

ON MEASURE-PRESERVING C^1 TRANSFORMATIONS
OF COMPACT-OPEN SUBSETS
OF NON-ARCHIMEDEAN LOCAL FIELDS

JAMES KINGSBERY, ALEX LEVIN, ANATOLY PREYGEL, AND CESAR E. SILVA

ABSTRACT. We introduce the notion of a *locally scaling* transformation defined on a compact-open subset of a non-archimedean local field. We show that this class encompasses the Haar measure-preserving transformations defined by C^1 (in particular, polynomial) maps, and prove a structure theorem for locally scaling transformations. We use the theory of polynomial approximation on compact-open subsets of non-archimedean local fields to demonstrate the existence of ergodic Markov, and mixing Markov transformations defined by such polynomial maps. We also give simple sufficient conditions on the Mahler expansion of a continuous map $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$ for it to define a Bernoulli transformation.

1. INTRODUCTION

The p -adic numbers have arisen in a natural way in the study of some dynamical systems; for example in the study of group automorphisms of solenoids in Lind and Schmidt [LS94]. Other situations in dynamics where the p -adic numbers come up are surveyed in Ward [War06]. At the same time there has been interest in the study of the dynamics (topological, complex, or measurable) of naturally arising maps (such as polynomials) defined on the p -adics; see for example Benedetto [Ben01], Khrennikov and Nilson [KN04], and Rivera-Letelier [RL03]. In particular, Bryk and Silva in [BS05] studied the measurable dynamics of simple polynomials on balls and spheres on the field \mathbb{Q}_p of p -adic numbers. The maps they studied are ergodic but not totally ergodic, and they asked whether there exist polynomials on \mathbb{Q}_p that define (Haar) measure-preserving transformations that are mixing. Woodcock and Smart in [WS98] show that the polynomial map $x \mapsto \frac{x^p - x}{p}$ defines a Bernoulli, hence mixing, transformation on \mathbb{Z}_p . A consequence of our work is a significant extension of the result for this map, placing it in a greater context (see in particular Example 8.6).

Rather than working on \mathbb{Q}_p we find that the natural setting for our work is over a non-archimedean local field K . We introduce a class of transformations, called locally scaling, and show in Lemma 4.4 that measure-preserving C^1 (in particular, polynomial) maps are locally scaling. In Section 5 we apply the theory of Markov

Received by the editors September 1, 2006.

2000 *Mathematics Subject Classification*. Primary 37A05; Secondary 37F10.

Key words and phrases. Measure-preserving, ergodic, non-archimedean local field.

©2008 American Mathematical Society
Reverts to public domain 28 years from publication

shifts to classify the dynamics of locally scaling transformations, decomposing the transformation into a disjoint union of ergodic Markov transformations and local isometries. We show that if a \mathcal{C}^1 transformation T is measure-preserving when restricted to a compact-open set X , then X can be written as a finite disjoint union of invariant compact-opens sets such that T restricted to each such set is either a local isometry or is topologically and measurably conjugate to an ergodic Markov transformation. In particular, we show that for a measure-preserving locally scaling transformation on a compact-open set, if it is weakly mixing, then it must be mixing, and that there are no invertible weakly mixing measure-preserving \mathcal{C}^1 transformations on compact-open sets, answering a question in [BS05]. We also show the existence of polynomials defining transformations exhibiting nearly the full range of behaviors possible for locally scaling transformations, such as ergodic Markov, mixing Markov, and Bernoulli transformations.

Given a polynomial defined on a compact-open subset of K , our work shows that a finite computation may check whether it defines a measure-preserving transformation and whether it defines a mixing transformation; the question of ergodicity is also answered, except in the case where the polynomial is 1-Lipschitz, which has been studied by Anashin in [Ana02]. Works that study the dynamics of isometric analytic maps include [AV94], [HY83], and [Lin04].

We briefly mention related works studying measurable dynamics of certain maps on spaces related to the p -adics. The works [RB], [FRL04], and [FRL06] construct a natural invariant measure for a wide-class of rational functions, as in existing constructions in complex dynamics. The natural domain for these constructions is the so-called Berkovich projective space, a space much larger than the ordinary p -adics.

We now indicate an outline of the rest of the paper. Section 2 reviews results on Markov shifts, Section 3 reviews preliminaries on non-archimedean local fields as well as some analytic definitions, and Section 6 recalls some of the theory of polynomial approximation on rings of integers of non-archimedean local fields.

Section 4 establishes the fact that measure-preserving \mathcal{C}^1 maps are locally scaling, and Section 5 proves our main structural results, in particular Proposition 5.7 and Theorem 5.8. Section 7, in particular Theorem 7.2, shows that polynomial maps are in a sense a representative class of locally scaling transformations, and demonstrates the existence of polynomial maps defining locally scaling transformation with various behaviors, including mixing. Section 8 and Section 9 are devoted to demonstrating two interesting classes of locally scaling maps on \mathbb{Z}_p that arise naturally in the study of polynomial approximations. Specifically, Section 8 studies maps which are isometrically conjugate to the natural realization of the (one-sided) Bernoulli shift, and shows for instance that the map $x \mapsto \binom{x}{p^e}$ on \mathbb{Z}_p is Bernoulli. Section 9 then studies similar binomial-coefficient maps which are locally scaling and so have very regular structures but fail to be Haar measure-preserving.

2. MARKOV SHIFTS

Let H be a finite non-empty set. By a *stochastic matrix on H* we mean a map $A : H^2 \rightarrow \mathbb{R}_{\geq 0}$ such that

$$\sum_{j \in H} A(i, j) = 1 \quad \text{for each } i \in H.$$

Putting H into a bijection with the set $\{0, \dots, \#H - 1\}$ we may regard A as a $\#H \times \#H$ matrix with non-negative entries, and the entries in each row summing to 1. In analogy with this case, we will refer to the sets $\{A(i, \cdot)\}$ and $\{A(\cdot, j)\}$ as *rows* and *columns* of A , respectively.

By a *row vector on H* we mean a map $w : H \rightarrow \mathbb{R}$. For w a row vector and A a stochastic matrix, we define their product as the row vector wA defined by

$$wA(j) = \sum_{i \in H} w(i)A(i, j).$$

We will say that w is non-negative (respectively, positive) if it takes values in $\mathbb{R}_{\geq 0}$ (respectively, $\mathbb{R}_{> 0}$).

To any stochastic matrix A we may associate the following symbolic dynamical system:

- (i) Let

$$X_A = \{x \in \prod_{i \geq 0} H : A(\pi_n(x), \pi_{n+1}(x)) \neq 0 \text{ for all } n \geq 0\},$$

where $\pi_n : \prod_{i \geq 0} H \rightarrow H$ is a projection to the n^{th} coordinate. Give each finite factor the discrete topology, and X_A the subspace topology inherited from the product topology.

- (ii) Let $T_A : X_A \rightarrow X_A$ be defined by $\pi_n \circ T_A = \pi_{n+1}$; that is, T_A is simply “shifting left.” Then (X_A, T_A) is a topological dynamical system.
- (iii) If in addition we are given a non-negative row vector w , then we may define a measure on X_A by

$$\mu_{A,w}([d_0 d_1 d_2 \dots d_\ell]) = w(d_0)A(d_0, d_1) \cdots A(d_{\ell-1}, d_\ell),$$

$$\text{where } [d_0 \dots d_\ell] \stackrel{\text{def}}{=} \bigcap_{n=0}^{\ell} \pi_n^{-1}(d_n).$$

We call a set of the form $[d_0 \dots d_\ell]$ a *cylinder set*; we may observe that the cylinder sets form a base for the topology on X_A . Note that if w is in fact positive, then $\mu_{A,w}$ assigns positive measure to each cylinder set and hence to each open set. We may check that if $w = wA$, then T_A is measure-preserving with respect to $\mu_{A,w}$.

We call such a dynamical system a *Markov shift*. We say that a dynamical system is *Markov* if it is isomorphic to some Markov shift.

We say that a stochastic matrix A is *irreducible* or *ergodic* if for each $i, j \in H$ there exists an $n \in \mathbb{N}$ such that $A^n(i, j) > 0$. This condition has a natural interpretation in terms of the connectedness of a certain directed graph associated with A , as we shall see in the proof of Proposition 2.1. We say that a stochastic matrix A is *primitive* if there exists an $n \in \mathbb{N}$ such that $A^n(i, j) > 0$ for all $i, j \in H$. For ergodic theory preliminaries refer to [Sil08] and [Wal82].

Using the Perron-Frobenius Theorem on non-negative irreducible and primitive matrices, along with a graph theoretic interpretation of the stochastic matrix, one may obtain an ergodic decomposition result for Markov shifts:

Proposition 2.1. *Let A be a stochastic matrix, and w a positive row vector such that $w = wA$.*

Then we may partition H into disjoint sets

$$H = \bigsqcup_{k=1}^n H_k$$

such that

- (i) $A(i, j) = 0$ for $i \in H_k, j \in H_\ell$ with $k \neq \ell$; and
- (ii) $A_k = A|_{H_k \times H_k}$ is irreducible for $k = 1, \dots, n$.

Then $w_k = w|_{H_k}$ satisfies $w_k = w_k A_k$. Also, we have the ergodic decomposition of (X_A, μ_A, w, T_A) as

$$(X_A, \mu_A, w, T_A) = \bigsqcup_{k=1}^n (X_{A_k}, \mu_{A_k, w_k}, T_{A_k}),$$

where X_{A_k} is viewed as a T_A -invariant subset of X_A , so that $\mu_{A_k, w_k} = \mu_A|_{A_k}$ and $T_{A_k} = T_A|_{A_k}$.

Moreover, the k^{th} summand is mixing if A_k is primitive.

Proof. Construct a graph on H as follows. We place a directed edge from $i \rightarrow j$ if and only if $A(i, j) > 0$. As w is strictly positive, this is equivalent to the condition that $w(i)A(i, j) > 0$. We say that the flow or flux associated to this edge is $w(i)A(i, j)$. Now, the flow out of i is

$$\sum_{j \in H} w(i)A(i, j) = w(i),$$

as A is a stochastic matrix. The flow into i is

$$\sum_{k \in H} w(k)A(k, i) = w(i),$$

as $w = wA$. So, we see that the flux into and out of i are both equal to $w(i)$.

This implies that for every finite subset of H , the in-flux and out-flux will be equal. For $i \in H$, let $R(i)$ be the set of points reachable from i , and $B(i)$ the set of points which can reach i . Note that $R(i)$ has out-flux 0 by construction and $B(i)$ has in-flux 0 by construction; as H is finite, these subsets are finite, so both have in-flux and out-flux equal to 0.

