Journal

Israel Journal of Mathematics, 251(1)

ISSN

0021-2172

Authors

Foreman, Matthew

Weiss, Benjamin

Publication Date

2022-12-01

DOI

10.1007/s11856-022-2439-7

ISRAEL JOURNAL OF MATHEMATICS **251** (2022), 327–364
DOI: 10.1007/s11856-022-2439-7

ODOMETER BASED SYSTEMS

BY

Matthew Foreman*

Department of Mathematics, University of California at Irvine
Irvine, CA 92717-0001, USA
e-mail: mforeman@math.uci.edu

AND

Benjamin Weiss

Einstein Institute of Mathematics, The Hebrew University of Jerusalem
Givat Ram, Jerusalem 91904, Israel
e-mail: weiss@math.huji.ac.il

ABSTRACT

Construction sequences are a general method of building symbolic shifts that capture cut-and-stack constructions and are general enough to give symbolic representations of Anosov–Katok diffeomorphisms. We show here that any finite entropy system that has an odometer factor can be represented as the limit of a special class of construction sequences, the odometer based construction sequences. These naturally correspond to those cut-and-stack constructions that do not use spacers. The odometer based construction sequences can be constructed to have the small word property and every Choquet simplex can be realized as the simplex of invariant measures of the limit of an odometer based construction sequence.

* The first author would like to acknowledge partial support from NSF grant DMS-2100367.

Received August 11, 2021 and in revised form January 19, 2022

Preface

This preface is due solely to the first author.

I had the privilege to meet Benjy Weiss in the early 1990s. I was surprised by his openness and ability to interact with a stranger in a different field. Indeed our joint work, which covers decades and hundreds of published pages (and many more pages of unpublished work), has been a tribute to his willingness to work with someone with a completely different background and training.

I humorously, but sincerely, categorize Benjy's roles as:

Inspiration: Benjy is the perfect mathematical role model. Unparalleled modesty, openness in sharing ideas and fairness combined with unlimited talent.

Oracle: Ask any question and his remarkable memory will nearly instantly retrieve the answer if it is known.

Speculator: He is completely willing to discuss paths that seem implausible at first and see where they lead. The fact that they are often empty doesn't seem to bother him.

Friend and Mentor: This is perhaps his most important role for me.

With these points in mind, I sincerely thank Benjy Weiss for the years of companionship and productivity.

1. Introduction

Construction sequences are a general method of building symbolic shifts that capture cut-and-stack constructions and are general enough to give symbolic representations of Anosov–Katok diffeomorphisms. This paper studies a special class of construction sequences, the odometer based construction sequences that corresponds to those cut-and-stack constructions that don't use spacers.

In [7] we show that there is a functorial isomorphism between the symbolic systems that are limits of odometer based construction sequences and symbolic systems that are limits of a class of construction sequences called circular systems. Many circular systems, in turn, can be realized as diffeomorphisms of the 2-torus. As a corollary the qualitative ergodic theoretic structure of the odometer based systems is reflected in the diffeomorphisms of the 2-torus. For example, in [9] it is shown that there are measure-distal diffeomorphisms of the torus of all countable ordinal heights and for all Choquet simplices $\mathcal{K}$, there is a Lebesgue measure preserving ergodic diffeomorphism of the torus that has $\mathcal{K}$ as its simplex of invariant measures. This uses Corollary 2 in this paper.

To use the functor defined in [7] to build diffeomorphisms with complicated behavior, one needs that the class of transformations isomorphic to limits of odometer based construction sequences exhibits quite rich ergodic phenomena. This is the point of the current paper.

It is a classical theorem of Krieger ([10]) that an ergodic system with finite entropy has a finite generating partition. This gives a symbolic representation for any such system and shows that the theory of finite entropy ergodic measure preserving systems coincides with the theory of finite valued ergodic stationary processes $\{X_n\}$. When studying stationary processes $\{X_n\}$ it is often useful to have a block structure, namely a way of dividing the indices into a hierarchy of blocks of lengths $k_1, k_1k_2, k_1k_2k_3, \ldots$ in a unique fashion. If this is possible then the process will have as a factor the odometer transformation corresponding to the sequence $\{k_n\}$. Our main theorem (Theorem 21) is that it is always possible to find such a symbolic representation with a rather simple form whenever this necessary condition is satisfied.

Theorem (in Section 4): *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system with finite entropy. Then X has an odometer factor if and only if X is isomorphic to an odometer based symbolic system.*

The class of ergodic transformations containing an odometer factor is easily characterized spectrally as those ergodic transformations whose associated unitary Koopman operator has infinitely many eigenvalues of finite multiplicative order. (We prove this in Proposition 6.)

If T is totally ergodic, i.e., all powers are ergodic, then the product of T with any odometer is ergodic. In general we have the following easy proposition which illustrates the ubiquity of ergodic transformations with odometer factors:

Proposition 1: *Given any ergodic transformation $\mathbb{X} = (X, \mathcal{B}, \mu, T)$ either:*

(1) *$\mathbb{X}$ has an odometer factor, or:*
(2) *there is an odometer $\mathfrak{O}$ such that $\mathbb{X} \times \mathfrak{O}$ is ergodic (and $\mathbb{X} \times \mathfrak{O}$ has finite entropy if $\mathbb{X}$ does).*

In particular, every finite entropy transformation is a factor of a finite entropy odometer based symbolic system and the finite entropy transformations that have an odometer factor are closed under finite entropy extensions.

Proof. If $\mathbb{X}$ does not already have an odometer factor, then the Koopman operator associated to $\mathbb{X}$ has finitely many eigenvalues with finite order. Let $\mathfrak{O}$ be an odometer such that the eigenvalues of the Koopman operator of $\mathfrak{O}$ are relatively prime to the orders of the eigenvalues associated to $\mathbb{X}$. It is then easily verified that $\mathbb{X} \times \mathfrak{O}$ is ergodic. ∎

We remark that we can put a pre-partial ordering $\leq_F$ on the set of ergodic transformations $\mathcal{E}_{\mathrm{FE}}$ with finite entropy by setting $(X, T) \leq_F (Y, S)$ if and only if there is a factor map $\pi : Y \to X$. A standard definition from the theory of pre-partial orderings is that a set $C \subseteq \mathcal{E}$ is a **cone** relative to $\leq_F$ iff and only if:

C is **closed upwards:** if $(X, T) \in C$ and $X \leq_F Y$ then $Y \in C$,

Extension: for all ergodic (X, T) with finite entropy there is a $Y \in C$ with $X \leq_F Y$.

The relevance of cones is that the set of cones generate a non-principal filter relative to $\leq_F$. Hence a cone can be viewed as a set of measure one for a finitely additive measure on the $\mathcal{E}_{FE}$ with the ordering $\leq_F$. In this rigorous sense Proposition 1 shows that the set of transformations of finite entropy with an odometer factor is a large set.

Corollary 2: *For all finite or countable ordinals α there is an ergodic measure distal, odometer based system of distal height α.*

Proof. Fix a finite or countable ordinal α. By the results of Beleznay and Foreman ([2]) there is a measure distal transformation T of distal height α. Since T is measure distal it has entropy 0. The Koopman operator corresponding to T has no eigenvalues of finite order. Hence we can choose a sequence $\langle k_n : n \in \mathbb{N}\rangle$ going to infinity and consider the odometer transformation $\mathcal{O}$ based on the sequence $\langle k_n : n \in \mathbb{N}\rangle$. (See section 2.2 for a formal definition.) By Proposition 1, $T \times \mathcal{O}$ is an ergodic odometer based transformation.

Using Lemma 2.8 of [2] (and the discussion surrounding it), one sees that $T \times \mathcal{O}$ has distal height α. ∎

We should point out that special symbolic processes with a block structure, called Toeplitz systems, have been well studied from the point of view of topological dynamics. (See, e.g., [4], [3], [13].)

The structure of this paper. Section 2 has the basic definitions used in the paper as well as properties of odometer based systems that we use in the construction. Section 4 contains the proof of our main theorem, Theorem 21.

It begins by pointing out a known fact that odometers cannot be represented topologically as symbolic shifts, in contrast to Theorem 21, which is in the measure category. As a precursor it then presents the odometer as an odometer based system, describes the plan of the proof and finally gives the proof in detail.

In Section 5 we discuss the connections with Toeplitz systems, showing how to augment a Toeplitz system to get an odometer based system while preserving the simplex of invariant measures. It then follows from a remarkable theorem of Downarowicz [3] (generalizing work of Williams [13]) saying that arbitrary simplices of invariant measures can be realized on Toeplitz sequences to see that arbitrary simplices of invariant measures can be realized on limits of odometer based construction sequences.

The applications of this paper require that the odometer based construction sequences in the domain of the isomorphism functor has the frequencies of words decreasing arbitrarily fast. We call this the small word property. In Section 6 we define the small word property and show that we can realize odometer based systems continuously in a sequence of small word requirements.

The road not taken. After a draft of this paper was circulated, it was pointed out that there is an alternate proof of our result about simplices of invariant measures. It is based on a result of Downarowicz and Lacroix ([4], theorem 8) which states that every transformation satisfying the hypothesis of our main theorem can be represented as the orbit closure of a Toeplitz sequence. Then applying Proposition 32 in this paper presents orbit closures of Toeplitz sequences as limits of odometer based construction sequences, giving an alternate proof. However, the proof that is given in [4] makes key use of a result in Weiss' paper, [11]. In that paper only a brief sketch of a more general theorem is given and the specific result that they need for the alternate proof is not even mentioned there. Moreover, the applications of the results in this paper in [9] use the specific representation given here that has the small word property. This idea is new to this paper and is completely missing from the alternate proof.

We also note the work of Williams presenting the odometer itself as a limit of a construction sequence (see Williams, [13]), as well as the recent work of Adams, Ferenczi, and Petersen [1], which realizes generalized odometers and indeed all rank one systems as "constructive symbolic rank one systems," in the terminology of [5].

2. Preliminaries

An **alphabet** Σ is a finite collection of symbols. A **word** in Σ is a finite sequence of elements of Σ. If $w \in \Sigma^{<\mathbb{N}}$ is a word, we denote its length by $|w|$. By $\Sigma^{\mathbb{Z}}$ we mean doubly infinite sequences of letters in Σ. This has a natural product topology induced by the discrete topology on Σ. This topology is compact if Σ is finite. For u a word in Σ and $k \in \mathbb{Z}$, we use the notation $\langle u \rangle_k$ for the basic open interval in $\Sigma^{\mathbb{Z}}$ consisting of $\{f \in \Sigma^{\mathbb{Z}} : f \restriction [k, k+|u|) = u\}$. If we omit the k, $\langle u \rangle$ means $\langle u \rangle_0$.

