

# MEASURABLE DYNAMICS OF SIMPLE $p$ -ADIC POLYNOMIALS

JOHN BRYK AND CESAR E. SILVA

## 1. INTRODUCTION.

The  $p$ -adic numbers have many fascinating properties that are different from those of the real numbers. These properties are a consequence of the fact that the distance in the  $p$ -adics is measured using a non-Archimedean absolute value or norm. In this article we study the dynamics of algebraically defined transformations on the  $p$ -adics and see that there is a strong connection between the topological property of minimality, which is easy to check for such transformations, and the measure-theoretic property of ergodicity.

We start with a section that introduces the  $p$ -adic numbers, including their topology and the relevant measures on them, and then define the basic notions from dynamics that we require. In section 4 we show that minimal isometries on subsets of the  $p$ -adics are defined on finite unions of balls and are never totally ergodic. In Theorem 4.4 we give a new short proof that for an invertible isometry of a compact open subset of the  $p$ -adics minimality implies unique ergodicity and that ergodicity implies minimality. (As we discuss, this theorem is known in a more general context.) Thus for isometries on compact open subsets of the  $p$ -adics, the properties of minimality, ergodicity, and unique ergodicity are equivalent.

In later sections we study classes of transformations defined by multiplication, translation, and monomial mappings. We demonstrate that they are minimal (hence uniquely ergodic) on balls or, in some examples, spheres in  $\mathbb{Q}_p$ . Many of these results are known, but we present a unified and different treatment of them.

Dynamics on the  $p$ -adics has been the focus of several researchers recently. While most of this work deals with topological or complex dynamical properties of the  $p$ -adics, for which the reader may refer to [1], [4], [13], [17], and the references cited in these articles, measurable dynamics on the  $p$ -adics has also received some exposure, particularly in [14], [2], [3] [12], [7], and [8].

## 2. THE $p$ -ADIC NUMBERS.

Let  $p$  be a prime number. The construction of the field  $\mathbb{Q}_p$  of  $p$ -adic numbers is analogous to that of the real number field  $\mathbb{R}$ . However, whereas  $\mathbb{R}$  is constructed by completing the field  $\mathbb{Q}$  of rational numbers with respect to the usual absolute value, we construct the  $p$ -adic numbers by using the  $p$ -adic norm. We start with an informal description that illustrates the

main ideas. If  $n$  is a positive integer, then we can write  $n$  in the form  $n = \sum_{i=0}^N c_i p^i$ , where the  $c_i$  are unique integers in  $\{0, \dots, p-1\}$  and  $N \geq 0$ . We call this the  $p$ -adic representation of  $n$ . Suppose now that we wish to extend this representation to any rational number. We start with an example, considering the case of  $p = 5$  and the rational number  $-3/20$ . Then we can write

$$\begin{aligned} -\frac{3}{20} &= \frac{1}{5} \cdot \frac{3}{-4} = \frac{1}{5} \cdot \frac{3}{1-5} = \frac{3}{5} \cdot \frac{1}{1-5} = \frac{3}{5} (5^0 + 5^1 + 5^2 + \dots) \\ &= 3 \cdot 5^{-1} + 3 \cdot 5^0 + 3 \cdot 5^1 + 3 \cdot 5^2 + \dots = \sum_{i=-1}^{\infty} 3 \cdot 5^i. \end{aligned}$$

At the moment we are not going to worry about issues of convergence but just consider the sum in a formal way. (We used this example because the formal expansion for  $1/(1-5)$  is well known, but similar expansions can be obtained for any rational number.) There are a couple of things we notice from this example. First, to obtain “ $p$ -adic representations” of rational numbers it seems that we need to allow infinite series of powers of  $p$ , possibly starting at a negative power such as  $i = -1$ . Second, we have to make sense of the infinite sum. In other words, we need to consider some metric under which we can show convergence of the partial sums of the infinite series, or more precisely, a metric under which the partial sums form Cauchy sequences.

To continue our informal description we consider a set  $X$  consisting of all formal sums  $x$  of the type

$$x = \sum_{i=k}^{\infty} c_i p^i,$$

where  $k$  is an integer,  $c_i = c_i(x)$  belongs to  $\{0, 1, \dots, p-1\}$  for  $i = k, k+1, \dots$ , and  $c_k \neq 0$ . Add to this set the point 0 represented as  $\sum_{i=0}^{\infty} c_i x^i$  where  $c_i = 0$  for  $i \geq 0$ . Using the formal manipulations illustrated by our example (with a general prime  $p$  replacing 5) we can identify each rational number with a unique element of  $X$ . A standard diagonalization argument shows that the set  $X$  is uncountable, so it necessarily contains elements that do not come from members of  $\mathbb{Q}$ . To make sense of the infinite expansions we introduce a notion of convergence in  $X$  motivated by the fact that we desire sequences of partial sums of the form we have discussed to form Cauchy sequences. We define a distance so that two infinite formal sums are “close” if a “large” number of their initial terms are the same. More precisely, define the distance between the formal sums  $\sum_{i=k}^{\infty} c_i p^i$  and  $\sum_{i=\ell}^{\infty} d_i p^i$  to be  $p^{-\min\{k,\ell\}}$  if  $k \neq \ell$ , and otherwise define it to be  $p^{-m}$ , where  $m$  is the first integer such that  $m \geq k = \ell$  and  $c_m \neq d_m$  (of course, the distance is 0 if  $c_m = d_m$  for all  $m$ ). Then the distance between the partial sums  $S_n$  and  $S_{n+m}$  ( $m > 0$ ) is  $p^{-(n+1)}$  (identifying the partial sum  $S_n = \sum_{i=-1}^n c_i p^i$  with the series  $\sum_{i=-1}^{\infty} c_i p^i$  in which  $c_i = 0$  when  $i > n$  and assuming that

$c_{n+1} \neq 0$ ). Therefore, under this distance the partial sums of our example would form a Cauchy sequence.

We now change gears and formally define the  $p$ -adic metric on  $\mathbb{Q}$ . The obvious place to start is to define the  $p$ -adic norm  $|x|_p$  of a rational number  $x$ . To do this when  $x$  is nonzero, we write  $x$  in the form  $x = p^\alpha(r/s)$ , where  $p$  divides neither  $r$  nor  $s$ , and define

$$|x|_p = p^{-\alpha}.$$

(And, of course, we set  $|0|_p = 0$ .) In particular, if we express a natural number  $n$  using its  $p$ -adic representation,  $n = \sum_{i=k}^N c_i p^i$  where  $c_k \neq 0$ , then  $|n|_p = p^{-k}$ , which agrees with the distance we used for our example. One can easily verify that  $|-x|_p = |x|_p$  and  $|xy|_p = |x|_p |y|_p$ . The first interesting property we notice about this norm is that it satisfies the *strong triangle inequality*:

$$(1) \quad |x + y|_p \leq \max\{|x|_p, |y|_p\},$$

with equality holding if  $|x|_p \neq |y|_p$ . Normed fields that satisfy inequality (1) are called *non-Archimedean*. (It is not hard to see that the strong triangle inequality implies that for any integer  $n$  different from 1 and any nonzero  $x$  in  $\mathbb{Q}$  it is the case that  $|nx|_p \leq |x|_p$ , clearly contradicting the Archimedean property of the reals.) An interesting consequence of the non-Archimedean property is that a series of rational numbers  $\sum_{i=0}^{\infty} a_i$  converges in  $\mathbb{Q}_p$  if and only if  $\lim_{i \rightarrow \infty} |a_i|_p = 0$  (for an explanation of this property and other consequences the reader is referred to [6]). The  *$p$ -adic metric* is the metric defined on  $\mathbb{Q}$  by  $d_p(x, y) = |x - y|_p$ . This metric also satisfies the strong triangle inequality. Given  $p$ , we can complete  $\mathbb{Q}$  in the standard way to arrive at a normed field  $(\mathbb{Q}_p, |\cdot|_p)$  such that  $\mathbb{Q}$  is dense in  $\mathbb{Q}_p$ ,  $|\cdot|_p$  extends the  $p$ -adic norm from  $\mathbb{Q}$  to  $\mathbb{Q}_p$ , and every Cauchy sequence in  $\mathbb{Q}_p$  (with respect to  $|\cdot|_p$ ) is convergent.

One can show that each nonzero  $x$  in  $\mathbb{Q}_p$  is uniquely representable in the manner

$$x = \sum_{n=k}^{\infty} c_n p^n,$$

where  $k$  is an integer,  $c_n = c_n(x)$  belongs to  $\{0, 1, \dots, p-1\}$  for  $n = k, k+1, \dots$ ,  $c_k \neq 0$ , and convergence is with respect to the  $p$ -adic norm. In fact it is easy to see that  $|x|_p = p^{-k}$ . (Of course  $x = 0$  admits a trivial representation of this type, say with  $k = 0$  and  $c_n = 0$  for  $n = 0, 1, \dots$ ) Addition in  $\mathbb{Q}_p$  is defined termwise modulo  $p$  with a carry to the right. For example,  $1 + (\sum_{i=0}^{\infty} (p-1) \cdot p^i) = \sum_{i=0}^{\infty} p \cdot p^i = 0$ , so the additive inverse of 1 is  $\sum_{i=0}^{\infty} (p-1) \cdot p^i$ . Multiplication is defined in a similar way.

One special subset of  $\mathbb{Q}_p$  that we shall consider quite often is the *ring of  $p$ -adic integers*  $\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leq 1\}$ . It is not difficult to see that  $\mathbb{Z}_p$  is the closure of  $\mathbb{Z}$  in  $\mathbb{Q}_p$  and consists of precisely those  $x$  in  $\mathbb{Q}_p$  admitting

representations of the type

$$x = \sum_{n=0}^{\infty} c_n p^n$$

with  $c_n$  in  $\{0, \dots, p-1\}$ . The ring  $\mathbb{Z}_p$  is not a field ( $p$  does not have a multiplicative inverse in  $\mathbb{Z}_p$ ), but it is an integral domain with a unique prime ideal, namely,  $p\mathbb{Z}_p$ . For each integer  $k$  we can talk about the factor ring  $\mathbb{Z}_p/p^k\mathbb{Z}_p$ , a ring that is isomorphic to  $\mathbb{Z}/p^k\mathbb{Z}$ . If  $n$  satisfies  $\gcd(n, p) = 1$ , then  $[n]$  is a unit in  $\mathbb{Z}/p^k\mathbb{Z}$  (here and later  $[x]$  signifies the equivalence class in the relevant quotient structure of an element  $x$ ). The set of all units in  $\mathbb{Z}_p/p^k\mathbb{Z}_p$  is a multiplicative group  $(\mathbb{Z}_p/p^k\mathbb{Z}_p)^\times$  that is isomorphic to  $(\mathbb{Z}/p^k\mathbb{Z})^\times$ .

The topology of  $\mathbb{Q}_p$  is generated by the metric induced by the  $p$ -adic norm. Given  $y$  in  $\mathbb{Q}_p$  and  $r \geq 0$ , we define the *closed ball*  $B_r(y)$  of radius  $r$  centered at  $y$  by  $B_r(y) = \{x \in \mathbb{Q}_p : |x - y|_p \leq r\}$ , the corresponding *open ball*  $B_r^\circ(y)$  by  $B_r^\circ(y) = \{x \in \mathbb{Q}_p : |x - y|_p < r\}$ , and the *sphere* of radius  $r$  centered at  $y$  by  $S_r(y) = \{x \in \mathbb{Q}_p : |x - y|_p = r\}$ .

