Topological **M**ethods in **N**onlinear **A**nalysis
Journal of the Juliusz Schauder University Centre
Volume 40, 2012, 415–427

AN APPLICATION OF THE ERGODIC THEOREM OF INFORMATION THEORY TO LYAPUNOV EXPONENTS OF CELLULAR AUTOMATA

Wojciech Bułatek — Maurice Courbage
Brunon Kamiński — Jerzy Szymański

Abstract. We prove a generalization of the individual ergodic theorem of the information theory and we apply it to give a new proof of the Shereshevsky inequality connecting the metric entropy and Lyapunov exponents of dynamical systems generated by cellular automata.

1. Introduction

The Kolomogorov–Sinai entropy plays an important role in the investigation of dynamical systems. In the theory of smooth dynamical systems there is another important quantity describing the dynamics, namely the Lyapunov exponents. These two quantities are connected by the Ruelle inequality and the Pesin equality (see for example [7]).

Among topological dynamical systems, cellular automata play a special role from the point of view of applications. They are generated by the iteration of a continuous endomorphism on the symbolic shift systems. There have appeared many papers concerning these systems during the last decades (see for example interesting reviews by Blanchard, Kurka, Maass [1], and Kurka [6]). In the eighties, Wolfram ([11]) studied the one-dimensional cellular automata

2010 *Mathematics Subject Classification.* Primary: 37A35, 37B15; Secondary: 28D20.
Key words and phrases. Cellular automata, dynamical systems, entropy, local entropy, Lyapunov exponents.

as models of pattern formation in spatially infinite lattice and introduced the concept of propagation of some local patterns that he called local perturbations. He introduced certain quantities in order to characterize the speed of propagation of those patterns under the dynamics and called them Lyapunov exponents.

Although the Kolmogorov–Sinai entropy is well defined in the case of cellular automata systems, the classical Lyapunov exponents cannot be considered because the phase space does not possess a differentiable structure. However there have been constructed mathematical analogues of Lyapunov exponents following the ideas of Wolfram in [11]. The first such definition has been made by Shereshevsky ([8]). In [10] it have been introduced another quantities, the so-called average Lyapunov exponents, that differ from those of Shereshevsky.

The Lyapunov exponents of cellular automata have been investigated and computed in several papers (see for example [4], [5]) and they are presently a subject of an active research in several examples. The paper [3] contains so called directional Lyapunov exponents which are generalizations of the Shereshevsky concept to space-time semi-group actions.

We shall consider in the sequel the concept of the Lyapunov exponents defined by Shereshevsky. Applying the Brin–Katok local entropy, Shereshevsky presented in [8] an interesting inequality connecting the entropy with Lyapunov exponents. It can be considered as an analogue of the Ruelle inequality. The proof of Shereshevsky is too brief and contains some errors. Therefore we give here our proof. Although its main idea is the same as that of Shereshevsky, the new tools used by us are Theorem 2.2 and the proof of the crucial Shereshevsky inclusion (Lemma 2.3).

2. Definitions and auxiliary results

2.1. An extension of the Breiman–McMillan–Shannon theorem. Let $(X,\mathcal{B},\mu)$ be a probability space and let T be its automorphism. We denote by $\mathcal{J}_T \subset \mathcal{B}$ the σ-algebra of T-invariant sets.

Let α be a finite measurable partition of X. For a given $x \in X$ we denote by $\alpha(x)$ the atom of α which contains x. Let $\mathcal{A}$ be the algebra of sets generated by α and let $\mathcal{A}_T^-$ denote the past σ-algebra generated by $\mathcal{A}$ and T, i.e.

$$\mathcal{A}_T^- = \bigvee_{i=1}^{+\infty} T^{-i}\mathcal{A}.$$

Now we recall the well known concept of the information of an algebra. Namely, we define

$$I(\mathcal{A}) = -\sum_{A\in\alpha} \log\mu(A)\cdot\chi_A$$

where χ_A is the indicator function of A, and we call it the information of $\mathcal{A}$.