Now, we can have no edges into or out of either of these two sets. But, if $t \in R(i)$ and $y \in B(i)$, then there is a path from y to t ; so we must have $t \in B(i)$ and $y \in R(i)$, and so $B(i) = R(i)$. So, $B(i) = R(i)$ is strongly connected, and there are no edges into or out of this set.

For $\ell > 0$, note that $A^\ell(i, j) > 0$ is equivalent to there being a path of length precisely ℓ from i to j . It follows that the collection

$$\{B(i) : i \in H\}$$

gives our desired decomposition of H .

We readily note that $w_k = w_k A_k$ for for $k = 1, \dots, n$. Then, as A_k is irreducible, [Wal82, Theorem 1.19] implies that the k^{th} summand is ergodic, from which the ergodic decomposition follows. Finally, [Wal82, Theorem 1.31] implies that the k^{th} summand is mixing if A_k is primitive. \square

3. ANALYTIC DEFINITIONS, PRELIMINARIES, AND NOTATION

Let K be a *non-archimedean local field*, which we take to be either a finite field extension of \mathbb{Q}_p or $\mathbb{F}_{p^n}((t))$ for some prime p .

Let $|\cdot|$ be a non-archimedean multiplicative valuation (sometimes called a “non-archimedean absolute value”) on K , such that $|\cdot|$ generates the topology on K . Denote $\mathcal{V} = |K^\times| = \{|x| : x \in K^\times\}$, $\mathcal{O} = \{x \in K : |x| \leq 1\}$ and $\mathfrak{p} = \{x \in K : |x| < 1\}$ (note that $\mathcal{O}, \mathfrak{p}$ are independent of the choice of valuation). It is the case that \mathcal{O} is a ring with maximal ideal \mathfrak{p} and that \mathcal{O}/\mathfrak{p} is a finite field (the *residue field*). Let $p = \text{char } \mathcal{O}/\mathfrak{p}$, $q = \#\mathcal{O}/\mathfrak{p}$, both finite with q a power of p .

We denote

$$B_r(x) = \{y \in K : |x - y| \leq r\},$$

and call such a set (for any value of r) a *ball*. A ball of radius precisely r will be called an *r -ball*. Let μ be Haar measure on K , normalized such that $\mu(\mathcal{O}) = 1$; define $\rho : \mathcal{V} \rightarrow \mathbb{R}_{>0}$ by $\rho(r) = \mu(B_r(0))$.

Now, we recall the following standard results:

- (i) \mathcal{O} is a discrete valuation ring with unique maximal ideal \mathfrak{p} .
- (ii) $\mathfrak{p} = \pi\mathcal{O}$ for any $\pi \in \mathfrak{p} \setminus \mathfrak{p}^2$; we call any such π a *uniformizing parameter*.
- (iii) \mathcal{V} is the discrete abelian (multiplicative) subgroup of \mathbb{Q} generated by $|\pi|$. In light of this, we may define a map $v : K \rightarrow \mathbb{Z} \cup \{+\infty\}$ defined by $v(0) = +\infty$ and $v(x) = \log_{|\pi|} |x|$ for $x \in K^\times$; this is the additive valuation (sometimes just “valuation”) on K .
- (iv) For $r = |\pi|^k$, $k \geq 0$, it is the case that

$$\rho(r) = \mu(B_r(0)) = (\#\mathcal{O}/\mathfrak{p}^k)^{-1} = q^{-k}.$$

Indeed, for $r \in \mathcal{V}$ we see that $\rho(r) = q^{-\log_{|\pi|} r}$.

- (v) A subset $X \subset K$ is compact-open if and only if X is a finite union of balls.

We direct the interested reader to [Ser62] for a thorough treatment of related topics.

We will continue to use the symbols $K, \mu, p, q, |\cdot|, \mathcal{O}, \mathfrak{p}, \pi, v, \mathcal{V}, \rho, B_r$ with these meanings below.

Let X be an open subset of K and $a \in X$. Then we say that a function $f : X \rightarrow K$ is *strictly differentiable* or C^1 at a (denoted $f \in C^1(a)$) if the limit

$$\lim_{\substack{(x,y) \rightarrow (a,a) \\ x \neq y}} \frac{f(x) - f(y)}{x - y}$$

exists. We write $f \in C^1(X)$ if $f \in C^1(a)$ for each $a \in X$. For more on this notion, see [Sch84] or [Rob00].

4. MEASURE-PRESERVING C^1 MAPS ON NON-ARCHIMEDEAN LOCAL FIELDS

Definition 4.1. For $X \subset K$ compact-open, we say that a transformation $T : X \rightarrow X$ is *locally scaling* for $r \in \mathcal{V}$ if X is a finite union of r -balls and if there exists a function $C : X \rightarrow \mathbb{R}_{\geq 1}$ such that

$$|x - y| \leq r \Rightarrow |T(x) - T(y)| = C(x)|x - y| \quad \text{for } x, y \in X.$$

We will refer to C as the *scaling function*.

Remark 4.2. Let us make the following observations about locally scaling transformations:

- (i) By the symmetry of x and y in the previous displayed equation, C is constant on cosets of $B_r(0)$. We will write $H = X/B_r(0)$ for the set of cosets of $B_r(0)$ contained in X (recall that X is a union of such cosets); we treat elements of H as subsets of X . Then C induces a map $C : H \rightarrow \mathbb{R}_{\geq 1}$.
- (ii) The terminology “locally scaling” is convenient but perhaps slightly misleading. For us, such transformations must not only locally scale distances, but must do so by a factor that is at least 1.

This definition is motivated by the ease of analyzing the structure of such maps together with the following easy lemma:

Lemma 4.3. *Let $X \subset K$ be open, and suppose $f \in \mathcal{C}^1(X)$ is such that $f'(x) \neq 0$ for $x \in X$. Then $|f'|$ is locally constant on X . If, moreover, X is compact, $f(X) \subset X$, and $|f'(a)| \geq 1$ for all $a \in X$, then the transformation $f : X \rightarrow X$ is locally scaling for some $r \in \mathcal{V}$.*

Proof. Fix $a \in X$. Since X is open and $f \in \mathcal{C}^1(a)$, there exists $r_a \in \mathcal{V}$ so that $B_{r_a}(a) \subset X$ and

$$\left| \frac{f(x) - f(y)}{x - y} - f'(a) \right| < |f'(a)| \quad \text{for } x, y \in B_{r_a}(a), x \neq y.$$

By the strong triangle inequality, it follows that $|f(x) - f(y)| = |f'(a)||x - y|$ for all $x, y \in B_{r_a}(a)$. In particular, we note that $|f'(x)|$ is constant on $B_{r_a}(a)$.

If X is compact, then there is a finite set $a_0, \dots, a_k \in X$ so that $B_{r_{a_0}}(a_0), \dots, B_{r_{a_k}}(a_k)$ is an open cover of X . Taking $r \leq \min_{i=0}^k r_{a_i}$, we see that X is a union of r -balls and that f restricted to each $B_r(a)$ scales distance by $|f'(a)| \geq 1$ by the previous equation. So, f is locally scaling for $r \in \mathcal{V}$ with scaling function $C(x) = |f'(x)|$. \square

Lemma 4.4. *Let $X \subset K$ be open. Let $f \in \mathcal{C}^1(X)$ be such that $f(X) \subset X$ and such that the transformation $f : X \rightarrow X$ is measure-preserving with respect to $\mu|_X$. Then $|f'(a)| \geq 1$ for all $a \in X$.*

Suppose further that X is compact. Then f is locally scaling for some $r \in \mathcal{V}$. If, in addition, f is invertible, then it is a local isometry and not weakly mixing.

Proof. Suppose there is an $a \in X$ with $|f'(a)| < 1$. Take $\alpha \in \mathcal{V}$ such that $|f'(a)| \leq \alpha < 1$. As X is open, there exists $r' \in \mathcal{V}$ such that $B_{r'}(a) \subset X$ and $B_{r'}(f(a)) \subset X$. Moreover, as $f \in \mathcal{C}^1(a)$, we may take $r \in \mathcal{V}$, with $r \leq r'$, such that

$$|f(x) - f(y) - f'(a)(x - y)| \leq \alpha |x - y| \quad \text{for any } x, y \in B_r(a).$$

It follows by the strong triangle inequality that $|f(x) - f(y)| \leq |x - y|\alpha$ for $x, y \in B_r(a)$. So, $B_{\alpha r}(f(a)) \subset X$ by construction, and moreover $f^{-1}(B_{\alpha r}(f(a))) \supseteq B_r(a)$. Taking measures we note that

$$\mu(B_{\alpha r}(f(a))) = \rho(\alpha r) < \rho(r) = \mu(B_r(a)) \leq \mu(f^{-1}(B_{\alpha r}(f(a)))) .$$

So, f is not measure-preserving.

This proves that $|f'(a)| \geq 1$ for all $a \in X$. If X is compact, then Lemma 4.3 applies.

If f is invertible and locally scales distances by a constant ≥ 1 , then that constant must be 1. So, f is a locally isometry. Taking $r \in \mathcal{V}$ small enough so that X is a

union or r -balls and f is locally scaling for r , we note that f descends to a quotient transformation $f : X/B_r(0) \rightarrow X/B_r(0)$. As f has a finite quotient, it cannot be weakly mixing. \square

Remark 4.5. Note that for $f \in K[x]$, $f \in C^1(K)$. So, if $f(X) \subset X$ induces a measure-preserving transformation $f : X \rightarrow X$, then Lemma 4.4 and Lemma 4.3 imply that f is locally scaling.

Example 4.6. Consider the map $f : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$ defined by

$$f(x) = \begin{pmatrix} x \\ 2 \end{pmatrix} = \frac{x(x-1)}{2}.$$

Then

$$|f(x) - f(y)| = \left| \frac{(x-y)(x+y-1)}{2} \right| = 2|x-y||x+y-1|.$$

So, f is locally scaling for $r = 1/2$, since $|x+y-1| = 1$ when $|x-y| \leq 1/2$. We will see in Section 8 that $f : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$ is actually measure-preserving and in fact Bernoulli.

Remark 4.7. Suppose $f \in K[x]$ defines a transformation $f : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$. Since f is polynomial, we have $f \in C^1(\mathbb{Z}_p)$ and thus, by Lemma 4.3, is locally scaling for some $r \in \mathcal{V}$. We can use the Taylor expansion of f to find such an r (this idea is similar to that in [KN04, p. 33, Lemma 1.6]). Specifically, writing

$$f(x+z) - f(x) = zf'(x) + \sum_{k=2}^{\deg f} \frac{z^k}{k!} f^{(k)}(x)$$

we note by the strong triangle inequality that it suffices to choose r so that

$$\left| \frac{f^{(k)}(x)}{k!} \right| r^{k-1} < |f'(x)|$$

for all $x \in \mathbb{Z}_p$ and $n \geq 2$.

