

ALMOST SURE CENTRAL LIMIT THEOREM FOR STRICTLY STATIONARY PROCESSES

EMMANUEL LESIGNE

(Communicated by Stanley Sawyer)

ABSTRACT. On any aperiodic measure preserving system, there exists a square integrable function such that the associated stationary process satisfies the Almost Sure Central Limit Theorem.

INTRODUCTION

The Almost Sure Central Limit Theorem (ASCLT), first formulated by Lévy in [9], has been studied by various authors at the end of the eighties ([6], [3], [10], [8]). This theorem gives conditions under which, for a sequence of random variables satisfying the Central Limit Theorem (CLT), the Gaussian asymptotic behaviour can be observed along individual trajectory of the process. In the Lacey and Philipp note [8], the ASCLT is stated under optimal hypotheses, and the proof is short and clear. Here is their result. (If x is a real number, notation $\delta(x)$ will be used for the Dirac mass at point x .)

Theorem. *Let $(X_n)_{n \geq 1}$ be an independent and identically distributed sequence of square integrable real random variables with $\mathbb{E}(X_n) = 0$ and $\mathbb{E}(X_n^2) = 1$.*

Almost surely, the sequence of probability distributions

$$\left(\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta \left(\frac{X_1 + X_2 + \cdots + X_k}{\sqrt{k}} \right) \right)_{n \geq 1}$$

converges weakly to the Gaussian law $N(0, 1)$.

Several authors, including Berkes and Dehling ([2]), Atlagh and Weber ([1]), and Lacey have observed that for i.i.d. sequences, the finite second moment condition is necessary for the ASCLT. So, in this context, necessary and sufficient conditions for the CLT and the ASCLT are the same.

This paper is a contribution to the study of the general case of strictly stationary sequences.

The question of the CLT for strictly stationary processes has been extensively studied, from various points of view. Given a probability measure preserving dynamical system $(\Omega, \mathcal{T}, \mu, T)$ and a real measurable function f on Ω we say that

Received by the editors June 14, 1998 and, in revised form, July 22, 1998.

1991 *Mathematics Subject Classification.* Primary 28D05, 60G10, 60F05.

Key words and phrases. Almost sure central limit theorem, triangular arrays, stationary processes, measure preserving dynamical systems, approximation by Gaussian processes.

the process $(f \circ T^n)_{n \geq 0}$ satisfies the CLT if the sequence $\left(\frac{1}{\sqrt{n}} \sum_{k < n} f \circ T^k\right)_{n > 0}$ converges in distribution to a Gaussian law. One can ask in particular:

(a) Under hypothesis on the dynamical system, implying some stochastic behaviour, is it possible to find class of functions f on Ω such that the process $(f \circ T^n)_{n \geq 0}$ satisfies a CLT ?

(b) What is the generic asymptotic behaviour of the distributions of $\frac{1}{\sqrt{n}} \sum_{k < n} f \circ T^k$ when the function f belongs, for example, to $L^2(\mu)$?

(c) Does there exist a function f on Ω such that the process $(f \circ T^n)_{n \geq 0}$ satisfies a CLT ?

We shall not discuss here question (a). Question (b) has been studied by Volny in [11]. A positive answer to question (c) has been given by Burton and Denker in [4]. Around question (c), interesting constructions and refinements have been given in particular by Lacey in [7] and by Volny in [12].

Each of these questions can be asked for the ASCLT and in this note we study question (c). We use Volny's construction of a stationary process well approximated by a Gaussian i.i.d. process on an arbitrary dynamical system, and we follow the proof of the ASCLT given by Lacey and Philipp. We prove that, on an arbitrary nontrivial measure preserving system, there exists a function f such that the sequence $\left(\frac{1}{\sqrt{n}} \sum_{k < n} f \circ T^k\right)$ satisfies the ASCLT.