It is clear that

$$I(\mathcal{A})(x) = -\log \mu(\alpha(x)), \quad x \in X.$$

Let $\mathcal{C} \subset \mathcal{B}$ be a given σ-algebra. The function

$$I(\mathcal{A}|\mathcal{C}) = -\sum_{A\in\alpha} \log \mu(\mathcal{A}\,|\,\mathcal{C}) \cdot \chi_A$$

where $\mu(\,\cdot\,|\,\mathcal{C})$ denotes the conditional measure determined by μ and $\mathcal{C}$, is said to be the conditional information of $\mathcal{A}$ given $\mathcal{C}$. We shall use in the sequel the conditional information $I_{\mathcal{A}} = I(\mathcal{A}\,|\,\mathcal{A}_T^-)$ of $\mathcal{A}$ given $\mathcal{A}_T^-$. The function $I_{\mathcal{A}}$ is integrable and

$$\int_X I_{\mathcal{A}}\, d\mu = h_\mu(\mathcal{A}, T)$$

where $h_\mu(\mathcal{A}, T)$ denotes the mean entropy of $\mathcal{A}$ w.r. to T.

Lemma 2.1. *For any finite algebra $\mathcal{A} \subset \mathcal{B}$ holds*

$$\lim_{m,n\to+\infty} \frac{1}{m+n+1} I\left(\bigvee_{i=-m}^{n} T^{-i}\mathcal{A}\right) = E(I_{\mathcal{A}}\,|\,\mathcal{J}_T) \quad a.e.$$

where $E(I_{\mathcal{A}}\,|\,\mathcal{J}_T)$ is the conditional expectation of $I_{\mathcal{A}}$ w.r. to $\mathcal{J}_T$.

Proof. Let $f, f_i\colon X \to \mathbb{R}$, $i \in \mathbb{N}$, be defined by

$$f = I_{\mathcal{A}}, \quad f_0 = I(\mathcal{A}), \quad f_k = I\left(\mathcal{A}\,\middle|\, \bigvee_{i=1}^{k} T^{-i}\mathcal{A}\right),$$

$k \geq 1$ and let $f^\star = E(f\,|\,\mathcal{J}_T)$.

Applying the chain rule we have

$$I\left(\bigvee_{i=-m}^{n} T^{-i}\mathcal{A}\right) = \sum_{i=-m}^{n} f_{n-i} \circ T^i, \quad n, m \geq 1.$$

It follows from the Birkhoff individual ergodic theorem applied to T and T^{-1}, respectively, that

$$\frac{1}{n+1}\sum_{i=0}^{n} f \circ T^i \to f^\star, \quad \frac{1}{n+1}\sum_{i=0}^{n} f \circ T^{-i} \to f^\star \quad \text{a.e.}$$

as $n \to +\infty$. Hence

$$\frac{1}{m+n+1}\sum_{i=-m}^{n} f \circ T^i \to f^\star \quad \text{a.e.} \tag{2.1}$$

as $m, n \to +\infty$.

We have

$$\left|\frac{1}{m+n+1} I\left(\bigvee_{i=-m}^{n} T^{-i}\mathcal{A}\right) - f^\star\right| \leq \Delta_{m,n}^{(1)} + \Delta_{m,n}^{(2)} \tag{2.2}$$

where

$$\begin{aligned}\Delta^{(1)}_{m,n} &= \left|\frac{1}{m+n+1}\sum_{i=-m}^{n} f\circ T^i - f^\star\right|,\\ \Delta^{(2)}_{m,n} &= \frac{1}{m+n+1}\sum_{i=-m}^{n} |f_{n-i}-f|\circ T^i, \quad m,n\geq 1.\end{aligned}$$

By (2.1) we have

$$\lim_{m,n\to+\infty}\Delta^{(1)}_{m,n}=0 \quad \text{a.e.}$$

In order to show that $\Delta^{(2)}_{m,n}\to 0$ almost everywhere when $m,n\to+\infty$ we consider the sequences $(g_i)_{i\in\mathbb{N}}$ and $(G_N)_{N\in\mathbb{N}}$ of functions defined by

$$g_i=|f_i-f|,\quad G_N=\sup_{i\geq N} g_i,\quad i,N\geq 0.$$

Since $\mathcal{A}$ is finite, the function f is integrable and so, by Lemma 4.26 of [9], G_N is also integrable for any $N\geq 0$.