Perhaps the thorniest issue in the paper is defining right vs left shifts, particularly in view of the cultural differences. We define the **shift** given $\mathbf{Sh} : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ by setting $\mathbf{Sh}(f)(n) = f(n+1)$. For this paper $\mathbf{Sh}$ is the **left shift** because f is being shifted left. The **right shift** is $\mathbf{Sh}^{-1}$. Similarly an occurrence of a word $u \in \Sigma^{<\mathbb{N}}$ in an $f \in \Sigma^{\mathbb{Z}}$ at a $k \in \mathbb{Z}$ is **to the left** of an occurrence $y \in \Sigma^{<\mathbb{N}}$ at l if $k < l$. A **symbolic system** is a closed, shift-invariant $\mathbb{K} \subseteq \Sigma^{\mathbb{Z}}$ for some Σ.

A collection of words $\mathcal{W}$ is **uniquely readable** if and only if whenever $u, v, w \in \mathcal{W}$ and $uv = pws$ then either p or s is the empty word.

We note that we can view both words and elements of $\Sigma^{\mathbb{Z}}$ as functions. If $f : A \to B$ and $A' \subseteq A$, the restriction of f to A' is denoted $f \restriction A'$.

Given a collection $\langle w_i : 0 \le i \le n \rangle$ of finite words in some alphabet, we let

$$\prod_{i=0}^{n} w_i$$

denote the concatenation $w_0 w_1 w_2 \cdots w_n$. Similarly for a single word w and $k \ge 1$ we let w^k denote the concatenation of k-copies of w. So, for example, $w^3 = www$.

2.1. Partitions and symbolic systems. Let $(X, \mathcal{B}, \mu), (Y, \mathcal{C}, \nu)$ be standard probability measure spaces and let $\pi : X \to Y$. Then ν is the measure **induced by μ and π** if and only if for all $C \in \mathcal{C}, \pi^{-1}(C) \in \mathcal{B}$ and $\nu(C) = \mu(\pi^{-1}(C))$. In symbols we write $\nu = \pi^*(\mu)$.

An **ordered partition** of X is a set $\mathcal{P} = \langle A_0, A_1, \ldots \rangle$ such that each $A_i \in \mathcal{B}$, $A_i \cap A_j = \emptyset$ if $i \neq j$, and $X = \bigcup_i A_i$. We allow our partitions to be finite or countable and identify two partitions $\mathcal{P} = \langle A_i \rangle$, $\mathcal{Q} = \langle B_j \rangle$ if for all i,

$$\mu(A_i \Delta B_i) = 0.$$

We will frequently refer to ordered countable measurable partitions simply as **partitions**. A partition is finite iff for all large enough $n, \mu(P_n) = 0$. If $\mathcal{P}$ and $\mathcal{Q}$ are partitions, then $\mathcal{Q}$ **refines** $\mathcal{P}$ iff the atoms of $\mathcal{Q}$ can be grouped into sets $\langle S_n : n \in \mathbb{N}\rangle$ such that

$$\sum_n \mu\bigg(P_n \Delta\bigg(\bigcup_{i\in S_n} Q_i\bigg)\bigg) = 0.$$

In this case we will write that $\mathcal{Q} \ll \mathcal{P}$. A **decreasing sequence of partitions** is a sequence $\langle \mathcal{P}_n : n \in \mathbb{N}\rangle$ such that for all $m < n, \mathcal{P}_n \ll \mathcal{P}_m$. If $A \in \mathcal{B}$ is a measurable set and $\mathcal{P}$ is a partition, then we let $\mathcal{P} \upharpoonright A$ be the partition of A defined as $\langle P_n \cap A : n \in \mathbb{N}\rangle$.

Definition 3: Let $(X, \mathcal{B}, \mu)$ be a measure space. We will say that a sequence of partitions $\langle \mathcal{P}_n : n \in \mathbb{N}\rangle$ **generates** (or generates $\mathcal{B}$) iff the smallest σ-algebra containing $\bigcup_n \mathcal{P}_n$ is $\mathcal{B}$ (modulo measure zero sets). If T is a measure preserving transformation we will write $T\mathcal{P}$ for the partition $\langle Ta : a \in \mathcal{P}\rangle$. In the context of a measure preserving $T : X \to X$ we will say that a partition $\mathcal{P}$ is a **generator** for T iff $\langle T^i\mathcal{P} : i \in \mathbb{Z}\rangle$ generates $\mathcal{B}$.

Given a measure preserving system $(X, \mathcal{B}, \mu, T)$ and a partition $\mathcal{P}$ of X, define a map $\phi : X \to \mathcal{P}^{\mathbb{Z}}$ by setting (for each $a \in \mathcal{P}$)

$$\phi(x)(n) = a \text{ if and only if } T^n x \in a.$$

The bi-infinite sequence $\phi(x)$ will be called the $\mathcal{P}$-name of x. The closure of $\phi(X) \subseteq \mathcal{P}^{\mathbb{Z}}$ is a symbolic system.

Define a measure on $\mathcal{P}^{\mathbb{Z}}$ by setting

$$\phi^*(\mu)(A) = \mu(\phi^{-1}[A]).$$

This is a Borel measure on the symbolic shift $\mathcal{P}^{\mathbb{Z}}$ and makes $(\mathcal{P}^{\mathbb{Z}}, \mathcal{C}, \nu, \mathbf{Sh})$ into a factor of $(X, \mathcal{B}, \mu, T)$ (where $\nu = \phi^*(\mu)$). This factor map is an isomorphism if and only if $\mathcal{B}$ is the smallest shift-invariant σ-algebra containing all of the sets in $\mathcal{P}$ (up to sets of measure zero); i.e., $\mathcal{P}$ is a generator for T. In general the support of ν is the closure of $\phi(X)$.

Remark 4: Let $\mathcal{P}, \mathcal{Q}$ be partitions of X. Then $\mathcal{P}$ and $\mathcal{Q}$ determine factors $Y_{\mathcal{P}}$ and $Y_{\mathcal{Q}}$. Define $\phi : X \to Y_{\mathcal{P}} \times Y_{\mathcal{Q}}$ by setting $\phi(x) = (s_p, s_q)$ where s_p is the $\mathcal{P}$-name of x and s_q is the $\mathcal{Q}$-name of x. Let $\eta = \phi^*(\mu)$. Then $(Y_{\mathcal{P}} \times Y_{\mathcal{Q}}, \mathcal{C}, \eta, \mathbf{Sh})$ is isomorphic to the smallest factor of X containing both $Y_{\mathcal{P}}$ and $Y_{\mathcal{Q}}$ as factors.

2.2. Basic facts about odometers. Let $\langle k_i : i \in \mathbb{N}\rangle$ be an infinite sequence of integers with $k_i \geq 2$. Then the sequence k_i determines an **odometer** transformation with domain the compact space[1]

$$\boldsymbol{O} \stackrel{\text{def}}{=} \prod_i \mathbb{Z}_{k_i}.$$

The space $\boldsymbol{O}$ is naturally a monothetic compact abelian group, with the operation of addition and "carrying right". We will denote the group element $(1, 0, 0, 0, \dots)$ by $\bar{1}$, and the result of adding $\bar{1}$ to itself j times by $\bar{j}$.

The Haar measure on this group can be defined explicitly. Define a measure ν_i on each $\mathbb{Z}_{k_i}$ that gives each point measure $1/k_i$. Then Haar measure μ is the product measure of the ν_i.

The odometer transformation $\mathcal{O} : \boldsymbol{O} \to \boldsymbol{O}$ is defined by taking an $x \in \prod_i \mathbb{Z}_{k_i}$ and adding the group element $\bar{1}$, More explicitly, $\mathcal{O}(x)(0) = x(0) + 1(\text{mod } k_0)$ and $\mathcal{O}(x)(1) = x(1)$ unless $x(0) = k_0 - 1$, in which case we "carry one" and set $\mathcal{O}(x)(1) = x(1) + 1(\text{mod } k_1)$, etc.

The map $\mathcal{O} : \boldsymbol{O} \to \boldsymbol{O}$ is a topologically minimal, uniquely ergodic, invertible homeomorphism that preserves the measure μ. When we are viewing the odometer as a measure preserving system we will denote it by $\mathfrak{O}$.

Define $U_{\mathfrak{O}} : L^2(\mathfrak{O}) \to L^2(\mathfrak{O})$ by setting $U_{\mathfrak{O}}(f) = f \circ \mathcal{O}$. Then $U_{\mathfrak{O}}$ is the canonical unitary operator associated with $\mathcal{O}$. The characters $\chi \in \hat{O}$ are eigenfunctions for the $U_{\mathfrak{O}}$ since

$$\chi(x + \bar{1}) = \chi(\bar{1})\chi(x).$$

Since the characters form a basis for $L^2(\mathfrak{O})$, the odometer map has discrete spectrum.

Here is an explicit description of the characters, which we call $\mathcal{R}_n$ here. Fix $n \geq 1$ and let

$$K_n = \prod_{i<n} k_i.$$

Let $A_0 \subset \prod_i \mathbb{Z}_{k_i}$ be the collection of points whose first $n+1$ coordinates are zero, and for $0 \leq k < K_n$ set $A_k = \mathcal{O}^k(A)$. Define

$$\mathcal{R}_n = \sum_{k=0}^{K_n - 1} (e^{2\pi i/K_n})^k \mathrm{Ch}_{A_k}$$

where Ch_A is the characteristic function of A.

[1] We write $\mathbb{Z}/\mathbb{Z}_k$ as $\mathbb{Z}_k$.

Then:

(1) $\mathcal{R}_n$ is an eigenvector of $U_{\mathfrak{O}}$ with eigenvalue $e^{2\pi i/K_n}$,
(2) $(\mathcal{R}_n)^{k_n} = \mathcal{R}_{n-1}$,
(3) $\{(\mathcal{R}_n)^k : 0 \leq k < K_n, n \in \mathbb{N}\}$ form a basis for $L^2(\prod_i \mathbb{Z}_{k_i})$.

