# Mixing on rank-one transformations 

by<br>Darren Creutz (Los Angeles, CA)<br>and Cesar E. Silva (Williamstown, MA)


#### Abstract

We prove that mixing on rank-one transformations is equivalent to "the uniform convergence of ergodic averages (as in the mean ergodic theorem) over subsequences of partial sums". In particular, all polynomial staircase transformations are mixing.


## 1. Introduction

1.1. Rank-one transformations. Rank-one transformations are transformations "well-approximated" by a sequence of discrete spectrum transformations, so it was very surprising when in 1970 Ornstein Orn72 showed the existence of rank-one mixing transformations. Rank-one mixing transformations are mixing of all orders [Kal84], Ryz93 and enjoy other remarkable properties, see e.g. [Kin88]. Ornstein's construction was stochastic in nature: there is a class of rank-one transformations so that almost surely a transformation in that class is mixing; however, it did not yield a deterministic procedure for constructing one.
1.2. Staircase transformations. A few years later, Smorodinsky conjectured that a specific rank-one transformation, the classical staircase transformation, is mixing. In 1992, Adams and Friedman [AF92] gave a deterministic algorithm involving a sequence of cutting and stacking constructions that produced a mixing rank-one transformation, and later Adams Ada98 proved that Smorodinsky's conjecture is true.

Informally, a staircase transformation is a cutting and stacking transformation with a sequence $\left\{r_{n}\right\}$ of natural numbers such that at the $n$th stage the $n$th column or stack is cut into $r_{n}$ subcolumns, and "spacers" (see Section 3) are placed in a staircase fashion on the subcolumns before stacking, i.e., the number of spacers in each subsequent subcolumn is in-

[^0]creased by 1. Adams showed that the resulting staircase transformation is mixing provided that $r_{n}^{2} / h_{n} \rightarrow 0$ as $n \rightarrow \infty$ (which also implies that the transformation is finite-measure-preserving), where $h_{n}$ denotes the number of levels, or height, of the $n$th column. He then asked whether the mixing property holds for every finite-measure-preserving staircase transformation simply under the assumption that $r_{n} \rightarrow \infty$.

In 2003, Ryzhikov wrote the authors a short email stating that in 2000 he gave a lecture where he proved that all staircases are mixing (Theorem 4.1 below) Ryz03 (giving a positive answer to Adams' question); the result presented here was developed independently of his work and indeed we did not receive his proof sketch until after sending a preprint of this paper to him in 2005. We would also like to thank Ryzhikov for asking a question that clarified our writing of the definition of polynomial staircase transformations. The application of our main theorem shows that polynomial staircase transformations are mixing (Theorem 7.3). Specializing to the case of linear polynomials shows that all staircase transformations are mixing.
1.3. Restricted growth. The $r_{n}^{2} / h_{n} \rightarrow 0$ condition, a restriction on the asymptotic growth of the spacers relative to the column height, was generalized to all rank-one transformations and called "restricted growth" in CS04. The staircase transformation of Smorodinsky's conjecture is obtained when $r_{n}=n+1$; verifying that it satisfies the restricted growth condition is straightforward. In [CS04], the authors proved an equivalence between mixing and a condition on the spacer sequence for rank-one transformations with restricted growth. It followed that restricted growth rank-one transformations with the sequence of spacers given by a polynomial satisfying some general conditions (including the staircases of Ada98) are mixing. Ornstein's result also follows from that theorem.
1.4. Our result. In this paper we lift the restricted growth condition from the theorems in [CS04]. We present a self-contained proof of a condition equivalent to mixing for rank-one transformations involving the uniform convergence of certain averages of partial sums of the spacer sequence. Staircase and polynomial staircase transformations satisfy this condition.

## 2. Mixing properties

2.1. Dynamical systems. For our study, a dynamical system shall mean a standard probability measure space $(X, \mathcal{B}, \mu)$, and a transformation $T: X \rightarrow X$ a map that is invertible, measurable and measure-preserving. Throughout the paper, $(X, \mu)$ is $[0,1)$ under Lebesgue measure and $\mathcal{B}$ is the algebra of Lebesgue measurable subsets.
2.2. Mixing. A transformation $T$ is mixing when for all $A, B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \mu\left(T^{n}(A) \cap B\right)-\mu(A) \mu(B)=0
$$

$\left\{t_{n}\right\}$ is a mixing sequence (with respect to $T$ ) when for all $A, B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \mu\left(T^{t_{n}}(A) \cap B\right)-\mu(A) \mu(B)=0
$$

2.3. Ergodicity. A transformation $T$ is ergodic when for all $A \in \mathcal{B}$, if $T^{-1}(A)=A$ then $\mu(A)=0$ or $\mu(A)=1$. The mean (von Neumann) ergodic theorem states that $T$ is ergodic if and only if for all $B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{-j}\right| d \mu=0
$$

where we write $\chi_{B}=\mathbb{1}_{B}-\mu(B)\left(\mathbb{1}_{B}\right.$ being the characteristic function of the set $B$ ). A transformation $T$ is totally ergodic when for any $\ell \in \mathbb{N}, \ell \neq 0$, the transformation $T^{\ell}$ is ergodic. (We use the notation $\mathbb{N}=\{0,1,2, \ldots\}$ for the natural numbers and $\mathbb{Z}_{N}=\{0,1, \ldots, N-1\}$ for the $N$-element subset of $\mathbb{N}$ with the usual order.)
2.4. Ergodic sequences. Unless otherwise stated, the term sequence shall mean a sequence in $\mathbb{N}$ that is strictly increasing. A sequence $\left\{a_{n}\right\}$ is an ergodic sequence (with respect to a transformation $T$ ) when for all $B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{-a_{j}}\right| d \mu=0
$$

2.5. Power ergodicity. We introduce the concept of power ergodicity, all powers of an ergodic transformation being "uniformly" ergodic in the sense that the ergodic averages converge uniformly (to the projection onto the constants). Earlier results on specific rank-one mixing used precursors to this notion, including the uniform Cesàro property in AF92 (and implicitly in Ada98]) and power uniform ergodicity in (CS04.

Definition 2.1. A transformation $T$ is power ergodic when for all $B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \sup _{k \in \mathbb{N}} \int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{-j k}\right| d \mu=0
$$

## 3. Rank-one transformations

3.1. Cutting and stacking. Begin with $[0,1)$, the only "level" in the initial "column". "Cut" it into $r_{0}$ "sublevels", pieces of equal length: $\left[0,1 / r_{0}\right),\left[1 / r_{0}, 2 / r_{0}\right), \ldots,\left[r_{0}-1 / r_{0}, 1\right)$. Place an interval of the same length "above" $\left[1 / r_{0}, 2 / r_{0}\right)$, i.e., place $\left[1,\left(r_{0}+1\right) / r_{0}\right)$ above $\left[1 / r_{0}, 2 / r_{0}\right)$. Likewise, place $j$ "spacer" sublevels above each piece. Now, "stack" the resulting sub-
columns from left to right by placing $\left[0,1 / r_{0}\right)$ at the bottom, $\left[1 / r_{0}, 2 / r_{0}\right)$ above it, the spacer level above that, $\left[2 / r_{0}, 3 / r_{0}\right)$ above the spacer and so on, ending with the topmost of the $r_{0}-1$ spacers. This stack of $h_{1}=r_{0}+\sum_{j=0}^{r_{0}-1} j$ levels (of length $1 / r_{0}$ ), the second column, defines a map

$$
T_{0}:\left[0,1+\frac{1}{r_{0}} \sum_{j=0}^{r_{0}-1} j-\frac{1}{r_{0}}\right) \rightarrow\left[\frac{1}{r_{0}}, 1+\frac{1}{r_{0}} \sum_{j=0}^{r_{0}-1} j\right)
$$

that sends points directly up one level.
Repeat the process: cut the entire new column into $r_{1}$ subcolumns of equal width $1 / r_{0} r_{1}$, preserving the stack map on each subcolumn; place $j$ spacers (intervals not yet in the space, the same width as the subcolumns) above each subcolumn $\left(j \in \mathbb{Z}_{r_{1}}\right)$; and stack the resulting subcolumns from left to right. Our new column defines a map $T_{1}$ that agrees with $T_{0}$ where it is defined and extends it to all but the topmost spacer of the rightmost subcolumn. Iterating this process leads to a transformation $T$ defined on all but a Lebesgue measure zero set.

The transformations obtained in this manner are called staircase transformations. More generally, one may place $s_{n, j}$ spacers above the $j$ th subcolumn at the $n$th stage in place of the $j$ spacers above. A transformation created by cutting and stacking as just described (with a single column resulting from each iteration) is a rank-one transformation. The reader is referred to [Fer97] and [Fri70] for more details. Rank-one transformations are measurable and measure-preserving under Lebesgue measure, and are completely defined by the doubly-indexed sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ where at the $n$th step we cut into $r_{n}$ pieces and place $s_{n, j}$ spacers above each subcolumn (for staircase transformations, $s_{n, j}=j$ ). This $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ is the spacer sequence for the transformation and $\left\{r_{n}\right\}$ is the cut sequence. The height sequence $\left\{h_{n}\right\}$ is the number of levels in each column: $h_{0}=1$ and $h_{n+1}=r_{n} h_{n}+\sum_{j=0}^{r_{n}-1} s_{n, j}$. It is well-known (and left to the reader) that if $\lim \inf r_{n}<\infty$ then the transformation will be partially rigid hence cannot be mixing. We shall assume from now on that $\lim r_{n}=\infty$.

We write $I_{n, i}$ to denote the $i$ th level in the $n$th stack $\left(i \in \mathbb{Z}_{h_{n}}\right)$ where $I_{n, 0}$ is the bottom level and $T\left(I_{n, i}\right)=I_{n, i+1}$, and write $C_{n}=\bigcup_{i=0}^{h_{n}-1} I_{n, i}$ to denote the $n$th column and $S_{n}=C_{n+1} \backslash C_{n}$ to denote the spacers added. We write $I_{n, i}^{[j]}$ for the $j$ th sublevel of the $i$ th level of the $n$th column, i.e., $I_{n, 0}^{[0]}$ is the leftmost sublevel of the bottom level $\left(I_{n, 0}^{[0]}=I_{n+1,0}\right.$ becomes the bottom level of the next column). Note that $T$ is defined on a finite-measure space if and only if $\sum_{n=0}^{\infty} \mu\left(S_{n}\right)<\infty$, and in that case $T$ is isomorphic to the transformation defined on $[0,1)$ obtained by cutting and stacking in the same fashion as $T$ but beginning with $C_{0}=[0,1 / K)$ where $K$ is the measure of the space the original $T$ is defined on.
4. Mixing on staircase transformations. We shall first prove directly that staircase transformations are mixing, to illustrate the techniques used. Staircases are the simplest mixing rank-one transformations and serve as a model for the general case. As mentioned in the introduction, the following theorem was proved by Adams under a growth restriction Ada98 and announced in full generality by Ryzhikov in 2000 (unpublished).

Theorem 4.1. Let $T$ be a staircase transformation. Then $T$ is mixing.
In what follows we shall first outline the ideas and then present the formal proof. Consider a level $I$ in the $n$th column defining a rank-one transformation that is at least $r_{n}$ above the bottom. Look at $T^{h_{n}}$ applied to the $j$ th sublevel of $I$. There are $j$ spacers added above that subcolumn so $T^{h_{n}}$ will map the $j$ th sublevel of $I$ to the $(j+1)$ th sublevel of the level $j$ below $I$. This means that $T^{h_{n}}$ maps $I$ to a progression of $1 / r_{n}$ sized parts of $r_{n}$ consecutive levels. So the characteristic function of $T^{h_{n}}(I)$ is in fact an average of characteristic functions of consecutive levels, which is to say that it is of the form $r_{n}^{-1} \sum_{j=0}^{r_{n}-1} \mathbb{1}_{I} \circ T^{-j}$. The ergodicity of $T$ then guarantees that this quantity tends to zero so in fact the sequence $\left\{h_{n}\right\}$ will be mixing.

In what follows, this idea of turning a sequence into an ergodic type average will be the primary ingredient. We first establish that for an arbitrary sequence, the mixing behavior is controlled by specific ergodic type averages. We then prove that for staircase transformations, any ergodic type average along arithmetic progressions tends to zero.

The following lemma formalizes our discussion of sublevels above:
Lemma 4.2. Let $p \in \mathbb{N}, i \in \mathbb{Z}_{h_{p}}, j \in \mathbb{Z}_{r_{p}}$, and $B$ a union of levels in $C_{p}$. Then
(i) $I_{p, i}=\bigcup_{j=0}^{r_{p}-1} I_{p, i}^{[j]}$;
(ii) $T^{k h_{p}+j k+\frac{1}{2} k(k-1)} I_{p, i}^{[j]}=I_{p, i}^{[j+k]}$;
(iii) $\mu\left(I_{p, i}^{[j]} \cap B\right)=\frac{1}{r_{p}} \mu\left(I_{p, i} \cap B\right)$.

Proof. (i) is immediate from the construction of rank-one transformations; (ii) follows from $k$ applications of the fact that $T^{h_{p}+j} I_{p, i}^{[j]}=I_{p, i}^{[j+1]}$; and for (iii), $B$ is a union of levels in $C_{p}$ so $I_{p, i} \subseteq B$ or $I_{p, i} \cap B=\emptyset$.