Note that  $B_1(0) = \mathbb{Z}_p$ . Moreover, observing that the image of  $\mathbb{Q}_p \setminus \{0\}$  under the mapping  $x \mapsto |x|_p$  is the discrete subset  $\{p^n : n \in \mathbb{Z}\}$  of  $\mathbb{R}$ , we see that closed and open balls of positive radius are equivalent, in the sense that  $B_r(y) = B_{r+\epsilon}^\circ(y)$  for all sufficiently small  $\epsilon > 0$ . Thus, for the remainder of the paper, whenever we say *ball* we mean a closed ball of positive radius. Furthermore, we assume that any radius  $r$  that occurs has the form  $r = p^{-\ell}$  for some integer  $\ell$ , which we may do because  $B_r(y) = B_{p^{-\ell}}(y)$  when  $p^{-\ell} \leq r < p^{-\ell+1}$ . For the same reasons,  $S_r(y)$  is nonempty if and only if  $r = p^{-\ell}$  for some  $\ell$ . One can verify that spheres, as well as closed and open balls, are both compact and open sets in the  $p$ -adic topology. (As in any metric space, a set  $S$  is *compact* if every open cover of  $S$  has a finite subcover, and  $S$  is *locally compact* if each point of  $S$  has a compact neighborhood.) Finally, every point inside of a ball is a “center” of that ball—that is, if  $B$  is a ball of radius  $r$ , then  $B_r(x) = B$  for all  $x$  in  $B$ . It follows that if two balls  $B$  and  $B'$  of equal radius have a nonempty intersection, then  $B = B'$ , a fact that we use in several of our proofs.

We conclude this section with a fascinating property of spheres that does not hold in the case of the reals: any sphere in  $\mathbb{Q}_p$  can be written as a finite union of pairwise disjoint balls! In fact, let  $x$  be in  $\mathbb{Q}_p$  and consider the sphere  $S = S_r(x)$  of radius  $r = p^n$  for some integer  $n$ . Clearly,  $S$  is contained in  $B_r(x)$ . Let  $s = p^{n-1}$ . We first show that  $B_r(x)$  can be written as the union of  $p$  disjoint balls of radius  $s$ . To see this note first that  $\mathbb{Z}_p (= B_1(0))$  is the disjoint union of the  $p$  cosets that constitute  $\mathbb{Z}_p/p\mathbb{Z}_p$ , and it is easy to see that each of these cosets is a subball of  $\mathbb{Z}_p$  of radius  $1/p$ . Since  $B_r(x) = x + p^{-n}\mathbb{Z}_p$ , we find that  $B_r(x)$  is the disjoint union of  $p$  subballs of radius  $s$ —namely, the images of the cosets of  $p\mathbb{Z}_p$  under the same dilation and translation. Let the centers of these  $p$  disjoint balls be  $y_i$ ,

with  $y_1 = x$ . Consider the set  $B^* = B_r(x) \setminus B_s(x)$ . Then  $B^*$  is the union of the  $p - 1$  pairwise disjoint balls  $B_s(y_i)$  with  $y_i$  not in  $B_s(x)$ . Now we claim that  $B^* = S_r(x)$ . Indeed, if  $y$  is in  $B^*$ , then  $|x - y|_p \leq r$  but  $|x - y|_p > s$ . Therefore  $|x - y|_p = r$  and  $y$  is in  $S_r(x)$ . For the converse, let  $y$  be a point in  $S_r(x)$ . Then  $y$  is in  $B_r(x)$ , but since  $|x - y|_p = r > s$ ,  $y$  is not in  $B_s(x)$ . Therefore  $y$  is in  $B^*$  and  $S_r(x) = B^*$ , showing that  $S_r(x)$  is the finite union of pairwise disjoint balls.

The  $p$ -adic numbers were introduced by Kurt Hensel in the late nineteenth century (1897) in order to study the properties of algebraic numbers. Hensel's original description of the  $p$ -adic numbers involved an analogy between the ring of integers and the ring of polynomials over the complex numbers, the crux of which was the development of a representation of rational numbers analogous to that of Laurent expansions of rational functions—namely, the  $p$ -adic expansion. The  $p$ -adic numbers have played a fundamental role in number theory, and are now finding wider applications in science (see, for example, [13]).

### 3. TOPOLOGICAL AND MEASURABLE DYNAMICS.

In its simplest form a (discrete time, abstract) dynamical system consists of a set  $X$  and a map or transformation  $T : X \rightarrow X$ . We will mainly be interested in the case when the transformation  $T$  is invertible. In fact, we mention only one example that is noninvertible (in section 4) and one lemma (Lemma 4.5) that treats the noninvertible case. Thus absent an explicit statement to the contrary, the reader may assume that any transformation encountered in this paper is invertible.

One of the first objects of interest is the *orbit* of point a point  $x$  in  $X$  under  $T$ , defined as  $\{T^n(x)\}_{n=-\infty}^{\infty}$  and denoted by  $\mathcal{O}(x)$  (here  $T^n$  denotes the  $n$ -fold iterate of  $T$ ,  $T^n = T \circ \dots \circ T$ , with  $n$  factors,  $T^0$  is the identity mapping of  $X$ , and  $T^{-n} = (T^{-1})^n$  for a positive integer  $n$ ).

To study properties of such iterations we impose some structure on the space and the transformations. In topological dynamics we require  $X$  to be a topological space and  $T$  to be a continuous transformation; in the interesting cases we assume that  $X$  is compact. In the invertible case  $T$  is assumed to be a homeomorphism. The important notion from topological dynamics for us is minimality. A transformation  $T : X \rightarrow X$  is *minimal* if the orbit  $\mathcal{O}(x)$  of each  $x$  in  $X$  is dense in  $X$ . It is clear from the definition that a homeomorphism  $T$  is minimal if and only if  $\bigcup_{n=-\infty}^{\infty} T^n(U) = X$  for each nonempty open set  $U$  in  $X$ . Using Zorn's lemma, it can be shown that for any homeomorphism  $T$  of a compact space  $X$  there exists a nonempty closed subset  $E$  of  $X$  such that  $E$  is  *$T$ -invariant* (i.e.,  $T(E) = E$ ) and the restriction of  $T$  to  $E$  is minimal. Minimal sets such as  $E$  were introduced in 1912 by Birkhoff in his study of the stability of dynamical systems [5, vol. 1, p. 660]. It is clear that a minimal homeomorphism is “indecomposable”

in the sense that it does not admit any nonempty proper closed invariant subset.

Before we can consider notions from measurable dynamics (or ergodic theory), we need to introduce a measurable structure on  $\mathbb{Q}_p$ . For this we first produce a (real-valued) measure  $\mu$  on  $\mathbb{Q}_p$ . This measure is defined in a way that is analogous to the definition of Lebesgue measure on  $\mathbb{R}$ . We start by declaring the *measure of a ball* to be equal to its radius. That is, we define  $\mu(B_r(x)) = r$  for each  $x$  in  $\mathbb{Q}_p$  and  $r$  in  $\{p^n : n \in \mathbb{Z}\}$ . Mimicking what happens in the real case, we define the outer measure  $\mu^*(A)$  of any subset  $A$  of  $\mathbb{Q}_p$  as one would expect:

$$(2) \quad \mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(B_n) : A \subset \bigcup_{n=1}^{\infty} B_n \right\},$$

where the infimum is taken over all countable collections of balls  $\{B_n\}_{n \in \mathbb{N}}$  covering  $A$ . It is an exercise to show that  $\mu^*(B) = \mu(B)$  for each ball  $B$ .

We say that a set  $A$  is *measurable* (or *Lebesgue measurable*) if for each  $\epsilon > 0$  there exists an open set  $G$  containing  $A$  such that  $\mu^*(G \setminus A) < \epsilon$ . It is clear that open sets are measurable, and one can show that the collection of measurable sets is closed under the operations of taking countable unions and complements, and hence forms a  $\sigma$ -algebra  $\mathcal{L}$  of subsets of  $\mathbb{Q}_p$ . The outer measure  $\mu^*$  when restricted to  $\mathcal{L}$  is a translation invariant measure that we denote by  $\mu$ . (Those with some experience in the area will realize that Lebesgue measure  $\mu$  is nothing but Haar measure on the locally compact group  $\mathbb{Q}_p$ .) The proofs of these properties are analogous to those in the case of the development of Lebesgue measure in the reals as in, for example, [16].

Let  $X$  be a compact subset of  $\mathbb{Q}_p$  with  $\mu(X) > 0$  (while one may also consider more general sets, such as infinite measure sets, our treatment of ergodicity here is restricted to the compact case, so  $\mu(X) < \infty$ ). A transformation  $T : X \rightarrow X$  is said to be a *measurable transformation* if  $T^{-1}(A)$  belongs to  $\mathcal{L}(X) = \{E \cap X : E \in \mathcal{L}\}$  whenever  $A$  does. Let  $\nu$  be a probability measure defined on  $\mathcal{L}(X)$ . (In all of our examples  $\nu$  will be normalized Lebesgue measure, i.e.,  $\nu(A) = \mu(A)/\mu(X)$  for all  $A$  in  $\mathcal{L}(X)$ .) We say that  $T$  is *measure-preserving* for  $\nu$  if  $T$  is measurable and  $\nu(A) = \nu(T^{-1}(A))$  holds for all  $A$  in  $\mathcal{L}(X)$ . (In this case we also say that the measure  $\nu$  is *invariant under  $T$*  or  *$T$ -invariant*.) The transformations that we consider (with the exception of the single Bernoulli example of section 4) are, either by assumption or by construction, invertible with both  $T$  and  $T^{-1}$  measurable and measure-preserving for the relevant measure.

To introduce relevant concepts from measurable dynamics we use the simple notion of an invariant set. We note that the existence of a nonempty proper invariant set  $A$  allows one to decompose the dynamics of  $T$  into simpler systems, for  $T|_A$ , the restriction of  $T$  to  $A$ , generates a dynamical system in its own right. Now let  $\nu$  be a probability measure defined on  $\mathcal{L}(X)$ , and let  $T : X \rightarrow X$  be a measure-preserving transformation with respect to

$\nu$ . The transformation  $T$  is *ergodic* with respect to  $\nu$  if  $\nu(A)$  is 0 or 1 for any  $T$ -invariant set  $A$  in  $\mathcal{L}(X)$ . (This notion was first called *metrical transitivity* and was introduced in 1928 by Birkhoff and Smith [5, vol. 2, p. 380].) From the point of view of measurable dynamics the ergodic transformations are the “indecomposable” transformations. The reason that this notion is important stems from the ergodic theorem, proved in 1931, first in the case of mean convergence by von Neumann and then in the stronger case of pointwise convergence by Birkhoff. (The reader may refer to [5, vol. 2, pp. 462-465] for a short history, to [22] for a modern account, and to [20] or [21] for the proofs of these theorems.) The pointwise convergence version of the ergodic theorem asserts that ergodicity is equivalent to the property that

$$(3) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{I}_A(T^n(x)) = \nu(A)$$

holds for each measurable set  $A$  in  $X$  and for each point  $x$  of  $X$  outside of a set of measure zero (in general depending on  $A$ ). (Here  $\mathbb{I}_A$  denotes the characteristic function of the set  $A$ .) In the full statement of the theorem  $\mathbb{I}_A$  is replaced by an integrable function  $f$  and  $\nu(A)$  is replaced by  $\int_A f d\nu$ ; however, this more general statement follows from (3) by standard approximation results. The quantity on the left of (3) is clearly (the limit of) the average number of times that images of  $x$  under iterates of  $T$  land in  $A$ . The ergodic theorem then asserts that for any measurable set  $A$  and for any point  $x$  outside some set  $E = E(A)$  of measure zero, the average number of times that the forward orbit of  $x$  visits  $A$  tends to the measure of  $A$  or, in other words, that the time-average is asymptotically equal to the space-average for all points outside some set of measure zero. This theorem solved an important problem that came from statistical mechanics: to give a condition under which the time-average and the space-average of a dynamical system would agree. While verification of the ergodicity property for actual physical systems has proved to be difficult, the notion of ergodicity has found numerous applications in mathematics. This is partly due to the fact that every finite-measure-preserving transformation admits a decomposition into “ergodic components.” Roughly stated, if  $\nu$  is an invariant measure for a transformation  $T$ , then  $\nu$  can be written as an integral of measures each of which is defined on some invariant subset of  $X$  and is ergodic with respect to  $T$ . (For the case of some of our examples this decomposition is illustrated in Figures 1, 2, and 3.) This theorem has a long history, starting with work of von Neumann in 1932 and culminating with independent developments by Maharam and Rohlin, for which the reader may refer to [15, sec. 8] and [20, chap. 3], respectively.

