Ergod. Th. & Dynam. Sys. (2014), **34**, 534–542 doi:10.1017/etds.2012.146
© *Cambridge University Press, 2012*

Quantitative ergodic theorems for weakly integrable functions

ALAN HAYNES

School of Mathematics, University of Bristol, Bristol, UK
(e-mail: alan.haynes@bristol.ac.uk)

(*Received* 7 *December* 2011 *and accepted in revised form* 12 *September* 2012)

Abstract. Under suitable hypotheses we establish a quantitative pointwise ergodic theorem which applies to trimmed Birkhoff sums of weakly integrable functions.

1. *Introduction*

Suppose that T is an ergodic measure-preserving transformation of a probability space (X, μ) and that f is a non-negative measurable function on X. If $f \in \mathrm{L}^1(X, \mu)$, then the Birkhoff ergodic theorem [**13**, Theorem 1.14] tells us that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) = \int_X f \, d\mu \quad \text{for almost every } x \in X.$$

Several mathematicians have contemplated, in different contexts, the possibility of extending this theorem to non-integrable functions. In this more general setup the problem is to determine whether there is a sequence $\{F_N\} \subseteq \mathbb{R}$ such that

$$\lim_{N\to\infty} \frac{1}{F_N} \sum_{n=0}^{N-1} f(T^n x) = 1 \quad \text{for } \mu\text{-almost every } x \in X. \tag{1}$$

Note that since T is ergodic, if the limit on the left-hand side of (1) exists then it must be constant almost everywhere [**13**, Theorem 1.6]. Aaronson showed in [**1**, Theorem 1] that if T is any ergodic transformation of a probability space and f is a non-negative measurable function for which (1) is satisfied for some choice of $\{F_N\}$, then f must be integrable.

On the other hand it was shown by Diamond and Vaaler [**6**] that for the special case of simple continued fractions (i.e. the Gauss map on $\mathbb{R}/\mathbb{Z}$), the ergodic theorem can be extended to weakly integrable functions by removing at most one term from each of the Birkhoff sums. This result was recently generalized by Aaronson, Nakada, and Natsui [**2, 10, 11**]. It is worth mentioning that these advances were preceded by analogous results due to Mori [**8, 9**] concerning the classical law of large numbers for independent and identically distributed (i.i.d.) random variables.

https://doi.org/10.1017/etds.2012.146 Published online by Cambridge University Press

The goal of this paper is to present a quantitative version of these results which applies to a large class of well-known dynamical systems. Our results do require estimates on the rate of mixing of the transformations involved, but such hypotheses are to be expected, as for many dynamical systems there is no hope of obtaining the types of theorems that we are presenting (cf. [**4, 5, 7**]). To strengthen this point we demonstrate in Theorem 4 that even the hypothesis of strong mixing alone is not enough to guarantee the results in our first two theorems.

Let $\mathcal{A} = \{A_i\}_{i\in\mathbb{N}}$ be a measurable partition of a probability space (X, μ) and let T be a measure-preserving transformation of the space. Further suppose that there is a function $g : \mathbb{N} \to [0, \infty)$ with $g(N) = o(N)$ and

$$\sum_{n=0}^{N}\left(\frac{\mu(A_i \cap T^{-n}A_j)}{\mu(A_i)\mu(A_j)} - 1\right) \le g(N) \quad \text{for all } i, j, N \in \mathbb{N}. \tag{2}$$

As a special case of one of our main results, we obtain the following.

THEOREM 1. *Suppose that (X, μ), T, $\mathcal{A}$, and g are as above and that $f : X \to [0, \infty)$ is measurable with respect to the σ-algebra generated by $\mathcal{A}$. Given $\epsilon > 0$ let*

$$F(N) = \int_{\{f \le N \log^{1/2+\epsilon} N\}} f\, d\mu \quad \text{and} \quad G(N) = \sum_{n=1}^{N} g(n).$$

If $g(N) = o(N^{1/2})$ and $F(N + 1) - F(N) \ll 1/N$ then there is a function $\delta_\epsilon : X \times \mathbb{N} \to \{0, 1\}$ such that for almost every $x \in X$,

$$\sum_{n=0}^{N-1} f(T^n x) = NF(N) + \delta_\epsilon(x, N) \cdot \max_{0\le n<N} f(T^n x) + O_\epsilon((N + G(N))^{2/3} \log^{5/3+\epsilon} N).$$