The first step in our proof of mixing is the well-established result on the height sequence:

Proposition 4.3 (Ada98]). Let $T$ be a staircase transformation with height sequence $\left\{h_{n}\right\}$. Then $\left\{h_{n}\right\}$ is mixing with respect to $T$.

This a direct consequence of Theorem 3 in CS04 applied to staircase transformations; alternatively, it follows from the proof of mixing on "D1" in Theorem 2.3 of Adams Ada98] (the restricted growth condition is not used in that part of the proof).

The following technical lemma, in combination with Lemma 4.2(ii), shows that if we look at a block of $\ell$ consecutive sublevels of some level $I$ then after applying $T^{k h_{n}}$ and dropping the levels at the bottom of the column we are left with an ergodic type average of powers of $T^{k}$ exactly as what occurs for $h_{n}$ but weighted by the size of the block $\ell / r_{n}$.

Lemma 4.4. For $p, k \in \mathbb{N}, \ell \in \mathbb{Z}_{r_{p}}$ and $c \in \mathbb{Z}_{r_{p}-\ell}$,

$$
\begin{aligned}
\sum_{i=0}^{h_{p}-1} \mid \sum_{j=0}^{\ell-1} \mu\left(T^{-j k} I_{p, i}^{[j+c]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) & \mu(B) \mid \\
& \leq \frac{\ell}{r_{p}} \int\left|\frac{1}{\ell} \sum_{j=0}^{\ell-1} \chi_{B} \circ T^{-j k}\right|+\frac{\ell^{2} k}{r_{p} h_{p}} .
\end{aligned}
$$

Proof. For $i \geq \ell k$,

$$
\mu\left(T^{-j k} I_{p, i}^{[j+c]} \cap B\right)=\frac{1}{r_{p}} \mu\left(T^{-j k} I_{p, i} \cap B\right)
$$

by Lemma 4.2 (iii). And for $i<\ell k$,

$$
\left|\sum_{j=0}^{\ell-1} \mu\left(T^{-j k} I_{p, i}^{[j+c]} \cap B\right)-\mu\left(I_{p, i}^{[j+c]}\right) \mu(B)\right| \leq \ell \mu\left(I_{p, i}^{[0]}\right) \leq \frac{\ell}{r_{p} h_{p}}
$$

since $\mu\left(C_{p}\right) \leq 1$.
Consider now the ergodic type average of $r_{n}$ powers of $T^{k(n)}$, that is, let $k$ vary with $n$. It is clear from the construction of rank-one transformations by columns that if $k(n) h_{p} \leq t_{n} \leq(k(n)+1) h_{p}$ then the behavior of $T^{t_{n}}$ is controlled by that of $T^{k(n) h_{p}}$ and $T^{(k(n)+1) h_{p}}$. The following proposition states that if we know a priori that the ergodic type averages of the powers of $T^{k(n)}$ tend to zero and furthermore that $k(n)$ is small enough that we may drop the bottom $k(n)$ levels (these levels have vanishing measure), then following the same procedure as above, $T^{k(n) h_{p}}$ will be mixing and so then will be the sequence $\left\{t_{n}\right\}$.

Definition 4.5. In the context of a rank-one transformation with heights $\left\{h_{n}\right\}$ and cuts $\left\{r_{n}\right\}$, given $t \in \mathbb{N}$, the unique $p$ and $k$ for $t$ are the unique numbers $p, k \in \mathbb{N}$ such that $h_{p} \leq t<h_{p+1}$ and $k h_{p} \leq t<(k+1) h_{p}$.

Note that since $k<h_{p+1} / h_{p}$ and by the finite-measure-preserving prop$\operatorname{erty} h_{p+1} / h_{p} r_{p}=\mu\left(C_{p+1}\right) / \mu\left(C_{p}\right) \rightarrow 1$, we have $k / r_{p}<h_{p+1} / h_{p} r_{p} \rightarrow 1$.

In what follows, when we say "choose $p$ and $k$ " we shall mean the above construction. In particular, we shall often assume that $k \leq r_{p}$ and leave to the reader to verify that in fact knowing $k \leq(1+\epsilon) r_{p}$ for small $\epsilon$ suffices.

Proposition 4.6. Let $T$ be a staircase transformation with cut sequence $\left\{r_{n}\right\}$ and height sequence $\left\{h_{n}\right\}$ and let $\left\{t_{n}\right\}$ be a sequence. Choose $p(n)$ and $k(n)$ such that $h_{p(n)} \leq t_{n}<h_{p(n)+1}$ and $k(n) h_{p(n)} \leq t_{n}<(k(n)+1) h_{p(n)}$. If $r_{p(n)} k(n) / h_{p(n)} \rightarrow 0$ and for any $B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \int\left|\frac{1}{r_{p(n)}} \sum_{j=0}^{r_{p(n)}-k(n)-1} \chi_{B} \circ T^{-j k(n)}\right| d \mu=0
$$

and likewise with $k(n)$ replacing $k(n)+1$, then $\left\{t_{n}\right\}$ is mixing with respect to $T$.

Proof. The idea here is to write $k(n)$ and $h_{p}$ as above; the proposition then follows from the fact that $T^{k_{n} h_{p}}$ sends a level in $C_{p}$ to $1 / r_{p}$ parts of a progression of $r_{p}-k_{n}$ levels that are all $k_{n}$ apart so the ergodic type average of the powers of $T^{k_{n}}$ is exactly the condition needed.

We now make this precise. Here, and in the proofs below, we shall sometimes drop the explicit dependence on $n$ : we write $p=p(n)$ and $k=k(n)$. Pick $m=m(n)$ such that $t_{n}=k h_{p}+m$ where $0 \leq m<h_{p}$. Let $A, B$ be unions of levels in $C_{N}$ for some $N \in \mathbb{N}$. For $n$ such that $p \geq N$, as in Proposition 4.3 ,

$$
\begin{aligned}
&\left|\mu\left(T^{t_{n}} A \cap B\right)-\mu(A) \mu(B)\right| \\
& \leq \sum_{i=0}^{h_{p}-1}\left|\mu\left(T^{k h_{p}+m} I_{p, i} \cap B\right)-\mu\left(I_{p, i}\right) \mu(B)\right| \\
&= \sum_{i=0}^{h_{p}-m-1}\left|\mu\left(T^{k h_{p}} I_{p, i+m} \cap B\right)-\mu\left(I_{p, i+m}\right) \mu(B)\right| \\
&+\sum_{i=h_{p}-m}^{h_{p}-1}\left|\mu\left(T^{(k+1) h_{p}} I_{p, i+m-h_{p}} \cap B\right)-\mu\left(I_{p, i+m-h_{p}}\right) \mu(B)\right| \\
& \leq \sum_{i=0}^{h_{p}-1}\left|\mu\left(T^{k h_{p}} I_{p, i} \cap B\right)-\mu\left(I_{p, i}\right) \mu(B)\right| \\
&+\sum_{i=0}^{h_{p}-1}\left|\mu\left(T^{(k+1) h_{p}} I_{p, i} \cap B\right)-\mu\left(I_{p, i}\right) \mu(B)\right| .
\end{aligned}
$$

Now by Lemma 4.2(i) and the triangle inequality,

$$
\begin{aligned}
\sum_{i=0}^{h_{p}-1} \mid \mu\left(T^{k h_{p}} I_{p, i} \cap B\right)- & \mu\left(I_{p, i}\right) \mu(B) \mid \\
\leq & \sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{r_{p}-k-1} \mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right| \\
& +\sum_{i=0}^{h_{p}-1} \sum_{j=r_{p}-k}^{r_{p}-1}\left|\mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right|
\end{aligned}
$$

and $T^{k h_{p}} I_{p, i}^{[j]}=T^{(k+j) h_{p}+\frac{1}{2} j(j-1)} I_{p+1, i}$, so for $j \geq r_{p}-k$ we have $(k+j) h_{p}+$ $\frac{1}{2} j(j-1) \geq r_{p} h_{p}=h_{p+1}-\frac{1}{2} r_{p}\left(r_{p}-1\right)$ so that in fact $T^{k h_{p}} I_{p, i}^{[j]}=T^{h_{p+1}} I_{p, a}^{[b]}$ for some $a$ and $b$, and $I_{p, a}^{[b]}$ is a level in $C_{p+1}$, hence

$$
\begin{aligned}
& \sum_{i=0}^{h_{p}-1} \sum_{j=r_{p}-k}^{r_{p}-1}\left|\mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right| \\
& \quad \leq \sum_{i=0}^{h_{p+1}-1}\left|\mu\left(T^{h_{p+1}} I_{p+1, i} \cap B\right)-\mu\left(I_{p+1, i}\right) \mu(B)\right|+\frac{1}{2} r_{p}\left(r_{p}-1\right) \mu\left(I_{p+1,0}\right) \\
& \quad \leq \sum_{i=r_{p+1}}^{h_{p+1}-1}\left|\frac{1}{r_{p+1}} \sum_{j=0}^{r_{p+1}-1} \mu\left(T^{-j} I_{p+1, i} \cap B\right)-\mu\left(I_{p+1, i}\right) \mu(B)\right| \\
& \quad+\sum_{i=r_{p+1}}^{h_{p+1}-1} \frac{2}{r_{p+1}} \mu\left(I_{p+1, i}\right)+\sum_{i=0}^{r_{p+1}-1} \mu\left(I_{p+1, i}\right)+\frac{1}{2} r_{p}\left(r_{p}-1\right) \mu\left(I_{p+1,0}\right) \\
& \quad \leq \int\left|\frac{1}{r_{p+1}} \sum_{j=0}^{r_{p+1}-1} \chi_{B} \circ T^{-j}\right| d \mu+\frac{2}{r_{p+1}}+\frac{r_{p+1}}{h_{p+1}}+\frac{1}{2} r_{p}\left(r_{p}-1\right) \mu\left(I_{p+1,0}\right)
\end{aligned}
$$

which approaches zero since $\frac{1}{2} r_{p}\left(r_{p}-1\right) \mu\left(I_{p+1,0}\right)=\mu\left(S_{p}\right) \rightarrow 0$ and $T$ is ergodic.

For $\frac{1}{2} k(k-1) \leq i<h_{p}$, using Lemma 4.2 (ii), we see that $T^{k h_{p}} I_{p, i}^{[j]}=$ $T^{-j k} I_{p, i-\frac{1}{2} k(k-1)}^{[j+k]}$ so by Lemma 4.4 ,

$$
\begin{aligned}
& \sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{r_{p}-k-1} \mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right| \\
& \quad \leq \int\left|\frac{1}{r_{p}} \sum_{j=0}^{r_{p}-k-1} \chi_{B} \circ T^{-j k}\right| d \mu+\frac{\left(r_{p}-k\right)^{2} k}{r_{p} h_{p}}+\frac{1}{2} k(k-1) \frac{r_{p}-k}{r_{p}} \mu\left(I_{p, 0}\right)
\end{aligned}
$$

Now

$$
\frac{\left(r_{p}-k\right)^{2} k}{r_{p} h_{p}} \leq \frac{r_{p} k}{h_{p}} \rightarrow 0 \quad \text { and } \quad \frac{1}{2} k(k-1) \frac{r_{p}-k}{r_{p}} \mu\left(I_{p, 0}\right) \leq \frac{k^{2}}{h_{p}} \leq \frac{k r_{p}}{h_{p}} \rightarrow 0
$$

as $n \rightarrow \infty$ by hypothesis and the integral goes to zero by the final hypothesis. The above repeats similarly for the $k+1$ part.

We now examine the case of a sequence where the $k(n)$ derived as above is not small enough to simply drop the bottommost levels:

Proposition 4.7. Let $T$ be a staircase transformation with cut sequence $\left\{r_{n}\right\}$ and height sequence $\left\{h_{n}\right\}$ and let $\left\{t_{n}\right\}$ be a sequence. Choose $p(n)$ and $k(n)$, as usual, such that $k(n) h_{p(n)} \leq t_{n}<(k(n)+1) h_{p(n)}$. If there exists a sequence $\{l(n)\}$ such that $l(n) \rightarrow \infty, l(n) k(n) / h_{p(n)} \rightarrow 0$ and $(l(n)) / r_{p(n)} \rightarrow 0$ and such that for any $\alpha(n, q) \in \mathbb{N}$ indexed over

$$
0 \leq q<Q_{n}=\left\lceil\frac{r_{p(n)}-k(n)}{l(n)}\right\rceil \quad \text { with } \quad \limsup _{n} \sup _{q} \frac{\alpha(n, q)}{k(n)}<1
$$

and $\alpha(n, q) \leq \alpha(n, q+1) \leq \alpha(n, q)+1$,

$$
\lim _{n \rightarrow \infty} \frac{1}{Q_{n}} \sum_{q=0}^{Q_{n}-1} \int\left|\frac{1}{l(n)} \sum_{j=0}^{l(n)-1} \chi_{B} \circ T^{-j(k(n)+2-\alpha(n, q))}\right| d \mu=0
$$

then $\left\{t_{n}\right\}$ is mixing with respect to $T$.
Proof. The idea here is to break the sublevels up into blocks of size small enough that we can drop the bottommost levels for each block separately and then apply our above methods to each block. Our lemma on blocks of sublevels tells us that each block will be weighted by the size of the block, so if each block's ergodic type average tends to zero then the entire quantity will tend to zero as well.

