On Mixing-Like Notions in Infinite Measure

Cesar E. Silva

Abstract. Measurable dynamical systems are defined on a measure space, such as the unit interval or the real line, with a transformation or map acting on the space. After discussing dynamical properties for probability spaces such as ergodicity, weak mixing, and mixing, we consider analogs of mixing and weak mixing in infinite measure, and present related examples and definitions that are the result of research with undergraduates. Rank-one transformations are introduced and used to construct the main examples.

1. INTRODUCTION. An abstract dynamical system is a set of points X together with a map or transformation defined on it that we denote by $T : X \to X$. If we were to study the planets moving around the sun, for example, we could represent each planet with a 3-dimensional position vector and a momentum vector, so with a point in $\mathbb{R}^6$; then a state of the system at a fixed moment in time is given by a point in $\mathbb{R}^{54}$ (assuming nine planets). The set of all possible states then is a certain subset of $\mathbb{R}^{54}$, and assuming Newton's laws of motion, there is a transformation T such that given a state x of the system at a certain moment, then $T(x)$ is the state after a unit of time, and $T(T(x))$ the state after two units of time, etc. Many dynamical systems come from differential equations, but in this article we will take the approach that we are already given an abstract set X and a map $T : X \to X$. We will only consider invertible transformations (one-to-one and onto outside a set of measure zero). We are interested in studying what dynamical behaviors are possible for such a system.

Given a dynamical system (X, T), one of the first things one is interested in is the orbits of a point, where the **orbit** of a point x is defined by $\mathcal{O}(x) = \{T^n(x) : n \in \mathbb{Z}\}$ (with T^n denoting the composition of T with itself n times). We also consider what we call the nth image of a set $A \subset X$ defined by

$$T^n A = \{T^n(x) : x \in A\},$$

for a fixed integer n; this tells us where a set of states is going under iteration by T. We start now with an example to illustrate some of these ideas.

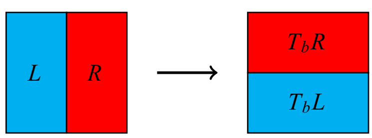

Figure 1. The baker's transformation T_b on the unit square.

Let X stand for the unit square. We will define a transformation T_b on X called the **baker's transformation**. This is a transformation on a 2-dimensional set, but it can be seen as a 2-dimensional version of the process of kneading dough; we do not discuss this here but the reader may refer to [**33**]. To define the transformation, partition

http://dx.doi.org/10.4169/amer.math.monthly.124.9.807
MSC: Primary 37A00

the unit square into a left piece L and a right piece R (see Figure 1). We think of pushing the left piece down (in the vertical direction) and stretching (in the horizontal direction), so that a point (x, y) on the left goes to the point $(2x, y/2)$. So L is sent to another rectangle $T_b L$ of the same area, but half the height and twice the width. The right rectangle R is similarly pushed down and stretched, but now the result is put on top of $T_b L$ as in Figure 1, which shows the transformation T_b, with the initial input and final result of one application of the transformation.

The first property we note of this transformation is that it is **measure-preserving**. There is a well-defined notion of area for rectangles in the unit square such as regions L and R in Figure 1; for example, the area of L is $1/2$. One can extend the notion of area to more general regions of the square by approximating the region by rectangles; we call this extension **2-dimensional Lebesgue measure**, or simply measure. More precisely, if A is a subset of the square, we define its measure, denoted by $\mu(A)$, as the infimum of all sums $\sum_i \text{area}(R_i)$, where the R_i are rectangles whose union (possibly countably infinite) covers A. So

$$\mu(A) = \inf\left\{\sum_i \text{area}(R_i) : A \subset \bigcup_i R_i\right\}.$$

It can be shown that this agrees with the area of rectangles: $\mu(R) = area(R)$ for all rectangles R. It is reasonable to require this new measure to satisfy some natural properties for a generalized notion of area, such as the fact that the measure of a countable union of disjoint sets, $\mu\left(\bigcup A_i\right)$, should be the sum of their measures, $\sum_i \mu(A_i)$. This fact can be shown for a large family of subsets of the square, called the **measurable** sets; they contain all rectangles, countable unions of them, and their complements. In terms of notation we remark that $\mu(A)$ has been defined for any set A, and in this case it is customary to call it the **outer measure** of A and to write it as $\mu^*(A)$ instead of $\mu(A)$. When A is measurable, one uses $\mu(A)$ instead, but its computation is the same.

The measure-preserving property for a transformation T can now be stated as saying that for all measurable sets A, the measure of A should be the same as the measure of TA, namely $\mu(A) = \mu(TA)$. (We are implicitly assuming that TA also has to be measurable whenever A is measurable and conversely; also, we are only concerned here with invertible transformations, such as T_b.) One can verify that for our baker's transformation, for each rectangle R', $T_b R'$ is another rectangle of the same area; for example this is easy to see for the rectangles L and R. Interestingly, it suffices to verify the measure-preserving property for all rectangles, and then it follows that it holds for all measurable sets, see, e.g., [**33**, 6.3.5]. From now on, all our sets will be measurable even if not explicitly stated.

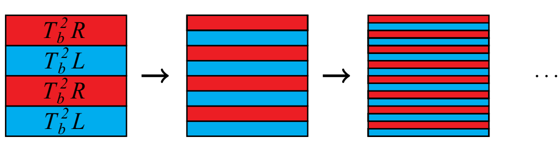

Figure 2. Second and further iterates of the baker's transformation T_b.

The following interesting dynamical property can be shown about this transformation. Imagine that we drop some raisins in the rectangle and the raisins comprise a measurable set A that occupies 5% of the rectangle. (Our raisins here also have to be 2-dimensional!) Then after iterating a large number of times, if we fix our gaze in any particular region of the space, we will see about 5% of this region occupied by raisins. To be more precise, say we fixed our gaze in a region called B. Here $T_b^n A$ denotes where the raisins are after n iterations of the transformation. The part of the raisins that is in B after n iterations is the set $T_b^n A \cap B$. Then we claim that the proportion of B that is made up of raisins after n iterations, namely $\mu(T_b^n A \cap B)/\mu(B)$, converges to $\mu(A)$, the original proportion of raisins.

More precisely, it can be shown that for all measurable sets A and B in the square X, it is the case that

$$\lim_{n\to\infty} \mu(T_b^n A \cap B) = \mu(A)\mu(B). \tag{1}$$

We call this property of the transformation T_b **mixing**. The reader may recall that two events A and B are said to be independent with respect to a probability if the probability of their intersection is the product of the respective probabilities. We can think of the mixing property of a transformation T as the situation where given any two measurable sets A and B, if we iterate one of them, $T^n A$, it becomes independent of the other in the limit: $\lim \mu(T^n A \cap B) = \mu(A)\mu(B)$. Figure 2 shows a few iterates of the transformation T_b. Imagine that A is the set L and B the set R. Then $T_b^3 A$ consists of 8 strips of the original left piece (cyan) and we see that they are beginning to spread uniformly over B. The reader may have noticed that the sets L and R already become independent under one iteration, namely $\mu(T_b L \cap R) = 1/4 = \mu(L)\mu(R)$. There is a large class of rectangles where this is true and it can be used to verify (1) for all rectangles, and to show mixing for this transformation.

Now we will be interested in verifying various dynamical properties, such as mixing, for all measurable sets. This introduces a theme that will be prevalent in all our examples. We first verify the dynamical property for a large class of sets, such as all rectangles or a special class of rectangles, and then use some sort of approximation to verify it for all measurable sets. In the case of mixing (i.e., (1)) we have already mentioned that it suffices to show condition (1) for all rectangles, and then it follows for all measurable sets, but this will not be the case for other properties.

The second important example we consider are rotations. A rotation T_α is a transformation on $[0, 1)$ given by

$$T_\alpha(x) = x + \alpha \bmod 1,$$

where α is a real number that may be assumed to be in $[0, 1)$. By mod 1 we mean that we identify 0 with 1, and every number $x \in \mathbb{R}$ is identified with its unique integer translate x' in $[0, 1)$ (i.e., x' is the unique element of $[0, 1)$ so that $x' + n = x$ for some integer n). As we identify 0 with 1 we can think of the interval $[0, 1]$ becoming a circle of length 1 and then T_α becomes a rotation by $2\pi\alpha$ (the result of translating by α radians). When α is rational, every point is periodic; for example, every point x under $T_{1/4}$ has **period** 4, i.e., $T_{1/4}^4(x) = x$. On the other hand, when α is irrational, it can be shown that T_α satisfies the ergodic property, mentioned presently.