For a large class of Markov maps and invariant measures this theorem gives an error term of $O(N^{2/3} \log^{5/3+\epsilon} N)$. This will certainly hold for any fibred system with continued fraction mixing, examples of which can be found in [**3**] and [**11**, §3]. However, the reader should note that exponential decay of correlations is not, *a priori*, a requirement for (2) to hold. What we are asking for is analogous to a weak-mixing version of the continued-fraction mixing property. Furthermore, we obtain a non-trivial error term in Theorem 1 even in cases when the sum of $g(n)$ does not converge. It would be interesting to produce examples of measure-preserving dynamical systems satisfying (2), which do not possess the continued fraction mixing property. However, as in the case of weak mixing maps which are not strong mixing, such examples do not appear to be trivial to construct. We will return to this in a later paper but for now we maintain focus on the problem at hand. As a prototypical application of Theorem 1 we give the following quantitative version of [**6**, Corollary 1] (see also [**12**]).

COROLLARY 1. *For any $\epsilon > 0$ there is a function $\delta_\epsilon : \mathbb{R} \times \mathbb{N} \to \{0, 1\}$ such that for almost every $x \in \mathbb{R}$ (with respect to Lebesgue measure),*

$$\sum_{n=1}^{N} a_n(x) = \frac{N}{\log 2} \sum_{n\le N \log^{1/2+\epsilon} N} n \log\left(1 + \frac{1}{n(n+2)}\right) + \delta_\epsilon(x, N) \cdot \max_{1\le n\le N} a_n(x) + O_\epsilon(N^{2/3} \log^{5/3+\epsilon} N),$$

https://doi.org/10.1017/etds.2012.146 Published online by Cambridge University Press

where $x = [a_0(x); a_1(x), a_2(x), \ldots]$ *denotes the simple continued fraction expansion of* x.

One interesting feature of these results is that the main terms vary with ϵ in a subtle way. For example, in the continued fractions result the first main term starts off as

$$\frac{N \log N}{\log 2} + \frac{(1/2 + \epsilon) N \log\log N}{\log 2},$$

but the error term is of much smaller order. The explanation is that this dependence on ϵ is balanced by the choice of the cutoff function δ_ϵ.

To state our main results in full generality, for any increasing function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ we define $\mathrm{L}^{\phi,w} = \mathrm{L}^{\phi,w}(X, \mu)$ to be the collection of measurable functions $f : X \to \mathbb{R}$ for which

$$K_\phi(f) = \sup_{\lambda > 0} \phi(\lambda) \cdot \mu\{x \in X : |f(x)| > \lambda\} < \infty.$$

If $\phi(\lambda) = \lambda$ then $\mathrm{L}^{\phi,w}$ is the usual space of weakly integrable functions.

THEOREM 2. *Suppose that* (X, μ), T, $\mathcal{A}$, *and* g *satisfy (2) and suppose that* $f \in \mathrm{L}^{\phi,w}$ *is non-negative and measurable with respect to* $\mathcal{A}$. *Given* $\epsilon > 0$ *set* $\tau_\phi(N) = \phi^{-1}(N \log^{1/2+\epsilon} N)$, *and let*

$$F_1(N) = \int_{\{f \le \tau_\phi(N)\}} f \, d\mu, \quad F_2(N) = \int_{\{f \le \tau_\phi(N)\}} f^2 \, d\mu,$$

$$G(N) = \sum_{n=1}^{N} g(n) \quad \text{and} \quad F_3(N) = F_1(N)^2 (N + G(N)) + F_2(N).$$

If

$$(N+1)F_1(N+1) - NF_1(N) \ll F_3(N+1) - F_3(N) \ll F_3(N)^{2/3}$$

then there is a function $\delta_\epsilon : X \times \mathbb{N} \to \{0, 1\}$ *such that for almost every* $x \in X$

$$\sum_{n=0}^{N-1} f(T^n x) = NF_1(N) + \delta_\epsilon(x, N) \cdot \max_{0 \le n < N} f(T^n x) + O_\epsilon(F_3(N)^{2/3} \log^{1/3+\epsilon} F_3(N)).$$

In practice, the hypotheses on F_1 and G are easy to check. The main differences between our results and those in [**2, 10, 11**] are that ours are quantitative and, although our hypotheses on $g(N)$ are in general stronger than those in [**2**, Theorem 1.1], we only require the removal of at most one term from each Birkhoff sum.