5. STRUCTURE OF LOCALLY SCALING TRANSFORMATIONS

Lemma 5.1. *Let $X \subset K$ be compact-open, with $T : X \rightarrow X$ locally scaling for $r \in \mathcal{V}$. Let $C : X \rightarrow \mathbb{R}_{\geq 1}$ be the scaling function of Definition 4.1. Then for each $a \in X$ and $r' \in \mathcal{V}$ with $r' \leq r$, the map*

$$T|_{B_{r'}(a)} : B_{r'}(a) \rightarrow B_{r'C(a)}(T(a))$$

is a bijection.

Proof. Denote $B = B_{r'}(a)$ and $B' = B_{r'C(a)}(T(a))$. As C is constant on $B_r(a)$, $|Tx - Ty| = C(x)|x - y| = C(a)|x - y|$ for all $x, y \in B \subset B_r(a)$. This implies that $T(B) \subset B'$, so our restriction is well-defined. It also implies that the restriction is injective.

For each $k \geq 0$ we may take coset representatives a_0, \dots, a_{q^k-1} for $B/B_{r'|\pi|^k}(0)$. Then for $i, j \in \{0, \dots, q^k-1\}$ we have

$$|T(a_i) - T(a_j)| = C(a)|a_i - a_j| > r'C(a)|\pi|^k.$$

Therefore, $T(a_0), \dots, T(a_{q^k-1})$ are precisely the q^k coset representatives for $B'/B_{r'C(a)|\pi|^k}(0)$. It follows that $T(B)$ is dense in B' .

Now, note that $T|_B$ is continuous. So, $T(B)$ is the continuous image of a compact set, thus compact, and so closed. So, $T(B) = B'$. This proves surjectivity, and the lemma is proved. \square

Corollary 5.2. *Let X, T, C be as in Lemma 5.1. Set $H = X/B_r(0)$. For any $i, j \in H$, $a \in j$, and $r' \in \mathcal{V}$ with $r' \leq r$, the set*

$$i \cap T^{-1}(B_{r'}(a))$$

is either the empty set or a ball of radius $r'/C(i)$, according to whether or not $i \cap T^{-1}(j)$ is empty.

Proof. Denote $B = B_{r'}(a)$. Assume $i \cap T^{-1}(j)$ is not empty, so there is a $y \in i \cap T^{-1}(j)$. Then the map

$$T|_{B_r(y)} : i = B_r(y) \rightarrow B_{rC(i)}(T(y)) \supseteq j \supseteq B$$

is a bijection. It follows that $i \cap T^{-1}(B)$ is non-empty, and we may in fact assume that $y \in i \cap T^{-1}(B)$.

Then, as

$$T|_{B_{r'/C(i)}(y)} : B_{r'/C(i)}(y) \rightarrow B_{r'}(T(y)) = B$$

is also a bijection, it follows that $i \cap T^{-1}(B) = B_{r'/C(i)}(y)$. \square

Definition 5.3. Let $X \subset K$ be compact-open, and let $T : X \rightarrow X$ be locally scaling for $r \in \mathcal{V}$. Let $H = X/B_r(0)$ and $C : H \rightarrow \mathbb{R}_{\geq 1}$ be the scaling function. Then we define the *associated transition matrix* to be the map $A : H^2 \rightarrow \mathbb{R}_{\geq 0}$ given by, for $i, j \in H$,

$$A(i, j) = \begin{cases} 0 & i \cap T^{-1}(j) = \emptyset, \\ \rho(1/C(i)) & \text{otherwise.} \end{cases}$$

Lemma 5.4. *Let $X \subset K$ be compact-open and $T : X \rightarrow X$ be locally scaling for $r \in \mathcal{V}$; let $H = X/B_r(0)$ and let $A : H^2 \rightarrow \mathbb{R}_{\geq 0}$ be the associated transition matrix. Then:*

(i) *For $S \subset X$ measurable and $i \in H$*

$$\mu(i \cap T^{-1}(S)) = \sum_{j \in H} \mu(S \cap j) A(i, j).$$

(ii) $A(i, j) = \frac{1}{\rho(r)} \mu(i \cap T^{-1}(j))$.

(iii) A is a stochastic matrix on H .

(iv) T is measure-preserving if and only if the sum of each column of A is 1.

Proof.

(i):

By disjoint additivity of μ , it suffices to prove the equality in the case $S \subset j$ for some $j \in H$. As the balls form a sufficient semi-ring in the Borel σ -algebra of X , we may in addition assume that S is a ball. Say $S = B_{r'}(a)$ for $r' \leq r$ and $a \in j$. Then, by Corollary 5.2, we know that $i \cap T^{-1}(S)$ is either the empty set or a ball of radius $r'/C(i)$, according to whether or not $i \cap T^{-1}(j)$ is empty. Taking measures

we get

$$\begin{aligned} \mu(i \cap T^{-1}(S)) &= \begin{cases} 0 & i \cap T^{-1}(j) = \emptyset, \\ \rho(r'/C(i)) & \text{otherwise} \end{cases} \\ &= \mu(S)A(i, j) = \sum_{j \in H} \mu(S \cap j)A(i, j). \end{aligned}$$

(ii):

Put $S = j$ in (i). Then we get

$$\mu(i \cap T^{-1}(j)) = \mu(j \cap j)A(i, j) = \rho(r)A(i, j).$$

(iii):

Note that for each $i \in H$, by disjoint additivity of μ along with (ii) we have

$$\rho(r) \sum_{j \in H} A(i, j) = \sum_{j \in H} \mu(i \cap T^{-1}(j)) = \mu(i \cap X) = \mu(i) = \rho(r).$$

(iv):

If T is measure-preserving, then for each $j \in H$ we have, by disjoint additivity of μ ,

$$\sum_{i \in H} A(i, j) = \frac{1}{\rho(r)} \sum_{i \in H} \mu(i \cap T^{-1}(j)) = \frac{1}{\rho(r)} \mu(X \cap T^{-1}(j)) = \frac{1}{\rho(r)} \mu(T^{-1}(j)) = 1.$$

For the converse we use (i) and disjoint additivity:

$$\mu(T^{-1}(S)) = \sum_{i \in H} \mu(i \cap T^{-1}(S)) = \sum_{i, j \in H} \mu(S \cap j)A(i, j) = \sum_{j \in H} \mu(S \cap j) = \mu(S). \quad \square$$

Remark 5.5. Given $f \in K[x]$ such that $f(X) \subset X$, Lemma 4.4 and Lemma 5.4 allow us to determine whether $f : X \rightarrow X$ is measure-preserving. By a linear change of variables, we may assume that $X \subset \mathcal{O}$. Then we first check whether $|f'(a)| \geq 1$ for all $a \in X$ (writing $f = h/\pi^k$ with $h \in \mathcal{O}[x]$ reduces this to checking that $|h'(a)| \geq 1/|\pi^k|$ for $a \in X$, which is a finite computation in $\mathcal{O}/\mathfrak{p}^k$). We may take $r \in \mathcal{V}$ with $r \leq |\pi|^k$ such that X is a finite union of r -balls, and f will be locally scaling for r . The associated transition matrix can be obtained in a finite computation.

Example 5.6. We consider a polynomial map on \mathbb{Z}_2 that defines a locally scaling but not measure-preserving transformation:

$$f(x) = \binom{x}{3} = \frac{x(x-1)(x-2)}{3!}.$$

We note that

$$|f'(x)| = \left| \frac{3x^2 - 6x + 2}{6} \right| = \begin{cases} 1, & x \in B_{1/2}(0), \\ 2, & x \in B_{1/2}(1), \end{cases}$$

where the final equality follows by computing the values of the numerator modulo 4. In particular, we see that $|f'(x)| \geq 1$ for all $x \in \mathbb{Z}_2$, and so f is locally scaling by Lemma 4.3. One can verify that f is locally scaling for $r = 1/2$ and that the

associated transition matrix is

$$A = (A(i, j))_{0 \leq i, j \leq 1} = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

identifying i with $i + 2\mathbb{Z}_2$ for $i \in \{0, 1\}$. By Lemma 5.4 f is not measure-preserving.

Proposition 5.7. *Let $X \subset K$ be compact-open, and let $T : X \rightarrow X$ be a locally scaling transformation for $r \in \mathcal{V}$, with $\Sigma = (X, \mu, T)$ the corresponding measurable dynamical system.*

Let $H = X/B_r(0)$, let $A : H^2 \rightarrow \mathbb{R}_{\geq 0}$ be the associated transition matrix, and let $w : H \rightarrow \mathbb{R}_{\geq 0}$ be the positive row vector given by $w(i) = \rho(r)$ for $i \in H$. Let $\Sigma' = (X_A, \mu_{A,w}, T_A)$ be the corresponding Markov shift.

Then there exists a continuous, measure-preserving surjection $\Phi : X \rightarrow X_A$ satisfying $\Phi \circ T = T_A \circ \Phi$. Moreover, the pre-image under Φ of a cylinder set is a ball of the same measure.

Proof. For each $n \geq 0$, let $\pi_n : X_A \rightarrow H$ denote projection to the n^{th} coordinate. Let $\phi : X \rightarrow H$ be the canonical projection. Consider the map $\Phi : X \rightarrow X_A$ defined by

$$\pi_n \circ \Phi = \phi \circ T^n;$$

i.e., the n^{th} slot in X_A denotes which element of H the point $T^n(x)$ is in. Then $\Phi \circ T = T_A \circ \Phi$ by construction.

Let $d_0, d_1, \dots \in H$. We will prove by induction on the number of slots specified (the ‘‘length’’ of the cylinder set $[d_0 \dots d_\ell]$) that the claim that the pre-image of the cylinder set $[d_0 \dots d_\ell]$ is a ball of the same measure as the cylinder set. Note that $\Phi^{-1}([d_0]) = d_0$ is a ball of the correct measure, as

$$\mu_{A,w}([d_0]) = w(d_0) = \rho(r) = \mu(d_0)$$

by construction. Now,

$$\Phi^{-1}([d_0 \dots d_\ell]) = \Phi^{-1}([d_0]) \cap T^{-1}\Phi^{-1}([d_1 \dots d_\ell]).$$

By the inductive hypothesis, this is the intersection of two balls, and is thus again a ball; so Φ is continuous.

