

Int. Journal of Math. Analysis, Vol. 7, 2013, no. 15, 699 - 712

## Bounds for Estimators of Ergodic Averages

Alejandro Mesón and Fernando Vericat

Instituto de Física de Líquidos y Sistemas Biológicos (IFLYSIB)  
CONICET–UNLP and Grupo de Aplicaciones Matemáticas y Estadísticas  
de la Facultad de Ingeniería (GAMEFI) UNLP, La Plata, Argentina

### Abstract

We consider different ergodic averages and estimate the measure of the set of points in which the averages apart from a given value. The cases considered are empirical measures of cylinders in symbolic spaces and averages of maps given a kind Lyapunov exponents, in a such spaces. Besides we obtain bounds for the fluctuations of ergodic averages from amenable action groups. The bounds obtained are valid for any “time”, not only, like in case of large deviations, for asymptotic values.

**Keywords:** ergodic averages, bound estimators, empirical measures

## 1 Introduction

In [3] was estimated the entropy of a “sample”  $x_0, x_1, \dots$ , corresponding to a discrete-time ergodic process  $\{X_n : n \in \mathbf{N}\}$ , i.e.  $X_n = X \circ f^n$ , where  $f$  is an ergodic measure-preserving transformation and  $X$  is a measurable function. The problem is analyzed in the setting of symbolic dynamical systems and so any sample may be seen as an element of symbolic space  $(\Sigma, \sigma)$  of infinite sequences in a finite alphabet  $\Omega$  and with  $\sigma : \Sigma \rightarrow \Sigma$  the usual shift. The so called “*plug-in estimator*” is related to the frequency of a sequence  $y_0 y_1 \dots y_{k-1}$  in a block sample  $x_0 x_1 \dots x_{n-1} \in \Omega^n$ , extending the block sample to a periodic sequence  $\bar{x} = x_0 x_1 \dots x_{n-1} x_0 x_1 \dots x_{n-1} \dots \in \Sigma$ , the plug-in can be considered as a  $\sigma$ -invariant measure in  $\Sigma$ . In [3] are estimated the fluctuations of the entropy of the plug-in for a class a Gibbs measures associated to real-valued maps on  $\Sigma$ . The main tool for the estimates is a theorem for separately Lipschitz  $n$ -variable functions  $\Phi : \Sigma^n \rightarrow \mathbf{R}$  proved in [5]. Separately Lipschitz means that

$$Lip_j(\Phi) := \sup_{x^{(0)}, \dots, x^{(n-1)} \in \Sigma} \sup_{x^{(j)} \neq y^{(j)}} \frac{|\Phi(x^{(0)}, \dots, x^{(j)}, \dots, x^{(n-1)}) - \Phi(x^{(0)}, \dots, y^{(j)}, \dots, x^{(n-1)})|}{d(x^{(j)}, y^{(j)})} < \infty, \tag{1}$$

for any  $j = 0, 1, \dots, n - 1$ .

**Theorem A** [5]: If  $\varphi : \Sigma \rightarrow \mathbf{R}$  is a potential which has a Gibbs measure  $\mu_\varphi$  and  $\Phi$  is a *separately Lipschitz* function of  $n$ -variables then holds:

$$\int \exp[\Phi(x, \sigma x, \dots, \sigma^{n-1}x)] d\mu_\varphi(x) \leq \exp\left[\int \Phi(y, \sigma y, \dots, \sigma^{n-1}y) d\mu_\varphi(y)\right] \exp\left[D \sum_{j=0}^{n-1} Lip_j^2(\Phi)\right],$$

where  $D$  is a constant depending only on  $\varphi$ . In [3] were proved the following estimations, as corollaries of the above formula:

$$\mu_\psi\left(\left\{x : \Phi(x, \sigma x, \dots, \sigma^{n-1}x) - \int \Phi(y, \sigma y, \dots, \sigma^{n-1}y) d\mu_\varphi(y) > \alpha\right\}\right) \leq 2 \exp\left[\frac{\alpha^2}{4D \sum_{j=0}^{n-1} Lip_j^2(\Phi)}\right], \tag{2}$$

and

$$\mu_\psi\left(\left\{x : \frac{1}{n} \sum_{i=0}^{n-1} \varphi(\sigma^i(x)) - \int \varphi d\mu_\psi > \alpha\right\}\right) \leq \exp(B\alpha n^2), \tag{3}$$

where  $\mu_\psi$  is the Gibbs state for the potential  $\psi$  and  $B$  is a constant depending on  $D$  and  $\varphi$ .

The variance of  $\Phi$  can be bounded by

$$\begin{aligned} Var(\Phi) &:= \int \left(\Phi(x, \sigma x, \dots, \sigma^{n-1}x) - \int \Phi(y, \sigma y, \dots, \sigma^{n-1}y) d\mu_\varphi(y)\right)^2 d\mu_\varphi(x) \\ &\leq 2D \sum_{j=0}^{n-1} Lip_j^2(\Phi) \end{aligned} \tag{4}$$

In some cases instead of function defined on  $\Sigma^n$  can be considered maps of  $n$ -variables  $\Phi$  defined on  $\Omega^n$ , in this case instead of *Lip* is taken the oscillation in each coordinate

$$\delta_j(\Phi) := \sup_{x_0, \dots, x_{n-1} \in \Sigma} \sup_{x_j \neq y_j} |\Phi(x_0, \dots, x_j, \dots, x_{n-1}) - \Phi(x_0, \dots, y_j, \dots, x_{n-1})|.$$

The earlier results are valid with  $\delta_j$  instead of *Lip* <sub>$j$</sub> [3].

