]. ∎

THEOREM 1. *If $\sigma^2 \neq 0$, then*

(i) $$\sup_{z \in \mathbf{R}} \left| \mu \left\{ \frac{1}{\sigma \sqrt{n}} S(n) \leqslant z \right\} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} \, dx \right| = O(n^{-\nu})$$

for some $\nu > 0$.

(ii) *Set for $t \geqslant 0$*

$$S(t) = \sum_{0 \leqslant i < t} (-\log J \circ f^i + h_\mu(f)).$$

Without changing its distribution one can redefine the process $(S(t))_{t \geqslant 0}$ on a richer probability space with standard Brownian motion $(B(t))_{t \geqslant 0}$ such that almost surely

$$|S(t)/\sigma - B(t)| = O(t^{1/2-\lambda}) \quad \textit{for some } \lambda > 0.$$

Proof. In view of (13) and Lemmas 1–4 all the assumptions of [5, Th. 7.1] and [1, Th. 7] are satisfied. ∎

Remark 1. Theorem 1 (ii) implies integral tests, log-log laws and weak invariance principles for the process $(-\log J \circ f^i + h_\mu(f))$ (see [5]).

3. Estimation of σ^2. In the previous section we have proved the central limit theorem and almost sure invariance principle for the process $(-\log J \circ f^i + h_\mu(f))_{i \in N}$. Both of these theorems hold under the condition $\sigma^2 > 0$. We will prove that for a large class of maps satisfying (i)–(vi) this condition is satisfied.

Let $B = \overline{\bigcup_{n=1}^{\infty} f^n(A)}$. This is the set of singularities of the density φ of μ, which is a Cantor set of μ-measure 0 (see [4]). Let V be a neighbourhood of B.

LEMMA 5. *Let H be the constant from [7, Lemma 4]. If $y \in A_n(x)$, $[f^n(x), f^n(y)] \cap V = \emptyset$ and $|f^n(x) - f^n(y)|$ is sufficiently small, then*

$$\left| \frac{(f^n)'(x)}{(f^n)'(y)} - 1 \right| \leqslant 2H |f^n(x) - f^n(y)|.$$

Proof. Let $y \in A_n(x)$, $[f^n(x), f^n(y)] \cap V = \emptyset$. In view of [7, Lemma 4]

$$\left| \frac{1}{(f^n)'(x)} - \frac{1}{(f^n)'(y)} \right| \leqslant H|x-y|.$$

Hence

$$\left| \frac{(f^n)'(y)}{(f^n)'(x)} - 1 \right| \leqslant H |(f^n)'(y)| \cdot |x-y|$$

and if $|f^n(x) - f^n(y)|$ is sufficiently small then the argument used in the proof of [7, Lemma 10] completes the proof. ∎

LEMMA 6. *Let $y \in A_n(x)$, $f^n(x)$, $f^n(y) \in K$, $[f^n(x), f^n(y)] \cap V = \emptyset$. If $|f^n(x) - f^n(y)|$ is sufficiently small and $\varphi(x) > 0$, then $\varphi(y) > 0$ and*

$$\left| \frac{\varphi(x)}{\varphi(y)} - 1 \right| \leqslant L |f^n(x) - f^n(y)|,$$

where L is a constant depending only on V.

Proof. We have assumed that μ is weakly mixing. Hence, in view of [4, Th. 6.2 (e)], for μ-almost every x, $\varphi(x)$ is the limit of the sequence $(f_*^k(1)(x))$ up to a constant, and the same is true for y instead of x (f_* is the Perron–Frobenius operator with respect to the Lebesgue measure). Hence, it suffices to estimate $f_*^k(1)(x)/f_*^k(1)(y)$ uniformly with respect to k. We have

$$f_*^k(1)(x) = \sum_{z \in f^{-k}(x)} \frac{1}{|(f^k)'(z)|}, \quad f_*^k(1)(y) = \sum_{w \in f^{-k}(y)} \frac{1}{|(f^k)'(w)|}.$$

Since $[f^n(x), f^n(y)] \cap V = \emptyset$, for every $z \in f^{-k}(x)$ we can find a corresponding $w \in f^{-k}(y)$ such that z, w belong to the same interval of monotonicity of f^k. Fix such a pair z, w. We have

$$\frac{|(f^k)'(w)|}{|(f^k)'(z)|} = \frac{|(f^{k+n})'(w)|}{|(f^{k+n})'(z)|} \cdot \frac{|(f^n)'(x)|}{|(f^n)'(y)|}.$$