Noting that $\Phi^{-1}([d_1 \dots d_\ell]) \subset \Phi^{-1}[d_1] = d_1$ and applying claim (i) of Lemma 5.4, along with the inductive hypothesis, we see that this ball has the correct measure

$$\begin{aligned} \mu(\Phi^{-1}([d_0 \dots d_\ell])) &= \sum_{j \in H} \mu(j \cap \Phi^{-1}([d_1 \dots d_\ell])) A(d_0, j) \\ &= A(d_0, d_1) \mu(\Phi^{-1}([d_1 \dots d_\ell])) \\ &= A(d_0, d_1) \mu([d_1 \dots d_\ell]) = \mu([d_0 \dots d_\ell]). \end{aligned}$$

As the cylinder sets are a sufficient semi-ring in the Borel σ -algebra of X_A , this shows that Φ is measure-preserving.

Note that Φ continuous and measure-preserving implies that Φ is surjective: X is compact and X_A is Hausdorff, so the image must be closed. However, the image must have full measure and so must be dense (w positive implies that all cylinder sets, hence all open sets, have strictly positive measure). \square

Theorem 5.8. *Let $X, T, H, \Sigma, \Sigma', \Phi$ be as in Proposition 5.7. Moreover, assume that Σ is measure-preserving, so that Σ' is as well.*

Now, let $H = \bigsqcup_{k=1}^n H_k$ be a decomposition of H in the sense of Proposition 2.1, so that

$$\Sigma' = \bigsqcup_{k=1}^n \Sigma'_k,$$

where $\Sigma'_k = (X_{A_k}, \mu_{A_k, w}, T_{A_k})$.

For $k = 1, \dots, n$ define

$$\tilde{\Sigma}_k = \begin{cases} \text{restriction of } \Sigma \text{ to } \Phi^{-1}(X_{A_k}) & \#X_{A_k} < \infty, \\ \Sigma'_k & \text{otherwise,} \end{cases}$$

then we have an isomorphism of topological and measurable dynamical systems

$$\Sigma \cong \bigsqcup_{k=1}^n \tilde{\Sigma}_k.$$

Moreover, each term in this decomposition is either locally an isometry or ergodic Markov, according to whether or not $\#X_{A_k} < \infty$.

Proof. By Lemma 5.4, Σ' is a quotient of Σ , so Σ measure-preserving implies Σ' measure-preserving.

For $k = 1, \dots, n$, denote $C_k = \Phi^{-1}(X_{A_k})$, $\mu_k = \mu|_{C_k}$, $T_k = T|_{C_k}$. The decomposition of Σ' induces the following decomposition of Σ :

$$\Sigma = \bigsqcup_{k=1}^n (C_k, \mu_k, T_k).$$

To complete the proof of the proposition, it suffices to show that $(C_k, \mu_k, T_k) \cong \tilde{\Sigma}_k$ for $k = 1, \dots, n$ as topological and measurable dynamical systems, and to classify them as being locally isometries and ergodic Markov in the two cases. We now handle the two cases separately:

Case 1: $\#X_{A_k} < \infty$.

If $\#X_{A_k} < \infty$, then the isomorphism $(C_k, \mu_k, T_k) \cong \tilde{\Sigma}_k$ follows by definition. Note that the measure on Σ'_k is necessarily atomic; as it is ergodic, it must in fact be the inverse orbit of a single atom. As T_A , hence T_{A_k} , is measure-preserving, each of the atoms must have equal measure. It follows that each element $x \in X_{A_k}$ is of the form

$$x = (d_0, d_1, \dots, d_\ell, d_0, \dots, d_\ell, d_0, \dots, d_\ell, \dots),$$

with $A(d_0, d_1) = A(d_1, d_2) = \dots = A(d_\ell, d_0) = 1$. Then $\Phi^{-1}(x) = d_0$, where $C(d_0) = 1$ (here, C is from the definition of locally scaling).

So, C_k must be a collection of r -balls with $C(x) = 1$ for $x \in C_k$. Then, for $x, y \in C_k$ with $|x - y| \leq r$, we have $|T(x) - T(y)| = C(x)|x - y| = |x - y|$. This shows that $\tilde{\Sigma}_k$ is locally an isometry, as desired.

Case 2: $\#X_{A_k} = \infty$.

If $\#X_{A_k} = \infty$, then we claim that Φ induces an isomorphism $(C_k, \mu_k, T_k) \cong \tilde{\Sigma}_k$. In a measure-preserving Markov shift, any atoms must have finite inverse orbit; so Σ'_k ergodic and $\#X_{A_k} = \infty$ implies that μ_{A_k, w_k} is non-atomic. Recall that Φ is surjective. We claim that it is also injective. For $x \in X$ let $d_n = \pi_H T^n(x)$ for $n = 0, 1, \dots$. Then

$$\Phi^{-1}(\Phi(x)) = \bigcap_{\ell \geq 0} \Phi^{-1}([d_0 \dots d_\ell]).$$

We have from Proposition 5.7 that each of these pre-images is a ball. Then $\Phi^{-1}(\Phi(x))$ is the intersection of a nested family of balls. If the intersection contains more than a single point, then the radii of the balls do not go to 0, and so the intersection has a non-empty interior and thus positive measure. Now, the measure on X_{A_k} is non-atomic, so $\mu_{A_k, w_k}(\Phi(x)) = 0$. As Φ is measure-preserving, this implies that $\mu(\Phi^{-1}(\Phi(x))) = 0$; by the above considerations this implies that $\Phi^{-1}(\Phi(x))$ contains at most one point. So, Φ is injective.

Then Φ is a continuous, measure-preserving bijection. Observe that Φ takes closed sets to closed sets by compactness, so its inverse is also continuous. This also implies that Φ^{-1} is measurable, and then Φ measure-preserving implies Φ^{-1} measure-preserving. So, Φ is an isomorphism of topological and measurable dynamical systems $(C_k, \mu_k, T_k) \cong \tilde{\Sigma}_k$, as desired. As the later is ergodic Markov, the former is as well. \square

Corollary 5.9. *Let $X \subset K$ be compact-open. Let $f \in \mathcal{C}^1(X)$ be such that $f(X) \subset X$ and such that $f : X \rightarrow X$ is measure-preserving with respect to $\mu|_X$. Then X consists of a finite disjoint union of compact-open sets such that T restricted to each such set is either topologically and measurably isomorphic to an ergodic Markov transformation or a local isometry.*

Proof. By Lemma 4.4 T must be a locally scaling transformation. Theorem 5.8 completes the proof. \square

Corollary 5.10. *Let $X \subset K$ be compact-open and $T : X \rightarrow X$ a measure-preserving locally scaling transformation. If T is ergodic, then it is either Markov or locally an isometry. In particular, if it is weakly mixing, then it also Markov and so mixing. So, for a measure-preserving locally scaling transformation on a compact-open X , weakly mixing implies mixing. Furthermore, if T is invertible it cannot be weakly mixing.*

Proof. If T is ergodic, then the decomposition in Theorem 5.8 must be trivial. So, T must be either Markov or locally an isometry. If it is locally an isometry, then it cannot be weakly mixing. So, weakly mixing implies weakly mixing Markov, which in turn implies mixing. If T is invertible it has no Markov factors, and by Proposition 5.7 it factors through a finite dynamical system and hence cannot be weakly mixing. \square

Corollary 5.11. *For a locally scaling transformation, the following properties depend only on the associated transition matrix:*

- (i) *Measure-preserving.*
- (ii) *Weakly mixing, mixing, exact, Bernoulli.*

Proof. By Lemma 5.4, the property of being measure-preserving depends only on the associated transition matrix.

Note that the decomposition in Theorem 5.8 depends only on the associated transition matrix. Given an associated transition matrix, we have the following cases:

- (i) The decomposition is trivial, and the system is a local isometry. Then it is not weakly mixing (or any of the stronger properties listed).
- (ii) The decomposition is trivial, and the system is ergodic Markov. In this case, the system is determined up to isomorphism by the matrix.

- (iii) The decomposition is not trivial. In this case, the system is not ergodic and cannot satisfy any of the stronger properties listed. \square

6. POLYNOMIAL APPROXIMATION IN \mathcal{O}

The above results dealt with C^1 functions, extending to polynomial maps as a special case. In the next sections we will be interested in finding polynomial maps with specified associated transition matrices. In preparation for this, we will need some results on the approximation of continuous maps $\mathcal{O} \rightarrow K$. For the reader's convenience, we will sketch here the definitions and results of [Ami64], slightly simplified for our applications.

Say $X \subset \mathcal{O}$ is compact-open. Moreover, assume that X is a finite union of r -balls for $r \in \mathcal{V}$. Then, for $r' \leq r$, each r' -ball contained in X is a union of precisely q balls of radius $|\pi|r'$ contained in X . In the terminology of [Ami64], this makes X a *regular valued compact* (*compact valué régulier* in the original French).

For $k \geq \log_{|\pi|} r$, we may define $H_k = X/B_{|\pi|^k}(0)$, and a projection map $\pi_k : X \rightarrow H_k$. Then we say that a sequence $\{u_k \in X : k \in \mathbb{N}\}$ is *very well distributed* (*très bien répartie*) if for each $k \geq \log_{|\pi|} r$, $h \in H_k$, and $m \geq 1$ we have

$$\#\{i < m \# H_k : u_i \in h\} = m.$$

That is, the terms of the sequence must be equally distributed among the possible values mod \mathfrak{p}^k for $k \geq \log_{|\pi|} r$. Note that the condition that the $\{u_k\}$ are very well distributed implies that they are distinct.

Now, given such a sequence $\{u_0, u_1, \dots\}$, we may define the *corresponding interpolating polynomials* for $k \geq 0$:

$$P_k(x) = (x - u_0)(x - u_1) \cdots (x - u_{k-1}) \quad \text{and} \quad Q_k(x) = \frac{P_k(x)}{P_k(u_k)}.$$

We may summarize some of the results of [Ami64, §II.6.2] as follows:

Theorem 6.1 (Amice). *Let $X \subset \mathcal{O}$ be compact-open, and let $\{u_k\}$ be a very well distributed sequence with values in X with P_k, Q_k the corresponding interpolating polynomials. Let $f : X \rightarrow K$ be continuous, and for $k \geq 0$ set*

$$a_k = P_k(u_k) \left(\sum_{j=0}^k \frac{f(u_j)}{P'_{k+1}(u_j)} \right).$$

Then:

- (i) $|a_k| \rightarrow 0$ as $k \rightarrow \infty$.
- (ii) $\sum_{k \geq 0} a_k Q_k(x) \rightarrow f(x)$ uniformly on X .
- (iii) The a_k are determined by (ii).
- (iv) $\sup_{x \in X} |f(x)| = \sup_{k \in \mathbb{N}} |a_k|$.