We often use another condition equivalent to ergodicity that is easier to check. Assume that  $T$  is invertible. Clearly the set  $\bigcup_{i=-\infty}^{\infty} T^i(A)$  is  $T$ -invariant for any given subset  $A$  of  $X$ . Therefore, if  $A$  has positive measure and  $T$  is ergodic, then  $\nu(X \setminus \bigcup_{i=-\infty}^{\infty} T^i(A)) = 0$ . It follows that if  $T$  is

ergodic, then for each pair of sets  $A$  and  $B$  of positive measure there exists an integer  $n$  such that  $\nu(T^n(A) \cap B) > 0$  (i.e., sets of positive measure “hit” other sets of positive measure under iteration by  $T$ ). Without too much additional work the reader can show that the converse is also true and that in fact  $n$  can be assumed to be positive. In general it does not follow from the ergodicity of  $T$  that  $T^2$  is also ergodic. If it happens that  $T^n$  is ergodic for all  $n$  different from 0 we say that  $T$  is *totally ergodic*.

One last notion of significance for what follows is that of “unique ergodicity,” which we define only when  $X$  is a compact subset of  $\mathbb{Q}_p$ . A homeomorphism  $T : X \rightarrow X$  is *uniquely ergodic* if there exists only one  $T$ -invariant probability measure  $\nu$  defined on  $\mathcal{L}(X)$ . (In all of our examples this measure  $\nu$  will be normalized Lebesgue measure on  $\mathcal{L}(X)$ .) Interestingly, this already implies that  $T$  is ergodic with respect to  $\nu$ . For if  $T$  were not ergodic with respect to  $\nu$ , one could construct another  $T$ -invariant probability measure  $\bar{\nu}$  different from  $\nu$ . Indeed, suppose that  $Y$  were a  $T$ -invariant set in  $\mathcal{L}(X)$  with  $\nu(Y)\nu(Y^c) > 0$ . Then the probability measure  $\bar{\nu}$  defined on  $\mathcal{L}(X)$  by  $\bar{\nu}(A) = \nu(A \cap Y)/\nu(Y)$  would be a  $T$ -invariant probability measure different from  $\nu$ , a contradiction. Therefore  $T$  must be ergodic with respect to  $\nu$ . In the uniquely ergodic case one has a refinement of the ergodic theorem. In fact,  $T$  is uniquely ergodic if and only for all continuous functions  $f : X \rightarrow \mathbb{R}$ ,

$$(4) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n(x)) = \int f d\nu,$$

where the convergence is uniform for all  $x$  in  $X$  (see [15, sec. 5.3]).

#### 4. MINIMAL ISOMETRIES AND PERMUTATIONS ON BALLS.

All of the transformations  $T$  we consider have the property that

$$|T(x) - T(y)|_p = |x - y|_p$$

for all  $x$  and  $y$  in some compact subspace  $X$  of  $\mathbb{Q}_p$  with  $\mu(X) > 0$ . (Any such  $T$  is said to be an *isometry with respect to  $|\cdot|_p$* .) For example, the transformation  $T : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  given by  $T(x) = x + 1$  is clearly an isometry.

Consider  $B_r(x)$  for any  $x$  in  $\mathbb{Q}_p$ . If  $T$  is an invertible isometry, then  $y$  belongs to  $B_r(T(x))$  if and only if  $T^{-1}(y)$  lies in  $B_r(x)$ . It follows that  $T(B_r(x)) = B_r(T(x))$ , so the image of a ball is a ball of equal size. This implies that  $T$  is measure-preserving for Lebesgue measure.

An important property that will surface in all of our examples is minimality. We warn the reader that there is another commonly used dynamical concept that is related to but weaker than minimality—it is called topological transitivity—where one requires only that the orbit of *some* point be dense. However, as one might expect, an invertible isometry that is topologically transitive turns out to be minimal [21, p. 131]. Since all of our examples are invertible isometries, and since showing that they are minimal

is in general no harder than showing that they are topologically transitive, we discuss only minimality. The following has a simple proof, but we state it as a theorem for future reference.

**Theorem 4.1.** *If  $T : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  is defined by  $T(x) = x + 1$ , then  $T$  is minimal.*

*Proof.* Since the class [1] generates each of the finite groups  $\mathbb{Z}_p/p^k\mathbb{Z}_p$ , given any  $\epsilon > 0$  and any points  $x$  and  $y$  of  $\mathbb{Z}_p$  we can find  $n$  such that  $|T^n(x) - y|_p = |x + n - y|_p \leq p^{-k} < \epsilon$ .  $\square$

We remark that while we proved that the *forward orbit*  $\{T^n(x)\}_{n=0}^\infty$  (often denoted  $\mathcal{O}_+(x)$ ) of every point  $x$  is dense, this is no stronger than minimality since for a homeomorphism  $T$  of a compact space  $X$  every orbit of  $T$  is dense if and only if every forward orbit of  $T$  is dense [21, p. 129].

Now minimality turns out to be an extremely useful property for our purposes: it is easy to verify for many algebraic transformations. In this section we first show in Lemma 4.2 that, if  $T : X \rightarrow X$  is a minimal isometry of a subset  $X$  of  $\mathbb{Q}_p$  that contains a nonempty open set, then  $X$  must be a finite union of balls, hence a compact open set. In particular, this rules out minimal isometries  $T : X \rightarrow X$  for subsets  $X$  of  $\mathbb{Q}_p$  that contain open sets and have  $\mu(X) = \infty$ . Because the transformations of interest to us are algebraic, requiring that their domains of definition contain open sets is not an unreasonable restriction.

Theorem 4.4 is the real engine of this paper. It states that for an invertible isometry of a compact open subset of  $\mathbb{Q}_p$  minimality is equivalent to unique ergodicity. Now, it is already well known that minimal isometries of compact metric spaces must be uniquely ergodic (see [15, sec. 5.8]). However, our proof of this fact for compact open subsets of  $\mathbb{Q}_p$  is much simpler than the standard proof. Also, in our context, we give a short proof that ergodicity implies minimality, hence unique ergodicity—a theorem also known in a more general setting [15].

We note in passing that minimality and unique ergodicity are distinct properties. For example, as shown by Furstenberg [9], there are minimal homeomorphisms of the two-dimensional torus that are not uniquely ergodic. Also, while a uniquely ergodic map is minimal on the support of its invariant measure [15, sec. 5.2], there exist uniquely ergodic homeomorphisms (on the unit circle in the complex plane for example) that are not minimal (see [21]). Furthermore, it is easy to construct ergodic isometries that are neither minimal nor uniquely ergodic. For example, let  $X = S_p(0) \cup \{0\}$ , and let  $T : X \rightarrow X$  be defined so that the restriction of  $T$  to  $S_p(0)$  is any ergodic isometry on  $S_p(0)$  (such as the maps of Theorem 7.2) and  $T(0) = 0$ . Let  $\nu$  be the probability measure defined on measurable subsets  $A$  of  $S_p(0)$  by  $\nu(A) = \mu(A)/\mu(S_p(0))$  and extend  $\nu$  to a measure on  $X$  by setting  $\nu(\{0\}) = 0$  (i.e.,  $\nu(A) = \nu(A \cap S_p(0))$ ). Then  $T$  is an ergodic isometry of  $X$  with respect to  $\nu$  that is not minimal and is not uniquely ergodic. (Note, however, that  $\nu$  is not positive on open sets of  $X$ , so does not satisfy Theorem 4.4(3).)

We start with a lemma that is of independent interest.

**Lemma 4.2.** *Let  $X$  be a subset of  $\mathbb{Q}_p$ , let  $T : X \rightarrow X$  be an invertible isometry, and let  $x$  be a point of  $X$ . If  $\mathcal{O}_+(x)$  is dense in  $X$ , then  $X$  is contained in a finite union of balls, hence  $X$  is of finite  $\mu$ -measure. Furthermore, if  $X$  contains some ball  $B$  around  $x$  of positive radius  $r$ , then  $X = \bigcup_{i=0}^{n-1} T^i(B)$  for some  $n > 0$ ; in particular,  $X$  is a compact open subset of  $\mathbb{Q}_p$ .*

*Proof.* Let  $r > 0$ . Then there exists a smallest  $n > 0$  such that  $d(T^n(x), x) < r$ . Thus  $x$  lies in  $B_r(T^n(x)) \cap X$  and  $B_r(T^n(x)) \cap B_r(x) \neq \emptyset$ . Therefore  $B_r(T^n(x)) = B_r(x)$ , which ensures that the set  $A = \bigcup_{i=0}^{n-1} B_r(T^i(x)) \cap X$  is  $T$ -invariant. Certainly  $A$  is contained in  $X$ . If  $z$  were a point of  $X \setminus A$ , then  $d(T^i(x), z) > r$  would hold for all nonnegative  $i$ , so  $z$  would not be a limit point of the forward orbit of  $x$ . This would be a contradiction. We conclude that  $X = A$ .

Now if  $B = B_r(x)$  lies in  $X$ , then we have  $X = \bigcup_{i=0}^{n-1} B_r(T^i(x))$ . To complete the proof we note that any two balls of equal radius must be either disjoint or equal and that  $B_r(T^i(x)) = T^i(B)$ .  $\square$

The reader may want to verify that every compact open subset of  $\mathbb{Q}_p$  is a finite unions of balls. We do not need this fact explicitly. However, in what follows one could substitute “compact open set” for “finite union of balls.” We also obtain the following interesting corollary:

**Corollary 4.3.** *Let  $X$  be a nonempty compact open set in  $\mathbb{Q}_p$  and let  $T : X \rightarrow X$  be a minimal invertible isometry. Then for any ball  $B$  of positive radius contained in  $X$  there exists an integer  $n > 0$  such that  $T^n(B) = B$ . Furthermore, if  $X = B_{p^{-\ell}}(a)$  for some  $a$  in  $\mathbb{Q}_p$  and  $\ell$  in  $\mathbb{Z}$ , then for each  $x$  in  $B_{p^{-\ell}}(a)$  and each  $j$  with  $j > \ell$  it is the case that  $T^{p^{j-\ell}}(B_{p^{-j}}(x)) = B_{p^{-j}}(x)$ . In particular,  $T$  is not totally ergodic for Lebesgue measure.*

*Proof.* We prove only the second part. From the proof of Lemma 4.2 we learn that there is an integer  $n$  such that

$$(5) \quad B_{p^{-\ell}}(x) = B_{p^{-j}}(x) \cup \dots \cup B_{p^{-j}}(T^{n-1}(x))$$

and such that the balls in the union are pairwise disjoint. We infer that  $T^n(B_{p^{-j}}(x)) = B_{p^{-j}}(x)$ , so  $T^n$  is not ergodic. Using the fact that the sets on each side of equation (5) have the same measure, we obtain  $p^{-\ell} = np^{-j}$  or, equivalently,  $n = p^{j-\ell}$ .  $\square$

We are now ready for the main theorem of this section.