Theorem 2 will be derived from the following quantitative result for sums of Birkhoff sums.

THEOREM 3. *Suppose that* (X, μ), T, *and* $\mathcal{A}$ *are as above and that (2) is satisfied. Let* $\{f_n\}_{n \in \mathbb{N}}$ *be a sequence of non-negative functions in* L^2, *measurable with respect to* $\mathcal{A}$, *and for each* $N \in \mathbb{N}$ *let*

$$F_1(N) = \int_X \sum_{n=1}^{N} f_n \, d\mu, \quad F_2(N) = \int_X \left(\sum_{n=1}^{N} f_n \right)^2 d\mu,$$

$$G(N) = \sum_{n=1}^{N} g(n) \quad \text{and} \quad F_3(N) = F_1(N)^2 (N + G(N)) + F_2(N).$$

https://doi.org/10.1017/etds.2012.146 Published online by Cambridge University Press

If

$$(N+1)F_1(N+1) - NF_1(N) \ll F_3(N+1) - F_3(N) \ll F_3(N)^{2/3} \tag{3}$$

then for any $\epsilon > 0$ and for almost every $x \in X$ we have that

$$\sum_{m=0}^{N-1}\sum_{n=1}^{N} f_n(T^m x) = NF_1(N) + O_\epsilon(F_3(N)^{2/3}\log^{1/3+\epsilon} F_3(N)).$$

From Theorem 3 we also obtain a quantitative version of the classical ergodic theorem.

COROLLARY 2. *If (2) is satisfied, if $f \in \mathrm{L}^2$ is measurable with respect to $\mathcal{A}$, and if*

$$g(N+1) \ll (N + G(N))^{2/3}, \tag{4}$$

then for almost every $x \in X$ we have that

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) = \int_X f\,d\mu + O_\epsilon(N^{-1}(N+G(N))^{2/3}\log^{1/3+\epsilon}(N+G(N))).$$

We conclude with a discussion of how our hypotheses on T compare with the usual notions of ergodicity and mixing. On the one hand it is easy to see that if $\mathcal{A}$ is finite, then (2) will be satisfied whenever T is ergodic. On the other hand when $\mathcal{A}$ is infinite, then even strong mixing is not enough to ensure that (2) is satisfied. Furthermore, as the following theorem demonstrates, the hypothesis of strong mixing by itself is not strong enough to draw the conclusion in Theorem 2.

THEOREM 4. *For any choice of $\{F_N\} \subseteq \mathbb{R}^+$ and $\delta : \mathbb{R}/\mathbb{Z} \times \mathbb{N} \to \{0, 1\}$, the quantities*

$$\frac{1}{F_N}\left(\sum_{n=0}^{N-1}\left\lfloor \frac{1}{\{2^n x\}}\right\rfloor - \delta(x, N)\cdot \max_{0\le n<N}\left\lfloor \frac{1}{\{2^n x\}}\right\rfloor\right)$$

either diverge for almost every $x \in \mathbb{R}/\mathbb{Z}$ or they converge to zero almost everywhere (with respect to Lebesgue measure), as $N \to \infty$.

This example shows the importance of the uniformity in hypothesis (2).

2. *Proofs of Theorem 3 and Corollary 2*

For each $N \in \mathbb{N}$ let

$$\Phi_N(x) = \sum_{m=0}^{N-1}\sum_{n=1}^{N} f_n(T^m x).$$

Since T is a measure-preserving transformation we have that

$$\int_X \Phi_N\,d\mu = NF(N).$$

For the second moments we have

$$\int_X \Phi_N^2\,d\mu = \sum_{n,n'=1}^{N}\sum_{m,m'=0}^{N-1}\int_X f_n(T^m x) f_{n'}(T^{m'} x)\,d\mu. \tag{5}$$

https://doi.org/10.1017/etds.2012.146 Published online by Cambridge University Press