The reason we have been dropping the bottommost levels above is that the behavior of the sublevels of those bottom levels is not to form an evenly spaced progression on the levels but in fact to form two or more such progressions by "coming back through the top" of the column. This coming back through the top can be avoided if we can ensure that the $k(n)$ we are using on a given block has the property that $j k(n)$ is small compared to $h_{n}$ for $j=0, \ldots, l$ where $l$ is the size of the block. The following proposition states and proves this formally: if the blocks we use are of size $l(n)$ and for any $k(n)-\alpha$ corresponding to the number of $h_{n}$ 's we need to remove when examining each block ( $\alpha$ depends on the block), then if the ergodic type average tends to zero, the original sequence will be mixing. The conditions required on the $l(n)$ are necessary to perform the step of dropping the bottommost levels for each block (independently).

Now we make this precise: as before, we shall drop the explicit mention of $n$ and write $p=p(n), k=k(n), l=l(n), Q=Q_{n}$ and $\alpha(\cdot)=\alpha(n, \cdot)$. Let $B$ be a union of levels in $C_{N}$ for some $N \in \mathbb{N}$. Then, as in the proof of Proposition 4.6, it suffices to show that

$$
\sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{r_{p}-k-1} \mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right| \rightarrow 0
$$

and similarly for $k+1$. Now

$$
\begin{aligned}
& \sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{r_{p}-k-1} \mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right| \\
& \quad \leq \sum_{i=0}^{h_{p}-1} \sum_{q=0}^{Q-1}\left|\sum_{j=0}^{l-1} \mu\left(T^{k h_{p}} I_{p, i}^{[j+q l]} \cap B\right)-\mu\left(I_{p, i}^{[j+q l]}\right) \mu(B)\right|+\sum_{i=0}^{h_{p}-1} l \mu\left(I_{p, i}^{[0]}\right)
\end{aligned}
$$

and $\sum_{i=0}^{h_{p}-1} l \mu\left(I_{p, i}^{[0]}\right)=\left(l / r_{p}\right) \mu\left(C_{p}\right) \rightarrow 0$ by hypothesis. For $0 \leq q<Q$, choose $\alpha(q)=\alpha(n, q)$ minimal such that $q l k+\frac{1}{2} k(k-1)<\alpha(q)\left(h_{p}+q l+\right.$ $\left.k-\frac{1}{2}\right)-\frac{1}{2} \alpha(q)^{2}$. This reasoning for this choice of $\alpha$ is that $T^{k h_{p}}$ acts like an average of $T^{-j(k-\alpha)}$ due to the fact that the spacer levels being added are not small compared to $h_{p}$, that is, $T^{k h_{p}}$ would push levels through $k$ times except that the spacers remove $\alpha$ of those times; we have chosen $\alpha$ so that $T^{k h_{p}} I_{p, i}^{[q l]}=T^{w} I_{p, i}^{[q l+k-\alpha(q)]}$ and $w>0$, and $\alpha$ is minimal. The $\alpha(q)$ are clearly nondecreasing and since $l r_{p} / h_{p} \rightarrow 0$ (by hypothesis) we have $\alpha(q+1)-\alpha(q) \leq 1$ for large $n$.

Since $\alpha(q)$ is chosen minimally, consider using $\frac{1}{2} k$ in the defining inequality for $\alpha(q)$ : since $q l \leq Q l \leq\left(\left(r_{p}-k\right) / l+1\right) l=r_{p}-k+l$ we have

$$
\begin{aligned}
\frac{1}{2} k\left(h_{p}+q l\right. & \left.+k-\frac{1}{2}\right)-\frac{1}{4} k^{2}-q l k-\frac{1}{2} k(k-1) \\
& =\frac{1}{2} k\left(h_{p}-q l-\frac{1}{2}-\frac{3}{4} k\right) \geq \frac{1}{2} k\left(h_{p}-r_{p}-\frac{1}{2}+\frac{1}{4} k-l\right)
\end{aligned}
$$

and since $l / r_{p} \rightarrow 0$ and $r_{p} / h_{p} \rightarrow 0$, in the limit this quantity is nonnegative. Therefore $\lim \sup \alpha(q) / k \leq 1 / 2$ uniformly over $q$.

Set $\beta(q)=\alpha(q) h_{p}-q l(k-\alpha(q))-\frac{1}{2}(k-\alpha(q))(k-\alpha(q)-1)$ so that $0 \leq \beta(q)<h_{p}+q l+k+1 / 2$. Then $\beta(q)$ roughly corresponds to the height in the column where applying $T^{k h_{p}}$ switches between acting like $T^{(k-\alpha(q)) h_{p}}$ and like $T^{(k-\alpha(q)-1) h_{p}}$. For $0 \leq j<l$ and $0 \leq i<h_{p}$, by Lemma 4.2 (ii),

$$
T^{k h_{p}} I_{p, i}^{[j+q l]}=T^{-j(k-\alpha(q))+\beta(q)} I_{p, i}^{[j+q l+k-\alpha(q)]}
$$

For $0 \leq i<h_{p}-\beta(q)$, we have

$$
T^{-j(k-\alpha(q))+\beta(q)} I_{p, i}^{[j+l+k-\alpha(q)]}=T^{-j(k-\alpha(q))} I_{p, i+\beta(q)}^{[j+q l+k-\alpha(q)]}
$$

and for $h_{p}-\beta(q)+q l-\alpha(q)+k \leq i<h_{p}$, we have

$$
\begin{aligned}
T^{-j(k-\alpha(q))+\beta(q)} I_{p, i}^{[j+l+k-\alpha(q)]} & =T^{-j(k-\alpha(q))+h_{p}} I_{p, i+\beta(q)-h_{p}}^{[j+q l+k-\alpha(q)]} \\
& =T^{-j(k-\alpha(q)+1)} I_{p, i+\beta(q)-h_{p}-q l-k+\alpha(q)}^{[j+q l+k-\alpha(q)+1]}
\end{aligned}
$$

Hence for each $q$,

$$
\begin{aligned}
& \sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{l-1} \mu\left(T^{k h_{p}} I_{p, i}^{[j+q l]} \cap B\right)-\mu\left(I_{p, i}^{[j+q l]}\right) \mu(B)\right| \\
& \quad \leq \sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{l-1} \mu\left(T^{-j(k-\alpha(q))} I_{p, i}^{[j+q l+k-\alpha(q)]} \cap B\right)-\mu\left(I_{p, i}^{[j+q l]}\right) \mu(B)\right| \\
& \quad+\sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{l-1} \mu\left(T^{-j(k-\alpha(q)+1)} I_{p, i}^{[j+q l+k-\alpha(q)+1]} \cap B\right)-\mu\left(I_{p, i}^{[j+q l]}\right) \mu(B)\right| \\
& \quad+(q l+k-\alpha(q)) \frac{l}{r_{p}} \mu\left(I_{p, 0}\right),
\end{aligned}
$$

and since $q l+k-\alpha(q) \leq r_{p}$,

$$
\sum_{q=0}^{Q-1}(q l+k-\alpha(q)) \frac{l}{r_{p}} \mu\left(I_{p, 0}\right) \leq \frac{Q l r_{p}}{r_{p} h_{p}} \leq \frac{r_{p}}{h_{p}} \rightarrow 0
$$

By Lemma 4.4 ,

$$
\begin{aligned}
& \sum_{q=0}^{Q-1} \sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{l-1} \mu\left(T^{-j(k-\alpha(q))} I_{p, i}^{[j+q l+k-\alpha(q)]} \cap B\right)-\mu\left(I_{p, i}^{[j+q l+k-\alpha(q)]}\right) \mu(B)\right| \\
& \leq \sum_{q=0}^{Q-1} \frac{l}{r_{p}} \int\left|\frac{1}{l} \sum_{j=0}^{l-1} \chi_{B} \circ T^{-j(k-\alpha(q))}\right| d \mu+\sum_{q=0}^{Q-1} \frac{l^{2}(k-\alpha(q))}{r_{p} h_{p}}
\end{aligned}
$$

Note that $Q \leq\left(r_{p}-k\right) / l+1 \leq 2 r_{p} / l$ by hypothesis, so

$$
\sum_{q=0}^{Q-1} \frac{l^{2}(k-\alpha(q))}{r_{p} h_{p}} \leq \frac{l^{2} Q k}{r_{p} h_{p}} \leq \frac{2 l k}{h_{p}} \rightarrow 0
$$

by hypothesis, and applying the above also to the $i \geq h_{p}-\beta(q)+q l+k-\alpha(q)$ case yields

$$
\begin{aligned}
& \sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{r_{p}-k-1} \mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right| \\
& \leq \frac{1}{Q} \sum_{q=0}^{Q-1} \int\left|\frac{1}{l} \sum_{j=0}^{l-1} \chi_{B} \circ T^{-j(k-\alpha(q))}\right| d \mu \\
&+\frac{1}{Q} \sum_{q=0}^{Q-1} \int\left|\frac{1}{l} \sum_{j=0}^{l-1} \chi_{B} \circ T^{-j(k-\alpha(q)+1)}\right| d \mu+\gamma
\end{aligned}
$$

where $\gamma=\gamma_{n} \rightarrow 0$ is the sum of all the terms above that tend to zero. Thus the above quantity goes to zero by the final hypothesis. Applying the entire argument again to the $k+1$ case implies that $\left\{t_{n}\right\}$ is mixing.

REMARK 4.8. The $\alpha(q)$ range from 0 to $R \approx h_{p} / Q$ (all quantities here of course depend on $n$ ) and each particular value is taken on with approximately the same density over $j$. This is due to the fact that $\alpha(q)$ is roughly $j k / h_{p}$.

Having obtained our above results that mixing on sequences will follow from convergence to zero of ergodic type averages of $T^{k}$ where $k$ is allowed to move with $n$ (the number of terms being averaged), we now show that such averages do in fact converge to zero.

We shall need the Block Lemma, our next statement, which tells us that an ergodic type average of powers of $T^{k}$ will be dominated by an ergodic type average of powers of $T^{k q}$ with proportionally fewer terms.

Lemma 4.9 (Block Lemma, Ada98). Let $T$ be a measure-preserving transformation and $B \in \mathcal{B}$. Then for any $R, L, q \in \mathbb{N}$,

$$
\int\left|\frac{1}{R} \sum_{j=0}^{R-1} \chi_{B} \circ T^{-j}\right| d \mu \leq \int\left|\frac{1}{L} \sum_{j=0}^{L-1} \chi_{B} \circ T^{-j q}\right| d \mu+\frac{q L}{R}
$$

Proof. The main idea is to split the sum into blocks of size $L p$ and then each $L p$ block into $L$ blocks and use the measure-preserving property to combine terms. Details are left to the reader.

The case when $k(n)$ is bounded follows directly from the total ergodicity of $T$ (which holds since $T$ has a mixing sequence). Our next proposition is that if $k(n) / n$ is bounded then the ergodic type averages of $n$ consecutive powers of $k(n)$ converge to zero.

Proposition 4.10. Let $T$ be a staircase transformation. Then for any $\left\{k_{n}\right\}$ such that $\lim \sup _{n} k_{n} / n<\infty$ and any $B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{j k_{n}}\right| d \mu=0
$$

This follows from the proof of Lemma 2.2 in Adams Ada98 (and appears in [CS04]). The idea is to use the Block Lemma to multiply $k_{n}$ by a number $q_{n}$ so that $k_{n} q_{n} \approx h_{p}$ for some $p$. Then the mixing behavior of $h_{n}$ will yield the convergence of the average of (any number of terms of) consecutive powers of $T^{k_{n} q_{n}}$, which by the Block Lemma dominates the average of powers of $T^{k_{n}}$. We will need that $k_{n} / n$ is bounded to ensure that the terms in the sum dropped by the Block Lemma (the $p L / R$ term) are small.

The following two lemmas also follow from the proof of Lemma 2.2 in Ada98 (using the "Blum-Hanson" trick). The reader is referred to the proof of Proposition 7.6 for a (combined) proof of Proposition 4.10 and the following two lemmas in a slightly more general context.

Lemma 4.11. Let $T$ be a staircase transformation with heights $\left\{h_{n}\right\}$ and let $\left\{t_{n}\right\}$ and $\left\{p_{n}\right\}$ be sequences such that $h_{p_{n}} \leq t_{n}<2 h_{p_{n}}$. Then for any $\epsilon>0$ there exists $L$ such that for all sufficiently large $n$,

$$
\int\left|\frac{1}{L} \sum_{d=0}^{L-1} \chi_{B} \circ T^{-d t_{n}}\right| d \mu<\epsilon
$$

LEMMA 4.12. Let $T$ be a staircase transformation and $\left\{t_{n}\right\}$ a sequence such that for any fixed $d$ the sequence $\left\{d t_{n}\right\}$ is mixing with respect to $T$. Then for any $\epsilon>0$, for sufficiently large $n$,

$$
\int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{-j t_{n}}\right| d \mu<\epsilon
$$

We now establish the final result needed to prove staircases are mixing: that the ergodic type average of powers of $T^{k_{n}}$ will converge to zero regardless of how fast $k_{n}$ grows. This property is of some independent interest and is referred to as power ergodicity.

Proposition 4.13. Let $T$ be a staircase transformation. Then $T$ is power ergodic.