**Theorem 4.4.** *Let  $X$  be a nonempty compact open set in  $\mathbb{Q}_p$ , and let  $T : X \rightarrow X$  be an invertible isometry. The following statements are equivalent:*

- (1)  $T : X \rightarrow X$  is minimal.
- (2)  $T : X \rightarrow X$  is uniquely ergodic.

- (3)  $T$  is ergodic for any  $T$ -invariant probability measure  $\nu$  on  $\mathcal{L}(X)$  that is positive on nonempty open sets (such as normalized Lebesgue measure).

*Proof.* (1)  $\Rightarrow$  (2): Let  $\bar{\mu}$  be the normalized Lebesgue measure on  $X$ . Since  $T$  is an isometry,  $\bar{\mu}$  is a  $T$ -invariant measure. Suppose now that  $T$  is minimal on  $X$ , and let  $\nu$  be any  $T$ -invariant probability measure on  $X$ . Consider a ball  $B$  of positive radius  $r$  contained in  $X$ . By Lemma 4.2,  $X$  can be written in the form  $X = \bigcup_{i=0}^{n-1} T^i(B)$ , where the sets in the union are pairwise disjoint and  $n > 0$  is an integer. Since  $\bar{\mu}$  is a  $T$ -invariant probability measure, this yields  $1 = n\bar{\mu}(B)$ . Similarly,  $1 = n\nu(B)$ . Then  $\bar{\mu}(B) = 1/n = \nu(B)$ . Thus  $\bar{\mu}$  and  $\nu$  agree on all balls contained in  $X$  and therefore  $\nu = \bar{\mu}$ , showing that  $T$  is uniquely ergodic. (The fact that  $\nu = \bar{\mu}$  can be seen by noting that any such measure  $\nu$  satisfies an equality similar to (2).)

(2)  $\Rightarrow$  (3): As we have already seen, if  $T$  is uniquely ergodic, then  $T$  is ergodic with respect to its unique invariant probability measure on  $\mathcal{L}(X)$ .

(3)  $\Rightarrow$  (1): Suppose that  $T$  is ergodic on  $X$  with respect to a probability measure  $\nu$  that is positive on nonempty open sets. Let  $\epsilon > 0$ , and consider balls  $B_1 = B_r(x)$  and  $B_2 = B_r(y)$  in  $X$  with  $r < \epsilon/2$ . As  $T$  is ergodic with respect to  $\nu$ , there exists  $n > 0$  such that  $\nu(T^n(B_1) \cap B_2) > 0$ . Since  $T$  is an isometry,  $T^n(B_1)$  is a ball of radius  $r$  centered at  $T^n(x)$ . Thus by the triangle inequality,  $|T^n(x) - y|_p \leq 2r < \epsilon$ . As this holds for all  $\epsilon > 0$ ,  $y$  is a limit point of  $\mathcal{O}_+(x)$ . Because  $x$  and  $y$  in  $X$  are arbitrary, this shows that  $T$  is minimal.  $\square$

Theorem 3.2 holds in the more general setting where  $X$  is a compact metric space (say with metric  $d$ ) and  $T : X \rightarrow X$  is an equicontinuous map, as was pointed out to the authors by J. Auslander. (A map  $T$  is said to be *equicontinuous* if for each  $\epsilon > 0$  there exists  $\delta > 0$  such that  $d(T^n(x), T^n(y)) < \epsilon$  holds for all  $n$  and for all  $x$  and  $y$  in  $X$  satisfying  $d(x, y) < \delta$ .) In this case the proof that (3)  $\Rightarrow$  (1) is the same as the one given for Theorem 4.4, (2)  $\Rightarrow$  (3) is immediate, and the implication (1)  $\Rightarrow$  (2) was demonstrated by Oxtoby [15, sec. 5.8]. Isometries are clearly equicontinuous; furthermore, it can be shown that any equicontinuous map  $T : X \rightarrow X$  is an isometry for some metric on  $X$  that is topologically equivalent to the given metric.

Once we have deduced that a map  $T$  is an invertible isometry on  $\mathbb{Q}_p$ , it is simply a matter of finding the sets on which it is minimal (provided they exist). Various criteria help us decide where to look. If it happens that  $d(x, T(x)) = r$  for all  $x$  in  $\mathbb{Q}_p$ , then we focus on balls of radius  $r$ . This is the case, for example, for the map  $T(x) = x + 1$ , where clearly  $|x - T(x)|_p = 1$ . Therefore we study this map on balls of radius 1, such as  $\mathbb{Z}_p$ . On the other hand, if  $T$  has a fixed point  $y$  (i.e.,  $T(y) = y$ ), then we have that  $|x - y|_p = |T(x) - T(y)|_p = |T(x) - y|_p$ . In this case all the points of the orbit of  $x$  are equidistant from  $y$ , so we focus our attention on spheres around  $y$ . An example is given by the map  $M_a : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  defined

by  $M_a(x) = ax$ , which has a fixed point at 0. So if  $M_a$  is an isometry (which is only possible if  $|a|_p = 1$ ), then we look at spheres around 0. In the end, showing that  $T$  is minimal on any of these spaces relies on the algebraic and number theoretic qualities of the  $p$ -adics. We intend to establish minimality only for a few distinguished maps and then deduce minimality for other maps using the notion of isomorphism that we study in section 5.

It is well known that isometries  $T : X \rightarrow X$  on compact spaces  $X$  cannot be weakly mixing [20, p. 54]. Recall that a transformation  $T$  preserving a probability measure  $\nu$  is *weakly mixing* if the product transformation  $T \times T : X \times X \rightarrow X \times X$  is ergodic with respect to the product measure  $\nu \times \nu$ . This implies, in particular, that  $T$  is totally ergodic. It is interesting to note that, in view of Corollary 4.3, a minimal isometry in  $\mathbb{Q}_p$  is not even totally ergodic (with respect to its unique invariant measure). This leads to the question: Can any probability-measure-preserving transformations on  $\mathbb{Q}_p$  defined by rational polynomial functions be weakly mixing (or even mixing) on an invariant subset of  $\mathbb{Q}_p$  having positive measure? In this context we note that if we are willing to consider noninvertible transformations, one can easily construct mixing, hence weakly mixing, examples on  $\mathbb{Z}_p$ . (A transformation  $T$  that preserves a finite measure  $\nu$  is *mixing* if for all measurable sets  $A$  and  $B$  with  $\nu(A) > 0$  it is true that  $\lim_{n \rightarrow \infty} \nu(T^{-n}(A) \cap B) / \nu(A) = \nu(B)$ ; in other words, the relative proportion of space that  $T^{-n}(A)$  occupies in an arbitrary set  $B$  is in the limit the measure of  $B$ .) These mixing transformations are the analogues of the Bernoulli transformations and are defined on  $\mathbb{Z}_p$  by  $T(0.c_0c_1c_2\dots) = 0.c_1c_2\dots$  (alternatively, by  $T(x) = px - c_0(x)p^{-1}$ , where  $c_0(x)$  is the coefficient of  $p^0$  in the representation of  $x$ .) Then  $T$  is a noninvertible finite-measure-preserving transformation that can be shown to be mixing by standard methods similar to those used to establish mixing for the Bernoulli transformations on  $[0, 1)$ . (The main part of the argument is to verify that for any balls  $B_1 = B_{p^{-\ell}}(c)$  and  $B_2 = B_{p^{-k}}(d)$  and for any  $n > k$  it is the case that  $\mu(T^{-n}(B_1) \cap B_2) = \mu(B_1)\mu(B_2)$ .)

This section concludes with a lemma that clarifies the role of invertibility in our presentation. In particular, it follows from the lemma that a measure-preserving isometry on  $\mathbb{Z}_p$  must be invertible, that transformations on  $\mathbb{Z}_p$  satisfying the Lipschitz condition in the lemma cannot be totally ergodic when they are invertible (Corollary 4.3), and that any noninvertible measure-preserving transformation on  $\mathbb{Z}_p$  must involve some sort of “expansion” with respect to the  $p$ -adic distance (such as the Bernoulli example). Lemma 4.5 is stated without proof in [3, Theorem 1.1], where the author also studies in detail transformations satisfying the Lipschitz condition of the lemma.

**Lemma 4.5.** *Let  $T : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  be a transformation satisfying a Lipschitz condition with constant 1:  $|T(x) - T(y)|_p \leq |x - y|_p$  for all  $x$  and  $y$  in  $\mathbb{Z}_p$ . Then  $T$  is surjective if and only if  $T$  preserves Lebesgue measure, in which case  $T$  is an invertible isometry.*

*Proof.* Suppose that  $T$  is surjective. Let  $n \geq 0$ , let  $x$  be a point in  $\mathbb{Z}_p$ , and write  $r = p^{-n}$ . Clearly,  $T(B_r(x)) \subseteq B_r(T(x))$ . (This is a direct consequence of the Lipschitz condition, since if  $z$  is in  $T(B_r(x))$ , then  $z = T(a)$  for some  $a$  in  $B_r(x)$  and  $|z - T(x)|_p = |T(a) - T(x)|_p \leq |a - x|_p \leq r$ .) Now suppose that  $T(B_r(x))$  is properly contained in  $B_r(T(x))$ , and let  $z$  be a point of  $B_r(T(x)) \setminus T(B_r(x))$ . Since  $T$  is surjective, there exists  $a$  in  $\mathbb{Z}_p$  such that  $T(a) = z$ . As  $T(a)$  is in  $B_r(T(x))$ ,  $B_r(T(x)) = B_r(T(a))$ . However, because  $a$  is not in  $B_r(x)$ , it follows that  $B_r(x) \cap B_r(a) = \emptyset$ . Now we can write  $\mathbb{Z}_p$  as a union pairwise disjoint balls in the form  $\mathbb{Z}_p = \bigcup_{i=1}^{p^n} B_r(x_i)$  for some points  $x_i$  in  $\mathbb{Z}_p$ , with  $x_1 = x$  and  $x_2 = a$ , say. Then the set  $T(\mathbb{Z}_p) = \bigcup_{i=1}^{p^n} T(B_r(x_i)) = \bigcup_{i=1}^{p^n} B_r(T(x_i))$  must be contained in at most  $p^n - 1$  balls of radius  $r$  (as the terms in the union corresponding to  $i = 1$  and  $i = 2$  are equal), contradicting the assumption that  $T$  is surjective. Therefore  $T(B_r(x)) = B_r(T(x))$ . Because this is true for each  $x$  in  $X$  and for  $r = p^{-n}$  for each positive integer  $n$ ,  $T$  is an isometry. Injectivity is an immediate consequence of the fact that  $T$  is an isometry. It also follows that  $T$  preserves Lebesgue measure.

