# TRIMMED SUMS FOR NON-NEGATIVE, MIXING STATIONARY PROCESSES.

JON AARONSON, HITOSHI NAKADA

ABSTRACT. We consider the effect of "trimming" ergodic sums of their maximal values on the strong law of large numbers for non-negative, non-integrable, mixing stationary processes, and to study the ?cusp visits? of a certain interval map.

## §0 INTRODUCTION

**Laws of large numbers and sum trimming.** We consider non-negative, $\mathbb{R}$-valued ergodic, stationary processes $(X_1, X_2, \dots)$. In case $E(X_1) = \infty$, there is no strong law of large numbers for the partial sums $S_n := \sum_{k=1}^{n} X_k$.

It is shown in [Aar77] (see also [Aar97] §2.3) that if $b_n > 0$ are constants then,

$$(\bigstar) \qquad \text{either } \varlimsup_{n \to \infty} \frac{1}{b_n} S_n = \infty \text{ a.s., or } \lim_{n \to \infty} \frac{1}{b_n} S_n = 0 \text{ a.s. } .$$

See [Fel46] and [CR61] for the original proofs in the i.i.d. case.

There may be a weak law of large numbers when $E(X_1) = \infty$. Feller ([Fel45]) showed that if $(X_1, X_2, \dots)$ are non-negative, i.i.d. random variables, the weak law of large numbers holds in the sense that

$$(\clubsuit) \qquad \exists \, b(n) \text{ constants such that } \frac{S_n}{b(n)} \xrightarrow{P} 1$$

(where $\xrightarrow{P}$ denotes stochastic convergence) iff $L(t) := E(X \wedge t)$ is slowly varying at $\infty$ (see below) and in this case $b(n) \sim nL(b(n))$.

The strong law here breaks down in a particular way: since $E(X) = \infty \Rightarrow E(b^{-1}(X)) = \infty$, we have (by the Borel-Cantelli lemma)

$$(\spadesuit) \qquad \varlimsup_{n \to \infty} \frac{S_n}{b(n)} \geq \varlimsup_{n \to \infty} \frac{X_n}{b(n)} = \infty \text{ a.s..}$$

1

The question arose as to whether the maximal terms of $\{X_1, \dots, X_n\}$ are "responsible" for the failure of the strong law, particularly in view of the fact that under the additional assumption that $L(t) \sim L(t \log \log t)$ (as shown in [KT77]) ,

($\heartsuit$) $$\varliminf_{n \to \infty} \frac{S_n}{b(n)} = 1 \text{ a.s..}$$

Mori studied strong laws for i.i.d. random variables when finitely many of these maximal terms are excluded (trimmed) from the sums $S_n$ and characterised (in terms of the distribution of the $X_k$ and the normalising constants) when a trimmed strong law holds (see [Mor76] and [Mor77]).

In this paper, we consider such trimming for dependent processes, extending a theorem of Mori's (theorem 1.1 below) to certain continued fraction mixing processes (see below), and exhibiting Markov chains (satisfying ($\clubsuit$), ($\spadesuit$) and ($\heartsuit$)) for which it fails.

For simplicity, we restrict attention to non-negative processes, as in the general $\mathbb{R}$-valued case, there may be interaction of the positive and negative parts causing strong laws which are spurious from the viewpoint of this paper.

**Regular variation.** Recall (from [Kar33],[BGT87], [Fel66]) that a measurable function $f : \mathbb{R}_+ \to \mathbb{R}_+$ is called *regularly varying* (at $\omega = 0, \infty$) if $\forall \lambda > 0$, $\exists \lim_{t \to \omega} \frac{f(\lambda t)}{f(t)} =: \ell(\lambda)$. In case $f$ is regularly varying, the function $\ell$ is necessarily of form $\ell(\lambda) = \lambda^\alpha$ for some $\alpha \in \mathbb{R}$ which is called the *index* (of regular variation of $f$).

The function $f : \mathbb{R}_+ \to \mathbb{R}_+$ is called *slowly varying* at $\omega$ if it is regularly varying at $\omega$ with index 0, i.e. $\frac{f(\lambda t)}{f(t)} \to_{t \to \omega} 1 \; \forall \; \lambda > 0$. Write $E(X \wedge t) =: L(t)$ and set $\epsilon(t) := t(\log^+ L)'(t) = \frac{tc(t)}{L(t)}$ for large $t$ enough that $L(t) > 1$, where $c(t) := P(X > t) = L'(t)$.

Both $L$ and log are increasing and concave whence so is $\log L$, and $\frac{\epsilon(t)}{t}$ decreases in $t$ for $t$ large.

By Karamata's representation theorem ([Kar33], see also [BGT87], [Fel66]) $L(t) = E(X \wedge t)$ is slowly varying at $\infty$ iff $\epsilon(t) \to 0$ as $t \to \infty$.

We'll call an increasing function $A : \mathbb{R}_+ \to \mathbb{R}_+$ *weakly regularly varying* if

$$A(2t) \ll A(t), \; \& \; A^{-1}(2t) \ll A^{-1}(t)$$

equivalently $\exists \; M > 1$ such that $A(2t) \le M A(t)$, $\& \; 2A(t) \le A(Mt)$. A decreasing function $B : \mathbb{R}_+ \to \mathbb{R}_+$ will be so called if the increasing function $\frac{1}{B}$ is weakly regularly varying.

It can be shown that a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ which is regularly varying at $\infty$ with nonzero index is weakly regularly varying, whereas a slowly varying function cannot be weakly regularly varying.

**Dependence.** The asymptotic behaviours ($\clubsuit$), ($\spadesuit$) and ($\heartsuit$) persist when the assumption of independence is relaxed to that of continued fraction mixing; the stationary process $(X_1, X_2, \dots)$ being called *continued fraction mixing* (c.f.-mixing) if $\vartheta(1) < \infty$ and $\vartheta(n) \downarrow 0$ where

$$\vartheta(n) := \sup\{|\tfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)\mathbb{P}(B)} - 1| : \ A \in \sigma_1^k, \ B \in \sigma_{k+n}^\infty, \ \mathbb{P}(A)\mathbb{P}(B) > 0, \ k \geq 1\}.$$

Any probability preserving Gibbs-Markov map is c.f.- mixing with $\vartheta(n) \downarrow 0$ exponentially (see [AD01] or §4.7 of [Aar97]).

The proof of ($\spadesuit$) in the c.f.-mixing case is the same as in the i.i.d. case, but uses the strong Borel-Cantelli lemma of Renyi ([Ren70] p. 391). See [Aar86] and §5 of [AD90] for ($\clubsuit$); and [AD89] for ($\heartsuit$).

**Results.** Let $(X_1, X_2, \cdots)$ be a non-negative, ergodic stationary process with $E(X \wedge t) =: L(t)$. Set $a(t) := \frac{t}{L(t)}$ and $b := a^{-1}$.

Write $\{X_k\}_{k=1}^n = \{r_{n,k}\}_{k=1}^n$ where $r_{n,1} \geq r_{n,2} \geq \cdots \geq r_{n,n}$ and set $M_n^{(\nu)} := \sum_{k=1}^\nu r_{n,k}$.

Let (for $r > 0$) $J_r := \sum_{n=1}^\infty \frac{\epsilon(n)^r}{n}$ and define

$$\mathfrak{N}_X := \begin{cases} \min\{\kappa \in \mathbb{N} : \ J_{\kappa+1} < \infty\} & \text{if } \exists \ \kappa, \ J_\kappa < \infty, \\ \infty & \text{else.} \end{cases}$$

Note that $\mathfrak{N}_X < \infty$ implies that $L(t) := E(X \wedge t)$ is slowly varying at $\infty$.