1. STATEMENTS

By $(\Omega, \mathcal{T}, \mu, T)$ we shall denote a dynamical system where T is an invertible measure preserving transformation of the probability space $(\Omega, \mathcal{T}, \mu)$. If f is a real function defined on Ω and if k is a positive integer, we shall denote $S_k f$ the Birkhoff

$$\text{sum } S_k f := \sum_{j=1}^k f \circ T^j.$$

If x is a real number, the notation $\delta(x)$ will be used for the Dirac mass at point x .

Theorem 1. *Let $(\Omega, \mathcal{T}, \mu, T)$ be an aperiodic and ergodic measure preserving dynamical system. There exists a square integrable and zero mean function f on the probability space $(\Omega, \mathcal{T}, \mu)$ such that, almost surely, the sequence of probability distributions*

$$\left(\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta\left(\frac{S_k f}{\sqrt{k}}\right)\right)_{n \geq 1}$$

converges weakly to the Gaussian law $N(0, 1)$.

Remarks.

- Using the fact that functions of the form $g \circ T - g$, with $g \in L^\infty(\mu)$, are dense in the space $L_0^2(\mu)$ of square integrable zero mean functions, we can deduce immediately that the set of functions f satisfying the conclusion of Theorem 1 is dense in $L_0^2(\mu)$.

- Replacing the \sqrt{k} by the right normalization, a similar result can be proved for stable laws with index $\alpha \in]1, 2[$ in place of the normal law.

The proof of Theorem 1 is based on a version of the ASCLT for triangular arrays of random variables and on a construction given by Dalibor Volny in his study of

invariance principles for strictly stationary processes. We state now an almost sure CLT for triangular arrays.

Theorem 2. *Let $(Z_{n,j})_{1 \leq j \leq n}$ be a triangular array of random variables satisfying the following five conditions:*

$$(1) \quad \text{for each } (n, j), \mathbb{E}(Z_{n,j}) = 0 ,$$

$$(2) \quad \text{random sequences } (Z_{n,1})_{n \geq 1}, (Z_{n,2})_{n \geq 2}, (Z_{n,3})_{n \geq 3}, \dots \text{ are mutually independent,}$$

$$(3) \quad \text{there exists a constant } C \text{ such that, for each } (n, j), \mathbb{E}(|Z_{n,j}|^2) \leq \frac{C}{n} ,$$

$$(4) \quad \lim_{n \rightarrow \infty} \sum_{j=1}^n \mathbb{E}(|Z_{n,j}|^2) = 1 ,$$

$$(5) \quad \text{for all } \epsilon > 0, \quad \lim_{n \rightarrow \infty} \sum_{j=1}^n \mathbb{E}(|Z_{n,j}|^2; |Z_{n,j}| > \epsilon) = 0 .$$

Then, almost surely, the sequence of probability distributions

$$\left(\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta \left(\sum_{j=1}^k Z_{k,j} \right) \right)_{n \geq 1}$$

converges weakly to the Gaussian law $N(0, 1)$.

By the classical Lindeberg CLT, conditions (1), (2), (4) and (5) imply that the sequence of random variables $\left(\sum_{j=1}^k Z_{k,j} \right)_{k \geq 1}$ converges in distribution to the Gaussian law. Notice that, in order to obtain this CLT, condition (2) can be replaced by the weaker following condition:

$$(6) \quad \text{for each } n, \text{ random variables } Z_{n,j}, (1 \leq j \leq n), \text{ are mutually independent.}$$

But in Theorem 2, condition (2) cannot be replaced by condition (6). Indeed it is possible to construct a triangular array $(Z_{n,j})_{1 \leq j \leq n}$ satisfying the five conditions

(1), (6), (3), (4), (5) and such that the sequence of random variables $\left(\sum_{j=1}^k Z_{k,j} \right)_{k \geq 1}$ is constant.

To deduce Theorem 1 from Volny's construction and Theorem 2, we shall use the following approximation result.