In view of Lemma 5 both

$$\left| \frac{(f^{k+n})'(w)}{(f^{k+n})'(z)} - 1 \right| \quad \text{and} \quad \left| \frac{(f^n)'(x)}{(f^n)'(y)} - 1 \right|$$

do not exceed $2H |f^n(x) - f^n(y)|$. Hence $|f_*^k(1)(x)/f_*^k(1)(y) - 1|$ does not exceed $|f^n(x) - f^n(y)|$ multiplied by a constant depending only on V. ∎

Let $(\tilde{I}, \tilde{f}, \tilde{\mu})$ be the natural extension of the system (I, f, μ) and $\pi\colon \tilde{I} \to I$ the natural projection. We consider the partition $\mathscr{C} = \{\pi^{-1}(x)\}_{x\in I}$ of the space $\tilde{I}$, with the canonical system of conditional measures $\{\tilde{\mu}_x\}_{x\in I}$ induced by $\tilde{\mu}$. If $[x, y] \cap V = \emptyset$, then we have the natural one-to-one transformation $i\colon \pi^{-1}(x) \to \pi^{-1}(y)$ such that for every $\tilde{z} \in \pi^{-1}(x)$ and $n \in N$, the points $\pi(\tilde{f}^{-n}(z))$, $\pi(\tilde{f}^{-n}(i(z)))$ belong to the same interval of monotonicity of f^n.

LEMMA 7. *If $x, y \in K$, $[x, y] \cap V = \emptyset$ and $|x-y|$ is sufficiently small, then for every measurable $\tilde{G} \subset \pi^{-1}(x)$ of positive $\tilde{\mu}_x$-measure we have*

$$\left|\frac{\tilde{\mu}_y(i(\tilde{G}))}{\tilde{\mu}_x(\tilde{G})} - 1\right| \leqslant P|x-y|,$$

where $P > 0$ depends only on V.

Proof. It suffices to observe that for every $z \in I$, $\tilde{\mu}_z\{\tilde{v} \in \pi^{-1}(z)\colon \pi \circ \tilde{f}^{-n}(\tilde{v}) = \alpha\} = (J_n(\alpha))^{-1}$ and apply Lemmas 5 and 6. ∎

LEMMA 8. *If*

$$(14)\qquad \log|f'| - h_\mu(f) = u \circ f - u \qquad \mu\text{-a.e.}$$

for some $u \in L^2(\mu)$, then $u|_{K\setminus V}$ is equal μ-almost everywhere to a continuous function on $K \setminus V$.

Proof. If $u \in L^2(\mu)$ then by the Lusin Theorem there exists a closed set $F_0 \subset I$ such that $u|_{F_0}$ is continuous and $\mu(F_0) > \frac{3}{4}$. Let $F \subset F_0$ be the subset on which (14) holds, $\mu(F_0) = \mu(F)$. Let U be an interval contained in $K \setminus V$. We put $\tilde{F} = \pi^{-1}(F_0)$.

By the ergodicity of μ, for μ-almost every $x \in I$, the sequence $(\tilde{f}^{-n}(\tilde{x}))_{n\in N}$ contains points from $\tilde{F}$ which appear with frequency at least $\frac{3}{4}$ (because $\tilde{\mu}(\tilde{F}) > \frac{3}{4}$). We can express the same in terms of our system of conditional measures $\{\tilde{\mu}_x\}_{x\in I}$. Namely, if we set

$$D_x = \{\tilde{y} \in \pi^{-1}(x)\colon (\tilde{f}^{-n}(\tilde{x}))_{n\in N} \text{ contains points from } \tilde{F} \text{ which appear with frequency at least } \tfrac{3}{4}\},$$

then

$$(15)\qquad \tilde{\mu}_x(D_x) = 1 \quad \text{for } \mu\text{-almost every } x \in I.$$