For the converse suppose that  $T$  is not surjective. Then there exists  $x$  in  $\mathbb{Z}_p \setminus T(\mathbb{Z}_p)$ . Since  $T$  is measure-preserving,

$$\mu(\mathbb{Z}_p \setminus T(\mathbb{Z}_p)) = \mu(T^{-1}(\mathbb{Z}_p \setminus T(\mathbb{Z}_p))) = \mu(\emptyset) = 0.$$

Therefore  $\mathbb{Z}_p \setminus T(\mathbb{Z}_p)$  does not contain any ball of positive radius. This implies the existence of a sequence  $\{x_n\}$  of points in  $T(\mathbb{Z}_p)$  converging to  $x$ . Select  $y_n$  in  $\mathbb{Z}_p$  with  $T(y_n) = x_n$ . Taking a subsequence  $y_{n_i}$  converging to some point  $y$  in  $\mathbb{Z}_p$ , we obtain  $T(y) = x$  (using the continuity of  $T$ ), a contradiction. Therefore  $T$  is surjective.  $\square$

## 5. ISOMORPHISMS.

Let  $X_1$  and  $X_2$  be finite unions of balls in  $\mathbb{Q}_p$ . Let  $T_i : X_i \rightarrow X_i$  ( $i = 1, 2$ ) be a continuous transformation. A map  $\phi : X_1 \rightarrow X_2$  is called a *topological isomorphism* of  $(X_1, T_1)$  and  $(X_2, T_2)$  if it is a homeomorphism of  $X_1$  onto  $X_2$  and the equality  $\phi(T_1(x)) = T_2(\phi(x))$  holds for all  $x$  in  $X_1$  (i.e.,  $\phi$  conjugates the actions of  $T_1$  and  $T_2$ ). One can easily verify that minimality is preserved under topological isomorphism.

If  $\nu_1$  and  $\nu_2$  are probability measures on  $\mathcal{L}(X_1)$  and  $\mathcal{L}(X_2)$ , respectively, a *measurable isomorphism* of the measurable dynamical systems  $(X_1, T_1, \nu_1)$  and  $(X_2, T_2, \nu_2)$  is a map  $\phi : X_1 \rightarrow X_2$  that is invertible and measure-preserving (i.e.,  $\nu_1(A) = \nu_2(\phi(A))$  for all sets  $A$  in  $\mathcal{L}(X_1)$ ) and has the property that  $\phi(T_1(x)) = T_2(\phi(x))$  for all  $x$  in  $X_1$  outside a subset of  $X_1$  of zero  $\nu_1$ -measure. Again, it is not hard to show that ergodicity is invariant under measurable isomorphism.

Now, we observe that in our case any topological isomorphism is a measurable isomorphism, so we henceforth refer to such maps simply as “isomorphisms.” In fact, as  $T_1$  and  $T_2$  are minimal isometries (hence, uniquely

ergodic by Theorem 4.4), if  $\nu_1$  and  $\nu_2$  are their invariant probability measures on  $\mathcal{L}(X_1)$  and  $\mathcal{L}(X_2)$ , respectively, and if  $\phi : X_1 \rightarrow X_2$  is a topological isomorphism, then the measure  $\nu_1 \circ \phi^{-1}$  is an invariant measure for  $T_2$ . It must therefore coincide with  $\nu_2$ . This shows that  $\phi$  is measure-preserving, making it a measurable isomorphism.

In our examples, the isomorphism property will follow from a standard lemma whose proof we include for the sake of completeness.

**Lemma 5.1.** *Let  $X_1$  and  $X_2$  be finite unions of balls in  $\mathbb{Q}_p$ , let  $T_1 : X_1 \rightarrow X_1$  and  $T_2 : X_2 \rightarrow X_2$  be continuous transformations, and let  $\phi : X_1 \rightarrow X_2$  be a surjective map such that  $\phi(T_1(x)) = T_2(\phi(x))$  for all  $x$  in  $X_1$ . If there is a constant  $C > 0$  such that*

$$(6) \quad |\phi(x) - \phi(y)|_p = C|x - y|_p$$

for all  $x$  and  $y$  in  $X_1$ , then  $\phi$  is an isomorphism of  $(X_1, T_1)$  and  $(X_2, T_2)$ . Under these circumstances  $T_1$  is an invertible isometry if and only if  $T_2$  is an invertible isometry, and  $T_1$  is minimal (hence, uniquely ergodic) if and only if  $T_2$  is minimal (hence, uniquely ergodic).

*Proof.* It is easy to see that (6) implies that  $\phi$  is one-to-one and continuous and that (6) holds for  $\phi^{-1}$  (with  $C^{-1}$  replacing  $C$ ), so  $\phi$  is an isomorphism. Now, if  $T_1$  is an isometry, then

$$\begin{aligned} |T_2(\phi(x)) - T_2(\phi(y))|_p &= |\phi(T_1(x)) - \phi(T_1(y))|_p \\ &= C|T_1(x) - T_1(y)|_p = C|x - y|_p \\ &= CC^{-1}|\phi(x) - \phi(y)|_p = |\phi(x) - \phi(y)|_p, \end{aligned}$$

whence by the surjectivity of  $\phi$  the map  $T_2$  is likewise an isometry. Noting that  $T_2^n = \phi^{-1} \circ T_1^n \circ \phi$  completes the proof.  $\square$

## 6. TRANSLATIONS ON $\mathbb{Q}_p$ .

We start our study of special maps by looking at the translation  $T_b : \mathbb{Q}_p \rightarrow \mathbb{Q}_p$  defined by  $T_b(x) = x + b$  for fixed  $b$  in  $\mathbb{Q}_p$ . We first consider the case when  $|b|_p \leq 1$ , in which event  $T_b : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ . We later show that the other cases can be reduced to this one using isomorphisms. (We note that  $T_1 : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$  is isomorphic to a well-known transformation, the so-called dyadic odometer.) The reader is referred to Anashin [2], [3] for a study of measure-preserving and ergodic maps on  $\mathbb{Z}_p$ .

**Theorem 6.1.** *If  $b$  belongs to  $\mathbb{Z}_p$  and  $T_b : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  is defined by  $T_b(x) = x + b$ , then  $T_b$  is an invertible isometry. Furthermore,  $T_b$  is minimal and uniquely ergodic on  $\mathbb{Z}_p$  if and only if  $|b|_p = 1$ .*

*Proof.* It is clear that  $T_b$  is invertible and is an isometry, for

$$|T_b(x) - T_b(y)|_p = |x + b - y - b|_p = |x - y|_p.$$

Assume that  $|b|_p = 1$  and define  $\phi : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  by  $\phi(x) = bx$ . Then

$$\phi(T_1(x)) = \phi(x + 1) = b(x + 1) = bx + b = T_b(\phi(x)).$$

Also,  $|\phi(x) - \phi(y)|_p = |b|_p|x - y|_p$ , which implies that  $\phi$  satisfies an equation of type (6). Finally, as  $|b|_p = 1$ ,  $b$  is invertible in  $\mathbb{Z}_p$  (see, for example, [18, p.5]) and therefore  $\phi$  is surjective. Combining Theorems 4.1 and 4.4 with Lemma 5.1, we conclude that  $T_b$  is minimal and uniquely ergodic on  $\mathbb{Z}_p$ .

For the converse, assume that  $T_b$  is minimal but that  $|b|_p < 1$ . Write  $b = p^\ell c$ , where  $|c|_p = 1$  and  $\ell > 0$ . Then by the first part of the proof  $T_c$  is minimal on  $\mathbb{Z}_p$ . By Corollary 4.3,

$$T_c^{p^\ell}(B_{p^{-\ell}}(0)) = B_{p^{-\ell}}(0).$$

On the other hand,  $T_c^{p^\ell}(x) = x + p^\ell c = x + b$ , which contradicts the minimality of  $T_b$ . Therefore  $|b|_p = 1$ .  $\square$

**Corollary 6.2.** *If  $p$  and  $q$  are different primes,  $b$  is in  $\mathbb{Z}_p$ , and  $c$  is in  $\mathbb{Z}_q$ , then  $(\mathbb{Z}_p, T_b)$  is not isomorphic to  $(\mathbb{Z}_q, T_c)$ .*

*Proof.* From Theorem 6.1 it follows that  $T_b^p = T_{pb}$  is not minimal on  $\mathbb{Z}_p$ , whereas  $T_c^p = T_{pc}$  is minimal on  $\mathbb{Z}_q$ . Therefore the two systems are not isomorphic.  $\square$

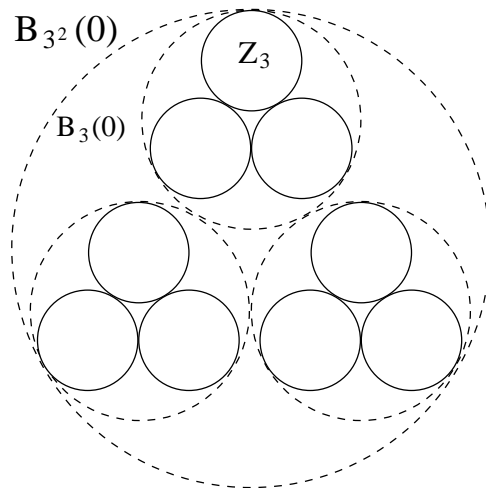


FIGURE 1. Ergodic decomposition of  $T_b$  ( $|b|_3 = 1$ ) on  $\mathbb{Q}_3$ .

We now characterize translations  $T_b : X \rightarrow X$  for subsets  $X$  of  $\mathbb{Q}_p$  and  $b$  in  $\mathbb{Q}_p$ . We consider only the case when  $X$  contains an open set and  $T_b$  is minimal on  $X$ . Invoking Lemma 4.2, we write  $X$  as a pairwise disjoint union of balls  $X = B_r(a) \cup \dots \cup B_r(T_b^{n-1}(a))$  for some  $a$  in  $X$ , with  $n \geq 1$  and  $r > 0$ . Next consider the ball  $B_s(a)$  of radius  $s = |b|_p$  and observe that  $B_s(a)$  is a  $T_b$ -invariant set. If it were the case that  $r < s$ , then  $B_s(a)$  would be properly contained in  $B_r(a)$ , contradicting the minimality of  $T$ . Hence we may assume that  $r \leq s$ . Now if it were the case that  $s < r$ , then  $B_s(a)$  would contain  $B_r(a)$ , and again as it is  $T_b$ -invariant, it would have to contain

$X$ . Finally we observe that Corollary 6.3 shows that  $T_b$  is minimal on  $B_s(a)$ . Since  $X$  is  $T$ -invariant, this will imply that  $X$  is equal to  $B_s(a)$ . Thus it suffices to analyze  $T_b$  on balls of radius  $s = |b|_p$ . We use isomorphisms again to obtain our next result, which together with Corollary 4.3 can be used to obtain the ergodic decomposition illustrated in Figure 1.