Rotations are defined on linear sets such as the unit interval; so we need to discuss a generalization of the notion of length on the interval, and more generally on the real line. One can define a measure on the unit interval, called **1-dimensional Lebesgue measure** in a way analogous to the 2-dimensional case. If A is a subset of the unit interval we define $\mu(A)$ as the infimum of all sums $\sum \text{length}(I_i)$, where the I_i are intervals whose union (possibly countably infinite) covers A. This generalizes in a natural way the notion of length; we will use the same notation μ as it should be clear from the context if we refer to 1-dimensional or 2-dimensional Lebesgue measure. We can now establish that rotations are measure-preserving by checking that intervals are sent to intervals of the same length.

There is one more extension we need to discuss. We will also be interested in a measure defined on the whole real line, or any interval of the real line. We can extend 1-dimensional Lebesgue measure to subsets of any interval in the real line in the same way as we did for subsets of the unit interval: just cover the set with a collection of intervals. It follows then that intervals such as $[0, \infty)$ have infinite measure. In what follows we will mainly consider two measure spaces: one case will be when $X = [0, 1)$, which we will also call a probability space, and the other when $X = [0, \infty)$, which we will call an infinite measure space.

To prove properties of dynamical systems for all measurable sets, we will first prove them for intervals and then we will need to do some approximation to verify the claim for all measurable sets. The following property will be important when doing the approximation to all measurable sets. If a set A has positive measure, then there has to be an interval I that is "very full" of A; more precisely, if A has positive measure, then for any $\delta > 0$ there exists an interval I (of positive length) such that

$$\mu(A \cap I) > (1 - \delta)\mu(I);$$

in this case we will say that I is **more than** $(1 - \delta)$**-full of** A. (We do not have control on the length of the interval I.) We give a brief argument for this. We can approximate a given measurable set A by a finite union of intervals to any desired accuracy, i.e., there is a set K that is a finite disjoint union of intervals so that $\mu(A \setminus K \cup K \setminus A) < \varepsilon$ for a given $\varepsilon > 0$; this follows from the definition of the measure as the infimum of sums of lengths of intervals. (Here $A \setminus K$ denotes the set of points in A that are not in K.) So A is "almost," in measure, a finite union of intervals. It is possible to choose ε sufficiently small in terms of δ so that at least one of the intervals is more than $(1 - \delta)$-full of A, see, e.g., [**33**, 3.7.1]. (The argument goes by contradiction: if all intervals were not sufficiently full of A then their union could not approximate A well.) Finally, we mention that this is possible for other nice collections of intervals. For example, one could just look at dyadic intervals, i.e., intervals whose endpoints are of the form $\frac{k}{2^n}$ for integers k and n (this is since arbitrary intervals can be approximated by finite unions of dyadic intervals to any desired accuracy).

We have said already that at its most basic, an abstract dynamical system consists of a set X and a transformation T. Often one puts an additional structure on X, and in this article it will be that of a measure space, which consists of a distinguished collection of so-called measurable sets, and a measure μ. We will refer to this as the measure space (X, μ), and to simplify our exposition we will not have a symbol for the collection of measurable sets; the reader may assume, though, that we are just using one of the spaces we have discussed.

A measure-preserving transformation T is said to be **ergodic** if whenever A is an **invariant** (measurable) set for T, i.e., $TA = A$, then either A or A^c has measure zero. If there were such an invariant set A, then T restricted to A and T restricted to A^c would be dynamical systems in their own right; in particular, the orbit of points that start in A would never reach A^c. If T is ergodic, then it cannot be decomposed in this way. Any rational rotation is not ergodic; for example, an invariant set for $T_{1/2}$ is $A = [1/4, 1/2) \cup [3/4, 1)$, and $\mu(A)\mu(A^c) \neq 0$. However, if α is irrational, then T_α is ergodic; we give an argument at the end of Section 3. Furthermore, if T is a mixing transformation and A is an invariant set, letting $B = A^c$ in (1) yields $0 = \mu(A)\mu(A^c)$, so T has to be ergodic. Thus the baker's transformation is ergodic since it is mixing. Interestingly, it is not sufficient to verify this property for all rectangles (i.e., to show that there are no invariant rectangles, or intervals) to show ergodicity, as shown at the end of Section 5.

As we will see in Section 2, it can be shown that for measure-preserving transformations on a probability space, ergodic is equivalent to "mixing on the average," namely that for all measurable sets A and B in X,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^i A \cap B) = \mu(A)\mu(B). \tag{2}$$

It is clear that condition (1) implies (2). There is a third condition between these two conditions. A measure-preserving transformation T is said to be **weakly mixing** if for all measurable sets A and B,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |\mu(T^i A \cap B) - \mu(A)\mu(B)| = 0. \tag{3}$$

To distinguish mixing from weak mixing, sometimes mixing has been called **strong mixing**, but we will simply use mixing.

At the level of numerical sequences $(a_i)_i$, we can write the following three limit conditions which are in strict order of implication:

$$\lim_{i\to\infty} a_i = 0; \tag{4}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |a_i| = 0; \tag{5}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} a_i = 0. \tag{6}$$

It is not hard to show that (4) implies (5), which in turn implies (6), nor to find bounded sequences showing that the converses of these implications do not hold.

If we let

$$a_i = \mu(T^i A \cap B) - \mu(A)\mu(B), \tag{7}$$

then (4), (5), and (6) correspond to mixing, weak mixing, and ergodic, respectively (if we require they hold for all measurable sets A, B). It is harder to see examples where the sequences are given by a transformation for which the converses do not hold.

Our main goal in this article is to explore mixing-like conditions for transformations defined on measure spaces where the total measure is infinite. This is accomplished by looking at work done with undergraduates in this area. In the references, the names of authors who were undergraduates at the time of the research are starred.

We will start by first exploring the ergodic property in more detail, where we also make a case for exploring analogs of the weak mixing condition rather than the mixing condition when considering infinite measure.

2. CONSEQUENCES OF ERGODICITY AND EXAMPLES.

If a transformation T (on probability spaces) satisfies condition (2), then it is ergodic by a simple argument similar to the one we gave showing that the mixing condition (1) implies ergodic. The converse is also true and follows from the ergodic theorem.

We mention a special case of the ergodic theorem, which is all that we need now; a proof can be found in any ergodic theory book, see, e.g., [**33**]. We first need some notation; let $\mathbb{I}_A$ denote the indicator function of a set A, that is $\mathbb{I}_A(x) = 1$ when x is in A, and otherwise $\mathbb{I}_A(x) = 0$. Then the expression $\sum_{i=0}^{n-1} \mathbb{I}_A(T^i x)$ denotes the number of visits to A of a point x under n iterations by T (counting from $i = 0$ to $i = n - 1$).

The ergodic theorem states that, when T is ergodic and $\mu(X) = 1$, for all measurable sets A, for all x in X outside a set of measure zero, the average number of visits to A of x, namely, $\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{I}_A(T^i x)$ converges to the measure of A. When $\mu(X) = \infty$, it states that for all measurable sets A of finite measure, the average number of visits to A of a point x, outside a set of measure zero, converges to 0. The ergodic theorem holds not only for characteristic functions $\mathbb{I}_A$ as stated, but for all integrable functions.

The following proposition can be seen as a direct consequence of the ergodic theorem.

Proposition 1. *Let (X, μ) be a measure space and $T : X \to X$ be an ergodic measure-preserving transformation. Let A and B be measurable sets on X. If $\mu(X) = 1$, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^i A \cap B) = \mu(A)\mu(B). \tag{8}$$

If $\mu(X) = \infty$ and $\mu(A) < \infty$, $\mu(B) < \infty$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^i A \cap B) = 0. \tag{9}$$

It follows from condition (9) that when T is ergodic and infinite measure-preserving, then for all sets of finite measure A and B, there must exist a sequence $\{n_k : k \in \mathbb{N}\}$ so that

$$\lim_{k\to\infty} \mu(T^{n_k} A \cap B) = 0.$$

For nice measure spaces, such as the ones we are working in, one can in fact show that there is a single sequence that works for all finite measure sets A, B. This ends our dream of finding mixing behavior in infinite measure: we cannot expect convergence of $\mu(T^n A \cap B)$ to any positive quantity.