Proof. For any $k$ we find the largest $p$ such that $k \approx x h_{p}$ for some positive integer $x$. The proof is by cases. In the case when $x / r_{p}$ is bounded away from zero, we may use the Block Lemma to multiply $k$ by $q=r_{p} / x$ so that $k q \approx h_{p+1}$. In the case when $x r_{p} / h_{p}$ is small, $x / r_{p}$ will be bounded so the averages of the powers of $T^{x}$ will vanish. Then $\left\{k_{n}\right\}$ will be mixing (Proposition 4.6) so its average must converge to zero.

Having disposed of these cases, if $x r_{p} / h_{p}$ is bounded then we can pick $l(n)$ satisfying the conditions of Proposition 4.7 and so that $x / l(n)$ is bounded. Then Proposition 4.10 tells us the average of powers of $T^{x}$ vanishes, so by Proposition 4.7, $\{k(n)\}$ is mixing. The final case is when $x r_{p} / h_{p} \rightarrow \infty$. We again find a sequence $l(n)$ and break the average into blocks of size $l(n)$ but
in this case the values of $\alpha$ we obtain (as in the statement of Proposition 4.7) will effectively take on all values. We then use that $T$ is weak mixing to show that the average of averages of powers of $T^{x-\alpha}$ tends to zero and so again $\left\{k_{n}\right\}$ is mixing.

Now we make this precise. Let $\left\{k_{n}\right\}$ be an arbitrary sequence. Write $k_{n}=k=x h_{p}+m$ as usual, dropping the $n$ and picking $p$ uniquely so that $h_{p} \leq k<h_{p+1}$ and then $x$ and $m$ accordingly.

Case 1a: $\liminf _{n} x / r_{p}=\delta>0$. Pick a large fixed $L$ and let $z=$ $\left\lceil h_{p+1} / k\right\rceil$. Then by the Block Lemma,

$$
\int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{j k}\right| d \mu \leq \int\left|\frac{1}{L} \sum_{l=0}^{L-1} \chi_{B} \circ T^{l z k}\right| d \mu+\frac{L z}{n}
$$

and $h_{p+1} / k \leq r_{p} h_{p} / x h_{p}=r_{p} / x \leq 1 / \delta$ so since $L$ is fixed the rightmost term tends to zero as $n \rightarrow \infty$. By Lemma 4.11 with $t=z k$ (so that $h_{p+1} \leq t \leq$ $\left.h_{p+1}+k\right)$ the above can be made arbitrarily small by taking $L$ sufficiently large.

Case 1b: $\lim _{n} x r_{p} / h_{p}=0$. Since $\lim \sup x / r_{p} \leq 1$ by the construction of the unique $x$ and $p$, by Proposition 4.10 (using $r_{p}-x$ in place of $n$ in the proposition) we have $\int\left|\left(1 / r_{p}\right) \sum_{j=0}^{r_{p}-x-1} \chi_{B} \circ T^{-j x}\right| d \mu \rightarrow 0$ (and also for $x+1$ in place of $x$ ) because we may assume $r_{p}-x \rightarrow \infty$ as when $\left(r_{p}-x\right) / r_{p} \rightarrow 0$ the claim is trivial (and of course $r_{p} \rightarrow \infty$ since $k \rightarrow \infty$ hence $p \rightarrow \infty$ ). Then by Proposition 4.6 (using the hypothesis for this case) we find that $\left\{k_{n}\right\}$ is mixing. For any fixed $l$ the same argument shows that $\left\{l k_{n}\right\}$ is a mixing sequence. The result follows from Lemma 4.11.

Case 2: $\lim \sup _{n} x r_{p} / h_{p}<\infty($ and not Cases 1a or 1 b$)$. Let $\Delta$ be the limit supremum and $\delta$ the limit infimum of $x r_{p} / h_{p}$ (so $0<\delta \leq \Delta<\infty$, the $\delta=0$ case being covered by 1 b and the $\Delta=\infty$ to be Case 3). Pick $\epsilon=\epsilon_{n} \rightarrow 0$ such that $\epsilon h_{p} / r_{p} \rightarrow \infty$ (possible since $T$ is finite-measurepreserving) and such that $\epsilon r_{p} / x \geq \alpha>0$ for some $\alpha>0$ (requires that $x / r_{p} \rightarrow 0$, which we may assume to be the case by dropping to a subsequence of the $n$ and noting that the subsequence of $n$ where this does not tend to zero is covered by Case 1a above; specifically, fix $\epsilon>0$ and split the sequence into the part where $r_{p} / h_{p}<\epsilon$ and where it is greater, handle each separately and then take $\epsilon \rightarrow 0$ ). Set $l=\epsilon h_{p} / x$ (modify $\epsilon$ slightly so that $l$ is an integer). Then $l=l_{n} \rightarrow \infty$ and $l / r_{p}=\epsilon h_{p} / x r_{p} \leq \epsilon / \delta \rightarrow 0$ and $l x / h_{p}=\epsilon \rightarrow 0$. Note that $x / l \leq x \Delta / \epsilon r_{p} \leq \infty$. Then by Proposition 4.10, $\int\left|(1 / l) \sum_{j=0}^{l-1} \chi_{B} \circ T^{j(x+2-\alpha(q))}\right| d \mu \rightarrow 0$ for any $\alpha(q)$ satisfying the requirements of Proposition 4.7. Proposition 4.7 in turn shows that $\left\{k_{n}\right\}$ is mixing. Clearly the same argument works for a constant multiple $\left\{d k_{n}\right\}$ so by Lemma 4.12 the result follows.

CASE 3: $\lim \sup _{n} x r_{p} / h_{p}=\infty($ and not Cases 1a, 1b or 2$)$. Since $x / h_{p} \leq$ $r_{p} / h_{p} \rightarrow 0$ there exists $l=l_{n} \rightarrow \infty$ such that $x l / h_{p} \rightarrow 0$. Then

$$
\frac{l}{r_{p}}=\frac{h_{p}}{x r_{p}} \frac{l x}{h_{p}} \rightarrow 0
$$

Set $Q=Q_{n}=\left\lceil\left(r_{p}-x\right) / l\right\rceil \rightarrow \infty$ (recall we have eliminated Case 1 b and may drop to a subsequence since the subsequence along which this is bounded will be covered by that case). Let $\alpha(q)$ be the sequence of values required in Proposition 4.7. By the Block Lemma,

$$
\begin{aligned}
& \frac{1}{Q} \sum_{q=0}^{Q-1} \int\left|\frac{1}{l} \sum_{j=0}^{l-1} \chi_{B} \circ T^{j(x+2-\alpha(q))}\right| d \mu \\
& \leq \frac{1}{Q} \sum_{q=0}^{Q-1} \int\left|\frac{1}{L} \sum_{j=0}^{L-1} \chi_{B} \circ T^{j(x+2-\alpha(q))}\right| d \mu+\frac{L}{l}
\end{aligned}
$$

For each fixed $d \neq 0$ the fact that $T$ is weakly mixing implies that $T^{d}$ is weakly mixing. There is then a density one sequence $J_{d}$ such that $T^{d}$ is mixing along $J_{d}$. Now $\alpha(q)$ takes on all integer values from 0 to $R_{n}$ where $R_{n} \rightarrow \infty$ since $Q=Q_{n} \rightarrow \infty$, and takes on each particular value with approximately the same density (see Remark 4.8). So for $m \leq R$,

$$
\begin{aligned}
& \#\left\{(q, r) \in \mathbb{Z}_{Q} \times \mathbb{Z}_{Q}: \alpha(q)-\alpha(r)=m\right\} \\
& \quad=\sum_{t=0}^{m} \#\left\{q \in \mathbb{Z}_{Q}: \alpha(q)=t\right\} \#\left\{r \in \mathbb{Z}_{Q}: \alpha(r)=t-m\right\} \approx m\left(\frac{Q}{R}\right)^{2} \leq \frac{Q^{2}}{R}
\end{aligned}
$$

Therefore

$$
\frac{1}{Q^{2}} \#\left\{(q, r) \in \mathbb{Z}_{Q} \times \mathbb{Z}_{Q}: \alpha(q)-\alpha(r) \notin J_{d}\right\} \lesssim \frac{\#\left(\mathbb{Z}_{R} \backslash J_{d}\right)}{R} \rightarrow 0
$$

Then for each fixed $d \neq 0$ there is some $Z_{d}$ such that

$$
\frac{1}{Q} \sum_{q, r=0}^{Q-1}\left|\mu\left(T^{d(\alpha(q)-\alpha(r))} B \cap B\right)-\mu(B) \mu(B)\right|<\epsilon
$$

for any $Q \geq Z_{d}$.
By the Hölder inequality and the measure-preserving property (repeatedly),

$$
\begin{aligned}
& \left(\frac{1}{Q} \sum_{q=0}^{Q-1} \int\left|\frac{1}{L} \sum_{j=0}^{L-1} \chi_{B} \circ T^{j(x+2-\alpha(q))}\right| d \mu\right)^{4} \\
& \quad \leq\left(\frac{1}{Q} \sum_{q=0}^{Q-1} \frac{1}{L^{2}} \sum_{j, t=0}^{L-1} \mu\left(T^{(j-t)(x+2-\alpha(q))} B \cap B\right)-\mu(B) \mu(B)\right)^{2}
\end{aligned}
$$

$$
\begin{aligned}
& \leq\left(\frac{1}{L^{2}} \sum_{j, t=0}^{L-1} \int\left|\frac{1}{Q} \sum_{q=0}^{Q-1} \chi_{B} \circ T^{-(j-t) \alpha(q)}\right| \circ T^{(j-t)(x+2)} d \mu\right)^{2} \\
& \leq \frac{1}{L^{2}} \sum_{j, t=0}^{L-1} \frac{1}{Q^{2}} \sum_{q, r=0}^{Q-1} \mu\left(T^{(j-t)(\alpha(q)-\alpha(r))} B \cap B\right)-\mu(B) \mu(B)
\end{aligned}
$$

Hold $L$ fixed and pick $n$ such that $Q \geq Z_{d}$ for all $-L \leq d \leq L$ (and $d \neq 0)$. Then the above is less than $\epsilon+1 / L$, which tends to zero. Hence by Proposition 4.7 we see that $\left\{k_{n}\right\}$ is mixing. In fact, so also is any constant multiple of $\left\{k_{n}\right\}$, and the result follows by Lemma 4.12 .

We are now ready to prove that staircases are mixing. By Proposition 4.7 this reduces to showing that ergodic type averages tend to zero and the above tells us this is the case.

Proof of Theorem 4.1. Let $\left\{t_{n}\right\}$ be a sequence. Let $\left\{r_{n}\right\}$ be the cut sequence and $\left\{h_{n}\right\}$ the height sequence for $T$ and set $p(n)$ and $k(n)$ as usual. Pick a sequence $\left\{l_{n}\right\}$ such that $l_{n} \rightarrow \infty, l_{n} k(n) / h_{p(n)} \rightarrow 0$ and $l_{n} / r_{p(n)} \rightarrow 0$ (possible since $k(n) / h_{p(n)} \leq r_{p(n)} / h_{p(n)} \rightarrow 0$ ). Since $T$ is power ergodic (Proposition 4.13), for any $B \in \mathcal{B}$,

$$
\int\left|\frac{1}{l_{n}} \sum_{j=0}^{l_{n}-1} \chi_{B} \circ T^{-j(k(n)+2-\alpha(n, q))}\right| d \mu \rightarrow 0
$$

for appropriate $\alpha(n, q)$. Then by Proposition 4.7, $\left\{t_{n}\right\}$ is mixing with respect to $T$.
5. General rank-one transformations. We now establish the results we obtained on staircases for general rank-one transformations. The methods follow the same strategy as above but with many additional technical complications due to the fact that the spacer sequence is now arbitrary and no longer arithmetical so tricks such as the Block Lemma are no longer straightforward. The propositions below are direct analogues of those above. The reader is encouraged to become familiar with the concepts in the case of staircase transformations before reading the details of the general case.
5.1. Dynamical sequences. Dynamical sequences are the natural representation of spacer sequences for rank-one transformations (see Section 3).

Definition 5.1. A dynamical sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ is a doubly indexed collection of integers $s_{n, j}$ for $n \in \mathbb{N}$ and $j \in \mathbb{Z}_{r_{n}}$ where $\left\{r_{n}\right\}$ is a given sequence, called the index sequence, which must have the property that $\lim _{n \rightarrow \infty} r_{n}=\infty$. The integer $s_{n, j}$ is the $j$ th element of the dynamical sequence at the $n$th stage.

### 5.2. Partial sums of dynamical sequences

Notation. Let $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ be a dynamical sequence and $n \in \mathbb{N}, j \in \mathbb{Z}_{r_{n}}$, $k \in \mathbb{Z}_{r_{n}-j}$. The $k$ th partial sum of the $j$ th element at the $n$th stage is

$$
s_{n, j}^{(k)}=\sum_{z=0}^{k-1} s_{n, j+z} .
$$

Definition 5.2. Let $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ be a dynamical sequence and $k \in \mathbb{N}$. The $k$ th partial sum dynamical sequence is the dynamical sequence $\left\{s_{n, j}^{(k)}\right\}_{\left\{r_{n}-k\right\}}$ ( $n$ "begins" at the smallest value such that $r_{n} \geq k$ ) whose elements are the $k$ th partial sums of $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$. Let $\left\{k_{n}\right\}$ be a sequence such that $k_{n}<r_{n}$ for all $n$. The $\left\{k_{n}\right\}$ th partial sum dynamical sequence of $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ is the dynamical sequence $\left\{s_{n, j}^{\left(k_{n}\right)}\right\}_{\left\{r_{n}-k_{n}\right\}}$.