In this article we shall consider ergodic averages which can be estimated by the above result, i.e. by mean of adequate functions  $\Phi$  defined accordingly to the problem. We also work in situations more general than symbolics, specifically we consider amenable action groups and analyze fluctuations of Lindenstrauss ergodic averages[6].

## 2 Previous definitions and notations

Let  $(\Sigma, \sigma)$  be the symbolic space of infinite sequences in a finite alphabet  $\Omega$  and with  $\sigma : \Sigma \rightarrow \Sigma$  the Bernoulli shift  $(\sigma x)_n = x_{n+1}$ . This space can be endowed with the metric

$d_t(x, y) = t^{\overline{N}}$ , where  $t \in (0, 1)$  and  $N = \max \{j : x_j = y_j, 0 \leq j < \overline{N}\}$ . More generally can be introduced metrics in  $\Sigma$  in this way:

let  $x \wedge y$  be the common prefix of  $x$  and  $y$  and  $N$  the unique integer such that  $x_j = y_j, 0 \leq j < N$  and  $x_N \neq y_N$ . Let  $\psi : \bigcup_{n \geq 0} \Omega^n \rightarrow \mathbf{R}$  be a non-increasing map such that  $\lim_{n \rightarrow \infty} \psi(x_0, \dots, x_{n-1}) = 0$ . Thus

$$d(x, y) = \psi(x \wedge y),$$

is a metric in  $\Sigma$ [4]. When  $\psi = t^{\overline{N}}$ , is obtained the metric  $d_t$ .

A map  $f : X \rightarrow X$  with a compact metric space is *conformal* if the function  $a : X \rightarrow [0, \infty)$  defined by

$$a(x) = \lim_{y \rightarrow x} \frac{d(f(x), f(y))}{d(x, y)} \tag{5}$$

is continuous. When  $0 < a(x) < \infty$ , for any  $x$ , it says that  $f$  is without critical points nor singularities.

If  $X$  is a Riemannian manifold then  $f$  is said to be conformal if

$$\|Df(x)\| = a(x)I(x) \text{ where } I \text{ is an isometry of } X.$$

If  $x = x_0 x_1 \dots x_{n-1} \dots \in \Sigma$ , then by  $[x]_n$  will be denoted the truncation to its first  $n$ -symbols, the cylinder of length  $n$  containing  $x$  is

$\mathcal{C}_n(x) = \{y : [x]_n = [y]_n\}$ . The entropy of a measure  $\nu$  is

$$h(\nu) = \lim_{n \rightarrow \infty} \left[ - \sum_{[x]_n \in \Omega^n} \nu(\mathcal{C}_n(x)) \log \nu(\mathcal{C}_n(x)) \right]$$

and holds  $h(\nu) \leq \log \Omega$ .

For conformal maps in metric spaces Climenahaga[4] gave a special a definition of Lyapunov exponents

$$\begin{aligned} \bar{\lambda}(x) &= \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \log(a(f^i(x))) \\ \underline{\lambda}(x) &= \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \log(a(f^i(x))), \end{aligned}$$

where the limit exists when is verified the condition of tempered contraction[4]. The Lyapunov exponent of the measure  $\nu$  is

$$\lambda(\nu) = \int \lambda(x) d\nu.$$

In the case of symbolic spaces the map  $a$  can be defined as

$$a(x) = \lim_{y \rightarrow x} \frac{d(\sigma(x), \sigma(y))}{d(x, y)} = \lim_{y \rightarrow x} \frac{\psi(\sigma(x) \wedge \sigma(y))}{\psi(x \wedge y)} = \frac{\psi(x_1, \dots, x_{n-1})}{\psi(x_0, \dots, x_{n-1})}, \quad (6)$$

where the case  $\psi = t^{\bar{N}}$  gives  $a(x) = t$  for any  $x$  and  $\lambda(x) = t$ .

Let  $x_0 x_1 \dots x_{n-1}$  be a sequence in  $\Omega^n$  and  $y_0 y_1 \dots y_{k-1}$  be a sequence in  $\Omega^k$ . From  $x_0 x_1 \dots x_{n-1}$  can be formed a periodic sequence  $x^n = x_0 x_1 \dots x_{n-1} x_0 x_1 \dots x_{n-1} \dots \in \Sigma$ , now let us consider the plug-in estimator

$$\mathcal{E}_k(x_0 x_1 \dots x_{n-1} : y_0 y_1 \dots y_{k-1}) = \frac{1}{n} \text{card} \{j = 0, \dots, n-1 : [\sigma^j x^n]_k = y_0 y_1 \dots y_{k-1}\}. \quad (7)$$