**Theorem 1.1**
*Suppose that $(X_1, X_2, \cdots)$ is c.f.- mixing, then*

$$\varlimsup_{n \to \infty} \sum_{k=1}^n 1_{[X_k > tb(n)]} = \mathfrak{N}_X \leq \infty \ \forall \ t > 0. \tag{i}$$

(ii)    *Suppose that $\sum_{n=1}^\infty \frac{\vartheta(n)}{n} < \infty$, and that $\mathfrak{N}_X < \infty$, then $\exists \ b_n = o(b(n))$ (depending only on the distribution of $X$) such that*

$$S_n - M_n^{(\mathfrak{N}_X)} \sim S_n^{(b_n)} \sim b(n) \ a.s. \ as \ n \to \infty.$$

*where $S_n^{(b)} := \sum_{k=1}^n X_k \wedge b$.*

**Remarks**

1) It follows from (i) of theorem 1.1, that $\overline{\lim}_{n\to\infty}\frac{1}{b(n)}(S_n-M_n^{(K)})=\infty$ a.s. $\forall\ K<\mathfrak{N}_X$ and it follows from (ii) of theorem 1.1, that $\frac{1}{b(n)}(S_n-M_n^{(K)})\to 1$ a.s. $\forall\ K\geq\mathfrak{N}_X$.

2) It is not hard to show using Birkhoff's theorem, that if $(X_1,X_2,\cdots)$ is an ergodic, stationary process with $E(|X|)<\infty$, then $\frac{1}{n}(S_n-M_n^{(K)})\to E(X)$ a.s. $\forall\ K\in\mathbb{N}$.

In case $(X_1,X_2,\cdots)$ are i.i.d.r.v.'s, theorem 1.1 follows from theorem 1 in [Mor77]. The proof of theorem 1.1 (given in §1) differs from that of theorem 1 in [Mor77] mainly in the estimation of large deviation probabilities of truncated sums. The use of log-moment generating functions in [Mor77] is not possible here due to the dependence. We use moment estimations. Also the truncations are different.

In §2, we present examples of mixing, non-negative Markov chains $(X_1,X_2,\dots)$ satisfying (♣), (♡), (♠) and $\mathfrak{N}_X=1$, but violating theorem 1.1 in that

$$\overline{\lim_{n\to\infty}}\ \frac{1}{b(n)}(S_n-M_n^{(K)})=\infty \text{ a.s. } \forall\ K\in\mathbb{N}.$$

§3 is an application of theorem 1.1 to modified continued fractions. Let $x=\frac{1}{b_1-\frac{1}{b_2-\frac{1}{\ddots}}}$, then ( as shown in [Aar86]) $\frac{1}{n}\sum_{k=1}^n b_k\overset{P}{\to}3$ with respect to Lebesgue measure on $[0,1]$. We show that $\frac{1}{n}\sum_{k=1}^n b_k\not\to$ a.s..

## §1 Proof of theorem 1.1

We'll use the (elementary) fact that if $A:\mathbb{R}_+\to\mathbb{R}_+$ is increasing, weakly regularly varying, and $h(n)\downarrow,\ \gamma>0$ then

$$\sum_{n=1}^\infty n^\gamma h(A(n))<\infty \text{ implies } \sum_{n=1}^\infty n^\gamma h(\epsilon A(n))<\infty\ \forall\ \epsilon>0$$

since if $K\in\mathbb{N}$ satisfies $\epsilon A(Kn)\geq A(n)$, then

$$\sum_{n=1}^\infty n^\gamma h(\epsilon A(n))=\sum_{j=0}^{K-1}\sum_{n=1}^\infty(Kn+j)^\gamma h(\epsilon A(Kn+j))$$

$$\leq K^{\gamma+1}\sum_{n=1}^\infty(n+1)^\gamma h(A(n)))<\infty.$$

Let $N_{n,b}:=\#\{k\leq n:\ X_k>b\}\ \ (b>0)$.

The following is a straightforward generalisation of lemma 3 in [Mor76] and lemma 2 in [Mor77] to the c.f.-mixing case, and we only give a sketch of the proof.

**Lemma 1**

*Suppose that $(X_1, X_2, \cdots)$ is c.f.-mixing and that $B : \mathbb{R}_+ \to \mathbb{R}_+$ is non-decreasing and satisfies $nc(B(n)) \to 0$, then for $\nu \in \mathbb{N}$,*

$$\varlimsup_{n \to \infty} N_{n,B(n)} \leq \nu \quad a.s. \quad \Longleftrightarrow \quad \sum_{n=1}^{\infty} n^\nu P(X > B(n))^{\nu+1} < \infty.$$

*In this case, if $B : \mathbb{R}_+ \to \mathbb{R}_+$ is weakly regularly varying, then*

$$\varlimsup_{n \to \infty} N_{n,cB(n)} \leq \nu \quad a.s. \quad \forall \ c > 0.$$

**Proof**

As above,

$$\sum_{n=1}^{\infty} n^\nu P(X > B(n))^{\nu+1} \asymp \sum_{n=1}^{\infty} n^\nu P(X > cB(n))^{\nu+1} \ \forall \ c > 0$$

in case $B : \mathbb{R}_+ \to \mathbb{R}_+$ is weakly regularly varying. The proof therefore splits into 2 parts:

$$P(N_{n,B(n)} \geq \nu) \asymp (nc(B(n)))^\nu \ \forall \ \nu \geq 1; \tag{1}$$

and

$$\varlimsup_{n \to \infty} N_{n,B(n)} \leq \nu \quad \text{a.s.} \quad \Longleftrightarrow \quad \sum_{n=1}^{\infty} \frac{P(N_{n,B(n)} \geq \nu+1)}{n} < \infty. \tag{2}$$

Set $N_n = N_{n,B(n)}$ for $n \geq 1$. To establish (1), suppose that $M$ is as in the definition of c.f. mixing and that $\vartheta(\kappa) < 1$.

$$P(N_n \geq \nu) \leq \sum_{K \subset \{1,\ldots,n\}, \ |K|=\nu} P(X_k > B(n) \ \forall \ k \in K)$$

$$\leq M^\nu \binom{n}{\nu} c(B(n))^\nu$$

$$\ll n^\nu c(B(n))^\nu.$$

Now fix $n \gg \kappa$ so that $nc(B(n)) < \frac{1}{2}$. For $1 \leq k \leq n$ let

$$A_k := \bigcap_{1 \leq j \leq n, \ |j-k| \geq \kappa} [X_k > B(n), \ X_j \leq B(n)],$$

then

$$\sum_{k=1}^{n} 1_{A_k} \leq \kappa 1_{[N_n \geq 1]}$$

and

$$P(A_k) \geq (1 - \vartheta(\kappa))^2 P(\bigcap_{j=1}^{k-\kappa} [X_j \leq B(n)]) c(B(n)) P(\bigcap_{j=k+\kappa}^{n} [X_j \leq B(n)])$$

$$\geq (1 - \vartheta(\kappa))^2 (1 - kc(B(n))) c(B(n)) (1 - (n-k)c(B(n)))$$

$$\geq \tfrac{1}{4} (1 - \vartheta(\kappa))^2 c(B(n))$$

whence

$$P(N_n \geq 1) \geq \tfrac{1}{\kappa} \sum_{k=1}^{n} P(A_k) \geq \tfrac{1}{4\kappa}(1 - \vartheta(\kappa))^2 nc(B(n)) =: \eta nc(B(n)).$$

It now follows that for $n \gg \nu\kappa$ so large that $nc(B(n)) < \tfrac{1}{2}$

$$P(N_n \geq \nu) \geq P\Big( \sum_{\ell=1}^{\frac{n}{\nu}-\kappa} 1_{[X_{j\frac{n}{\nu}+\ell}>B(n)]} \geq 1 \ \forall \ \ 0 \leq j \leq \nu - 1 \Big)$$

$$\geq (1 - \vartheta(\kappa))^\nu P(N_{\frac{n}{\nu}-\kappa,B(n)} \geq 1)^\nu \geq (1 - \vartheta(\kappa))^\nu (\eta(\tfrac{n}{\nu} - \kappa)c(B(n)))^\nu$$

$$\gg n^\nu c(B(n))^\nu.$$

This establishes (1).