We have

$$\Delta^{(2)}_{m,n}=\frac{1}{m+n+1}\sum_{i=0}^{m+n} g_i\circ T^{n-i},\quad m,n\geq 1.$$

Using the Doob individual martingale convergence theorem we have

$$g_i\to 0 \quad \text{a.e. as } i\to+\infty.$$

Therefore it is easy to see that

(2.3) $$G_N\to 0 \quad \text{a.e. as } N\to+\infty.$$

For fixed $N\geq 0$ and m,n with $m+n>N$, we have

$$\begin{aligned}\Delta^{(2)}_{m,n} &\leq \frac{1}{m+n+1}\sum_{i=0}^{N} g_i\circ T^{n-i}+\frac{1}{m+n+1}\sum_{i=N+1}^{m+n} G_N\circ T^{n-i}\\ &\leq \frac{n}{m+n+1}\cdot\frac{1}{n}\left(\sum_{i=0}^{N} g_i\circ T^{-i}\right)\circ T^n+\frac{1}{m+n+1}\sum_{i=-m}^{n} G_N\circ T^{n-i}.\end{aligned}$$

Since the first summand tends to 0 almost everywhere as $m,n\to+\infty$, then using the ergodic theorem we get

$$\limsup_{m,n\to+\infty}\Delta^{(2)}_{m,n}\leq \lim_{m,n\to+\infty}\frac{1}{m+n+1}\sum_{i=-m}^{n} G_N\circ T^i = E(G_N\,|\,\mathcal{J}_T) \quad \text{a.e.}$$

Integrating the both sides of the above inequality we obtain

(2.4) $$\int_X \limsup_{m,n\to+\infty}\Delta^{(2)}_{m,n}\,d\mu\leq \int_X G_N\,d\mu,\quad N\geq 1.$$

Since $G_N \le G_0 \in L^1(X,\mu)$, $N \ge 0$, then by (2.3) and the Lebesgue dominated convergence theorem we have

$$\lim_{N\to+\infty} \int_X G_N \, d\mu = 0.$$

Thus taking in (2.4) the limit as $N \to +\infty$, we get

$$\lim_{m,n\to+\infty} \Delta^{(2)}_{m,n} = 0 \quad \text{a.e.}$$

what completes the proof by (2.2). □

Theorem 2.2. *For any finite algebra $\mathcal{A} \subset \mathcal{B}$, any strictly positive measurable functions $r, s\colon X \to \mathbb{R}$ and any $p \in \mathbb{N}$ we have*

$$\lim_{n\to+\infty} \frac{1}{n+1} I\Big(\bigvee_{i=-[nr(x)]-p}^{[ns(x)]+p} T^{-i}\mathcal{A}\Big)(x) = (r(x)+s(x))E(I_{\mathcal{A}} \mid \mathcal{J}_T)(x)$$

for almost all $x \in X$.

Proof. First we introduce the following notation:

$$\begin{aligned} r(n,x) &= [nr(x)] + p, & s(n,x) &= [ns(x)] + p, \\ t(n,x) &= r(n,x) + s(n,x) + 1, & t(x) &= r(x) + s(x), \quad n \ge 1,\ x \in X. \end{aligned}$$

Since $t(n,x) \le r(n,x) + s(n,x) + 2p + 1$, $n \ge 1$, we have, for $n > p$,

$$\begin{aligned} &\bigg|\frac{1}{n+1} I\Big(\bigvee_{i=-r(n,x)}^{s(n,x)} T^{-i}\mathcal{A}\Big)(x) - t(x)\cdot E(I_{\mathcal{A}} \mid \mathcal{J}_T)(x)\bigg| \\ &\qquad = \frac{t(n,x)}{n+1}\bigg|\frac{1}{t(n,x)} I\Big(\bigvee_{i=-r(n,x)}^{s(n,x)} T^{-i}\mathcal{A}\Big)(x) - E(I_{\mathcal{A}} \mid \mathcal{J}_T)(x)\bigg| \\ &\qquad\quad + \bigg|\frac{t(n,x)}{n+1} - t(x)\bigg| \cdot E(I_{\mathcal{A}} \mid \mathcal{J}_T)(x) \\ &\qquad \le (t(x)+1)\bigg|\frac{1}{t(n,x)} I\Big(\bigvee_{i=-r(n,x)}^{s(n,x)} T^{-i}\mathcal{A}\Big)(x) - E(I_{\mathcal{A}} \mid \mathcal{J}_T)(x)\bigg| \\ &\qquad\quad + \frac{t(x)+1}{n} \cdot E(I_{\mathcal{A}} \mid \mathcal{J}_T)(x). \end{aligned}$$

Let $X_0 \subset X$ be such that for $x \in X_0$ Lemma 2.1 holds. Hence $\mu(X_0) = 1$.

Let $x \in X_0$ and let $\varepsilon > 0$ be arbitrary. By Lemma 2.1 there exists $N_0(x)$ such that

$$\bigg|\frac{1}{k+l+1} I\Big(\bigvee_{i=-k}^{l} T^{-i}\mathcal{A}\Big)(x) - E(I_{\mathcal{A}} \mid \mathcal{J}_T)(x)\bigg| < \frac{\varepsilon}{2(t(x)+1)}$$

for $k, l > N_0(x)$.