For a fixed n, the sets $\{A_i : 0 \leq i < K_n\}$ form a tower which will play a special role in our proofs. More generally, if $(X, \mathcal{B}, \mu, T)$ is an ergodic measure preserving system and $\pi : X \to \mathfrak{O}$ is a factor map, we set $B_n^i = \pi^{-1} A_i$. Then $\{B_n^i : 0 \leq i < K_n\}$ is a partition of X that forms a tower in the sense that $T[B_n^i] = B_n^{i+1}$ for $i < K_n - 1$ and $T[B_n^{K_n-1}] = B_n^0$.

Definition 5: We will call the tower $\mathcal{T}_n = \{B_n^i : 0 \leq i < K_n\}$ the **n-tower associated with $\mathfrak{O}$**.

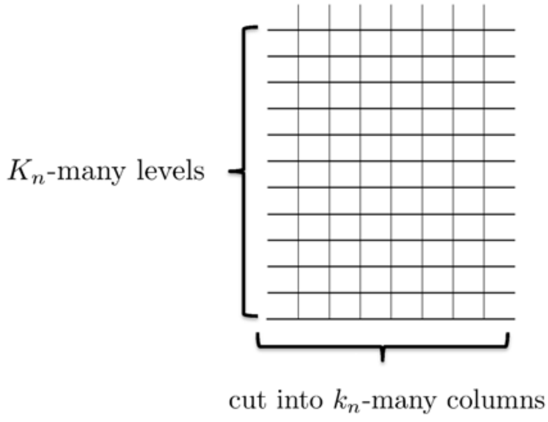

Figure 1. The tower $\mathcal{T}_n$.

Figure 1 illustrates the n^{th} tower, which we will denote $\mathcal{T}_n$. The horizontal lines represent the levels of the tower. The "$n+1^{st}$-digit" of points in $\boldsymbol{O}$ determine k_n many vertical cuts through $\mathcal{T}_n$. Enumerating the levels according to their lexicographic order in $\prod_{j<n+1} \mathbb{Z}_{k_j}$ amounts to stacking the post-cut columns of $\mathcal{T}_n$. Figure 2 illustrates the result $\mathcal{T}_{n+1}$ of the cutting and stacking.

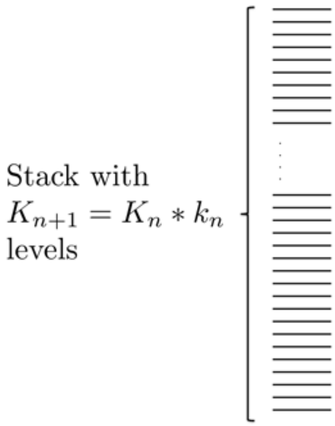

Figure 2. The tower $\mathcal{T}_{n+1}$.

Spectral characterization. Here is a standard spectral characterization of transformations with an odometer factor. Suppose that $(X, \mathcal{B}, \mu, T)$ is an ergodic measure preserving system. Let $U_T : L^2(X) \to L^2(X)$ be defined by

$$U_T(f) = f \circ T.$$

Let G be the group of eigenvalues of U_T that have finite multiplicative order (as elements of $\mathbb{C}$).

Suppose that G is infinite. Then there is a sequence of generators $\{g_n : n \in \mathbb{N}\}$ of G so that

$$o(g_n) \mid o(g_{n+1}).$$

The dual $\hat{G}$ of G is the odometer based on $\langle k_n : n \in \mathbb{N} \rangle$, where $k_n = o(g_n)$. We have outlined the proof of:

Proposition 6: *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving system. Then X has an odometer factor if and only if U_T has infinitely many eigenvalues of finite multiplicative order.*

We will use the following:

Proposition 7: *Let $\hat{G}$ be the measure preserving system described above and $\pi : X \to \hat{G}$ be the canonical factor map. If $\phi : X \to Y$ is any factor map of X to an odometer, then there is a $\pi' : \hat{G} \to Y$ such that $\phi = \pi' \circ \pi$ almost everywhere.*

Here is a useful remark.

Proposition 8: *Let $\langle k_n : n \in \mathbb{N}\rangle$ determine an odometer transformation $\mathfrak{O}$ and $K_n = \prod_{i<n} k_i$. Then for any infinite subsequence of $\langle K_{n_j} : j \in \mathbb{N}\rangle$ of $\langle K_n : n \in \mathbb{N}\rangle$, if we set $k'_0 = K_{n_0}$ and for $j \geq 1, k'_i = K_{n_j}/K_{n_{j-1}}$, then the odometer $\mathfrak{O}'$ determined by $\langle k'_j : j \in \mathbb{N}\rangle$ is isomorphic to $\mathfrak{O}$.*

In particular, an arbitrary odometer $\mathfrak{O}$ has a presentation where $\sum 1/k_n < \infty$.

2.3. Invariant measures. Let X be a compact separable metric space and $T : X \to X$ be a homeomorphism. Then the collection of T-invariant probability measures on the Borel subsets of $X, \mathcal{M}(X,T)$, endowed with the weak topology, forms a Choquet simplex $\mathcal{K}$: a compact, metrizable subset of a locally convex space such that for each $\mu \in \mathcal{K}$ there exists a unique measure concentrated on the extremal points of $\mathcal{K}$ which represents μ. Since the extreme points of the invariant measures are the ergodic measures, this is a restatement of the Ergodic Decomposition Theorem.

We emphasize that for μ to belong to $\mathcal{M}(X,T)$, μ must be defined on all Borel subsets of X.

3. Construction sequences and their limits

Here is the general definition of a construction sequence and its limit. In this section we briefly explain and prove some general statements in earlier papers. This paper is concerned with a special case, the odometer based construction sequences.

Definition 9: A **construction sequence** in a finite alphabet Σ is a sequence of non-empty collections of words $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ with the properties that:

(1) $\mathcal{W}_0 = \Sigma$,
(2) all of the words in each $\mathcal{W}_n$ have the same length q_n and the collection $\mathcal{W}_n$ is uniquely readable (in the sense of the definition in Section 2),
(3) each $w \in \mathcal{W}_n$ occurs at least once as a subword of every $w' \in \mathcal{W}_{n+1}$,

(4) there is a summable sequence $\langle \epsilon_n : n \in \mathbb{N} \rangle$ of positive numbers such that for each n, every word $w \in \mathcal{W}_{n+1}$ can be uniquely parsed into segments[2]

$$u_0 w_0 u_1 w_1 \cdots w_l u_{l+1} \tag{1}$$

such that each $w_i \in \mathcal{W}_n$, $u_i \in \Sigma^{<q_n}$ and for this parsing

$$\frac{\sum_i |u_i|}{q_{n+1}} < \epsilon_{n+1}. \tag{2}$$

We call the elements of $\mathcal{W}_n$ n**-words**, and let $s_n = |\mathcal{W}_n|$.

Definition 10: Let $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ be a construction sequence in an alphabet Σ. The limit of $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ is defined to be the collection $\mathbb{K}$ of $x \in \Sigma^{\mathbb{Z}}$ such that for all finite intervals $I \subseteq \mathbb{Z}$ there is a $w \in \mathcal{W}_n$ and $J \subseteq [0, q_n - 1)$ for some n such that $x \restriction I = w \restriction J$. Suppose $x \in \mathbb{K}$ is such that for some $a_n \leq 0 < b_n$ and $x \restriction [a_n, b_n) \in \mathcal{W}_n$. Then $w = x \restriction [a_n, b_n)$ is the **principal** n**-subword** of x. We set $r_n(x) = |a_n|$, which is the position of $x(0)$ in w.

Definition 11: Let $\langle k_n : n \in \mathbb{N} \rangle$ be a coefficient sequence, A construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ is **odometer based** if and only if $\mathcal{W}_{n+1} \subseteq \mathcal{W}_n^{k_n}$. A symbolic system $\mathbb{K}$ is **odometer based** if it has a construction sequence that is odometer based. For an odometer based construction sequence and $n \geq 1$ we let $K_n = \prod_{m<n} k_m$.[3]

Informally: odometer based construction sequences are those built without spacers.

One way of defining certain elements of $\mathbb{K}$ is illustrated in the following Lemma that is stated in Remark 2.15 in [7] section 2.

LEMMA 12: *Let $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ be a construction sequence. Let $\langle w_n : n \geq k \rangle$ be a sequence of words with $w_n \in \mathcal{W}_n$. Let $\langle r_n : n \geq k \rangle$ be a sequence of natural numbers such that*

(1) *$r_n \in [0, q_n)$ and both $r_n \to \infty, q_n - r_n \to \infty$,*
(2) *the r_{n+1}^{th} letter in w_{n+1} is in the position of the r_n^{th}-letter of w_n.*

Then there is a unique $s \in \mathbb{K}$ such that for all $n \geq k$, $r_n(s) = r_n$ and the principal n-subword of s is w_n.

[2] We assume that $l \geq 1$.

[3] K_n will be equal to the q_n in Definition 9.

It follows immediately from Lemma 12 that if $w \in \mathcal{W}_n$ and an $r \in [0, q_n)$ there is an $x \in \mathbb{K}$ such that w is the principal n-subword of x and $r_n(w) = r$.

The next definition is [7, Definition 2.9].

Definition 13: Suppose that $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ is a construction sequence for a symbolic system $\mathbb{K}$. We define the set S to be the collection of $x \in \mathbb{K}$ such that there are sequences of natural numbers $\langle a_m : m \in \mathbb{N} \rangle$, $\langle b_m : m \in \mathbb{N} \rangle$ going to infinity such that for all m there is an $n, x \upharpoonright [-a_m, b_m) \in \mathcal{W}_n$.

For clarity we make the following stronger definition:

Definition 14: Define the set S' to be the collection of $x \in \mathbb{K}$ such that there are sequences $a_n, b_n \geq 0$ going to infinity such that for all $n, x \upharpoonright [-a_n, b_n) \in \mathcal{W}_n$.

Note that taking $k = 0$ in Lemma 12, the given s belongs to S'.

Let $s \in \mathbb{K}$ have principal n and $n + 1$-subwords. Then the following are easy to check:

$$r_{n+1} \geq r_n, \tag{3}$$

$$q_{n+1} - r_{n+1} \geq q_n - r_n. \tag{4}$$

Zig-zagging. We now describe a method for using Lemma 12 to build elements $s \in S'$ we will call **zig-zagging**. Suppose we have $\langle w_m : m \leq n \rangle$ and $\langle r_m : m \leq n \rangle$. The operations left zig and right zig each can be used to extend the construction 2-stages to $\langle w_m : m \leq n + 2 \rangle$ and $\langle r_m : m \leq n + 2 \rangle$. A zig-zag consists of a left zig followed by a right zig, or a right zig followed by a left zig. Thus a zig-zag extends the construction 4 stages—from n to $n + 4$.