This estimator is related to the empirical frequency of the sequence  $y_0 y_1 \dots y_{k-1}$  in the sample  $x_0 x_1 \dots x_{n-1}$ . By  $\mathcal{M}(X)$  is denoted the space of measures on the space  $X$ . Due to the extension of the sample to an infinite sequence the plug-in can be seen as empirical measure

$$\mathcal{E}_n(x) = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{\sigma^i x} \in \mathcal{M}(\Sigma),$$

where  $\delta$  is the point-mass measure. If  $y_0 y_1 \dots y_{k-1}$  is the  $k$ -truncation of a sequence  $y \in \Sigma$ , then the plug-in (7) is seen as  $\mathcal{E}_n(x)(\mathcal{C}_k(y))$ . For any ergodic measure  $\nu$  holds

$$\lim_{k \rightarrow \infty} \mathcal{E}_k(x^n)(\mathcal{C}_k(y)) = \nu(\mathcal{C}_k(y))$$

for any  $x^n$  in a set of  $\nu$ -measure one. By Ornstein and Weiss (see e.g. [8]) can be taken  $k$  depending on  $n$  and setting  $\mathcal{E}_{k(n)}(x^n)$  and the limit with  $n \rightarrow \infty$ . In [7] were obtained large-deviation results for empirical measures, in that article was described a large deviation process for

$\nu(\{x : \mathcal{E}_k(x) \in K\})$ , with  $K \subset \mathcal{M}(\Sigma)$ . Herein we analyze the measure of the fluctuations  $\{x : \mathcal{E}_{k(n)}(x^n)(\mathcal{C}_{k(n)}(y)) - \nu(\mathcal{C}_{k(n)}(y)) > \alpha\}$ , for any  $n$ .

For Lyapunov exponents in symbolic spaces we shall study the measure of the fluctuations  $\frac{1}{n} \sum_{i=0}^{n-1} \log(a(\sigma^i(x))) - \lambda(x)$ , for any  $n$ , with  $a(x)$  defined in (6), this can be worked out from (2). However, the bound depends on the variation of  $a(x)$ , this value is contained in the constant  $B$  in (3), and in case of using the metric  $d_t$  recall that  $a$  is constant, so, for our purposes, it must be considered another metrics from the map  $\psi$  as described above.

Of course in a more general situation than symbolics the techniques based on the oscillation of separately Lipschitz functions cannot be used. We consider here as a more general situation amenable action groups and analyze the fluctuations of ergodic averages from the action. Let us recall basic definitions in this setting:

Let  $(X, \nu)$  be a measured space and  $\Gamma$  a locally compact groups acting on it, by  $m_\Gamma$  is denoted the Haar measure on  $\Gamma$ . A group  $\Gamma$  is *amenable* if for any compact  $K \subset \Gamma$  and  $\delta > 0$  there is a set  $F \subset \Gamma$  such that

$$\frac{m_\Gamma(F \Delta K F)}{m_\Gamma(F)} \leq \delta.$$

A such set  $F$  is called  $(K, \delta)$ -invariant. A sequence  $(F_n)$ ,  $F_n \subset \Gamma$ , is a *Følner sequence* if each  $F_n$  is  $(K, \delta)$ -invariant, for any compact  $K \subset \Gamma$  and  $\delta > 0$  and

$$\lim_{n \rightarrow \infty} \frac{m_\Gamma(F_n \Delta K F_n)}{m_\Gamma(F_n)} = 0.$$

Any amenable group has a Følner sequence[6]. From the action  $(x, \gamma) \rightarrow \gamma x$  and for a Følner sequence  $(F_n)$  can be defined an ergodic average, let  $\varphi : X \rightarrow \mathbf{R}$  and set

$$S_{F_n, \varphi}(x) = \frac{1}{m_\Gamma(F_n)} \int_{F_n} \varphi(\gamma x) dm_\Gamma(\gamma). \tag{8}$$

When  $\Gamma = \mathbf{Z}$ ,  $F_n = \{1, 2, \dots, n\}$  and  $f : X \rightarrow X$  obtains the usual ergodic average

$$\frac{1}{n} \sum_{i=0}^{n-1} \varphi(ix), \text{ where } ix = f^i(x).$$

A sequence  $(F_n)$  of subsets of  $\Gamma$  is *tempered* if there is a constant  $C > 0$  such that

$$m_\Gamma \left( \bigcup_{k=1}^{n-1} F_k^{-1} F_n \right) \leq C m_\Gamma(F_n). \text{ Any Følner sequence has a tempered subsequence[6].}$$

An important result of Lindenstrauss[6] is that if  $\varphi \in L^1(X, \nu)$  then  $S_{F_n, \varphi}(x)$  converges to a  $\Gamma$ -invariant map  $\bar{\varphi}$ , for any  $\nu - a.e.x$  and when the action is ergodic the map is constant  $\nu - a.e.$ , this generalizes the Birkhoff ergodic theorem.

In this article we shall estimate the measure of sets  $\{x : S_{F_n, \varphi}(x) - \bar{\varphi}(x) > \alpha\}$ .