Let
$$N_1(x) = \max\left(\frac{N_0(x)}{r(x)}, \frac{N_0(x)}{s(x)}\right)$$
and let $N_2(x)$ be such that
$$\frac{t(x)+1}{n}\, E(I_{\mathcal{A}} \mid \mathcal{J}_T)(x) < \frac{\varepsilon}{2}$$
for $n > N_2(x)$. Therefore taking $n > \max(N_1(x), N_2(x))$ we obtain
$$\left|\frac{1}{n+1} I\biggl(\bigvee_{i=-r(n,x)}^{s(n,x)} T^{-i}\mathcal{A}\biggr)(x) - t(x)E(I_{\mathcal{A}} \mid \mathcal{J}_T)(x)\right| < \varepsilon$$
which completes the proof. □

2.2. Cellular automata and Lyapunov exponents. Let $X = S^{\mathbb{Z}}$, $S = \{0, 1, \dots, M-1\}$, $M \geq 2$ and let $\mathcal{B}$ be the σ-algebra generated by cylindric sets. We equip X with the distance d defined as follows (cf. [8])
$$d(x,y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x_0 \neq y_0, \\ e^{-N(x,y)} & \text{if } x \neq y,\ x_0 = y_0 \end{cases}$$
where $N(x,y) = \sup\{n \geq 0;\ x_i = y_i,\ |i| \leq n\}$, $x, y \in X$. We denote by σ the shift transformation of X and by f the automaton transformation of X generated by an automaton local rule F, i.e.
$$(\sigma x)_i = x_{i+1}, \quad (fx)_i = F(x_{i-r}, \dots, x_{i+r}), \quad i \in \mathbb{Z},$$
$$F\colon S^{2r+1} \longrightarrow S, \quad r \in \mathbb{N}.$$
For any $p, q \in \mathbb{Z}$, $p \leq q$ and $x \in X$ we denote by $\widetilde{F}(x_{p-r}, \dots, x_{q+r})$ the concatenation
$$\widetilde{F}(x_{p-r}, \dots, x_{q+r}) = F(x_{p-r}, \dots, x_{p+r})F(x_{p+1-r}, \dots, x_{p+1+r}) \dots F(x_{q-r}, \dots, x_{q+r}).$$
It is obvious that
$$f(x)(p,q) \stackrel{\text{def}}{=} f(x)_p f(x)_{p+1} \dots f(x)_q = \widetilde{F}(x_{p-r}, \dots, x_{q+r}).$$

By an interval in $\mathbb{Z}$ we mean a set which consists of all integers which belong to an interval in $\mathbb{R}$.

Let $I \subset \mathbb{Z}$ be an interval and let $x = (x_i)$, $y = (y_i) \in X$. We shall write $x = y(I)$ if $x_i = y_i$, $i \in I$.

Let $x \in X$. Let $s, p, q \in \mathbb{Z}$ be such that $p \leq q$. Following Shereshevsky ([8]) we put

$$\begin{aligned} W_s^+(x) &= \{y \in X;\, y = x(s, +\infty)\}, \\ W_s^-(x) &= \{y \in X;\, y = x(-\infty, -s)\}, \\ C_p^q(x) &= \{y \in X;\, y = x(p, q)\}. \end{aligned}$$

For a given $n \geq 1$ one defines

$$\widetilde{\Lambda}_n^{\pm}(x) = \inf\{s \geq 0;\, f^n(W_0^{\pm}(x)) \subset W_s^{\pm}(f^n x)\}$$

and

$$\widetilde{l}_n^{\pm}(x) = \inf\{s \geq 0;\, f^i(W_0^{\pm}(x)) \subset W_s^{\pm}(f^i x) \text{ for all } 0 \leq i \leq n\}.$$

It is clear that

$$\widetilde{l}_n^{\pm}(x) = \max(\widetilde{\Lambda}_1^{\pm}(x), \dots, \widetilde{\Lambda}_n^{\pm}(x)).$$

We put

$$\Lambda_n^{\pm}(x) = \sup_{j \in \mathbb{Z}} \widetilde{\Lambda}_n^{\pm}(\sigma^j x), \qquad l_n^{\pm}(x) = \sup_{j \in \mathbb{Z}} \widetilde{l}_n^{\pm}(\sigma^j x).$$

It easy to see that

$$l_n^{\pm}(x) = \max(\Lambda_1^{\pm}(x), \dots, \Lambda_n^{\pm}(x)).$$

Let μ be a Borel probability measure on $\mathcal{B}$ invariant with respect to σ and f.