Given the word $w_n \in \mathcal{W}_n$ and the location $r_n \in [0, q_n)$, let $w_{n+2} \in \mathcal{W}_{n+2}$. Then w_{n+2} contains two subwords w^l_{n+1} and w^r_{n+1} that belong to $\mathcal{W}_{n+1}$ with w^l_{n+1} occurring to the left of w^r_{n+1}. (More precisely, if we view w_{n+2} as a string of letters and the first letter of w^l_{n+1} is at location l_0 and the first letter of w^r_{n+1} is at l_1, then $l_0 < l_1$. We include the locations of the words w^l_{n+1}, w^r_{n+1} in w_{n+2} as part of the information in the superscripts l, r.)

We define the **left zig**. The right zig is defined analogously except that **right zig** uses w^r_{n+1}. Let w_{n+1} be the word w^l_{n+1}. Consider the left most occurrence of w_n in w^l_{n+1}. Then the r_n^{th} letter of this occurrence of w_n occurs in some location k in w^l_{n+1}. Let r_{n+1} be this k. The letter in w^l_{n+1} in position r_{n+1} occurs at some k^* in w_{n+2}. Let $r_{n+2} = k^*$.

Because the w^l_{n+1} occurs to the left of w^r_{n+1} we must have

$$q_{n+2} - r_{n+2} \geq q_{n+1}.$$

By equations (3) and (4), we have $r_{n+2} \geq r_{n+1} \geq r_n$. Thus, if we are doing an inductive construction using Lemma 12 the left zig extends it from n to $n+2$ and we have:

$$r_{n+2} \geq r_{n+1},$$
$$q_{n+2} - r_{n+2} \geq q_{n+1}.$$

We describe a **left-right zigzag**. We are given a word w_n and a location r_n. Choose $w_{n+2} \in \mathcal{W}_{n+2}$ and do a left zig. This gives a w_{n+1} and locations r_{n+1}, r_{n+2}. Now choose $w_{n+4} \in \mathcal{W}_{n+4}$ and do a right zig. This gives w_{n+3} and locations r_{n+3}, r_{n+4}.

The words $w_{n+1}, w_{n+2}, w_{n+3}, w_{n+4}$ and locations $r_{n+1}, r_{n+2}, r_{n+3}, r_{n+4}$ satisfy the conditions of Lemma 12, and the following inequalities:

$$r_{n+4} \geq q_{n+3}, \tag{5}$$
$$q_{n+2} - r_{n+2} \geq q_{n+1}. \tag{6}$$

In particular, both the locations and the distances from the ends of the words strictly grow in these four steps.

The **right-left zigzag** is done the same way except that moving from n to $n+2$ is a right zig and going from $n+2$ to $n+4$ is a left zig. The analogous equations to (5) and (6) hold with the roles of $n+2$ and $n+4$ swapped and $n+1$ and $n+3$ swapped.

Example 15: Fix a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ and let $w_0 \in \mathcal{W}_0$. Let $f \in \{0,1\}^{\mathbb{N}}$. Build an $s \in \mathbb{K}$ by induction on n. The construction inductively chooses words $\langle w_m : m \leq 4n \rangle$ and locations $\langle r_m : m \leq 4n \rangle$. To pass from stage $4n$ to stage $4(n+1)$ one does a left-right zig-zag if $f(n) = 0$ and a right-left zig-zag if $f(n) = 1$.

The result is a sequence of words $\langle w_n : n \in \mathbb{N} \rangle$ and locations $\langle r_n : n \in \mathbb{N} \rangle$ that satisfy the hypotheses of Lemma 12, and hence determine an $s \in \mathbb{K}$. Equations (5) and (6) show that $s \in S'$.

PROPOSITION 16: *Let S, S' be defined as in Definitions 13 and 14. Then*

(1) *S is a shift invariant dense $\mathcal{G}_\delta$ subset of $\mathbb{K}$.*

(2) *S' contains perfect set.*

Proof. We first show that S is dense. Let $U = \langle u \rangle_k$ be a non-empty basic open set in $\mathbb{K}$. By making u longer and thus shrinking U we can assume that U sits on an interval $[c, d)$ with $c \leq 0 < d$. Applying Definition 13, we can find an n, $w \in \mathcal{W}_n$, $a \leq c < d \leq b$ and an $x \in \mathbb{K}$ such that $x \upharpoonright [a, b) = w$, $x \upharpoonright [c, d) = u$. Build a sequence of words $\langle w_m : n \leq m \rangle$ with $w_n = w$, $r_n = -a$ satisfying the hypothesis of Lemma 12. If s is the limit of these words given by Lemma 12, then $s \in U \cap S$.

To see S is $\mathcal{G}_\delta$, for each $b > 0$,

$$S_b = \{x \in \mathbb{K} : \text{for some } a_m, b_m \text{ with } a_m < -b \text{ and } b < b_m,\ x \upharpoonright [-a_m, b_m) \in \mathcal{W}_n\}$$

is an open set. Hence

$$S = \bigcap_{b \in \mathbb{N}} S_b$$

is a $\mathcal{G}_\delta$ set.

To see that S' contains a perfect set, we use Example 15. For each $f \in \{0,1\}^{\mathbb{N}}$ the example builds an $s(f) \in S'$. If $f \neq g$ and $f(m) \neq g(m)$ then

$$r_m(s(f)) \neq r_m(s(g)).$$

Moreover, the map $f \mapsto s(f)$ is continuous. Hence the $\{s(f) : f \in \{0,1\}^{\mathbb{N}}\}$ is a compact perfect set. ∎

Remark 17: Suppose that $s = s(f)$ is constructed in the manner of the proof of Proposition 16 and f is not eventually constant. Then for infinitely many n, the principal n-subword of s has an n-subword occurring to the left of it in the principal $n+1$ subword of s and similarly for words to the right of the principal n-subword.

The upshot of this discussion is the following:

Proposition 18: *The set $\mathbb{K}$ is a non-empty, closed, shift invariant, topologically transitive subshift of $\Sigma^{\mathbb{Z}}$.*

Proof. $\mathbb{K}$ is clearly non-empty since it contains the set S' defined in Proposition 16.

We first show that $\mathbb{K}$ is shift invariant. Let $y \in \mathbb{K}$ and consider $\mathbf{Sh}^n(y)$. For I a finite interval, let $u = \mathbf{Sh}^n(y) \upharpoonright I$. Then $u = y \upharpoonright \mathbf{Sh}^{-n}(I)$. Hence for some n, $J \subseteq [0, q_n)$, $w \in \mathcal{W}_n, u = w \upharpoonright J$. But then $\mathbf{Sh}^n(y) \upharpoonright I = w \upharpoonright J$. Hence $\mathbf{Sh}^n(y) \in \mathbb{K}$.

We next show that $\mathbb{K}$ is closed. Suppose that $\langle y_n : n \in \mathbb{N}\rangle \subseteq \mathbb{K}$ and $y_n \to_n y$. Let $I \subseteq \mathbb{Z}$ be a finite interval. Then for all large enough n, $y_n \upharpoonright I = y \upharpoonright I$. Hence there is an n and a $w \in \mathcal{W}_n$ such that $y \upharpoonright I$ is a subsequence of w. Hence $y \in \mathbb{K}$.

Finally we show that if $s = s(f)$ where f is not eventually constant (as in Remark 17), then both its forward and backwards orbits are dense in $\mathbb{K}$. We do the forward orbit; the argument for the backwards orbit is the same reversing "left" and "right". Let $x \in \mathbb{K}$ and O be an open set containing x. Then we can choose a basic open interval $\langle w\rangle_k \subseteq O$ with $x \in \langle w\rangle_k$. From the defining property of $\mathbb{K}$ we can find a word $w_m \in \mathcal{W}_m$ and a location k^* such that $\langle w_m\rangle_{k^*} \subseteq O$. Let $n > m+1$ be such that $f(n) = 0$. Then s is built between $4n$ and $4(n+1)$ by a left-right zig-zag. Hence the $4n+2$ principal subword of s contains an $n+1$-word $u \in \mathcal{W}_{n+1}$ to the right of its principal $n+1$-subword.

By clause 3 of the definition of a construction sequence (Definition 9), w_m occurs in u. Hence after finitely many shifts, the principal subword of $\mathbf{Sh}^r(s)$ is w_m and $r_m(\mathbf{Sh}^r(s)) = k^*$. Hence

$$\begin{aligned}\mathbf{Sh}^r(s) &\in \langle w_m\rangle_{k^*}\\ &\subseteq \langle w\rangle_k\\ &\subseteq O. \qquad ∎\end{aligned}$$

We now give an "external" characterization of $\mathbb{K}$:

Proposition 19: *The system $\mathbb{K}$ is the smallest closed, shift-invariant set that intersects all intervals of the form $\langle w\rangle_0$ for $w \in \mathcal{W}_n$.*

Proof. Let C be the intersection of all closed shift-invariant sets that intersect all intervals of the form $\langle w\rangle_0$. It suffices to show that $\mathbb{K} \subseteq C$.

Let X be an arbitrary closed, shift-invariant set intersecting the appropriate intervals. We first show that $X \cap \mathbb{K} \neq \emptyset$.

Let $s \in S' \subseteq \mathbb{K}$. Let $r_n(s) = a_n$ and $w_n = s \upharpoonright [a_n, b_n)$ be its principal subword. By the assumptions on X, there is a $y_n \in X$ such that $y_n \in [w_n]_{-a_n}$. Then the sequence $\langle y_n : n \in \mathbb{N}\rangle$ converges to s since $s \in S'$. So $s \in X$. If we take $s = s(f)$ for an f that is not eventually constant, then by Proposition 18, the orbit of s is dense in $\mathbb{K}$. Since X is closed, $\mathbb{K} \subseteq X$. ∎

Uniform construction sequences. Not every symbolic shift in a finite alphabet can be built as a limit of a construction sequence, however this method directly codes cut-and-stack constructions of transformations on probability spaces.

In this section we use Definition 9. In particular that every $w \in \mathcal{W}_{n+1}$ can be written uniquely in the form $u_0 w_0 u_1 w_1 \dots w_l u_{l+1}$ with the $w_i \in \mathcal{W}_n$ and $u_i \in \Sigma^{<q_n}$.