Now for each i, $n \in \mathbb{N}$ let us denote the value of f_n on A_i by $\alpha_{n,i}$, so that

$$\begin{aligned}
&\sum_{m=0}^{N-1}\sum_{m'=m}^{N-1}\int_X f_n(T^m x) f_{n'}(T^{m'}x)\,d\mu \\
&\quad=\sum_{m=0}^{N-1}\sum_{m'=0}^{N-1-m}\int_X f_n(x) f_{n'}(T^{m'}x)\,d\mu \\
&\quad=\sum_{\substack{i,j=1\\ i\neq j}}^{\infty}\alpha_{n,i}\alpha_{n',j}\sum_{m=0}^{N-1}\sum_{m'=0}^{N-1-m}\mu(A_i\cap T^{-m'}A_j) \qquad (6)\\
&\qquad+\sum_{i=1}^{\infty}\alpha_{n,i}\alpha_{n',i}\sum_{m=0}^{N-1}\sum_{m'=0}^{N-1-m}\mu(A_i\cap T^{-m'}A_i). \qquad (7)
\end{aligned}$$

Using (2) in (6) gives

$$\begin{aligned}
&\sum_{\substack{i,j=1\\ i\neq j}}^{\infty}\alpha_{n,i}\alpha_{n',j}\sum_{m=0}^{N-1}\sum_{m'=0}^{N-1-m}\mu(A_i\cap T^{-m'}A_j)\\
&\quad=\sum_{\substack{i,j=1\\ i\neq j}}^{\infty}\alpha_{n,i}\alpha_{n',j}\mu(A_i)\mu(A_j)\sum_{m=0}^{N-2}(N-m+O(g(N-m))). \qquad (8)
\end{aligned}$$

It is worth pointing out that here we have also used the fact that the summand corresponding to $m = N-1$ is zero, since $i \neq j$. For (7) we have that

$$\begin{aligned}
&\sum_{i=1}^{\infty}\alpha_{n,i}\alpha_{n',i}\sum_{m=0}^{N-1}\sum_{m'=0}^{N-1-m}\mu(A_i\cap T^{-m'}A_i)\\
&\quad=\sum_{i=1}^{\infty}\alpha_{n,i}\alpha_{n',i}\mu(A_i)^2\sum_{m=0}^{N-2}(N-m+O(g(N-m)))+\int_X f_n f_{n'}\,d\mu. \qquad (9)
\end{aligned}$$

Combining (8) and (9) gives

$$\begin{aligned}
&\sum_{m=0}^{N-1}\sum_{m'=m}^{N-1}\int_X f_n(T^m x) f_{n'}(T^{m'}x)\,d\mu\\
&\quad=\int_X f_n\,d\mu\cdot\int_X f_{n'}\,d\mu\cdot\left(\frac{N^2+N}{2}+O(G(N))\right)+\int_X f_n f_{n'}\,d\mu,
\end{aligned}$$

and adding this to the complementary sum in (5) (i.e. with $m' \leq m$) we have that

$$\begin{aligned}
&\int_X(\Phi_N - NF_1(N))^2\,d\mu\\
&\quad\leq\sum_{n,n'=1}^{N}\left(\sum_{\substack{m,m'=0\\ m\leq m'}}^{N-1}+\sum_{\substack{m,m'=0\\ m'\leq m}}^{N-1}\right)\int_X f_n(T^m x) f_{n'}(T^{m'}x)\,d\mu - N^2F_1(N)^2\\
&\quad\ll F_1(N)^2(N+G(N))+F_2(N)=F_3(N).
\end{aligned}$$

https://doi.org/10.1017/etds.2012.146 Published online by Cambridge University Press

Now for each $j \in \mathbb{N}$ let N_j be the smallest integer with

$$F_3(N_j) > j^3 \log^{1+\epsilon}(j+1).$$

By Chebyshev's inequality and hypothesis (3) we have that

$$\sum_{j=1}^{\infty} \mu\{x \in X : |\Phi_{N_j}(x) - N_j F_1(N_j)| > j^2 \log^{1+\epsilon}(j+1)\} \ll \sum_{j=1}^{\infty} \frac{1}{j \log^{1+\epsilon}(j+1)} < \infty,$$

and the Borel–Cantelli lemma allows us to conclude that for almost every $x \in X$

$$\Phi_{N_j}(x) = N_j F_1(N_j) + O_\epsilon(j^2 \log^{1+\epsilon}(j+1)) \quad \text{as } j \to \infty. \tag{10}$$