Proposition 1. *Let $(V_n)_{n \geq 1}$ and $(W_n)_{n \geq 1}$ be two sequences of random variables such that:*

1. *the sequence (V_n) satisfies the ASCLT, that is to say, almost surely, the sequence of probability measures*

$$\left(\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta(V_k) \right)_{n \geq 1}$$

converges weakly to the Gaussian law $N(0, 1)$;

2. for all $\epsilon > 0$, there exists $\delta > 0$ such that

$$\mathbb{P}(|V_n - W_n| > \epsilon) = O\left(\frac{1}{(\log n)^\delta}\right).$$

Then the sequence (W_n) satisfies the ASCLT.

2. PROOF OF THEOREM 2

This proof follows the lines of the proof of the almost sure CLT given by Lacey and Philipp in [8]. We shall denote $S_n := \sum_{j=1}^n Z_{n,j}$.

We want to prove, under conditions (1)-(5), that almost surely, for every compactly supported continuous function f on \mathbb{R} we have

$$(7) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} f(S_k) = \int_{\mathbb{R}} f d\nu$$

where ν is the Gaussian law.

The space of compactly supported continuous functions, equipped with the uniform norm, contains a dense denumerable subset of lipschitz functions. Thus it is enough to prove that, for every lipschitz bounded function f on \mathbb{R} , property (7) is almost surely true. Let us fix a function f , bounded by one, and a positive constant L such that for all real numbers x and y , we have $|f(x) - f(y)| \leq L|x - y|$. The Lindeberg CLT just tells us that

$$\lim_{k \rightarrow \infty} \mathbb{E}[f(S_k)] = \int_{\mathbb{R}} f d\nu.$$

This implies that

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbb{E}[f(S_k)] = \int_{\mathbb{R}} f d\nu$$

and we shall prove that, almost surely,

$$(8) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \left(f(S_k) - \mathbb{E}[f(S_k)] \right) = 0.$$

This law of large numbers will be obtained by an estimation of the correlations of random variables S_k . It will be a direct consequence of the next three lemmas.

Lemma 1. *There is a constant D , such that if $1 \leq k < l$, then*

$$\left| \text{cov}\left(f(S_k), f(S_l)\right) \right| \leq D\sqrt{\frac{k}{l}}.$$

Lemma 2. *If $(R_k)_{k \geq 1}$ is a uniformly bounded sequence of centered random variables such that, for $1 \leq k < l$,*

$$\left| \mathbb{E}[R_k R_l] \right| \leq D\sqrt{\frac{k}{l}}$$

then, almost surely,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left(\sum_{j=4^k}^{j=4^{k+1}} \frac{1}{j} R_j \right) = 0.$$

Lemma 3. Let $(r_k)_{k \geq 1}$ be a bounded sequence of real numbers. If

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left(\sum_{j=4^k}^{j=4^{k+1}} \frac{r_j}{j} \right) = 0,$$

then

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{r_k}{k} = 0.$$

Proof of Lemma 1. Thanks to the independence hypothesis (2) we have

$$\text{cov}\left(f(S_k), f(S_l)\right) = \text{cov}\left(f(S_k), f(S_l) - f\left(\sum_{j=k+1}^l Z_{l,j}\right)\right).$$

With our hypothesis on f this implies

$$\left| \text{cov}\left(f(S_k), f(S_l)\right) \right| \leq 2L\mathbb{E}\left[\sum_{j=1}^k Z_{l,j}\right].$$

Using Cauchy-Schwarz, (1),(2) and (3), we deduce that

$$\left| \text{cov}\left(f(S_k), f(S_l)\right) \right| \leq 2L\sqrt{C}\sqrt{\frac{k}{l}}.$$