**Corollary 6.3.** *If  $b$  in  $\mathbb{Q}_p$  has  $|b|_p = p^{-\ell}$  ( $\ell \in \mathbb{Z}$ ) and  $T_b$  is defined by  $T_b(x) = x + b$ , then for each  $a$  in  $\mathbb{Q}_p$  the map  $T_b : B_{p^{-\ell}}(a) \rightarrow B_{p^{-\ell}}(a)$  is an invertible isometry. Furthermore,  $T_b$  is minimal and uniquely ergodic on  $B_{p^{-\ell}}(a)$ .*

*Proof.* We define a map  $\phi : \mathbb{Z}_p \rightarrow B_{p^{-\ell}}(a)$  by  $\phi(x) = bx + a$ . It is easy to see that this map is invertible, and we have

$$|\phi(x) - \phi(y)|_p = |bx + a - by + a|_p = |b|_p|x - y|_p.$$

Moreover,

$$\begin{aligned} \phi(T_1(x)) &= \phi(x + 1) = b(x + 1) + a \\ &= bx + a + b = \phi(x) + b = T_b(\phi(x)), \end{aligned}$$

showing that  $\phi$  satisfies the hypotheses of Lemma 5.1. By Theorem 4.1  $T_1$  is minimal, and uniquely ergodic, on  $\mathbb{Z}_p$ . Lemma 5.1 asserts that  $T_b$  is an invertible isometry that is minimal and uniquely ergodic on  $B_{p^{-\ell}}(a)$ .  $\square$

It is interesting to compare translations on  $\mathbb{Z}_p$  with their real analogs. Translations on the entire real line are plainly not ergodic with respect to Lebesgue measure. Translations by rational numbers on, say,  $[0, 1)$  with endpoints identified (i.e., on  $\mathbb{R}/\mathbb{Z}$ , the reals modulo 1) are periodic and not ergodic. On the other hand, translations on  $\mathbb{R}/\mathbb{Z}$  by irrational numbers are totally ergodic and minimal (see, for example, [20] or [21]). This stands in sharp contrast to the situation in  $\mathbb{Q}_p$ , where *all* translations are ergodic on certain subsets of  $\mathbb{Q}_p$ , but none are totally ergodic.

## 7. MULTIPLICATION IN $\mathbb{Q}_p$ .

While the first polynomials we have considered have direct analogs in the reals as minimal transformations, the ones we consider for the remainder of this article take advantage of special properties of  $\mathbb{Q}_p$ . In this section we consider the transformation  $M_a : \mathbb{Q}_p \rightarrow \mathbb{Q}_p$  defined by  $M_a(x) = ax$  for some fixed member  $a$  of  $\mathbb{Q}_p$ , a multiplication (or dilation) by an element of  $\mathbb{Q}_p$ .

The first study of ergodic properties of transformations on the  $p$ -adics that we are aware of is in [14], where Oselies and Zieschang considered the continuous automorphisms of the group  $\mathbb{Z}_p$ . These automorphisms are of the form  $M_a$  with  $a$  in  $S_1(0)$ , and in [14] Oselies and Zieschang obtained their ergodic decomposition on  $\mathbb{Z}_p$ . The multiplication transformations  $M_a$  acting on the group of units  $S_1(0)$  were first studied by Coelho and Parry in [7]. In particular they proved that  $M_a$  is ergodic on  $S_1(0)$  if and only

if  $a$  is a primitive root modulo  $p$  and  $a^{p-1} \not\equiv 1 \pmod{p^2}$ . This is similar to Theorem 7.2, but we note that for  $\ell \neq 0$  the spheres  $S_{p^{-\ell}}$  are not multiplicative groups. The motivation of Coelho and Parry was to use these transformations to obtain information on the distribution of sequences such as the Fibonacci sequence, which we describe later in this section.

To make  $M_a$  a measure-preserving transformation, and in fact an isometry, it is clear that we require  $|a|_p = 1$ .

One difference between these last two sections and the previous sections is that we now demonstrate minimality of our transformations on spheres rather than on balls. However, we briefly return to balls in Theorem 8.5.

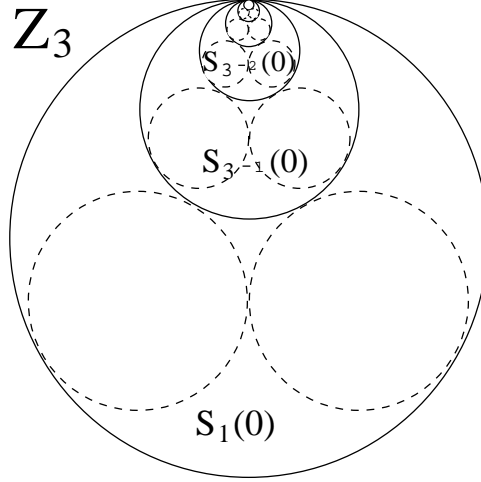
In any metric space an isometry with a fixed point defines a dynamical system on spheres centered at this point, for these spheres are invariant sets for the isometry. For example, because 0 is a fixed point of  $M_a$ ,

$$|M_a(x) - 0|_p = |ax|_p = |a|_p|x|_p = |x - 0|_p,$$

so we see that  $M_a$  is a well-defined bijective transformation of the spheres  $S_{p^{-\ell}}(0)$  for each integer  $\ell$ . In contrast with the real case, where spheres in  $\mathbb{R}^n$  have measure zero with respect to  $n$ -dimensional Lebesgue measure, in the case of the  $p$ -adics spheres have positive measure—in fact, a sphere of radius  $r$  about a point  $x$  has measure greater than or equal to the open ball of radius  $r$  centered there. Thus we can use the Lebesgue measure that we defined for the  $p$ -adics to analyze dynamical systems on these spheres.

We observe immediately that there are certain choices of  $a$  that guarantee that  $M_a$  is not minimal on any sphere. For example, when  $p \geq 3$  the ring  $\mathbb{Z}_p$  contains the  $(p-1)$ th roots of unity, so  $M_a^{p-1}$  is the identity transformation for each such root  $a$ . This means that  $M_a$  is periodic and thus minimal only on finite sets. Also, recall that for  $p = 3$  we may write the sphere  $S_1(0)$  as a disjoint union of balls  $S_1(0) = B_{1/3}(1) \cup B_{1/3}(2)$ . Consider  $M_a : \mathbb{Z}_3 \rightarrow \mathbb{Z}_3$  with  $a$  in  $B_{1/3}(1)$ . Then  $a$  has the form  $a = 1 \cdot p^0 + c_1 \cdot p + c_2 \cdot p^2 + \dots$  for some  $c_0, c_1, \dots$ . If  $x$  is a point in  $B_{1/3}(1)$ , then it is straightforward to verify that  $ax$  is also in  $B_{1/3}(1)$ . It follows that  $B_{1/3}(1)$  (and, likewise,  $B_{1/3}(2)$ ) is an invariant set under  $M_a$ , so  $M_a$  is not minimal on  $S_1(0)$ . (In this case we could refine our analysis to treat subsets of spheres, as illustrated in Figure 2, but we will deal here only with transformations that are ergodic on spheres. For a complete classification of the ergodic decomposition of  $M_a$ , consult [7].)

To see what restrictions we need on  $a$  we focus on  $S_1(0)$  for the moment. If  $M_a$  is minimal on this sphere, then every integer that is not divisible by  $p$  is a limit point of the orbit of 1, the set  $\{a^n\}_{n=0}^\infty$ . Let  $\epsilon = p^{-k}$  with  $k$  in  $\mathbb{N}$ , and let  $x$  in  $\mathbb{Z}$  be such that  $p$  does not divide  $x$ . Thus there must exist  $n \geq 0$  such that  $|a^n - x|_p < p^{-k}$  or, in other words, such that  $a^n \equiv x \pmod{p^k}$  (i.e.,  $a^n - x$  is in  $p^k\mathbb{Z}_p$ ). Since this holds for every  $x$  in  $\mathbb{Z}$  not divisible by  $p$ , we see that  $[a]$  generates the multiplicative group of units  $(\mathbb{Z}_p/p^k\mathbb{Z}_p)^\times$  for each  $k$  in  $\mathbb{N}$ . In fact, this is exactly the property that we want  $a$  to satisfy. Furthermore, there is an easy way to characterize the  $a$  that enjoy

FIGURE 2. Ergodic decomposition of  $M_a$  on  $\mathbb{Q}_3$ .

this property using the following well-known lemma from group theory (for a proof of the first part the reader may refer to [19, Theorem 6.7], and for the second part to [19, Theorem 5.44]).

**Lemma 7.1.** *Let  $p$  be an odd prime, and let  $a$  be a member of  $\mathbb{Z}_p$ . Then  $[a]$  generates the group of units  $(\mathbb{Z}_p/p^k\mathbb{Z}_p)^\times$  for each  $k$  in  $\mathbb{N}$  if and only if  $[a]$  generates the group of units  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$ . Furthermore,  $(\mathbb{Z}_2/2^k\mathbb{Z}_2)^\times$  is not cyclic when  $k \geq 3$ .*

Notice that Lemma 7.1 implies that  $M_a$  can never be minimal (hence never uniquely ergodic) on the sphere  $S_1(0)$  of  $\mathbb{Z}_2$ . Thus we bid farewell to the even prime and consider only odd primes  $p$  for the rest of this paper. We are now ready for our characterization of the minimality of  $M_a$ . As mentioned earlier, a theorem similar to Theorem 7.2 (for  $\ell = 0$ ) was proved in [7].

**Theorem 7.2.** *Let  $p$  be an odd prime, let  $a$  in  $\mathbb{Q}_p$  have  $|a|_p = 1$ , and let  $M_a : S_{p^{-\ell}}(0) \rightarrow S_{p^{-\ell}}(0)$  with  $\ell$  in  $\mathbb{Z}$  be defined by  $M_a(x) = ax$ . Then  $[a]$  generates  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$  if and only if  $M_a$  is minimal and uniquely ergodic on  $S_{p^{-\ell}}(0)$ .*

*Proof.* Assume initially that  $[a]$  generates  $(\mathbb{Z}_p/p^{k-\ell}\mathbb{Z}_p)^\times$ . Fix  $x$  in  $S_{p^{-\ell}}(0)$ . We want to show that, for any  $y$  in  $S_{p^{-\ell}}(0)$  and any  $\epsilon > 0$ , there exists  $n$  such that  $|M_a^n(x) - y|_p = |a^n x - y|_p < \epsilon$ . Since  $\ell$  is fixed, we may assume that  $\epsilon \leq p^{-\ell}$ . Choose  $k$  such that  $p^{-k} < \epsilon$ . Then  $k > \ell$ , so  $k - \ell > 0$ . Note that  $|a^n x - y|_p = |x|_p |a^n - yx^{-1}|_p$  for all  $n$ . The fact that  $|yx^{-1}|_p = |y|_p/|x|_p = 1$ , shows that  $yx^{-1}$  lies on  $S_1(0)$ . Thus  $[yx^{-1}]$  belongs to  $(\mathbb{Z}_p/p^{k-\ell}\mathbb{Z}_p)^\times$ . Since  $[a]$  generates  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$ , by Lemma 7.1  $[a]$  must generate  $(\mathbb{Z}_p/p^{k-\ell}\mathbb{Z}_p)^\times$ , meaning that there exists a positive integer  $n$  such

that  $a^n \equiv yx^{-1} \pmod{p^{k-\ell}}$  or, stated differently, that  $|a^n - yx^{-1}|_p < p^{\ell-k}$ . This leads to the conclusion that

$$|M_a^n(x) - y|_p = |x|_p |a^n - yx^{-1}|_p < p^{-\ell} p^{\ell-k} = p^{-k} < \epsilon.$$

Therefore  $M_a$  is minimal and, by Theorem 4.4, uniquely ergodic on  $S_{p^{-\ell}}(0)$ .

