

Thermodynamic Formalism for a family of cellular automata and duality with the shift

Artur O. Lopes, Elismar R. Oliveira and Marcelo Sobottka

July 5, 2024

Abstract

We will consider a family of cellular automata $\Phi : \{1, 2, \dots, r\}^{\mathbb{N}} \circlearrowleft$ that are not of algebraic type. Our first goal is to determine conditions that result in the identification of probabilities that are at the same time σ -invariant and Φ -invariant, where σ is the full shift. Via the use of versions of the Ruelle operator $\mathcal{L}_{A,\sigma}$ and $\mathcal{L}_{B,\Phi}$ we will show that there is an abundant set of measures with this property; they will be equilibrium probabilities for different Lipschitz potentials A, B and for the corresponding dynamics σ and Φ . Via the use of a version of the involution kernel W for a (σ, Φ) -mixed skew product $\hat{\Phi} : \{1, 2, \dots, r\}^{\mathbb{Z}} \circlearrowleft$, given A one can determine B , in such way that the integral kernel e^W produce a duality between eigenprobabilities ρ_A for $(\mathcal{L}_{A,\sigma})^*$ and eigenfunctions ψ_B for $\mathcal{L}_{B,\Phi}$. In another direction, considering the non-mixed extension $\hat{\Phi}_n : \{1, 2, \dots, r\}^{\mathbb{Z}} \circlearrowleft$ of Φ , given a Lipschitz potential $\hat{A} : \{1, 2, \dots, r\}^{\mathbb{Z}} \rightarrow \mathbb{R}$, we can identify a Lipschitz potential $A : \{1, 2, \dots, r\}^{\mathbb{N}} \rightarrow \mathbb{R}$, in such way that relates the variational problem of $\hat{\Phi}_n$ -Topological Pressure for \hat{A} with the Φ -Topological Pressure for A . We also present a version of Livsic's Theorem. Whether or not Φ (or $\hat{\Phi}$) can eventually be conjugated with another shift of finite type is irrelevant in our context.

1 Introduction

Denote the alphabet by $\mathcal{A} = \{1, 2, \dots, r\}$, $\Omega = \mathcal{A}^{\mathbb{N}}$, and take two Lipschitz potentials $B_1, B_2 : \Omega \rightarrow \mathbb{R}$.

Let's first consider a more general setting that involves two commuting transformations $\Phi_1 : \Omega \rightarrow \Omega$ and $\Phi_2 : \Omega \rightarrow \Omega$, and where some of the results we will describe can be applied. A particular case of interest is when Φ_1 is

the unilateral shift $\sigma : \Omega \rightarrow \Omega$ and $\Phi_2 : \Omega \rightarrow \Omega$ is a cellular automata to be defined later. As usual

$$\sigma(x_1, x_2, \dots, x_n, \dots) = (x_2, x_3, \dots, x_n, \dots).$$

The bilateral shift is denoted by $\hat{\sigma} : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$.

Consider two r to 1 surjective expanding transformation Φ_1, Φ_2 , which commute, and for $j = 1, 2$, the two corresponding topological pressure problems

$$P_{\Phi_j}(B_j) := \sup_{\rho \in \mathcal{M}(\Phi_j)} \{h_{\Phi_j}(\rho) + \int B_j(x) d\rho(x)\}, \quad (1.1)$$

where, $h_{\Phi_j}(\rho)$ is the Shannon-Kolmogorov entropy of ρ for the dynamics of Φ_j , and $\mathcal{M}(\Phi_j)$ and $\mathcal{M}(\Phi_1, \Phi_2)$ are the set of probabilities that are invariant, respectively, for Φ_j and both Φ_1, Φ_2 . From the above problems we get, respectively, the solutions μ_{B_j, Φ_j} , $j = 1, 2$, which are called the equilibrium probabilities for respectively B_j, Φ_j , $j = 1, 2$. The value $P_{\Phi_j}(B_j)$ is called the topological pressure for the potential B_j and the dynamics of Φ_j , $j = 1, 2$.

We will show here that $\mathcal{M}(\Phi_1, \Phi_2)$ is not empty, which will imply that $\mathcal{M}(\sigma, \Phi)$ is also not empty (and has an abundance of elements as described for instance in Example 2.5).

The Ruelle operator is defined as $\mathcal{L}_j(f) = \mathcal{L}_{B_j, \Phi_j}(f) = g$, via

$$g(y) = \mathcal{L}_j(f)(y) = \sum_{\Phi_j(x)=y} e^{B_j(x)} f(x). \quad (1.2)$$

The Ruelle Theorem is valid for both $\mathcal{L}_1, \mathcal{L}_2$, and we get, respectively, the eigenvalues λ_{B_j, Φ_j} , eigenfunctions ψ_{B_j, Φ_j} , and eigenprobabilities ρ_{B_j, Φ_j} for \mathcal{L}_j^* , $j = 1, 2$. It is well known that $\mu_{B_j, \Phi_j} = \psi_{B_j, \Phi_j} \rho_{B_j, \Phi_j}$, when normalized. A particular case of interest is when $\Phi_1 = \sigma$ and $\Phi_2 = \Phi$ is the CA to be defined in the introduction.

A general reference for classical Thermodynamic Formalism is [31] and for IFS Thermodynamic Formalism see [11], [45], or [6].

It is well known that the eigenfunction ψ_{B_j, Φ_j} , and eigenprobability ρ_{B_j, Φ_j} are dual objects; as the duality of functions and measures, but by this we mean a relation not in the sense of Riesz Theorem. They are obtained respectively via $\mathcal{L}_{B_j, \Phi_j}$ and $(\mathcal{L}_{B_j, \Phi_j})^*$. We are interested here, among other things, in conditions on B_1 and B_2 that relate ψ_{B_2, Φ_1} and ρ_{B_1, Φ_2} .

In another direction, we ask: is there B_1, B_2 such that $\mu_{B_1, \Phi_1} = \mu_{B_2, \Phi_2}$?

Following [20], the answer to this last question regards considering a certain special relation between B_1 and B_2 .

Below, in order to be in accordance with the notation in [20], we will take $B_1 = A_2$ and $B_2 = A_1$.

The authors in [20] introduced the condition

$$1) A_1 - A_1 \circ \Phi_1 = A_2 - A_2 \circ \Phi_2, \quad (1.3)$$

or a more general one, where it is assumed that there exists a Lipschitz function w such

$$2)(A_1 - A_1 \circ \Phi_1) - (A_2 - A_2 \circ \Phi_2) = (w - w \circ \Phi_1) - (w - w \circ \Phi_1) \circ \Phi_2, \quad (1.4)$$

which is equivalent to

$$(A_1 - A_1 \circ \Phi_1) = [A_2 + (w - w \circ \Phi_1)] - [A_2 + (w - w \circ \Phi_1)] \circ \Phi_2.$$

In [20] it was shown that if A_1, A_2 satisfies condition 2), then $\mu_{A_2, \Phi_1} = \mu_{A_1, \Phi_2}$ (see our Theorem 3.7 for a more detailed proof).

However, [20] does not provide an example of potentials satisfying such conditions 1) or 2). An interesting question is to find an example where $\Phi_1 = \sigma$ and $\Phi_2 = \Phi$, and this will be provided later in Example 2.5.

In Section 2 we introduce the local rule that will define the class of cellular automata Φ (and we present examples) that will be the focus of Sections 3 and 4.

In Theorem 3.7 in Section 3 we exhibit conditions on potentials A_1 and A_2 such that the respective equilibrium probabilities for σ and Φ , are the same. Theorem 3.9 is a kind of reciprocal of Theorem 3.7.

In Section 4 via the use of a version of the involution kernel W for a (σ, Φ) -mixed skew product $\hat{\Phi} : \{1, 2, \dots, r\}^{\mathbb{Z}} \circlearrowleft$, we show that given a Lipschitz potential A one can determine another Lipschitz potential B , in such way that the integral kernel $e^W : \{1, 2, \dots, r\}^{\mathbb{Z}} \rightarrow \mathbb{R}$ produce a duality between eigenprobabilities ρ_A for $(\mathcal{L}_{A, \sigma})^*$ and eigenfunctions ψ_B for $\mathcal{L}_{B, \Phi}$. In another direction, considering the non-mixed extension $\hat{\Phi}_n : \{1, 2, \dots, r\}^{\mathbb{Z}} \circlearrowleft$ of Φ , given a Lipschitz potential $\hat{A} : \{1, 2, \dots, r\}^{\mathbb{Z}} \rightarrow \mathbb{R}$, we can identify a Lipschitz potential $A : \{1, 2, \dots, r\}^{\mathbb{N}} \rightarrow \mathbb{R}$, in such way that relates the variational problem of $\hat{\Phi}_n$ -Topological Pressure for \hat{A} with the Φ -Topological Pressure for A . Properties for the $\hat{\Phi}_n$ -equilibrium probability for $\hat{A} : \{1, 2, \dots, r\}^{\mathbb{Z}} \rightarrow \mathbb{R}$ can be derived in this way.

In Section 5 we present a version of Livsic's Theorem for our setting.

In Section 6, with the aim of putting our results in context, we present a review of previous results that are at the interface of cellular automata and ergodic theory. In particular questions that are in one way or another related to Furstenberg's conjecture.

Finally, in the Appendix we present the proof of some technical results mentioned before on the text.

2 A family of Cellular Automata

Remember that the alphabet is $\mathcal{A} = \{1, 2, \dots, r\}$ and $\Omega = \mathcal{A}^{\mathbb{N}}$.

In Ω we consider the metric d such that for

$$x = (x_1, x_2, \dots), y = (y_1, y_2, \dots) :$$

a) If $x_1 \neq y_1$, then, $d(x, y) = 2^{-1} = 1/2$, and otherwise,

$$d(x, y) = \begin{cases} 0, & x = y \\ 2^{-\{\min_j x_j \neq y_j\}}, & x \neq y \end{cases}$$

$x = (x_1, x_2, \dots), y = (y_1, y_2, \dots)$.

Note that, if $x_1 = y_1$, then $d(x, y) \leq 1/4$.

We define a particular local rule $\phi : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$, where we assume that for any fixed a , the law

$$b \in \mathcal{A} \rightarrow \phi(a, b) \text{ is bijective.} \quad (2.1)$$

Then, for each a we get that $b \rightarrow \phi(a, b)$ is a permutation on d symbols.

Since \mathcal{A} is finite, the local rule is completely described by a matrix $M := (\phi(i, j))_{r \times r}$ such that $\phi(a, b) = M_{a,b}$, thus each row is a permutation of \mathcal{A} . For example, for $\mathcal{A} = \{1, 2, 3\}$ we may choose,

$$M := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

So, in this case $\phi(3, 2) = M_{3,2} = 1$.

We will consider the cellular automata $\Phi : \Omega \rightarrow \Omega$ given by

$$\Phi(x_1, x_2, x_3, x_4, \dots, x_n, \dots) = (\phi(x_1, x_2), \phi(x_2, x_3), \phi(x_3, x_4), \dots). \quad (2.2)$$

We avoid the case of the trivial cellular automata: for each a , $\phi(a, b) = b$, in order that Φ is not σ . Here we call **permutative** the CA Φ obtained from the ϕ of (2.1). Our examples do not necessarily fit the ones in the family of algebraic cellular automata.

Note that $\sigma \circ \Phi = \Phi \circ \sigma$. The Φ defined by (2.1) is in some sense the simpler of all possible CA.

We point out that our main goal is to determine properties that result in the identification of probabilities that are at the same time σ -invariant and Φ -invariant. If A_2 and A_1 are such that

$$(A_1 - A_1 \circ \sigma) - (A_2 - A_2 \circ \Phi) = (w - w \circ \sigma) - (w - w \circ \sigma) \circ \Phi, \quad (2.3)$$

for some Lipschitz function w , then, in Theorem 3.7 we will show that the classical σ -equilibrium probability for the pressure of A_2 (see [31]) will be also Φ -invariant.

Therefore, whether or not Φ can eventually be conjugated with another shift is irrelevant in our context. We will work with information and properties obtained directly from A_1, A_2, σ and Φ (see for instance expressions (2.8) and (7.1)).

The next result will show that Φ is such that for each $x \in \Omega$, the set of preimages $\Phi^{-n}(x)$, $n \in \mathbb{N}$ is dense in Ω .

Proposition 2.1. *Let $\phi : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ a local rule (we assume in this section that*

$$j \rightarrow \phi(a, j) \text{ is a bijection for any } a \in \mathcal{A}. \quad (2.4)$$

Consider the map $\Phi : \Omega \rightarrow \Omega$, where $\Omega = \{1, \dots, r\}^{\mathbb{N}}$, associated to this rule. For any $x \in \Omega$ the set of preimages by Φ is dense.

For proof see Proposition 7.1 in the Appendix.

We denote by $\mathcal{M}(\sigma), \mathcal{M}(\Phi)$ and $\mathcal{M}(\sigma, \Phi)$ the set of probabilities that are invariant, respectively, for σ, Φ and simultaneously for σ , and Φ .

The function Φ is continuous, expanding at rate 2 (see Lemma 5.1), and we will show that the set $\mathcal{M}(\Phi, \sigma)$ is not empty and with cardinality bigger than one in several cases (see Example 2.5).

Given a , we denote $u_a(b) = c$, the element c such that $\phi(a, c) = b$.

Given $y = (b_1, b_2, \dots, b_n, \dots)$, **there exist r points** $x = (a_1, a_2, \dots, a_n, \dots)$, such that $\Phi(x) = y$. Indeed, given $a_1 \in \mathcal{A}$, take

$$x = (a_1, u_{a_1}(b_1), u_{u_{a_1}(b_1)}(b_2), \dots). \quad (2.5)$$

Given $j \in \mathcal{A}$, we denote by $\tau_j : \Omega \rightarrow \Omega$ the function such that given $x = (x_1, x_2, \dots, x_n, \dots)$

$$\tau_j(x) := (j, u_j(x_1), u_{u_j(x_1)}(x_2), \dots).$$

For each j we get that τ_j has a Lipschitz constant equal to $1/2$. Indeed, note that if $w_1 = (x_1, x_2, \dots, x_n, x_{n+1}, x_{n+2}, \dots)$ and $w_2 = (x_1, x_2, \dots, x_n, y_{n+1}, y_{n+2}, \dots)$, where $x_{n+1} \neq y_{n+1}$ (which means $d(w_1, w_2) = 2^{-n}$), then, as $\phi(x_n, x_{n+1}) \neq \phi(x_n, y_{n+1})$ (by bijectivity on the second variable)

$$\begin{aligned} d(\Phi(w_1), \Phi(w_2)) &= \\ d((\phi(x_1, x_2), \dots, \phi(x_n, x_{n+1}), \dots), (\phi(x_1, x_2), \dots, \phi(x_n, y_{n+1}), \dots)) &= \\ 2^{-(n-1)} = 2 \cdot 2^{-n} = 2 d(w_1, w_2). \end{aligned}$$

Remark 2.2. The map $\Phi : \Omega \rightarrow \Omega$ of (2.4), where $\Omega = \{1, \dots, r\}^{\mathbb{N}}$, has a transitive orbit (see Theorem 5.9). In Theorem 5.10 in Subsection 5.1 we show that each periodic point x for Φ with period m is a solution of $\tau_{j_1} \circ \dots \circ \tau_{j_m}(x) = x$, for some choice of $j_1, \dots, j_m \in \{1, \dots, r\}$. For each m there exists r^m points of period m . In Subsection 5.1 we exhibit the periodic points of periods two and three in a particular example.