□

Proof of Lemma 2. Let us note $U_k := \sum_{j=4^k}^{j=4^{k+1}-1} \frac{1}{j} R_j$. For $1 \leq k < l$, we have

$$\mathbb{E}[U_k U_l] \leq D \sum_{j=4^k}^{j=4^{k+1}-1} \sum_{i=4^l}^{i=4^{l+1}-1} \frac{1}{ji} \sqrt{\frac{j}{i}} \leq D \left(\sum_{j < 4^{k+1}} \frac{1}{\sqrt{j}} \right) \left(\sum_{i < 4^{l+1}} \frac{1}{i\sqrt{i}} \right) = O(2^{k-l}).$$

The law of large numbers for bounded orthogonal sequences of square integrable random variables is a classical result (see e.g. [5], thm 5.1.2). The proof can be easily extended to the case of bounded sequences satisfying a weak correlation property like the sequence (U_n) . Thus, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n U_k = 0 \quad \text{a.s.}$$

and Lemma 2 is proved.

The proof of Lemma 3 is an easy exercise, and Theorem 2 is a direct application of Lemmas 2 and 3 to $R_k := f(S_k) - \mathbb{E}[f(S_k)]$.

3. PROOF OF PROPOSITION 1

Let us fix $\epsilon > 0$ and show that hypothesis 2 of Proposition 1 implies that

$$(9) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{1}_{(|V_k - W_k| > \epsilon)} = 0 \quad \text{a.s.}$$

By hypothesis 2, in which we can suppose $0 < \delta < 1$, we have

$$\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbb{P}(|V_k - W_k| > \epsilon) = O\left(\frac{1}{\log n} \sum_{k=2}^n \frac{1}{k} (\log k)^\delta\right) = O\left(\frac{1}{(\log n)^\delta}\right).$$

We fix an integer p such that $p\delta > 1$ and we have

$$\mathbb{E}\left(\frac{1}{n^p} \sum_{k=1}^{2^{n^p}} \frac{1}{k} \mathbf{1}_{(|V_k - W_k| > \epsilon)}\right) = O\left(\frac{1}{n^{p\delta}}\right),$$

$$\sum_{n > 0} \mathbb{E}\left(\frac{1}{n^p} \sum_{k=1}^{2^{n^p}} \frac{1}{k} \mathbf{1}_{(|V_k - W_k| > \epsilon)}\right) < +\infty;$$

thus

$$(10) \quad \lim_{n \rightarrow \infty} \frac{1}{n^p} \sum_{k=1}^{2^{n^p}} \frac{1}{k} \mathbf{1}_{(|V_k - W_k| > \epsilon)} = 0 \quad \text{a.s.}$$

If N is a positive integer, let n be the integer such that $2^{n^p} \leq N < 2^{(n+1)^p}$. We have

$$\begin{aligned} \frac{1}{\log N} \sum_{k=1}^N \frac{1}{k} \mathbf{1}_{(|V_k - W_k| > \epsilon)} &\leq \frac{1}{n^p} \sum_{k=1}^{2^{(n+1)^p}} \frac{1}{k} \mathbf{1}_{(|V_k - W_k| > \epsilon)} \\ &\leq \left(\frac{n+1}{n}\right)^p \frac{1}{(n+1)^p} \sum_{k=1}^{2^{(n+1)^p}} \frac{1}{k} \mathbf{1}_{(|V_k - W_k| > \epsilon)} \end{aligned}$$

and, thanks to (10), this implies (9).

Let a and b be two rational numbers with $a < b$. We have

$$\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{1}_{[a,b]}(W_k) \leq \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{1}_{[a-\epsilon, b+\epsilon]}(V_k) + \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{1}_{(|V_k - W_k| > \epsilon)}.$$

Using (9) and hypothesis 1, and denoting by \mathcal{N} the distribution function of the Gaussian law, we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{1}_{[a,b]}(W_k) \leq \mathcal{N}(b + \epsilon) - \mathcal{N}(a - \epsilon) \quad \text{a.s.}$$

and, letting ϵ go to zero,

$$\limsup_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{1}_{[a,b]}(W_k) \leq \mathcal{N}(b) - \mathcal{N}(a) \quad \text{a.s.}$$