The proof of (2) is that of lemma 3 of [Mor76], but using the strong Borel-Cantelli lemma of Renyi (see [Ren70] p. 391) which is valid for c.f.-mixing processes instead of the classical one (which is only valid for i.i.d.r.v.'s).          □


*Proof of* (i) *of theorem 1.1*

By lemma 1 , a.s.,

$$\varlimsup_{n\to\infty} \sum_{k=1}^{n} 1_{[X_k>b(n)]} = \min\Big\{\kappa \geq 1 : \ \sum_{n=1} n^\kappa c(b(n))^{\kappa+1} < \infty\Big\}$$

Using $c(x) = \frac{\epsilon(x)L(x)}{x}$ and $b(n+1) - b(n) \asymp L(b(n)) = \frac{b(n)}{n}$, we have for $r > 0$,

$$\sum_{n=1}^{\infty} n^{r-1}c(b(n))^r = \sum_{n=1}^{\infty} \frac{\epsilon(b(n))^r}{n} \asymp \sum_{n=1}^{\infty} (b(n+1) - b(n)) \frac{\epsilon(b(n))^r}{b(n)}$$

$$\asymp \sum_{n=1}^{\infty} \sum_{b(n)\leq k<b(n+1)} \frac{\epsilon(k)^r}{k} = J_r.$$

Thus, $\min\{\kappa \geq 1 : \ \sum_{n=1} n^\kappa c(b(n))^{\kappa+1} < \infty\} = \mathfrak{N}_X$ establishing (i).          □

*Proof of* (ii) *of theorem 1.1*

The main ingredient here is the estimation of moments of truncated sums in claim 1.

Define $\Delta(b) := \frac{1}{L(b)} \int_0^1 \epsilon(bt)L(bt)dt$, then $\Delta(b) \underset{b\to\infty}{\longrightarrow} 0$.

As in [Mor77] (but with $\Delta$ in place of $\epsilon$), define

$$\phi(x) := \frac{a(x)}{\sqrt{\Delta(x)}}.$$

We claim that $\phi(x) \uparrow \infty$ as $x \uparrow \infty$. Indeed

$$\tfrac{1}{\phi(x)^2} = \tfrac{\Delta(x)}{a(x)^2} = \tfrac{L(x)}{x} \cdot \tfrac{1}{x^2} \int_0^x tc(t)dt \downarrow 0.$$

Set $b_n := \phi^{-1}(n)$.

**Claim 1**

$$E\left(\left|\frac{S_n^{(b_n)}}{b(n)} - 1\right|^Q\right) \ll \Delta(b_n)^{\frac{Q+1}{2}} + \frac{1}{n}\sum_{k=1}^{n}\vartheta(k) \quad \forall \ Q \in 2\mathbb{N}. \qquad (1).1$$

*Proof* Set $b_n := \phi^{-1}(n)$. Fix $n \geq 1$ and set $Y_k := X \wedge b_n - L(b_n)$, then

$$E\left(|S_n^{(b_n)} - nL(b_n)|^Q\right) = E\left(\left(\sum_{k=1}^{n} Y_k\right)^Q\right) = \sum_{1 \leq k_1,\ldots,k_Q \leq n} E\left(\prod_{i=1}^{Q} Y_{k_i}\right).$$

The latter sums need further organisation before estimation.

Given $1 \leq k_1,\ldots,k_Q \leq n$ let $K := \{k_1,\ldots,k_Q\} = \{\kappa_1,\ldots,\kappa_\nu\}$ where $\nu \leq Q$ and $\kappa_1 < \cdots < \kappa_\nu$, and define $f : \{1,\ldots,\nu\} \to \mathbb{N}$ by $f(j) := \#\{1 \leq i \leq Q : k_i = \kappa_j\}$, then $\sum_{j=1}^{\nu} f(j) = Q$ and it follows that

$$\sum_{1 \leq k_1,\ldots,k_Q \leq n} E\left(\prod_{i=1}^{Q} Y_{k_i}\right) = \sum_{\nu=1}^{Q} \sum_{1 \leq \kappa_1 < \cdots < \kappa_\nu \leq n} \sum_{f \in E_\nu^{(Q)}} E\left(\prod_{j=1}^{\nu} Y_{\kappa_j}^{f(j)}\right)$$

where

$$E_\nu^{(Q)} := \left\{f : \{1,\ldots,\nu\} \to \mathbb{N}, \ \sum_{j=1}^{\nu} f(j) = Q = 2p\right\}.$$

There are two cases: $f \geq 2$ and $\min_{1 \leq k \leq \nu} f(k) = 1$. Given $1 \leq \nu \leq Q$ let

$$F_\nu^{(Q)} := \{f \in E_\nu^{(Q)} : \ f \geq 2\}, \ G_\nu^{(Q)} := \{f \in E_\nu^{(Q)} : \ \min_{1 \leq k \leq \nu} f(k) = 1\}.$$

It follows that

$$E\left(|S_n^{(b_n)} - nL(b_n)|^Q\right) \leq \sum_{\nu=1}^{Q} \sum_{f \in E_\nu^{(Q)}} \sum_{1 \leq \kappa_1 < \cdots < \kappa_\nu \leq n} \left|E\left(\prod_{j=1}^{\nu} Y_{\kappa_j}^{f(j)}\right)\right|$$

$$(\ddagger) \qquad = \sum_{\nu=1}^{Q} \sum_{f \in F_\nu^{(Q)}} + \sum_{\nu=1}^{Q} \sum_{f \in G_\nu^{(Q)}}$$

Since $F_\nu^{(Q)} = \varnothing$ for $\nu > p$, we have by c.f.-mixing that

$$\sum_{\nu=1}^{Q} \sum_{f \in F_\nu^{(Q)}} = \sum_{\nu=1}^{p} \sum_{f \in F_\nu^{(Q)}}$$

$$\ll \sum_{\nu=1}^{p} \sum_{f \in F_\nu^{(Q)}} \sum_{1 \leq \kappa_1 < \cdots < \kappa_\nu \leq n} \prod_{j=1}^{\nu} E\left(|Y_{\kappa_j}|^{f(j)}\right)$$

For $r \geq 2$ we have

$$E(|Y|^r) \leq 2^r E((X \wedge b_n)^r) = 2^r r \int_0^{b_n} x^{r-2} \epsilon(x) L(x) dx$$

$$= r 2^r b_n^{r-1} \int_0^1 t^{r-2} \epsilon(b_n t) L(b_n t) dt = r 2^r b_n^{r-1} L(b_n) \Delta(b_n)$$

so for $1 \leq \kappa_1 < \cdots < \kappa_\nu \leq n$ and $f \in F_\nu^{(Q)}$:

$$\prod_{j=1}^\nu E(|Y_{\kappa_j}|^{f(j)}) \ll \prod_{k=1}^\nu (b_n^{f(k)-1} \Delta(b_n) L(b_n)) = b_n^Q \left( \frac{L(b_n)\Delta(b_n)}{b_n} \right)^\nu.$$

Now $\frac{L(x)}{x} \sim \frac{1}{a(x)} = \frac{1}{\phi(x)\sqrt{\Delta(x)}}$ whence $\frac{L(b_n)}{b_n} = \frac{1}{\phi(b_n)\sqrt{\Delta(b_n)}} = \frac{1}{n\sqrt{\Delta(b_n)}}$ and

$$\prod_{j=1}^\nu E(|Y_{\kappa_j}|^{f(j)}) \ll b_n^Q \frac{\Delta(b_n)^{\frac{\nu}{2}}}{n^\nu}.$$

Thus:

$$\sum_{\nu=1}^Q \sum_{f \in F_\nu^{(Q)}} \ll \sum_{\nu=1}^p \binom{n}{\nu} b_n^Q \frac{\Delta(b_n)^{\frac{\nu}{2}}}{n^\nu} \asymp \sum_{\nu=1}^p b_n^Q \Delta(b_n)^{\frac{\nu}{2}} \sim b_n^Q \sqrt{\Delta(b_n)}.$$

We now turn to the estimation of $\sum_{f \in G_\nu^{(Q)}}$ in ($\ddagger$). Although $E(|X \wedge b_n|^r) = o(b_n^{r-1} L(b_n)) \ \forall \ r \geq 2$, we have $E(|X \wedge b_n|) = L(b_n)$, which is too large, and we must use c.f.-mixing more delicately in this case.