Since the functions f_n are non-negative we have that $\Phi_{N_j} \le \Phi_N \le \Phi_{N_{j+1}}$ whenever $N_j \le N \le N_{j+1}$. Using (3) we see that

$$\begin{aligned} &N_{j+1}F_1(N_{j+1}) - N_j F_1(N_j) \\ &\quad \ll F_3(N_{j+1}) - F_3(N_j) \\ &\quad = (j+1)^3 \log^{1+\epsilon}(j+2) - j^3 \log^{1+\epsilon}(j+1) + O(j^2 \log^{2/3(1+\epsilon)}(j+1)) \\ &\quad \ll j^2 \log^{1+\epsilon}(j+1). \end{aligned}$$

Thus, we can interpolate between N_j and N_{j+1} in (10) to obtain

$$\Phi_N(x) = N F_1(N) + O_\epsilon(j^2 \log^{1+\epsilon}(j+1)), \quad N_{j-1} \le N < N_j,$$

for almost every $x \in X$. Finally for $N \ge N_{j-1}$ we have that

$$j^2 \log^{1+\epsilon}(j+1) \ll F_3(N)^{2/3} \log^{1/3+\epsilon} F_3(N),$$

which finishes the proof.

Corollary 2 follows directly from Theorem 3 by setting $f_1 = f$ and $f_n = 0$ for $n \ge 2$. In this case $F_1(N)$ is constant and thus (3) is satisfied if and only if (4) is.

3. *Proofs of Theorems 1 and 2 and Corollary 1*

The proof of Theorem 2 requires one more important piece of information.

LEMMA 1. *Assume that the hypotheses of Theorem 2 are satisfied and that $f \notin \mathrm{L}^1$. Then for almost every $x \in X$ there are at most finitely many $N \in \mathbb{N}$ for which*

$$\#\{0 \le n < N : f(T^n x) > \tau_\phi(N)\} > 1.$$

Proof. For each $N \in \mathbb{N}$ let

$$B_N = \sum_{\substack{m,n=0 \\ m \ne n}}^{2N-1} \mu\{x \in X : f(T^m x) > \tau_\phi(N),\ f(T^n x) > \tau_\phi(N)\}.$$

Also if $N \in \mathbb{N}$, then let $\{i_j\}_{j \in \mathbb{N}}$ be the sequence of distinct integers for which

$$\{x \in X : f(x) > \tau_\phi(N)\} = \bigcup_{j=1}^{\infty} A_{i_j}.$$

https://doi.org/10.1017/etds.2012.146 Published online by Cambridge University Press

By our hypotheses on T we have that

$$\begin{aligned} B_N &= 2\sum_{m=0}^{2N-2}\sum_{n=m+1}^{2N-1}\mu\{f(x)>\tau_\phi(N),\ f(T^{n-m}x)>\tau_\phi(N)\} \\ &= 2\sum_{j,k=1}^{\infty}\sum_{m=0}^{2N-2}\sum_{n=1}^{2N-1-m}\mu\{A_{i_j}\cap T^{-n}A_{i_k}\} \\ &= 2\sum_{j,k=1}^{\infty}\mu(A_{i_j})\mu(A_{i_k})\sum_{m=0}^{2N-2}(2N-m-1+O(g(2N-m))) \\ &\ll \mu\{f(x)>\tau_\phi(N)\}^2N^2. \end{aligned}$$

Since $f\in \mathrm{L}^{\phi,w}$ we deduce that

$$B_N \ll (\log N)^{-(1+\epsilon)}.$$

Now since

$$\sum_{k=1}^{\infty} B_{2^k} < \infty,$$

the Borel–Cantelli lemma guarantees that for almost every $x\in X$ there are only finitely many $k\in\mathbb{N}$ for which $f(T^m x)>\tau_\phi(2^k)$ and $f(T^n x)>\tau_\phi(2^k)$ for two distinct integers $0\le m<n<2^{k+1}$. Now the proof is finished with the observation that if $2^k\le N<2^{k+1}$, then $\tau_\phi(N)\ge\tau_\phi(2^k)$. □