By a similar argument we can show that the liminf is almost surely greater than $\mathcal{N}(b) - \mathcal{N}(a)$. Thus, almost surely, for every rational interval I ,

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \mathbf{1}_I(W_k) = \nu(I)$$

and this is enough to prove the weak convergence of the sequence

$$\left(\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta(W_k)\right).$$

4. PROOF OF THEOREM 1

Let $(\Omega, \mathcal{T}, \mu, T)$ be an aperiodic and ergodic measure preserving dynamical system. From [12] we can deduce the following claim: there exist a square integrable random variable f , and an array $(Y_{n,j})_{1 \leq j \leq n}$ of random variables defined on $(\Omega, \mathcal{T}, \mu)$, such that

$$(11) \quad \text{for each } n, Y_{n,1}, Y_{n,2}, \dots, Y_{n,n} \text{ are identically distributed,}$$

$$(12) \quad \mathbb{E}(Y_{n,j}) = 0 \text{ and } \lim_{n \rightarrow \infty} \mathbb{E}((Y_{n,j})^2) = 1 ,$$

$$(13) \quad \text{random sequences } (Y_{n,1})_{n \geq 1}, (Y_{n,2})_{n \geq 2}, (Y_{n,3})_{n \geq 3}, \dots \text{ are mutually independent,}$$

and, with the notation $V_n := \frac{1}{\sqrt{n}} \sum_{j=1}^n Y_{n,j}$ and $W_n := \frac{1}{\sqrt{n}} \sum_{j=1}^n f \circ T^j$,

$$(14) \quad \mathbb{E}(|V_n - W_n|^2) = O\left(\frac{1}{\log n}\right) .$$

Perhaps some words of explanation are necessary here: using Volny's notation from [12] we can define

$$Y_{n,j} = \sum_{k: 2^k < n < 3^k} p_k (X_{k,j} - X_{k,j+d_k})$$

and

$$f = \sum_{k=1}^{\infty} \sum_{i=0}^{p_k-1} (\overline{f_k} \circ T^i - \overline{f_k} \circ T^{i+d_k})$$

where the $X_{k,j}$, $1 \leq k, 1 \leq j < 2 \times 3^k$, are mutually independent centered random variables defined on Ω , whose distributions do not depend on j , which insures (11). Variances of these random variables are chosen so that (12) is satisfied. For each j , the random sequence $(Y_{n,j})_{n \geq j}$ is measurable with respect to $\sigma(X_{k,j}, X_{k,j+d_k} \mid j < 3^k)$, and the sets $\{(k, j), (k, j + 3^k) \mid j < 3^k\}$ are two by two disjoint when j varies. This proves that condition (13) is satisfied.

Functions $\overline{f_k}$ are in $L_0^2(\mu)$ and satisfies, for any $k \geq 1$ and $0 \leq i \leq 2 \times 3^k$,

$$\mathbb{E}(|X_{k,i} - \overline{f_k} \circ T^i|^2) < 6^{-6k}$$

and affirmation (14) is contained in Theorem 1 of [12].

The claim is justified.

Conditions (11), (12), (13) imply that the variables V_n converge in distribution to a Gaussian law. Thanks to (14), we can conclude that the sequence (W_n) converges in distribution to a Gaussian law.

Conditions (11), (12), (13) allow us to apply Theorem 2 to the variables $Z_{n,j} := \frac{1}{\sqrt{n}} Y_{n,j}$. Thus, the sequence (V_n) satisfies the ASCLT. Condition (14) implies hypothesis 2 of Proposition 1; therefore the sequence (W_n) satisfies the ASCLT, which is the conclusion of Theorem 1.

5. A REMARK AND A QUESTION

From the main theorem of [11] Volny and Weber deduced the following result which goes in the opposite direction of our Theorem 1.