Definition 20: Let $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ be a construction sequence. Then $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ is **uniform** if there are $\langle d_n : n \in \mathbb{N} \rangle$, where $d_n : \mathcal{W}_n \to (0,1)$ and a sequence $\langle \epsilon_n : n \in \mathbb{N} \rangle$ going to zero such that for each n all words $w \in \mathcal{W}_n$ and $w' \in \mathcal{W}_{n+1}$ if $f(w, w')$ is the number of i such that $w = w_i$

$$\sum_{w \in \mathcal{W}_n} \left| \frac{f(w,w')}{q_{n+1}/q_n} - d_n(w) \right| < \epsilon_{n+1}. \tag{7}$$

If $f(w, w')$ is a constant (depending on n) for all $w \in \mathcal{W}_n, w' \in \mathcal{W}_{n+1}$ we can take

$$d_n(w) = \frac{f(w,w')}{q_{n+1}/q_n}$$

and satisfy Definition 20. In this case we call the construction sequence and $\mathbb{K}$ **strongly uniform**.

Lemma 11 of [8] shows that uniform construction sequences are uniquely ergodic.

4. Odometer based construction sequences

In this section we prove:

THEOREM 21: *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system with finite entropy. Then X has an odometer factor if and only if X is measure isomorphic to a topologically minimal odometer based symbolic system.*

Proposition 19 of [6] shows that the limit of an odometer based construction sequence is topologically minimal. This is not true for general construction sequences. For example, there are circular systems with singleton orbits. (See [7, Definition 3.17].)

We use the following observation when working with odometer based sequences and systems. Viewing the odometer based systems as isomorphic to cut-and-stack constructions with no spacers, it is immediate. However, we give a symbolic proof.

LEMMA 22: *Let $\mathbb{K}$ be an odometer based system with construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$. Then for all $s \in \mathbb{K}$ and all n, s has a principal n-subword. Moreover, $\mathbb{K} = S'$.*

Proof. Fix an n and let $a = -q_n$ and $b = q_n$. Then from the definition of $\mathbb{K}$, there is an $m > n$, a word $w \in \mathcal{W}_m$ and an interval $J \subseteq [0, q_m - 1)$ such that $s \restriction [-a.b) = w \restriction J$. Because the system is odometer based w_m can be uniquely written as a concatenation of a sequence of subwords from $\mathcal{W}_n$.[4] The interval J can be extended to an interval J' by adding at most q_n letters so that $w \restriction J'$ is a concatenation of words from $\mathcal{W}_n$, say $w \restriction J' = w_0 w_1 \dots w_k$. The $(q_n + 1)^{st}$ element of J must lie inside one of the w_i's. Let a_n be the position of w_i. Then the principal subword of s is w_i and if $x \restriction [-a_n, b_n) = w_i$, then a_n, b_n go to ∞. ∎

The name odometer based system is motivated by the following proposition:

PROPOSITION 23: *Suppose that $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ is an odometer based construction sequence for a symbolic system $\mathbb{K}$. Let K_n be the length of the words in $\mathcal{W}_n$, $k_0 = K_1$ and for $n > 0$, $k_n = K_{n+1}/K_n$. Then the odometer $\mathfrak{O}$ determined by $\langle k_n : n \in \mathbb{N}\rangle$ is canonically a factor of $\mathbb{K}$.*

Proof. Let $s \in \mathbb{K}$. By the unique readability, for each n, s can be uniquely parsed into a bi-infinite sequence of n words. For each n, there is an c_n such that the principal n-block of s is the c_n^{th} n-word in the principal $n+1$-block of s.

Define a map $\phi : \mathbb{K} \to \prod_n \mathbb{Z}/k_n\mathbb{Z}$ by setting

$$\phi(s) = \langle c_n : n \in \mathbb{N}\rangle.$$

It is easy to check that $\phi(\mathbf{Sh}(s)) = \mathcal{O}(\phi(s))$ (where "**Sh**" stands for the shift map). ∎

Remark 24: We note two things:

(1) Hypothesis 2 of Lemma 12 simplifies in the case of odometer based systems to being the requirement that $r_{n+1} \equiv_{K_n} r_n$.

(2) Our notation is inconsistent in that for general construction sequences we use q_n for the lengths of the n-words, but for odometer based systems we use K_n.

[4] This is shown by induction on $m \geq n + 1$.

4.1. Odometers are not topological subshifts. Theorem 21 says that all ergodic measure preserving transformations with a non-trivial odometer factor are measure theoretically isomorphic to an odometer based symbolic system. In contrast, it is well known that as topological dynamical systems, odometers are not homeomorphic to symbolic shifts. For background we give a very brief proof of this fact.

Definition 25: Let (X, d) be a metric space. A map $T : X \to X$ is **expansive** if there is an $\epsilon > 0$ so that for all $x \neq y$ in X there is an n, such that

$$d(T^n x, T^n y) \geq \epsilon.$$

The following is easy to verify:

Proposition 26: *Let (X, d) be a compact metric space and $T : X \to X$.*

(1) *If T is an isometry, then T is not expansive unless X is finite.*
(2) *If $X \subseteq \Sigma^{\mathbb{Z}}$ is a compact subshift, and T is the shift map, then T is expansive.*

Proof. The first proposition is trivial. To see the second, note that we can assume Σ is finite. Let c be the minimum distance between cylinder sets $\langle i \rangle$ and $\langle j \rangle$ based at 0. Then if $x \neq y$, we can find an $n, x(n) \neq y(n)$. It follows that $d(T^n x, T^n y) \geq c$. ∎

Because the odometer is a rotation on a compact metric group it is an isometry, hence is not expansive.

For the reader wanting concrete details, in view of Proposition 26, to see that an odometer cannot be presented as a topological subshift it suffices to show that, viewed as metric systems, odometer transformations are isometries. Let $\boldsymbol{O} = \prod_0^\infty \mathbb{Z}/k_n\mathbb{Z}$ be an odometer and T be the odometer map $\mathcal{O}$.

For $x, y \in \boldsymbol{O}$, define $\Delta(x, y)$ to be the least n such that $x(n) \neq y(n)$ and

$$d(x, y) = \frac{1}{2^{\Delta(x,y)}}.$$

Then d is a complete metric yielding the product topology on $\boldsymbol{O}$ and is invariant under $\mathcal{O}$. If $\boldsymbol{O}$ were homeomorphic to a subshift $X \subseteq \Sigma^{\mathbb{Z}}$ then d could be copied over to a shift invariant metric d_X on X. But then d_X must be expansive contradicting the first item of Proposition 26.

Thus, by Proposition 26, it follows that the odometer is not isometric to a subshift of $\Sigma^{\mathbb{Z}}$ for any finite Σ.

4.2. Presenting the odometer. In this section we show that odometers are measure theoretically isomorphic to odometer based systems. In Section 6 we will be concerned with the rates of descent of the measures of the basic open intervals $\langle w\rangle$ determined by $w \in \mathcal{W}_n$ as n increases. Lemma 42, a generalization of the next lemma, is a key tool for showing that the small word property defined in Section 6 can be achieved for odometer based systems.

Lemma 27: *If $\mathfrak{O}$ is an odometer determined by $\langle k_n : n \in \mathbb{N}\rangle$ with $k_n \geq 2$, then there is a uniform odometer based construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ such that the associated symbolic system $\mathbb{K}$ is topologically minimal, uniquely ergodic and measure theoretically conjugate to $\mathfrak{O}$.*

Proof. By Proposition 8, we can assume that $\sum 1/k_n < \infty$. We define an odometer based construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ such that each $\mathcal{W}_n = \{a_n, b_n\}$ has exactly two words in it.

- Let $\Sigma = \{a, b\}$ and $\mathcal{W}_0 = \Sigma$.
- Suppose that we are given $\mathcal{W}_n = \{a_n, b_n\}$. Let $\mathcal{W}_{n+1} = \{a_{n+1}, b_{n+1}\}$ with $a_{n+1}, b_{n+1} \in \mathcal{W}_n^{k_n}$ where:

$$a_{n+1} = a_n a_n a_n b_n b_n b_n a_n b_n a_n b_n \cdots x,$$
$$b_{n+1} = b_n b_n b_n a_n a_n a_n a_n b_n a_n b_n \cdots x,$$

  where x is either a_n or b_n, depending on whether k_n is even or odd.

The number of alternations of a_n and b_n is determined by k_n, so the second item is well-defined.

It is easy to verify inductively that the a_n's and b_n's are uniquely readable (look for patterns of the form $a_n a_n a_n$ and $b_n b_n b_n$) and that $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ is uniform. Let $\mathbb{K}$ be the associated symbolic system. Then $\mathbb{K}$ is uniquely ergodic, with an invariant measure μ.

Let

$$G = \{x \in \boldsymbol{O} : \text{for all large enough } n, 10 \leq x(n) \leq k_n - 10\}.$$

Since $\sum 1/k_n < \infty$, the Borel–Cantelli Lemma implies that G has measure one for $\mathfrak{O}$. Further, G is invariant under $\mathcal{O}^{\pm 1}$.

We define $\psi : G \to \mathbb{K}$. By Lemma 12, we can determine $\psi(x)$ by defining a suitable sequence $\langle r_n : n \geq k\rangle$ and $\langle w_n : n \geq k\rangle$.

Let $x \in G$ and suppose that for all $n \geq k$, both $x(n) \geq 10$ and $x(n) \leq k_n - 10$. For $n \geq k$, let

$$r_n = x(0) + x(1)K_1 + x(2)K_2 + \cdots + x(n)K_n.$$

Since $x(n) \geq 10$, either for all $n+1$-words $w \in \mathcal{W}_{n+1}$, the $x(n)^{th}$ n-subword in w is a_n, or for all $n+1$-words $w \in \mathcal{W}_{n+1}$, the $x(n)^{th}$ n-subword in w is b_n. Let w_n be either a_n or b_n accordingly.

Let $\psi(x)$ be the element s of $\mathbb{K}$ determined by $\langle r_n : n \geq k\rangle$ and $\langle w_n : n \geq k\rangle$, as in Lemma 12. From the definition of G,

(1) $\psi(x)$ is well-defined and $\psi(x) \in S$,
(2) ψ is one-to-one and continuous,
(3) $\psi(\mathcal{O}^{\pm 1}(x)) = \mathbf{Sh}^{\pm 1}(\psi(x))$,
(4) if $\phi : \mathbb{K} \to \mathfrak{O}$ is the factor map given in Proposition 23, then $\phi \circ \psi$ is the identity map.

Because ψ is one-to-one, continuous and G is Borel, the image of G under ψ is a Borel set.