### 5.3. Spread out dynamical sequences

Definition 5.3. A dynamical sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ is spread out, which will be written as $s_{n, j} \rightarrow \infty$, when for every fixed $M \in \mathbb{N}$,

$$
\lim _{n \rightarrow \infty} \frac{1}{r_{n}^{2}} \#\left\{j, k \in \mathbb{Z}_{r_{n}}:\left|s_{n, j}-s_{n, k}\right|<M\right\}=0 .
$$

This property was referred to as a dynamical sequence being monotonic in CS04.

### 5.4. Ergodic dynamical sequences

Definition 5.4. A dynamical sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ is an ergodic dynamical sequence (with respect to a transformation $T$ ) when for all $B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \int\left|\frac{1}{r_{n}} \sum_{j=0}^{r_{n}-1} \chi_{B} \circ T^{-s_{n, j}}\right| d \mu=0
$$

5.5. Mixing and ergodic dynamical sequences. The following is a standard generalization of the Blum-Hanson theorem from sequences to dynamical sequences.

Theorem 5.5. A transformation $T$ is mixing if and only if every spread out dynamical sequence is ergodic with respect to $T$.

Proof. Let $T$ be a mixing transformation and $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ a spread out dynamical sequence. For any $B \in \mathcal{B}$ and any $\epsilon>0$ there exists $M \in \mathbb{N}$ such that for all $m \geq M$ or $m \leq-M$ we have $\left|\mu\left(T^{m}(B) \cap B\right)-\mu(B) \mu(B)\right|<\epsilon$. As $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ is spread out, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $\#\left\{(j, l) \in \mathbb{Z}_{r_{n}} \times \mathbb{Z}_{r_{n}}:\left|s_{n, j}-s_{n, l}\right|<M\right\}<\epsilon r_{n}^{2}$. Then

$$
\begin{aligned}
\int \left\lvert\, \frac{1}{r_{n}}\right. & \left.\sum_{j=0}^{r_{n}-1} \chi_{B} \circ T^{-s_{n, j}}\right|^{2} d \mu \\
& =\frac{1}{r_{n}^{2}} \sum_{j, l=0}^{r_{n}-1} \mu\left(T^{s_{n, j}-s_{n, l}} B \cap B\right)-\mu(B) \mu(B) \\
& \leq \frac{1}{r_{n}^{2}} \sum_{\left|s_{n, j}-s_{n, l}\right| \geq M}\left|\mu\left(T^{s_{n, j}-s_{n, l}} B \cap B\right)-\mu(B) \mu(B)\right|+\frac{1}{r_{n}^{2}} \epsilon r_{n}^{2} \mu(B) \\
& <\frac{1}{r_{n}^{2}} \sum_{\left|s_{n, j}-s_{n, l}\right| \geq M} \epsilon+\epsilon \mu(B)<\epsilon(1+\mu(B)) .
\end{aligned}
$$

Conversely, suppose $T$ is not mixing. As mixing is equivalent to Rényi mixing there exist $B \in \mathcal{B}, \delta>0$ and a sequence $\left\{t_{m}\right\}$ such that $\mu\left(T^{t_{m}} B \cap B\right)$ $-\mu(B) \mu(B) \geq \delta$ for all $m$. Define $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ by $r_{n}=n$ and $s_{n, j}=t_{j}$. Then

$$
\begin{aligned}
\int\left|\frac{1}{r_{n}} \sum_{j=0}^{r_{n}-1} \chi_{B} \circ T^{-s_{n, j}}\right| d \mu & \geq\left|\frac{1}{n} \sum_{j=0}^{n-1} \int_{B} \chi_{B} \circ T^{-t_{j}} d \mu\right| \\
& =\left|\frac{1}{n} \sum_{j=0}^{n-1} \mu\left(T^{t_{j}} B \cap B\right)-\mu(B) \mu(B)\right| \geq \delta
\end{aligned}
$$

Since $\left\{t_{m}\right\}$ is strictly increasing, $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ is spread out.

### 5.6. Levels of rank-one transformations

Lemma 5.6. Let $T$ be a rank-one transformation with levels $\left\{I_{n, i}\right\}$, heights $\left\{h_{n}\right\}$, and spacers $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$. Let $p \in \mathbb{N}, i \in \mathbb{Z}_{h_{p}}, j \in \mathbb{Z}_{r_{p}}, k \in \mathbb{Z}_{r_{p}-j}$, $t \in \mathbb{Z}_{h_{p}-i}$ and $B$ a union of levels in $C_{p}$. Then the following hold:
(i) $I_{p, i}=\bigcup_{j=0}^{r_{p}-1} I_{p, i}^{[j]}$;
(ii) $T^{k h_{p}+s_{p, j}^{(k)}}\left(I_{p, i}^{[j]}\right)=I_{p, i}^{[j+k]}$;
(iii) $\mu\left(T^{t}\left(I_{p, i}^{[j]}\right) \cap B\right)=\frac{1}{r_{p}} \mu\left(T^{t}\left(I_{p, i}\right) \cap B\right)$.

Lemma 5.7. For any $p \in \mathbb{N}$, any $\Lambda \subseteq \mathbb{Z}_{h_{p}}$, any $\Gamma \subseteq \mathbb{N}$, any union $B$ of levels in $C_{p}$ and any maps $f: \Gamma \rightarrow \mathbb{Z}$ and $g: \Gamma \rightarrow \mathbb{Z}_{r_{p}}$,

$$
\begin{aligned}
\sum_{i \in \Lambda} \mid \sum_{j \in \Gamma} \mu\left(T^{f(j)}\left(I_{p, i}^{[g(j)]}\right)\right. & \cap B)-\mu\left(I_{p, i}^{[g(j)]}\right) \mu(B) \mid \\
& \leq \int\left|\frac{1}{r_{p}} \sum_{j \in \Gamma} \chi_{B} \circ T^{f(j)}\right| d \mu+\left(\sup _{j \in \Gamma} f(j)\right) \frac{1}{h_{p}} \frac{\# \Gamma}{r_{p}}
\end{aligned}
$$

### 5.7. Mixing sequences

Proposition 5.8. Let $T$ be a rank-one transformation with spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ and heights $\left\{h_{n}\right\}$ and let $k \in \mathbb{N}$. If $\left\{s_{n, j}^{(k)}\right\}_{\left\{r_{n}-k\right\}}$ is ergodic (with respect to $T$ ) then $\left\{k h_{n}\right\}$ is mixing (with respect to $T$ ).

Proof. Let $B$ be a union of levels in $C_{N}$ for some fixed $N \in \mathbb{N}$. For any sets $J_{n} \subseteq \mathbb{Z}_{r_{n}-k}$, apply Lemmas 5.6(i) then 5.6(ii) and finally 5.7, for any $n \geq N$, to get

$$
\begin{aligned}
& \sum_{i=0}^{h_{n}-1}\left|\mu\left(T^{k h_{n}}\left(I_{n, i}\right) \cap B\right)-\mu\left(I_{n, i}\right) \mu(B)\right| \\
& \quad \leq \sum_{i=0}^{h_{n}-1}\left|\sum_{j \in \mathbb{Z}_{r_{n}-k} \backslash J_{n}} \mu\left(T^{k h_{n}}\left(I_{n, i}^{[j]}\right) \cap B\right)-\mu\left(I_{n, i}^{[j]}\right) \mu(B)\right|+\frac{k}{r_{n}}+\frac{\# J_{n}}{r_{n}} \\
& \quad=\sum_{i=0}^{h_{n}-1} \left\lvert\, \sum_{j \in \mathbb{Z}_{r_{n}-k} \backslash J_{n}} \mu\left(T^{\left.-s_{n, j}^{(k)}\left(I_{n, i}^{[j+k]}\right) \cap B\right)-\mu\left(I_{n, i}^{[j+k]}\right) \mu(B) \left\lvert\,+\frac{k}{r_{n}}+\frac{\# J_{n}}{r_{n}}\right.}\right.\right. \\
& \quad \leq \int\left|\frac{1}{r_{n}} \sum_{j=0}^{r_{n}-k} \chi_{B} \circ T^{-s_{n, j}^{(k)}}\right| d \mu+\frac{2 k}{r_{n}}+2 \frac{\# J_{n}}{r_{n}}+\frac{1}{h_{n}} \sup _{j \in \mathbb{Z}_{r_{n}-k \backslash J_{n}}} s_{n, j}^{(k) .}
\end{aligned}
$$

As $\left\{s_{n, j}^{(k)}\right\}_{\left\{r_{n}-k\right\}}$ is ergodic with respect to $T$, we need only show that there exist sets $J_{n} \subseteq \mathbb{Z}_{r_{n}-k}$ such that $\# J_{n} / r_{n} \rightarrow 0$ and $h_{n}^{-1} \sup _{j \notin J_{n}} s_{n, j}^{(k)} \rightarrow 0$. Suppose not. Then there exists $\delta>0$ such that $s_{n, j}^{(k)} \geq \delta h_{n}$ for at least $\delta r_{n}$ values of $j$ (for infinitely many $n$ ). But then at least $k^{-1} \delta r_{n}$ values of $j$ are such that $s_{n, j} \geq(\delta / k) h_{n}$ so $\mu\left(S_{n}\right) \geq\left(\delta^{2} / k^{2}\right) r_{n} h_{n} \mu\left(I_{n+1,0}\right)=\left(\delta^{2} / k^{2}\right) \mu\left(C_{n}\right)$, contradicting that $T$ is defined on a finite-measure space.

Proposition 5.9. Let $T$ be a rank-one transformation with height sequence $\left\{h_{n}\right\}$ and spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ and let $\left\{t_{n}\right\}$ be a sequence. Set $p(n)$ and $k(n)$ such that $h_{p(n)} \leq t_{n}<h_{p(n)+1}$ and $k(n) h_{p(n)} \leq t_{n}<$ $(k(n)+1) h_{p(n)}$. If there exists $J_{n} \subseteq \mathbb{Z}_{r_{p(n)}-k(n)}$ such that $\# J_{n} / r_{p(n)} \rightarrow 0$ and $h_{p(n)}^{-1} \sup _{j \notin J_{n}} s_{n, j}^{(k(n))} \rightarrow 0$ and for all $B \in \mathcal{B}$,

$$
\int\left|\frac{1}{r_{p(n)}} \sum_{j=0}^{r_{p(n)}-k(n)-1} \chi_{B} \circ T^{-s_{p(n), j}^{(k(n))}}\right| d \mu \rightarrow 0
$$

then $\left\{t_{n}\right\}$ is mixing with respect to $T$.
Proof. Write $t_{n}=k h_{p}+m$ (with $k=k(n), p=p(n)$ and $m=m(n)$ as before) and let $A, B$ be unions of levels in $C_{N}$ for some $N \in \mathbb{N}$. For sufficiently large $n$, as in Proposition 4.3,

$$
\begin{aligned}
\left|\mu\left(T^{t_{n}} A \cap B\right)-\mu(A) \mu(B)\right| \leq & \sum_{i=0}^{h_{p}-1}\left|\mu\left(T^{k h_{p}} I_{p, i} \cap B\right)-\mu\left(I_{p, i}\right) \mu(B)\right| \\
& +\sum_{i=0}^{h_{p}-1}\left|\mu\left(T^{(k+1) h_{p}} I_{p, i} \cap B\right)-\mu\left(I_{p, i}\right) \mu(B)\right| .
\end{aligned}
$$

By Lemma 5.6(i),

$$
\begin{aligned}
\sum_{i=0}^{h_{p}-1} \mid \mu\left(T^{k h_{p}} I_{p, i} \cap B\right)- & \mu\left(I_{p, i}\right) \mu(B) \mid \\
\leq & \sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{r_{p}-k-1} \mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right| \\
& +\sum_{i=0}^{h_{p}-1} \sum_{j=r_{p}-k}^{r_{p}-1}\left|\mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right|
\end{aligned}
$$

and the second summand tends to zero as in Proposition 4.6.
By Lemmas 5.6(ii) and 5.7,

$$
\begin{aligned}
\sum_{i=0}^{h_{p}-1} \mid \sum_{j=0}^{r_{p}-k-1} & \mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B) \mid \\
& \leq \int\left|\frac{1}{r_{p}} \sum_{j \notin J_{n}} \chi_{B} \circ T^{-s_{p, j}^{(k)}}\right| d \mu+\left(\sup _{j \notin J_{n}} s_{p, j}^{(k)}\right) \frac{1}{h_{p}}+\frac{\# J_{n}}{r_{p}} \\
& \leq \int\left|\frac{1}{r_{p}} \sum_{j=0}^{r_{p}-k-1} \chi_{B} \circ T^{-s_{p, j}^{(k)}}\right| d \mu+\left(\sup _{j \notin J_{n}} s_{p, j}^{(k)}\right) \frac{1}{h_{p}}+2 \frac{\# J_{n}}{r_{p}} .
\end{aligned}
$$

The same reasoning holds for $k+1$.
DEfinition 5.10. Fix $T$ a rank-one transformation with spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$. Let $Q, p, k$ be positive integers and $\left\{\Gamma_{q}\right\}_{q=0}^{Q-1}$ be a partition of $\mathbb{Z}_{r_{p}-k}$ (that is, $\Gamma_{q} \subseteq \mathbb{Z}_{r_{p}-k}$ are disjoint and $\bigcup \Gamma_{q}=\mathbb{Z}_{r_{p}-k}$ ). We say $\left(Q, p, k,\left\{\Gamma_{q}\right\}\right)$ respects the spacing arrangement at $k$ of $T$ when $s_{p, j}^{(k)} \leq s_{p, j^{\prime}}^{(k)}$ for all $j \in \Gamma_{q}$ and $j^{\prime} \in \Gamma_{q^{\prime}}$ whenever $q \leq q^{\prime}$.