If $x, y \in U$, $\tilde{\mu}_x(D_x) = 1$, and x, y are close enough, then in view of Lemma 7, $\tilde{\mu}_y(i(D_x)) = 1$. If additionally $\tilde{\mu}_y(D_y) = 1$, then we have $\tilde{\mu}_y(i(D_x) \cap D_y) = 1$, in particular $i(D_x) \cap D_y \neq \emptyset$. Take $\tilde{y}_0 \in i(D_x) \cap D_y$, and put $\tilde{x}_0 = i^{-1}(\tilde{y}_0)$. We have $\tilde{x}_0 = (x, x_1, x_2, \ldots)$, $\tilde{y}_0 = (y, y_1, y_2, \ldots)$, where x_n, y_n belong to the same interval of monotonicity of f^n. By the definition of D_x and D_y, we can choose increasing sequences of positive integers $(n_\alpha)_{\alpha\in N}$, $(m_\alpha)_{\alpha\in N}$ such that $x_{n_\alpha}, y_{m_\alpha} \in F$ for every $\alpha \in N$ and each of these sequences has density at least $\frac{3}{4}$. This density is sufficient to find a common increasing subsequence $(k_\alpha)_{\alpha\in N}$ of both sequences. Hence, if we set $x_\alpha = x_{k_\alpha}$, $y_\alpha = y_{k_\alpha}$, we obtain from (14)

$$\begin{aligned}(16)\qquad u(x) - u(y) &= u(f^{k_\alpha} x_\alpha) - u(f^{k_\alpha} y_\alpha)\\ &= \sum_{j=1}^{k_\alpha} \big(u(f^j x_\alpha) - u(f^{j-1} x_\alpha)\big) + u(x_\alpha)\\ &\quad - \sum_{j=1}^{k_\alpha} \big(u(f^j y_\alpha) - u(f^{j-1} y_\alpha)\big) - u(y_\alpha)\\ &= \log\frac{(f^{k_\alpha})'(x_\alpha)}{(f^{k_\alpha})'(y_\alpha)} + u(x_\alpha) - u(y_\alpha).\end{aligned}$$

From (15), (16), Lemma 5, [7, Lemma 4] and the uniform continuity of u on F, we obtain the uniform continuity of u on a subset of U of full μ-measure. But U is an arbitrary interval contained in $K \setminus V$, so we have proved the uniform continuity of u on a subset of $K \setminus V$ of full μ-measure. In particular, this subset is dense in $K \setminus V$, so we can extend u restricted to this subset to a continuous function on $K \setminus V$. ∎

Let A_0 denote the set of those critical points of f which are contained in K.

LEMMA 9. *Let $u \in L^2(\mu)$ be as in Lemma 8. If $J \subset K$ is an interval such that J does not contain critical points of f and $u|_J$ is equal μ-a.e. to a continuous function on J, then $u|_{f(J)}$ is equal μ-a.e. to a continuous function on $f(J)$.*

Proof. This follows immediately from (i) and (14). ∎

LEMMA 10. *Let $u \in L^2(\mu)$ be as in Lemma 8. If $a \in A_0$, then for every $n \in N$ there exist constants $d_1, d_2 \in \mathbf{R}$, $\alpha \in (0, 1)$ such that for μ-almost all y sufficiently close to $f^n(a)$*

$$d_1 + \log|y - f(a)|^\alpha \leqslant u(y) \leqslant d_2 + \log|y - f^n(a)|^\alpha.$$

Proof. Let $U \subset I \setminus V$ be an open interval. By the topological exactness of the system (f, K, μ) [4, Prop. 5.7] there exists $k \geqslant 0$ such that $f^k(U) \supset K$. In particular, there exists $z \in U$ such that $f^k(z) = a$. By Lemma 8, we can assume that u is continuous on $K \setminus V$. Take the first critical point on the trajectory of z. We can assume that this is a. By Lemma 9, u is continuous in a neighbourhood of a. By (vi) and (14) the assertion of the lemma holds for $n = 1$. We obtain it for any n by induction, using (14) and also (vi) whenever we pass through other critical points. ∎

LEMMA 11. *$\log\varphi \in L^p(\mu)$ for every $p \geqslant 1$.*

Proof. The proof is analogous to the proof of [8, Lemma 11]. ■

Proposition 1. *If there exists $a \in A_0$ such that for some $n \in N$, $f^{-1}(f^n(a))$ intersects $K \setminus \bigcup_{n=0}^{\infty} f^i(A_0)$, then $\sigma^2 > 0$.*

Proof. In view of [6, Lemma 1], $\sigma^2 > 0$ if and only if $-\log J + h_\mu(f)$ is not homologous to 0 in $L^2(\mu)$. But $J = (\varphi \circ f)|f'|/\varphi$ and in view of Lemma 11 the functions $\log J$ and $\log|f'|$ are homologous in $L^2(\mu)$. Hence it suffices to show that under the assumptions of the proposition $\log|f'| - h_\mu(f)$ is not homologous to zero in $L^2(\mu)$.