A very well distributed sequence $\{u_k\}$ is said to be *well ordered* (*bien ordonnée*) if $|u_n - u_m| = |\pi|^{v_q(n-m)}$ for all $n, m \geq 0$, where $v_q(n-m)$ is the exact power of q dividing $n-m \in \mathbb{Z}$. Following our sources, we will call such a sequence *T.B.R.B.O.* (*très bien répartie bien ordonnée*). This allows us to state results of Helmsmoortel and Barsky, characterizing Lipschitz and C^1 functions on \mathcal{O} in terms of the coefficients in their expansions. This result may be found in [Bar73].

Theorem 6.2 (Helsmoortel, Barsky). *Let $\{u_k\}$ be a T.B.R.B.O. sequence with values in \mathcal{O} , with P_k, Q_k the corresponding interpolating polynomials. Let $f : \mathcal{O} \rightarrow K$ be continuous, with*

$$f(x) = \sum_{k \geq 0} a_k Q_k(x)$$

the expansion of f in the sense of Theorem 6.1. For $k \geq 1$, define

$$\kappa_k = |\pi|^{-\lfloor \log_q k \rfloor}.$$

Then:

- (i) *f is r -Lipschitz if and only if $r \leq \kappa_k |a_k|$ for all $k \geq 1$;*
- (ii) *$f \in \mathcal{C}^1(\mathcal{O})$ if and only if $\kappa_k |a_k| \rightarrow 0$ as $k \rightarrow \infty$.*

Example 6.3. Note that $\{0, 1, 2, \dots\} \subset \mathbb{Z}_p$ satisfies the conditions for being a very well distributed sequence, and is in fact trivially T.B.R.B.O. Then

$$Q_k(x) = \frac{x(x-1)\cdots(x-k+1)}{k \cdot (k-1) \cdots 1} = \binom{x}{k}.$$

So, in this case the above reduces to the Mahler expansion.

More generally: Let a_0, \dots, a_{q-1} be a complete set of coset representatives for \mathcal{O}/\mathfrak{p} . For $k \in \mathbb{N}$, we will define u_k in terms of the base- q expansion of k :

$$k = \sum_{i=0}^{\ell} k_i q^i \mapsto \sum_{i=0}^{\ell} a_{k_i} \pi^i = u_k.$$

Then, say we have $n, m \in \mathbb{N}$ with $n = \sum_{i \geq 0} n_i q^i$ and $m = \sum_{i \geq 0} m_i q^i$. Let $\ell = v_q(i-j) = \min\{i : n_i \neq m_i\}$. Then

$$|u_n - u_m| = |\pi|^\ell = |\pi|^{v_q(n-m)},$$

and $\{u_k\}$ is T.B.R.B.O. In particular, this implies that there is always a T.B.R.B.O. sequence for \mathcal{O} and corresponding interpolating polynomials, such that the results cited in this section hold.

Now, we establish a lemma that will be of particular interest to us:

Lemma 6.4. *Let $\{u_n\}$ be a T.B.R.B.O. sequence in \mathcal{O} with corresponding interpolating polynomials P_k, Q_k . Then, for $k \in \mathbb{N}$:*

- (i) $Q_k(\mathcal{O}) \subset \mathcal{O}$.
- (ii) Q_k is κ_k -Lipschitz, with κ_k as in Theorem 6.2.
- (iii) If $k = q^\ell$ for some $\ell \geq 0$, then

$$|Q_k(x) - Q_k(y)| = \kappa_k |x - y| \text{ for all } x, y \in \mathcal{O} \text{ with } |x - y| \leq 1/\kappa_k.$$

Proof. Claim (i) follows by applying Theorem 6.1(iv) with $f = Q_k$ (so that $a_i = 1$ for $i = k$ and 0 otherwise). Claim (ii) follows similarly from Theorem 6.2.

Assume $k = q^\ell$. Fix $x, y \in \mathcal{O}$ with $|x - y| \leq 1/\kappa_k$; as the $\{u_n\}$ are very well distributed there is some $m \in \{0, \dots, k-1\}$ such that $|x - u_m| \leq 1/\kappa_k$ (hence also $|y - u_m| \leq 1/\kappa_k$).

Define a polynomial

$$S_j(y) = \sum_{0 \leq i_1 < i_2 < \dots < i_j < k} \frac{(y - u_0) \cdots \widehat{(y - u_{i_1})} \cdots \widehat{(y - u_{i_j})} \cdots (y - u_{k-1})}{(u_k - u_{k-1}) \cdots (u_k - u_0)}.$$

Then we may observe that

$$\begin{aligned} Q_k(x') - Q_k(y') &= \frac{(x' - u_{k-1}) \cdots (x' - u_0) - (y' - u_{k-1}) \cdots (y' - u_0)}{(u_k - u_{k-1}) \cdots (u_k - u_0)} \\ &= \sum_{j=1}^k (x' - y')^j S_j(y') \end{aligned}$$

for any $x', y' \in \mathcal{O}$. Then

$$\begin{aligned} Q_k(x) - Q_k(y) &= (Q_k(x) - Q_k(u_m)) - (Q_k(y) - Q_k(u_m)) \\ &= (x - y)S_1(u_m) + \sum_{j=2}^k S_j(u_m) ((x - u_m)^j - (y - u_m)^j). \end{aligned}$$

We will prove the following two statements, which together with the strong triangle inequality and the previous expression imply our desired result:

- (i) $|S_1(u_m)| = \kappa_k$;
- (ii) $|S_j(u_m)| < \kappa_k^j$ for $1 < j \leq k$.

That this suffices is clear, for the $j = 1$ term will dominate in valuation.

Observe that

$$S_1(u_m) = \frac{(y - u_0) \cdots (\widehat{y - u_m}) \cdots (y - u_{k-1})}{(u_k - u_0) \cdots (u_k - u_{k-1})}$$

and that

$$S_j(u_m) = \sum_{\substack{0 \leq i_1 < i_2 < \cdots < i_j < k \\ m \notin \{i_1, \dots, i_j\}}} \frac{(u_m - u_0) \cdots (\widehat{u_m - u_{i_1}}) \cdots (\widehat{u_m - u_{i_j}}) \cdots (u_m - u_{k-1})}{(u_k - u_{k-1}) \cdots (u_k - u_0)}.$$

Suppose $\{u_n\}$ is a very well distributed sequence. Then it is easy to check that $\{v_0, \dots, v_{k-1}\}$ must contain precisely $q^{\ell'}$ elements bounded by $\pi^{\ell'}$ for each $\ell' \leq \ell$. Now, observe that both $\{u_m - u_n : n \in \mathbb{N}\}$ and $\{u_k - u_n : n \in \mathbb{N}\}$ are very well distributed. Let m' be the unique index in $\{0, \dots, k-1\}$ such that $|u_k - u_{m'}| \leq 1/\kappa_k$; the very well distributed property of $\{u_n\}$ implies that this is in fact an equality. Our previous count implies that we must have

$$\left| \frac{(y - u_0) \cdots (\widehat{y - u_m}) \cdots (y - u_{k-1})}{(u_k - u_0) \cdots (\widehat{u_k - u_{m'}}) \cdots (u_k - u_{k-1})} \right| = 1.$$

So,

$$|S_1(u_m)| = \left| \frac{1}{u_k - u_{m'}} \right| = \kappa_k.$$

Now,

$$\left| \frac{S_j(u_m)}{S_i(u_m)} \right| = \left| \sum_{\substack{0 \leq i_1 < i_2 < \cdots < i_{j-1} < k \\ m \notin \{i_1, \dots, i_{j-1}\}}} \frac{1}{(u_m - u_{i_1}) \cdots (u_m - u_{i_{j-1}})} \right| < \kappa_k^{j-1},$$

for $|u_m - u_{i_1}|, \dots, |u_m - u_{i_{j-1}}| > 1/\kappa_k$, as $\{u_n\}$ is very well distributed (and so the first k elements must be in disjoint $1/\kappa_k$ -balls). This completes our proof. \square

7. POLYNOMIAL MAPS ON \mathcal{O} REALIZING LOCALLY SCALING TRANSFORMATIONS

Sections 4 and 5 characterize measure-preserving polynomial transformations on \mathcal{O} in terms of locally scaling transformations. However, we have shown the existence of only a handful of such maps. In this section, we will show that in fact the polynomials, in a sense, provide a representative class among the measure-preserving locally scaling maps.

We begin with a lemma giving sufficient conditions for two maps to have the same associated transition matrices.

Lemma 7.1. *Suppose $T : \mathcal{O} \rightarrow \mathcal{O}$ is locally scaling for $r \in \mathcal{V}$. Set $H = \mathcal{O}/B_r(0)$, let $C : H \rightarrow \mathbb{R}_{\geq 1}$ be the scaling function for T , and let $A : H^2 \rightarrow \mathbb{R}_{\geq 0}$ be the associated transition matrix for T . Suppose in addition that $S : \mathcal{O} \rightarrow \mathcal{O}$ is a transformation such that the difference $R = T - S$ satisfies*

- (i) $|R(x) - R(y)| < C(x)|x - y|$ whenever $0 < |x - y| \leq r$;
- (ii) $|R(x)| \leq rC(x)$ for all x .

Then S is locally scaling for $r \in \mathcal{V}$, with scaling function C and associated transition matrix A .

Proof. For $0 < |x - y| \leq r$ we have

$$|S(x) - S(y)| = |T(x) - T(y) + R(x) - R(y)| = |T(x) - T(y)| = C(x)|x - y|$$

by the strong triangle inequality. Indeed, $|T(x) - T(y)| = C(x)|x - y|$, which by (i) is strictly greater than $|R(x) - R(y)|$. So, S is locally scaling for $r \in \mathcal{V}$, with scaling function C .

Now, it remains to verify that $i \cap T^{-1}(j) = \emptyset \Leftrightarrow i \cap S^{-1}(j) = \emptyset$ for $i, j \in H$. For this, it suffices to show that $T(B_r(x)) = S(B_r(x))$ for all $x \in \mathcal{O}$. Indeed, applying Lemma 5.1 and (ii) yields

$$T(B_r(x)) = B_{rC(x)}(T(x)) = B_{rC(x)}(S(x)) = S(B_r(x)). \quad \square$$

For $S \subset K$ we say that $T : S \rightarrow K$ is *affine* if it is given by $x \mapsto ax + b$ for some constants $a, b \in K$. We say that $T : \mathcal{O} \rightarrow K$ is *locally affine* if for each $x \in \mathcal{O}$ there exists an $r \in \mathcal{V}$ such that $T|_{B_r(x)}$ is affine.

Theorem 7.2. *Let $r \in \mathcal{V}$ and $H = \mathcal{O}/B_r(0)$. Let A be a stochastic matrix on H . Then, let*

$$\mathcal{T}_A = \{T \text{ locally scaling for } r : A \text{ is the associated transition matrix for } T\}.$$

If \mathcal{T}_A is non-empty, then:

- (i) \mathcal{T}_A contains a locally affine transformation;
- (ii) \mathcal{T}_A contains infinitely many polynomials.