It is shown in [8] that the limits

$$\lambda^{\pm}(x) = \lim_{n \to \infty} \frac{1}{n} \Lambda_n^{\pm}(x)$$

exist μ-a.e. and they are f and σ-invariant and integrable.

The limit λ^+ (resp. λ^-) is called the right (left) Lyapunov exponent of f.

It is easy to show that

$$0 \leq \lambda^{\pm}(x) \leq r \quad \text{and} \quad \lambda^{\pm}(x) = \lim_{n \to \infty} \frac{l_n^{\pm}(x)}{n}.$$

Lemma 2.3. *For any natural numbers n, p, i such that $n \geq 0$, $0 \leq i \leq n$, $p \geq r$ and $x \in X$ we have*

$$f^i\Big(C_{-p-l_n^+(x)}^{p+l_n^-(x)}(x)\Big) \subset C_{-p}^{p}(f^i x).$$

Prrof. The arguments here have some similarity with the arguments of Tisseur given for average Lyapunov exponents ([10]). We shall use in the sequel the abbreviation $l_n^{\pm} = l_n^{\pm}(x)$.

Let $y \in C_{-p-l_n^+}^{p+l_n^-}(x)$. We have to show that

$$f^i(y) = f^i(x)(-p, p) \quad \text{for all } 0 \leq i \leq n. \tag{2.5}$$

It is clear that the sets $W^+_{-p-l_n^+}(x) \cap W^-_{-p-l_n^-}(y)$, $W^+_{-p-l_n^+}(y) \cap W^-_{-p-l_n^-}(x)$ consist of single elements. Let us denote them by z and w, respectively. Thus we have

$$z = x(-p-l_n^+, +\infty), \qquad z = y(-\infty, p+l_n^-), \tag{2.6}$$

$$w = x(-\infty, p+l_n^-), \qquad w = y(-p-l_n^+, +\infty). \tag{2.7}$$

Let $0 \le i \le n$. From (2.6) and (2.7) it follows that

$$f^i(z) = f^i(x)(-p, +\infty), \tag{2.8}$$

$$f^i(w) = f^i(x)(-\infty, p). \tag{2.9}$$

Indeed, applying the formula (cf. [3])

$$\sigma^a W^\pm_c(\sigma^b x) = W^\pm_{c\mp a}(\sigma^{a+b}x), \quad a, b, c \in \mathbb{Z}, \ x \in X,$$

we get

$$\begin{aligned} f^i(W^+_{-p-l_n^+}(x)) &= \sigma^{p+l_n^+} f^i(W_0^+(\sigma^{-p-l_n^+}x)) \\ &\subset \sigma^{p+l_n^+} W^+_{\widetilde{l}_n^+(\sigma^{-p-l_n^+}x)}(\sigma^{-p-l_n^+} f^i x) \\ &\subset \sigma^{p+l_n^+} W^+_{l_n^+}(\sigma^{-p-l_n^+} f^i x) = W^+_{-p}(f^i x), \quad 0 \le i \le n. \end{aligned}$$

This means that (2.8) is satisfied.

Similarly we show the inclusion

$$f^i\Big(W^-_{-p-l_n^-}(x)\Big) \subset W^-_{-p}(f^i x), \quad 0 \le i \le n$$

what gives (2.9).

Now we shall show that if

$$f^k(x) = f^k(y)(-p, p) \quad \text{for all } 0 \le k \le i \tag{2.10}$$

then

$$f^i(y) = f^i(z)(-p-r, r), \tag{2.11}$$

$$f^i(y) = f^i(w)(-r, p+r). \tag{2.12}$$

In fact we shall prove more, namely that for all $0 \le k \le i$

$$f^k(y) = f^k(z)(-p-r(i+1-k), r), \tag{2.13}$$

$$f^k(y) = f^k(w)(-r, p+r(i+1-k)). \tag{2.14}$$

We only prove (2.13) because the proof of (2.14) is similar.

We argue by induction w.r. to $k \in \{0, \dots, i\}$. The validity of (2.13) for $k = 0$ follows at once from the inequalities $p \ge r$, $l_n^-, l_n^+ \ge 0$.