Fix $\nu \leq Q$, $f \in G_\nu^{(Q)}$ and suppose that $1 \leq J \leq \nu$ satisfies $f(J) = 1$. We'll do the "generic" (difficult) case $2 \leq J \leq \nu - 1$ ($\Rightarrow \ \nu \geq 3$).

$$\sum_{1 \leq \kappa_1 < \cdots < \kappa_\nu \leq n} |E(\prod_{i=1}^\nu Y_{\kappa_i}^{f(i)})|$$

$$= \sum_{L=1}^n \sum_{1 \leq \kappa_1 < \cdots < \kappa_{J-1} \leq L-1} \sum_{L+1 \leq \kappa_{J+1} < \cdots < \kappa_\nu \leq n} |E(\prod_{i=1}^{J-1} Y_{\kappa_i}^{f(i)} Y_L \prod_{i=J+1}^\nu Y_{\kappa_i}^{f(i)})|$$

Fix $\kappa_1 < \cdots < \kappa_{J-1} < L < \kappa_{J+1} < \cdots < \kappa_\nu \leq n$. By c.f.–mixing and $E(Y_L) = 0$,

$$|E(\prod_{i=1}^{J-1} Y_{\kappa_i}^{f(i)} Y_L \prod_{i=J+1}^\nu Y_{\kappa_i}^{f(i)})|$$

$$\leq E(\prod_{i=1}^{J-1} |Y_{\kappa_i}|^{f(i)}) E(|Y_L|) E(\prod_{i=J+1}^\nu |Y_{\kappa_i}^{f(i)}|)(\vartheta(L - \kappa_{J-1}) + \vartheta(\kappa_{J+1} - L))$$

$$\ll b_n^{Q-\nu} L(b_n)^\nu (\vartheta(L - \kappa_{J-1}) + \vartheta(\kappa_{J+1} - L)),$$

whence, by the above

$$\sum_{1 \le \kappa_1 < \cdots < \kappa_\nu \le n} |E(\prod_{i=1}^{\nu} Y_{\kappa_i}^{f(i)})| \ll$$

$$b_n^{Q-\nu} L(b_n)^\nu \sum_{1 \le K < L < K' \le n} \binom{K-1}{J-2}\binom{n-K'-1}{\nu-J-1}(\vartheta(L-K)+\vartheta(K'-L))$$

$$\le b_n^{Q-\nu} L(b_n)^\nu n^{\nu-3} \sum_{1 \le K < L < K' \le n} (\vartheta(L-K)+\vartheta(K'-L))$$

$$\le 2 b_n^{Q-\nu} L(b_n)^\nu n^{\nu-3} n^2 \sum_{k=1}^{n} \vartheta(k)$$

$$\ll n^{\nu-1} b_n^{Q-\nu} L(b_n)^\nu \sum_{k=1}^{n} \vartheta(k) = \frac{b_n^Q}{n}\left(\frac{1}{\Delta(b_n)}\right)^{\frac{\nu}{2}} \sum_{k=1}^{n} \vartheta(k)$$

It follows that

$$\sum_{\nu=1}^{Q} \sum_{f \in E_\nu^{(Q)}} \sum_{1 \le \kappa_1 < \cdots < \kappa_\nu \le n} |E(\prod_{k \in K} Y^{f(k)})| \ll \frac{b_n^Q}{n} \sum_{\nu=1}^{Q}\left(\frac{1}{\Delta(b_n)}\right)^{\frac{\nu}{2}} \sum_{k=1}^{n} \vartheta(k)$$

$$\sim \frac{b_n^Q}{n}\left(\frac{1}{\Delta(b_n)}\right)^{\frac{Q}{2}} \sum_{k=1}^{n} \vartheta(k).$$

Putting things together:

$$E(|S_n^{(b_n)} - nL(b_n)|^Q) \ll b_n^Q\left(\sqrt{\Delta(b_n)} + \frac{1}{n}\left(\frac{1}{\Delta(b_n)}\right)^{\frac{Q}{2}} \sum_{k=1}^{n} \vartheta(k)\right).$$

Next, note that $\phi(x) = \frac{a(x)}{\sqrt{\Delta(x)}}$ whence
$a(\phi^{-1}(x)) = x\sqrt{\Delta(\phi^{-1}(x))}$, $a(b_n) = n\sqrt{\Delta(b_n)}$ and

$$E(|\frac{S_n^{(b_n)}}{nL(b_n)} - 1|^Q) \ll \left(\frac{b_n}{nL(b_n)}\right)^Q\left(\sqrt{\Delta(b_n)} + \frac{1}{n}\left(\frac{1}{\Delta(b_n)}\right)^{\frac{Q}{2}} \sum_{k=1}^{n} \vartheta(k)\right)$$

$$= \Delta(b_n)^{\frac{Q+1}{2}} + \frac{1}{n} \sum_{k=1}^{n} \vartheta(k) \to 0.$$

Thus $\frac{S_n^{(b_n)}}{nL(b_n)} \xrightarrow{P} 1$. Since $nc(b_n) \to 0$, we have $\frac{S_n}{nL(b_n)} \xrightarrow{P} 1$, whence $nL(b_n) \sim b(n)$ and

$$E(|\frac{S_n^{(b_n)}}{b(n)} - 1|^Q) \ll \Delta(b_n)^{\frac{Q+1}{2}} + \frac{1}{n} \sum_{k=1}^{n} \vartheta(k)$$

which is (1.1) and the claim is established. $\qquad\qquad\square$

**Claim 2**

$$\sum_{n=1}^{\infty} \frac{1}{n} P([|\frac{S_n^{(b_n)}}{b(n)} - 1| > \epsilon]) < \infty \ \forall \ \epsilon > 0. \qquad\qquad (1).2$$

*Proof*

By the Chebyshev-Markov inequality, $P([|\frac{S_n^{(b_n)}}{b(n)} - 1| > \epsilon]) \ll E(|\frac{S_n^{(b_n)}}{b(n)} -$

$1|^Q)$, $\forall\ Q > 1$, so by claim 1, (1.2) will follow from $\sum_{n=1}^{\infty} \frac{\Delta(b_n)^{\frac{Q+1}{2}}}{n} < \infty$
for some $Q > 1$ and $\sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{k=1}^{n} \vartheta(k) < \infty$. The latter follows form the
assumptions on $\{\vartheta(n)\}_{n\geq1}$ as

$$\sum_{n=1}^{\infty} \tfrac{1}{n^2} \sum_{k=1}^{n} \vartheta(k) = \sum_{k=1}^{\infty} \vartheta(k) \sum_{n=k}^{\infty} \tfrac{1}{n^2} \asymp \sum_{k=1}^{\infty} \tfrac{\vartheta(k)}{k} < \infty.$$

We'll show that

$$\sum_{n=1}^{\infty} \frac{\Delta(b_n)^{\kappa}}{n} \asymp J_{\kappa} \ \forall\ \kappa > 0. \tag{1.3}$$

The proof of (1.3) is in two parts.