Since Hopf's book was published in 1937 [**22**], perhaps the first book in ergodic theory, there has been a lot of interest in formulating conditions for infinite measure-preserving transformations that capture the notion of mixing in finite measure. One of them is the following. A transformation T is **Koopman mixing** if for all sets A, B of finite measure,

$$\lim_{n\to\infty} \mu(T^n A \cap B) = 0. \tag{10}$$

One could argue that this is a reasonable notion of mixing in infinite measure, since condition (1) in the case of a finite measure space becomes $\lim_{n\to\infty} \mu(T^n A \cap B) = \mu(A)\mu(B)/\mu(X)$, and one could think of (10) as the limit of this equation when

$\mu(X) \to \infty$. There is a formulation of this condition using an operator called a Koopman operator. When that formulation is interpreted in the finite case, it agrees with mixing in finite measure. On the other hand, Koopman mixing, in infinite measure, does not even imply ergodic (just consider two disjoint copies of a transformation that is Koopman mixing), and when the transformation is ergodic, Koopman mixing has been shown to be independent of many of the mixing-like conditions we study below. Koopman mixing in infinite measure was introduced by Hajian and Kakutani, who called it **zero type**, and they showed that an ergodic infinite transformation is either of zero type or of **positive type** (i.e., $\liminf \mu(T^n A \cap A) > 0$) [**16**]. A further refinement of positive type was studied more recently by the authors in [**32**].

We will now study conditions in infinite measure that we think exhibit some mixing-like behavior in the sense of resembling conditions that are equivalent to weak mixing in the finite case. We will consider in particular the notions of power weak mixing, and weak double ergodicity. The examples will be rank-one transformations and we start by describing this class of transformations. Related notions and other examples will be explored in more detail in an upcoming survey with T. Adams. The reader may also refer to [**13**]; other more recent notions of mixing for infinite measure that we do not consider have been explored in [**26, 28**].

3. RANK-ONE CONSTRUCTIONS. We introduce a family of transformations, called rank-one, which has been used to construct a wide variety of interesting examples and counterexamples in ergodic theory—for example, the first weak mixing and nonmixing transformations. All rank-one transformations will be invertible and defined on the unit interval or the positive real line, with Lebesgue measure. They will be measure-preserving and ergodic with respect to this measure. All our examples in the remainder of the article will be rank-one.

Rank-one transformations have several definitions and we present a geometric, constructive way of obtaining them. The constructions are inductive, by a process called cutting and stacking. At each stage of the construction there is one column, justifying the reason they are called rank-one, as the number of columns in the construction gives rise to the rank. A **column** or **Rohlin column** consists of a finite list of intervals of the same length, that we call **levels**. The first interval is the base of the column and we think of levels as placed on top of each other; we call the number of levels the **height** of the column. A column defines a **column map**, which on each interval is defined as the translation that sends that interval to the one above. For example, the left part in Figure 4 shows a column consisting of two intervals; this column defines a transformation on the first interval as the map that sends x to $x + 1/2$ (this is determined by saying that $[0, 1/2)$ is sent to $[1/2, 1)$); the column map remains undefined on the top interval of a column.

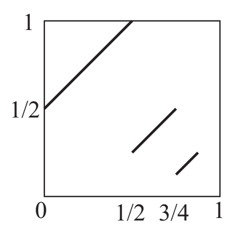

Figure 3. The odometer transformation T_d on the unit interval.

Our first example is the **dyadic odometer**, also called the Kakutani–von Neumann transformation; it was originally defined as a piecewise translation on infinitely many

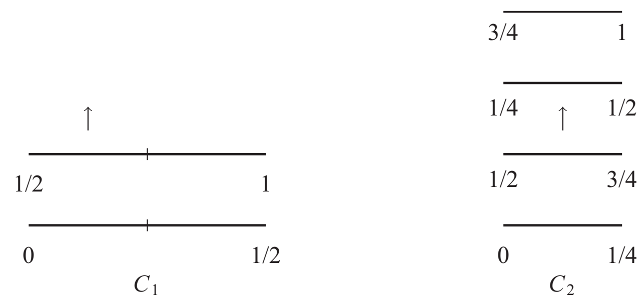

Figure 4. Columns C_1 and C_2 in the dyadic odometer.

subintervals of the unit interval; this is illustrated in Figure 3. Let $I_0, = [0, 1/2)$, $I_1 = [1/2, 3/4)$, $I_2 = [3/4, 7/8)$, etc. be the subintervals shown on the x-axis in Figure 3, and let $J_0 = [1/2, 1)$, $J_1 = [1/4, 1/2)$, $J_2 = [1/8, 1/4), \ldots$ be the subintervals in the y-axis of the same figure. Note that, for each $n = 0, 1, 2, \ldots,$ I_n is of the same length as J_n, and that one can see in Figure 3 that the transformation sends the interval I_n by translation onto the interval J_n. It was later realized that this transformation can be given by a rank-one cutting and stacking construction. We describe that method now.

We define the dyadic odometer T_d as a rank-one construction using the method of cutting and stacking. As we said, we use an inductive process to give a sequence of columns. The base case defines the column C_0 to be the unit interval, so $C_0 = \{[0, 1)\}$. The height of this column is $h_0 = 1$. We give the next step in detail before proceeding to the inductive step. Cut the interval in C_0 in half, which gives the intervals $I_0 = [0, 1/2)$ and $J_0 = [1/2, 1)$. Then put J_0 on top to construct a column C_1, consisting of the two intervals I_0 and J_0. We see that T_d sends interval I_0 to interval J_0, so T_d on $[0, 1/2)$ is the map that sends x to $x + 1/2$, agreeing with the definition of the transformation in Figure 3. Column C_1 is shown in Figure 4.

Now observe that $I_1 = [1/2, 3/4)$ is the left half of the top level of column C_1, and $J_1 = [1/4, 1/2)$ is the right half of the bottom level of column C_1. To signify that I_1 is sent to J_1 by T_d, what we do now is cut each level of column C_1 in half, resulting in two subcolumns, and then we place the right subcolumn on top of the left subcolumn, resulting in column C_2 as depicted in Figure 4. We see that this shows that I_1 is sent to J_1, and also note that the definition of T_d on I_0 has not changed. To define T_d on I_2, we analogously note that I_2 is the left half of the top level of C_2, so we cut C_2 into two subcolumns and place the right subcolumn on top of the left. This defines T_d on I_2. We continue in this way subdividing columns in half.

Now we give the inductive step, for which we assume we are given column C_n, consisting of h_n levels. To obtain C_{n+1} from C_n, cut each level in C_n into two subintervals, yielding two subcolumns. Then place, or stack, the right subcolumn on top of the left subcolumn to obtain C_{n+1} of height $h_{n+1} = 2h_n$. Note that C_{n+1} extends the definition of the map to the left half of the top level of C_n. This is illustrated in Figure 5, which shows column C_n.

We introduce one more definition. We see that a level I in C_n is cut into two halves and each half becomes a new level in C_{n+1}; we call these the **descendants** of I in C_{n+1}. In this way we see that a level I in C_n has two descendants in C_{n+1}, four descendants in C_{n+2}, eight in C_{n+3}, and so forth.

Since each x in $[0, 1)$ is in a level of some column that is not the top level, the transformation is defined at x; thus T_d is defined on all of $[0, 1)$. Before discussing other properties of T_d, we note that we can regard it as an invertible transformation. There are some points with no pre-image, but they are all in the positive orbit of 0,

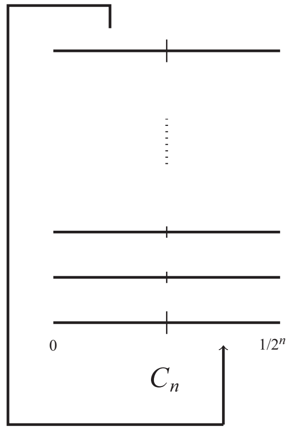

Figure 5. Column C_n in the dyadic odometer.