Suppose now that

$$f^k(z) = f^k(y)(-p-r(i+1-k), r) \quad \text{for all } 0 \le k \le i-1. \tag{2.15}$$

We shall show that

$$f^{k+1}(z)=f^{k+1}(y)(-p-r(i-k),r). \tag{2.16}$$

We have

$$\begin{aligned} f^{k+1}(z)(-p-r(i-k),r) &= \widetilde{F}(f^k(z)(-p-r(i+1-k),2r)) \\ &= \widetilde{F}(f^k(z)(-p-r(i+1-k),r))\widetilde{F}(f^k(z)(-r+1,2r)). \end{aligned} \tag{2.17}$$

The assumption (2.15) tells that

$$f^k(z)=f^k(y)(-p-r(i+1-k),r).$$

We claim that

$$\widetilde{F}(f^k(z)(-r+1,2r))=\widetilde{F}(f^k(y)(-r+1,2r)).$$

Indeed, the equality (2.8) gives $f^i(z)=f^i(x)(-p,+\infty)$ for any $0\le i\le n$. Hence in particular $f^k(z)=f^k(x)(-p,+\infty)$ and so, since $p\ge r$, we get

$$f^k(z)=f^k(x)(-r+1,2r). \tag{2.18}$$

We have $k\le i-1$, i.e. $k+1\le i$ and therefore applying (2.10) we have

$$f^{k+1}(y)=f^{k+1}(x)(-p,p).$$

Hence by (2.18) and $p\ge r$ we get

$$\begin{aligned} \widetilde{F}(f^k(z)(-r+1,2r)) &= \widetilde{F}(f^k(x)(-r+1,2r))=f^{k+1}(x)(1,r) \\ &= f^{k+1}(y)(1,r)=\widetilde{F}(f^k(y)(-r+1,2r)). \end{aligned} \tag{2.19}$$

Therefore (2.17) implies

$$\begin{aligned} f^{k+1}(z)(-p-r(i-k),r) &= \widetilde{F}(f^k(y)(-p-r(i+1-k),r))\widetilde{F}(f^k(y)(-r+1,2r)) \\ &= f^{k+1}(y)(-p-r(i-k),0)f^{k+1}(y)(1,r) \\ &= f^{k+1}(y)(-p-r(i-k),r) \end{aligned}$$

which gives (2.16) and so (2.13).

Substituting $k=i$ in (2.13) and (2.14) we get (2.11) and (2.12).

Now we prove (2.5) by induction w.r. to $i\in\{0,\dots,n\}$.

The property (2.5) is obviously true for $i=0$ because $l_n^{\pm}\ge 0$.

Let us now suppose that for some $0\le i\le n-1$ the following statement is true

$$f^k(y)=f^k(x)(-p,p)\quad \text{for all } 0\le k\le i. \tag{2.20}$$

We shall show that

$$f^{i+1}(y)=f^{i+1}(x)(-p,p).$$

It follows from (2.11) and (2.12) that

$$f^i(y) = f^i(z)(-p-r,r), \tag{2.21}$$

$$f^i(y) = f^i(w)(-r,p+r). \tag{2.22}$$

Applying (2.8), (2.9), (2.21), (2.22) we have

$$\begin{aligned} f^{i+1}(y)(-p,p) &= \widetilde{F}(f^i(y)(-p-r,p+r)) \\ &= \widetilde{F}(f^i(y)(-p-r,r))\widetilde{F}(f^i(y)(-r+1,p+r)) \\ &= \widetilde{F}(f^i(z)(-p-r,r))\widetilde{F}(f^i(w)(-r+1,p+r)) \\ &= f^{i+1}(z)(-p,0)f^{i+1}(w)(1,p) \\ &= f^{i+1}(x)(-p,0)f^{i+1}(x)(1,p) = f^{i+1}(x)(-p,p) \end{aligned}$$

which gives the desired result. □

3. Entropy and Lyapunov exponents for cellular automata transformations

Theorem 3.1 ([8]). *For any Borel probability measure μ invariant w.r.t. to σ and f it holds*

$$h_\mu(f) \le \int_X (\lambda^+(x) + \lambda^-(x))h_\mu(\sigma,x)\mu(dx)$$

where $h_\mu(\sigma,x)$ is the local entropy of σ at $x \in X$ (cf. [2]).