Firstly, for $\kappa,\ \gamma > 0$ and writing $\gamma' = \phi(\gamma)$, we have

$$\sum_{n=1}^{\infty} \frac{\Delta(b_n)^{\kappa}}{n} \asymp \int_{\gamma'}^{\infty} \frac{\Delta(\phi^{-1}(x))^{\kappa}dx}{x} = \int_{\gamma'}^{\infty} \frac{a(\phi^{-1}(x)))^{2\kappa}dx}{x^{2\kappa+1}} \leftarrow_{t\to\infty} \int_{\gamma}^{t} \frac{a(y)^{2\kappa}\phi'(y)dy}{\phi(y)^{2\kappa+1}}$$

$$= \left[\frac{-a(y)^{2\kappa}}{2\kappa\phi(y)^{\kappa}}\right]_{\phi^{-1}(\gamma)}^{t} + \int_{\gamma}^{t} \frac{a(y)^{2\kappa-1}a'(y)dy}{\phi(y)^{2\kappa}} = \int_{\gamma}^{t} \frac{L(y)\Delta(y)^{\kappa}a'(y)dy}{y} + o(1)$$

$$\asymp \int_{\gamma}^{\infty} \frac{\Delta(y)^{\kappa}dy}{y}.$$

Next, we show that $\int_{\gamma}^{\infty} \frac{\Delta(y)^{\kappa}dy}{y} \asymp J_{\kappa}$.

We start with $J_{\kappa} \ll \int_{c}^{\infty} \frac{\Delta(x)^{\kappa}dx}{x}$ because $\epsilon \ll \Delta$. To see this, recall
that $\frac{\epsilon(x)}{x} \downarrow$ whence $\epsilon(by) \geq y\epsilon(b) \ \forall\ b > 0,\ 0 < y < 1$ and

$$\Delta(b) = \tfrac{1}{L(b)} \int_{0}^{1} \epsilon(by)L(by)dt \geq \tfrac{\epsilon(b)}{L(b)} \int_{0}^{1} yL(by)dt \sim \tfrac{\epsilon(b)}{2}.$$

To show $\int_{c}^{\infty} \frac{\Delta(x)^{\kappa}dx}{x} \ll J_{\kappa}$:

$$\int_{1}^{\infty} \frac{\Delta(b)^{\kappa}db}{b} = \int_{1}^{\infty} \tfrac{1}{b}\left(\int_{0}^{1} \epsilon(bt)\tfrac{L(bt)dt}{L(b)}\right)^{\kappa}db \stackrel{\text{Jensen's ineq.}}{\to} \leq \int_{1}^{\infty} \tfrac{1}{b} \int_{0}^{1} \epsilon(bt)^{\kappa}\tfrac{L(bt)dt}{L(b)}db$$

$$\leq \int_{0}^{1} \int_{1}^{\infty} \frac{\epsilon(bt)^{\kappa}dbdt}{b} \stackrel{y:=bt}{\to} = \int_{0}^{1} \int_{t}^{\infty} \frac{\epsilon(y)^{\kappa}dydt}{y}$$

$$= \int_{1}^{\infty} \frac{\epsilon(y)^{\kappa}dy}{y} + \int_{0}^{1} \int_{t}^{1} \frac{\epsilon(y)^{\kappa}dydt}{y} = \int_{1}^{\infty} \frac{\epsilon(y)^{\kappa}dy}{y} + \int_{0}^{1} \epsilon(y)^{\kappa}dy$$

$$= J_{\kappa} + O(1),$$

(1.3) and claim 2 are established. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

**Claim 3** $\frac{S_n^{(b_n)}}{b(n)} \to 1\ a.s..$

*Proof*   From claim 2 by condensation,

$$\sum_{j=1}^{\infty} P([|\frac{S^{(b_{[\lambda^j]})}_{[\lambda^j]}}{b([\lambda^j])} - 1| > \epsilon]) < \infty \ \forall \ \epsilon > 0, \ \lambda > 1$$

whence $\frac{S^{(b_{[\lambda^j]})}_{[\lambda^j]}}{b([\lambda^j])} \to 1$ a.s. $\forall \ \lambda > 1$. By monotonicity, $\forall \ \lambda > 1$, a.s.,

$$\frac{1}{\lambda} = \lim_{j\to\infty} \frac{S^{(b_{[\lambda^{j-1}]})}_{[\lambda^{j-1}]}}{b([\lambda^j])} \le \varliminf_{n\to\infty} \frac{S^{(b_n)}_n}{b(n)} \le \varlimsup_{n\to\infty} \frac{S^{(b_n)}_n}{b(n)} \le \lim_{j\to\infty} \frac{S^{(b_{[\lambda^{j+1}]})}_{[\lambda^{j+1}]}}{b([\lambda^j])} = \lambda \ \text{a.s.}$$

showing that $\frac{S^{(b_n)}_n}{b(n)} \to 1$ a.s.. □

**Claim 4**

$$\varlimsup_{n\to\infty} \sum_{k=1}^{n} 1_{[X_k > b_n]} \le 2\mathfrak{N} + 2 \ \ a.s..$$

*Proof*   By lemma 1, it suffices to show

$$\sum_{n=1}^{\infty} n^{2\mathfrak{N}+1} c(b_n)^{2\mathfrak{N}+2} < \infty.$$

For $\kappa = 2\mathfrak{N} + 2$,

$$\sum_{n=1}^{\infty} n^{\kappa-1} c(b_n)^{\kappa} = \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{n\epsilon(b_n)}{a(b_n)}\right)^{\kappa} \ll \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{n\Delta(b_n)}{a(b_n)}\right)^{\kappa}$$

$$= \sum_{n=1}^{\infty} \frac{\Delta(b_n)^{\frac{\kappa}{2}}}{n} \overset{(1.3)}{\to} \asymp J_{\frac{\kappa}{2}} = J_{\mathfrak{N}+1} < \infty.$$

□

**Claim 5**

$$S_n - M_n^{(\mathfrak{N}_X)} \sim b(n) \ \ a.s..$$

*Proof*

$\forall \ \eta > 0$, a.s. for $n$ large

$$S_n - M_n^{(\mathfrak{N}_X)} = S_n^{(\eta b(n))} = S_n^{(b_n)} \pm (2\mathfrak{N} + 2)\eta b(n)$$

whence

$$1 - (2\mathfrak{N} + 2)\eta \le \varliminf_{n\to\infty} \frac{S_n - M_n^{(\nu)}}{b(n)} \le \varlimsup_{n\to\infty} \frac{S_n - M_n^{(\nu)}}{b(n)} \le 1 + (2\mathfrak{N} + 2)\eta.$$

□ This finishes the proof of theorem 1.1.

**Example**

If $\epsilon(t) \to 0$, $\epsilon(t) = \frac{1}{(\log t)^{o(1)}}$ as $t \to \infty$ (e.g. ), then $\mathfrak{N}_X = \infty$.

If $\epsilon(t) = o(\frac{1}{\log\log\log t})$, then $L(t) \sim L(t \log\log t)$ and ($\heartsuit$) holds.

Both conditions are satisfied for $L(t) = e^{\frac{\log(t+30)}{\log\log(t+30)}}$. Thus there are processes (i.i.d.r.v.'s) $(X_1, X_2, \dots)$ satisfying ($\heartsuit$), but for which $\mathfrak{N}_X = \infty$ and trimming of any bounded number of maxima will not ensure a.s. convergence.

## §2 Markov chains with no trimmed strong law

In this section we construct examples showing that theorem 1 fails for general mixing Markov chains.

### Examples

*There are non-negative, mixing Markov chains $(Y_1, Y_2, \dots)$ satisfying $E(Y) = \infty$, $\mathfrak{N}_Y = 1$, ($\spadesuit$), ($\spadesuit$) and ($\heartsuit$) with normalising constants $b(n) = nE(Y \wedge b(n))$; but such that*

$$\varlimsup_{n \to \infty} \frac{(S_n - M_n^{(K)})}{b(n)} = \infty \ \ a.s. \ \ \forall \ K \in \mathbb{N}.$$

For convenience, we construct the Markov chains over probability preserving transformations. Let $S$ be an ergodic probability preserving transformation of the standard probability space $(\Omega, \mathcal{A}, p)$ and $f : \Omega \to \mathbb{N}$ be measurable, integrable and so that $\{f \circ S^n : n \geq 0\}$ are independent (e.g. $\Omega = \mathbb{N}^\mathbb{N}$, $S$ = shift, $f(x) = x_1$ and $p$ is a product measure).