Given sequences $\left\{p_{n}\right\},\left\{k_{n}\right\}$ and $\left\{Q_{n}\right\}$, a sequence of partitions $\left\{\Gamma_{n, q}\right\}_{Q_{n}}$ that each individually respect the spacing arrangement of $T$ at $k_{n}$ and also have the property that $Q_{n} / r_{p_{n}} \rightarrow 0$ is said to respect the spacing arrangement for $T$ with block sizes $\left\{b_{n}\right\}$ when $\# \Gamma_{n, q} \geq b_{n}$ for a density one set of $q$.

Proposition 5.11. Let $T$ be a rank-one transformation with spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ and heights $\left\{h_{n}\right\}$ and let $\left\{t_{n}\right\}$ be a sequence. Set $p(n)$ and $k(n)$ such that $h_{p(n)} \leq t_{n}<h_{p(n)+1}$ and $k(n) h_{p(n)} \leq t_{n}<(k(n)+1) h_{p(n)}$. If there exists a sequence $\left\{L_{n}\right\}$ such that

$$
\frac{L_{n}}{h_{p_{n}}} \frac{1}{r_{p_{n}}} \sum_{j=0}^{r_{p_{n}}-1} s_{p_{n}, j} \rightarrow 0 \quad \text { and } \quad L_{n} \rightarrow \infty
$$

that also has the property that every sequence of partitions that respects the spacing arrangement of $T$ with block sizes $\left\{h_{p_{n}} / L_{n}\right\}$ is ergodic with respect to $T$, meaning

$$
\lim _{n \rightarrow \infty} \sum_{q=0}^{Q_{n}-1} \int\left|\frac{1}{r_{p(n)}} \sum_{j \in \Gamma_{n, q}} \chi_{B} \circ T^{-s_{p(n), j}^{(k(n)+2-\alpha(n, q))}}\right| d \mu=0
$$

then $\left\{t_{n}\right\}$ is mixing with respect to $T$.
Proof. Let $B$ be a union of levels in $C_{N}$ for some $N \in \mathbb{N}$. Let $n \in \mathbb{N}$ be such that $t_{n} \geq h_{N}$. Set $m=t_{n}-k h_{p}$ (writing $k=k(n)$ and $p=p(n)$ and so forth) so $m \in \mathbb{Z}_{h_{p}}$. Set $\epsilon=\epsilon_{n}=1 / L_{n}$. Let $\Psi=\Psi_{n}: \mathbb{Z}_{r_{p}-k} \rightarrow \mathbb{Z}_{r_{p}-k}$ be a map such that $s_{p, \Psi(j)}^{(k)} \leq s_{p, \Psi(j+1)}^{(k)}$ for all $j \in \mathbb{Z}_{r_{p}-k-1}$. Set $l(0)=0$ and $\alpha(-1)=0$ and then proceed inductively to define $l(q+1)=l_{n}(q+1)$ and $\alpha(q)=\alpha_{n}(q)$ given $l(q)$ and $\alpha(q-1)$ as follows: choose $l(q+1)$ to be the smallest positive integer less than $r_{p}-k$ such that

$$
s_{p, \Psi(l(q+1))}^{(k-\alpha(q-1)+1)}-s_{p, \Psi(l(q))}^{(k-\alpha(q-1))} \geq \epsilon h_{p}
$$

if such an integer exists and choose $l(q+1)=r_{p}$ and set $Q_{n}=q+1$ if not. Note that

$$
Q_{n} \leq \frac{s_{p, \Psi\left(r_{p}-k-1\right)}^{(k)}}{\epsilon h_{p}} \leq \frac{s_{p, 0}^{\left(r_{p}\right)}}{\epsilon h_{p}}=\frac{r_{p} \mu\left(S_{p}\right)}{\epsilon \mu\left(C_{p}\right)}
$$

so $Q_{n} / r_{p} \rightarrow 0$. Choose $\alpha(q)$ such that
(i) $(\alpha(q)-1) h_{p}+s_{p, \Psi(l(q))+k-\alpha(q)+1}^{(\alpha(q)-1)}<s_{p, \Psi(l(q))}^{(k)}$;
(ii) $s_{p, \Psi(l(q))}^{(k)} \leq \alpha(q) h_{p}+s_{p, \Psi(l(q))+k-\alpha(q)}^{(\alpha(q))}$.

The $\alpha(q)$ are clearly nondecreasing over $q$. Since

$$
\frac{L}{h_{p}} \frac{1}{r_{p}} \sum_{j=0}^{r_{p}-1} s_{p, j} \rightarrow 0
$$

for large $n$ we have $\alpha(q+1)-\alpha(q) \leq 1$.
Set $\beta(q)=\alpha(q) h_{p}+s_{p, \Psi(l(q))+k-\alpha(q)}^{(\alpha(q))}-s_{p, \Psi(l(q))}^{(k)}$ and note that $0 \leq \beta(q)$ $<h_{p}+s_{p, \Psi(l(q))+k-\alpha(q)}$. Set $\beta^{\prime}(q)=\beta(q)-s_{p, \Psi(l(q))+k-\alpha(q)}$ and note that
$\beta^{\prime}(q)<h_{p}$. For all $q \in \mathbb{Z}_{q}$, define the sets

$$
\Gamma_{q}=\Gamma_{n, q}=\left\{j \in \mathbb{Z}_{r_{p}}: s_{p, \Psi(l(q))}^{(k)} \leq s_{p, j}^{(k)}<s_{p, \Psi(l(q+1))}^{(k)}\right\}
$$

and the maps $\Psi(q)=\Psi(n, q): \mathbb{Z}_{l(q+1)-l(q)} \rightarrow \Gamma_{q}$ by $\Psi(q)(j)=\Psi(j+l(q))$ for all $j \in \mathbb{Z}_{l(q+1)-l(q)}$. As in Proposition 4.7, it suffices to show that

$$
\sum_{i=0}^{h_{p}-1}\left|\sum_{j=0}^{r_{p}-1} \mu\left(T^{k h_{p}} I_{p, i}^{[j]} \cap B\right)-\mu\left(I_{p, i}^{[j]}\right) \mu(B)\right| \rightarrow 0
$$

and likewise for $k+1$. Note that

$$
\begin{aligned}
\sum_{i=0}^{h_{p}-1} \mid \sum_{j=0}^{r_{p}-k-1} & \mu\left(T^{-s_{p, j}^{(k)}}\left(I_{p, i}^{[j+k]}\right) \cap B\right)-\mu\left(I_{p, i}^{[j+k]}\right) \mu(B) \mid \\
& =\sum_{i=0}^{h_{p}-1}\left|\sum_{q=0}^{Q-1} \sum_{j \in \Gamma_{q}} \mu\left(T^{-s_{p, j}^{(k)}}\left(I_{p, i}^{[j+k]}\right) \cap B\right)-\mu\left(I_{p, i}^{[j+k]}\right) \mu(B)\right|
\end{aligned}
$$

Now for any $q \in \mathbb{Z}_{q}$, any $i \in \mathbb{Z}_{h_{p}}$ and any $j \in \Gamma_{q}$, using Lemma 5.6(ii),

$$
\begin{aligned}
T^{-s_{p, j}^{(k)}}\left(I_{p, i}^{[j+k]}\right) & =T^{-\left(s_{p, j}^{(k)}-s_{p, \Psi(l(q))}^{(k)}\right)-s_{p, \Psi(l(q))}^{(k)}}\left(I_{p, i}^{[j+k]}\right) \\
& =T^{-\left(s_{p, j}^{(k)}-s_{p, \Psi(l(q))}^{(k)}\right)-s_{p, \Psi(l(q))}^{(k)}+\alpha(q) h_{p}+s_{p, j+k-\alpha(q)}^{(\alpha(q))}\left(I_{p, i}^{[j+k-\alpha(q)]}\right)} \\
& =T^{-\left(s_{p, j}^{(k)}-s_{p, \Psi(l(q))}^{(k)}\right)+\beta(q)}\left(I_{p, i}^{[j+k-\alpha(q)]}\right) .
\end{aligned}
$$

If $i<h_{p}-\beta(q)$ then

$$
T^{-s_{p, j}^{(k)}}\left(I_{p, i}^{[j+k]}\right)=T^{-\left(s_{p, j}^{(k-\alpha(q))}-s_{p, \Psi(l(q))}^{(k-\alpha(q))}\right)}\left(I_{p, i+\beta(q)}^{[j+k-\alpha(q)]}\right)
$$

If $i \geq h_{p}-\beta^{\prime}(q)$ then

$$
\begin{aligned}
& T^{-s_{p, j}^{(k)}}\left(I_{p, i}^{[j+k]}\right) \\
& \quad=T^{-\left(s_{p, j}^{(k-\alpha(q))}-s_{p, \Psi(l(q))}^{(k-\alpha(q))}\right)+\beta(q)-h_{p}-s_{p, j+k-\alpha(q)}\left(I_{p, i}^{[j+k-\alpha(q)+1]}\right)} \\
& \quad=T^{-\left(s_{p, j}^{(k-\alpha(q)+1)}-s_{p, \Psi(l(q))}^{(k-\alpha(q)+1)}\right)+\beta(q)-h_{p}-s_{p, \Psi(l(q))+k-\alpha(q)}\left(I_{p, i}^{[j+k-\alpha(q)+1]}\right)} \\
& \quad=T^{-\left(s_{p, j}^{(k-\alpha(q)+1)}-s_{p, \Psi(l(q))}^{(k-\alpha(q)+1)}\right)}\left(I_{p, i+\beta^{\prime}(q)-h_{p}}^{[j+k-\alpha(q)+1]}\right)
\end{aligned}
$$

as $i+\beta^{\prime}(q)-h_{p} \geq 0$ because $i \geq h_{p}-\beta^{\prime}(q)$ and $i+\beta^{\prime}(q)-h_{p}<i<h_{p}$ because $\beta^{\prime}(q)<h_{p}$.

If $h_{p}-\beta(q) \leq i<h_{p}-\beta^{\prime}(q)$ then, as above,

$$
\begin{aligned}
T^{-s_{p, j}^{(k)}}\left(I_{p, i}^{[j+k]}\right) & =T^{-\left(s_{p, j}^{(k-\alpha(q))}-s_{p, \Psi(l(q))}^{(k-\alpha(q))}\right)+\beta(q)-h_{p}-s_{p, j+k-\alpha(q)}}\left(I_{p, i}^{[j+k-\alpha(q)+1]}\right) \\
& =T^{-\left(s_{p, j}^{(k-\alpha(q)+1)}-s_{p, \Psi(l(q))}^{(k-\alpha(q))}\right)}\left(I_{p, i+\beta(q)-h_{p}}^{[j+k-\alpha(q)+1]}\right) .
\end{aligned}
$$

Applying the first case and Lemma 5.7, then the fact that $T$ is measurepreserving, we get

$$
\begin{aligned}
& \sum_{q=0}^{Q_{n}-1} \sum_{i=0}^{h_{p}-\beta(q)}\left|\sum_{j \in \Gamma_{q}} \mu\left(T^{-s_{p, j}^{(k)}}\left(I_{p, i}^{[j+k]}\right) \cap B\right)-\mu\left(I_{p, i}^{[j+k]}\right) \mu(B)\right| \\
& \leq \sum_{q=0}^{Q_{n}-1} \int\left|\frac{1}{r_{p}} \sum_{j \in \Gamma_{q}} \chi_{B} \circ T^{-\left(s_{p, j}^{(k-\alpha(q))}-s_{p, \Psi(l(q))}^{(k-\alpha(q))}\right)}\right| d \mu \\
&+\sum_{q=0}^{Q_{n}-1}\left(\sup _{j \in \Gamma_{q}} s_{p, j}^{(k-\alpha(q))}-s_{p, \Psi(l(q))}^{(k-\alpha(q))}\right) \frac{1}{h_{p}} \frac{\# \Gamma_{q}}{r_{p}} \\
& \leq \sum_{q=0}^{Q_{n}-1} \int\left|\frac{1}{r_{p}} \sum_{j \in \Gamma_{n, q}} \chi_{B} \circ T^{-s_{p, j}^{(k-\alpha(q))}}\right| d \mu+\epsilon
\end{aligned}
$$

Similarly for the second and third cases above (with $k+1$ replacing $k$ ). Combining these three cases and letting $n \rightarrow \infty$, we have the result.