Suppose (14) holds for some $u \in L^2(\mu)$. Let $a \in A_0$ be such that there exist $n \in N$ and $z \in K \setminus \bigcup_{n=0}^{\infty} f^i(A_0)$ such that $f(z) = f^n(a)$. In view of Lemma 10 there exist constants $d \in \boldsymbol{R}$, $\alpha \in (0, 1)$ such that for μ-almost all y sufficiently close to $f^n(a)$

$$d + \log|y - f^n(a)|^\alpha \geqslant u(y). \tag{17}$$

Let $U = K \setminus \bar{V}$. By the topological exactness of f there exists a positive integer k such that $f^k(U) \supset K$. In particular, there exists $y \in U$ such that $f^k(y) = z$ and $y, f(y), \ldots, f^k(y) \notin A_0$. The set A is finite, so there is a small closed interval $J \subset U$ containing y such that J does not contain any critical point of f^{k+1} and f^{k+1} is continuous on J. In view of Lemmas 8 and 9, $u|_{f^{k+1}(J)}$ is equal μ-a.e. to a continuous function on $f^{k+1}(J)$. But this function is uniformly continuous, because $f^{k+1}(J)$ is a closed interval, in particular this function is bounded near $f^{k+1}(y) = f^n(a)$, which contradicts (17). ■

Proposition 2. *If the set $\bigcup_{n=1}^{\infty} f^i(A_0)$ is infinite, then $\sigma^2 > 0$.*

Proof. Suppose $\bigcup_{n=1}^{\infty} f^i(A_0)$ is infinite. In view of Propositon 1 it suffices to show that $f^{-1}(\bigcup_{n=1}^{\infty} f^i(A_0)) \cap K$ contains a point which does not belong to $\bigcup_{n=0}^{\infty} f^i(A_0)$.

Let $a \in A_0$ have infinite forward trajectory. Such an a exists because A_0 is finite. Let $\omega(a)$ be the set of its ω-limit points.

The set $\omega(a)$ is a closed set of μ-measure zero ([4, Lemma 3.7]), so we can find an open interval $U \subset K$ such that $\operatorname{dist}(U, \omega(a)) > \varepsilon$ for a small $\varepsilon > 0$. Let $n_0 \in N$ be so large that for every $n \geqslant n_0$, $\operatorname{dist}(f^n(a), \omega(a)) < \varepsilon$. By the topological transitivity there exists $m \in N$ such that $f^m(U) \supset K$. We claim that $f^{m+n_0}(a)$ has at least 2 preimages under f^m. Indeed, one of them belongs to U and the second is just $f^{n_0}(a)$ (which is close to $\omega(a)$ and far from U). Moreover, both of them belong to K.

Now, it is easy to see that one of the points $f^{n_0+j}(a)$, $j = 1, \ldots, m$, has at least 2 f-preimages, both in K. One of them does not lie on the trajectory of a (otherwise this trajectory would be finite), and we denote this point by b. But b can lie on the trajectory of $a_1 \in A_0$, $a_1 = f^n(a)$, $n \geqslant 0$. In this case we have $\omega(a_1) = \omega(a)$ and we can start with a_1 instead of a and find b_1 instead of b, and so on. This procedure must stop after a finite number of steps because A_0 is finite. ■

Theorem 2. *Assume f is unimodal. Then the following conditions are equivalent:*

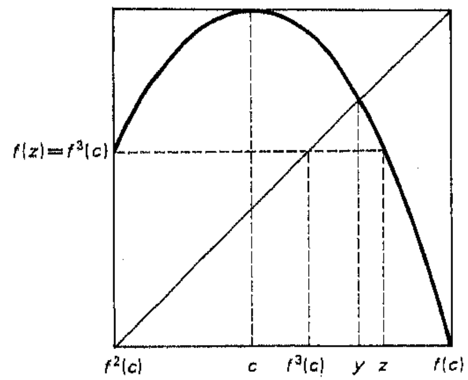
(A) $\sigma^2 = 0$.

(B) *(f, μ) is isomorphic to the one-sided Bernoulli shift with the initial distribution* $(\frac{1}{2}, \frac{1}{2})$.

(C) $h_\mu(f) = \log 2$.