Example 2.3. We will present examples of periodic orbits for $\Phi : \{1, 2\}^{\mathbb{N}} \rightarrow \{1, 2\}^{\mathbb{N}}$ in Subsection 5.1. For example, when

$$M = \begin{pmatrix} \phi(1, 1) & \phi(1, 2) \\ \phi(2, 1) & \phi(2, 2) \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

The points $x = (1, 2, 2, \dots)$ and $(2, 2, 2, \dots)$ are fixed points.

The point $x = (2, 1, 2, 2, 2, 2, \dots)$ has period two: $\Phi(x) = (1, 1, 2, 2, 2, \dots)$.

The next example shows that the sets $\mathcal{M}(\sigma)$ and $\mathcal{M}(\Phi)$ are different.

Example 2.4. Consider $r = 3$ and a case where lines can repeat as in

$$M = \begin{pmatrix} \phi(1, 1) & \phi(1, 2) & \phi(1, 3) \\ \phi(2, 1) & \phi(2, 2) & \phi(2, 3) \\ \phi(3, 1) & \phi(3, 2) & \phi(3, 3) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

In this case

$$\Phi^{-1}(\overline{1}) = \overline{33} \cup \overline{21} \cup \overline{11},$$

$$\Phi^{-1}(\overline{2}) = \overline{32} \cup \overline{22} \cup \overline{12}$$

and

$$\Phi^{-1}(\overline{3}) = \overline{31} \cup \overline{23} \cup \overline{13}.$$

Moreover, $\Phi^{-2}(\overline{1}) \cup \Phi^{-2}(\overline{2}) \cup \Phi^{-2}(\overline{3})$ is a total of 3^3 different cylinders.

The independent probability μ on $\{1, 2, 3\}^{\mathbb{N}}$ associated to the weights

$$(1/7, 2/7, 4/7)$$

is σ -invariant but not Φ -invariant. Indeed,

$$\mu(\Phi^{-1}(\overline{1})) = \mu(\overline{33}) + \mu(\overline{21}) + \mu(\overline{11}) =$$

$$4/7 \quad 4/7 + 2/7 \quad 1/7 + 1/7 \quad 1/7 \neq 4/7 \quad 1/7 + 2/7 \quad 1/7 + 1/7 \quad 1/7 = 1/7 = \mu(\overline{1}).$$

In this case $\mathcal{M}(\sigma) \neq \mathcal{M}(\Phi)$.

If μ_0 is the measure of maximal entropy (for the shift σ), then,

$$\mu_0(\overline{b_1, b_2, \dots, b_n}) = 3^{-n}, \quad (2.6)$$

and therefore from the above μ_0 is also Φ -invariant, that is $\mu_0 \in \mathcal{M}(\sigma, \Phi)$.

A particular case of ϕ as in (2.1) is the **bipermutative** CA Φ where we assume the extra assumption on ϕ : for each b , the law

$$a \in \mathcal{A} \rightarrow \phi(a, b) \text{ is bijective.} \quad (2.7)$$

Most of our results are for the general case of the permutative CA Φ . We point out that in several papers on the area what is called a bipermutative CA is always described by a local rule ϕ derived from some kind of group structure on the set of symbols $\{1, 2, \dots, r\}$; the algebraic cellular automata (see for instance [33] and [15]).

Example 2.5. We will provide an example where equation (1.3) is true for $\Phi_1 = \sigma$ and $\Phi_2 = \Phi$, $r = 2$, for functions that depends on the first two coordinates.

Consider ϕ such that

$$M = \begin{pmatrix} \phi(1, 1) & \phi(1, 2) \\ \phi(2, 1) & \phi(2, 2) \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}.$$

This example of ϕ defines a permutative Φ that is not bipermutative.

Take the functions A_1 and A_2 depending on the two first coordinates satisfying

$$A_1(x_1, x_2, \dots) = Q_{x_1, x_2}$$

and

$$A_2(x_1, x_2, \dots) = C_{x_1, x_2}.$$

Assume that

$$\begin{aligned} Q_{2,1} &= Q_{1,2}, C_{1,2} = C_{1,1} + Q_{1,2} - Q_{2,2}, \\ C_{2,1} &= C_{1,1} - Q_{1,1} + Q_{1,2}, C_{2,2} = C_{1,1}. \end{aligned} \quad (2.8)$$

Then,

$$A_1 - A_1 \circ \sigma = A_2 - A_2 \circ \Phi.$$

From [20] (see also Theorem 3.7) this implies that the σ -equilibrium for A_2 is equal to the Φ -equilibrium for A_1 . Then, $\mathcal{M}(\sigma, \Phi)$ is not empty.

Without loss of generality we can assume that $C_{12} = 1 - C_{11}$ and $C_{22} = 1 - C_{21}$ (see [22]). The σ -equilibrium probability for A_2 is a Markov stationary measure μ (see [22] or [43]). The set $\mathcal{M}(\sigma, \Phi)$ contains the one-parameter family of stationary Markov probabilities μ index by $-1 < \lambda < 1$, obtained

by symmetric line stochastic matrices, where $C_{2,1} - C_{1,1} = \lambda$. In this case $Q_{1,1} = Q_{1,2} - \lambda$ and $Q_{2,2} = Q_{1,1}, Q_{1,2} = Q_{2,1}$.

For the line stochastic matrix

$$\begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix},$$

the left invariant probability vector is $(\Pi_1, \Pi_2) = (1/2, 1/2)$.

In this case

$$\Phi^{-1}(\bar{1}) = \bar{1}, \bar{2} \cup \bar{2}, \bar{2},$$

and

$$\Phi^{-1}(\bar{2}) = \bar{1}, \bar{1} \cup \bar{2}, \bar{1}.$$

Then,

$$\mu(\Phi^{-1}(\bar{1})) = \mu(\bar{1}, \bar{2}) + \mu(\bar{2}, \bar{2}) = 1/2 \cdot 2/3 + 1/2 \cdot 1/3 = 1/2 = \mu(\bar{1}).$$

In the same way $\mu(\Phi^{-1}(\bar{2})) = \mu(\bar{2})$. This, computation highlights the property that Φ preserves the Markov probability μ .

If

$$M = \begin{pmatrix} \phi(1, 1) & \phi(1, 2) \\ \phi(2, 1) & \phi(2, 2) \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

there are also solutions for A_1 and A_2 when

$$\begin{aligned} Q_{2,1} &= 2Q_{1,2} - Q_{1,1}, Q_{2,2} = Q_{1,2}, C_{1,2} = C_{1,1} + Q_{1,1} - Q_{1,2}, \\ C_{2,1} &= C_{1,1} - Q_{1,1} + Q_{1,2}, C_{2,2} = C_{1,1}. \end{aligned}$$

In this case, one can deduce (after a simple manipulation) that the independent probability with weights $(1/2, 1/2)$ (the maximal entropy measure) is the only one that can be obtained as Φ -equilibrium for potentials that depend on two coordinates.

This example of ϕ defines a bipermutative CA Φ , and it also fits the case of an algebraic CA on two symbols (see [32] and [33]).

The property that we just show for this bipermutative CA (the maximal entropy measure is the only one in $\mathcal{M}(\sigma, \Phi)$, for potentials that depend on two coordinates) is in some sense a very particular example of a series of important cases covered by the setting of algebraic cellular automata (see [42], [41], [7] and [32]).

For instance in [17] the authors show that the only Φ -invariant, σ -ergodic measure m with with positive Φ -entropy is the maximal entropy measure. In another direction for affine cellular automata the authors in [27] show that the only measure with complete connections and summable decay, that is simultaneously invariant by the cellular automata and the shift map, is the measure of maximum entropy.

We will provide an example where equation (1.3) is true for $\Phi_1 = \sigma$ and $\Phi_2 = \Phi$, $r = 2$, for functions that depend on the first three coordinates in Example 7.3 in the Appendix. This will show the existence of more complex examples in $\mathcal{M}(\sigma, \Phi)$.

Proposition 2.6. *Consider $\mathcal{K} := \{\bar{1}, \bar{2}, \dots, \bar{r}\}$, a partition of Ω . Then, for all $n \geq 0$ we have*

$$\mathcal{K} \vee \Phi^{-1}(\mathcal{K}) \vee \dots \vee \Phi^{-n}(\mathcal{K}) = \bigsqcup_{a_1 \dots a_{n+1} \in \mathcal{A}} \overline{a_1 \dots a_{n+1}}.$$

Moreover, give n , $x \in \Omega$, for each $y = (a_1, a_2, \dots, a_n, y_{n+1}, y_{n+2}, \dots) \in \Phi^{-n}(x)$, there exist a unique z of the form $z = (a_1, a_2, \dots, a_n, z_{n+1}, z_{n+2}, \dots)$ in $\sigma^{-1}(x)$, and vice versa.

For proof see Proposition 7.2 in the Appendix.

Remark 2.7. *Given the natural partition $\mathcal{K} = \{\bar{1}, \bar{2}, \dots, \bar{r}\}$*

$$\mathcal{K}_n = \mathcal{K} \vee \Phi^{-1}\mathcal{K} \vee \dots \vee \Phi^{-n}\mathcal{K}$$

suppose that it is equal to the partition $\{\overline{a_1, a_2, \dots, a_n}, a_j \in \mathcal{A}, j = 1, 2, \dots, n\}$, for every $n \in \mathbb{N}$. This is the case we consider here as shown in Proposition 2.6.

In the case $\mu \in \mathcal{M}(\sigma)$, then the entropy according to σ

$$h(\mu) = h_\sigma(\mu) = - \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{(a_1, a_2, \dots, a_m) \in \{1, 2, \dots, d\}^m} \mu(\overline{a_1, a_2, \dots, a_m}) \log(\mu(\overline{a_1, a_2, \dots, a_m})). \quad (2.9)$$

In the case $\mu \in \mathcal{M}(\Phi)$, then the entropy according to Φ is

$$h(\mu) = h_\Phi(\mu) = - \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{S \in \mathcal{K}_n} \mu(S) \log(\mu(S)). \quad (2.10)$$

In this case, if $\mu \in \mathcal{M}(\sigma, \Phi)$, then $h_\Phi(\mu) = h_\sigma(\mu)$.

3 Thermodynamic Formalism for Φ

First will state some results on Thermodynamic Formalism for the Φ defined via (2.1). We point out that the proofs of these results follow from a simple adaptation of the ones claiming analogous results (which are well known) for

the shift σ , as in [31] (Φ is an expanding map as shown in Lemma 5.1). We point out that the IFS Thermodynamic Formalism (as in [18], [1] and [3]) can also be applied to Φ (the similar results that we will use here are also true)

The classical Ruelle Theorem is valid for σ ; the same proof also works replacing σ for Φ . In order to make a distinction we can write $\mathcal{L}_{A_2, \sigma}$ and $\mathcal{L}_{A_1, \Phi}$.

Given a Lipschitz potential $A : \Omega = \{1, 2, \dots, r\}^{\mathbb{N}} \rightarrow \mathbb{R}$, we interested in a probability μ_A which maximizes the pressure

$$P_{\Phi}(A) := \sup_{\rho \in \mathcal{M}(\Phi)} \{h_{\Phi}(\rho) + \int A(x) d\rho(x)\} = \quad (3.1)$$

$$h_{\Phi}(\mu_A) + \int A d\mu_A,$$

where A is the potential and $h_{\Phi}(\rho)$ is the Φ -entropy of ρ (defined before).

Such μ_A is unique and will be called the Φ -equilibrium for A .

Given a Lipschitz potential $A : \Omega \rightarrow \mathbb{R}$, define $\mathcal{L}_A(f) = \mathcal{L}_{A, \Phi}(f) = g$, via

$$g(y) = \mathcal{L}_A(f)(y) = \mathcal{L}_{A, \Phi}(f)(y) = \sum_{\Phi(x)=y} e^{A(x)} f(x). \quad (3.2)$$

Note that for each $n > 1$

$$g(y) = \mathcal{L}_A^n(f)(y) = \sum_{\Phi^n(x)=y} e^{\sum_{j=0}^{n-1} A(\Phi^j(x))} f(x). \quad (3.3)$$

Consider a Lipschitz-continuous potential A with Lipschitz constant c , and $x, y \in \mathcal{A}^{\mathbb{N}}$, then given $\mathbf{a} \in \mathcal{A}^k$, there exist just two points $\mathbf{a}x, \mathbf{a}y$ in the same cylinder set such that $\Phi^n(\mathbf{a}x) = x$, and $\Phi^n(\mathbf{a}y) = y$. An important property for the validity of the Ruelle Theorem (that we can also use for the dynamics of Φ) is

$$|A(\mathbf{a}x) - A(\mathbf{a}y)| \leq c d(\mathbf{a}x, \mathbf{a}y) \leq c \frac{1}{2^k} d(x, y) \quad (3.4)$$

Above we are using the second claim in Proposition 2.6.

Note that given a Lipschitz-continuous potential A , there exists $b > 0$, such that, $\forall n$, and $\forall \mathbf{j} = (j_1, j_2, \dots, j_n) \in \{1, 2, \dots, r\}^n = \mathcal{A}^n$

$$\frac{\prod_{j=0}^{n-1} e^{A(\Phi^j(x))}}{\prod_{j=0}^{n-1} e^{A(\Phi^j(y))}} < b, \quad (3.5)$$

$\forall x, y \in \overline{j_1, j_2, \dots, j_n} = \bar{\mathbf{j}}$.

From (3.4) follows that for a Lipschitz-continuous potential A , given $x, y \in \mathcal{A}^{\mathbb{N}}$, we get that

$$|S_{n,A}(\mathbf{a}x) - S_{n,A}(\mathbf{a}y)| \quad (3.6)$$

is bounded independently of n and the size of the word $\mathbf{a} \in \mathcal{A}^k$, $k > 0$.

The above property (3.5) (or (3.6)) is called the **bounded distortion property** for the Lipschitz-continuous potential A . This property is the key ingredient for analyzing the asymptotic of (3.3) when $n \rightarrow \infty$.

We denote by $\mathcal{C} = C(\Omega, \mathbb{R})$ the space of continuous functions $f : \Omega \rightarrow \mathbb{R}$.

For the proof of the next theorem see for instance [31] (adapted to our setting).

Theorem 3.1. *There exists a strict positive Lipschitz eigenfunction $\psi_A = \psi_{A,\Phi}$ for $\mathcal{L}_{A,\Phi} : \mathcal{C} \rightarrow \mathcal{C}$, associated to a unique strictly positive eigenvalue $\lambda_A = \lambda_{A,\Phi}$. The eigenvalue is simple (and isolated from the rest of the spectrum when \mathcal{L}_A acts on the set of Lipschitz functions) and it is equal to the supremum of the modulus of the values of the spectrum. Moreover, $\log \lambda_{A,\Phi} = P_\Phi(A)$.*

The pressure $P_\Phi(-\log r) = 0$ because in this case, the eigenvalue is equal to 1.