If ν is the measure on $\boldsymbol{O}$ giving the odometer system, then ψ induces a shift-invariant measure $\nu^* = \psi^*\nu$ on the Borel subsets of $\mathbb{K}$. Since $\mathbb{K}$ is uniquely ergodic, $\nu^* = \mu$. Hence ψ is a measure isomorphism between $\mathfrak{O}$ and $\mathbb{K}$. ∎

Remark 28: In Proposition 32 we use properties 1–4 of the proof and the fact that $\psi[G]$ is Borel.

4.3. The plan. In this section we explain the idea of the proof of Theorem 21; the details will follow in the next section. To show that a given transformation with an odometer factor is isomorphic to a symbolic system built from an odometer based construction sequence, we build a generating partition so that the names of points on the bases of the n-towers in Definition 5 form an odometer based construction sequence.

Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving system with an odometer factor $\mathfrak{O}$. By Lemma 27, $\mathfrak{O}$ is isomorphic to an odometer based system in the alphabet $\Sigma = \{a, b\}$. Call the resulting construction sequence $\langle \mathcal{W}_n^{\mathcal{O}} : n \in \mathbb{N}\rangle$. If $\mathbb{K}$ is the symbolic system associated with this construction sequence we have

$$X \xrightarrow[\pi]{} \mathcal{O} \xrightarrow[\psi]{} \mathbb{K}$$

where π and ψ are defined on full measure sets.

Let $\mathcal{Q} = \{Q_0, Q_1\}$ be the partition of X corresponding to the basic open intervals $\langle a\rangle, \langle b\rangle$ in $\mathbb{K}$ (so $Q_0 = (\psi \circ \pi)^{-1}\langle a\rangle$ and $Q_1 = (\psi \circ \pi)^{-1}\langle b\rangle$). Then $\mathcal{Q}$ generates the subalgebra of $\mathcal{B}$ corresponding to the factor $\mathfrak{O}$.

Suppose that $C \subseteq X$ is a set of positive measure. Let $T_C : C \to C$ be the induced map: $T_C(c) = d$ if and only if for the least $k > 0$, with $T^k(c) \in C$ one has $T^k(c) = d$. Suppose that $\mathcal{P}_0 = \{P_1, P_2, \ldots, P_a\}$ is a generator for T_C, where $a \in \mathbb{N}$. Let $D = X \setminus C$ and $\mathcal{P} = \mathcal{P}_0 \cup \{D\}$. Then for $x \in X$, the $\mathcal{P}$-name of x uniquely determines x, and thus $\mathcal{P}$ is a generator for X.

For a typical x, the combined $\mathcal{P}_0, \mathcal{Q}$-name of x can be visualized as in Figure 3. The elements of $\mathcal{Q}$ parse the x-orbit into n-words which measure the duration an orbit stays in D, while the elements of $\mathcal{P}_0$ determine the orbit of x inside C. Since $\mathcal{P}_0$ and $\mathcal{Q}$ determine x, in building an odometer based symbolic representation of $(X, \mathcal{B}, \mu, T)$, one has complete freedom to fill in symbols in the parts of the x-orbit that lie in D. This allows our word construction to satisfy the definition of an odometer based construction sequence.

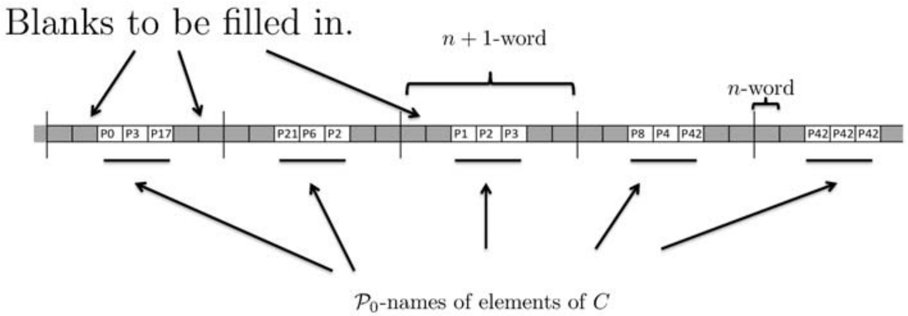

Figure 3. The $\mathcal{P}_0$-name of x punctuated by the odometer.

In terms of partitions, this can be restated as saying that we can modify the atoms of the partition $\mathcal{P}_0$ by adding elements of D in any arbitrary way, as long as the restriction of each atom of $\mathcal{P}_0$ to C remains the same. If $\mathcal{P}_0' = \{P_1', P_2', \ldots, P_a'\}$ is such a modification of $\mathcal{P}_0$, then any partition refining $\mathcal{P}_0'$ and $\mathcal{Q}$ still forms a generator for T. Hence, as in Remark 4, the symbolic system consisting of pairs $(s_{\mathcal{P}_0'}, s_{\mathcal{Q}})$ of $\mathcal{P}_0'$ and $\mathcal{Q}$-names is isomorphic to $(X, \mathcal{B}, \mu, T)$.

Of course, Figure 3 is an over-simplification of the possibilities for the orbit: it assumes that the set C fits coherently with the odometer factor. In other words, C must be chosen to be measurable with respect to the sub-σ-algebra of $\mathcal{B}$ generated by the odometer factor.

4.4. THE PROOF. Suppose that $(X, \mathcal{B}, \mu, T)$ has entropy less than $\frac{1}{2}\log c$ for some integer constant $c \geq 2$. By Proposition 8, we can assume that

$$K_1 = k_0 > 10, \quad K_n = \prod_{i<n} k_i$$
$$\text{and} \quad k_n > 4c^{K_n}10^{n+1}.$$

Let $B_0, B_1, \ldots$ be the bases of the n-towers in X associated with $\mathfrak{O}$ by the factor map π; in the notation of Definition 5, $B_n = B_n^0$ and $\mathcal{T}_n$ is the tower with base B_n. Let $d_n = 4K_{n-1}c^{K_{n-1}}$ and define

$$D_n = \bigcup_{0 \leq i \leq d_n} B_n^i$$

and

$$D = \bigcup_1^\infty D_n.$$

Thus D_n consists of the first d_n levels of the n-tower. Since all of the levels of the tower have the same measure, the measure of D_n is

$$\begin{aligned}\frac{d_n}{K_n} &= \frac{4K_{n-1}c^{K_{n-1}}}{K_n} \\ &= \frac{4K_{n-1}c^{K_{n-1}}}{K_{n-1}k_{n-1}} \\ &< \frac{4K_{n-1}c^{K_{n-1}}}{K_{n-1}4c^{K_{n-1}}10^n} \\ &= 10^{-n}.\end{aligned}$$

Set $C = X \setminus D$. Clearly C is measurable with respect to the odometer factor, since it is a union of levels of the odometer towers. Moreover, $\mu(C) > 3/4$, and hence the entropy of T_C is less than $(2/3)\log c$. By Krieger's Theorem [10] there is a generating partition $\mathcal{P}_0 = \{P_1, P_2, \ldots, P_c\}$ for T_C that has c elements.

Figure 4 is a graphical representation of $\mathcal{T}_n$ showing:

(1) C as whitespace.

(2) D_n lightly shaded as an initial segment of the levels of $\mathcal{T}_n$.

(3) The sets D_m for $m < n$ are initial segments of earlier $\mathcal{T}_m$ and hence get stacked as bands across $\mathcal{T}_n$. They are given an intermediate shading in Figure 4.

(4) Because each D_m is an initial segment of $\mathcal{T}_m$, at the previous stage the points in D_m have to be in the leftmost columns of $\mathcal{T}_{m-1}$. Moreover, for $m < m'$, K_m divides $d_{m'}$. Thus $D_{m'}$ is made up of whole columns of $\mathcal{T}_m$. Consequently $\bigcup_{m>n} D_m$ forms a contiguous rectangle on the left side of $\mathcal{T}_n$. This region is indicated by the darkest shading.

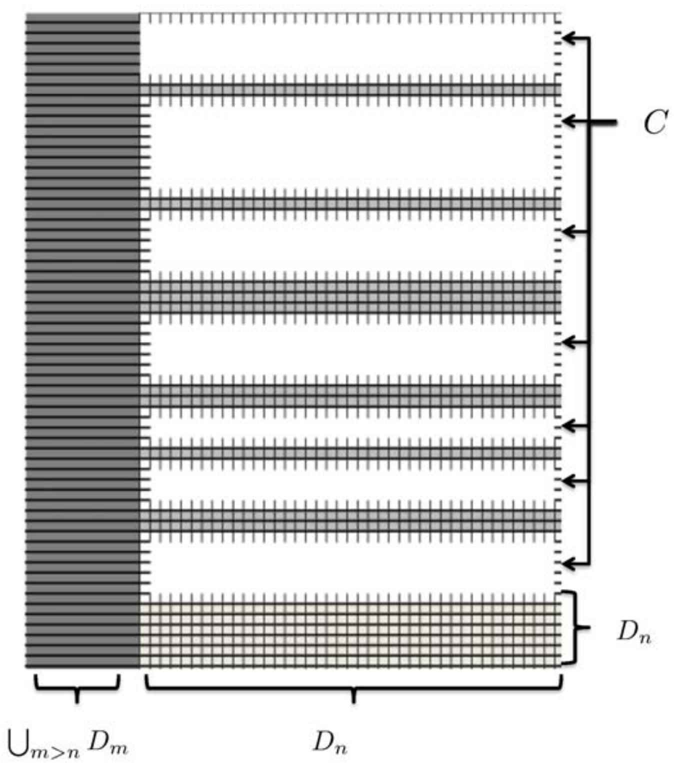

Figure 4. The n^{th} stage of the construction. The shaded horizontal bands are elements of D_m for $m < n$.

We construct $\mathcal{P}_0'$ in the manner described in Section 4.3: we add points from D to each P_j to get a final partition $\mathcal{P}_0' = \{P_1', P_2', \dots, P_a'\}$. For each i we build an increasing sequence $\langle P_i(n) : n \in \mathbb{N}\rangle$ and let

$$P_i' = \bigcup_n P_i(n).$$

The construction sequence will use the alphabet $\Sigma = \{P_1', P_2' \dots, P_a'\} \times \mathcal{Q}$, where $\mathcal{Q} = \{Q_0, Q_1\}$ is the partition generating the odometer; $\mathcal{W}_0 = \Sigma$ and $\mathcal{W}_n$ will consist of the Σ-names of points that occur in the base of $\mathcal{T}_n$. Thus the construction is completely determined by the manner we add points to the P_i to get $P_i(n)$.