Note that the proof of Lemma 1 used both (2) and the hypothesis that $f\in \mathrm{L}^{\phi,w}$. It appears that neither of these conditions can be weakened. For the proof of Theorem 2 we now apply Theorem 3 with

$$f_n(x)=\begin{cases} f(x) & \text{if } \tau_\phi(n-1)<f(x)\le\tau_\phi(n),\\ 0 & \text{otherwise.}\end{cases}$$

Then the functions $F_1(N)$ in the statements of both theorems are the same. If $f\in \mathrm{L}^1$, then Theorem 2 follows directly from Theorem 3. If $f\notin \mathrm{L}^1$, then Theorem 2 is a consequence of Theorem 3 together with Lemma 1.

For the proof of Theorem 1 first note that under the hypotheses there,

$$\begin{aligned} &N(\log^{1/2+\epsilon}N)\cdot\mu\{N\log^{1/2+\epsilon}N<f\le(N+1)\log^{1/2+\epsilon}(N+1)\} \\ &\quad\le F(N+1)-F(N) \\ &\quad\ll 1/N. \end{aligned}$$

It follows easily from this that $f\in \mathrm{L}^{\lambda,w}$ and that

$$\int_{\{f\le N\log^{1/2+\epsilon}N\}} f\,d\mu \ll \log N$$

and

$$\int_{\{f\le N\log^{1/2+\epsilon}N\}} f^2\,d\mu \ll N\log^{1/2+\epsilon}N.$$

https://doi.org/10.1017/etds.2012.146 Published online by Cambridge University Press

These observations together with the assumption that $g(N) = o(N^{1/2})$ easily imply that both inequalities in (3) are satisfied. Theorem 1 then follows directly from Theorem 2.

Corollary 1 follows Theorem 1 by considering the Birkhoff sums of the function $f_g : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$, $f_g(x) = a_1(x)$ with respect to the Gauss map $T_g : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$, $T_g(x) = [0; a_2(x), a_3(x), \ldots]$, which is ergodic with respect to the absolutely continuous Borel measure μ_g defined by

$$\mu_g(E) = \frac{1}{\log 2} \int_E \frac{dx}{1+x}.$$

In this case it is well known that $g(N) \ll 1$, and the result follows immediately.

4. *Proof of Theorem 4*

Let $X = \mathbb{R}/\mathbb{Z}$ be equipped with Lebesgue measure, and let $T : X \to X$ be the map $T(x) = 2x$. This is a strong mixing measure-preserving transformation of $\mathbb{R}/\mathbb{Z}$ [**13**, Theorems 1.10, 1.28]. Let $f = f_g$ be the function defined at the end of the previous section and note that this function is in weak-L^1 but not in L^1. Therefore, by Aaronson's theorem [**1**, Theorem 1], for any $\{F_N\} \subseteq \mathbb{R}^+$ we have that

$$\begin{aligned} &\liminf_{N\to\infty} \frac{1}{F_N} \sum_{n=0}^{N-1} f(T^n x) = 0 \text{ almost everywhere} \quad \text{or} \\ &\limsup_{N\to\infty} \frac{1}{F_N} \sum_{n=0}^{N-1} f(T^n x) = \infty \text{ almost everywhere (or both).} \end{aligned} \tag{11}$$

Now suppose that $\{F_N\} \subseteq \mathbb{R}$ and $\delta : \mathbb{R}/\mathbb{Z} \times \mathbb{N} \to \{0, 1\}$ are chosen so that

$$\frac{1}{F_N}\left(\sum_{n=0}^{N-1} f(T^n x) - \delta(x, N) \cdot \max_{0\le n<N} f(T^n x)\right) \tag{12}$$