Proof.

(i):

Suppose $T \in \mathcal{T}_A$. Let $C : H \rightarrow \mathbb{R}_{\geq 1}$ be the scaling function in the definition of locally scaling and observe that its image is contained in \mathcal{V} . Let $S : H \rightarrow K$ be any function satisfying $|S(h)| = C(h)$ for all $h \in H$. Then, for each $h \in H$, we observe that $T(h) = \{T(x) : x \in h\}$ and $S(h)h = \{S(h)x : x \in h\}$ are both balls of radius $C(h)r$. So, there exists a function (indeed, many functions) $M : H \rightarrow K$ such that $T(h) = S(h)h + M(h)$. Regarding S and M as functions with domain \mathcal{O} via the

quotient map $\mathcal{O} \rightarrow H$, we may define a locally affine transformation $T_{S,M} : \mathcal{O} \rightarrow \mathcal{O}$ by the formula $T_{S,M}(x) = S(x)x + M(x)$.

Since S, M are constant on elements of H and $|S(h)| = C(h)$, it follows at once that $T_{S,M}$ is locally scaling for r with a scaling function C . By construction, $T_{S,M}(h) = T(h)$ (which is also why $T_{S,M}\mathcal{O} \subset \mathcal{O}$), so that $T_{S,M} \in \mathcal{T}_A$ by the definition of the associated transition matrix.

(ii):

Let $T \in \mathcal{T}_A$. By (i), we may assume that T is locally affine and hence strictly differentiable. Let $\{u_k\}$ be a T.B.R.B.O. sequence in \mathcal{O} (which must exist by Example 6.3). Let

$$T = \sum_{k \geq 0} a_k Q_k$$

be the decomposition of T in the sense of Theorem 6.1.

Take $\alpha \in \mathcal{V}$ with $\alpha < 1$. By Theorem 6.2, $\kappa_k |a_k| \rightarrow 0$ as $k \rightarrow \infty$, so there exists an $N \in \mathbb{N}$ such that for $k > N$ we have $\kappa_k |a_k| \leq \alpha < 1$. Moreover, take N such that $N > \#H$.

Let

$$f(x) = \sum_{k=0}^N a_k Q_k(x).$$

Note that f is a polynomial. Set

$$R(x) = T(x) - f(x) = \sum_{k > N} a_k Q_k(x).$$

Our bound on $\kappa_k |a_k|$ along with the choice $N > \#H = q^{\log_{|\pi|} r}$ implies that $|a_k| < 1/\kappa_N \leq r$ for $k > N$; so Theorem 6.1(iv) implies that $|R(x)| \leq r$ for all $x \in \mathcal{O}$. Lemma 6.4 implies that R is α -Lipschitz, so that $|R(x) - R(y)| \leq \alpha|x - y|$ for $x, y \in \mathcal{O}$. Noting that $\alpha < 1 \leq C(x)$ for all $x \in \mathcal{O}$, we observe that we may apply Lemma 7.1 to conclude that $f \in \mathcal{T}_A$.

Note that if

$$g(x) = \sum_{k > N} b_k Q_k(x)$$

is a polynomial such that $\kappa_k |b_k| \leq \alpha$, then the above argument also shows that $f + g \in \mathcal{T}_A$. So, there are indeed infinitely many polynomials in \mathcal{T}_A . \square

In particular, Theorem 7.2 shows the existence of measure-preserving mixing transformations on the p -adics given by polynomial maps. We can also use this method to compute explicit examples of such maps, but it is not particularly enlightening to do so.

8. POLYNOMIAL BERNOULLI MAPS ON \mathcal{O}

The construction of the preceding section gives infinite classes of measure-preserving polynomials with different kinds of measurable dynamics. Among these maps are Markov mixing maps. We will now study the class of such polynomials whose associated transition matrix has all entries equal, in which case the Markov transformation is in fact Bernoulli. The main upshot of this study is a class of explicitly given and relatively simple measure-preserving Bernoulli polynomial maps.

Definition 8.1. We say that a measure-preserving locally scaling map $T : \mathcal{O} \rightarrow \mathcal{O}$ is *isometrically Bernoulli* for $r \in \mathcal{V}$ if it is locally scaling for $r \in \mathcal{V}$ and all entries of the associated transition matrix are equal.

Let $V = \mathcal{O}/\mathfrak{p}^\ell \cong \mathbb{F}_q^\ell$ and define

$$B_V = \left(\prod_{i \geq 0} V, \mu_V, T_V \right),$$

where μ_V is the product probability measure and T_V is the left-shift. We may let d'_V be the quotient metric on V . Then we may define a metric d_V on $\prod_{i \geq 0} V$ by

$$d_V((a_0, a_1, a_2, \dots), (b_0, b_1, b_2, \dots)) = |\pi|^{-\ell(m-1)} d'_V(a_m, b_m),$$

where $m = \min\{i : a_i \neq b_i\}$.

We give two justifications for this metric:

- (i) View elements of V as ℓ -tuples under the isomorphism $\mathbb{F}_q^\ell \cong V$ corresponding to π -adic expansion (i.e., the isomorphism induced by the map shown in (ii)). Then, expanding each element in the product to a ℓ -tuple, d_V is just the dictionary metric (with base $|\pi|$).
- (ii) For each $a \in V$ we may let $\bar{a} \in \mathcal{O}$ be a coset representative for the quotient. Then the map

$$(a_0, a_1, \dots) \mapsto \sum_{i \geq 0} \bar{a}_i \pi^{\ell i}$$

gives a bijection $\prod_{i \geq 0} V \rightarrow \mathcal{O}$. This metric is the unique metric making this map an isometry.

Now, the term isometrically Bernoulli is partially motivated by the following:

Lemma 8.2. *Let $T : \mathcal{O} \rightarrow \mathcal{O}$ be a transformation and let $\ell \geq 1$. Then the following are equivalent:*

- (i) T is isometrically Bernoulli for $r = |\pi|^\ell$.
- (ii) For all $x, y \in \mathcal{O}$ satisfying $|x - y| \leq |\pi|^\ell$,

$$|T(x) - T(y)| = |\pi|^{-\ell} |x - y|.$$

- (iii) Let $V = \mathcal{O}/\mathfrak{p}^\ell$. There exists an invertible isometry $\Phi : \mathcal{O} \rightarrow \prod_{i \geq 0} V$ such that $\Phi \circ T = T_V \circ \Phi$; that is, (\mathcal{O}, μ, T) is metrically isomorphic to B_V .

Proof. (i) \Rightarrow (ii):

Let $H = \mathcal{O}/B_r(0)$, and let $A : H^2 \rightarrow \mathbb{R}_{\geq 0}$ be the associated transition matrix. Note that if T is isometrically Bernoulli, then each entry of A must be equal, and hence must be equal to $\frac{1}{\#H} = \rho(r)$. Now, T must be locally scaling for $r \in \mathcal{V}$, so $|T(x) - T(y)| = C(x)|x - y|$ for $|x - y| \leq r = |\pi|^\ell$. But, we must have $\rho(1/C(x)) = \rho(r)$, so $C(x) = 1/r = |\pi|^{-\ell}$.

(ii) \Rightarrow (iii):

Let $V = \mathcal{O}/\mathfrak{p}^\ell$. Now, (ii) implies that T is locally scaling for r . Letting A be the associated transition matrix, we readily note that all non-zero entries of A must be equal to $\rho(|\pi|^\ell)$. As A is a stochastic matrix, this implies that all entries of A are non-zero.

Now, let $\Sigma' = (X_A, \mu_{A,w}, T_A)$ be as in Theorem 5.8. We see that $B_V = \Sigma'$. We observed above that all entries of A are non-zero; then, A is irreducible and

Theorem 5.8 gives us a topological and measurable isomorphism $\Phi : \mathcal{O} \rightarrow X_A$. Note that the balls of X_A with respect to d_V are just the cylinder sets. Moreover, one may check that for each $m \geq 0$, X_A is a disjoint union of q^m balls of radius $r = |\pi|^m$, which must then each have measure $q^{-m} = \rho(r)$. Then Proposition 5.7 implies that Φ^{-1} takes balls of a given radius to balls of the same radius. Moreover, Φ^{-1} must take each of the q^m distinct balls of radius $|\pi|^m$ in X_A to a distinct ball of radius $|\pi|^m$ in \mathcal{O} . So each ball of radius $|\pi|^m$ in \mathcal{O} must be the pre-image of precisely one ball of the same radius in X_A . It follows that Φ and Φ^{-1} are both isometries.

(iii) \Rightarrow (i):

Note that for $x, y \in \mathcal{O}$ we have

$$|T(x) - T(y)| = d_V(\Phi(T(x)), \Phi(T(y))) = d_V(T_V(\Phi(x)), T_V(\Phi(y))).$$

Then, for $d_V(\Phi(x), \Phi(y)) = |x - y| \leq |\pi|^\ell$, we compute

$$d_V(T_V(\Phi(x)), T_V(\Phi(y))) = |\pi|^{-\ell} d_V(\Phi(x), \Phi(y)) = |\pi|^{-\ell} |x - y|. \quad \square$$

Remark 8.3. Note that item (ii) of Lemma 8.2 implies that T is $|\pi|^{-\ell}$ -Lipschitz.

Now, we may combine Lemma 6.4 with Lemma 8.2 to get:

Corollary 8.4. *Let $\{u_k\}$ be a T.B.R.B.O. sequence with values in \mathcal{O} , with corresponding interpolating polynomials P_k, Q_k . Say $T : \mathcal{O} \rightarrow \mathcal{O}$ is given by the expansion, in the sense of Theorem 6.1:*

$$T(x) = \sum_{k \geq 0} a_k Q_k(x), \text{ with } a_k \in \mathcal{O}, |a_k| \rightarrow 0.$$

Assume that

- (i) $M = \max_{k \geq 0} \kappa_k |a_k|$ exists, where κ_k is as in Theorem 6.2.
- (ii) There is a unique $k_M \geq 0$ attaining this maximum, and moreover it is of the form $k_M = q^\ell$ for some $\ell \geq 1$.
- (iii) $|a_{k_M}| = 1$ (hence, $M = \kappa_{k_M}$).

Then T is isometrically Bernoulli for $r = 1/M \in \mathcal{V}$.