## 6. Mixing theorem

Theorem 6.1. For a rank-one transformation $T$, the following are equivalent:
(i) $T$ is a mixing transformation;
(ii) $T$ is "rank-one uniform mixing": for all $B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \sup _{h_{n} \leq t<h_{n+1}} \sum_{i=0}^{h_{n}-1}\left|\mu\left(T^{t} I_{n, i} \cap B\right)-\mu\left(I_{n, i}\right) \mu(B)\right|=0
$$

(iii) every sequence of partitions that respects the spacing arrangement of $T$ with block sizes tending to infinity is ergodic with respect to $T$.

Proof. (ii) $\Rightarrow$ (i) follows as in the proofs of our propositions. (iii) $\Rightarrow$ (ii) follows as in Proposition 5.11. Finally, (i) $\Rightarrow$ (iii) follows since if not then for some sequence of partitions that respect the spacing arrangement of $T$ there is some $\{q(n)\}$ with $\# \Gamma_{n, q(n)} \rightarrow \infty\left(\right.$ as $\left.Q_{n} / r_{n} \rightarrow 0\right)$ and the dynamical sequence $\left\{s_{n, j}^{(k(n))}\right\}_{\Gamma_{n, q(n)}}$ is not ergodic with respect to $T$ so, by Theorem 5.5, $T$ cannot be mixing.

## 7. Polynomial staircase transformations

7.1. Polynomial power ergodicity. The term polynomial shall mean polynomials with rational coefficients that map integers to integers.

Proposition 7.1. Let $T$ be a power ergodic transformation. Then for any sequence of polynomials $\left\{p_{n}\right\}$ of bounded degree and all $B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{-p_{n}(j)}\right| d \mu=0
$$

Lemma 7.2 (van der Corput's inequality). For any complex numbers $a_{n}$ such that $\left|a_{n}\right| \leq 1$ and any $N, L \in \mathbb{N}$,

$$
\left|\frac{1}{N} \sum_{n=0}^{N-1} a_{n}\right|^{2} \leq \frac{N+L}{N}\left(\frac{1}{L}+2 \operatorname{Re}\left[\frac{1}{L} \sum_{l=1}^{L-1} \frac{L-l}{L} \frac{1}{N} \sum_{n=0}^{N-l-1} a_{n+l} a_{n}\right]\right)
$$

Proof of Proposition 7.1. Induct on the degree $D$ of the polynomials. Assume the condition holds for all $\left\{p_{n}\right\}$ of degree less than $D$. Note that the $D=1$ case corresponds to power ergodicity. Let $\left\{p_{n}\right\}$ be a sequence of polynomials of degree $D$. Let $B \in \mathcal{B}$. Fix $\epsilon>0$. Fix $L \in \mathbb{N}$ such that $1 / L<\epsilon$. For $n \in \mathbb{N}$ and $0<l<L$, set $P_{n, l}(j)=p_{n}(j+l)-p_{n}(j)$. Write $p_{n}(j)=$ $\sum_{a=0}^{D} c_{n, a} j^{a}$ for $c_{n, a} \in \mathbb{Q}$. Then $P_{n, l}(j)=\sum_{b=0}^{D-1}\left(\sum_{a=b+1}^{D} c_{n, a}\binom{a}{b} l^{a-b}\right) j^{b}$ is a polynomial in $j$ of degree $D-1$. By the inductive hypothesis, for each $0<l<L$, there exists $N_{l} \in \mathbb{N}$ such that for all $n \geq N_{l}$,

$$
\int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{-P_{n, l}(j)}\right| d \mu<\epsilon .
$$

Set $N_{0}=\lceil L / \epsilon\rceil$ so $l / n<\epsilon$ for all $0<l<L$ and $n \geq N_{0}$. For $n \geq$ $\max _{0 \leq l<L} N_{l}$,

$$
\begin{aligned}
\left\lvert\, \frac{1}{n} \sum_{j=0}^{n-l-1}\left(\mu \left(T^{-p_{n}(j+l)+p_{n}(j)} B\right.\right.\right. & \cap B)-\mu(B) \mu(B)) \mid \\
& \leq \int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{-P_{n, l}(j)}\right| d \mu+\frac{l}{n}<2 \epsilon
\end{aligned}
$$

By Hölder's inequality and van der Corput's inequality,

$$
\begin{aligned}
& {\left[\int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{-p(j)}\right| d \mu\right]^{2}} \\
& \quad \leq \frac{n+L}{n}\left(\frac{1}{L}+\frac{2}{L} \sum_{l=1}^{L-1} \frac{L-l}{L n} \sum_{j=0}^{n-l-1}\left(\mu\left(T^{-p_{n}(j+l)+p_{n}(j)} B \cap B\right)-\mu(B) \mu(B)\right)\right) \\
& \quad<(1+\epsilon)\left(\epsilon+\frac{2}{L} \sum_{l=1}^{L-1} \frac{L-l}{L} 2 \epsilon\right) \leq 5 \epsilon(1+\epsilon) .
\end{aligned}
$$

7.2. Polynomial staircase transformations. Let $\left\{p_{n}\right\}$ be a sequence of polynomials with bounded degree and uniformly bounded coefficients. A rank-one transformation with spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ given by $s_{n, j}=$ $p_{n}(j)$ is a polynomial staircase transformation when the polynomials are such that for every $L \in \mathbb{N}$,

$$
\limsup _{n \rightarrow \infty} \frac{1}{r_{n}} \#\left\{j \in \mathbb{Z}_{r_{n}}: L \text { divides } p_{n}(j+1)-p_{n}(j)\right\}<1 .
$$

A staircase transformation is then a linear polynomial staircase transformation.

Theorem 7.3. Polynomial staircase transformations are mixing.
Proposition 7.4 (Furstenberg). Let $\left\{p_{n}\right\}$ be a sequence of polynomials of bounded degree and uniformly bounded coefficients. Then a dynamical sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ given by $s_{n, j}=p_{n}(j)$ is ergodic with respect to any totally ergodic transformation.

Proof. Since

$$
\int\left|\frac{1}{r_{n}} \sum_{j=0}^{r_{n}-1} \chi_{B} \circ T^{-p_{n}(j)}\right| d \mu=\int\left|\frac{1}{r_{n}} \sum_{j=0}^{r_{n}-1} z^{p_{n}(j)}\right| d \sigma_{B}(z)
$$

where $\sigma_{B}$ is the spectral measure of $B$ under $T$, the result follows from Weyl's theorem on the equidistribution of polynomial sequences (that $n^{-1} \sum_{j=0}^{n-1} z^{p_{n}(j)} \rightarrow 0$ for all $\left.z \in S^{1} \backslash \mathbb{Q}\right)$.

Proposition 7.5. Let $T$ be a rank-one transformation with spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ such that for each $L \in \mathbb{N}, L \neq 1$,

$$
\limsup _{n \rightarrow \infty} \frac{1}{r_{n}} \#\left\{j \in \mathbb{Z}_{r_{n}-1}: L \text { divides } s_{n, j+1}-s_{n, j}\right\}<1
$$

Then $T$ is totally ergodic.
Proof. Suppose $T^{L}$ is not ergodic for some $L \in \mathbb{N}, L \neq 0,1$. Then there exists $B \in \mathcal{B}, 0<\mu(B)<1$, such that $T^{L} B=B$. We may assume $L$ is minimal so there exists $\delta>0$ such that $\mu\left(T^{l} B \cap B\right)<(1-\delta) \mu(B)$ for all $0<$ $l<L$. Define $\rho_{n}: \mathbb{Z}_{L} \rightarrow[0,1]$ by $\rho_{n}(l)=r_{n}^{-1} \#\left\{j \in \mathbb{Z}_{r_{n}}: s_{n, j}=l \bmod L\right\}$. By hypothesis there exists $\delta_{1}>0$ such that $\#\left\{j \in \mathbb{Z}_{r_{n}}: L\right.$ divides $s_{n, j+1}-$ $\left.s_{n, j}\right\} \geq \delta_{1} r_{n}$ for all sufficiently large $n$. So for at least $\delta_{1} r_{n}$ values of $j$, $s_{n, j+1} \neq s_{n, j} \bmod L$. Hence $\rho_{n}(l) \leq 1-\delta_{1}$ for all $0 \leq l<L$. Let $\left\{h_{n}\right\}$ be the height sequence for $T$. For any $\epsilon>0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ there exists a union $B_{n}$ of levels in $C_{n}$ such that $\mu\left(B \triangle B_{n}\right)<\epsilon$ (here $\triangle$ denotes symmetric difference). Write $B_{n}=\bigcup_{i \in \Lambda_{n}} I_{n, i}$. Since $L$ is fixed, $L / h_{n} \rightarrow 0$ so we may assume that $T^{l} B_{n}, 0 \leq l<L$, is a union of levels in $C_{n}$. Set $w_{n}, v_{n}$ such that $h_{n}=w_{n} L+v_{n}$ for some $0 \leq v_{n}<L$. Set
$B_{n}^{\prime}=T^{v_{n}} B_{n}$. Then $B_{n}^{\prime}$ is a union of levels in $C_{n}$. Hence

$$
\begin{aligned}
\mu\left(T^{w_{n} L} B \cap B\right) & <\mu\left(T^{h_{n}} B_{n} \cap T^{v_{n}} B_{n}\right)+2 \epsilon=\sum_{i \in \Lambda_{n}} \sum_{j=0}^{r_{n}-1} \mu\left(T^{h_{n}} I_{n, i}^{[j]} \cap B_{n}^{\prime}\right)+2 \epsilon \\
& \leq \frac{1}{r_{n}} \sum_{j=0}^{r_{n}-1} \mu\left(T^{-s_{n, j}} B_{n} \cap B_{n}^{\prime}\right)+2\left(\epsilon+\frac{1}{r_{n}}+\mu\left(S_{n}\right)\right) \\
& =\sum_{l=0}^{L-1} \rho_{n}(l) \mu\left(T^{-l-v_{n}} B_{n} \cap B_{n}\right)+2\left(\epsilon+\frac{1}{r_{n}}+\mu\left(S_{n}\right)\right) \\
& <\sum_{l=0}^{L-1} \rho_{n}(l) \mu\left(T^{-l} B \cap B\right)+2\left(2 \epsilon+\frac{1}{r_{n}}+\mu\left(S_{n}\right)\right) \\
& \leq\left(1-\delta_{1}\right) \mu(B)+\delta_{1}(1-\delta) \mu(B)+2\left(2 \epsilon+\frac{1}{r_{n}}+\mu\left(S_{n}\right)\right)
\end{aligned}
$$

since $\rho_{n}(l) \leq 1-\delta_{1}$ for $0 \leq l<L$ and $\mu\left(T^{l} B \cap B\right) \leq(1-\delta) \mu(B)$ for $0<l<L$. Then $\mid \mu\left(T^{w_{n} L} B \cap B\right)<\left(1-\delta \delta_{1}\right) \mu(B)+4 \epsilon+2 / r_{n}+2 \mu\left(S_{n}\right)$. But $T^{w_{n} L} B=B$ is then a contradiction.

Proposition 7.6. Let $T$ be a rank-one transformation with spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ such that for each fixed $k$, the partial sum sequence $\left\{s_{n, j}^{(k)}\right\}_{\left\{r_{n}-k\right\}}$ is ergodic with respect to $T$. Then for any sequence $\left\{k_{n}\right\}$ such that $\limsup k_{n} / n<\infty$ and any $B \in \mathcal{B}$,

$$
\lim _{n \rightarrow \infty} \int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{j k_{n}}\right| d \mu=0
$$

Proof. First, let $\left\{t_{n}\right\}$ be a sequence such that $h_{n} \leq t_{n} \leq d h_{n}$ for some fixed integer $d$. Write $t_{n}=x_{n} h_{n}+m_{n}$ for some $0 \leq m_{n}<h_{n}$ and $1 \leq x_{n} \leq d$. Then for any level $I_{n, i}$ with $i<h_{n}-m_{n}$ we see that

$$
T^{t_{n}} I_{n, i}=T^{x_{n} h_{n}} I_{n, i+m_{n}}
$$

and for $I_{n, i}$ with $i \geq h_{n}-m_{n}$,

$$
T^{t_{n}} I_{n, i}=T^{\left(x_{n}+1\right) h_{n}} I_{n, i+m_{n}-h_{n}}
$$

For any union $B$ of levels in some $C_{N}$ and any $n \geq N$ let $B_{x}$ be the levels in $I \in C_{n} \cap B$ such that $T^{t_{n}} I=T^{x h_{n}} I^{\prime}$ for some level $I^{\prime}$. Then $B$ is the disjoint union of $B_{x}$ for $1 \leq x \leq d+1$. By Proposition 5.8 , the sequence $\left\{x h_{n}\right\}$ is mixing since the partial sum sequence $\left\{s_{n, j}^{(k)}\right\}_{\left\{r_{n}-k\right\}}$ is ergodic with respect to $T$ for each fixed $x$; in fact, as can be seen from the proof of Proposition 5.8,

$$
\mu\left(T^{x h_{n}} A \cap B\right)-\mu(A) \mu(B) \rightarrow 0
$$

uniformly over all collections $A$ of levels in $C_{n}$. Since there are only $d$ choices of $x$ this means that $\left\{t_{n}\right\}$ is mixing.