$\mathcal{O}^+(0) = \{0, T_d(0), T_d^2(0), \ldots\}$. If we delete $\mathcal{O}^+(0)$, a measure zero set, then T_d is an invertible transformation on $[0, 1) \setminus \mathcal{O}^+(0)$. We are now ready to prove two basic properties of T_d.

Lemma 2. *The dyadic odometer transformation is measure-preserving.*

Proof. Let T_d be the dyadic odometer and A a measurable set. We will prove that $T_d A$ is measurable and of the same measure as A. For each $n > 0$, given column C_n, let C_n' consist of C_n except its top level. Then $X = [0, 1)$ is equal to the union $\bigcup_{n \geq 1} C_n'$. (By abuse of notation we use C_n' to denote both the column and the union of its levels.) For each $n > 0$, set $A_n = A \cap C_n'$. It is clear that A is the increasing union of the sets A_n. Also, as T_d is a translation on each level of C_n', $T_d A_n$ is measurable and of the same measure as A_n. As $T_d A$ is also equal to the increasing union of the $T_d A_n$, it follows that $T_d A$ is measurable and of the same measure as A. (By a similar argument, T_d^{-1} is also measure-preserving.) ■

The following lemma gives a useful property equivalent to ergodicity that we use to prove that the dyadic odometer is ergodic in Proposition 4.

Lemma 3. *An invertible transformation T is ergodic if and only if for all sets A and B of positive measure there exists $n \in \mathbb{Z}$ so that $\mu(T^n A \cap B) > 0$.*

Proof. Suppose that T is ergodic and let A and B be sets of positive measure. Set $A^* = \bigcup_{n \in \mathbb{Z}} T^n A$. Then $T(A^*) = A^*$. By ergodicity, the complement of A^* has measure zero, which means that there must be some n such that $\mu(T^n A \cap B) > 0$ as B has positive measure. Conversely, let A be an invariant set of positive measure and B be the complement of A. If B has positive measure, then there is an integer n such that $\mu(T^n A \cap A^c) > 0$, contradicting that A is invariant. Therefore the complement of A has measure zero and T is ergodic. ■

Proposition 4. *The dyadic odometer transformation is ergodic.*

Proof. Let A and B be sets of positive measure. Since the levels of columns form the dyadic intervals, there exist intervals I in column C_n and J in column C_m, for some n and m, so that I is more than 1/2-full of A, and that J is more than 1/2-full of B:

$$\mu(A \cap I) > \frac{1}{2}\mu(I) \text{ and } \mu(B \cap J) > \frac{1}{2}\mu(J).$$

Suppose $n > m$. We note that when J is divided in halves, by the pigeon-hole principle, at least one of its two descendants in C_{m+1} must continue being more than 1/2-full of B. In this way we can continue subdividing J until we reach a descendant of J in C_n that is more than 1/2-full of B; let us call it J'. As I is in C_n, there is an integer k so that $T_d^k I = J'$. Next we observe that T_d is a translation on each level, so if a level is more than 1/2-full of A, then its translation continues being more than 1/2-full of the translation of A. So $T_d^k I = J'$ is more than 1/2-full of $T_d^k A$. As J' is more than 1/2-full of B, it follows that $\mu(T_d^k A \cap B) > 0$. Now if $n < m$, we subdivide I in halves so that at least one descendant is more than 1/2-full of A. We continue subdividing I until we reach a descendant I' of I that is in C_m and is more than 1/2-full of A. Then again there is an integer k so that $T_d^k I' = J$ and it follows that $\mu(T_d^k A \cap B) > 0$. If $n = m$, we already have the levels in the same column and there is an integer k so that $T_d^k I = J$. Therefore T_d is ergodic by Lemma 3. ■

Remark. The proofs we have presented that the dyadic odometer is measure-preserving and that it is ergodic hold in essentially the same form for any rank-one construction; that is the main reason we have presented them in detail. These properties remain true for all the examples we will see later. There are rank-one constructions where columns are subdivided into many pieces; in some cases this number many increase with n and be unbounded. In some cases extra intervals, called spacers, may be added to some columns before stacking to obtain the next column. In all constructions where the total measure is infinite, extra intervals are added to the columns. If I is an interval that is more than 1/2-full of A, when we subdivide it, independently of the number of subintervals, at least one subinterval in the subdivision will continue being more than 1/2-full of A. It follows that for all later columns, at least one level will continue being more than 1/2-full of A. The measure-preserving argument also holds in the general case. The reader is encouraged to check that in all the rank-one constructions below (and in fact for all rank-one transformations), essentially the same proofs show measure-preserving and ergodic; further details can be found in [**33**].

If T is measure-preserving on a finite measure space, a direct consequence is that the transformation is **conservative**; that is, for all sets A of positive measure there is an integer $n \neq 0$ so that $\mu(T^n A \cap A) > 0$. This result, which is equivalent to the Poincaré recurrence theorem, follows from the fact that as the sets $\{T^n A : n \in \mathbb{N}\}$ have the same measure, they cannot all be disjoint. We recall that the Poincaré recurrence theorem states that for all sets A of positive measure, there is a set N of measure zero, so that for all $x \in A \setminus N$ there is an integer $n > 0$ so that $T^n(x) \in A \setminus N$ (see, e.g., [**33**]). In infinite measure, though, there are measure-preserving transformations that are not conservative, as can be seen by considering the transformation $T(x) = x + 1$ on $\mathbb{R}$. There are even ergodic transformations that are not conservative, as the shift $n \to n + 1$ on the integers $\mathbb{Z}$ with counting measure shows (the counting measure of a set is its number of elements).

However, when the measure is not atomic, as is the case with Lebesgue measure, if the transformation is invertible and ergodic, then it is conservative. To see this, suppose that $\mu(T^n A \cap A) = 0$ for all $n \neq 0$. The fact that μ is not atomic means that there exists a measurable set $B \subset A$ such that $0 < \mu(B) < \mu(A)$. Let $B^* = \bigcup_{n \in \mathbb{Z}} T^n B$. Then B^* is an invariant set of positive measure, but B^* is disjoint from $A \setminus B$, a set of positive measure. So the complement of B^* does not have measure zero and T could not be ergodic. Therefore T is conservative.

Observe that for the dyadic odometer, if $A = [0, 1/2)$, then $T^2(A) = A$. So while T is ergodic, T^2 is not. This implies directly that T cannot be weakly mixing. We give one proof using the Cartesian product.

Given a transformation T on a measure space (X, μ), we can define the Cartesian product space $X \times X$ as the set consisting of all ordered pairs (x, y) when $x, y \in X$. We can define a measure $\mu \times \mu$ by first defining it on product sets, i.e., sets of the form $A \times B$, where A, B are measurable sets in X, by $\mu \times \mu(A \times B) = \mu(A)\mu(B)$; then it can be extended to a class of measurable sets in a way similar to previous definitions, but now the role of intervals is played by product sets. The product transformation $T \times T$ on $X \times X$ is defined by

$$T \times T(x, y) = (T(x), T(y)).$$

We will see later that when T is finite, measure-preserving, and weakly mixing, then $T \times T$ is ergodic. Now we show that if T^2 is not ergodic, then $T \times T$ is not ergodic. In fact let A be a set such that $T^2 A = A$ and $\mu(A) > 0$, $\mu(A^c) > 0$. Then one can verify that $(A \times TA) \cup (TA \times A)$ is an invariant set for $T \times T$, so $T \times T$ is not ergodic. This works both in finite and in infinite measure.

We now consider several modifications of the dyadic odometer construction. First, instead of using dyadic intervals we could use triadic intervals. This is equivalent to starting with C_0 consisting of the unit interval and then defining C_{n+1} from C_n but subdividing each subinterval of C_n into three, and then stacking the subcolumns from right over left. This again yields an ergodic transformation T such that while T^2 is ergodic, T^3 is still not ergodic, so T is still not weakly mixing.