converges almost everywhere to $f^* : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$, as $N \to \infty$. Then f^* is measurable and T invariant and it follows from [**13**, Theorem 1.6] that $f^*(x) = c$ almost everywhere for some $c \in \mathbb{R}$. Suppose that $c \neq 0$, and let $B \subseteq \mathbb{R}/\mathbb{Z}$ be the set of x for which (11) holds but (12) converges to c. Then B is a set of measure one and since $c \neq 0$ it must be the second of the two options from (11) which holds for all $x \in B$. Choose $x \in B$ and let $1 \le N_1 < N_2 < \cdots$ be the sequence of positive integers N for which $\delta(x, N)$ takes the value one. Then we must have that

$$\limsup_{j\to\infty} \frac{\max_{0\le n<N_j} f(T^n x)}{F_{N_j}} = \infty.$$

Since $F_N \gg N$ we have that

$$\max_{0\le n<N_j} f(T^n x) > 2$$

for all sufficiently large j. Thus, if we let $0 \le n_j < N_j$ denote the integer for which

$$f(T^{n_j} x) = \max_{0\le n<N_j} f(T^n x)$$

we have that

$$f(T^{n_j+1} x) = (1/2) \max_{0\le n<N_j} f(T^n x).$$

https://doi.org/10.1017/etds.2012.146 Published online by Cambridge University Press

Finally it is not difficult to show that $n_j < N_j - 1$ for infinitely many j (otherwise, we would have to have that $c = 0$). This means that for infinitely many $j \in \mathbb{N}$,

$$\frac{1}{F_{N_j}}\left(\sum_{n=0}^{N_j-1} f(T^n x) - \delta(x, N_j) \cdot \max_{0\le n<N_j} f(T^n x)\right) \ge \frac{\max_{0\le n<N_j} f(T^n x)}{2F_{N_j}}.$$

However, the right-hand side tends to infinity as $j \to \infty$, and this is a contradiction. Therefore, if (12) converges almost everywhere, then it must converge to zero.

Acknowledgements. The author would like to thank Jeffrey Vaaler, Tom Ward, and Manfred Einsiedler for helpful conversations concerning the results in this paper. Also thanks go to the referee for several helpful suggestions. This research was supported by EPSRC grant EP/J00149X/1.

References

[1] J. Aaronson. On the ergodic theory of non-integrable functions and infinite measure spaces. *Israel J. Math.* **27**(2) (1977), 163–173.
[2] J. Aaronson and H. Nakada. Trimmed sums for non-negative, mixing stationary processes. *Stochastic Process. Appl.* **104**(2) (2003), 173–192.
[3] J. Aaronson and H. Nakada. On the mixing coefficients of piecewise monotonic maps. *Israel J. Math.* **148** (2005), 1–10.
[4] J. Avigad. The metamathematics of ergodic theory. *Ann. Pure Appl. Logic* **157**(2–3) (2009), 64–76.
[5] J. Avigad, P. Gerhardy and H. Towsner. Local stability of ergodic averages. *Trans. Amer. Math. Soc.* **362**(1) (2010), 261–288.
[6] H. G. Diamond and J. D. Vaaler. Estimates for partial sums of continued fraction partial quotients. *Pacific J. Math.* **122**(1) (1986), 73–82.
[7] A. G. Kachurovskii. Rates of convergence in ergodic theorems. *Uspekhi Mat. Nauk* **51**(4(310)) (1996), 73–124, Engl. transl. *Russian Math. Surveys* **51**(4) (1996) 653–703.
[8] T. Mori. The strong law of large numbers when extreme terms are excluded from sums. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **36**(3) (1976), 189–194.
[9] T. Mori. Stability for sums of i.i.d. random variables when extreme terms are excluded. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **40**(2) (1977), 159–167.
[10] H. Nakada and R. Natsui. Some metric properties of α-continued fractions. *J. Number Theory* **97** (2002), 287–300.
[11] H. Nakada and R. Natsui. On the metrical theory of continued fraction mixing fibred systems and its application to Jacobi–Perron algorithm. *Monatsh. Math.* **138**(4) (2003), 267–288.
[12] W. Philipp. A conjecture of Erdos on continued fractions. *Acta Arith.* **28**(4) (1975/76), 379–386.
[13] P. Walters. *An Introduction to Ergodic Theory (Graduate Texts in Mathematics, 79)*. Springer, New York, 1982.

https://doi.org/10.1017/etds.2012.146 Published online by Cambridge University Press