Build $(X, \mathcal{B}, q, T)$ the tower transformation over $S$ with height function $f$ (see [Kak43] or §1.5 of [Aar97]). This is an ergodic probability preserving transformation :

$$X := \{(x, n) : \ 1 \leq n \leq f(x)\}, \ \ q(A \times \{n\}) := \tfrac{p(A)}{E(f)},$$

$$T(x, n) := \left\{ \begin{array}{ll} (x, n+1) & n < f(x)), \\ (Sx, 1) & n = f(x). \end{array} \right.$$

Now define $g : X \to \mathbb{N}$ by $g(x, n) := n$.

Our examples will be of form $(Y_1, Y_2, \dots) := (g, g \circ T, g \circ T^2, \dots)$. A calculation indeed shows that the ergodic stationary process $(g, g \circ T, g \circ T^2, \dots)$ is a Markov chain (a renewal process) whose joint distributions are given by

$$q([g = s_0, g \circ T = s_1, \dots, g \circ T^n = s_n]) = \pi_{s_0} p_{s_0, s_1} \dots p_{s_{n-1}, s_n}$$

where $\pi_s := \tfrac{p([f \geq s])}{E(f)}$ and

$$p_{j,k} = \left\{ \begin{array}{l} \frac{p([f=j])}{E(f)\pi_j} \ \text{if} \ j \in \mathbb{N}, \ k = 1 \\ \frac{\pi_{j+1}}{\pi_j} \ \text{if} \ j \in \mathbb{N}, \ k = j+1, \\ 0 \ \text{else}. \end{array} \right.$$

This chain is mixing if (e.g.) $p([f = n]) > 0 \ \forall \ n \geq 1$ large.

**Proposition 2.1**  ([Tan74])

$$\frac{g \circ T^n}{n} \underset{n \to \infty}{\to} \ 0 \ a.s.$$

**Proof**  Since $E(f) < \infty$, we have $\frac{f \circ S^n}{n} \to 0$ a.s. on $\Omega$. Next, for a.e. $x \in \Omega$ and $\forall\ n$ large, $\exists\ 0 \le k_n \le n$ such that $g(T^n x) \le f(S^{k_n} x)$ whence $\frac{g \circ T^n}{n} \to 0$ a.s. on $\Omega$. The proposition follows from the $T$-invariance of $\overline{\lim}_{n \to \infty} \frac{g \circ T^n}{n}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Next, we investigate the asymptotic behaviour of $g_n = g_n^{(T)} := \sum_{k=0}^{n-1} g \circ T^k$. To this end, let

$$\mathcal{L}(t) := E\left(\left(\frac{f(f+1)}{2}\right) \wedge t\right).$$

**Lemma 2.2**
  (1) If $\mathcal{L}(t)$ is slowly varying at $\infty$ and $E(f^2) = \infty$, then

$$\mathcal{L}(t) \sim \tfrac{1}{2} E(f^2 \wedge t) \text{ as } t \to \infty.$$

  (2) If $p([f \ge u]) \sim \frac{h(u)}{u^2}$ where $\int_1^\infty \frac{h(u) du}{u} = \infty$ and $h$ is slowly varying at $\infty$, then $E(g) = \infty$, $\mathcal{L}$ is slowly varying at $\infty$ and

$$L_g(t) := E(g \wedge t) \sim \tfrac{1}{E(f)} \mathcal{L}(t^2) \text{ as } t \to \infty.$$

**Proof**

$$\tfrac{1}{2} E(f^2 \wedge t) = E\left(\tfrac{f^2}{2} \wedge t\right) \le \mathcal{L}(t) \sim \mathcal{L}\left(\tfrac{t}{2}\right) = \tfrac{1}{2} E(f(f+1) \wedge t) \sim \tfrac{1}{2} E(f^2 \wedge t). \quad (1)$$

To establish 2), we first note that $\forall\ \epsilon > 0$, $\int_1^t \frac{h(u) du}{u} \ge \int_{\epsilon t}^t \frac{h(u) du}{u} \sim h(t) \log \frac{1}{\epsilon}$ as $t \to \infty$, whence $h(t) = o(\int_1^t \frac{h(u) du}{u})$ as $t \to \infty$. It follows that $\int_1^t \frac{h(u) du}{u}$ is slowly varying at $\infty$ (because $\int_t^{\lambda t} \frac{h(u) du}{u} \sim h(t) \log \lambda$ as $t \to \infty$). Next

$$\tfrac{1}{2} E(f^2 \wedge t) = \tfrac{1}{2} E((f \wedge \sqrt{t})^2) = \int_0^{\sqrt{t}} s p([f \ge s]) ds \sim \int_1^{\sqrt{t}} \frac{h(u) du}{u}$$

which latter is slowly varying at $\infty$. Analogously to the proof of 1), we see that $\mathcal{L}(t)$ is slowly varying at $\infty$. Next,

$$q(g \ge u) = \tfrac{1}{E(f)} \sum_{\nu=u}^\infty p(f \ge \nu) \sim \frac{h(u)}{E(f) u}$$

whence

$$L_g(t) = \sum_{k=1}^t q(g \ge k) \sim \tfrac{1}{E(f)} \sum_{u=1}^t \frac{h(u)}{u} \sim \tfrac{1}{E(f)} \mathcal{L}(t^2).$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$$

We use the notation $g_n = g_n^{(T)} := \sum_{k=0}^{n-1} g \circ T^k$.

**Proposition 2.3**

*1) Suppose that $E(g) = \infty$, $\mathcal{L}$ is slowly varying and let $\beta(n) = n\mathcal{L}(\beta(n))$, then*

$$\frac{g_n}{\beta(n)} \xrightarrow{q} \frac{1}{E(f)}, \quad \overline{\lim_{n\to\infty}} \frac{g_n}{\beta(n)} = \infty \ a.s.,$$

*and, in case $\mathcal{L}(t) \sim \mathcal{L}(t \log \log t)$:*

$$\overline{\lim_{n\to\infty}} \frac{g_n}{\beta(n)} = \frac{1}{E(f)} \ a.s..$$

*2) Under the assumptions of lemma 2.2 and $\mathcal{L}(t^2) \sim \mathcal{L}(t)$; $(g, g \circ T, \dots)$ satisfies ($\clubsuit$), ($\spadesuit$) and ($\heartsuit$).*

**Proof**  Note that $T_\Omega = T^f = S$ whence $T_\Omega^n = T^{f_n^{(S)}}$ where $f_n = f_n^{(S)} := \sum_{k=0}^{n-1} f \circ S^k$. It follows that on $\Omega$:

$$g_{f_n^{(S)}}^{(T)} = h_n^{(S)}$$

where

$$h := g_f^{(T)} = \sum_{k=0}^{f-1} g \circ T^k = \frac{f(f+1)}{2}.$$

Since $\{h \circ S^n : \ n \geq 1\}$ are independent, by ($\clubsuit$), ($\spadesuit$) and ($\heartsuit$):

$$\frac{h_n^{(S)}}{\beta(n)} \xrightarrow{q} 1, \quad \overline{\lim_{n\to\infty}} \frac{h_n^{(S)}}{\beta(n)} = \infty \ a.s.,$$

and, in case $\mathcal{L}(t) \sim \mathcal{L}(t \log \log t)$:

$$\overline{\lim_{n\to\infty}} \frac{h_n^{(S)}}{\beta(n)} = 1 \ a.s..$$

By the PET, $f_n \sim E(f)n$ a.s. on $\Omega$, whence, a.s. on $\Omega$ (!):

$$\frac{g_{E(f)n}}{\beta(n)} \xrightarrow{q} 1, \quad \overline{\lim_{n\to\infty}} \frac{g_{E(f)n}}{\beta(n)} = \infty \text{ and, in case } \mathcal{L}(t) \sim \mathcal{L}(t \log \log t), \ \overline{\lim_{n\to\infty}} \frac{g_{E(f)n}}{\beta(n)} = 1.$$

Using the 1-regular variation of $\beta(n)$, and ergodicity of $T$, we establish 1) from which 2) follows since $\mathcal{L}(t^2) \sim \mathcal{L}(t)$ implies $\beta(n) \sim E(f)b(n)$ where $b(n) = nE(g \wedge b(n))$. □

**Remark**  Note that $\mathcal{L}(t^2) \sim \mathcal{L}(t)$ if $\epsilon(t) := t(\log^+ L)'(t) = o(\frac{1}{\log t})$ as $t \to \infty$.