Examples of weakly mixing transformations that are not mixing are harder to construct, and one of the first such examples is a transformation defined by Chacón [**6**]. There are two classic Chacón examples: the first and better known one appeared in [**17**]; we will discuss the one that appeared in [**6**], which we call the second Chacón transformation. Perhaps the simplest way to show weak mixing for the Chacón transformations is to verify a property called continuous spectrum, for which the reader could consult [**18**].

To define the second Chacón transformation transformation T_c, we need the concept of **spacers**. Spacers are new intervals that are introduced to a column before stacking to produce the next column; in this way the measure of the columns increases. For the second Chacón transformation we start with $C_0 = \{[0, 1/2)\}$. The union of the totality of spacers that are added will be the interval $[1/2, 1)$, so that in the end the transformation is defined on $[0, 1)$. To clarify the construction we show the first steps in detail. To construct C_1, cut $[0, 1/2)$ into two subintervals $I_0 = [0, 1/4)$, $I_1 = [1/4, 1/2)$ and consider a new subinterval of length 1/4 abutting $[0, 1/2)$, namely $J_1 = [1/2, 3/4)$, called a spacer. Then place I_1 above I_0, and J_1 above I_1, resulting in column C_1 consisting of the intervals I_0, I_1, J_1. So T maps I_0 to I_1 and I_1 to J_1. The first part of Figure 6 shows column C_0 with the intermediate step where the level is cut in half and a spacer (denoted by a dashed line and consisting of the interval J_1) is placed on the top right. The second part in Figure 6 shows column C_1, consisting of the three levels I_0, I_1, J_1, and the intermediate step where each level is cut in half and a spacer is put on the top right half. After stacking, this will yield column C_2 consisting of seven levels. For the inductive step, given column C_n consisting of h_n levels, cut each level in half, choose a new interval J_n, of the same length as half the levels of column C_n, and so that J_n abuts the union of the intervals making up C_n. Then stack the subcolumns so that the left half of the top level of C_n is sent to the right half of the bottom level, and the right half of the top level is sent to the spacer J_n. This produces a column C_{n+1}

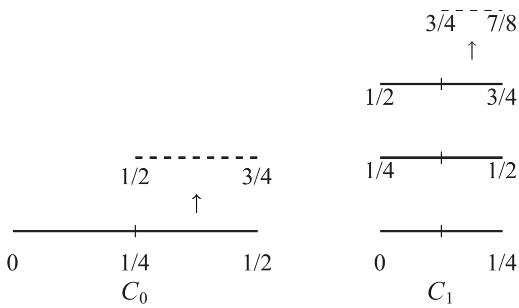

Figure 6. First steps in the construction of the second Chacón transformation T_c.

of height $h_{n+1} = 2h_n + 1$. It can be shown that this transformation is weakly mixing but not mixing [**6**]. Conditions stronger than weak mixing have been shown for this transformation in [**1, 19**]. The proof that it is not mixing is similar to the proof we will show in Section 5 that a construction that has a bounded number of cuts at every stage is not mixing.

We end this section with a the proof that irrational rotations are ergodic. (It can also be shown they are rank-one [**23**].) We start by proving that when α is irrational, every x in $[0, 1)$ has a dense orbit under T_α: for all $x \in [0, 1)$ and for all open intervals I in $[0, 1)$ there exists $n > 0$ such that $T_\alpha^n(x) \in I$. This fact is known as Kronecker's theorem and its proof has two parts. First we show that all the points in the positive orbit $\{T_\alpha^n(x) : n = 0, 1, 2, \ldots\}$ are distinct; in fact, if $T_\alpha^n(x) = T_\alpha^m(x)$, for $n \neq m$, then one can show α must be rational. Now for the second part we start subdividing $[0, 1)$ into two equal-length subintervals. One of the halves must contain infinitely many distinct points of the orbit; subdivide that half again so that one of its halves contains infinitely many distinct points of the orbit. Continue in this way until we have a half J of length less than ε that contains infinitely many distinct points of the orbit. Then J must contain two points of the form $T_\alpha^k(x)$ and $T_\alpha^{k+\ell}(x)$ for some $k > 0$, $\ell > 0$. Then when we translate by the transformation T_α^ℓ, points move a distance less than ε. So as we iterate by T_α^ℓ, the point x must eventually visit I.

Now suppose we are given sets A and B of positive measure in $[0, 1)$. We know we can find dyadic intervals I and J that are more than 3/4-full of A and B, respectively: $\mu(A \cap I) > \frac{3}{4}\mu(I)$ and $\mu(B \cap J) > \frac{3}{4}\mu(J)$. If the interval I is larger than J in length, we subdivide it in half and one of its halves must continue being more than 3/4-full of A. Continuing in this way if necessary, we can ensure I and J have the same length. Now let a be the left endpoint of I and b the left endpoint of J. As the orbit of a is dense, there exists an integer n so that $T_\alpha^n(x)$ is very close to b, say at a distance less than $1/8\mu(I)$. Then $T_\alpha^n I$ must overlap J a lot, and as $T_\alpha^n I$ must be more than 3/4-full of A, and J is more than 3/4-full of B, it follows that $\mu(T_\alpha^n A \cap B) > 0$. Therefore T_α is ergodic.

4. POWER WEAK MIXING.

When a transformation is finite measure-preserving, it can be shown that if T is weakly mixing, then for every ergodic finite measure-preserving transformation S, the Cartesian product $T \times S$ is ergodic, and also that when $T \times T$ is ergodic, then T is weakly mixing. (In fact, when $T \times T$ is ergodic one also has that $T \times T$ is weakly mixing.) A consequence of this is that if T is weakly mixing, then all finite Cartesian products $T \times \cdots \times T$ are ergodic. Kakutani and Parry proved in 1963 [**24**] that this is not the case in infinite measure. They constructed infinite, measure-preserving, invertible transformations T such that $T \times T$ is ergodic but $T \times T \times T$ is not conservative, hence not ergodic. They also constructed infinite

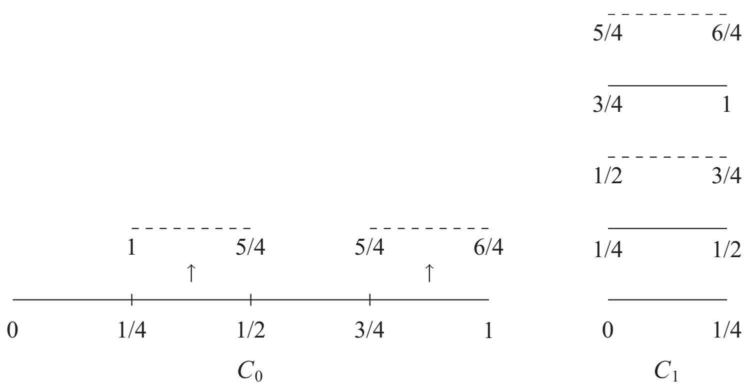

Figure 7. The first step in the construction of T_p.

measure-preserving transformations such that all their finite Cartesian products are ergodic and called this property **infinite ergodic index**. Their examples are Markov shifts, which we will not cover here but we will simply note that they are very different from rank-one transformations.

In work with undergraduates in Day et al. [**14**], a rank-one transformation T was constructed such that for every sequence of nonzero integers $k_1, \ldots, k_r$, the product transformation $T^{k_1} \times \cdots \times T^{k_r}$ is ergodic; a transformation satisfying this condition was said to be **power weakly mixing**. In [**2**], Adams et al. constructed a rank-one transformation T that has infinite ergodic index but such that $T \times T^2$ is not conservative, hence not ergodic and not power weakly mixing. So power weakly mixing is a strictly stronger condition than infinite ergodic index.

Soon after these examples were constructed, Bergelson asked if there existed a transformation T that has infinite ergodic index, but with $T \times T^{-1}$ not ergodic. More recently, it was shown by Clancy et al. [**7**] that there exists a transformation T such that $T \times T$ is ergodic but $T \times T^{-1}$ is not, but Bergelson's full question remains open. These results have been extended to more general group actions in [**11**].