Proof. The implication (B) $\Rightarrow$ (A) is obvious. We will prove (A) $\Rightarrow$ (C). Assume $\sigma^2 = 0$. It is known that $K = [f^2(c), f(c)]$, where c is the critical point of f (all the time we assume that the system is weakly mixing). In view of Proposition 2 the trajectory of c is finite, i.e. $f^k(c) = f^{k+n}(c)$ for some positive integers k, n. If k is the smallest positive integer with this property, then in view of Proposition 1, $f^k(c)$ is the unique point on the trajectory of c which has two f-preimages.

We will show $f^2(c) = f^3(c)$. If $f^2(c) \neq f^3(c)$ (see the figure), then $f^3(c)$



has two f-preimages, $f^2(c)$ and $z \neq f^2(c)$. If $z \leqslant f^3(c)$ then $f([f^2(c), y]) = [y, f(c)]$ and $f([y, f(c)]) = [f^2(c), y]$, where y is a fixed point. This is impossible by our assumption of weak mixing. Therefore $z > f^3(c)$. Then z also has two f-preimages and we have a contradiction. Hence $f^2(c) = f^3(c)$.

We have shown in the proof of Proposition 1 that the assumption $\sigma^2 = 0$ implies (14) for some $u \in L^2(\mu)$. Hence we have

$$\log|(f^n)'| = nh_\mu(f) + u \circ f^n - u \qquad \mu\text{-a.e.}$$

or

$$|(f^n)'| = \exp(nh_\mu(f))\exp(u \circ f^n)\exp(-u) \qquad \mu\text{-a.e.}$$

It follows from Lemmas 9, 10 that $\exp u$ is bounded and $\exp(-u)$ is integrable on the whole interval I. Therefore there exists $c>0$ such that

$$\int |(f^n)'|\, d\lambda \leqslant c \exp\big(nh_\mu(f)\big).$$

Hence

$$n^{-1}\log\int |(f^n)'|\, d\lambda \leqslant n^{-1}\log c + h_\mu(f).$$

The left side converges to the topological entropy of f which is equal to $\log 2$, while the right side converges to $h_\mu(f)$, which completes the proof of (C).

The implication (C) $\Rightarrow$ (B) is obvious. ■

THEOREM 3. *Assume that f satisfies* (i)–(vi) *and μ is weakly mixing. Then μ is a measure with maximal entropy for $f|_K$ iff $\sigma^2=0$.*

Proof. $\Rightarrow$ Assume that μ is a measure with maximal entropy. As noticed in the proof of [4, Th. 8.1], $h_\mu(f|_K)=h(f|_K)=\log\beta$ for some $\beta>1$, $f|_K$ is conjugate to a piecewise linear map g such that $|g'|=\beta$, and there exists a unique g-invariant probability measure ν absolutely continuous with respect to the Lebesgue measure λ. Denote the conjugacy by τ. Let I be the measure-theoretic jacobian of ν. Recall that the jacobian of μ is denoted by J. It is proved in the proof of [4, Th. 8.1] that under the assumption that μ is a measure with maximal entropy, we have $J=I\circ\tau$. But then $\log J$ is homologous to $h_\mu(f|_K)$, because I is obviously homologous to $h_\mu(f|_K)$ which is equal to $\log\beta$. Hence $\sigma^2=0$.

$\Leftarrow$ Assume $\sigma^2=0$. Then, in view of Proposition 2, the trajectory of the set of critical points in K is finite. Hence, by Lemmas 8–10, the function u satisfying (14) is such that $e^{u\circ f}$ is bounded and e^{-u} is integrable. Now the argument used in the proof of Theorem 2 yields

$$h(f|_K)=\lim_n n^{-1}\log\int |(f^n)'|\, d\lambda \leqslant h_\mu(f|_K). \ ■$$

Theorems 2 and 3 yield

COROLLARY. *For unimodal maps satisfying* (i)–(vi), *if the system (f, K, μ) is weakly mixing (i.e. the kneading sequence of f is indecomposable) and the absolutely continuous measure μ is also a measure with maximal entropy, then f is of the "fully developed chaos" type, i.e. $h(f)=\log 2$.*

This is a generalization of the results of [4, Section 9].