For the converse, suppose that  $[a]$  does not generate  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$ . Then there exists  $b$  in  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$  such that  $a^n$  is not congruent to  $b$  modulo  $p^2$  for any integer  $n$ . Now let  $x = p^\ell$  and  $y = bp^\ell$  with  $\ell$  in  $\mathbb{Z}$ . Clearly  $x$  and  $y$  are points of  $S_{p^{-\ell}}(0)$ , as  $|b|_p = 1$ . Also  $M_a^n(x) = a^n p^\ell$ . We have

$$|M_a^n(x) - y|_p = |a^n p^\ell - bp^\ell|_p = p^{-\ell} |a^n - b|_p \geq p^{-\ell-1}$$

for each nonnegative integer  $n$ , revealing that  $y$  is not a limit point of the orbit of  $x$ . Since  $x$  does not have a dense orbit in  $S_{p^{-\ell}}(0)$ ,  $M_a$  is not minimal on this space.  $\square$

A theorem equivalent to Theorem 7.2 is used in an intriguing way in [7] to give the distribution in  $\mathbb{Z}_p$  of the Fibonacci sequence. We outline the main ideas to give a flavor of the argument in [7]. We consider only the case when  $\sqrt{5}$  is an element of  $\mathbb{Z}_p$  (which, by quadratic reciprocity and a direct application of Hensel's lemma, which is stated in section 8, occurs when  $p \equiv 1$  or  $4 \pmod{5}$ ); when 5 is not a square modulo  $p$  one considers as in [7] an appropriate subgroup of  $\mathbb{Z}_p[\sqrt{5}]$ . Now let  $\{f_n\}$  be the Fibonacci sequence (i.e.,  $f_{n+1} = f_{n-1} + f_n$ ,  $f_0 = 0, f_1 = 1$ ). The idea is to find a function  $f$  and a number  $\alpha$  in  $S_1(0)$  so that the sequence  $f_n$  can be expressed in the form  $f_n = f \circ M_\alpha^n(1)$ . Once this is achieved it is possible to apply the ergodic theorem to the function  $f$  and the transformation  $M_\alpha$ . In the case of the Fibonacci sequence we obtain  $\alpha$  from Binet's formula: it is well known that  $f_n = (1/\sqrt{5})[\alpha^n - (-1/\alpha)^n]$  for  $\alpha = (1 + \sqrt{5})/2$ . The required function  $f$  is defined by  $f(z) = (1/\sqrt{5})[z - N(z)/z]$ , where  $N(z) = 1$  if  $z$  is a square modulo 5 and  $N(z) = -1$  otherwise. One can verify that indeed  $f_n = f(\alpha^n) = f \circ M_\alpha^n(1)$ . We further assume that  $M_\alpha$  is ergodic (if it is not ergodic, one must deal with the ergodic decomposition of  $M_\alpha$ , which is described in [7]). Using the fact that  $M_\alpha$  is uniquely ergodic on  $S_1(0)$ , we conclude that for any open and closed subset  $A$  of  $\mathbb{Z}_p$ ,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^N \mathbb{I}_A(f_n) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^N \mathbb{I}_A(f \circ M_\alpha^n(1)) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^N \mathbb{I}_{f^{-1}(A)}(M_\alpha^n(1)) \\ &= \int \mathbb{I}_{f^{-1}(A)} d\mu = \mu(f^{-1}(A)), \end{aligned}$$

which gives us the distribution of the Fibonacci sequence mod  $p^k$  in terms of the measure  $\mu \circ f^{-1}$ . In [7] the authors discuss this measure in detail.

The next natural transformation we consider is the affine mapping  $T_{a,b} : \mathbb{Q}_p \rightarrow \mathbb{Q}_p$  defined by  $T_{a,b}(x) = ax + b$  for  $a$  and  $b$  in  $\mathbb{Q}_p$  with  $|a|_p = 1$  and  $b \neq 0$ . The first thing to notice is that this transformation has a fixed point (assuming  $a \neq 1$ : when  $a = 1$  these transformations are translations, which we covered earlier) at  $c = b/(1 - a)$ . Thus we analyze the dynamical properties of  $T_{a,b}$  on the spheres  $S_{p^{-\ell}}(c)$  centered at  $c$ . We start by showing that affine transformations are in fact isomorphic to the multiplications we have already discussed and then obtain information on the dynamics of  $T_{a,b}$ .

**Theorem 7.3.** *Let  $p$  be an odd prime, and let  $a$  and  $b$  be points of  $\mathbb{Q}_p$  such that  $|a|_p = 1$  but  $a \neq 1$ . If  $T_{a,b}$  and  $M_a$  are transformations defined by  $T_{a,b}(x) = ax + b$  and  $M_a(x) = ax$ , respectively, and if  $c = b/(1 - a)$  is the fixed point of  $T_{a,b}$ , then  $(S_{p^{-\ell}}(c), T_{a,b})$  is isomorphic to  $(S_{p^{-\ell}}(0), M_a)$  for each integer  $\ell$ . Furthermore,  $T_{a,b} : S_{p^{-\ell}}(c) \rightarrow S_{p^{-\ell}}(c)$  is minimal and uniquely ergodic on  $S_{p^{-\ell}}(c)$  if and only if  $[a]$  generates  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$ .*

*Proof.* Define  $\phi : S_{p^{-\ell}}(c) \rightarrow S_{p^{-\ell}}(0)$  by  $\phi(x) = x - c$ . This map is certainly invertible. Furthermore, if  $x$  is a point of  $S_{p^{-\ell}}(c)$ , then

$$\begin{aligned} M_a(\phi(x)) &= ax - ac = ax - ab/(1 - a) \\ &= ax - (ab - b)/(1 - a) - b/(1 - a) = ax + b - c \\ &= \phi(T_{a,b}(x)). \end{aligned}$$

Finally, it is clear that  $|\phi(x) - \phi(y)|_p = |x - y|_p$ . An appeal to Lemma 5.1 shows that the two dynamical systems in question are isomorphic. To complete the proof we apply Theorem 7.2.  $\square$

We note that an interesting property of the  $p$ -adic numbers follows from Theorem 7.3.

**Corollary 7.4.** *Let  $[a]$  generate  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$ , where  $a$  is an element of  $\mathbb{Q}_p$  with  $|a|_p = 1$  but  $a \neq 1$ , let  $b$  be a point of  $\mathbb{Z}_p$ , and let  $c = b/(1 - a)$ . If  $x$  lies on  $S_1(c)$ , then the set*

$$\left\{ a^n x + b \sum_{i=0}^{n-1} a^i \right\}_{n=0}^{\infty}$$

*has exactly  $p^{k-1}(p - 1)$  distinct elements modulo  $p^k$ , and it is equal to  $S_1(c)$  modulo  $p^k$ .*

*Proof.* The hypotheses ensure that  $a \not\equiv 1 \pmod{p}$ . Therefore  $|1 - a|_p = 1$ ,  $c$  belongs to  $\mathbb{Z}_p$ , and  $S_1(c)$  is contained in  $\mathbb{Z}_p$ . We invoke Theorem 7.3 to complete the proof.  $\square$

## 8. SIMPLE POLYNOMIAL TRANSFORMATIONS IN $\mathbb{Q}_p$ .

For our last examples we restrict attention to the monomial mappings  $M_{a,n} : \mathbb{Q}_p \rightarrow \mathbb{Q}_p$  defined by  $M_{a,n}(x) = ax^n$  for  $n$  in  $\mathbb{N}$ . (These are simple generalizations of the maps  $\psi_n(x) = x^n$  that are studied in [12]). We assume

in addition that  $p$  is an odd prime,  $a$  lies in  $B_{p^{-1}}(1)$ , and  $n \not\equiv 1 \pmod{p}$ . The reasons for imposing these restrictions will soon become apparent.

We first note that, although 0 is a fixed point of  $M_{a,n}$ , it is immediate that we do not want to consider the action of  $M_{a,n}$  on arbitrary spheres centered at 0. If  $|x|_p < 1$ , then  $M_{a,n}^i(x) \rightarrow 0$  as  $i \rightarrow \infty$ , so  $M_{a,n}$  cannot be an isometry on any space containing two such points. If  $|x|_p > 1$ , then  $|M_{a,n}^i(x)|_p \rightarrow \infty$  as  $i \rightarrow \infty$ , so any invariant set containing a ball around  $x$  must have infinite measure. If  $M_{a,n}$  were to be isometric, this would prevent  $M_{a,n}$  from being minimal.

However,  $S_1(0)$  is an invariant subspace, since for all  $x$  in  $S_1(0)$

$$|M_{a,n}(x)|_p = |ax^n|_p = |a|_p |x|_p^n = 1.$$

Furthermore, we can show that, modulo certain restrictions on  $n$ ,  $M_{a,n}$  is an isometry on this sphere.

Our approach is to use isomorphisms to reduce  $M_{a,n}$  to a map that we have already considered. In order to do this we first use Hensel's lemma to establish the existence of a fixed point for  $M_{a,n}$ . Hensel's lemma is one of the early results in  $p$ -adic analysis and can be viewed as a version of Newton's method (for a proof the reader may consult [11] or [18]).

**Lemma 8.1 (Hensel's Lemma).** *Let  $f(x)$  belong to  $\mathbb{Z}_p[x]$ . Suppose that there exists  $\alpha_0$  in  $\mathbb{Z}_p$  such that  $|f(\alpha_0)|_p < |f'(\alpha_0)|_p^2$ , where  $f'(x)$  signifies the formal derivative of  $f(x)$ . Then there exists a unique  $\alpha$  in  $\mathbb{Z}_p$  such that  $f(\alpha) = 0$  and  $|\alpha - \alpha_0|_p = |f(\alpha_0)|_p / |f'(\alpha_0)|_p$ .*

**Lemma 8.2.** *Let  $p$  be an odd prime, let  $n$  be a positive integer such that  $n \not\equiv 1 \pmod{p}$ , and let  $a$  be a point of  $B_{p^{-1}}(1)$ . If  $M_{a,n} : \mathbb{Q}_p \rightarrow \mathbb{Q}_p$  is defined by  $M_{a,n}(x) = ax^n$ , then there exists a unique point  $x_0$  of  $B_{p^{-1}}(1)$  such that  $M_{a,n}(x_0) = x_0$ .*

*Proof.* Consider the polynomial  $f(x) = ax^n - x$  in  $\mathbb{Z}_p[x]$ . Its formal derivative is given by  $f'(x) = nax^{n-1} - 1$ . We want to apply Lemma 8.1 with  $\alpha_0 = 1$ . We see that  $f(1) = a - 1$ ,  $|a - 1|_p \leq p^{-1}$ , and  $f'(1) = na - 1$ . Since  $na \not\equiv 1 \pmod{p}$ ,  $|na - 1|_p = 1$ . Thus

$$|f(1)|_p \leq p^{-1} < 1 = |f'(1)|_p^2,$$

so by Hensel's lemma there exists a unique  $x_0$  in  $\mathbb{Z}_p$  such that  $f(x_0) = 0$  (in other words, such that  $M_{a,n}(x_0) = ax_0^n = x_0$ ) and such that  $|x_0 - 1|_p = |f(1)|_p / |f'(1)|_p \leq p^{-1}$ . Therefore  $x_0$  is the unique fixed point of  $M_{a,n}$  in  $B_{p^{-1}}(1)$ .  $\square$

We next show that  $M_{a,n}$  is isomorphic to  $T_{n,a'}$  for a value of  $a'$  to be specified later. The isomorphism that we use is the  $p$ -adic logarithm, which can be defined by its power series expansion:

$$\log_p x = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} (x-1)^i,$$

where convergence is with respect to  $|\cdot|_p$ . The region of convergence of the series is  $B_{p^{-1}}(1)$ . (We note that since  $B_{p^{-1}}(1) = 1 + p\mathbb{Z}_p$ , it follows that  $B_{p^{-1}}(1)$  is a multiplicative subgroup of  $\mathbb{Z}_p^\times = \{x \in \mathbb{Z}_p : |x|_p = 1\}$ .) The  $p$ -adic logarithm retains some of the properties of its real analog:  $\log_p(xy) = \log_p(x) + \log_p(y)$  and  $\log_p(x^n) = n \log_p x$  for all  $x$  and  $y$  in  $B_{p^{-1}}(1)$  and all integers  $n$ . We will also use the fact that  $\log_p : B_{p^{-1}}(1) \rightarrow B_{p^{-1}}(0)$  is a bijective isometry (see [18, sec. 4.4.2], [12]). We chose  $a$  in  $B_{p^{-1}}(1)$  and study the action of  $M_{a,n}$  on spheres in  $B_{p^{-1}}(1)$ .