**Proposition 2.4** *If $E(g) = \infty$, then*

$$\overline{\lim_{n\to\infty}} \frac{(g_n^{(T)} - M_n^{(K)})}{\beta(n)} = \infty \ a.s. \ \forall \ K \in \mathbb{N}.$$

**Proof**  $r_{n,1}(x) = g \circ T^{k_n(x)}(x)$ for some $0 \leq k_n(x) \leq n - 1$. Thus,

$$M_n^{(K)} \leq K r_{n,1}(x) = K g \circ T^{k_n(x)}(x) = o(n)$$

as $n \to \infty$ by proposition 2.1. On the other hand, $E(g) = \infty$, so $\frac{g_n}{n} \to \infty$ and $M_n^{(K)} = o(g_n)$ a.s.. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

**The advertised examples.** If $p([f \geq t]) \sim \frac{h(t)}{t^2}$ as $t \to \infty$ where $\frac{1}{h(t)} = \prod_{j=1}^{r} \log(t + e_j)$ for some $r \in \mathbb{N}$ where $e_1 := e$, $e_{j+1} := e^{e_j}$, then $\mathcal{L}(t) \sim \log^{r+1}(t) \sim \mathcal{L}(t^2)$ as $t \to \infty$ where $\log^1(t) := \log(t)$ and $\log^{r+1}(t) := \log(\log^r(t))$.

Thus, $E(g) = \infty$, $\mathfrak{N}_g = 1$, and $(g, g \circ T, \dots)$ satisfies ($\clubsuit$), ($\spadesuit$) and ($\heartsuit$) with normalising constants $b(n) = nE(Y \wedge b(n))$ but $\overline{\lim}_{n \to \infty} \frac{(g_n^{(T)} - M_n^{(K)})}{b(n)} = \infty$ a.s. $\forall\ K \in \mathbb{N}$..

## §3 Applications

**3.1 Modified continued fractions.** Let $x = \cfrac{1}{b_1 - \cfrac{1}{b_2 - \cfrac{1}{\ddots}}}$, then $b_n(x) = \left[\frac{1}{V^{n-1}x}\right] + 1$ where $Vx := 1 - \{\frac{1}{x}\}$. The transformation $V : [0, 1] \to [0, 1]$ has an infinite, invariant measure $\mu$ with density $\frac{d\mu}{dm}(x) = \frac{1}{1-x}$ with respect to which the function $b(x) = [\frac{1}{x}] + 1$ is not integrable. Nevertheless ( as shown in [Aar86])

$$A(n) := \frac{1}{n} \sum_{k=1}^{n} b_k \overset{P}{\to} 3.$$

We prove here that a.s.,

($\maltese$) $\qquad\qquad \underline{\lim}_{n \to \infty} A(n) = 2, \ \& \ \overline{\lim}_{n \to \infty} A(n) = \infty.$

As shown in [DK00],

$$A\Big(\sum_{k=1}^{n} a_{2k-1}\Big) = 2 + \frac{\sum_{k=1}^{n} a_{2k}}{\sum_{k=1}^{n} a_{2k-1}}$$

where $x = 1/a_1 + 1/a_2 + 1/\dots$. The regular continued fraction process $(a_1, a_2, \dots)$ is given by $a_n(x) := a(U^{n-1}x)$ where $a(x) := [\frac{1}{x}]$ and $U : (0, 1) \to (0, 1)$ is defined by $Ux := \{\frac{1}{x}\}$. Gauß' measure $d\mathbb{P}(x) = \frac{dx}{\log 2(1+x)}$ is $U$-invariant on $[0, 1]$. As shown in [Doe40], it is c.f.-mixing with $\vartheta(n) = O(\theta^n)$ for some $0 < \theta < 1$.

Theorem 1.1 holds with $\mathfrak{N}_a = 1$. The trimmed strong law for the regular continued fraction process was first established in [DV86].

Thus, ($\maltese$) follows from the following lemma.

**Lemma 3.1**

*Let $\{X_k\}_{k \geq 1}$ be a non-negative, stationary process with $\sum_{k=1}^{\infty} \frac{\vartheta(k)}{k} < \infty$, and suppose that $\mathfrak{N}_X < \infty$, then for $d \geq 2$ and $0 \leq i \neq j < d$,*

$$\varliminf_{n \to \infty} \frac{\sum_{k=1}^{n} X_{dk+i}}{\sum_{k=1}^{n} X_{dk+j}} = 0 \quad \& \quad \varlimsup_{n \to \infty} \frac{\sum_{k=1}^{n} X_{dk+i}}{\sum_{k=1}^{n} X_{dk+j}} = \infty \quad a.s..$$

**Proof**

Since $\mathfrak{N}_X < \infty$, $L$ is slowly varying at $\infty$, whence $b(t)$ defined by $b(t) = tL(b(t))$ is regularly varying at $\infty$ with index $1$ . We claim first that $\exists \ \beta_n = o(b(n))$ such that $\varlimsup_{n \to \infty} \sum_{k=1}^{n} 1_{[Z_k > \beta_n]} = \mathfrak{N}_X$ a.s. for any stationary process $\{Z_n\}$ with $\sum_{n=1}^{\infty} \frac{\vartheta(n)}{n} < \infty$ and dist $Z =$ dist $X$.

By lemma 1, $\sum_{n \in \mathbb{N}} n^{\mathfrak{N}_X} c\left(\frac{b(n)}{k}\right)^{\mathfrak{N}_X + 1} < \infty$. To obtain such a sequence $\{\beta_n\}$, fix $m_k \uparrow$ such that

$$\sum_{n \geq m_k} n^{\mathfrak{N}_X} c\left(\frac{b(n)}{k}\right)^{\mathfrak{N}_X + 1} < \frac{1}{2^k} \ \forall \ k \geq 1$$

and set $\beta_n := \frac{b(n)}{k}$ for $n \in \mathbb{N}$, $m_k \leq n < m_{k+1}$. Evidently, $\beta_n = o(b(n))$ and $\sum_{n \in \mathbb{N}} n^{\mathfrak{N}_X} c(\beta_n)^{\mathfrak{N}_X + 1} < \infty$, whence $\varlimsup_{n \to \infty} \sum_{k=1}^{n} 1_{[Z_k > \beta_n]} = \mathfrak{N}_X$ a.s..

By theorem 1.1, $S_n^{(\beta_n)} \sim b(n)$ a.s., and to see $\varlimsup_{n \to \infty} \frac{\sum_{k=1}^{n} X_{dk+i}}{\sum_{k=1}^{n} X_{dk+j}} = \infty$ a.s., fix $M > 0$ large and note that a.s., $\exists \ n_\ell \to \infty$ and $B_\ell \subset \{dk + i\}_{k=1}^{n_\ell}$, $|B_\ell| = \mathfrak{N}_X$ such that
(i) $X_k > Mb(n_\ell) \ \forall \ k \in B_\ell$, and (ii) $X_k \leq \beta_{n_\ell} \ \forall \ k \notin B_\ell, \ k \leq (d+1)n_\ell$. It follows that

$$\sum_{k=1}^{n_\ell} X_{dk+j} = \sum_{k=1}^{n_\ell} X_{dk+j} \wedge \beta_{n_\ell} \sim b(n_\ell) \text{ a.s.,}$$

whereas

$$\sum_{k=1}^{n_\ell} X_{dk+i} \geq M \mathfrak{N}_X b(n_\ell)$$

with the conclusion that

$$\varlimsup_{n \to \infty} \frac{\sum_{k=1}^{n} X_{dk+i}}{\sum_{k=1}^{n} X_{dk+j}} \geq \varlimsup_{\ell \to \infty} \frac{\sum_{k=1}^{n_\ell} X_{dk+i}}{\sum_{k=1}^{n_\ell} X_{dk+j}} \geq \lim_{\ell \to \infty} \frac{M \mathfrak{N}_X b(n_\ell)}{\sum_{k=1}^{n_\ell} X_{dk+j} \wedge \beta_{n_\ell}} = M \mathfrak{N}_X.$$

$\square$

**3.2 Visits to cusps.** Define $W : [0,1] \to [0,1]$ by $W(x) = \frac{x}{1-x}$ $\left(0 < x < \frac{1}{2}\right)$ and $W(1-x) = 1 - W(x)$.