4. Refinement of the Shannon–McMillan–Breiman Theorem. In this section we will prove the main result of the paper. Put

$$R_t(x)=\log\mu\big(A_{[t]}(x)\big)+[t]\,h_\mu(f).$$

THEOREM 4. *If σ^2 is as in the previous sections and $\sigma^2>0$, then without changing its distribution one can redefine the process $(R_t)_{t\geqslant 0}$ on a richer probability space with standard Brownian process $(B_t)_{t\geqslant 0}$ such that almost surely*

$$|R_t/\sigma-B(t)|=O(t^{1/2-\lambda}) \quad \textit{for some } \lambda>0.$$

Proof. In view of Theorem 1 it suffices to prove that $|R_t-S(t)|=O(t^{1/2-\lambda})$ almost surely, where λ and $S(t)$ are as in Theorem 1. We will prove even more:

$$|R_t-S(t)|=O(t^\alpha) \quad \mu\text{-a.e. for every } \alpha\in(0, \tfrac{1}{2}).$$

Fix $\alpha\in(0,\frac12)$. We have

$$R_t(x)-S(t)(x)=\log\big(J_{[t]}(x)\,\mu(A_{[t]}(x))\big),$$

so it suffices to prove that for μ-almost every x

$$\log\big(J_n(x)\,\mu(A_n(x))\big)=O(n^\alpha) \quad \forall n\in N.$$

In [7] we have defined for any $m, M\in N$ the family $\mathscr{A}_{m,M}$ of "good atoms" of $\bigvee_{i=0}^{m+M-1} f^{-i}\mathscr{A}'$, where $\mathscr{A}'$ is some partition of I finer that $\mathscr{A}$ [7, Def. 4]. We can assume that the endpoints of atoms of $\mathscr{A}'$ belong to $\bigcup_{i=0}^{N-1} f^{-i}(A)$ for some $N\in N$ (cf. [8, Section 4]). Hence the partition $\bigvee_{i=0}^{m+M+N-1} f^{-i}\mathscr{A}$ is finer than $\bigvee_{i=0}^{m+M-1} f^{-i}\mathscr{A}'$ and we can define the family $\mathscr{B}_{m,M+N}$ of "good atoms" of $\bigvee_{i=0}^{m+M+N-1} f^{-i}\mathscr{A}$ in the natural way as the family of all atoms of $\bigvee_{i=0}^{m+M+N-1} f^{-i}\mathscr{A}$ which are contained in atoms of $\mathscr{A}_{m,M}$.

In view of [7, Lemmas 10, 12, 13] there exist constants $L>0$, $\eta\in(0,1)$ such that for every $\Delta\in\mathscr{B}_{m,M+N}$

$$\left|\frac{J_m(x)}{J_m(y)}-1\right|\leqslant L\eta^{M+N} \quad \forall x, y\in\Delta, \tag{18}$$

$$\mu\big(\bigcup\mathscr{B}_{m,M+N}\big)\geqslant 1-L\eta^{M+N}. \tag{19}$$

Let n be such that $n^\alpha>N$. We have

$$\big|\log\big(J_n(x)\,\mu(A_n(x))\big)\big|\leqslant\left|\log\frac{J_{n-[n^\alpha]}(x)\,\mu(A_n(x))}{\mu\big(A_{[n^\alpha]}(f^{n-[n^\alpha]}(x))\big)}\right|$$
$$+\big|\log\mu\big(A_{[n^\alpha]}(f^{n-[n^\alpha]}(x))\big)\big|+\big|\log J_{[n^\alpha]}(f^{n-[n^\alpha]}(x))\big|.$$

If $A_n(x)$ is a "good atom", namely belongs to $\mathscr{B}_{n-[n^\alpha],[n^\alpha]}$, then $f^{n-[n^\alpha]}(A_n(x))=A_{[n^\alpha]}(f^{n-[n^\alpha]}(x))$ and in view of (18) the first summand does not exceed $\eta^{[n^\alpha]}$ multiplied by a positive constant.

To estimate the second summand notice that for a good atom $A_n(x)$

$$\mu\big(A_{[n^\alpha]}(f^{n-[n^\alpha]}(x))\big)\geqslant d\lambda\big(A_{[n^\alpha]}(f^{n-[n^\alpha]}(x))\big),$$

where $d = \inf_K \varphi > 0$, and by the definition of good atoms, $A_{[n^\alpha]}(f^{n-[n^\alpha]}(x))$ is of the form (v, w), where $f^k(v), f^l(w) \in A$ for some $k, l \in \{0, 1, \ldots, [n^\alpha]-1\}$ (see [7, Def. 4 (c)]). Let $s = \max\{k, l\}$. In view of (v) the length of $f^s(A_{[n^\alpha]}(f^{n-[n^\alpha]}(x)))$ is not smaller than some positive constant R. Therefore

$$\lambda\big(A_{[n^\alpha]}(f^{n-[n^\alpha]}(x))\big) \geqslant \frac{R}{(\sup_I |f'|)^s} \geqslant \frac{R}{(\sup_I |f'|)^{n^\alpha}}$$

and the second summand does not exceed n^α multiplied by a positive constant.