### 3 Results

**Theorem 1:** Let  $(\Sigma, \sigma)$  be the symbolic space of infinite sequences in a finite alphabet  $\Omega$ , let  $\varphi : \Sigma \rightarrow \mathbf{R}$  be a map having an unique Gibbs state  $\mu_\varphi$  ( for instance a Hölder continuous function) and  $\theta \in (0, 1)$ , then for any  $n, y$  holds:

$$\mu_\varphi \left( \{ \mathcal{E}_{k(n)}(x^n) (\mathcal{C}_{k(n)}(y)) - \nu(\mathcal{C}_{k(n)}(y)) > \alpha \} \right) \leq \exp \left[ - \frac{\alpha^2 n^{1-\theta}}{4D \log n} \right],$$

where  $D$  is the constant of the eq. (3) and with  $k(n) \leq \frac{\theta \log n}{2 \log(\text{card}\Omega)}$ . Besides the variance of  $\mathcal{E}_{k(n)}$  is estimated by

$$\text{Var} \left( \mathcal{E}_{k(n)}(\mathcal{C}_{k(n)}(y)) \right) \leq 2Dn^{1-\theta} \log n.$$

*Proof:* Let  $\Phi = \Phi_y(x_0 x_1 \dots x_{n-1}) = \mathcal{E}_{k(n)}(x^n) (\mathcal{C}_{k(n)}(y))$ . Now to obtain the bounds we must estimate  $\delta_j(\Phi)$ . We have

$$\mathcal{E}_{k(n)}(\mathcal{C}_{k(n)}(y)) \leq \frac{k(n)}{n} \text{ and so}$$

$$\sum_{j=0}^{n-1} \delta_j^2(\Phi) \leq \frac{\alpha^2 \log n^2}{4n (\log(\text{card}\Omega))^2} \leq \frac{\alpha^2 (\log n^\theta)^2}{4n (\log(\text{card}\Omega))^2} \leq \frac{n^\theta \log n^\theta}{n}. \text{ Hence}$$

$$\exp \left[ -\frac{\alpha^2}{4D \sum_{j=0}^{n-1} \delta_j^2(\Phi)} \right] \leq \exp \left[ -\frac{\alpha^2}{4D \log n} n^{1-\theta} \right],$$

and by the theorem A and its consequence (eq. 3) we obtain the desired result. For the variance, in the same way, we have

$$\text{Var} \left( \mathcal{E}_{k(n)} \left( \mathcal{C}_{k(n)}(y) \right) \right) \leq 2D \sum_{j=0}^{n-1} \delta_j^2(\Phi) \leq 2D n^{1-\theta} \log n.$$

■

The next analysis is for Lyapunov exponents, we analyze the concentration of the ergodic average of  $\log(a(x))$  around its  $\mu_\varphi$ -average, i.e. the Lyapunov exponent of  $\mu_\varphi$ , recall that the Gibbs states are ergodic. The function  $a(x)$  is a map defined from a map  $\psi : \bigcup_{n \geq 0} \Omega^n \rightarrow \mathbf{R}$ . Let  $\varphi : \Sigma \rightarrow \mathbf{R}$  be a continuous map, set

$V_{n,\psi}(\varphi) = \sup \left\{ \frac{|\varphi(x) - \varphi(y)|}{\psi(x_1, \dots, x_n)} : x_i = y_i, i = 0, 1, \dots, n - 1 \right\}$ . For the next result we shall consider a map  $\psi$  such that

- $\psi \leq 1$
- $\psi(x \wedge y) = 1$  when  $x_0 \neq y_0$ .
- $\psi(x \wedge y) \geq \psi(\sigma x \wedge \sigma y) \geq \dots \geq \psi(\sigma^{n-1}x \wedge \sigma^{n-1}y)$ , for any  $n$ .

**Proposition 1:** Let  $(\Sigma, \sigma)$  be the symbolic space of infinite sequences in a finite alphabet  $\Omega$ , let  $\varphi : \Sigma \rightarrow \mathbf{R}$  be a map having an unique Gibbs state  $\mu_\varphi$ . Let  $a(x) =$

$$\lim_{y \rightarrow x} \frac{\psi(\sigma(x) \wedge \sigma(y))}{\psi(x \wedge y)}$$

and  $S_n(\log a(x)) = \sum_{i=0}^{n-1} \log(a(\sigma^i x))$ , then for any  $n$

$$\mu_\varphi \left( \left\{ x \frac{1}{n} S_n(\log a(x)) - \lambda(\mu_\varphi) \right\} \right) \leq \exp \left[ -\frac{\alpha^2 n}{2D (V_{n,\psi}(\log a(x)))^2} \right].$$

*Proof:* Define

$$\Phi(x, \sigma x, \dots, \sigma^{n-1}x) = \frac{1}{n} S_n(\log a(x)) = \log a(x) + \log a(\sigma x) + \dots + \log a(\sigma^{n-1}x),$$

thus

$$\text{Lip}_j(\Phi) := \sup_{x, \sigma x, \dots, \sigma^{n-1}x} \sup_{\sigma^j x \neq \sigma^j y} \left\{ \frac{|\log a(\sigma^j x) - \log a(\sigma^j y)|}{\psi(\sigma(x) \wedge \sigma(y))} \right\},$$

and so, by the conditions of  $\psi$ , we get  $\sum_{j=0}^{n-1} Lip_j^2(\Phi) \leq n (V_{n,\psi}(\log a(x)))^2$ . Now, again by the consequence of theorem A, the conclusion of the proposition is obtained. ■