Remark 3.2. *If a continuous function $f > 0$ satisfies for some $\lambda > 0$*

$$\mathcal{L}_A(f) = \lambda f,$$

then, λ is the main eigenvalue and f is the main eigenfunction (see [31]).

Given a continuous potential $A : \Omega \rightarrow \mathbb{R}$, we can define the dual operator $\mathcal{L}_A^* = \mathcal{L}_{A,\Phi}^*$ on the space of the Borel finite measures on Ω , as the operator that sends a measure ν to the measure $\mathcal{L}_A^*(\nu)$, defined by

$$\int \psi d\mathcal{L}_A^*(\nu) = \int \mathcal{L}_A(\psi) d\nu. \quad (3.7)$$

for any $\psi \in \mathcal{C}$.

The operator \mathcal{L}_A^* acts on the space of all probabilities in Ω . Note that if ν is Φ -invariant, then, not necessarily the probability $\mathcal{L}_A^*(\nu)$ is Φ -invariant (the same claim is valid for σ).

Lemma 3.3. *Suppose $A : \Omega \rightarrow \mathbb{R}$ is a Lipschitz potential. Then, there exist a probability $\rho_A = \rho_{A,\Phi}$ on Ω and a real positive eigenvalue $\tilde{\lambda}_A = \tilde{\lambda}_{A,\Phi}$ such that*

$$\mathcal{L}_A^*(\rho_A) = \tilde{\lambda}_A \rho_A. \quad (3.8)$$

Moreover, $\tilde{\lambda}_A = \lambda_A$.

We call such ρ_A the eigenprobability for A .

Remark 3.4. *The eigenprobability ρ_A and the eigenfunction ψ_A are dual entities, one is obtained from \mathcal{L}_A^* and the other is obtained from \mathcal{L}_A ; both for the same eigenvalue λ_A .*

One can show (see [31]) that $\psi_{A,\Phi} \rho_{A,\Phi}$ is the equilibrium probability for $P_\Phi(A)$. Moreover, if w is such that $B = A + w - w \circ \Phi$, then the equilibrium probability for B and A are equal and $P_\Phi(A) = P_\Phi(B)$ (see [31]).

In [20] the authors considered transformations $\Phi_1, \Phi_2 : \Omega \rightarrow \Omega$ which commute and we will adapt their results for our setting: $\Phi_1 = \sigma$ and $\Phi_2 = \Phi$. For generality, we will state some claims in terms of Φ_1 and Φ_2 (assuming they satisfy the proper conditions).

Remark 3.5. *Denote by Π either Φ_1 or Φ_2 . We say that the Lipschitz potential B is normalized for Π if $\mathcal{L}_{B,\Pi}(1) = 1$.*

In this case

$$\lim_{n \rightarrow \infty} \mathcal{L}_{B,\Pi}^n(f) = \int f d\mu,$$

where μ is the equilibrium potential for B (see [31]). Moreover, $\mathcal{L}_{B,\Pi}^*(\mu) = \mu$.

If B is normalized it is usual to denote by $J : \Omega \rightarrow (0, 1)$ the function such that $B = \log J$. We call J the Jacobian of the probability μ such that $\mathcal{L}_{B,\Pi}^*(\mu) = \mathcal{L}_{\log J,\Pi}^*(\mu) = \mu$.

In the case of Φ , if J is a Jacobian, then for any x we have

$$\sum_{\Phi(y)=x} J(y) = 1.$$

The Jacobian in some sense plays the role of a stochastic matrix. In Example 8 in [22] this is properly described for the case of the shift.

Given a Lipschitz potential A , the associated eigenvalue λ_A and the corresponding eigenfunction ψ_A , the potential

$$\bar{A} = A + \log \psi_A - \log(\psi_A \circ \Phi) - \log \lambda_A$$

is normalized, has pressure zero and it is called the normalization of A . The equilibrium probability for A is the measure of maximal entropy, if and only if $\bar{A} = -\log r$ (see the end of Section 2 in [22]); this is true for Φ and σ .

An interesting result claims the following (see Lemma 2.1 in [20]):

Lemma 3.6. *Suppose Φ_1 and Φ_2 commute, then*

$$\mathcal{L}_{A_1, \Phi_2} \circ \mathcal{L}_{A_2, \Phi_1} = \mathcal{L}_{A_2, \Phi_1} \circ \mathcal{L}_{A_1, \Phi_2},$$

if and only if,

$$A_1 - A_1 \circ \Phi_1 = A_2 - A_2 \circ \Phi_2.$$

Below, in Theorem 3.7, we will fill in some details missing on the proof of Theorem 2.2 in [20] that we believe would be appropriate.

The next theorem contemplates the case

$$A_1 - A_1 \circ \sigma = A_2 - A_2 \circ \Phi,$$

taking $w = 0$. Therefore, $\mathcal{M}(\Phi, \sigma)$ is not empty for the case of Example 2.5.

Given Lipschitz potentials A_1 and A_2 , we ask: what are sufficient conditions on A_1, A_2 in such way that $\mu_{A_2, \Phi_1} = \mu_{A_1, \Phi_2}$?

Theorem 3.7. *Suppose Φ_1 and Φ_2 commute, and assume that A_1 and A_2 are Lipschitz functions. If for some Lipschitz function w we get*

$$(A_1 - A_1 \circ \Phi_1) - (A_2 - A_2 \circ \Phi_2) = (w - w \circ \Phi_1) - (w - w \circ \Phi_1) \circ \Phi_2, \quad (3.9)$$

Then, $\mu_{A_2, \Phi_1} = \mu_{A_1, \Phi_2}$.

Proof. Note that if $(B_1 - B_1 \circ \Phi_1) = (B_2 - B_2 \circ \Phi_2)$, then for any f

$$\mathcal{L}_{B_1, \Phi_2}(\mathcal{L}_{B_2, \Phi_1}(f)) = \mathcal{L}_{B_2 + B_1 \circ \Phi_1, \Phi_2 \circ \Phi_1}(f)(y) = \mathcal{L}_{B_2, \Phi_1}(\mathcal{L}_{B_1, \Phi_2}(f)). \quad (3.10)$$

Therefore, the Ruelle operators commute.

Condition (3.9) implies that $\mathcal{L}_{A_1, \Phi_2}$ and $\mathcal{L}_{A_2 + (w - w \circ \Phi_1), \Phi_1}$ commute.

Suppose that $\varphi > 0$ is eigenfunction for $\mathcal{L}_{A_1, \Phi_2}$, that is $\mathcal{L}_{A_1, \Phi_2}(\varphi) = \lambda\varphi$, and $\psi > 0$ is eigenfunction for $\mathcal{L}_{A_2, \Phi_1}$, that is $\mathcal{L}_{A_2, \Phi_1}(\psi) = \beta\psi$.

Moreover, we set $\mathcal{L}_{A_2, \Phi_1}^*(\rho) = \beta\rho$ and $\mathcal{L}_{A_1, \Phi_2}^*(\nu) = \lambda\nu$. The equilibrium probability for A_2 is $\psi\rho$.

Without loss of generality, we can assume that $\beta = \lambda = 1$. Indeed. Replacing A_1 by $A_1 - \log \lambda$ and A_2 by $A_2 - \log \beta$ we keep the relation (3.9) for the same w (and, respectively, the equilibrium properties will not change). We can also assume that A_2 is Φ_1 normalized, that is $\psi = 1$, which means $\rho = \mu_{A_2, \Phi_1, 1}$.

The eigenfunction for $\mathcal{L}_{A_2 + (w - w \circ \Phi_1), \Phi_1}$ is $f = e^{-w}$, indeed

$$\begin{aligned} \mathcal{L}_{A_2 + (w - w \circ \Phi_1), \Phi_1}(f)(x) &= \\ \mathcal{L}_{A_2 + (w - w \circ \Phi_1), \Phi_1}(e^{-w})(x) &= \sum_{\Phi_1(y)=x} e^{A_2(y)} \frac{e^{w(y)}}{e^{w \circ \Phi_1(y)}} e^{-w(y)} = \end{aligned}$$

$$\sum_{\Phi_1(y)=x} e^{A_2(y)} \frac{e^{w(y)}}{e^{w(x)}} e^{-w(y)} = \sum_{\Phi_1(y)=x} e^{A_2(y)} \frac{1}{e^{w(x)}} = \frac{1}{e^{w(x)}}.$$

Now, note that

$$\begin{aligned} \mathcal{L}_{A_2+(w-w \circ \Phi_1), \Phi_1}(\mathcal{L}_{A_1, \Phi_2}(\frac{1}{e^w})) &= \mathcal{L}_{A_1, \Phi_2}(\mathcal{L}_{A_2+(w-w \circ \Phi_1), \Phi_1}(\frac{1}{e^w})) = \\ &= \mathcal{L}_{A_1, \Phi_2}(\frac{1}{e^w}). \end{aligned}$$

Therefore, $g = \mathcal{L}_{A_1, \Phi_2}(\frac{1}{e^w}) > 0$ is eigenfunction for $\mathcal{L}_{A_2+(w-w \circ \Phi_1), \Phi_1}$, and associated to the eigenvalue 1. From Remark 3.2 we get that g is colinear with $\frac{1}{e^w}$. That is $\mathcal{L}_{A_1, \Phi_2}(\frac{1}{e^w}) = \gamma \frac{1}{e^w}$, for some $\gamma > 0$. Once more from Remark 3.2 we get that $\gamma = 1$, and the eigenfunction φ for $\mathcal{L}_{A_1, \Phi_2}$ is colinear with $\frac{1}{e^w}$. This shows that $w = -\log \varphi + c$.

This shows that (3.9) is true replacing w by $-\log \varphi$.

Then, $\mathcal{L}_{A_1, \Phi_2}$ and $\mathcal{L}_{A_2+((-\log \varphi)-(-\log \varphi) \circ \Phi_1), \Phi_1}$ commute.

Expression (3.9) is equivalent to

$$\begin{aligned} (A_2 - A_2 \circ \Phi_2) - (A_1 - A_1 \circ \Phi_1) &= \\ ((-w) - (-w) \circ \Phi_2) - ((-w) - (-w) \circ \Phi_2) \circ \Phi_1. \end{aligned} \quad (3.11)$$

Therefore, $\mathcal{L}_{A_2, \Phi_1}$ and $\mathcal{L}_{A_1-w+w \circ \Phi_2, \Phi_2}$ commute.

Note that the potential $A_1 - w + w \circ \Phi_2 - \log \gamma = A_1 - w + w \circ \Phi_2$ is Φ_2 normalized.

From Remark 3.5 we get for any continuous function h

$$\lim_{n \rightarrow \infty} \mathcal{L}_{A_2, \Phi_1}^n(\mathcal{L}_{A_1-w+w \circ \Phi_2-\log \gamma, \Phi_2})(h) = \int \mathcal{L}_{A_1-w+w \circ \Phi_2-\log \gamma, \Phi_2}(h) d\rho,$$

From the commutative property, we get from Remark 3.5

$$\lim_{n \rightarrow \infty} \mathcal{L}_{A_1-w+w \circ \Phi_2-\log \gamma, \Phi_2}(\mathcal{L}_{A_2, \Phi_1}^n(h)) = \mathcal{L}_{A_1-w+w \circ \Phi_2-\log \gamma, \Phi_2}(\int h d\rho) = \int h d\rho.$$

This shows that $\mathcal{L}_{A_1-w+w \circ \Phi_2, \Phi_2}^*(\rho) = \rho$. Therefore, we were able to show that ρ is the Φ_2 -equilibrium probability for A_1 . That is, $\mu_{A_2, \Phi_1, 1} = \mu_{A_1, \Phi_2, 1}$. \square

Remark 3.8. $(A - A \circ \sigma)$ is in some sense the discrete time version of the σ -derivative of the “function” A . Note that in the case μ is σ -invariant for the shift we get $\int (A - A \circ \sigma) d\mu = 0$, which corresponds to $\int_a^b f'(x) dx = 0$, when f is periodic in $[a, b]$.

The hypothesis

$$(A_1 - A_1 \circ \sigma) - (A_2 - A_2 \circ \sigma) = (w - w \circ \sigma) - (w - w \circ \sigma) \circ \sigma,$$

is a discrete version of the postulation

$$f' - g' = \frac{d^2}{dx^2} w.$$

The hypothesis

$$(A_1 - A_1 \circ \Phi_1) - (A_2 - A_2 \circ \Phi_2) = (w - w \circ \Phi_1) - (w - w \circ \Phi_1) \circ \Phi_2,$$

is in some sense a kind of mixed derivatives expression.

The next result is a kind of converse of Theorem 3.7.

Theorem 3.9. *Suppose Φ_1 and Φ_2 commute, and assume that A_1 and A_2 are Lipschitz functions. Also assume that $f d\mu = g d\nu$, where f, μ are, respectively, the eigenfunction and eigen-probability for $\mathcal{L}_{A_1, \Phi_2}$, and g, ν are, respectively, the eigenfunction and eigenprobability for $\mathcal{L}_{A_2, \Phi_1}$.*

Then, the Lipschitz function $w = \log g - \log f$ satisfies

$$(A_1 - A_1 \circ \Phi_1) - (A_2 - A_2 \circ \Phi_2) = (w - w \circ \Phi_1) - (w - w \circ \Phi_1) \circ \Phi_2. \quad (3.12)$$

For proof see Theorem 2.2 in [20].

Remark 3.10. *When asking properties derived from the equality between two equilibrium probabilities for respectively, A_1, Φ_2 and A_2, Φ_1 , we can assume in the hypothesis of Theorem 3.9 that $f = 1 = g$ (that is, A_1 is Φ_2 normalized and A_2 is Φ_1 normalized). In this way, from the hypotheses $\mu = \nu$, we derive the property*

$$(A_1 - A_1 \circ \Phi_1) = (A_2 - A_2 \circ \Phi_2). \quad (3.13)$$

4 The involution Kernel: σ and Φ duality

The eigenfunction and the eigenprobability are dual concepts (see Remark 3.4) and in this section, we will address results in some way related to this claim.

We will adapt the reasoning of [31], Proposition 1.2 in Section 1, to show that a natural skew dynamics associated with cellular automata (and the shift) have the special property that a potential $f(x, y)$, initially defined at the skew structure $\{1, 2, \dots, k\}^{\mathbb{N}} \times \{1, 2, \dots, k\}^{\mathbb{N}}$, is co-homologous to another

one $g(x, y)$ that depends only of the “future” coordinates x . This will generalize the well known result for the shift $\hat{\sigma}$ acting on $\{1, 2, \dots, k\}^{\mathbb{Z}}$, as in [31]. Here we represent

$$\{1, 2, \dots, k\}^{\mathbb{Z}} \simeq \{1, 2, \dots, k\}^{\mathbb{N}} \times \{1, 2, \dots, k\}^{\mathbb{N}} := \{(x, y) | x, y \in \{1, 2, \dots, k\}^{\mathbb{N}}\}$$

to avoid the identification of the coordinate correspondent to zero when we define our dynamics. In this notation, x is the “future” and y is the “past” because we could relabel y as the negative coordinates in $\{1, 2, \dots, k\}^{\mathbb{Z}}$.