The words must satisfy Definition 9. Clause 1 is automatic. Clause 2 holds because all words have length equal to the height of $\mathcal{T}_n$. Unique readability is immediate since the odometer based presentation of $\mathfrak{O}$ uses uniquely readable words in the language $\mathcal{Q}$. Clause 4 is vacuous since we have no spacers u_i occurring anywhere in the words: elements of $\mathcal{W}_{n+1}$ are simply concatenations of words from $\mathcal{W}_n$.

Proposition 19 of [6] shows that the limit of any odometer based construction sequence is minimal. In this construction we do more: each word in $\mathcal{W}_n$ occurs at least twice in each word in $\mathcal{W}_{n+1}$, a property which is stronger than clause 3. We satisfy this by "painting" the words from $\mathcal{W}_n$ onto D_{n+1}.

Let $P_i'(n)$ be the collection of points in P_i' at stage n, and $P_i'(0) = P_i$. Inductively we will assume that at stage n:

(1) $\bigcup_{n<m} D_m \cap P_i'(n) = \emptyset$ for all i, and

(2) $(\mathcal{T}_n \setminus \bigcup_{n<m} D_n) \subseteq \bigcup_i P_i'(n)$.

For $n = 1$, we consider $D_1 \setminus \bigcup_{m>1} D_m$. At stage 1 the minimality requirement says that each pair $(P_i'(0), j)$ for $1 \le i \le c$ and $j \in \{Q_0, Q_1\}$ occurs at least twice. Each of Q_0, Q_1 occur equally often in the $\mathcal{Q}$-names of the first d_1 letters of each $\mathcal{Q}$-name and $d_1 = 4c$. Hence it is possible to assign the levels in $D_1 \setminus \bigcup_{m>1} D_m$ to $\{P_1'(1), \dots, P_c'(1)\}$ in such a way that each element of the alphabet Σ occurs at least twice. Concretely this means that for each i, j there are two levels of $S_1 \setminus \bigcup_{m>1} D_m$ that belong to Q_j and are put into $P_i'(1)$.

To pass from n to $n+1$ in the construction, we know inductively that no elements of D_{n+1} have been assigned to any P_i' at earlier stages. Moreover, $\mathcal{W}_n$ consists of the Σ-names of the words in $B_0 \setminus \bigcup_{m>n} D_m$, where B_0 is the base of $\mathcal{T}_n$. There are at most $2c^{K_n}$ such words in the language Σ. Each such word has length K_n.

Since $d_{n+1} = 4K_n c^{K_n}$ there are ample levels in D_{n+1} that each level can be added to some $P'_i(n+1)$ in a manner that each word in $\mathcal{W}_n$ occurs at least twice as a Σ-name of an element in the first d_{n+1} levels of $\mathcal{T}_{n+1}$. ∎

Remark 29: The construction in the proof of Theorem 21 used a particular presentation of $\mathfrak{O}$ as an odometer based system in a language $\mathcal{Q} = \{a, b\}$ to build a language $\Sigma = \{P'_1, P'_2, \ldots, P'_c\} \times \mathcal{Q}$. If we were given another odometer based presentation $\langle \mathcal{W}_n^{\mathfrak{O}} : n \in \mathbb{N}\rangle$ of $\mathfrak{O}$ in a different finite language with letters $\{a_1, \ldots, a_k\}$ we could take $\Sigma = \{P'_1, P'_2, \ldots, P'_c\} \times \{a_1, \ldots, a_k\}$ and repeat the same construction over this presentation. We will call this the odometer based presentation of X **built over** $\langle \mathcal{W}_n^{\mathfrak{O}} : n \in \mathbb{N}\rangle$.

5. Toeplitz systems

In this section we use a result of Downarowicz ([3]) to show that every compact metrizable Choquet simplex is affinely homeomorphic to the simplex of invariant measures of an odometer based system. Williams showed that the orbit closure of every Toeplitz sequence in a finite language Σ is a minimal symbolic shift $\mathbb{L}$ with a continuous map to an odometer factor $\mathfrak{O}$. If $\pi^{\mathfrak{O}} : \mathbb{L} \to \mathfrak{O}$ is this factor map, it would be tempting to try to argue that the words occurring on $\pi^{\mathfrak{O}}$-pullbacks of the levels of the n-towers form an odometer based construction sequence. However, we don't know this in general; in particular, we don't know that the words constructed this way are uniquely readable.

To make the words uniquely readable we need to add information without changing the collection of invariant measures. To do this we introduce the notion of an augmented symbolic system.

Definition 30: Let (Z, σ, S) and (X, τ, T) be minimal compact topological systems and $\pi : Z \to X$ be a continuous factor map. Then (π, Z) is an **augmentation** of X if there is an S-invariant Borel set $A \subseteq Z$ such that if

$$L = \{x : \text{there is exactly one } y \in A \text{ with } \pi(y) = x\},$$

then for all T-invariant μ on X,

$$\mu(L) = 1.$$

We use this idea as follows:

Proposition 31: *Suppose that (π, Z) is an augmentation of X. Then there is a canonical affine homeomorphism from $\mathcal{M}(Z,S)$ to $\mathcal{M}(X,T)$.*

Proof. The map $\phi : M(Z,S) \to M(X,T)$ given by

$$\phi(\mu) = \pi^*(\mu)$$

is a continuous affine map.

If ν is an invariant measure on X, then the pullback of ν by π is an invariant measure ν' on a sub-σ-algebra of the Borel subsets of Z. Standard arguments show that ν' can be extended to an S-invariant measure μ on Z such that $\pi^*(\mu) = \nu$, hence ϕ is surjective.

We claim that A has measure one for all invariant measures μ on Z. Otherwise suppose that $\mu(A) < 1$. Consider $\nu = \phi(\mu)$. Let $B \subseteq L$ be ν-measurable and such that $\nu(B) = 1$. Then $\pi^{-1}(B) \subseteq A$ and has μ-measure one since $\nu = \pi^*(\mu)$.

We need to see that ϕ is one-to-one. Clearly ϕ takes ergodic measures to ergodic measures. Suppose that $\mu \neq \nu$ are ergodic measures on Z. Then there are disjoint invariant sets $B, C \subseteq Z$ such that $\mu(B) = \nu(C) = 1$. Let B' and C' be the images of B and C under the map π. Then, from the properties of L,

$$\pi^{-1}(L \cap B') \subseteq B \quad \text{and} \quad \pi^{-1}(L \cap C') \subseteq C.$$

Hence $L \cap B'$ and $L \cap C'$ are disjoint and have measure one for $\phi(\mu)$ and $\phi(\nu)$ respectively. Hence $\phi(\mu)$ and $\phi(\nu)$ are distinct.

Since ϕ is affine, continuous and one-to-one on the ergodic measures, it is a one-to-one map. Finally, since the set of invariant measures on Z is a compact space ϕ is a homeomorphism. ∎

To prove Proposition 32, we use:

Theorem (Downarowicz, [3, Theorem 5]): *For every compact metric Choquet simplex K there is a dyadic Toeplitz flow whose set of invariant measures is affinely homeomorphic to K.*

Proposition 32: *Let $\mathbb{L}$ be the orbit closure of a Toeplitz sequence x, $\mathfrak{O}$ be its maximal odometer factor based on a sufficiently fast growing sequence $\langle k_n \rangle$ and $\mathbb{K}$ be the odometer based presentation of $\mathfrak{O}$ defined in Lemma 27. Then there is an odometer based system $\mathbb{L}^* \subseteq \mathbb{L} \times \mathbb{K}$ such that if $\pi : \mathbb{L}^* \to \mathbb{L}$ is the projection to the first coordinate, then $(\pi, \mathbb{L}^*)$ is an augmentation of $\mathbb{L}$.*

Thus, as an immediate consequence of Downarowicz' theorem and Propositions 31 and 32:

Corollary 33: *For every compact metrizable Choquet simplex K there is an odometer based symbolic shift $\mathbb{L}^*$ whose set of invariant measures is affinely homeomorphic to K.*

Proof of Proposition 32. We use the language of Williams [13]. Let x be a Toeplitz sequence in a finite language Σ. Let $\mathbb{L}$ be the orbit closure of x under the shift map and $\mathfrak{O}$ be the associated odometer system.

As in [13] we can choose a sequence $\langle K_n : n \in \mathbb{N}\rangle$ of essential periods for x. By choosing the K_n's to grow fast enough we can assume that

(a) $K_n | K_{n+1}$,
(b) $\bigcup_n \mathrm{Per}_{K_n}(x) = \mathbb{Z}$.

Choosing a further subsequence we can also assume that

(c) if $k \equiv 0 \pmod{K_n}$, then there is an $i \equiv 0 \pmod{K_n}$ with $i < K_{n+1}$ and $x \restriction [k, k+K_n) = x \restriction [i, i+K_n)$.

Given n_0, for large enough $n, x \restriction [0, K_{n_0})$ is a subset of the K_n-skeleton of x. Since the K_n-skeleton is K_n-periodic, every subword of the K_n-skeleton is repeated K_{n+1}/K_n times in $x \restriction [0, K_{n+1})$. Thus by again thinning the K_n's we can assume that:

(d) for each n and $i \equiv 0 \pmod{K_n}$ and each word $w \in \Sigma^{K_n}$ occurring as $x \restriction [i, i+K_n)$, w occurs at least twice in $x \restriction [0, K_{n+1})$.

Let $\mathfrak{O}$ be the odometer with coefficient sequence $\langle k_n : n \in \mathbb{N}\rangle$, where

$$k_n = K_{n+1}/K_n.$$

Let $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ be the odometer based construction sequence in the presentation of $\mathfrak{O}$ given in Lemma 27. Let w_n^0, w_n^1 be the two words in $\mathcal{W}_n$. We define an odometer based construction sequence by setting $\mathcal{V}_n$ to be the collection of words v in the alphabet $\Sigma \times \{a, b\}$ of the form

$$(x \restriction (i, i+K_n), w_n^j)$$

where $x \in \mathbb{L}$, $i < K_{n+1}, i \equiv 0 (\mathrm{mod}\ K_n)$ and $j \in \{0, 1\}$.

To see that this is an odometer based construction sequence we check Definitions 9 and 11.

Unique readability of the words $v \in \mathcal{V}_n$ follows immediately from the fact that the w_n^j are. The fact that each $v \in \mathcal{V}_n$ occurs at least twice as a subword of each $v' \in \mathcal{V}_{n+1}$ follows immediately from item (d) of the properties of the essential periods of x. From item (c) and the structure of the word construction from the Toeplitz sequence each word in $\mathcal{V}_{n+1}$ is a concatenation of words in $\mathcal{V}_n$.