Finally we try to get concentration bounds for ergodic averages from amenable action groups. Let  $\Gamma$  be an amenable group and let  $\{F_1, F_2, \dots, F_N\}$  be a tempered sequence in  $\Gamma$  with constant  $C$  (definitions given in the earlier section). Recall that tempered sequences can be obtained as subsequences of Følner sequences. Let  $F$  be a compact subset of  $\Gamma$ , a collection  $\overline{\mathcal{F}}$  of subsets of  $\Gamma$  can be specified by sets  $A_j$ , such that  $F_j A_j \subset F$ ,  $j = 1, 2, \dots, N$ , and

$$\overline{\mathcal{F}} = \{F_j a : a \in A_j, j = 1, 2, \dots, N\}.$$

Random subsets of  $\Gamma$  can be given by a measurable map

$$\mathcal{F} : \Lambda \rightarrow \mathcal{P}(\overline{\mathcal{F}}),$$

where  $\Lambda$  is a probability space, which is called the sample space. In [6] are introduced the norm

$$\|\mathcal{F}(\omega)\| := \sum_{B \in \mathcal{F}(\omega)} m_\Gamma(B),$$

and the counting function  $F_\omega : \Gamma \rightarrow \mathbf{N}$

$$F_\omega(\gamma) = \sum_{B \in \mathcal{F}(\omega)} I_B(\gamma),$$

where  $I_B$  is the characteristic function of  $B$ .

Let  $Z_{n,\alpha} = Z_{F_n,\alpha} = \{x : S_{F_n,\varphi}(x) > \alpha\}$  where

$$S_{F_n,\varphi}(x) = \frac{1}{m_\Gamma(F_n)} \int_{F_n} \varphi(\gamma x) dm_\Gamma(\gamma).$$

The following proposition is like to the maximal ergodic inequality in [6], but without taking supremum in  $S_{F_n,\varphi}(x)$ .

**Proposition 2:** Let  $\nu$  be a measure on  $X$ , and  $(F_n) \subset \Gamma$ , be a tempered sequence of compact sets in  $\Gamma$ , (amenable with finite Haar measure). Then holds



$$\nu(Z_{n,\alpha}) \leq \frac{M}{\alpha} \|\varphi\|_1,$$

where  $M$  is a constant depending only on  $F_n$ .

*Proof:* Since  $\Gamma$  is amenable and each  $F_n$  is compact, can be found, for each  $n$ , a compact set  $\tilde{F}_n \subset \Gamma$ , such that if  $\varepsilon > 0$  and

$$G_n = F_n \tilde{F}_n \text{ then } m_\Gamma(G_n) \leq (1 + \varepsilon) m_\Gamma(\tilde{F}_n), \text{ we shall show that holds}$$

$$\int_{\tilde{F}_n} I_{Z_{n,\alpha}}(\gamma x) dm_\Gamma(\gamma) \leq \frac{M}{\alpha} \int_{F_n} |\varphi(\gamma x)| dm_\Gamma(\gamma), \text{ for some } M > 0. \tag{9}$$

Let  $A_{n,\alpha} = \{\gamma \in \tilde{F} : S_{F_n,\varphi}(\gamma x) > \alpha\}$ ,  $F_n A_{n,\alpha} \subset F$ , and let  $\mathcal{F}(\omega)$  be a random collection of subsets of  $F$  with counting function  $F_\omega$ . For  $a \in A_{n,\alpha}$  we have

$$S_{F_n,\varphi}(ax) = \frac{1}{m_\Gamma(F_n)} \int_{F_n} \varphi(\gamma ax) dm_\Gamma(\gamma) > \alpha, \text{ and}$$

$$\alpha m_\Gamma(F_n A_{n,\alpha}) \leq \frac{m_\Gamma(F_n a)}{m_\Gamma(F_n)} \int_{F_n} \varphi(\gamma ax) dm_\Gamma(\gamma). \text{ Thus if } \Delta \text{ is the modular function on}$$

$\Gamma$  defined by

$$\int \varphi(\gamma_1 \gamma_2) dm_\Gamma(\gamma_1) = \frac{1}{\Delta(\gamma_2)} \int \varphi(\gamma_1) dm_\Gamma(\gamma_1), \text{ then}$$

$$\alpha m_\Gamma(F_n A_{n,\alpha}) \leq \Delta(a) \int_{F_n} \varphi(\gamma x) dm_\Gamma(\gamma) = \int_{F_n a} \varphi(\gamma x) dm_\Gamma(\gamma). \text{ Then}$$

$$\alpha \sum_{A_{n,\alpha}} m_\Gamma(F_n A_{n,\alpha}) = \alpha \|\mathcal{F}(\omega)\| \leq \sum_{B \in \mathcal{F}(\omega)} \int_B \varphi(\gamma x) dm_\Gamma(\gamma).$$