The inverse branch $\tau_j : \Omega \rightarrow \Omega$ is given by the formula

$$\tau_j(x) = y \text{ if } \Phi(y) = x \text{ and } y_1 = j.$$

In the next result, we assume that $x \rightarrow A(x)$ is a function of the future variables x and $y \rightarrow A^*(y)$ is a function of the past variable y . One of the purposes of Subsection 4.1 is to derive such result from a more general reasoning which is of interest in itself.

Proposition 4.1. *Assume that A depending just of the future coordinates x , and A^* which depends on past coordinates y , are such that, there exist $W : \Omega \times \Omega$ satisfying for $i = 1, 2, \dots, r$, $x, y \in \Omega$*

$$(A + W)(ix, y) = (A^* + W)(x, \tau_i(y)). \quad (4.1)$$

Then, if ρ_{A^} is the eigenprobability for the Φ -Ruelle operator $\mathcal{L}_{A^*, \Phi}$ of the potential A^* , then*

$$\phi(x) = \int e^{W(x, y)} d\rho_{A^*}(y)$$

is the eigenfunction for the σ -Ruelle operator $\mathcal{L}_{A, \sigma}$ of the potential A .

Moreover, given a Lipschitz function $y \rightarrow A^(y)$, there exists a Lipschitz function $x \rightarrow A(x)$ and W as in (4.1).*

W will be called an involution kernel for A , and A is called the dual potential of A^* . The integral kernel $K(x, y) = e^{W(x, y)}$ relates eigenprobabilities and eigenfunctions.

We will present the proof of such results in Propositions 4.3 and 4.5. In this direction, it will be necessary to introduce the mixed skew product $\hat{\Phi}$ to be defined later (see (4.2)). Proposition 4.1 relates dual objects: the eigenprobabilities for the CA Φ and eigenfunctions for σ . Proposition 4.3 is similar, but different for Proposition 8 item (2) in [9], where it is considered $\hat{\sigma}$ instead of $\hat{\Phi}$.

Given $0 < \theta < 1$, consider the maximum metric \hat{d}_θ defined by

$$\hat{d}_\theta((x, y), (x', y')) := \max(d_\theta(x, x'), d_\theta(y, y')),$$

where

$$d_\theta(x, x') = \theta^N$$

for N being the first coordinate where $x_i \neq x'_i$ (and zero if $x = x'$). The metric \hat{d}_θ is equivalent to the one defined in [31].

In the general setting, we get a matrix M of zeroes and ones and deal with the subshifts of finite type

$$\begin{aligned} \hat{\Sigma}_M &:= \\ \{(x, y) | x, y \in \{1, 2, \dots, k\}^\mathbb{N}, M(x_i, x_{i+1}) = 1, M(y_{i-1}, y_i) = 1, M(y_1, x_1) = 1\}, \\ \Sigma_M &:= \{x | x \in \{1, 2, \dots, k\}^\mathbb{N}, M(x_i, x_{i+1}) = 1\}. \end{aligned}$$

As usual, the “future” and the “past” projections are well defined respectively as $\pi_1, \pi_2 : \hat{\Sigma}_M \rightarrow \Sigma_M$ by

$$\pi_1(x, y) = x \text{ and } \pi_2(x, y) = y.$$

In this setting we define an immersion function $\varphi : \Sigma_M \rightarrow \hat{\Sigma}_M$ by

$$\varphi(x) = (x, b(x_1))$$

where $M(b_1, x_1) = 1$ and $b(x_1) \in \Sigma_M$ is a fixed sequence.

Notice that the above formalism is necessary when we are dealing with actual subshifts of finite type. From now on, we assume we are considering **just the full shift (for simplification of the argument)**. In this case $\Sigma_M = \{1, 2, \dots, k\}^\mathbb{N}$, $\hat{\Sigma}_M = \{1, 2, \dots, k\}^\mathbb{N} \times \{1, 2, \dots, k\}^\mathbb{N}$ and the immersion function assume a very simple form

$$\varphi(x) = (x, y')$$

where $y' \in \{1, 2, \dots, k\}^\mathbb{N}$ is a fixed sequence, for all x .

In order to clarify the notation we denote $\Omega := \Sigma_M = \{1, 2, \dots, k\}^\mathbb{N}$, where it is necessary.

Assume $f : \Omega^2 \rightarrow \mathbb{R}$ is a Lipschitz potential. We will consider two skew homeomorphisms (the mixed one in Subsection 4.1 and the non-mixed one in Subsection 4.2), and our main goal is to relate f (via a cohomological equation) to a new potential $g : \Omega^2 \rightarrow \mathbb{R}$ such that $g(x, y) = g(x, z)$, for all $x, y, z \in \Omega$, that is, g depends only of the “future” coordinates.

4.1 The mixed skew product $\hat{\Phi}$

Consider $\hat{\Phi} : \Omega^2 \rightarrow \Omega^2$ defined by

$$\hat{\Phi}(x, y) = (\sigma(x), \tau_{x_1}(y)), \quad (4.2)$$

where the inverse branch $\tau_j : \Omega \rightarrow \Omega$ is given by the formula

$$\tau_j(x) = y \text{ if } \Phi(y) = x \text{ and } y_1 = j.$$

We introduce the n -branch

$$\tau_{n,x}(y) = \pi_2(\hat{\Phi}^n(x, y)).$$

For a fixed $z \in \Omega$ we define the immersion function $\varphi(x) = (x, z)$ and for any $x, y \in \Omega$ we define the **mix** W -kernel as the *formal* correspondence $W : \Omega^2 \rightarrow \mathbb{R}$ by

$$W(x, y) := \sum_{n \geq 0} f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(\varphi(x))) = \sum_{n \geq 0} f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x, z)). \quad (4.3)$$

The next result generalizes the result from [31] if we choose the trivial cellular automata (that is $\phi(x, y) = y$).

Theorem 4.2. *In the above conditions, we have the following properties.*

1. *The skew map $\hat{\Phi}$ is a homeomorphism.*
2. *The series defining $W(x, y)$ is absolutely convergent.*
3. *Given a Lipschitz function $f : \Omega^2 \rightarrow \mathbb{R}$, there exists a Lipschitz function $g : \Omega^2 \rightarrow \mathbb{R}$ such that*

$$f(x, y) = g(x, y) + W(x, y) - W(\hat{\Phi}(x, y)), \quad (4.4)$$

and $g(x, y) = g(x, z)$ for all $x, y, z \in \Omega$.

4. *W is a Lipschitz function with respect to the metric $d_{\sqrt{\theta}}$ in Ω^2 and $\text{Lip}(W) \leq 2\text{Lip}(f) \left(\frac{1+\theta}{1-\theta}\right)$. In particular, g is also a Lipschitz function with respect to the metric $d_{\sqrt{\theta}}$.*

Proof. (1) $\hat{\Phi}$ is obviously continuous because each coordinate is. To see that it is a bijection suppose

$$\hat{\Phi}(x, y) = (\sigma(x), \tau_{x_1}(y)) = (\sigma(x'), \tau_{x'_1}(y')) = \hat{\Phi}(x', y').$$

Since $\sigma(x') = \sigma(x)$, the strings x, x' coincide, except eventually by the first coordinate. On the other hand $z' = \tau_{x'_1}(y') = \tau_{x_1}(y) = z$ meaning that $\Phi(z) = y$ and $z_1 = x_1$ and $\Phi(z') = y'$ and $z'_1 = x'_1$ thus $x_1 = x'_1$. From the equations $\Phi(z) = y$, $\Phi(z') = y'$ and $z = z'$, we get $y = y'$ and $x = x'$.

(2) Consider the absolute value series

$$\sum_{n \geq 0} |f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x, z))| \leq \sum_{n \geq 0} \text{Lip}(f) \theta^n d(y, z) < \infty.$$

(3) One just needs to transform the equation

$$\begin{aligned} W(x, y) - W(\hat{\Phi}(x, y)) &= \\ &= \sum_{n \geq 0} f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x, z)) - \sum_{n \geq 0} f(\hat{\Phi}^{n+1}(x, y)) - f(\hat{\Phi}^n(\sigma(x), z)) = \\ &= f(x, y) - f(x, z) + f(\hat{\Phi}(x, y)) - f(\hat{\Phi}(x, z)) + f(\hat{\Phi}^2(x, y)) - \\ &\quad f(\hat{\Phi}^2(x, z)) + f(\hat{\Phi}^3(x, y)) - f(\hat{\Phi}^3(x, z)) + \\ &\quad + \dots - \left[f(\hat{\Phi}(x, y)) - f(\hat{\Phi}(\sigma(x), z)) + f(\hat{\Phi}^2(x, y)) - \right. \\ &\quad \left. f(\hat{\Phi}^2(\sigma(x), z)) + f(\hat{\Phi}^3(x, y)) - f(\hat{\Phi}^3(\sigma(x), z)) + \dots \right] = \\ &= f(x, y) - \left[f(x, z) + \sum_{n \geq 1} f(\hat{\Phi}^n(x, z)) - f(\hat{\Phi}^n(\sigma(x), z)) \right]. \end{aligned}$$

Thus, taking

$$g(x, y) = f(x, z) + \sum_{n \geq 1} f(\hat{\Phi}^n(x, z)) - f(\hat{\Phi}^n(\sigma(x), z))$$

we get the cohomological equation (4.4), where W is given by (4.3). We notice that $g(x, y)$ depends only on the future coordinates (given by x).

(4) Take any $x, x', y, y' \in \Omega$.

If, for a fixed $N > 0$ we suppose $d((x, y), (x', y')) = \theta^{2N}$, then

$$\begin{aligned} W(x, y) - W(x', y') &= \\ &= \sum_{n \geq 0} f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x, z)) - \sum_{n \geq 0} f(\hat{\Phi}^n(x', y')) - f(\hat{\Phi}^n(x', z)), \end{aligned}$$

can be evaluated as follows.

For $n \leq N$ we have

$$|f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x', y'))| \leq |f(\sigma^n(x), \tau_{n,x}(y)) - f(\sigma^n(x'), \tau_{n,x'}(y'))| \leq$$

$$\begin{aligned} & \text{Lip}(f) \max(d_\theta(\sigma^n(x), \sigma^n(x')), d_\theta(\tau_{n,x}(y), \tau_{n,x'}(y'))) \leq \\ & \leq \text{Lip}(f) \max(\theta^{-n} d_\theta(x, x'), \theta^n d_\theta(y, y')) \leq \text{Lip}(f) \theta^{2N-n}. \end{aligned}$$

A similar reasoning shows that

$$|f(\hat{\Phi}^n(x, z)) - f(\hat{\Phi}^n(x', z))| \leq \text{Lip}(f) \theta^{2N-n}.$$

Analogously, for every $n \geq 1$ we get

$$\begin{aligned} & |f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x, z))| = |f(\sigma^n(x), \tau_{n,x}(y)) - f(\sigma^n(x), \tau_{n,x}(z))| \leq \\ & \leq \text{Lip}(f) \max(d_\theta(\sigma^n(x), \sigma^n(x)), d_\theta(\tau_{n,x}(y), \tau_{n,x}(z))) \leq \\ & \leq \text{Lip}(f) \max(0, \theta^n d_\theta(y, y')) \leq \text{Lip}(f) \theta^n. \end{aligned}$$

A similar reasoning shows that

$$|f(\hat{\Phi}^n(x', y)) - f(\hat{\Phi}^n(x', z))| \leq \text{Lip}(f) \theta^n.$$

From these four inequalities, we can write

$$\begin{aligned} & |W(x, y) - W(x', y')| \leq \\ & \leq \left| \sum_{0 \leq n \leq N} f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x, z)) - \sum_{0 \leq n \leq N} f(\hat{\Phi}^n(x', y')) - f(\hat{\Phi}^n(x', z)) \right| + \\ & \left| \sum_{n > N} f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x, z)) - \sum_{n > N} f(\hat{\Phi}^n(x', y')) - f(\hat{\Phi}^n(x', z)) \right| \leq \\ & \leq \sum_{0 \leq n \leq N} |f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x', y'))| + \sum_{0 \leq n \leq N} |f(\hat{\Phi}^n(x, z)) - f(\hat{\Phi}^n(x', z))| + \\ & \sum_{n > N} |f(\hat{\Phi}^n(x, y)) - f(\hat{\Phi}^n(x, z))| + \sum_{n > N} |f(\hat{\Phi}^n(x', y')) - f(\hat{\Phi}^n(x', z))| \leq \\ & \leq \sum_{0 \leq n \leq N} \text{Lip}(f) \theta^{2N-n} + \sum_{0 \leq n \leq N} \text{Lip}(f) \theta^{2N-n} + \\ & \sum_{n > N} \text{Lip}(f) \theta^n + \sum_{n > N} \text{Lip}(f) \theta^n \leq \\ & 2\text{Lip}(f) \left(\theta^{2N} \sum_{0 \leq n \leq N} \theta^{-n} + \theta^{N+1} \frac{1}{1-\theta} \right) = \\ & = 2\text{Lip}(f) \left(\theta^N \sum_{0 \leq n \leq N} \theta^{-n} + \frac{\theta}{1-\theta} \right) \theta^N = C d_{\sqrt{\theta}}((x, y), (x', y')), \end{aligned}$$

for

$$C := 2\text{Lip}(f) \left(\frac{1+\theta}{1-\theta} \right) < \infty.$$

Here we used the fact that

$$\theta^N \sum_{0 \leq n \leq N} \theta^{-n} = \sum_{0 \leq n \leq N} \theta^n \leq \frac{1}{1-\theta}.$$

□

Proposition 4.3. *Assume that A depending just of the future coordinates x and A^* which depends on past coordinates y are such that, there exist $W : \Omega \times \Omega$ satisfying*

$$(A + W)(ix, y) = (A^* + W)(x, \tau_i(y)). \quad (4.5)$$

Then, if ρ_{A^} is the eigenprobability for the Φ -Ruelle operator $\mathcal{L}_{A^*, \Phi}$ of the potential A^* , then*

$$\phi(x) = \int e^{W(x,y)} d\rho_{A^*}(y)$$

is the eigenfunction for the σ -Ruelle operator $\mathcal{L}_{A, \sigma}$ of the potential A .

Moreover, given A^ , there exist A and W as in (4.5).*

Proof. Note that

$$\begin{aligned} \mathcal{L}_A^\sigma(\phi)(x) &= \mathcal{L}_A^\sigma \left(\int e^{W(\cdot, y)} d\rho_{A^*}(y) \right)(x) = \\ &= \int \mathcal{L}_A^\sigma(e^{W(\cdot, y)})(x) d\rho_{A^*}(y) = \int \sum_i e^{A(ix)} e^{W(ix, y)} d\rho_{A^*}(y) = \\ &= \int \sum_i e^{A^*(\tau_i(y))} e^{W(x, iy)} d\rho_{A^*}(y) = \int \mathcal{L}_{A^*}^\Phi(e^{W(x, \cdot)}) d\rho_{A^*}(y) = \\ &= \lambda_{A^*}^\phi \int e^{W(x, y)} d\rho_{A^*}(y). \end{aligned}$$

From Remark 3.2 we get that $\lambda_{A^*}^\phi = \lambda_A^\sigma$.