Now we will estimate the third summand. We have

$$\log J_{[n^\alpha]}(f^{n-[n^\alpha]}(x)) = \sum_{j=n-[n^\alpha]}^{n-1} J \circ f^j(x).$$

In view of Lemma 1 all moments of the variables $J \circ f^j$ are bounded uniformly with respect to j. Therefore, by the Borel–Cantelli lemma, for every $\varepsilon \in (0, 1)$

$$J \circ f^j(x)/j^\varepsilon \to 0 \quad \text{for } \mu\text{-almost every } x \in I.$$

Take $\varepsilon > 0$ such that $\alpha + \varepsilon < \frac{1}{2}$. For μ-a.e. x there exists a positive constant $Z(x)$ such that $J \circ f^j(x)/j^\varepsilon$ is bounded by $Z(x)$ for sufficiently large j. Hence, if n is sufficiently large, then for μ-almost every x

$$\sum_{j=n-[n^\alpha]}^{n-1} J \circ f^j(x) = n^\varepsilon \sum_{j=n-[n^\alpha]}^{n-1} J \circ f^j(x)/n^\varepsilon$$

$$\leqslant n^\varepsilon \sum_{j=n-[n^\alpha]}^{n-1} J \circ f^j(x)/j^\varepsilon \leqslant n^{\varepsilon+\alpha} Z(x).$$

We have proved that for μ-almost every $x \in \bigcup \mathscr{B}_{n-[n^\alpha],[n^\alpha]}$, $|\log(J_n(x)\,\mu(A_n(x)))|$ does not exceed $n^{\alpha+\varepsilon}$ multiplied by a positive constant depending only on x. In view of (19) and the Borel–Cantelli lemma, the same is true for μ-almost every $x \in I$. ■

Remark 2. Theorem 4 implies integral tests, weak invariance principles and log-log laws for the process $(R_t)_{t\geqslant 0}$. Hence, if $\sigma^2 > 0$ then $(R_t)_{t\geqslant 0}$ satisfies the law of iterated logarithm which implies that the sequence $-n^{-1}\log\mu(A_n(x))$ converges to $h_\mu(f)$ not faster than $\sqrt{(\log\log n)/n}$. This is so e.g. for all maps f with infinite trajectory of critical points (Corollary 1) and for all unimodal maps except of fully developed chaos with $h_\mu(f) = \log 2$ (Theorem 2).

References

[1] M. Denker and G. Keane, *Finitary codes and the law of the iterated logarithm*, Z. Wahrsch. Verw. Gebiete 52 (1980), 321–331.
[2] F. Hofbauer and G. Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Math. Z. 180 (1982), 119–140.
[3] I. A. Ibragimov, *Some limit theorems for stationary processes*, Teor. Veroyatnost. i Primenen. 7 (1962), 361–392 (in Russian).
[4] M. Misiurewicz, *Absolutely continuous measures for certain maps of an interval*, Publ. Math. IHES 53 (1981), 17–51.
[5] W. Philipp and W. F. Stout, *Almost sure invariance principles for sums of weakly dependent random variables*, Mem. Amer. Math. Soc. 161 (1975).
[6] F. Przytycki, M. Urbański and A. Zdunik, *Harmonic, Gibbs and Hausdorff measures on repellers for holomorphic maps*, preprint, Univ. of Warwick, 1986.
[7] K. Ziemian, *Almost sure invariance principle for some maps of an interval*, Ergodic Theory Dynamical Systems 5 (1985), 625–640.
[8] —, *Rate of convergence of conditional entropies for some maps of an interval*, this volume.

INSTYTUT MATEMATYKI UNIWERSYTETU WARSZAWSKIEGO
INSTITUTE OF MATHEMATICS, WARSAW UNIVERSITY
PKiN IX p., 00-901 Warszawa, Poland

Received February 1, 1988 (2403)