Proof. As k_M is the unique value attaining the maximum, the strong triangle inequality and Lemma 6.4 imply that

$$|T(x) - T(y)| = \left| \sum_{k \geq 0} a_k [Q_k(x) - Q_k(y)] \right| = \kappa_{k_M} |a_{k_M}| |x - y| = M |x - y|$$

for $|x - y| \leq 1/\kappa_{k_M} = 1/M$. Then our claim follows by Lemma 8.2. \square

Example 8.5. Let $K = \mathbb{Q}_p$ and $\mathcal{O} = \mathbb{Z}_p$. Then $\{0, 1, 2, \dots\}$ is a T.B.R.B.O. sequence, and letting P_k, Q_k be the corresponding interpolating polynomials we can check that $Q_k(x) = \binom{x}{k}$ (cf. Example 6.3). In this case, $\kappa_k = p^{\lfloor \log_p k \rfloor}$. Therefore, we can rewrite the sufficient conditions in Corollary 8.4 as follows. Given $T : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ defined by

$$T(x) = \sum_{k \geq 0} a_k \binom{x}{k}$$

assume that

- (i) $M = \max_{k \geq 0} |a_k| p^{\lfloor \log_p k \rfloor}$ exists.

- (ii) There is a unique $k_M \geq 0$ attaining this maximum, and moreover it is of the form $k_M = p^\ell$ for some $\ell \geq 1$.
- (iii) $|a_{k_M}| = 1$ (thus, $M = p^\ell$).

Then T is isometrically Bernoulli for $r = p^{-\ell}$. In particular, the polynomials $\binom{x}{p^\ell}$ for $\ell > 0$ clearly satisfy these conditions, and so each defines a Bernoulli transformation on \mathbb{Z}_p .

Note, in particular, that this criterion applies to any map $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$ defined by $u \binom{x}{p} + F(x)$ with $u \in \mathbb{Z}_p^\times$ and $F \in \mathbb{Z}_p[x]$. An example of such a map is that which is given by $\frac{x^p - x}{p}$ from [WS98].

In this context, the polynomials $\binom{x}{p}$ and $\frac{x^p - x}{p}$ are in a sense the most natural isometrically Bernoulli maps:

Example 8.6. Take a set of coset representatives for $\mathbb{Z}_p/p\mathbb{Z}_p$. Then, using Example 6.3, we may form a T.B.R.B.O. sequence, and then the p^{th} corresponding interpolating polynomial (and unit multiples of it) will be Bernoulli by Corollary 8.4.

Let's look at the two most common sets of coset representatives for the quotient $\mathbb{Z}_p/p\mathbb{Z}_p$:

- (i) Take as coset representatives $0, 1, 2, \dots, p-1$. The resulting T.B.R.B.O. sequence is $\{0, 1, \dots\}$. Then $P_p(x) = x(x-1) \cdots (x-p+1)$, and $Q_p(x) = \binom{x}{p}$ is the p^{th} corresponding interpolating polynomial.
- (ii) Take as coset representatives 0 , and the $(p-1)^{\text{st}}$ roots of unity (there are exactly $p-1$ by Hensel's Lemma); these are called the "Teichmüller representatives." Then $P_p(x) = x^p - x$ and

$$Q_p(x) = \frac{P_p(x)}{P_p(p)} = \frac{1}{p^{p-1} - 1} \frac{x^p - x}{p}.$$

So, the polynomials $\binom{x}{p}$ and $\frac{x^p - x}{p}$ (up to unit) are analogs, arising by the same construction from the two most common choices for the coset representatives of $\mathbb{Z}_p/p\mathbb{Z}_p$.

Example 8.7. Let $K = \mathbb{F}_q((t))$ and $\mathcal{O} = \mathbb{F}_q[[t]]$. We may construct a T.B.R.B.O. sequence as in Example 6.3, having $0, 1, 2, \dots, q-1, t$ as its first $q+1$ terms. Then

$$Q_q(x) = \frac{x(x-1) \cdots (x-q+1)}{t(t-1)(t-2) \cdots (t-q+1)},$$

which is isometrically Bernoulli by Corollary 8.4. However, $t-1, t-2, \dots, t-q+1$ are all units in \mathcal{O} ; thus

$$t(t-1) \cdots (t-q+1)Q_q(x) = \frac{x(x-1) \cdots (x-q+1)}{t}$$

defines a Bernoulli transformation as well.

Now, we will give two examples of isometrically Bernoulli polynomial maps on the rings of integers of finite extensions of \mathbb{Q}_p . First, we briefly motivate our choice of examples. For K a finite extension of \mathbb{Q}_p , let $n = [K : \mathbb{Q}_p]$, $f = [\mathcal{O}/\mathfrak{p} : \mathbb{F}_p]$, and $e = \log_{|\pi|} |p|$. It is a standard result that $ef = n$. It is evident that the nature of how \mathcal{O} compares to \mathbb{Z}_p depends on the values of e and f . The two extreme

cases are $f = 1, e = n$ (in which case we say that the extension is *totally ramified*) and $e = 1, f = n$ (in which case we say that the extension is *unramified*). We give an example from each of these two extremes. For more background on the relevant theory, including the “standard” results invoked in this paragraph and in the following two examples; see [Ser62], particularly Ch. I §7, 8, Ch. III §5, Ch. IV §4.

Example 8.8. Take $p > 2$ and let

$$K = \mathbb{Q}_p(\zeta_p), \text{ where } \zeta_p \text{ is a primitive } p^{\text{th}} \text{ root of unity.}$$

It is a standard result that $1 - \zeta_p$ may be taken as a uniformizing parameter and that the extension is totally ramified, and so the set $\{0, \dots, p - 1\}$ gives a complete set of coset representatives for \mathcal{O}/\mathfrak{p} . We may construct a T.B.R.B.O. sequence as in Example 6.3. The first p terms would be just $0, \dots, p - 1$, with the next term $1 - \zeta_p$. The first p^2 terms would be $\{i + j(1 - \zeta_p)\}$ for $0 \leq i, j < p$, with the next term $(1 - \zeta_p)^2$. Noting that $q = p$ and applying Corollary 8.4 shows that the transformations defined by the polynomials

$$\frac{x(x - 1) \dots (x - p + 1)}{1 - \zeta_p}$$

and

$$\frac{1}{(1 - \zeta_p)^3} \prod_{0 \leq i, j < p} (x - i - j(1 - \zeta_p))$$

are isometrically Bernoulli for $r = |\pi|^1 = |p|^{\frac{1}{p-1}}$ and $r = |\pi|^2 = |p|^{\frac{2}{p-1}}$, respectively.

Example 8.9. Take $f > 1$ and let

$$K = \mathbb{Q}_p(\zeta), \text{ where } \zeta \text{ is a primitive } (p^f - 1)^{\text{th}} \text{ root of unity.}$$

It is a standard result that K is the unique unramified extension of \mathbb{Q}_p of degree f . So, p may be taken as a uniformizing parameter. Let $\bar{\zeta} \in \mathcal{O}/\mathfrak{p}$ be the image of ζ under the quotient map. We note that $\bar{\zeta}$ must generate the residue field extension, i.e., $\mathcal{O}/\mathfrak{p} = \mathbb{F}_p(\bar{\zeta}) = \mathbb{F}_p[\bar{\zeta}]$. So, $S = \{a_0 + a_1\bar{\zeta} + \dots + a_{f-1}\bar{\zeta}^{f-1}\}$, with $0 \leq a_0, a_1, \dots, a_{f-1} < p$, is a complete set of coset representatives for \mathcal{O}/\mathfrak{p} . Applying the construction of Example 6.3 we may construct a T.B.R.B.O. sequence whose first $q = p^f$ terms are precisely the elements of S , with the following term being p . Then applying Corollary 8.4 shows that the transformation defined by the polynomial

$$\frac{1}{p} \prod_{0 \leq a_0, a_1, \dots, a_{f-1} < p} (x - a_0 - a_1\bar{\zeta} - \dots - a_{f-1}\bar{\zeta}^{f-1})$$

is isometrically Bernoulli for $r = |p|$.

Remark 8.10. Define, as usual,

$$\widehat{\mathbb{Z}} = \varprojlim_{n, |} \mathbb{Z}/n\mathbb{Z} \cong \prod_p \mathbb{Z}_p.$$

We briefly note that the results of this section allow us to produce examples of maps $f : \mathbb{N} \rightarrow \mathbb{Z}$ which extend to Bernoulli maps on \mathbb{Z}_p for each p , and hence to a Bernoulli map on $\widehat{\mathbb{Z}}$. More explicitly, we obtain the following proposition:

Proposition 8.11. *Let a_0, a_1, \dots be a sequence of integers satisfying the following conditions for each rational prime:*

- (i) $|a_k|_p = 1$ for $k = p$;
- (ii) $|a_k|_p < p^{-\lfloor \log_p k \rfloor}$ for $k > p$.

Define $f : \mathbb{N} \rightarrow \mathbb{Z}$ by

$$f(n) = \sum_{k=0}^n a_k \binom{n}{k}.$$

Then f extends to an isometrically Bernoulli transformation $f : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ for each prime p .

Proof. For each prime p , note that the quantity $p^{\lfloor \log_p k \rfloor} |a_k|_p$ attains its maximum for $k = p$ (and for no other k) and that $|a_k|_p = 1$. Then the result is immediate by Corollary 8.4. \square

9. POLYNOMIAL ALMOST BERNOULLI MAPS ON \mathbb{Z}_p

An important condition shared by isometrically Bernoulli polynomials maps is that their derivatives must have constant valuation. We may use this observation to come up with a class of interesting non-examples. Our non-examples will be polynomial maps whose derivatives have constant valuation but that are not measure-preserving.