By the Lindenstrauss covering lemma[6], we have

$\alpha \frac{1}{1+C} m_\Gamma(A_{n,\alpha}) \leq \alpha \mathbf{E}(\|\mathcal{F}(\omega)\|)$ , where  $\mathbf{E}$  is the expectation value and  $C$  is the constant for the tempered sequence

$$\alpha \frac{1}{1+C} \int_{\tilde{F}_n} I_{Z_{n,\alpha}}(\gamma x) dm_\Gamma(\gamma) = \alpha \frac{1}{1+C} m_\Gamma(A_{n,\alpha}) \leq \alpha \mathbf{E}(\|\mathcal{F}(\omega)\|). \text{ Therefore}$$

$$\mathbf{E} \left( \sum_{B \in \mathcal{F}(\omega)} \int_B \varphi(\gamma x) dm_\Gamma(\gamma) \right) = \mathbf{E} \left( \int_{F_n} F_\omega(\gamma) \varphi(\gamma x) dm_\Gamma(\gamma) \right) =$$

$$\int_{F_n} \mathbf{E} F_\omega(\gamma) \varphi(\gamma x) dm_\Gamma(\gamma) \leq 2 \int_{F_n} |\varphi(\gamma x)| dm_\Gamma(\gamma). \text{ So}$$

$$\int_{\tilde{F}_n} I_{Z_{n,\alpha}}(\gamma x) dm_\Gamma(\gamma) \leq 2 \frac{1}{(1+C)\alpha} \int_{F_n} |\varphi(\gamma x)| dm_\Gamma(\gamma), \text{ and the inequality (9) holds}$$

with  $M = \frac{2}{1+C}$ .

Then integrating on the whole  $X$ , with respect to the measure  $\nu$

$$\nu(Z_{n,\alpha}) = \frac{1}{m_\Gamma(\tilde{F}_n)} \int_X \left( \int_{\tilde{F}_n} I_{Z_{n,\alpha}}(\gamma x) dm_\Gamma(\gamma) \right) d\nu(x) \leq$$

$$\frac{M}{\alpha} \int_X \left( \int_{F_n} |\varphi(\gamma x)| dm_\Gamma(\gamma) \right) d\nu(x) = \frac{M}{\alpha} \frac{m_\Gamma(G_n)}{m_\Gamma(\tilde{F}_n)} \int_X |\varphi(x)| d\nu(x) \leq \frac{M}{\alpha} (1 + \varepsilon) \|\varphi\|_1.$$

Since  $\varepsilon$  is arbitrary the result follows. ■

**Definition:** Let  $\nu$  be a probability measure on  $X$ , the space  $L^2(X, \nu)$  is formed by the maps  $\varphi : X \rightarrow \mathbf{R}$ , such that

$$\int_X |\varphi(x) - t|^2 d\nu(x) < \infty, \text{ for some } t \in \mathbf{R}.$$

As a metric in  $L^2(X, \nu)$  can be taken

$$d_2(\varphi, \psi) = \sqrt{\int_X |\varphi(x) - \psi(x)|^2 d\nu(x)}$$

**Definition:** A measure  $m$  on  $\mathbf{R}$  has *second finite moment* if

$$\int_{\mathbf{R}} (t - u)^2 dm(u) < \infty, \text{ for any } t \in \mathbf{R}.$$

Let us denote by  $\mathcal{M}_2(\mathbf{R})$  to the set of measures in  $\mathcal{M}(\mathbf{R})$  with second finite moment.

Let us consider, for any Følner sequence  $(F_n)$ , the empirical measures  $\mathcal{E}_{F_n, \varphi}(x) \in \mathcal{M}(\mathbf{R})$ , given by

$$\mathcal{E}_{F_n, \varphi}(x)(E) = \frac{1}{m_\Gamma(F_n)} m_\Gamma(\{\gamma : \varphi(\gamma x) \in E\}). \tag{10}$$

In the classic case  $\Gamma = \mathbf{Z}$ ,  $F_n = \{1, 2, \dots, n\}$  and  $f : X \rightarrow X$  obtains

$$\mathcal{E}_{n, \varphi}(x) = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{\varphi(f^i(x))}.$$

**Definition:** Let  $m \in \mathcal{M}_2(\mathbf{R})$ , the *barycenter* of  $m$  is the real number  $\bar{t}$  which minimizes the map  $F(t) = \int_{\mathbf{R}} (t - u)^2 dm(u)$ .

For  $m \in \mathcal{M}_2(\mathbf{R})$  such a minimum does exist[?]

Thus we have a map  $bar : \mathcal{M}_2(\mathbf{R}) \rightarrow \mathbf{R}$ , and holds

$$bar(\mathcal{E}_{F_n, \varphi}(x)) = S_{F_n, \varphi}(x).$$

By the Lindenstrauss ergodic convergence theorem for groups the sequence  $\{bar(\mathcal{E}_{F_n, \varphi}(x))\} = \{S_{F_n, \varphi}(x)\}$  converges *a.e.* to a  $\Gamma$ -invariant map  $\bar{\varphi}(x)$ .