We now describe the rank-one transformation from [**14**] that is power weakly mixing and illustrate ideas to prove ergodicity of products. Column C_0 consists of the unit interval; this will be the case for all infinite measure constructions (and as before, all our intervals will be left-closed, right-open). Before giving the inductive step, to illustrate the idea of the construction, we show in detail how column C_1 is obtained. Cut the single level of C_0 into four equal subintervals; then choose a new subinterval of length $1/4$ to place above the second subinterval of I_0, and another new subinterval to place above the fourth subinterval of I_0. These new subintervals are chosen to the right of $[0, 1)$, so we choose $[1, 5/4)$ and $[5/4, 6/4)$. Then stack from right over left to obtain column C_1. Figure 7 illustrates column C_0 with the cuts and spacers, and the result of stacking the sublevels and spacers is shown in column C_1 of the same figure.

Now for the inductive step, given column C_n of height h_n, cut each level into four equal pieces to yield four subcolumns each of height h_n. Place h_n new levels, spacers, over the second subcolumn, and a single spacer over the last subcolumn. Then stack right over left to obtain C_{n+1} of height $h_{n+1} = 5h_n + 1$. This is illustrated in first part of Figure 8. We let X be the union of the columns, and as $\mu(C_{n+1}) > \frac{5}{4}\mu(C_n)$, we see that X has infinite measure, in fact $X = [0, \infty)$.

To begin to understand the dynamics of T_p we look at its action on a level. Let I be a level in column C_n, as in the left part of Figure 8. First we describe how C_n "sits" in C_{n+1}. Figure 8 shows column C_n subdivided into four subcolumns, a column of h_n

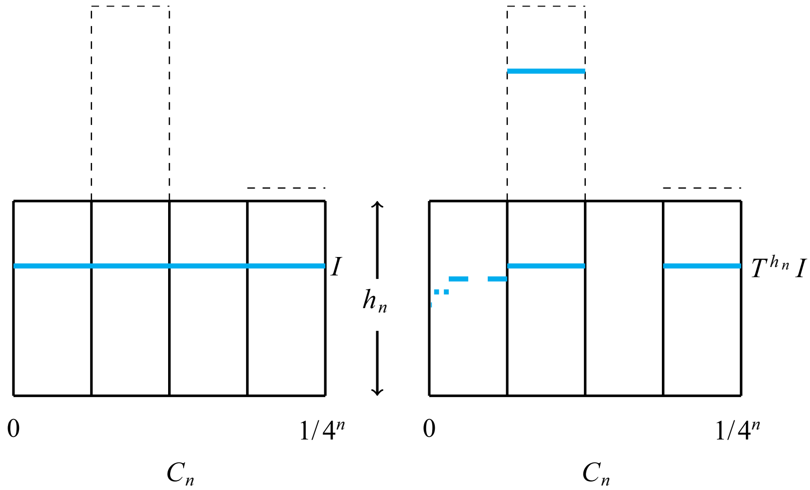

Figure 8. Column C_n in the construction of the power weakly mixing transformation in [**14**]. The solid levels on the right show $T_p^{h_n} I$ for a level I in C_n.

new levels, or spacers, above the second subcolumn, and a single new level above the fourth subcolumn. To construct C_{n+1} we send the top of the first subcolumn to the bottom of the second subcolumn, the top of the second to the bottom of the column of spacers, the top of the column of spacers to the bottom of the third subcolumn, the top of the third to the bottom of the fourth, and finally the top of the fourth to the single spacer; this forms a new column of height $5h_n + 1$ and with a fourth of the width of C_n. We can view column C_{n+1} as consisting of five **sections**, each of height h_n, and an additional single level at the top. We number these five sections of C_{n+1} from 1 to 5. The first, second, fourth, and fifth section each come from one of the four subcolumns of C_n. The third section of C_{n+1} consists of h_n spacers that were added to C_n. Now we see that $T^{h_n} I$ will consist of a level of length $\frac{1}{4}\mu(I)$ in each of sections 2, 3, and 4, and also of a **crescent** in section 1 (see the second part of Figure 8). The crescent is the result of I being pushed through all the spacers above the fourth subcolumn (we have to think of the spacers that are added in all later iterations of the construction). We see then that $T^{h_n} I$ intersects levels below I with a lower bound depending on how far down the other level is; for example, if J is two levels below I (i.e., $J = T^{-2} I$), then $T^{h_n} I$ intersects J in measure more than $1/2 \cdot 1/4^2 \cdot \mu(J)$. Higher iterates such as $T^{2h_n} I, T^{3h_n} I, T^{4h_n} I$ have a similar pattern, where we see the crescent moving to the right and a new crescent being born; for example, the crescent of $T^{h_n} I$ moves to the second subcolumn in $T^{2h_n} I$, to the third subcolumn in $T^{3h_n} I$, and to the fourth subcolumn in $T^{4h_n} I$. So we know a lower bound of the intersection of $T^{kh_n} I, k = 1, \ldots, 4$, with a level J that is below I in C_n, and the bound depends on how far down J is from I.

For iterates past $4h_n$ we use the structure of h_n. For example, we can write $5h_n = h_{n+1} - 1$, showing that $T^{5h_n} I$ is like $T^{h_{n+1}} I$ but a level below; this can be done for kh_n. Using these ideas one can have a lower bound on the measure of the intersection of levels. There are two points to note here. The interval J cannot be too far down if we want a certain lower bound; at the same time, depending on k, the crescent of T^{kh_n} moves down, so if J is too close to I it might be missed by the crescent. In summary, for a given range of k, one can control a lower bound for the measure of the intersection of $T^{kh_n} I$ with J provided it is not too close and not too far down from I, with bounds that we can calculate.

To show that T_p is power weakly mixing we need to show that for any sequence of nonzero integers $\{k_1, \ldots, k_r\}$, for any positive measure sets A and B in the r-fold Cartesian product, there is an integer m so that

$$\mu^{(r)}([T_p^{k_1} \times \cdots \times T_p^{k_r}]^m A \cap B) > 0. \tag{11}$$

(Here $\mu^{(r)}$ denotes the measure in the Cartesian product of r copies of X.)

The proof of condition (11) follows by first verifying it when A and B are Cartesian products of intervals (using the ideas we have outlined), and then using an approximation argument to verify it for all measurable sets, a theme we have introduced before.

We note, though, that to prove power weak mixing it is not sufficient to verify (11) when A and B are products of levels, not even when A and B are products of arbitrary measurable sets; we will see this in Section 5. The following theorem was proved using these ideas.

Theorem 5 ([14]). *The rank-one infinite measure-preserving transformation T_p satisfies the property that for any sequence of nonzero integers $k_1, \ldots, k_r$ the transformation $T^{k_1} \times \cdots \times T^{k_r}$ is ergodic.*

The power weak mixing property was shown for some $\mathbb{Z}^d$ actions in [**29**] and for some group extensions in [**15**], and for many other examples in [**10, 12, 25**].

We end by outlining an argument that when T is finite measure-preserving and weakly mixing, then $T \times T$ is ergodic. Let a_i be the sequence given by the formula in equation (7) for fixed measurable sets A, B. If (a_i) satisfies condition (5), the weak mixing condition, then there exists a subsequence (a_{i_k}) so that $\lim_k a_{i_k} = 0$. Indeed, for each $k > 0$ there must exist an element a_{i_k} such that $|a_{i_k}| < \frac{1}{k}$. This then means that $\lim_k \mu(T^{i_k} A \cap B) = \mu(A)\mu(B)$. Then it is possible to show that there is a single sequence that works for all measurable sets A, B. Using a standard approximation argument, this then can be used to show that (a_{i_k}) is a mixing sequence for measurable sets in $X \times X$, i.e., the mixing condition is satisfied for T along that sequence. This in turn implies that $T \times T$ is ergodic (in fact, it also implies that it is weakly mixing).

5. WEAK DOUBLE ERGODICITY.

We have known since the work of Parry and Kakutani [**24**] that many properties that are equivalent to the weak mixing property in the finite case do not remain equivalent in the infinite measure case. One can subdivide these properties as those that are stronger than Cartesian square ergodic (such as power weak mixing), and those that are weaker than Cartesian square ergodic, such as weak double ergodicity. (We say that T has ergodic Cartesian square when $T \times T$ is ergodic.)