**Lemma 8.3.** *Let  $p$  be an odd prime, let  $n$  be a nonnegative integer such that  $n \not\equiv 1 \pmod{p}$ , and let  $a$  be a point of  $B_{p^{-1}}(1)$ . If  $x_0$  is the unique fixed point of  $M_{a,n}$  in  $B_{p^{-1}}(1)$ , then  $(S_{p^{-\ell}}(x_0), M_{a,n})$  is isomorphic to  $(S_{p^{-\ell}}(\log_p x_0), T_{n, \log_p a})$  for  $\ell = 1, 2, \dots$ . In particular,  $M_{a,n}$  is an invertible isometry on  $S_{p^{-\ell}}(x_0)$ .*

*Proof.* Consider  $\log_p : S_{p^{-\ell}}(x_0) \rightarrow S_{p^{-\ell}}(\log_p x_0)$ . This is a bijective isometry (i.e.,

$$|\log_p(x) - \log_p(y)|_p = |x - y|_p,$$

for  $x$  and  $y$  on  $S_{p^{-\ell}}(x_0)$ ). Now consider  $T_{n, \log_p a} : S_{p^{-\ell}}(\log_p x_0) \rightarrow S_{p^{-\ell}}(\log_p x_0)$ . For  $x$  on  $S_{p^{-\ell}}(x_0)$  we compute

$$\log_p(M_{a,n}(x)) = \log_p(ax^n) = n \log_p x + \log_p a = T_{n, \log_p a}(\log_p x).$$

By Lemma 5.1,  $\log_p$  is an isomorphism.  $\square$

The following theorem, including the equivalence of minimality, ergodicity, and unique ergodicity, was proved by Gundlach, Khrennikov, and Lindahl [12] for the case of the transformation  $M_{1,n}$  (i.e.,  $a = 1$ ,  $x_0 = 1$ ). As also noted in [12], it is interesting to observe that the spheres  $S_{p^{-\ell}}(x_0)$  are not multiplicative groups.

**Theorem 8.4.** *Let  $p$  be an odd prime, let  $n$  be a nonnegative integer such that  $n \not\equiv 1 \pmod{p}$ , let  $a$  be a point of  $B_{p^{-1}}(1)$ , and let  $x_0$  be the unique fixed point of  $M_{a,n}$  in  $B_{p^{-1}}(1)$ . Then for  $\ell = 1, 2, \dots$  the transformation  $M_{a,n} : S_{p^{-\ell}}(x_0) \rightarrow S_{p^{-\ell}}(x_0)$  is minimal and uniquely ergodic if and only if  $[n]$  generates  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$ .*

*Proof.* Using the isomorphism from Lemma 8.3 in Lemma 5.1, it follows that  $M_{a,n}$  is an invertible isometry on  $S_{p^{-\ell}}(x_0)$ . Furthermore,  $M_{a,n}$  is minimal and uniquely ergodic if and only if  $T_{n, \log_p a}$  is minimal and uniquely ergodic, which by Theorem 7.3 occurs if and only if  $[n]$  generates  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$ .  $\square$

While we do not discuss this in detail, the ergodic decomposition of  $M_{a,n}$  is illustrated in Figure 3. In closing, we discuss briefly the case when  $M_{a,n}$  has at least one other nonzero fixed point, call it  $x_1$ . Then  $x_1$  is a point of  $S_1(0)$ , for  $|x_1|_p = |ax_1^n|_p = |a|_p|x_1|_p^n = 1$ . One can use an isomorphism to show that  $M_{a,n}$  is ergodic on the spheres centered at  $x_1$  that are contained within  $B_{p^{-1}}(x_1)$ . Indeed, an isomorphism  $\phi : S_{p^{-\ell}}(x_1) \rightarrow S_{p^{-\ell}}(x_0)$  is defined by  $\phi(x) = xx_0x_1^{-1}$ . It is also possible to use isomorphisms to characterize  $M_{a,n}$

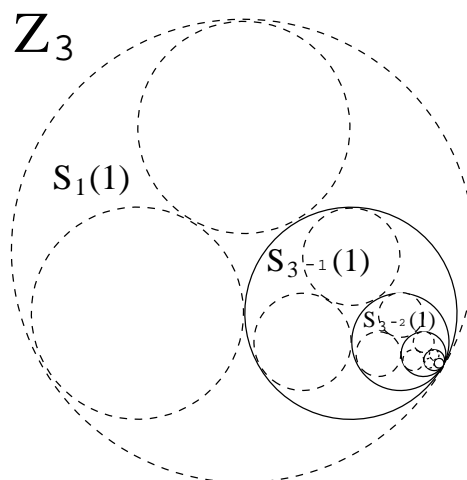


FIGURE 3. Ergodic decomposition of  $M_{a,n}$  ( $a = 1$ ) on  $\mathbb{Z}_3$ .

on spheres around other fixed points for the case of values of  $a$  outside the ball  $B_{p^{-1}}(1)$ .

We conclude with a result about the minimality of  $M_a$  on balls (rather than on spheres), remarking that it holds without requiring the condition (as in Theorem 7.2) that  $[a]$  generate  $(\mathbb{Z}_p/p^2\mathbb{Z}_p)^\times$ .

**Theorem 8.5.** *Let  $p$  be an odd prime number. If  $a$  is a point of  $B_{p^{-1}}(1)$ ,  $\ell$  is an integer satisfying  $|\log_p(a)|_p = p^{-\ell}$ , and  $c$  is a point of  $\mathbb{Q}_p$ , then  $M_a$  is minimal and uniquely ergodic on  $B_{p^{-\ell}}(c)$ .*

*Proof.* By the proof of Lemma 8.3,  $M_a$  acting on  $B_{p^{-\ell}}(c)$  is isomorphic to  $T_{\log_p a}$  acting on  $B_{p^{-\ell}}(\log_p a)$ . By Corollary 6.3,  $M_a$  is minimal on  $B_{p^{-\ell}}(c)$ .  $\square$

**ACKNOWLEDGMENTS.** The authors would like to acknowledge support from a National Science Foundation REU grant and a Bronfman Science Center grant from Williams College as part of the SMALL program. We would like to thank Frank Morgan for an early reading of our paper and Joe Auslander for remarking to us that Theorem 4.4 holds in the more general setting of equicontinuous maps. We are indebted to the referees for comments and suggestions that improved the exposition.

#### REFERENCES

- [1] S. Albaverio, A. Khrennikov, B. Tirostsi, and S. de Shmedt,  $p$ -adic dynamical systems, *Teoret. Mat. Fiz.* **114** (1998) 349–365.
- [2] V. S. Anashin, Uniformly distributed sequences of  $p$ -adic integers, *Mat. Zametki* **55** (1994) 3–46, 188; translation in *Math. Notes* **55** (1994) 109–133.
- [3] —, Uniformly distributed sequences of  $p$ -adic integers, *Diskret. Mat.* **14** (2002) 3–64; translation in *Discrete Math. Appl.* **12** (2002) 527–590.

- [4] R. L. Benedetto, Hyperbolic maps in  $p$ -adic dynamics, *Ergod. Th. & Dynam. Sys.* **21** (2001) 1-11.
- [5] G. D. Birkhoff, *Collected Mathematical Papers*, 3 vols., American Mathematical Society, New York, 1950.
- [6] E. B. Burger and T. Struppeck, Does  $\sum_{n=0}^{\infty} \frac{1}{n!}$  really converge? Infinite series and  $p$ -adic analysis, *Amer. Math. Monthly* **103** (1996) 565-577.
- [7] Z. Coelho and W. Parry, Ergodicity of  $p$ -adic multiplications and the distribution of Fibonacci numbers, in *Topology, Ergodic Theory, Real Algebraic Geometry*, Amer. Math. Soc. Transl. Ser. 2, no. 202, American Mathematical Society, Providence, 2001, pp. 51-70.
- [8] S. Eigen, A. Hajian, and S. Kakutani, Weak complements of integers (in preparation).
- [9] H. Furstenberg, Strict ergodicity and transformations on the torus, *Amer. J. Math.* **83** (1961) 573-601.
- [10] —, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, Princeton, 1981.
- [11] F. Gouvêa,  *$p$ -adic Numbers*, Springer-Verlag, New York, 2000.
- [12] M. Gundlach, A. Khrennikov, and K.-O. Lindahl, On ergodic behavior of  $p$ -adic dynamical systems, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **4** (2001) 569-577.
- [13] A. Khrennikov, *Non-Archimedean Analysis: Quantum Paradoxes, Dynamical Systems and Biological Models*, Kluwer, Dordrecht, 1997.
- [14] R. Oselies and H. Zieschang, Ergodische Eigenschaften der Automorphismen  $p$ -adischer Zahlen, *Arch. Math.* **26** (1975) 144-153.
- [15] J. Oxtoby, Ergodic sets, *Bull. Amer. Math. Soc.* **58** (1952) 116-136.
- [16] —, *Measure and Category*, 2nd ed., Springer-Verlag, New York, 1980.
- [17] J. Rivera-Letelier, Espace hyperbolique  $p$ -adique et dynamique des fonctions rationnelles, *Compositio Math.* **138** (2003) 199-231.
- [18] A. Robert, *A Course in  $p$ -adic Analysis*, Springer-Verlag, New York, 2000.
- [19] J. Rotman, *An Introduction to the Theory of Groups*, Springer-Verlag, New York, 1995.
- [20] D. Rudolph, *Fundamentals of Measurable Dynamics*, Clarendon Press, Oxford, 1990.
- [21] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, New York, 1982.
- [22] J. D. Zund, George David Birkhoff and John von Neumann: A question of priority and the ergodic theorems, 1931-1932, *Historia Mathematica* **29** (2002) 138-156.

*Department of Mathematics, Hill Center, Busch Campus, 110 Frelinghuysen Rd., Rutgers University, Piscataway, NJ 08854-8019 USA*  
*jbryk@math.rutgers.edu*

*Department of Mathematics, Williams College, Williamstown, MA 01267, USA*  
*csilva@williams.edu*