Proof. Let $G \subset X$ be a measurable set with $\mu(G) = 1$ on which $\lambda^\pm$ are defined. We take $x \in G$ and we fix two parameters $p \in \mathbb{N}$, $p \ge r$ and $\delta > 0$. We consider the following sequences of positive integers

$$\lambda_n^\pm(\delta,x) = [(\lambda^\pm(x)+\delta)n]+1$$

and the sets

$$B_n(f,x,\varepsilon_p) = \{y \in X : d(f^k y, f^k x) < \varepsilon_p,\ 0 \le k \le n\}$$

where $\varepsilon_p = e^{-p}$, $n \in \mathbb{N}$.

We show that

$$B_n(f,x,\varepsilon_p) \supset \bigg(\bigvee_{m\in A_n(x)} \sigma^m \overline{\mathcal{P}}\bigg)(x) \tag{3.1}$$

where $\overline{\mathcal{P}}$ is the zero-time partition of X and

$$A_n(x) = \{m \in \mathbb{Z} : -p-\lambda_n^+(\delta,x) \le m \le p + \lambda_n^-(\delta,x)\}.$$

Since

$$\lim_{n\to\infty} \frac{l_n^\pm(x)}{n} = \lambda^\pm(x) \tag{3.2}$$

there exists $N = N_\delta$ such that

$$\frac{l_n^{\pm}(x)}{n} \le \lambda^{\pm}(x) + \delta \quad \text{for } n > N. \tag{3.3}$$

Now Lemma 2.3 implies

$$f^k\Big(C^{p+l_n^-(x)}_{-p-l_n^+(x)}(x)\Big) \subset C^p_{-p}(f^k x) \quad \text{for any } 0 \le k \le n. \tag{3.4}$$

It follows from (3.3) that $l_n^{\pm}(x) \le \lambda_n^{\pm}(\delta, x)$, $n > N$. It is clear that

$$C^{p+\lambda_n^-(\delta,x)}_{-p-\lambda_n^+(\delta,x)}(x) = \Big(\bigvee_{m\in A_n(x)} \sigma^m\overline{\mathcal{P}}\Big)(x), \quad n > N.$$

Therefore we have to show the inclusion

$$B_n(f, x, \varepsilon_p) \supset C^{p+\lambda_n^-(\delta,x)}_{-p-\lambda_n^+(\delta,x)}(x), \quad x \in G. \tag{3.5}$$

Let $y \in C^{p+\lambda_n^-(\delta,x)}_{-p-\lambda_n^+(\delta,x)}(x)$. Hence by (3.4) we get

$$f^k(y) \in f^k\Big(C^{p+\lambda_n^-(\delta,x)}_{-p-\lambda_n^+(\delta,x)}(x)\Big) \subset f^k\Big(C^{p+l_n^-(x)}_{-p-l_n^+(x)}(x)\Big) \subset C^p_{-p}(f^k x),$$

for $0 \le k \le n$, $n > N$. This means that $(f^k y)_m = (f^k x)_m$, $-p \le m \le p$ and so $N(f^k y, f^k x) \ge p$, i.e.

$$d(f^k y, f^k x) \le e^{-p}, \quad 0 \le k \le n.$$

In other words $y \in B_n(f, x, \varepsilon_p)$ which proves (3.5) and so (3.1).

It follows from Theorem 2.2 that for μ almost all x holds

$$\lim_{n\to+\infty} -\frac{1}{n}\log\Big(\mu\Big(\Big(\bigvee_{m\in A_n(x)} \sigma^m\overline{\mathcal{P}}\Big)(x)\Big)\Big) \tag{3.6}$$
$$= \lim_{n\to+\infty} \frac{1}{n} I\Big(\Big(\bigvee_{m\in A_n(x)} \sigma^m\mathcal{P}\Big)(x)\Big) = (\lambda^+(x) + \lambda^-(x) + 2\delta)E(I_{\mathcal{P}} \mid \mathcal{J}_\sigma)(x)$$

where $\mathcal{P}$ is the σ-algebra generated by $\overline{\mathcal{P}}$.