Proposition 9.1. *Given $n \in \mathbb{N}$, the map $f : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ defined by*

$$f(x) = \binom{x}{n}$$

satisfies $|f'(x)| = C$ for some $C \in \mathcal{V}$ and for all $x \in \mathbb{Z}_p$ if and only if $n = ap^\ell$, with $1 \leq a < p$ and $\ell \in \mathbb{Z}_{\geq 0}$, and

$$\frac{1}{u} + \dots + \frac{1}{u+a-1} \not\equiv 0 \pmod{p}$$

for each $u \in \{1, 2, \dots, p-a\}$, where inverses mod p are taken in \mathbb{F}_p^ .*

Proof. Write $n = ap^\ell + r$, $a < p$, $r < p^\ell$. We first wish to show that $r = 0$. We may compute

$$|f'(a)| = \left| \binom{n-1}{a} \right| / |n| \quad \text{for } 0 \leq a \leq n-1.$$

If $r \geq 1$, then $(1+t)^{r-1} \equiv (1+tp^\ell)^a (1+t)^{r-1} \pmod{p}$ has no $t^{p^\ell-1}$ term, so we see that $|f'(p-1)| < |f'(0)|$. Thus, $r = 0$. Note that $|f'(x)|$ is constant on \mathbb{Z}_p iff it is constant on \mathbb{Z} , since $\mathbb{Z} \subset \mathbb{Z}_p$ is dense. For $x \in \mathbb{Z}$, we observe that the set $\{x, x-1, \dots, x-n+1\}$ contains precisely n/p^k terms divisible by p^k for $k \leq \ell$, and either zero or one term divisible by some higher power of p . By the strong triangle inequality, we note that $|f'(x)| \leq p^\ell$, with equality if there is a term divisible by $p^{\ell+1}$. If no term is divisible by $p^{\ell+1}$, then let $up^\ell, (u+1)p^\ell, \dots, (u+a-1)p^\ell$ be the

terms divisible by p^ℓ and observe that $u, u + 1, \dots, u + a - 1$ are coprime to p . An easy computation shows that $|f'(x)| = p^\ell$ if and only if

$$\sum_{i=0}^{a-1} (u+i)^{-1} \not\equiv 0 \pmod{p}. \quad \square$$

Remark 9.2. Suppose the hypotheses of the proposition hold. Then we may compute the value of $f(x) \pmod{p}$ by performing a careful but easy computation involving cancelling corresponding powers in $x(x-1)\cdots(x-n+1)$ (henceforth, “the numerator”) and $n!$. The only terms which are not obviously matched are those corresponding to terms divisible by p^ℓ . Suppose $up^\ell, \dots, (u+a-1)p^\ell$ are the terms in the numerator divisible by p^ℓ , where now we need not assume that $u, \dots, u+a-1$ are coprime to p . Then

$$f(x) \equiv \frac{\prod_{i=0}^{a-1} (u+i)}{a!} \pmod{p}.$$

The simplest family of maps satisfying the hypotheses of Proposition 9.1 is that of the following corollary:

Corollary 9.3. *Let $p \geq 3$; then $n = (p-2)p^\ell$ satisfies the conditions of Proposition 9.1, and so $|f'(x)| = p^\ell$ for all $x \in \mathbb{Z}_p$.*

Proof. We may verify the last condition of Proposition 9.1 by an easy computation. □

Example 9.4. Suppose $p \geq 3$, $\ell \geq 1$, and $n = (p-2)p^\ell$. Define a transformation $f : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ by

$$f(x) = \binom{x}{n}.$$

By the corollary, we have the fact that $|f'(x)| = p^\ell$ for all $x \in \mathbb{Z}_p$. Mimicking the proof of Lemma 6.4 we can show that f is locally scaling for $r = p^{-(\ell+1)}$. Moreover, since $(p-1)! \equiv -1 \pmod{p}$, the computation of the previous remark yields

$$f(x) \pmod{p} \equiv \begin{cases} 0 & \text{if } x \equiv 0, 1, \dots, n-1 \pmod{p^{\ell+1}}, \\ 1 & \text{if } x \equiv n, n+1, \dots, n+p^\ell-1 \pmod{p^{\ell+1}}, \\ -1 & \text{if } x \equiv n+p^\ell, \dots, p^{\ell+1}-1 \pmod{p^{\ell+1}}. \end{cases}$$

Identifying $i \in \{0, 1, \dots, p^{\ell+1}-1\}$ with $i + p^{\ell+1}\mathbb{Z}_p$, the associated transition matrix for f is thus

$$\begin{aligned} A &= (A(i, j))_{0 \leq i, j < p^{\ell+1}} \\ &= \begin{cases} \frac{p-2}{p} & \text{if } i \equiv 0, 1, \dots, n-1 \pmod{p^{\ell+1}} \text{ and } j \equiv 0 \pmod{p}, \\ \frac{1}{p} & \text{if } i \equiv n, n+1, \dots, n+p^\ell-1 \pmod{p^{\ell+1}} \text{ and } j \equiv 1 \pmod{p}, \\ \frac{1}{p} & \text{if } i \equiv n+p^\ell, \dots, p^{\ell+1}-1 \pmod{p^{\ell+1}} \text{ and } j \equiv -1 \pmod{p}. \end{cases} \end{aligned}$$

So, if $p \neq 3$, we see that f is not measure-preserving. However, A does have a left-eigenvector of eigenvalue 1:

$$w = (w(i))_{0 \leq i < p^{\ell+1}} = \begin{cases} \frac{p-2}{p} & \text{if } i \equiv 0 \pmod{p}, \\ \frac{1}{p} & \text{if } i \equiv \pm 1 \pmod{p}. \end{cases}$$

This corresponds to an f -invariant measure $\tilde{\mu}$ on \mathbb{Z}_p defined on any μ -measurable set $S \subset \mathbb{Z}_p$ by

$$\tilde{\mu}(S) = \frac{p-2}{p}\mu(S \cap B_r(0)) + \frac{1}{p}\mu(S \cap B_r(1)) + \frac{1}{p}\mu(S \cap B_r(-1)).$$

Then the map Φ defined in Proposition 5.7 gives a measurable isomorphism of $(\mathbb{Z}_p, \tilde{\mu}, f)$ with $(X_A, \mu_{A,w}, T_A)$, where the latter dynamical system is mixing Markov.

ACKNOWLEDGEMENTS

This paper is based on research by the Ergodic Theory group of the 2005 SMALL summer research project at Williams College. Support for the project was provided by National Science Foundation REU Grant DMS - 0353634 and the Bronfman Science Center of Williams College. The authors would like to thank several anonymous referees for careful readings of the paper and valuable suggestions, and for bringing our attention to [AV94], [HY83], [RB], [FRL04], and [FRL06].

REFERENCES

- [Ami64] Yvette Amice, *Interpolation p -adique*, Bull. Soc. Math. France **92** (1964), 117–180. MR0188199 (32:5638)
- [Ana02] V. S. Anashin, *Uniformly distributed sequences of p -adic integers*, Diskret. Mat. **14** (2002), no. 4, 3–64. MR1964120 (2004a:11077)
- [AV94] David K. Arrowsmith and Franco Vivaldi, *Geometry of p -adic Siegel discs*, Phys. D **71** (1994), no. 1-2, 222–236. MR1264116 (95d:11162)
- [Bar73] Daniel Barsky, *Fonctions k -lipschitziennes sur un anneau local et polynômes à valeurs entières*, Bull. Soc. Math. France **101** (1973), 397–411. MR0371863 (51:8080)
- [Ben01] Robert L. Benedetto, *Hyperbolic maps in p -adic dynamics*, Ergodic Theory Dynam. Systems **21** (2001), no. 1, 1–11. MR1826658 (2002c:11163)
- [BS05] John Bryk and Cesar E. Silva, *Measurable dynamics of simple p -adic polynomials*, Amer. Math. Monthly **112** (2005), no. 3, 212–232. MR2125384 (2005m:37018)
- [FRL04] Charles Favre and Juan Rivera-Letelier, *Théorème d'équidistribution de Brolin en dynamique p -adique*, C. R. Math. Acad. Sci. Paris **339** (2004), no. 4, 271–276. MR2092012 (2005f:37090)
- [FRL06] ———, *Équidistribution quantitative des points de petite hauteur sur la droite projective*, Math. Ann. **335** (2006), no. 2, 311–361. MR2221116 (2007g:11074)
- [HY83] M. Herman and J.-C. Yoccoz, *Generalizations of some theorems of small divisors to non-Archimedean fields*, Geometric dynamics (Rio de Janeiro, 1981), Lecture Notes in Math., vol. 1007, Springer, Berlin, 1983, pp. 408–447. MR730280 (85i:12012)
- [KN04] Andrei Yu. Khrennikov and Marcus Nilson, *p -adic deterministic and random dynamics*, Mathematics and its Applications, vol. 574, Kluwer Academic Publishers, Dordrecht, 2004. MR2105195 (2005h:37102)
- [Lin04] Karl-Olof Lindahl, *On Siegel's linearization theorem for fields of prime characteristic*, Nonlinearity **17** (2004), no. 3, 745–763. MR2057125 (2005a:37078)
- [LS94] Douglas Lind and Klaus Schmidt, *Bernoullicity of solenoidal automorphisms and global fields*, Israel J. Math. **87** (1994), no. 1-3, 33–35. MR1286813 (95e:28013)
- [RB] Robert Rumely and Matthew H. Baker, *Analysis and dynamics on the berkovich projective line*, <http://arxiv.org/abs/math/0407433>, 1–150.
- [RL03] Juan Rivera-Letelier, *Dynamique des fonctions rationnelles sur des corps locaux*, Astérisque (2003), no. 287, xv, 147–230, Geometric methods in dynamics. II. MR2040006 (2005f:37100)
- [Rob00] Alain M. Robert, *A course in p -adic analysis*, Graduate Texts in Mathematics, vol. 198, Springer-Verlag, New York, 2000. MR1760253 (2001g:11182)
- [Sch84] W. H. Schikhof, *Ultrametric calculus*, Cambridge Studies in Advanced Mathematics, vol. 4, Cambridge University Press, Cambridge, 1984, An introduction to p -adic analysis. MR791759 (86j:11104)

- [Ser62] Jean-Pierre Serre, *Corps locaux*, Publications de l'Institut de Mathématique de l'Université de Nancago, VIII, Actualités Sci. Indust., No. 1296. Hermann, Paris, 1962. MR0150130 (27:133)
- [Sil08] C.E. Silva, *Invitation to ergodic theory*, Student Mathematical Library, vol. 42, American Mathematical Society, Providence, RI, 2008. MR2371216
- [Wal82] Peter Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York, 1982. MR648108 (84e:28017)
- [War06] Thomas Ward, *Review of "p-adic deterministic and random dynamical systems" by Andrei Khrennikov and Marcus Nilsson, Mathematics and its Applications Vol. 574, Kluwer Academic Publishers (2004)*, Bull. Amer. Math. Soc. (N.S.) **43** (2006), no. 1, 133–137.
- [WS98] Christopher F. Woodcock and Nigel P. Smart, *p-adic chaos and random number generation*, Experiment. Math. **7** (1998), no. 4, 333–342. MR1678087 (99k:11123)

DEPARTMENT OF MATHEMATICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MASSACHUSETTS 01267
E-mail address: `06jck@williams.edu`

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MASSACHUSETTS 02138
E-mail address: `alex.levin@post.harvard.edu`
Current address: Department of Mathematics, MIT, Cambridge, Massachusetts 02139
E-mail address: `levin@mit.edu`

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MASSACHUSETTS 02138
E-mail address: `preygel@post.harvard.edu`
Current address: Department of Mathematics, MIT, Cambridge, Massachusetts 02139
E-mail address: `preygel@mit.edu`

DEPARTMENT OF MATHEMATICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MASSACHUSETTS 01267
E-mail address: `csilva@williams.edu`