For the proof of the existence of A and W see Proposition 4.5 and Remark 4.4. □

Remark 4.4. *A particular situation of the previous result occurs when f is defined in the following way: consider a Lipschitz potential $A^* : \Omega \rightarrow \mathbb{R}$ depending only on the past, and take*

$$f(x, y) = A^*(\pi_2(\hat{\Phi}(x, y))) = A^*(\pi_2(\sigma(x), \tau_{1,x}(y))) = A^*(\tau_{1,x}(y)) = A^*(\tau_{x_1}(y)).$$

In this case, the cohomological equation produces

$$A^*(\tau_{1,x}(y)) = g(x, y) + W(x, y) - W(\hat{\Phi}(x, y)),$$

where $g(x, y)$ does not depend on y , meaning that there exists a potential $A : \Omega \rightarrow \mathbb{R}$, given by

$$\begin{aligned} A(x) &= f(x, z) + \sum_{n \geq 1} f(\hat{\Phi}^n(x, z)) - f(\hat{\Phi}^n(\sigma(x), z)) = \\ &A^*(\tau_{1,x}(z)) + \sum_{n \geq 1} A^*(\tau_{n+1,x}(z)) - A^*(\tau_{n+1,\sigma(x)}(z)), \end{aligned}$$

also Lipschitz (w.r.t. the appropriated metric) such that

$$A(x) = A^*(\pi_2(\hat{\Phi}(x, y))) + W(x, y) - W(\hat{\Phi}(x, y)),$$

for some W .

In this situation we say that A and A^* are dual potentials w.r.t. the W -kernel W chosen by fixing z .

From this, we get a simpler cohomological equation

$$A^*(\tau_{1,x}(y)) = A(x) + W(x, y) - W(\hat{\Phi}(x, y)),$$

or, equivalently

$$A^*(\tau_{1,\sigma(x)}(y)) + W(\hat{\Phi}(x, y)) = A(x) + W(x, y). \quad (4.6)$$

From the above Remark, we get:

Proposition 4.5. *Given a Lipschitz function $y \rightarrow A^*(y)$, there exists a Lipschitz function $x \rightarrow A(x)$, and a bi-Lipschitz function $(x, y) \rightarrow W(x, y)$, such that for all $i = 1, 2, \dots, r$ and x, y*

$$(A + W)(ix, y) = (A^* + W)(x, \tau_i(y)). \quad (4.7)$$

In a similar fashion, given A one can find A^* and W such that (4.7) is true.

Proof. Replacing x by $ix = (i, x_1, x_2, \dots)$ in (4.6) we get

$$(A + W)(ix, y) = (A^* + W)(\sigma(ix), \tau_{1,ix}(y)) = (A^* + W)(x, \tau_i(y)). \quad (4.8)$$

□

In the above we replace $\hat{\Phi}$ by the shift $\hat{\sigma}$, and we exchange the variables x and y above, then we get the similar expression described by Definition 6 obtained in [9].

Finally, we notice that the dual relation can be reversed, that is,

$$A(x) = A^*(\pi_2(\hat{\Phi}(x, y))) + W(x, y) - W(\hat{\Phi}(x, y)),$$

is equivalent to

$$A(\pi_1(\hat{\Phi}^{-1}(x, y))) = A^*(y) + W(\hat{\Phi}^{-1}(x, y)) - W(x, y),$$

or

$$A^*(y) = A(\pi_1(\hat{\Phi}^{-1}(x, y))) + W(x, y) - W(\hat{\Phi}^{-1}(x, y)),$$

where $\hat{\Phi}^{-1}(x, y) = (y_1x, \Phi(y))$.

4.2 The non-mixed skew product $\hat{\Phi}_n$

In this subsection we consider the non-mixed skew product $\hat{\Phi}_n : \Omega^2 \rightarrow \Omega^2$ defined by

$$\hat{\Phi}_n(x, y) = (\Phi(x), \tau_{x_1}(y)),$$

where the inverse branch $\tau_j : \Omega \rightarrow \Omega$ is given by the formula

$$\tau_j(x) = y \text{ if } \Phi(y) = x \text{ and } y_1 = j.$$

We introduce the k -branch

$$\tau_{k,x}(y) = \pi_2(\hat{\Phi}_n^k(x, y)).$$

For a fixed $z \in \Omega$ we define the immersion function $\varphi(x) = (x, z)$ and for any $x, y \in \Omega$ we define the **non-mixed** involution kernel W as the *formal* correspondence $W : \Omega^2 \rightarrow \mathbb{R}$ given by

$$W_n(x, y) := \sum_{k \geq 0} f(\hat{\Phi}_n^k(x, y)) - f(\hat{\Phi}_n^k(\varphi(x))) = \sum_{k \geq 0} f(\hat{\Phi}_n^k(x, y)) - f(\hat{\Phi}_n^k(x, z)).$$

In a similar way as in the last subsection, we can show that such W_n is well-defined. Moreover, given a Lipschitz function $f : \Omega^2 \rightarrow \mathbb{R}$, there exists a Lipschitz function $g : \Omega^2 \rightarrow \mathbb{R}$, such that,

$$f(x, y) = g(x, y) + W_n(x, y) - W_n(\hat{\Phi}_n(x, y)), \quad (4.9)$$

and $g(x, y) = g(x, z)$ for all $x, y, z \in \Omega$.

A similar version of Theorem 4.2 for $\hat{\Phi}_n$ is true (the reasoning is similar to the one in the last subsection).

Denote by $\mathcal{M}(\hat{\Phi}_n)$ the set of Borel probabilities on $\Omega \times \Omega$ which are invariant by $\hat{\Phi}_n$. Given a Lipschitz potential $\hat{A} : \Omega \times \Omega \rightarrow \mathbb{R}$, consider the topological pressure problem

$$P(\hat{A}) := \sup_{\hat{\mu} \in \mathcal{M}(\hat{\Phi}_n)} \{h(\hat{\mu}) + \int \hat{A} d\hat{\mu}\}, \quad (4.10)$$

where $h(\hat{\mu})$ is the Shannon-Kolmogorov entropy of $\hat{\mu}$. A probability $\hat{\mu}_{\hat{A}}$ attaining the supremum value $P(\hat{A})$ will be called an $\hat{\Phi}_n$ -equilibrium probability for \hat{A} .

Of course, if $\hat{B} : \Omega \times \Omega \rightarrow \mathbb{R}$ is such that there exist a continuous $\hat{C} : \Omega \times \Omega \rightarrow \mathbb{R}$ satisfying

$$\hat{B} = \hat{A} + \hat{C} - \hat{C} \circ \hat{\Phi}_n,$$

then an equilibrium probability for \hat{A} is an equilibrium probability for \hat{B} , and vice-versa.

The reasoning showing the validity of expression (4.9) implies that we can replace in the above Pressure problem (4.10) the potential $\hat{A} : \Omega \times \Omega \rightarrow \mathbb{R}$ by a Lipschitz potential $\hat{B} : \Omega \rightarrow \mathbb{R}$, that depends just on future coordinates x . The equilibrium probabilities for \hat{A} and $\hat{B} := B$ will be the same. As B depends just on future coordinates one can define the Ruelle operator $\mathcal{L}_{B,\Phi}$ and take advantage of the Ruelle Theorem. Suppose $\mu_{B,\Phi}$ is the Φ -equilibrium probability for B . This will define the probability $\mu_{B,\Phi}(\overline{a_1, a_2, \dots, a_m})$, $a_j \in \{1, 2, \dots, r\}$, $j = 1, 2, \dots, m$, for any cylinder $\overline{a_1, a_2, \dots, a_m}$. This means that we are in fact defining probabilities for sets of the form

$$\Omega \times \overline{a_1, a_2, \dots, a_m} \subset \Omega \times \Omega.$$

In (4.10) we are interested only in probabilities on $\mathcal{M}(\hat{\Phi}_n)$. There is only way to extend $\mu_{B,\Phi}$ for a $\hat{\Phi}_n$ -invariant probability $\hat{\mu}$ on $\Omega \times \Omega$. We set

$$\mu(\hat{\Phi}_n^{-k}(\Omega \times \overline{a_1, a_2, \dots, a_m})) = \mu_{B,\Phi}(\overline{a_1, a_2, \dots, a_m}).$$

The probability μ is called the natural extension of $\mu_{B,\Phi}$. Note that for our Φ (obtained from the local rule ϕ defined on section 2) we get that $\hat{\Phi}_n^{-k}(\Omega \times \overline{a_1, a_2, \dots, a_m})$ will exhaust the class of all possible cylinders in $\Omega \times \Omega$, changing k , the cylinders $\overline{a_1, a_2, \dots, a_m}$, etc. These cylinders will generate the Borel sigma-algebra of $\Omega \times \Omega$. In this way, we can identify the $\hat{\Phi}_n$ -equilibrium probability for \hat{A} via such μ . The several ergodic properties, for instance, exponential decay of correlations, for the Φ -equilibrium probability for B are transferred for the $\hat{\Phi}_n$ -equilibrium probability for \hat{A} . In other words, we can take advantage of the properties described in Section 3 for Φ , but now for the $\hat{\Phi}_n$ -equilibrium probability for \hat{A} .

5 Livsic's theorem for cellular automata

In this section, we will present a version of Livsic's Theorem for the permutative CA Φ satisfying (2.1) of Section 2. This result is based on estimates of the Birkhoff averages of the potential A over periodic orbits. Among other things, we are interested in criteria to be able to decide whether a certain potential A of Holder class has, or not, as equilibrium probability the measure of maximum entropy. We also present a criteria for finding periodic orbits for Φ .

We recall that, for given $0 < \theta < 1$, the usual metric is defined by

$$d_\theta(x, x') = \theta^N,$$

for N being the first coordinate where $x_i = x'_i$ (and zero otherwise), makes Φ be a uniform expanding homeomorphism (by θ^{-1}) and the inverse branches, $\tau_j : \Omega \rightarrow \Omega$ are given by the formula

$$\tau_j(x) = y \text{ if } \Phi(y) = x \text{ and } y_1 = j,$$

and are uniform contractions by θ .

Lemma 5.1. *The endomorphism $\Phi : \Omega \rightarrow \Omega$ is expanding, that is, there exist $\epsilon > 0$ and $L := \theta^{-1} > 1$ such that $d(\Phi(x), \Phi(y)) \geq Ld(x, y)$ for all $x, y \in \Omega$ such that $d(x, y) < \epsilon$.*

Proof. Indeed, if $d(x, y) = \theta^N < \epsilon$, $x \neq y$ then $x = (a_1, \dots, a_{N-1}, x_N, \dots)$ and $y = (a_1, \dots, a_{N-1}, y_N, \dots)$ with $x_N \neq y_N$. As

$$\Phi(x) = (\phi(a_1, a_2), \phi(a_2, a_3), \dots, \phi(a_{N-2}, a_{N-1}), \phi(a_{N-1}, x_N), \dots)$$

and

$$\Phi(y) = (\phi(a_1, a_2), \phi(a_2, a_3), \dots, \phi(a_{N-2}, a_{N-1}), \phi(a_{N-1}, y_N), \dots)$$

we have $\phi(a_{N-1}, x_N) \neq \phi(a_{N-1}, y_N)$ meaning that $d(\Phi(x), \Phi(y)) = \theta^{N-1} = \theta^{-1}d(x, y)$. Thus we can take $L := \theta^{-1} > 1$. \square

Theorem 5.2. *The map Φ satisfies the closing lemma property, that is, for every $\epsilon > 0$ there exists $\delta > 0$ such that if $x \in \Omega$ and $n \geq 0$ are such that $d(\Phi^n(x), x) < \delta$, then there exists $y \in \Omega$ such that $\Phi^n(y) = y$ and $d(\Phi^k(y), \Phi^k(x)) < \epsilon$ for all $0 \leq k \leq n - 1$.*

Proof. Consider the orbit $\Phi^k(x)$ for $0 \leq k \leq n$ there is always j_k such that $\tau_{j_k}(\Phi^k(x)) = \Phi^{k-1}(x)$ thus

$$\tau_{j_1}(\cdots \tau_{j_n}(\Phi^n(x))) = x.$$

Since $\tau_{j_1} \circ \cdots \circ \tau_{j_n}$ is a contraction by $\theta^n \ll 1$ we get a unique fixed point $y \in \Omega$ such that $\tau_{j_1}(\cdots \tau_{j_n}(y)) = y$. It is easy to see that $\Phi^n(y) = y$.

In this way,

$$\begin{aligned} d(y, x) &= d(\tau_{j_1}(\cdots \tau_{j_n}(y)), \tau_{j_1}(\cdots \tau_{j_n}(\Phi^n(x)))) = \theta^n d(y, \Phi^n(x)) \leq \\ &\leq \theta^n (d(y, x) + d(x, \Phi^n(x))) < \theta^n (d(y, x) + \delta) \end{aligned}$$

or

$$d(y, x) < \frac{\theta^n \delta}{1 - \theta^n}.$$

So, if $\frac{\theta^n \delta}{1 - \theta^n} < \varepsilon$ or

$$\delta < \frac{1 - \theta^n}{\theta^n} \varepsilon = (\theta^{-n} - 1) \varepsilon$$

then the proof is concluded since $\theta^{-n} - 1 > 0$ is bounded away from zero for all n . \square

Recall that, for a potential $A : \Omega \rightarrow \mathbb{R}$, the sum

$$S_n A(x) := \sum_{k=0}^{n-1} A(\Phi^k(x))$$

is well-defined.

The Walters property (w.r.t. Φ) is: for every $\zeta > 0$ there exist $\varepsilon > 0$ such that if $x, y \in \Omega$ and $n \geq 0$ are such that $d(\Phi^k(x), \Phi^k(y)) < \varepsilon$ for all $0 \leq k \leq n - 1$, then $|S_n A(x) - S_n A(y)| < \zeta$ (see [8]).