The measure $\nu \sim m$ with $\frac{d\nu}{dm}(x) = \frac{1}{x(1-x)}$ is $W$-invariant, and as shown in [Tha83] (see also [Aar97]), $([0,1], m, W)$ is conservative and ergodic.

The invariant measure density $\nu$ has "cusps" at $0$ and $1$ in the sense $\mu([0,\epsilon)) = \mu([1-\epsilon,1)) = \infty \ \forall \ \epsilon > 0$, but $\mu((a,b)) < \infty \ \forall \ 0 < a < b < 1$ and it is natural to ask about the frequency of visits to these "cusps".

It was shown in [Tha02] that

$$\frac{1}{n} \sum_{k=0}^{n-1} 1_{[0,\frac{1}{2})} \circ W^k \xrightarrow{m} \to \frac{1}{2}, \quad \text{whence} \quad \frac{\sum_{k=0}^{n-1} 1_{[0,\frac{1}{2})} \circ W^k}{\sum_{k=0}^{n-1} 1_{[\frac{1}{2},1)} \circ W^k} \xrightarrow{m} \to 1. \qquad (\dagger)$$

We show, using ($\maltese$) that

$$\varliminf_{n\to\infty} \frac{\sum_{k=0}^{n-1} 1_{[0,\frac{1}{2})}(W^k x)}{\sum_{k=0}^{n-1} 1_{[\frac{1}{2},1)}(W^k x)} = 0, \quad \varlimsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} 1_{[0,\frac{1}{2})}(W^k x)}{\sum_{k=0}^{n-1} 1_{[\frac{1}{2},1)}(W^k x)} = \infty \qquad (\ddagger)$$

(c.f. [Ino97] and [Ino01]).

Define $K : [0,1] \to \mathbb{Z}_+$ by $K(x) := \min\{j \geq 0 : W^j x > \frac{1}{2}\}$ and $\tilde{W} : [0,1] \to [0,\frac{1}{2}] \times \{0,1\}$ by $\tilde{W}(x) := W^{K(x)+1}(x)$. It turns out that $K(x) = b(x) - 2 := [\frac{1}{x}] - 1$, $W(x) = V(x) := 1 - \{\frac{1}{x}\}$ ($b$, $V$ as above), whence by ($\maltese$), $\varliminf_{n\to\infty} \frac{K_n(x)}{n} = 0$ and $\varlimsup_{n\to\infty} \frac{K_n(x)}{n} = \infty$ a.s. where $K_n := \frac{1}{n} \sum_{k=0}^{n-1} K \circ V^k$.

This proves ($\ddagger$) as

$$\sum_{k=0}^{K_n(x)-1} 1_{[0,\frac{1}{2})}(W^k x) = K_n(x) \quad \text{and} \quad \sum_{k=0}^{K_n(x)-1} 1_{[\frac{1}{2},1)}(W^k x) = n.$$

## References

[Aar77]   Jon Aaronson. On the ergodic theory of non-integrable functions and infinite measure spaces. *Israel J. Math.*, 27(2):163–173, 1977.

[Aar86]   Jon. Aaronson. Random $f$-expansions. *Ann. Probab.*, 14(3):1037–1057, 1986.

[Aar97]   Jon Aaronson. *An introduction to infinite ergodic theory*, volume 50 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.

[AD89]    Jon Aaronson and Manfred Denker. Lower bounds for partial sums of certain positive stationary processes. In *Almost everywhere convergence (Columbus, OH, 1988)*, pages 1–9. Academic Press, Boston, MA, 1989.

[AD90]    Jon Aaronson and Manfred Denker. Upper bounds for ergodic sums of infinite measure preserving transformations. *Trans. Amer. Math. Soc.*, 319(1):101–138, 1990.

[AD01]    Jon Aaronson and Manfred Denker. Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps. *Stoch. Dyn.*, 1(2):193–237, 2001.

[BGT87]   N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1987.

[CR61]    Y. S. Chow and Herbert Robbins. On sums of independent random variables with infinite moments and "fair" games. *Proc. Nat. Acad. Sci. U.S.A.*, 47:330–335, 1961.

[DK00]   Karma Dajani and Cor Kraaikamp. "The mother of all continued fractions". *Colloq. Math.*, 84/85(part 1):109–123, 2000. Dedicated to the memory of Anzelm Iwanik.

[Doe40]  W. Doeblin. Remarques sur la théorie métrique des fractions continues. *Compositio Math.*, 7:353–371, 1940.

[DV86]   Harold G. Diamond and Jeffrey D. Vaaler. Estimates for partial sums of continued fraction partial quotients. *Pacific J. Math.*, 122(1):73–82, 1986.

[Fel45]  W. Feller. Note on the law of large numbers and "fair" games. *Ann. Math. Statistics*, 16:301–304, 1945.

[Fel46]  W. Feller. A limit theorem for random variables with infinite moments. *Amer. J. Math.*, 68:257–262, 1946.

[Fel66]  William Feller. *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons, Inc., New York-London-Sydney, 1966.

[Ino97]  Tomoki Inoue. Ratio ergodic theorems for maps with indifferent fixed points. *Ergodic Theory Dynam. Systems*, 17(3):625–642, 1997.

[Ino01]  Tomoki Inoue. Correction to: "Ratio ergodic theorems for maps with indifferent fixed points" [Ergodic Theory Dynam. Systems **17** (1997), no. 3, 625–642; MR1452184 (98e:58109)]. *Ergodic Theory Dynam. Systems*, 21(4):1273, 2001.

[Kak43]  Shizuo Kakutani. Induced measure preserving transformations. *Proc. Imp. Acad. Tokyo*, 19:635–641, 1943.

[Kar33]  J. Karamata. Sur un mode de croissance régulière. Théorèmes fondamentaux. *Bull. Soc. Math. France*, 61:55–62, 1933.

[KT77]   Michael Klass and Henry Teicher. Iterated logarithm laws for asymmetric random variables barely with or without finite mean. *Ann. Probability*, 5(6):861–874, 1977.

[Mor76]  Toshio Mori. The strong law of large numbers when extreme terms are excluded from sums. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 36(3):189–194, 1976.

[Mor77]  T. Mori. Stability for sums of i.i.d. random variables when extreme terms are excluded. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 40(2):159–167, 1977.

[Ren70]  A. Renyi. *Probability theory.* North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1970. Translated by László Vekerdi, North-Holland Series in Applied Mathematics and Mechanics, Vol. 10.

[Tan74]  David Tanny. A zero-one law for stationary sequences. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 30:139–148, 1974.

[Tha83]  Maximilian Thaler. Transformations on [0, 1] with infinite invariant measures. *Israel J. Math.*, 46(1-2):67–96, 1983.

[Tha02]  Maximilian Thaler. A limit theorem for sojourns near indifferent fixed points of one-dimensional maps. *Ergodic Theory Dynam. Systems*, 22(4):1289–1312, 2002.

(Aaronson) School of Math. Sciences, Tel Aviv University, 69978 Tel Aviv, Israel.

*Email address*: aaro@tau.ac.il

(Nakada) Dept. Math., Keio University,Hiyoshi 3-14-1 Kohoku, Yoko-hama 223, Japan

*Email address*: `nakadamath.keio.ac.jp`