A *coupling* of two measures  $m_1, m_2 \in \mathcal{M}_2(\mathbf{R})$  is a measure  $\lambda \in \mathcal{M}_2(\mathbf{R}^2)$  which on the first and second factor projects on  $m_1, m_2$  respectively. The 2 – *Wasserstein metric* is defined as

$$W_2(m_1, m_2) = \inf_{\substack{\lambda \text{ coupling of} \\ \mu_1, \mu_2 \in \mathcal{M}(\mathbf{R}^2)}} \sqrt{\int_{\mathbf{R}^2} (t - u)^2 d\lambda(t, u)},$$

**Lemma 1:** Is valid

$$|S_{F_n, \varphi}(x) - S_{F_n, \psi}(x)| \leq W_2(\mathcal{E}_{F_n, \varphi}(x), \mathcal{E}_{F_n, \psi}(x)).$$

*Proof:* We have

$(t - S_{F_n, \psi}(x))^2 \leq \int_{\mathbf{R}} (t - u)^2 d\mathcal{E}_{F_n, \psi}(x)(u)$ , integrating with respect to  $d\mathcal{E}_{F_n, \varphi}(x)$  obtains

$$\int_{\mathbf{R}} (t - S_{F_n, \psi}(x))^2 d\mathcal{E}_{F_n, \varphi}(x)(t) \leq \int_{\mathbf{R}^2} (t - u)^2 d\mathcal{E}_{F_n, \varphi}(x)(t) d\mathcal{E}_{F_n, \psi}(x)(u). \text{ Besides}$$

$$(S_{F_n, \varphi}(x) - u)^2 \leq \int_{\mathbf{R}} (t - u)^2 d\mathcal{E}_{F_n, \varphi}(x)(t). \text{ Putting } u = S_{F_n, \psi}(x) \text{ we get}$$

$$(S_{F_n, \varphi}(x) - S_{F_n, \psi}(x))^2 \leq \int_{\mathbf{R}} (t - S_{F_n, \psi}(x))^2 d\mathcal{E}_{F_n, \varphi}(x)(t) \leq \int_{\mathbf{R}^2} (t - u)^2 d\mathcal{E}_{F_n, \varphi}(x)(t) d\mathcal{E}_{F_n, \psi}(x)(u).$$

Let  $\lambda_n(x) = \mathcal{E}_{F_n, \varphi}(x) \times \mathcal{E}_{F_n, \psi}(x)$ , thus

$$(S_{F_n, \varphi}(x) - S_{F_n, \psi}(x))^2 \leq \inf \int_{\mathbf{R}^2} (t - u)^2 d\mathcal{E}_{F_n, \varphi}(x)(t) d\mathcal{E}_{F_n, \psi}(x)(u) = W_2(\mathcal{E}_{F_n, \varphi}(x), \mathcal{E}_{F_n, \psi}(x)).$$

■

In a similar way can be proved the more general result

$$|S_{F_n, \varphi}(x) - bar(m)| \leq W_2(\mathcal{E}_{F_n, \varphi}(x), m), \text{ for any } m \in \mathcal{M}_2(\mathbf{R}).$$

The objective is to study the fluctuations set  $\{x : |S_{F_n, \varphi}(x) - \bar{\varphi}(x)| > \alpha\}$ . We firstly consider the case of when the map  $\varphi$  takes only finite values  $\{t_1, t_2, \dots, t_k\}$ . In this situation the sequence  $\{bar(\mathcal{E}_{F_n, \varphi}(x))\} = \{S_{F_n, \varphi}(x)\}$  converges to a map  $\bar{\varphi}(x) = bar(m(x))$ , with

$$m(x) = \sum_{i=1}^k \nu_x(\varphi^{-1}(t_i)) \delta_{t_i},$$

with  $\{\nu_x\}$  the ergodic decomposition of  $\nu$  with respect to the action. In a more general case any map can be approximated by functions taking finite values, so that the fluctuations for general maps may be estimated by the bounds obtained for the case of finite values.

**Theorem 2:** Let  $\varphi : X \rightarrow \mathbf{R}$  taking values  $\{t_1, t_2, \dots, t_k\}$ , if  $\nu \in \mathcal{M}(X)$  then

$$\nu(\{x : |S_{F_n, \varphi}(x) - \text{bar}(m(x))| > \alpha\}) \leq \frac{M}{\alpha^2} \|\Psi\|_1,$$

where  $\Psi(x) = \sum_{i=1}^k \nu_x(\varphi^{-1}(t_i)) (\varphi(x) - t_i)^2$ , and  $M = M(F_n)$  is the constant of the proposition 2.