Theorem 5.3. *If $A : \Omega \rightarrow \mathbb{R}$ is a Lipschitz (Hölder) potential then it has the Walters property w.r.t. Φ .*

Proof. Consider $\zeta > 0$, and $x, y \in \Omega$ and $n \geq 0$ are such that

$$d(\Phi^k(x), \Phi^k(y)) < \varepsilon,$$

for all $0 \leq k \leq n - 1$. As Φ is expanding by θ^{-1} we know that

$$d(\Phi^{n-1}(x), \Phi^{n-1}(y)) < \varepsilon,$$

only if $d(x, y) < \theta^{n-1} \varepsilon$ thus

$$\begin{aligned} |S_n A(x) - S_n A(y)| &\leq \sum_{k=0}^{n-1} |A(\Phi^k(x)) - A(\Phi^k(y))| \leq \text{Lip}(A) \sum_{k=0}^{n-1} \theta^{-k} d(x, y) \leq \\ &< \text{Lip}(A) \sum_{k=0}^{n-1} \theta^{-k} \theta^{n-1} \varepsilon \leq \frac{\text{Lip}(A)}{1-\theta} \varepsilon < \zeta \end{aligned}$$

if

$$\varepsilon < \frac{1-\theta}{\text{Lip}(A)} \zeta,$$

concluding our proof (the Hölder case is similar). \square

Let X be a compact metric space, $T : X \rightarrow X$ a continuous map, and $A : X \rightarrow \mathbb{R}$. We recall that $A : X \rightarrow \mathbb{R}$ is a coboundary if $A = h - h \circ T$ for some continuous map $h : X \rightarrow \mathbb{R}$ and A is cohomologous to $B : X \rightarrow \mathbb{R}$ if $A - B$ is a coboundary, that is, $A = B + h - h \circ T$. In particular, any coboundary is cohomologous to zero.

An important result is

Theorem 5.4. (*Livsic Theorem - see Section 19.2 in [19]*) *Let X be a compact metric space, $T : X \rightarrow X$ a continuous map satisfying the Closing Lemma and possessing a point whose orbit is dense, and $A : X \rightarrow \mathbb{R}$ a continuous function satisfying the Walters Property. Then A is a coboundary ($A = h - h \circ T$) if and only if for every periodic point $x = T^n(x) \in X$, we have $S_n A(x) = 0$.*

Obviously, $X = \Omega$ is a compact metric space and $T = \Phi$ is a continuous map. Moreover, each Lipschitz potential is continuous, and theorems 5.2 and 5.3 ensure that the Closing lemma and Walters property are true for Φ . Thus we have the following corollary.

Corollary 5.5. *If $A : \Omega \rightarrow \mathbb{R}$ is a Lipschitz (Hölder) potential and Φ is transitive, then A is a coboundary ($A = h - h \circ \Phi$) if and only if for every periodic point $x = \Phi^n(x)$, we have $S_n A(x) = 0$.*

The map Φ of expression (2.4) is transitive as shown in Theorem 5.9. Therefore, Corollary 5.5 is true for such Φ .

Remark 5.6. *Note that given Lipschitz (Holder) potentials A, B , and the Φ -equilibrium probabilities m_A and m_B , there exists normalized potentials \bar{A} and \bar{B} , such that the equilibrium probabilities for \bar{A} and \bar{B} are respectively m_A and m_B . It follows from the end of Section 2 in [22] that $m_A \neq m_B$,*

if and only if $\bar{A} \neq \bar{B}$. It follows that for a given Hölder potential A , the equilibrium probability m_A is the maximal entropy measure, if and only if, \bar{A} is constant.

Proposition 5.7. *If A is a Lipschitz (Hölder) potential such that for some constant c , is true that for any n periodic orbit $x = \Phi^n(x) \in X$, we have that $S_n(A+c)(x) = 0$, then the equilibrium probability for A is the maximal entropy measure.*

Proof. If for any n periodic orbit $x = \Phi^n(x) \in X$, we have that $S_n(A+c)(x) = 0$, then from Theorem 5.4 we get that there exists Lipschitz (Hölder) function g such that

$$A + c = g - g \circ \Phi.$$

This shows that A is Φ coboundary to a constant, and therefore the Φ -equilibrium probability for A is the maximal entropy measure. \square

Proposition 5.8. *If A is a Lipschitz (Hölder) potential such that for a certain n periodic orbit $x = \Phi^n(x)$, we have that $S_n(A)(x) = nc_1$, and for another m periodic orbit $x = \Phi^m(x)$, we have that $S_m(A)(x) = mc_2$, where $c_1 \neq c_2$, then the equilibrium probability for A is not the maximal entropy measure.*

Proof. A Lipschitz (Hölder) potential A such the Φ -equilibrium probability μ_A is the maximal entropy measure is of the form

$$A = g - g \circ \Phi + c,$$

where c is a constant.

Therefore, for any k periodic orbit $x = \Phi^k(x)$, we have that $S_k(A)(x) = kc$. In this way, under the above hypotheses, if $c_1 \neq c_2$, we reach a contradiction. \square

Theorem 5.9. *If the correspondence $i \rightarrow \phi(a, i)$ is bijective for any $a \in \{1, \dots, r\}$ then the map Φ is transitive, that is, there exists $y \in \Omega$ such that the orbit of y is dense in Ω .*

Proof. Consider the set of maps $\tau_j : \Omega \rightarrow \Omega$. We already know that under the hypothesis $i \rightarrow \phi(a, i)$ is bijective for any $a \in \{1, \dots, r\}$, is true that $\Omega = \bigcup_j \tau_j(\Omega)$. By extension, for any $k \geq 1$ we get

$$\Omega = \bigcup_{j_1 \dots j_k} \tau_{j_1} \circ \dots \circ \tau_{j_k}(\Omega)$$

a cover of Ω of diameter smaller than $\theta^k \rightarrow 0$ when $k \rightarrow \infty$. To simplify the notation we denote $\tau_{j_1 \dots j_k} := \tau_{j_1} \circ \dots \circ \tau_{j_k}$ and define recursively

$$Z_0 \in \Omega, Z_k := \tau_{r \dots r} \circ \dots \circ \tau_{1 \dots 21} \circ \tau_{1 \dots 11}(Z_{k-1})$$

where the composition is take over all k -uples $j_1 \dots j_k$.

We claim that for any $j_1 \dots j_k$ and any $x \in \tau_{j_1} \circ \dots \circ \tau_{j_k}(\Omega)$ and $w \in \Omega$ we get $d(x, \tau_{j_1 \dots j_k}(w)) < \theta^k$. Which is evident because each τ_j contracts by θ and $x \in \tau_{j_1} \circ \dots \circ \tau_{j_k}(\Omega)$ means that $x = \tau_{j_1} \circ \dots \circ \tau_{j_k}(w')$.

Consider the sequence $Z_k, k \geq 0$ and, by the compactness of Ω a point $y \in \Omega$ for which $Z_{k_i} \rightarrow y$ when $i \rightarrow \infty$.

We claim that the orbit of y by Φ is dense. To see that take any $x \in \Omega$ and $\varepsilon > 0$. Choose k_i big enough to ensure $\theta^{k_i} < \varepsilon/2$. Notice that, by continuity,

$$\Phi^{m k_i}(y) = \Phi^{m k_i}(\lim_{i \rightarrow \infty} Z_{k_i}) = \lim_{i \rightarrow \infty} \Phi^{m k_i}(Z_{k_i}).$$

Let $x \in \tau_{j_1}^0 \circ \dots \circ \tau_{j_{k_i}}^0(\Omega)$, as $\Phi \circ \tau_j = Id$ we can choose $0 \leq m \leq r^{k_i}$ in such way that

$$\Phi^{m k_i}(Z_{k_i}) = \tau_{j_1}^0 \circ \dots \circ \tau_{j_{k_i}}^0(w')$$

and, by the above result we get $d(\Phi^{m k_i}(Z_{k_i}), x) < \theta^{k_i} < \varepsilon/2$.

If we consider additionally k_i big enough to $d(\Phi^{m k_i}(y), \Phi^{m k_i}(Z_{k_i})) < \varepsilon/2$, we obtain

$$d(\Phi^{m k_i}(y), x) \leq d(\Phi^{m k_i}(y), \Phi^{m k_i}(Z_{k_i})) + d(\Phi^{m k_i}(Z_{k_i}), x) < \varepsilon,$$

proving the density. □

5.1 Periodic points

In this subsection, we describe an algorithmic procedure to compute the fixed points for $\Phi^m, m \in \mathbb{N}$. This characterization is helpful to be able to apply Livsic's Theorem.

We just assume that $i \rightarrow \phi(a, i)$ is bijective for any $a \in \{1, \dots, r\}$.

The inverse branches of Φ are, $\tau_j : \Omega \rightarrow \Omega$ are given by the formula

$$\tau_j(x) = y \text{ if } \Phi(y) = x \text{ and } y_1 = j$$

and are uniform contractions (by θ).

Theorem 5.10. *Suppose that $i \rightarrow \phi(a, i)$ is bijective for any $a \in \{1, \dots, r\}$. Then,*

1. Each periodic point x for Φ with period m is a solution of $\tau_{j_1} \circ \dots \circ \tau_{j_m}(x) = x$ for some choice of $j_1, \dots, j_m \in \{1, \dots, r\}$;
2. The periodic point, above defined, is the solution of the following recurrence

$$\left\{ \begin{array}{l} x_1 = j, z_1^1 = i_1, \dots, z_1^{m-1} = i_{m-1}, \\ x_{k+1}, z_{k+1}^j, 1 \leq j \leq m-1 \text{ solving} \end{array} \right\} \left\{ \begin{array}{l} \phi(x_k, x_{k+1}) = z_k^1 \\ \phi(z_k^1, z_{k+1}^1) = z_k^2 \\ \phi(z_k^2, z_{k+1}^2) = z_k^3 \\ \dots \\ \phi(z_k^{m-1}, z_{k+1}^{m-1}) = x_k \end{array} \right.$$

3. The map $(j, i_1, \dots, i_{m-1}) \rightarrow x$, defined as the solution of the above recurrence, is a bijection. In particular, Φ has exactly r^m periodic points of period m .

Proof. (1) It is evident that each composition $\tau_{j_1} \circ \dots \circ \tau_{j_m}$ has a unique fixed point x such that $\tau_{j_1} \circ \dots \circ \tau_{j_m}(x) = x$ because it is a contraction. Those are in fact periodic points of Φ , because $\Phi^m(x) = x$. Reciprocally, $\Phi^m(x) = x$ taking $\Phi(\Phi^m(x)) = x$ we can find $j_1 = x_1$ such that $\Phi^m(x) = \tau_{j_1}(x)$. Repeating this procedure we obtain $\tau_{j_1} \circ \dots \circ \tau_{j_m}(x) = x$.

(2) To avoid extremely complex notation we consider the case $m = 3$. The general case is obtained by the same reasoning. We must have

$$\tau_n \tau_i \tau_j(x) = x \text{ iff } \Phi(x) = \tau_i \tau_j(x) \text{ and } x_1 = n.$$

Let us to introduce new variables $z := \tau_i \tau_j(x)$ and $w := \tau_j(x)$ then

$$\tau_i(w) = z \text{ iff } \Phi(z) = w \text{ and } z_1 = i,$$

$$\tau_j(x) = w \text{ iff } \Phi(w) = x \text{ and } w_1 = j$$

and

$$\Phi(x) = z \text{ and } x_1 = n,$$

which is equivalent to the recursive system

$$\phi(i, z_2) = w_1 = j, \phi(z_2, z_3) = w_2, \phi(z_3, z_4) = w_3, \text{ etc.}$$

$$\phi(j, w_2) = x_1 = n, \phi(w_2, w_3) = x_2, \phi(w_3, w_4) = x_3, \text{ etc.}$$

$$\phi(n, x_2) = z_1 = i, \phi(x_2, x_3) = z_2, \phi(x_3, x_4) = z_3, \text{ etc.}$$

So we must solve the implicit recurrence

$$\left\{ \begin{array}{l} x_1 = n, z_1 = i, w_1 = j, \\ x_{k+1}, z_{k+1}, w_{k+1} \text{ solving} \end{array} \right\} \left\{ \begin{array}{l} \phi(x_k, x_{k+1}) = z_k \\ \phi(z_k, z_{k+1}) = w_k \\ \phi(w_k, w_{k+1}) = x_k \end{array} \right.$$

(3) By (2) we know that, in order to find the periodic points we must solve the previous recurrence with the following initial conditions

$$x_1 = j, z_1^1 = i_1, \dots, z_1^{m-1} = i_{m-1}.$$

As we have a finite set, is enough to show that the correspondence is injective. Suppose, by contradiction, that for a different set of initial conditions

$$x_1 = j, w_1^1 = h_1, \dots, w_1^{m-1} = h_{m-1}$$

we produce the same periodic point x . Using the respective equations we get for $k \geq 1$

$$\phi(x_k, x_{k+1}) = z_k^1 \text{ and } \phi(w_k^1, w_{k+1}^1) = w_k^1$$

thus $z_k^1 = w_k^1, k \geq 1$.

$$\phi(z_k^1, z_{k+1}^1) = z_k^2 \text{ and } \phi(w_k^1, w_{k+1}^1) = w_k^2$$

thus $z_k^2 = w_k^2, k \geq 1$.

And so on. In particular

$$z_1^1 = w_1^1, \dots, z_1^{m-1} = w_1^{m-1}$$

contradicting $(i_1, \dots, i_{m-1}) \neq (h_1, \dots, h_{m-1})$. □

Example 5.11. Let Φ be the map obtained by the local interaction

$$M := \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Fixed points ($m = 1$) are obtained by $x_1 = 1$ and x_{k+1} solving $\phi(x_k, x_{k+1}) = x_k$, that is,

$$\phi(1, x_2) = 1 \text{ then } x_2 := 2;$$

$$\phi(2, x_3) = 2 \text{ then } x_3 := 2;$$

$$\phi(2, x_4) = 2 \text{ then } x_4 := 2;$$

and so on, thus the first fixed point is $(1, 2, 2, 2, 2, 2, \dots)$. The second one is obtained by $x_1 = 2$ and x_{k+1} solving $\phi(x_k, x_{k+1}) = x_k$, that is,

$$\phi(2, x_2) = 2 \text{ then } x_2 := 2;$$

$$\phi(2, x_3) = 2 \text{ then } x_3 := 2;$$

$$\phi(2, x_4) = 2 \text{ then } x_4 := 2;$$

and so on, thus the first fixed point is $(2, 2, 2, 2, 2, 2, \dots)$.

For period 2 we have four points

a) $x_1 = 1, z_1 = 1$ and x_{k+1}, z_{k+1} solving $\phi(x_k, x_{k+1}) = z_k, \phi(z_k, z_{k+1}) = x_k$.

So,

$\phi(1, x_2) = 1, \phi(1, z_2) = 1$ then $x_2 = 2$ and $z_2 = 2$;

$\phi(2, x_3) = 2, \phi(2, z_3) = 2$ then $x_3 = 2$ and $z_3 = 2$;

and so on. Thus $(1, 2, 2, 2, 2, \dots)$ is our first point (it is also a fixed point).

b) $x_1 = 1, z_1 = 2$ and x_{k+1}, z_{k+1} solving $\phi(x_k, x_{k+1}) = z_k, \phi(z_k, z_{k+1}) = x_k$. So,

$\phi(1, x_2) = 2, \phi(2, z_2) = 1$ then $x_2 = 1$ and $z_2 = 1$;

$\phi(1, x_3) = 1, \phi(1, z_3) = 1$ then $x_3 = 2$ and $z_3 = 2$;

and so on. Thus $(1, 1, 2, 2, 2, \dots)$ is our second point.

The remaining points are $(2, 1, 2, 2, 2, \dots)$ and $(2, 2, 2, 2, 2, \dots)$.