Proposition 2 ([13]). *Let $(\Omega, \mathcal{T}, \mu, T)$ be an aperiodic and ergodic measure preserving dynamical system. Let $(a(n))_{n \in \mathbb{N}}$ be a real sequence such that $\lim_{n \rightarrow \infty} a(n) = +\infty$ and $\lim_{n \rightarrow \infty} \frac{a(n)}{n} = 0$.*

There exists a dense G_δ set of functions f in $L_0^2(\mu)$ satisfying: for any Borel probability measure ν on \mathbb{R} , there exists an increasing sequence of integers (n_k) such that, almost surely, the sequence

$$\left(\frac{1}{K} \sum_{k=1}^K \delta \left(\frac{1}{a(n_k)} S_{n_k} f \right) \right)$$

converges weakly to ν .

Of course, the convergence of these ordinary Cesaro means implies the convergence of logarithmic means

$$\left(\frac{1}{\log K} \sum_{k=1}^K \frac{1}{k} \delta \left(\frac{1}{a(n_k)} S_{n_k} f \right) \right)$$

to the same limit.

In the same vein we can ask the following question: does there exist a function f on Ω such that for any Borel probability measure ν on \mathbb{R} , there exists an increasing sequence of integers (N_k) such that, almost surely,

$$\lim_{k \rightarrow \infty} \frac{1}{\log N_k} \sum_{n=1}^{N_k} \frac{1}{n} \delta \left(\frac{S_n f}{\sqrt{n}} \right) = \nu \quad ?$$

ACKNOWLEDGEMENT

The author thanks Dalibor Volny for many instructive conversations and Yves Derriennic who brings reference [9] to his attention.

REFERENCES

- [1] Atlagh, M. & Weber, M. *Une nouvelle loi forte des grands nombres*. Convergence in Ergodic Theory and Probability, Eds.:Bergelson/March/Rosenblatt, Walter de Gruyter & Co. (1996) MR **97i**:60034
- [2] Berkes, I & Dehling H. *On the almost sure central limit theorem for random variables with infinite variance*. J. Theor. Probab. **7**, p.667-680. (1994) MR **95g**:60041
- [3] Brosamler, G.A. *An almost everywhere central limit theorem*. Math. Proc. Cambridge Phil. Soc. **104**, p.561-574. (1988) MR **89i**:60045
- [4] Burton, R. & Denker, M. *On the central limit theorem for dynamical system*. Transactions Amer. Math. Soc. **302**, p.715-726. (1987) MR **88i**:60039
- [5] Chung, K.L. : A course in probability theory. (1975)
- [6] Fisher, A. *Convex-invariant means and a pathwise central limit theorem*. Advances in Math. **63**, p.213-246. (1987) MR **88g**:60058
- [7] Lacey, M. *On central limit theorems, modulus of continuity and Diophantine type for irrational rotations*. J. d'Analyse Math. **61**, p.47-59. (1993) MR **95a**:60054
- [8] Lacey, M. & Philipp, W. *A note on the almost sure central limit theorem*. Statist. Probab. Letters **9**, p.201-205. (1990) MR **91e**:60100
- [9] Lévy, P. : Théorie de l'addition des variables aléatoires (chapitre VIII). (1937)

- [10] Schatte, P. *On strong versions of the central limit theorem*. Math. Nachr. **137**, p.249-256. (1988) MR **89i**:60070
- [11] Volny, D. *On limit theorems and category for dynamical systems*. Yokohama Math. J. **38**, p.29-35. (1990) MR **92c**:28014
- [12] Volny, D. *Invariance principles and Gaussian approximation for strictly stationary processes*. Preprint (1996), to be published in Transactions Amer. Math. Soc. CMP 98:13
- [13] Volny, D. & Weber, M. : personal communication. (1996)

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ FRANÇOIS RABELAIS, PARC DE GRANDMONT,
37200 TOURS, FRANCE

E-mail address: `lesigne@univ-tours.fr`