*Proof:* It is valid, by lemma 1 and proposition 2, that

$$\{x : |S_{F_n, \varphi}(x) - \text{bar}(m(x))| > \alpha\} \subset \{x : W_2(\mathcal{E}_{F_n, \varphi}(x), m(x)) > \alpha\},$$

and

$$\nu(\{x : S_{F_n, \phi}(x) > \alpha\}) \leq \frac{M}{\alpha^2} \|\phi\|_1, \text{ for any function } \phi.$$

Thus

$$\begin{aligned} W_2^2(\mathcal{E}_{F_n, \varphi}(x), m(x)) &\leq \int_{\mathbf{R}^2} (t - u)^2 d\mathcal{E}_{F_n, \varphi}(x)(t) dm(x)(u) = \\ &\int_{\mathbf{R}} \left[ \frac{1}{m_\Gamma(F_n)} \int_{F_n} (\varphi(\gamma x) - u)^2 dm_\Gamma(\gamma) \right] dm(x)(u). \end{aligned} \tag{11}$$

Let  $\phi_u(x) = (\varphi(x) - u)^2$ , we have that if  $T : \mathbf{R} \rightarrow \mathbf{R}$  then  $\int T dm(x) = \int T d\left(\sum_{i=1}^k \nu_x(\varphi^{-1}(t_i)) \delta_{t_i}\right)$ , therefore

$$\begin{aligned} \int_{\mathbf{R}} \left[ \frac{1}{m_\Gamma(F_n)} \int_{F_n} (\varphi(\gamma x) - u)^2 dm_\Gamma(\gamma) \right] dm(x)(u) &= \\ &= \frac{1}{m_\Gamma(F_n)} \int_{F_n} \sum_{i=1}^k \nu_x(\varphi^{-1}(t_i)) \phi_{t_i}(\gamma x)^2 dm_\Gamma(\gamma). \end{aligned}$$

Hence

$$\nu(\{x : |S_{F_n, \varphi}(x) - \text{bar}(m(x))| > \alpha^2\}) \leq \nu(\{x : S_{F_n, \Psi}(x) > \alpha^2\}) \leq \frac{M}{\alpha^2} \|\Psi\|_1, \text{ with}$$

$$\Psi(x) = \sum_{i=1}^k \nu_x(\varphi^{-1}(t_i)) (\varphi(x) - t_i)^2.$$

■

For general maps  $\varphi : X \rightarrow \mathbf{R}$ , holds that for a given  $\delta > 0$ , there is function  $\psi$  taking finite values such that  $d_2(\varphi, \psi) < \delta^2$ , in fact this is valid for functions valued in general separable spaces. Thus

$$\{x : |S_{F_n, \varphi}(x) - S_{F_n, \psi}(x)| > \delta\} \subset \{x : W_2(\mathcal{E}_{F_n, \varphi}(x), \mathcal{E}_{F_n, \psi}(x)) > \delta\}. \text{ We have}$$

$$W_2^2(\mathcal{E}_{F_n, \varphi}(x), \mathcal{E}_{F_n, \psi}(x)) \leq \int_{\mathbf{R}^2} (t - u)^2 d\mathcal{E}_{F_n, \varphi}(x)(t) d\mathcal{E}_{F_n, \psi}(x)(u) = \frac{1}{m_\Gamma(F_n)} \int_{F_n} (\varphi(\gamma x) - \psi(\gamma x))^2 dm_\Gamma(\gamma). \text{ So}$$

$$\nu(\{x : |S_{F_n, \varphi}(x) - S_{F_n, \psi}(x)| > \delta\}) \leq \frac{M}{\delta^2} \|\phi\|_1, \text{ with } \phi(x) = (\varphi(x) - \psi(x))^2. \text{ Since } \|\phi\|_1 = d_2^2(\varphi, \psi), \text{ is obtained that}$$

$\nu(\{x : |S_{F_n, \varphi}(x) - S_{F_n, \psi}(x)| > \delta\}) \leq M\alpha^2$ , for arbitrary  $\delta$ . Therefore the sequences of the ergodic averages for  $\varphi$  and  $\psi$  asymptotically oscillate by at most  $\delta$ , and where  $S_{F_n, \psi}(x)$  converges *a.e.* . This means that  $S_{F_n, \varphi}(x)$  converges outside a set of negligible measure to an invariant function  $\bar{\varphi}(x)$

Now the fluctuations for  $S_{F_n, \varphi}(x)$  may be estimated by the bounds obtained for a measure which takes finite values and which approximates  $\varphi$ .

## References

- [1] T. Austin, A  $CAT(0)$ -valued pointwise ergodic theorem, *J. of Topology and Analysis*, **3**, no. 2, 145–152 (2011).
- [2] E. Cartan, *Leçons sur la géométrie des espaces de Riemann*, Guathier-Villars, (1951) 2d. ed.
- [3] J.R. Chazottes and C. Maldonado, Concentration bounds for entropy estimation of one-dimensional Gibbs measures, *Nonlinearity*, **24**, 2371-2381 (2011).
- [4] V. Climenhaga, Bowen equation in the non-uniform setting, *Ergod. Th. and Dynam. Sys.* **31**, 1163-1182 (2011).

- [5] P. Collet, S. Martinez and B. Schmitt, Exponential inequalities for dynamical measures of expanding maps of the interval., *Probab. Theory Relat. Filelds* **123**, 301-322 (2002)
- [6] E. Lindenstrauss, Pointwise theorems for amenable groups, *Inven. Math.*, **146**, 259-295 (2001).
- [7] C. E. Pfister and W.G. Sullivan, Onthe topological entropy of saturated sets, *Ergod. Th. and Dynam. Sys.* **27**, 1-29 (2007).
- [8] P.C. Shields, *Ergodic Theory of Discrete Sample Paths (Graduate texts in Mathematics vol. 13)* , Providence R.L. AMS.

**Received: October, 2012**