Now let us remark that

$$E(I_{\mathcal{P}}|\mathcal{J}_\sigma)(x) = h_\mu(\sigma, x), \quad x \in X.$$

Indeed, for any $n, p \ge 1$ we have

$$\begin{aligned} B_n(\sigma, x, \varepsilon_p) &= \{y \in X : d(\sigma^i y, \sigma^i x) < e^{-p},\ 0 \le i \le n\} \\ &= \{y \in X : N(\sigma^i y, \sigma^i x) > p,\ 0 \le i \le n\} \\ &= \{y \in X : y_j = x_j,\ -p \le j \le n+p\} = \Big(\bigvee_{i=-p}^{p+n} \sigma^i\,\overline{\mathcal{P}}\Big)(x). \end{aligned}$$

Therefore by the Breimann–McMillan–Shannon Theorem

$$\begin{aligned} h_\mu(\sigma,x) &= \lim_{p\to+\infty}\Big(\limsup_{n\to+\infty} -\frac{1}{n}\log\mu(B_n(\sigma,x,\varepsilon_p))\Big) \\ &= \lim_{n\to+\infty} -\frac{1}{n}\log\mu\bigg(\bigg(\bigvee_{i=-p}^{p+n}\sigma^i\,\overline{\mathcal{P}}\bigg)(x)\bigg) = E(I_{\mathcal{P}}\mid\mathcal{J}_\sigma)(x) \end{aligned}$$

for almost all $x\in X$. Using this in (3.6) we obtain

$$\lim_{n\to+\infty} -\frac{1}{n}\log\bigg(\mu\bigg(\bigg(\bigvee_{m\in A_n(x)}\sigma^m\,\overline{\mathcal{P}}\bigg)(x)\bigg)\bigg) = (\lambda^+(x)+\lambda^-(x)+2\delta)h_\mu(\sigma,x).$$

Therefore applying (3.1) we get

$$\begin{aligned} h_\mu(f,x) &= \lim_{p\to+\infty}\Big(\limsup_{n\to+\infty} -\frac{1}{n}\log\mu(B_n(f,x,\varepsilon_p))\Big) \\ &\le (\lambda^+(x)+\lambda^-(x)+2\delta)h_\mu(\sigma,x). \end{aligned}$$

Hence taking the limit as $\delta\to 0$ and integrating the both sides we get by the Brin–Katok theorem ([2])

$$h_\mu(f) \le \int_X (\lambda^+(x)+\lambda^-(x))h_\mu(\sigma,x)\mu(dx),$$

which completes the proof. □

References

[1] F. Blanchard, P. Kurka and A. Maass, *Topological and measure-theoretic properties of one-dimensional cellular automata*, Lattice Dynamics (Paris, 1995), Phys. D **103** (1997).

[2] M. Brin and A. Katok, *On local entropy*, Geometric Dynamics (Rio de Janeiro, 1981), Lecture Notes in Math., vol. 1007, Springer, Berlin, 1983, pp. 30–38.

[3] M. Courbage and B. Kamiński, *Space-time directional Lyapunov exponents for cellular automata*, J. Stat. Phys. **124** (2006), 1499–1509.

[4] M. D'amico, G. Manzini and L. Margara, *On computing the entropy of cellular automata*, Theoret. Comput. Sci. **290** (2003), 1629–1646.

[5] M. Finelli, G. Manzini and L. Margara, *Lyapunov exponents versus expansivity and sensitivity in cellular automata*, J. Complexity **14** (1998), 210–233.

[6] P. Kurka, *Topological dynamics of cellular automata*, Encyclopedia of Complexity and Systems Science, Springer, part **20**, 2009, pp. 9246–9268.

[7] R. Mañe, *Ergodic Theory and Differentiable Dynamics*, Springer–Verlag, Berlin, 1987.

[8] M.A. Shereshevsky, *Lyapunov exponents for one-dimensional automata*, J. Nonlinear Sci. **2** (1992), 1–8.

[9] M. Smorodinsky, *Ergodic Theory, Entropy*, Lecture Notes in Math., vol. 214, Springer–Verlag, Berlin–Heidelberg–New York, 1971.

[10] P. Tisseur, *Cellular automata and Lyapunov exponents*, Nonlinearity **13** (2000), 1547–1560.

[11] S. Wolfram, *Cellular Automata and Complexity*, Addison–Wesley Publishing Company, 1974.

Manuscript received October 19, 2011

Wojciech Bułatek, Brunon Kamiński and Jerzy Szymański
Faculty of Mathematics and Computer Science
Nicolaus Copernicus University
Chopina 12/18
87-100 Toruń, POLAND

E-mail address: bulatek@mat.uni.torun.pl
bkam@mat.uni.torun.pl
jerzy@mat.uni.torun.pl

Maurice Courbage
Laboratoire Matière et Systèmes Complexes (MSC)
Université Paris 7 – Diderot
Case 7056, Bâtiment Condorcet, porte 718A
10, rue Alice Domon et Léonie Duquet, FRANCE

E-mail address: maurice.courbage@univ-paris-diderot.fr