Finally, we compute a periodic point of period 3. For that we choose $x_1 = 1, z_1 = 2, w_1 = 1$ and use the equations,

$\phi(x_k, x_{k+1}) = z_k, \phi(z_k, z_{k+1}) = w_k, \phi(w_k, w_{k+1}) = x_k$. So

$\phi(1, x_2) = 2, \phi(2, z_2) = 1, \phi(1, w_2) = 1$ then $x_2 = 1, z_2 = 1$ and $w_2 = 2$;

$\phi(1, x_3) = 1, \phi(1, z_3) = 2, \phi(2, w_3) = 1$ then $x_3 = 2, z_3 = 1$ and $w_3 = 1$;

$\phi(2, x_4) = 1, \phi(1, z_4) = 1, \phi(1, w_4) = 2$ then $x_4 = 1, z_4 = 2$ and $w_4 = 1$;

$\phi(1, x_5) = 2, \phi(2, z_5) = 1, \phi(1, w_5) = 1$ then $x_5 = 1, z_5 = 1$ and $w_5 = 2$;

and so on, thus $(1, 1, 2, 1, 1, 2, 1, 1, 2, 1, 1, 2, 1, 1, \dots)$ is a periodic point of period 3.

6 Measure rigidity review

The *measure rigidity* problem for a pair of maps, that is, the problem of showing the uniqueness of the measure (generally the Parry measure) that is invariant for some pair of maps, is related to an early result by Furstenberg who proved that for relative primes $p, q \in \mathbb{Z}$ the unique infinite set of the torus $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ which is invariant for the maps $T_p : \mathbb{T} \rightarrow \mathbb{T}$ and $T_q : \mathbb{T} \rightarrow \mathbb{T}$, given by $T_p(x) = px$ and $T_q(x) = qx$, is the whole torus itself [14]. Such result leads Furstenberg to conjecture that the Lebesgue probability measure is the unique continuous probability measure on the torus which is simultaneously invariant for the maps T_p and T_q (see [24]).

Since each point of a shift space is a sequence that carries explicitly all the information about its topological location in the space, it follows that there is a direct relationship between the emergence of patterns in a given point of the shift space under the action of some map and the trajectory of this point for that map. From the statistical point of view, by supposing $\Phi : \Omega \rightarrow \Omega$ is a cellular automaton and μ is a Φ -invariant probability measure on the Borelians of Ω , it means that the probability of a μ -randomly chosen point $x \in \Omega$ being such that $\Phi^n(x)$ has some pattern $w_1 w_2 \dots w_k$ on first k

positions is equal to $\mu(\overline{w_1 w_2 \dots w_k})$, for all $n \geq 0$.

If, additionally, μ is assumed to be also σ -invariant, then the probability that $\Phi^n(x)$ exhibits the pattern $w_1 w_2 \dots w_k$ at any fixed k consecutive positions, for a point $x \in \Omega$ randomly chosen according to μ , is equal to $\mu(\overline{w_1 w_2 \dots w_k})$. In other words, considering the double-indexed sequence $(\Phi^n(x)_i)_{i,n \geq 0}$, measures that are simultaneously invariant under σ and a cellular automaton Φ enable the identification of patterns that are spatially (in the index i) and temporally (in the index n) statistically invariant through the orbit of x for Φ .

An early result characterizing measures that are (σ, Φ) -invariant is due to Lind [23] who examined the scenario where \mathcal{A} is the group \mathbb{Z}_2 , $\Omega = \mathcal{A}^{\mathbb{Z}}$, $\Phi : \Omega \rightarrow \Omega$ has local rule $\phi(a, b) = a + b \pmod{2}$, and μ is any Bernoulli probability measure with full support on Ω . Under these conditions, Lind demonstrated that the Cesro mean distribution $\mathcal{C}_N^\Phi(\mu) = N^{-1} \sum_{n=0}^{N-1} \mu \circ \Phi^{-n}$ converges to the Haar measure, which here is the uniform Bernoulli probability measure λ . As consequence, since any initial Bernoulli measure is σ -invariant, and λ is (σ, Φ) -invariant, it follows that the uniform Bernoulli probability measure is the unique Bernoulli probability measure (σ, Φ) -invariant¹.

Latter, Schmidt [40] considered \mathcal{A} being any Abelian group, and extend the group operation from \mathcal{A} to $\Omega = \mathcal{A}^{\mathbb{Z}^d}$ as a component-wise operation. Among other important results provided by the author, one can use [40, Corollary 29.5, p. 289] to find out that if $\mathcal{A} = \mathbb{Z}_2$ and $\phi : \mathcal{A}^H \rightarrow \mathcal{A}$, the local rule of Φ , is defined with $H \subset \mathbb{Z}$ such that $|H| \geq 2$, then the unique (σ, Φ) -invariant probability measure with full support on $\Omega = \mathcal{A}^{\mathbb{Z}}$ which holds a certain mixing property (called H -mixing) is the uniform Bernoulli probability measure. This result can be extend for measures with full support on a subshift $G \subset \mathcal{A}^{\mathbb{Z}^d}$, where it is assumed that G is a subgroup of $\mathcal{A}^{\mathbb{Z}^d}$ and $\Phi(G) = G$. In such a case, the unique (σ, Φ) -invariant probability measure with full support on G is the Haar measure on G - see [37, Proposition 29]).

As Lind [23] and Schmidt [40], as the subsequent works that have addressed the problem of measure rigidity in cellular automata, have always considered an algebraic structure on the alphabet which induces an algebraic structure on $\mathcal{A}^{\mathbb{Z}^d}$, Ω being a subshift and a subgroup of $\mathcal{A}^{\mathbb{Z}^d}$, and cellular automata $\Phi : \Omega \rightarrow \Omega$ which are endomorphisms for that algebraic structure.

¹Note that given any σ -invariant measures μ , if the Cesro mean distributions $N^{-1} \sum_{n=0}^{N-1} \mu \circ \Phi^{-n}$ converges to some probability measure $\hat{\mu}$ in the weak* topology, then $\hat{\mu}$ is (σ, Φ) -invariant. Furthermore, it is interesting to notice that the uniform Bernoulli measures is invariant for a cellular automaton $\Phi : \mathcal{A}^{\mathbb{Z}^d} \rightarrow \mathcal{A}^{\mathbb{Z}^d}$ if, and only if, Φ is onto.

In particular, several works have considered the particular classes of cellular automata whose local rules have algebraic origin, that is, cellular automata $\Phi : \Omega \rightarrow \Omega$ with local rule $\phi : \mathcal{A}^H \rightarrow \mathcal{A}$ for some H finite subset of $\subset \mathbb{Z}^d$, such that for all $x = (x_i)_{i \in H} \in \mathcal{A}^H$ we have

$$\phi(x) = \sum_{i \in H} \eta_i(x_i) + c,$$

where: the sum is with respect to the operation considered in \mathcal{A} ; $\eta_i : \mathcal{A} \rightarrow \mathcal{A}$ is an endomorphism for each $i \in H$; $\eta_i \circ \eta_j = \eta_j \circ \eta_i$ for all $i, j \in H$; and c is a fixed symbol of \mathcal{A} . A cellular automaton in this form is said to be an *affine cellular automaton* and, in the particular case that $c = 0$ (the identity element of \mathcal{A}) it is said to be a *linear cellular automaton*, while if $\eta_i : \mathcal{A} \rightarrow \mathcal{A}$ is the identity map for all $i \in H$ and $c = 0$ it is said to be a *group cellular automaton*.

We remark that the algebraic origin of the cellular automata considered in all of these works is due to the need for a local rule for which one can explicitly compute successive iterates.

Following Linds seminal work [23], several papers have found out measure rigidity results (see Table 1). By assuming that \mathcal{A} is some group (with some specific features), Ω is some specific shift space on \mathcal{A} , μ is a measure on Ω (taken within some class of measures), and $\Phi : \Omega \rightarrow \Omega$ belongs to some specific class of endomorphic cellular automata, these works proved the convergence of the Cesaro mean distribution of μ under the dynamics of Φ .

On the other hand, following Schmidts approach and considering some conditions on the entropy and ergodicity, it was proved that the Haar measure (in this case the uniform Bernoulli measure) is the unique (σ, Φ) -invariant measure if $\Phi : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ is a bipermutative endomorphic cellular automaton with local rule $\phi : \mathcal{A}^H \rightarrow \mathcal{A}$ for some $H \subset \mathbb{Z}$, and: $\mathcal{A} = \mathbb{Z}/\mathbb{Z}_p$ for some p prime, $|H| = 2$, and Φ is an affine cellular automaton [17]; \mathcal{A} is any finite Abelian group, and $|H| = 2$ [33]; \mathcal{A} is any Abelian group, and $|H| > 2$ [39]. The Haar measure also was proved to be the unique (σ, Φ) -invariant measure exhibiting some entropic and ergodic properties for some classes endomorphic cellular automata $\Phi : \mathcal{A}^{\mathbb{Z}^d} \rightarrow \mathcal{A}^{\mathbb{Z}^d}$ on Abelian groups [10, 34].

It is worth noting that Furstenberg's conjecture on measure rigidity for maps on the torus was proved by Lyons in [24], four years after the result by Lind [23] for the cellular automaton, by considering the additional hypothesis that the measure is T_p exact. Furthermore, the author also found sufficient conditions under which a probability measure μ is such that $\nu \circ T_q^{-n}$ converges to the Lebesgue measure in the weak* topology. Still considering the case

of the torus, in [38] and [2] it was proved that the Lebesgue measure is the unique (T_p, T_q) -invariant measure in a wider class of measures. In [2], measure rigidity results were also proved for (T, S) -invariant measures where T and S are general differentiable expansive maps on the torus.

Shift space	Local rule type	Measure class	Proof	Ref.
\mathcal{A} is the group \mathbb{Z}_2 $\Omega = \mathcal{A}^{\mathbb{Z}}$	Group C. A.	Bernoulli measures	Harmonic analysis	[23]
		Markov measures		[12, 25]
\mathcal{A} is any finite Abelian group $\Omega = \mathcal{A}^{\mathbb{M}}$, \mathbb{M} any monoid	Diffusive linear C. A.	Harmonic mixing		[35, 36]
	Dispersive linear C. A.	Dispersion mixing		[36]
	Affine C. A. (right-permutative Ψ -associative and N -scaling)	Measures with complete connections and summable decay		[17]
\mathcal{A} is the group \mathbb{Z}_{p^s} $\Omega = \mathcal{A}^{\mathbb{Z}}$	Linear C. A.			[13]
$\Omega \subseteq (\bigoplus_{i=1}^s (\mathbb{Z}_{p^i})^{n_i})^{\mathbb{Z}^d}$		Markov measures	Renewal theory	[27, 28]
$\Omega \subseteq (\bigoplus_{i=1}^s (\mathbb{Z}_{p^i})^{n_i})^{\mathbb{Z}}$	Affine C. A.	Measures with complete connections and summable decay	Measure conjugacy	[29]
	Structurally-compatible C. A. (right-permutative Ψ -associative and N -scaling)			[41]

Table 1: Different settings where the measure rigidity results have been obtained through the Cesaro mean convergence of an initial σ -invariant measure. In the first column of the table we state the shift space and the group structure, in the second column the type of cellular automata, in the third column the class of initial measure, in the fourth column the method used to prove the convergence, and in the last column the paper where the result is proved.

7 Appendix

Proposition 7.1. *Let $\phi : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ the local rule we assumed in Section 2. Consider the map $\Phi : \Omega \rightarrow \Omega$ associated with this rule. Then, for any $x \in \Omega$ the set of preimages by Φ is dense.*

Proof. Let $\Lambda_k(x) := \{y \in \Omega \mid \Phi^k(y) = x\}$ be the set of preimages of x of order k and $\Lambda(x) := \bigcup_{k \geq 1} \Lambda_k(x)$ the set of all preimages of x .

We will use an IFS approach to simplify the reasoning. For each $j \in \mathcal{A}$ we define the inverse branch $\tau_j : \Omega \rightarrow \Omega$ by the formula

$$\tau_j(x) = y \text{ if } \Phi(y) = x \text{ and } y_1 = j.$$

In this way, the family τ_j , $j \in \mathcal{A}$ is a iterated function system.

Considering the fractal operator, $F(B) = \bigcup_{j \in \mathcal{A}} \tau_j(B)$, defined for the set of compact not empty parts of Ω , and the fact that each τ_j , $j \in \mathcal{A}$ is a Lipschitz contraction, there exists a unique compact set K such that $F(K) = K$ and $F^k(B) \rightarrow K$ for any B .

First, we observe that choosing $B = \{x\}$ we get $F^k(B) = \Lambda_k(x)$ thus $\Lambda(x)$ is dense in K .

To conclude our proof we claim that $K = \Omega$. To do that, we will show that $F(\Omega) = \Omega$ and use the uniqueness of such set K . As $F(\Omega) \subseteq \Omega$, by definition, we just need to show the opposite inclusion. Take $y \in \Omega$ and $j = y_1$ we need to find $x \in \Omega$ such that $\tau_j(x) = y$, in other words,

$$\phi(j, y_2) = x_1, \phi(y_2, y_3) = x_2, \phi(y_3, y_4) = x_3, \dots$$

which is always possible. Thus, $y \in F(\Omega)$ and, since y is arbitrary $F(\Omega) \supseteq \Omega$. So we have the equality. □

Proposition 7.2. *Consider $\mathcal{K} := \{\bar{1}, \bar{2}, \dots, \bar{r}\}$, a partition of Ω . Then, for all $n \geq 0$ we have*

$$\mathcal{K} \bigvee \Phi^{-1}(\mathcal{K}) \bigvee \dots \bigvee \Phi^{-n}(\mathcal{K}) = \bigsqcup_{a_1 \dots a_{n+1} \in \mathcal{A}} \overline{a_1 \dots a_{n+1}}.$$

Moreover, give n , $x \in \Omega$, for each $y = (a_1, a_2, \dots, a_n, y_{n+1}, y_{n+2}, \dots) \in \Phi^{-n}(x)$, there exist a unique z of the form $z = (a_1, a_2, \dots, a_n, z_{n+1}, z_{n+2}, \dots)$ in $\sigma^{-1}(x)$, and vice versa.

Proof. The proof is by induction w.r.t. n . The bases, $n = 0$ is the identity

$$\mathcal{K} = \bigsqcup_{a_1 \in \mathcal{A}} \overline{a_1}.$$

Now suppose, by hypothesis, that

$$\mathcal{K} \bigvee \Phi^{-1}(\mathcal{K}) \bigvee \dots \bigvee \Phi^{-n}(\mathcal{K}) = \bigsqcup_{a_1 \dots a_{n+1} \in \mathcal{A}} \overline{a_1 \dots a_{n+1}}.$$

We consider a generic cylinder $\overline{b_1 \dots b_{n+1} b_{n+2}}$ and by applying Φ we obtain

$$\Phi(\overline{b_1 \dots b_{n+1} b_{n+2}}) \subset \overline{\phi(b_1, b_2) \dots \phi(b_{n+1}, b_{n+2})}.$$

From our hypothesis of induction,

$$\overline{\phi(b_1, b_2) \dots \phi(b_{n+1}, b_{n+2})} \in \mathcal{K} \bigvee \Phi^{-1}(\mathcal{K}) \bigvee \dots \bigvee \Phi^{-n}(\mathcal{K}),$$

that is, $\overline{b_1 \dots b_{n+1} b_{n+2}} \in \Phi^{-1}(\mathcal{K}) \bigvee \dots \bigvee \Phi^{-n-1}(\mathcal{K})$.