Now, let $\left\{k_{n}\right\}$ be a (not necessarily increasing) sequence such that $k_{n} / n$ is bounded. Fix $p, q$ and $x$ such that $h_{p} \leq k<h_{p+1} \leq k q<2 h_{p+1}$ and $x h_{p} \leq k<(x+1) h_{p}$ (again we drop the $n$ subscripts for clarity). Let $B \in \mathcal{B}$ and $\epsilon>0$. Choose $L$ such that $1 / L<\epsilon(L$ not depending on $n)$. Note that

$$
\frac{q}{n}=\frac{q k^{2}}{n k^{2}} \leq \frac{2 h_{p+1}}{h_{p}^{2}} \frac{k}{n} \rightarrow 0
$$

since $h_{p+1} / h_{p}^{2} \approx r_{p} / h_{p} \rightarrow 0$ as $T$ is finite-measure-preserving. By the Block Lemma,

$$
\int\left|\frac{1}{n} \sum_{j=0}^{n-1} \chi_{B} \circ T^{-j k}\right| d \mu \leq \int\left|\frac{1}{L} \sum_{l=0}^{L-1} \chi_{B} \circ T^{-l k q}\right| d \mu+\frac{q L}{n}
$$

Now $t_{n}=x k_{n} q_{n}$ satisfies the requirements above so $\left\{x k_{n} q_{n}\right\}$ is a mixing sequence for each fixed $x$. Then, by Hölder's inequality,

$$
\left(\int\left|\frac{1}{L} \sum_{l=0}^{L-1} \chi_{B} \circ T^{-l k q}\right| d \mu\right)^{2}=\frac{1}{L} \sum_{x=-L+1}^{L-1} \frac{L-x}{L}\left(\mu\left(T^{x k_{n} q_{n}} B \cap B\right)-\mu(B) \mu(B)\right)
$$

and each term where $x \neq 0$ tends to zero by the mixing of the sequence, and the $x=0$ term contributes at most $1 / L<\epsilon$.

Proof of Theorem 7.3. Let $T$ be a polynomial staircase transformation. Then $T$ is totally ergodic by Proposition 7.5. Let $\left\{p_{n}\right\}$ be the polynomials defining the spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ of degree at most $D \in \mathbb{N}$ and let $\left\{c_{n, a}\right\}$ for $a \in \mathbb{Z}_{D+1}$ be the coefficients of the $\left\{p_{n}\right\}$. Then, for any $j \in \mathbb{Z}_{r_{n}}$ and $k \in \mathbb{Z}_{r_{n}-j}$,

$$
\begin{aligned}
s_{n, j}^{(k)} & =\sum_{z=0}^{k-1} p_{n}(j+z)=\sum_{z=0}^{k-1} \sum_{a=0}^{D} c_{n, a}(j+z)^{a}=\sum_{z=0}^{k-1} \sum_{a=0}^{D} \sum_{b=0}^{a} c_{n, a}\binom{a}{b} j^{b} z^{a-b} \\
& =\sum_{b=0}^{D}\left(\sum_{z=0}^{k-1} \sum_{a=b}^{D} c_{n, a}\binom{a}{b} z^{a-b}\right) j^{b}=p_{n, k}(j)
\end{aligned}
$$

are polynomials of degree at most $D$ in $j$ with lead coefficients $k c_{n, D}$.
Assume now that $T$ is power ergodic. Then $T$ is polynomial power ergodic by Proposition 7.1. For any $\left\{k_{n}\right\}$ and $\left\{\Gamma_{n, q}\right\},\left\{\Psi_{n, q}\right\},\left\{\alpha_{n, q}\right\}_{\left\{Q_{n}\right\}}$ as in Proposition 5.11, each $\left\{s_{n, \Psi_{n, q}(j)}^{\left(k_{n}-\alpha_{n, q}\right)}\right\}_{\left\{\# \Gamma_{n, q}\right\}}$ is itself a polynomial sequence (details are left to the reader) of degree at most $D$. Since $Q_{n} / r_{n} \rightarrow 0$, we see that $\# \Gamma_{n, q} \rightarrow \infty$ uniformly over a density one set of $q$; consequently, the averages over the sets $\Gamma_{n, q}$ tend uniformly to zero by polynomial power
ergodicity. Proposition 5.11 then gives the result. It only remains to show that $T$ is power ergodic.

Proposition 7.7. Let $T$ be a polynomial staircase transformation. Then $T$ is power ergodic.

Proof. For each fixed $k \in \mathbb{N}$ the $p_{n, k}$ are ergodic (Proposition 7.4) since the coefficients of $p_{n, k}$ are uniformly bounded when $k$ is fixed (see the proof above for an expression for the coefficients). Then by Proposition 7.6 the case when $\lim \sup _{n} k_{n} / n<\infty$ is done. Write $k=k_{n}=x h_{p}+m$ as usual. Cases 1a and 1b from Proposition 4.13 carry over directly (using Proposition 7.6). Hence we may assume that $\liminf _{n} x r_{p} / h_{p}>0$ and $x / r_{p} \rightarrow 0$. Pick $l_{n}$ as in Case 3 of Proposition 4.13 and $Q_{n}$ and $\alpha(q)$ as well (note that $\lim \sup _{n} h_{p} / x r_{p}<\infty$ is enough for that construction).

Then, using the techniques of the proof of Proposition 7.1 (applying van der Corput's lemma $d-1$ times where $d$ is the degree of the polynomials defining the spacer sequence), for any $\Gamma_{n, q}$ an appropriate subset of $\left\{s_{p, j}^{(x)}\right\}_{\left\{r_{p}-x\right\}}$ and any appropriate $\alpha(q)=\alpha(n, q)$,

$$
\begin{aligned}
& \int\left|\frac{1}{\# \Gamma_{n, q}} \sum_{j \in \Gamma_{n, q}} \chi_{B} \circ T^{-s_{p, j}^{(x+2-\alpha(q))}}\right| d \mu \\
& \leq\left(\int\left|\frac{1}{l_{n}} \sum_{j=0}^{l_{n}-1} \chi_{B} \circ T^{j(x+2-\alpha(q)) z(q)}\right| d \mu\right)^{\delta}+\epsilon
\end{aligned}
$$

for some $\delta>0$, integers $z(q)=z(n, q)$ and $\epsilon=\epsilon_{n} \rightarrow 0$. By the Block Lemma we may replace $l_{n}$ by some large fixed $L$ (with error $L / l_{n} \rightarrow 0$ ).

Now $\alpha(q)$ take on all integer values in an interval of length arbitrarily large, so as in the proof of Proposition 4.13 Case 3 , using that $T$ is weak mixing (so $T^{d}$ is weak mixing for all fixed $d \neq 0$ ) we find that the quantity above tends to zero when averaged over $q$. The result then follows as in Proposition 4.13 Case 3 (details here have been left to the reader as everything follows from techniques previously used in the paper).

## 8. Specific examples

### 8.1. Criterion for finite measure on rank-one transformations

Proposition 8.1. A rank-one transformation with heights $\left\{h_{n}\right\}$ and spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ is defined on a finite measure space if and only if

$$
\sum_{n=0}^{\infty} \frac{\bar{s}_{n}}{h_{n}}<\infty \quad \text { where } \quad \bar{s}_{n}=\frac{1}{r_{n}} \sum_{j=0}^{r_{n}-1} s_{n, j}
$$

Proof. Let $T,\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}},\left\{\bar{s}_{n}\right\}$, and $\left\{h_{n}\right\}$ be as above and let $(X, \mu)$ be the space $T$ is defined on. Let $\left\{C_{n}\right\}$ denote the columns of the construction as sets, $\left\{I_{n}\right\}$ the base levels of the columns, and $\left\{S_{n}\right\}$ the spacers added (so $S_{n}=C_{n+1} \backslash C_{n}$ and $\left.S_{0}=C_{0}\right)$. We see that

$$
\mu\left(S_{n}\right)=\sum_{j=0}^{r_{n}-1} s_{n, j} \mu\left(I_{n+1}\right)=\left(\frac{1}{r_{n}} \sum_{j=0}^{r_{n}-1} s_{n, j}\right) \mu\left(I_{n}\right)=\frac{\bar{s}_{n}}{h_{n}} \mu\left(C_{n}\right)
$$

and so

$$
\sum_{n=0}^{\infty} \mu\left(S_{n}\right)=\mu\left(C_{0}\right) \prod_{n=0}^{\infty}\left(1+\frac{\bar{s}_{n}}{h_{n}}\right)
$$

### 8.2. Specific examples of mixing transformations

Theorem 8.2. For $D \in \mathbb{N}$ and $\delta \in \mathbb{R}^{+}$, let $T_{D, \delta}$ be the rank-one transformation with spacer sequence $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ given by $s_{n, j}=j^{D}$ and $r_{n}=$ $\left\lfloor h_{n}^{1 /(D+\delta)}\right\rfloor\left(\right.$ where $\left\{h_{n}\right\}$ is the heights for $\left.T_{D, \delta}\right)$. Then $T_{D, \delta}$ is a mixing transformation.

Proof. By Theorem 7.3, we need only show the transformations are defined on a finite measure space. Let $D \in \mathbb{N}$ and $\delta \in \mathbb{R}^{+}$. Let $\left\{s_{n, j}\right\}_{\left\{r_{n}\right\}}$ be the spacers and $\left\{h_{n}\right\}$ the heights for $T_{D, \delta}$. Now, $\sum_{j=0}^{r_{n}-1} s_{n, j}=\sum_{j=0}^{r_{n}-1} j^{D} \approx r_{n}^{D+1}$. Since $h_{n+1}=r_{n} h_{n}+\sum_{j=0}^{r_{n}-1} s_{n, j}$, we see that $h_{n} \geq \prod_{z=0}^{n-1} r_{z} \geq 2^{n}$. Then

$$
\begin{aligned}
\sum_{n=0}^{\infty} \frac{\bar{s}_{n}}{h_{n}} & \approx \sum_{n=0}^{\infty} \frac{1}{r_{n} h_{n}} r_{n}^{D+1}=\sum_{n=0}^{\infty} \frac{r_{n}^{D}}{h_{n}} \approx \sum_{n=0}^{\infty} h_{n}^{(D) /(D+\delta)-1} \\
& \leq \sum_{n=0}^{\infty}\left(2^{\delta /(D+\delta)}\right)^{-n}<\infty
\end{aligned}
$$

by the convergence of geometric series. Proposition 8.1 completes the proof.
8.3. Ornstein's transformation. The original construction of rankone mixing transformations, due to Ornstein Orn72], involved placing spacer levels randomly using a uniform distribution so that almost surely the resulting transformation is mixing. The uniform distribution can be equally well interpreted as meaning that the spacer sequence almost surely has the property that the averages in Theorem 6.1 tend uniformly to zero so that mixing for these transformations follows from our theorem. The reader is referred to CS04 for details.
8.4. A note on restricted growth. The restricted growth condition in CS04] is equivalent to $r_{n} \bar{s}_{n} / h_{n} \rightarrow 0$ for polynomial staircase transformations (details are left to the reader) and to Adams' condition $r_{n}^{2} / h_{n} \rightarrow 0$ for staircase transformations. For $T_{D, \delta}$, we see that $r_{n} \bar{s}_{n} / h_{n} \approx r_{n}^{D+1} / h_{n} \approx$
$h_{n}^{(D+1) /(D+\delta)-1}=h_{n}^{(1-\delta) / D+\delta}$ and so $T_{D, \delta}$ has restricted growth if and only if $\delta>1$. Hence, the theorems of CS04 and Ada98 apply only to $T_{D, \delta}$ with $\delta>1$.

Acknowledgments. The authors would like to thank the referee for many helpful suggestions, and in particular for suggesting that first proving the special case for staircases and then generalizing makes for a more readable paper than the original exposition which directly proved the general case.

## References

[Ada98] T. Adams, Smorodinsky's conjecture on rank-one mixing, Proc. Amer. Math. Soc. 126 (1998), 739-744.
[AF92] T. Adams and N. Friedman, Staircase mixing, preprint, 1992.
[CS04] D. Creutz and C. E. Silva, Mixing on a class of rank-one transformations, Ergodic Theory Dynam. Systems 24 (2004), 407-440.
[Fer97] S. Ferenczi, Systems of finite rank, Colloq. Math. 73 (1997), 35-65.
[Fri70] N. A. Friedman, Introduction to Ergodic Theory, Van Nostrand, Princeton, NJ, 1970.
[Kal84] S. A. Kalikow, Twofold mixing implies threefold mixing for rank one transformations, Ergodic Theory Dynam. Systems 4 (1984), 237-259.
[Kin88] J. L. King, Joining-rank and the structure of finite rank mixing transformations, J. Anal. Math. 51 (1988), 182-227.
[Orn72] D. S. Ornstein, On the root problem in ergodic theory, in: Proc. Sixth Berkeley Sympos. Mathematical Statistics and Probability (Berkeley, CA, 1970-1971), Vol. 2, Univ. of California Press, 1972, 347-356.
[Ryz93] V. V. Ryzhikov, Joinings and multiple mixing of the actions of finite rank, Funktsional. Anal. i Prilozhen. 27 (1993), no. 2, 63-78 (in Russian).
[Ryz03] -, private communication, 2003.
Darren Creutz
Cesar E. Silva
Department of Mathematics
University of California, Los Angeles
Box 951555
Los Angeles, CA 90095-1555, U.S.A.
E-mail: dcreutz@math.ucla.edu


[^0]:    2010 Mathematics Subject Classification: Primary 37A25; Secondary 28D05.
    Key words and phrases: mixing, rank-one, staircase transformations.