A transformation T is called **weakly doubly ergodic** [**5**] if for all sets A, B of positive measure there exists an integer $n > 0$ so that

$$\mu(T^n A \cap A) > 0 \text{ and } \mu(T^n A \cap B) > 0.$$

The condition says that given two sets of positive measure, there is an iterate of the first set that intersects itself in positive measure, and also intersects the other set in positive measure. The crucial part here is that it is the same iterate in both cases. This property of course implies that the transformation is conservative and ergodic. It is interesting that in finite measure this condition already implies that $T \times T$ is ergodic, but it does not in infinite measure, as shown in work with undergraduates in [**5**]. The equivalence of this condition to weak mixing in finite measure was shown by Furstenberg [20], where we find the weak double ergodicity condition but without an explicit name. While we do not prove this equivalence, a proof that weak double ergodicity implies weak mixing is based on the fact that irrational rotations, while ergodic, are not weakly doubly ergodic. On the other hand, it can be easily verified, both in the finite and infinite measure cases, that if $T \times T$ is ergodic, then it is weakly doubly ergodic: if A and B are sets of positive measure, then there exists an integer n so that $(T \times T)^n(A \times A)$ intersects $A \times B$ in positive measure. It follows that $T^n A$ intersects both A and B in positive measure.

When α is irrational we have seen that T_α is ergodic. To see that T_α is not weakly doubly ergodic we choose two small intervals that are far apart (say $I = [0, 1/8)$ and $J = [1/2, 5/8)$); then for any n, when $T_\alpha^n I$ intersects I, as T_α acts by translating the interval mod 1 and does not break it, $T_\alpha^n I$ cannot intersect J, so T_α is not weakly doubly ergodic. It is easy to see that a mixing transformation T is weakly doubly ergodic as one can choose N so that $\mu(T^n A \cap A) > 0$ and $\mu(T^n A \cap B) > 0$ for all $n > N$.

The weak double ergodic condition was defined for infinite measure-preserving transformations in [5]. (It was also defined for nonsingular transformations—a transformation that, while not necessarily measure-preserving, still sends sets of positive measure to positive measure.) The condition was originally called "double ergodicity," but double ergodicity has also been used to denote the ergodicity of the Cartesian square (see the discussion in [21]), and to distinguish this notion we now use weak double ergodicity instead.

We now note that weak double ergodicity already implies a "higher version" of it. In contrast to this, in infinite measure it can happen that $T \times T$ is ergodic and $T \times T \times T$ is not. Furthermore, the 3-fold product can be ergodic while the 4-fold product is not, etc. [24]. More recently, it was shown that this behavior can also happen in the class of rank-one transformations [4]. On the other hand, in the case of weak double ergodicity it is interesting to note that the 2-fold case already implies the k-fold case as shown in [5]: If T is a transformation on X, then T is weakly doubly ergodic if and only if for all $A_i, B_i \subset X$, $1 \leq i \leq k$, of positive measure, there exists an integer $n > 0$ such that

$$\lambda(T^n A_i \cap B_i) > 0 \text{ for } i = 1, \ldots, k.$$

In [5], it is shown that in infinite measure there exist transformations that are weakly doubly ergodic, but do not have ergodic Cartesian square. A consequence of this is the advertised counterexample of Section 4. There exist infinite measure-preserving transformations T such that for all sets of positive measure A, B, C, D there is an integer $n > 0$ so that

$$\mu \times \mu[(T \times T)^n(A \times B) \cap (C \times D)] > 0$$

but $T \times T$ is not ergodic (in fact not even conservative [5]).

The examples are rank-one transformations and to describe them we first introduce the notion of staircase transformations.

It was very surprising when in 1972 Ornstein [31] found rank-one transformations (in finite measure) that are mixing. Ornstein's construction, though, was not explicit, but probabilistic in nature. He showed that there is a class of rank-one transformations so that almost surely a transformation in that class is mixing, but no explicit example was exhibited. It was much later when Adams [3] proved that a specific rank-one transformation, called the staircase transformation, is mixing. It is surprising to see mixing in the rank-one class because there is a way one can see rank-one transformations as closely approximated by rotations, and rotations are far from mixing. Also, we could

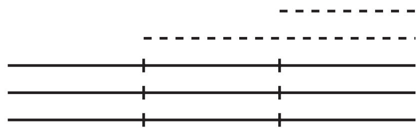

Figure 9. Column C_1 with spacers in the staircase transformation T_s.

think of mixing in a dynamical system as a measure of "randomness"; another measure of randomness is entropy (which we do not discuss here but roughly speaking is related to the rate at which orbits diverge), and rank-one transformations have zero entropy, so they are not random in the entropy sense.

We now define staircase transformations. Column C_0 consists of a single interval whose length is chosen so that in the end after all the spacers are added we end up with the unit interval. For the classical **staircase**, given column C_n cut each level into $r_n = n + 2$ subintervals. Number the subcolumns from 0 to $r_n - 1$; before stacking place i spacers (extra levels) above the ith subcolumn. So we start with column 0 consisting of a single level. Cut this level into two and place 0 spacers above the first half and 1 spacer above the second half, resulting in a column C_1 consisting of three intervals. Then cut C_1 into three subintervals, placing 0, 1, and 2 spacers above each subcolumn before stacking, etc. Figure 9 shows column C_1 with the spacers that are used to construct column C_2. We think of C_n with spacers placed on top of it in a staircase fashion. In the classical staircase case the number of subcolumns that C_n is cut into to construct C_{n+1} is $r_n = n + 2$; one can verify that the total measure of the space is finite. These ideas were used to construct other rank-one mixing examples in [**9**]. We note here that if r_n grows very fast, for example if $r_n = 2^{2^n}$, then the total measure of the space is infinite.

We note now that if r_n is bounded in a rank-one construction, then the transformation cannot be mixing. We give the argument as it is not complicated and illustrates some of the ideas used later. For concreteness, suppose $r_n = 3$ for all n. Then, for each n, column C_n is cut into three subcolumns, some number of spacers may be added above each of the subcolumns, and they are stacked right over left to form column C_{n+1}. Let s_n be the number of spacers that is added over the first subcolumn at stage n (it may be that $s_n = 0$). Let I be a level in C_n and let I_1, I_2, I_3 be the subintervals it is subdivided into when constructing the next column C_{n+1}. Then as I moves up the column under iteration by T, we note that after I_1 goes through the s_n spacers, it moves to the second subcolumn and after $h_n + s_n$ iterations ends up at I so that $T^{h_n+s_n} I_1 = I_2$. Then we see that $T^{h_n+s_n} I$ intersects I in measure at least $1/3\mu(I)$. For example, in Figure 9, if I is the bottom level of column C_1, we see that its left third, after three iterations (here there are no spacers over the left third, so $s_1 = 0$ and $h_1 = 3$) moves to the middle third of I. This pattern continues in later columns: level I in C_n has three descendants in C_{n+1} and each of these returns to itself under $T^{h_{n+1}+s_{n+1}}$ in measure at least 1/3 the measure of the descendant. Thus one can argue that for all $k \geq n$, if we set $\ell_k = h_k + s_k$, then $\mu(T^{\ell_k} I \cap I) \geq \frac{1}{3}\mu(I)$. As we can choose I so that $\mu(I) < 1/3$, then $\mu(T^{\ell_k} I \cap I)$ converges to a number bounded below by $\frac{1}{3}\mu(I) > \mu(I)^2$, so $\{\ell_k\}$ cannot be a mixing sequence for T; therefore the transformation is not mixing.

A transformation was defined in [**5**] to be a **tower staircase** if it is a rank-one staircase except that there is no restriction on the number of spacers that are placed in the last subcolumn. The staircase part of the construction is used to show that all tower staircases are weakly doubly ergodic. Furthermore, it is possible to choose a very large number of spacers at the last subcolumn so that levels are delayed a long time there and this can be used to show that the resulting transformation has nonconservative Cartesian product; of course, the large number of spacers forces infinite measure. Tower staircases that are Koopman mixing and power weakly mixing were constructed in [**12**]. More recently, it was shown in [**27**] that there are weakly doubly ergodic transformations with conservative but nonergodic Cartesian product.