Choosing $\bar{b}_1 \in \mathcal{K}$ we get, $\overline{b_1 \dots b_n b_{n+1}} \in \mathcal{K} \bigvee \Phi^{-1}(\mathcal{K}) \bigvee \dots \bigvee \Phi^{-n-1}(\mathcal{K})$. Thus $\mathcal{K} \bigvee \Phi^{-1}(\mathcal{K}) \bigvee \dots \bigvee \Phi^{-n-1}(\mathcal{K}) \supseteq \bigsqcup_{a_1 \dots a_{n+2} \in \mathcal{A}} \overline{a_1 \dots a_{n+2}}$.

To show the opposite relation we take any $B \in \mathcal{K} \bigvee \Phi^{-1}(\mathcal{K}) \bigvee \dots \bigvee \Phi^{-n-1}(\mathcal{K})$. We claim the B is a cylinder of length $(n+1)+1 = n+2$. By the induction hypothesis $\mathcal{K} \bigvee \Phi^{-1}(\mathcal{K}) \bigvee \dots \bigvee \Phi^{-n}(\mathcal{K}) = \bigsqcup_{a_1 \dots a_{n+1} \in \mathcal{A}} \overline{a_1 \dots a_{n+1}}$ so

$$B = \overline{a_1 \dots a_{n+1}} \cap \Phi^{-n-1}(\mathcal{K}).$$

Notice that

$$\Phi(y_1, y_2, y_3, \dots) = (\phi(y_1, y_2), \phi(y_2, y_3), \dots)$$

$$\Phi^2(y_1, y_2, y_3, \dots) = (\phi(\phi(y_1, y_2), \phi(y_2, y_3)), \phi(\phi(y_3, y_4), \phi(y_4, y_5)), \dots)$$

and so on. In this way $\Phi^m(y) = (\psi(y_1, \dots, y_{m+1}), \dots)$, where $\psi(y_1, \dots, y_{m+1})$ is a function of the first $m+1$ coordinates.

As $\mathcal{K} := \{\bar{1}, \bar{2}, \dots, \bar{r}\}$ and $\Phi^{n+1}(y) \in \mathcal{K}$ we must have \bar{j} such that $\Phi^{n+1}(y) \in \bar{j}$. At the same time $y_1 = a_1 \dots y_{n+1} = a_{n+1}$ because $y \in \overline{a_1 \dots a_{n+1}}$. In other words,

$$\psi(y_1, \dots, y_{(n+1)+1}) = j$$

$$\psi(a_1, \dots, a_{n+1}, y_{n+2}) = j.$$

Let a_{n+2} be the unique solution of the above equation. Thus,

$$B = \overline{a_1 \dots a_{n+1}} \cap \Phi^{-n-1}(\mathcal{K}) = \overline{a_1 \dots a_{n+1} a_{n+2}},$$

which concludes our proof. \square

Example 7.3. An example where equation (1.3) is true for $\Phi_1 = \sigma$ and $\Phi_2 = \Phi$, $r = 2$, for functions that depend on the first three coordinates is the following. We will have many more choices of possible functions A_1 and A_2 when compared with the last example.

Consider ϕ such that

$$M = \begin{pmatrix} \phi(1, 1) & \phi(1, 2) \\ \phi(2, 1) & \phi(2, 2) \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}.$$

Take the functions A_1 and A_2 depending on the three first coordinates satisfying

$$A_1(x_1, x_2, x_3, \dots) = Q_{x_1, x_2, x_3}$$

and

$$A_2(x_1, x_2, x_3, \dots) = C_{x_1, x_2, x_3}.$$

Then,

$$A_1 - A_1 \circ \sigma = A_2 - A_2 \circ \Phi,$$

means

$$Q_{x_1, x_2, x_3} - Q_{x_2, x_3, x_4} = C_{x_1, x_2, x_3} - C_{\phi(x_1, x_2), \phi(x_2, x_3), \phi(x_3, x_4)}.$$

We get a solution for the system if we assume that

- $C_{1,1,1} = C_{2,2,2}$
- $C_{1,1,2} = Q_{2,2,1} - Q_{2,2,2} + C_{2,2,2}$
- $C_{1,2,1} = Q_{2,1,2} - Q_{2,1,1} + Q_{1,2,2} - Q_{2,2,2} + C_{2,2,2}$
- $C_{1,2,2} = Q_{1,2,2} - Q_{2,2,2} + C_{2,2,2}$
- $C_{2,1,1} = Q_{2,2,1} - Q_{1,1,2} + Q_{1,2,2} - Q_{2,2,2} + C_{2,2,2}$
- $C_{2,1,2} = Q_{2,1,2} + Q_{2,2,1} - Q_{1,1,2} - Q_{2,2,2} + C_{2,2,2}$
- $C_{2,2,1} = Q_{2,2,1} - Q_{2,1,1} + Q_{1,2,2} - Q_{2,2,2} + C_{2,2,2}$
- $C_{2,2,2} = \text{free}$
- $Q_{1,1,1} = -Q_{2,2,1} + Q_{2,1,1} + Q_{1,1,2} - Q_{1,2,2} + Q_{2,2,2}$
- $Q_{1,1,2} = \text{free}$
- $Q_{1,2,1} = \text{free}$

- $Q_{1,2,2} = \text{free}$
- $Q_{2,1,1} = \text{free}$
- $Q_{2,1,2} = \text{free}$
- $Q_{2,2,1} = \text{free}$

$$\bullet Q_{2,2,2} = \text{free}. \quad (7.1)$$

Therefore, under the above hypotheses the σ -equilibrium for A_2 is equal to the Φ -equilibrium for A_1 .

We get above a system of 16 linear equations, therefore, given any choice of the 8 parameters $C_{i,j,k}$ (defining A_2) we can get values $Q_{i,j,k}$ (defining A_1) satisfying the system. The bottom line is that any σ -equilibrium probability for a potential A_2 , depending on the first three coordinates, can be realized as a Φ -invariant probability.

A particular solution of the above system would be

$$\begin{aligned} C_{1,1,1} &= \frac{1}{2}, C_{1,1,2} = -\frac{1}{2}, C_{1,2,1} = -\frac{1}{2}, C_{1,2,2} = -\frac{1}{2}, C_{2,1,1} = -\frac{1}{2}, C_{2,1,2} = -\frac{1}{2}, \\ C_{2,2,1} &= -\frac{1}{2}, C_{2,2,2} = \frac{1}{2}, Q_{1,1,1} = 1, Q_{1,1,2} = 0, Q_{1,2,1} = 0, Q_{1,2,2} = 0, \\ Q_{2,1,1} &= 0, Q_{2,1,2} = 0, Q_{2,2,1} = 0, Q_{2,2,2} = 1. \end{aligned}$$

The above potential is not a potential that depends on the first two coordinates.

References

- [1] M. F. Barnsley and S. Demko, Iterated function systems and the global construction of fractals. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 399 (1817): 243–275 (1985)
- [2] M. CARVALHO, Contributions to a rigidity conjecture. *Acta Applicandae Mathematicae* **53**, 265–295. (1998)
- [3] M. F. Barnsley, S. G. Demko, J. H. Elton, and J. S. Geronimo, Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities. *Annales de l’IHP Probabilités et statistiques*, volume 24, pages 367–394 (1988)
- [4] F. Blanchard and A. Maass, Dynamical Properties of Expansive One-Sided Cellular Automata, *Israel Journal of Mathematics*, 99, 149–174 (1997)

- [5] R. Bowen, Equilibrium states and the ergodic theory of Anosov diffeomorphisms. Lecture Notes in Mathematics 470, Springer (1975)
- [6] J. E. Brasil, E. R. Oliveira and R. R. Souza, Thermodynamic Formalism for General Iterated Function Systems with Measures. *Qual. Theory Dyn. Syst.* 22, 19 (2023).
- [7] T. Ceccherini-Silberstein and M. Coornaert, Cellular automata and groups. Springer Verlag (2010)
- [8] L. Cioletti, M. Denker, A. O. Lopes and M. Stadlbauer, Spectral Properties of the Ruelle Operator for Product Type Potentials on Shift Spaces, *Journal of the London Mathematical Society - Volume 95, Issue 2*, 684–704 (2017)
- [9] G. Contreras, A. O. Lopes and Ph. Thieullen. Lyapunov minimizing measures for expanding maps of the circle, *Ergodic Theory and Dynamical Systems* Vol 21, 1379-1409 (2001)
- [10] M. EINSIEDLER, Isomorphism and measure rigidity for algebraic actions on zero-dimensional groups. *Monatsh. Math.* **144**, 1, 39–69 (2005)
- [11] A. H. Fan and Ka-Sing Lau. Iterated function system and Ruelle operator. *J. Math. Anal. Appl.* 231(2):319–344 (1999)
- [12] P. A. Ferrari, A. Maass, and S. Martínez, Cesàro mean distribution of group automata starting from Markov measures. Preprint (1999).
- [13] P. A. Ferrari, A. Maass, S. Martínez and P. Ney, Cesàro Mean Distribution of Group Automata Starting from Measures with Summable Decay, *Ergodic Theory and Dynamical Systems*, 20, 6, 1657–1670 (2000)
- [14] H. Furstenberg, Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1**, 1–49 (1967)
- [15] J. Hawkins, *The Mathematics of Cellular Automata*, AMS (2024)
- [16] F. V. Haeseler, H.-D. Peitgen, G. Skordev, Linear Cellular Automata, Substitutions, Hierarchical Iterated Function Systems and Attractors *Function Systems and Attractors, in Fractal Geometry and Computer Graphics*, pp 3-23 (1992)
- [17] B. Host, A. Maass and S. Martínez, Uniform Bernoulli Measure in Dynamics of Permutative Cellular Automata with Algebraic Local Rules, *Discrete and Continuous Dynamical Systems*, 9, 6, 1423–1446 (2003)
- [18] J. E. Hutchinson. Fractals and self similarity, *Indiana University Mathematics Journal*, 30(5):713–747 (1981).
- [19] A. Katok and B. Hasselblatt, Introduction to the modern theory of dynamical systems. Cambridge University Press, Cambridge (1995)
- [20] Y-O Kim and J. Lee, On the Gibbs measures of commuting one-sided subshifts of finite type, *Osaka J. Math.* 37, 175-18 (2000)

- [21] A. O. Lopes and E. R. Oliveira, Entropy and variational principles for holonomic probabilities of IFS, *Disc. and Cont. Dyn. Systems, (Series A)* Vol 23, N 3, 937–955 (2009)
- [22] A. O. Lopes, Thermodynamic Formalism, Maximizing Probabilities and Large Deviations, notes on line
- [23] D. A. Lind, Applications of Ergodic Theory and Sofic Systems to Cellular Automata, *Physica D*, 10, 1-2, 36-44 Cellular automata (1984).
- [24] R. Lyons, On measures simultaneously 2- and 3-invariant. *Israel Journal of Mathematics* **61**, 2, 219–233 (1988)
- [25] A. Maass, and Martínez, On Cesàro limit distribution of a class of permutative cellular automata. *J. Statist. Phys.* **90**, 1-2, 435–452 (1998)
- [26] A. Maass, Some Dynamical Properties of One-Dimensional Cellular Automata, in *Dynamics of Complex Interacting Systems (Santiago, 1994)*. *Nonlinear Phenom. Complex Systems*. Vol. 2, Dordrecht: Kluwer Academic Publishers, pp. 35–80 (1996)
- [27] A. Maass, S. Martínez, M. Pivato and R. Yassawi, Asymptotic randomization of subgroup shifts by linear cellular automata. *Ergodic Theory and Dynamical Systems* **26**, 4, 1203–1224 (2006)
- [28] A. Maass, S. Martínez, M. Pivato and R. Yassawi, Attractiveness of the Haar measure for linear cellular automata on Markov subgroups. *IMS Lecture Notes - Monograph Series, Dynamics and Stochastics* **48**, 100–108. *Festschrift in honour of Michael Keane*. (2006)
- [29] A. Maass, S. Martinez, M. Sobotka, Limit measures for affine cellular automata on topological Markov subgroups, *Nonlinearity*, Volume 19, Number 9 (2006)
- [30] J. Milnor, On the Entropy Geometry of Cellular Automata, *Complex Systems*, 2, 3, 357–385 (1988)
- [31] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque* Vol 187-188 (1990)
- [32] M. Pivato, The ergodic theory of cellular automata, *International Journal of General Systems* Vol. 41, No. 6, 583–594 (2012)
- [33] M. Pivato, Invariant Measures for Bipermutative Cellular Automata, *Discrete and Continuous Dynamical Systems*, 12, 4, 723–736 (2005)
- [34] M. Pivato, Module Shifts and Measure Rigidity in Linear Cellular Automata, *Ergodic Theory and Dynamical Systems*, 28, 6, 1945–1958 (2008)
- [35] M. Pivato and R. Yassawi, Limit measures for affine cellular automata. *Ergodic Theory and Dynamical Systems* **22**, 4, 1269–1287 (2002)

- [36] M. Pivato and R. Yassawi, Limit measures for affine cellular automata II. *Ergodic Theory and Dynamical Systems* **24**, 6, 1961–1980 (2004)
- [37] M. Pivato, Ergodic Theory of Cellular Automata. In Meyers, R. (eds) *Encyclopedia of Complexity and Systems Science*. Springer, New York, NY-634-669 (2009)
- [38] D. J. Rudolph, $\times 2$ and $\times 3$ invariant measures and entropy. *Ergod. Th. & Dynam. Sys.* **10**, 395–406 (1990)
- [39] M. Sablik, Measure rigidity for algebraic bipermutative cellular automata. *Ergodic Theory and Dynamical Systems* **27**, 6, 1965–1990 (2007)
- [40] K. Schmidt, (1995). *Dynamical systems of algebraic origin*. Progress in Mathematics, Vol. **128**. Birkhäuser Verlag, Basel (1995)
- [41] M. Sobottka, Right-permutative cellular automata on topological Markov chains. *Disc. and Cont. Dynamic. Systems*, **20**, 4, 1095–1110 (2008)
- [42] M. Sobottka, Teoria Ergódica para Autômatos Celulares Algébricos, XXVII Coloq. Bras. Mat. - IMPA (2009)
- [43] F. Spitzer. A Variational characterization of finite Markov chains. *The Annals of Mathematical Statistics*. (43): N.1 303-307 (1972)
- [44] A. Walker, Entropy and Applications of Cellular Automata, Internet Archive (2023)
- [45] Y-L.Ye, Ruelle operator with weakly contractive iteration function systems, *Erg. Theo and Dyn. Syst.* 33 - 1265-1290 (2013)