Theorem 6 ([5]). *All tower staircase transformations are weak doubly ergodic. Furthermore, there exist tower staircases with nonconservative Cartesian product.*

Now we go back to the question of approximation and construct an example showing that approximation with intervals is not sufficient for ergodicity as characterized by Lemma 3. The reader may remember the construction of Cantor sets of positive measure. Using such sets one can construct the following interesting set in the unit interval. There is a set $M \subset [0, 1]$ such that every positive length interval I in $[0, 1]$ intersects both M and its complement M^c in positive measure: $\mu(I \cap M) > 0$ and $\mu(I \cap M^c) > 0$; see, e.g., [**30**] for a construction of M. Furthermore, there is a measurable bijection ϕ from M to the unit interval. The bijection ϕ can be given by $\phi(x) = \mu(M \cap [0, x))/\mu(M)$ (see [**30**] or [**8**, 8.3.6] for a general theorem). Using ϕ one can construct an ergodic transformation T_1 on M and another T_2 on M^c. Then let T_3 be the transformation on $[0, 1]$ that is T_1 on M and T_2 on M^c. Clearly T_3 is not ergodic (as M is an invariant set). We claim that for every positive length interval I, and measurable set B of positive measure, there exists an integer n so that $\mu(T^n I \cap B) > 0$. This follows since I intersects both M and M^c in positive measure, and B has to intersect at least one of M or M^c in positive measure. If $\mu(B \cap M) > 0$ we use the ergodicity of T_1 to get an n so that $\mu(T_1^n I \cap B \cap M) > 0$, and similarly for T_2 when $\mu(B \cap M^c) > 0$. So T_3 satisfies the ergodicity condition of Lemma 3 on intervals but is not ergodic. It is interesting to note that if T satisfies the ergodicity condition (2) on intervals, it does follow that it is ergodic. One can modify T_3 to construct an ergodic infinite measure-preserving transformation that is weakly doubly ergodic on intervals but not weakly doubly ergodic [**5**].

ACKNOWLEDGMENTS. The author would like to thank National Science Foundation REU grants, most recently grant DMS-1347804, and the Williams College Science Center, for support of the SMALL undergraduate research program at Williams College. I am also indebted to Terry Adams and Frank Morgan and to the referees for comments and suggestions.

REFERENCES

1. T. Adams, Uniformly sweeping out does not imply mixing, *Illinois J. Math.* **37** (1993) 525–536.
2. T. Adams, N. Friedman, C. E. Silva, Rank-one power weakly mixing non-singular transformations, *Ergodic Theory Dynam. Systems* **21** (2001) 1321–1332.
3. T. M. Adams, Smorodinsky's conjecture on rank-one mixing, *Proc. Amer. Math. Soc.* **126** (1998) 739–744.
4. T. M. Adams, C. E. Silva, On infinite transformations with maximal control of ergodic two-fold product powers, *Israel J. Math.* **209** (2015) 929–948.
5. A. Bowles*, L. Fidkowski*, A. E. Marinello*, C. E. Silva, Double ergodicity of nonsingular transformations and infinite measure-preserving staircase transformations, *Illinois J. Math.* **45** (2001) 999–1019.
6. R. V. Chacon, Weakly mixing transformations which are not strongly mixing, *Proc. Amer. Math. Soc.* **22** (1969) 559–562.
7. J. Clancy*, R. Friedberg*, I. Kasmalkar*, I. Loh*, T. Pădurariu*, C. E. Silva, S. Vasudevan*, Ergodicity and conservativity of products of infinite transformations and their inverses, *Colloq. Math.* **143** (2016) 271–291.
8. D. L. Cohn, *Measure Theory*. Birkhäuser Boston, Boston, MA, 1993. Reprint of the 1980 original.
9. D. Creutz*, C. E. Silva, Mixing on a class of rank-one transformations, *Ergodic Theory Dynam. Systems* **24** (2004) 407–440.

10. A. I. Danilenko, Funny rank-one weak mixing for nonsingular abelian actions, *Israel J. Math.* **121** (2001) 29–54.
11. ———, Finite ergodic index and asymmetry for infinite measure preserving actions, *Proc. Amer. Math. Soc.* **144** (2016) 2521–2532.
12. A. I. Danilenko, V. V. Ryzhikov, Mixing constructions with infinite invariant measure and spectral multiplicities, *Ergodic Theory Dynam. Systems* **31** (2011) 853–873.
13. A. I. Danilenko, C. E. Silva, Ergodic theory: Non-singular transformations, in *Mathematics of Complexity and Dynamical Systems. Vols. 1–3*, Springer, New York, 2012. 329–356.
14. S. L. Day*, B. R. Grivna*, E. P. McCartney*, C. E. Silva, Power weakly mixing infinite transformations, *New York J. Math.* **5** (1999) 17–24 (electronic).
15. C. Dodd*, P. Jeasakul*, A. Jirapattanakul*, D. M. Kane*, B. Robinson*, N. D. Stein*, C. E. Silva, Ergodic properties of a class of discrete abelian group extensions of rank-one transformations, *Colloq. Math.* **119** (2010) 1–22.
16. S. Eigen, A. Hajian, Y. Ito, V. Prasad, *Weakly Wandering Sequences in Ergodic Theory*. Springer Monographs in Mathematics. Springer, Tokyo, 2014.
17. N. A. Friedman, *Introduction to Ergodic Theory*. Van Nostrand Reinhold, New York, 1970.
18. ———, Replication and stacking in ergodic theory, *Amer. Math. Monthly* **99** (1992) 31–41.
19. N. A. Friedman, J. L. King, Rank one lightly mixing, *Israel J. Math.* **73** (1991) 281–288.
20. H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton Univ. Press, Princeton, NJ, 1981, M. B. Porter Lectures.
21. E. Glasner, B. Weiss, Weak mixing properties for non-singular actions, *Ergodic Theory Dynam. Systems* **36** (2016) 2203–2217.
22. E. Hopf, *Ergodentheorie*. Ergebnisse der Mathematik und ihrer Grenzgebiete. Vol. 5, no. 2, Springer, Berlin, 1937.
23. A. del Junco, Transformations with discrete spectrum are stacking transformations, *Canad. J. Math.* **28** (1976) 836–839.
24. S. Kakutani, W. Parry, Infinite measure preserving transformations with "mixing," *Bull. Amer. Math. Soc.* **69** (1963) 752–756.
25. Z. Kosloff, The zero-type property and mixing of Bernoulli shifts, *Ergodic Theory Dynam. Systems* **33** (2013) 549–559.
26. M. Lenci, On infinite-volume mixing, *Comm. Math. Phys.* **298** (2010) 485–514.
27. I. Loh*, C. E. Silva, Strict doubly ergodic infinite transformations (2016), https://arxiv.org/abs/1512.09340.
28. I. Melbourne, D. Terhesiu, Operator renewal theory and mixing rates for dynamical systems with infinite measure, *Invent. Math.* **189** (2012) 61–110.
29. E. J. Muehlegger*, A. S. Raich*, C. E. Silva, M. P. Touloumtzis*, B. Narasimhan*, W. Zhao*, Infinite ergodic index $\mathbf{Z}^d$-actions in infinite measure, *Colloq. Math.* **82** (1999) 167–190.
30. E. J. Muehlegger*, A. S. Raich*, C. E. Silva, W. Zhao*, Lightly mixing on dense algebras, *Real Anal. Exchange* **23** (1997/98) 259–265.
31. D. S. Ornstein, On the root problem in ergodic theory, in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*. Univ. California Press, Berkeley, 1972. 347–356.
32. T. Pădurariu*, C. E. Silva, E. Zachos*, On v-positive type transformations in infinite measure, *Colloq. Math.* **140** (2015) 149–170.
33. C. E. Silva, *Invitation to Ergodic Theory*. Student Mathematical Library, Vol. 42. American Mathematical Society, Providence, RI, 2008.

CESAR E. SILVA works in ergodic theory, a field he learned from Dorothy Maharam, his doctoral advisor at the University of Rochester. He is the Hagey Family Professor of Mathematics at Williams College. Silva is a co-founder of the SMALL undergraduate research program at Williams, funded by the NSF as an REU site. His interests outside mathematics include photography and piano.
Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267
csilva@williams.edu