By [13], there is a continuous factor map

$$\theta : \mathbb{L} \to \mathfrak{O}.$$

From the proof of Lemma 27 (see Remark 28) we see that there is an invariant Borel set $G \subset \mathfrak{O}$ of measure one and a one-to-one continuous map $\psi : G \to \mathbb{K}$. Let $\mathbb{L}^*$ be the limit of $\langle \mathcal{V}_n : n \in \mathbb{N}\rangle$. Note that for each $x \in \mathbb{L}$ there is an z such that $(x, z) \in \mathbb{L}^*$. Let

$$A = \{(y, \psi \circ \theta(y)) : \theta(y) \in G\} \subseteq \mathbb{L}^*.$$

Let μ be an invariant measure on $\mathbb{L}$. Then

$$\mu(\theta^{-1}(G)) = 1,$$

and for $y \in \theta^{-1}(G)$ there is a unique $z, (y, z) \in A$.

Let ρ be an invariant measure on $\mathbb{L}^*$. Let $\rho^{\mathbb{L}}$ be the $\mathbb{L}$ marginal. Then

$$\rho^{\mathbb{L}}(\theta^{-1}(G)) = 1.$$

If $y \in \theta^{-1}(G)$ and $(y, z) \in \mathbb{L}^*$, then $z = \psi \circ \theta(y)$. Hence $\mu(A) = 1$. ∎

Remark: The well-known Thue–Morse minimal system is an example of an odometer based system which is not Toeplitz.

6. The small word property and rates of descent

The applications of the representation theorem and Proposition 32 require that for all invariant measures on the limit system $\mathbb{K}$, the basic open intervals determined by words in $\mathcal{W}_{n+1}$ have measure much smaller than the measures of basic open intervals determined by words in $\mathcal{W}_n$. We show how to arrange this for odometer based systems by taking subsequences.

6.1. Empirical distributions and frequencies. In this section we introduce Empirical Distributions and the special case of Frequencies.

Suppose we are given a (general) construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$, and a word in $w' \in \mathcal{W}_{k+l}$ that can be uniquely written as a word

$$w' = u_0 w_0 u_1 w_1 \cdots w_J u_{J+1}$$

with $w_i \in \mathcal{W}_k$. We define the **empirical distribution**[5] of $\mathcal{W}_k$-words in w' by $\mathrm{EmpDist}_k(w')$. Formally:

$$\mathrm{EmpDist}_k(w')(w) = \frac{|\{0 \leq j \leq J : w_j = w\}|}{J+1}, \quad w \in \mathcal{W}_k.$$

Lemma 34: *Let $w \in \mathcal{W}_k$. If for all $w' \in \mathcal{W}_{k+1}$, $\eta_0 < \mathrm{EmpDist}(w')(w) < \eta_1$, and no $\mathcal{W}_k$-word occurs as subword of a spacer u_i, then for $k+l > k, w' \in \mathcal{W}_{k+l}$ we have $\eta_0 < \mathrm{EmpDist}(w')(w) < \eta_1$.*

Proof. We prove this in the case we use: odometer based construction sequences, and comment at the end how to give the general proof. If $w' \in \mathcal{W}_{k+l}$, then w' is a concatenation $w_0 w_1 \cdots w_{K_{k+l}/K_{k+1}-1}$ of K_{k+l}/K_{k+1} many words from $\mathcal{W}_{k+1}$. The number of occurrences $\mathrm{Occ}(w', w)$ of w in w' is the sum of the number of occurrences of w in the w_i's. We see that

$$\begin{aligned}\mathrm{EmpDist}_k(w')(w) &= \frac{\mathrm{Occ}(w', w)}{K_{k+l}/K_k} = \frac{\sum_i \mathrm{Occ}(w_i, w)}{K_{k+l}/K_k} \\ &= \frac{(K_{k+1}/K_k)\sum \mathrm{EmpDist}_k(w_i, w)}{K_{k+l}/K_k} \\ &= \frac{\sum \mathrm{EmpDist}_k(w_i, w)}{K_{k+l}/K_{k+1}}.\end{aligned}$$

Since the empirical distributions of $\mathcal{W}_k$ words in $\mathcal{W}_{k+l}$ words are the averages of the empirical distributions of the $\mathcal{W}_k$ words in $\mathcal{W}_{k+1}$ words, the lemma follows for odometer based construction sequences.

For general construction sequences the $\mathcal{W}_k$ words have spacers in them and the number of occurrences of $\mathcal{W}_n$-words in $\mathcal{W}_{n+1}$-words may vary (but only by a small amount). Nonetheless, a small variation of the argument just given shows that the empirical distributions of the k words in $k+l$-words is the weighted average of the distributions of k-words in $k+1$-words, and the lemma follows as before. ∎

[5] See [7].

Empirical distributions are related to measures via generic sequences. We summarize some results in [12] and [7].

Let μ be a shift invariant measure on a symbolic system $\mathbb{K}$ defined by a uniquely readable construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ in a finite language Σ. Assume that q_n is the length of the words in $\mathcal{W}_n$. By μ_m we will denote the discrete measure on the finite set Σ^m given by $\mu_m(u) = \mu(\langle u\rangle)$. By $\hat{\mu}_n(w)$ we will denote the discrete probability measure on $\mathcal{W}_n$ defined by

$$\hat{\mu}_n(w) = \frac{\mu_{q_n}(\langle w\rangle)}{\sum_{w'\in\mathcal{W}_n} \mu_{q_n}(\langle w'\rangle)}. \tag{8}$$

Thus $\hat{\mu}_n(w)$ is the relative measure of $\langle w\rangle_0$ among all $\langle w'\rangle_0, w' \in \mathcal{W}_n$. The denominator is a normalizing constant to account for spacers at stages $m > n$ and for the measures of sets $\langle w'\rangle_k$ where $0 < k < q_n$.

In an odometer based system, the normalizing denominator is the measure of $\bigcup_{w\in\mathcal{W}_n}\langle w\rangle_0$. This consists of all words in $\mathbb{K}$ with $r_n(s) = 0$. Because the length of the words in $\mathcal{W}_n$ is K_n, and the words are uniquely readable, $\mathbb{K}$ is the disjoint union of

$$\left\{\mathbf{Sh}^k\left(\bigcup_{w\in\mathcal{W}_n}\langle w\rangle_0\right) : 0 \le k < K_n\right\}.$$

Since μ is shift invariant for each k, $\mathbf{Sh}^k(\bigcup_{w\in\mathcal{W}_n}\langle w\rangle_0)$ has the same measure as $\bigcup_{w\in\mathcal{W}_n}\langle w\rangle_0$. Thus the denominator of equation (8) is exactly $1/K_n$.

Thus by the shift invariance of μ this is exactly $1/K_n$. Hence

$$\hat{\mu}_n(w) = K_n\mu(\langle w\rangle_0). \tag{9}$$

Definition 35: A sequence $\langle v_n \in \mathcal{W}_n : n \in \mathbb{N}\rangle$ is a **generic sequence of words** if and only if for all k and $\epsilon > 0$ there is an N such that for all $m, n > N$,

$$\|\mathrm{EmpDist}_k(v_m) - \mathrm{EmpDist}_k(v_n)\|_{\mathrm{var}} < \epsilon.$$

The sequence is generic for a measure μ if for all k

$$\lim_{n\to\infty} \|\mathrm{EmpDist}_k(v_n) - \hat{\mu}_k\|_{\mathrm{var}} = 0,$$

where $\| \ \|_{\mathrm{var}}$ is the variation norm on probability distributions.

It follows that if $\langle v_n : n \in \mathbb{N}\rangle$ is a generic sequence of words, then it is generic for a unique measure μ. Even though Definition 35 involves only the measures $\hat{\mu}_k$ it is easy to see (using the Ergodic Theorem) that for any $u \in \Sigma^k$, if $\langle v_n : n \in \mathbb{N}\rangle$ is generic then the density of the occurrences of u in the v_n will converge to $\mu(\langle u\rangle)$.

The following is Proposition 2.20 in [7]:

Proposition 36: *Let $\mathbb{K}$ be an ergodic symbolic system with construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ and measure μ. Then for any generic $s \in \mathbb{K}$ the sequence of principal subwords of s, $\langle w_n : n \in \mathbb{N}\rangle$, is generic for μ. In particular, generic sequences for μ exist.*

There is also a finitary notion of an **ergodic sequence**. For generic ergodic sequences $\hat{\mu}$ is defined to be the limit of the empirical distributions, and determines a shift invariant ergodic measure μ.

Thus empirical distributions capture the notion of ergodicity in a finitary way, and every generic point for an ergodic measure is a limit of empirical distributions along subwords.

Remark 37: In fact more than Lemma 34 is true (again, see the arguments in [7]). In Lemma 29, we can change $k+1$ to an arbitrary $k' > k$: Let $w \in \mathcal{W}_k$ and fix $k' > k$. If for all $w' \in \mathcal{W}_{k'}$, we have $\eta_0 < \mathrm{EmpDist}(w')(w) < \eta_1$, and no k-word occurs as subword of a spacer u_i in a word in $w' \in \mathcal{W}_{k'+l}$ with $k'+l > k$. Then for all $w' \in \mathcal{W}_{k'+l}$ we have

$$\eta_0 < \mathrm{EmpDist}(w')(w) < \eta_1.$$

In particular, for all shift invariant measures μ on the limit $\mathbb{K}$ of $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ we have $\eta_0 \leq \hat{\mu}_k(w) \leq \eta_1$.

6.2. The small word property.

Definition 38: Let $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ be a construction sequence. Let

$$f_n = \sup\{\mathrm{EmpDist}(w')(w) : w \in \mathcal{W}_n, w' \in \mathcal{W}_{n+1}\}$$

be the supremum of the empirical distributions of the n-words in $n+1$-words. The sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ has the **small word property** with respect to a sequence $\langle \delta_n : n \in \mathbb{N}\rangle$ if and only if for all n,

$$f_n < \delta_n < 1.$$

Odometer based systems are simpler. For odometer based construction sequences we write the lengths of the n-words, q_n, as K_n (so $K_n = \prod_{m<n} k_m$). With this notation the definition of empirical distribution simplifies. For $n < m$, $w \in \mathcal{W}_n$, $w' \in \mathcal{W}_m$, the empirical distribution $\mathrm{EmpDist}(w')(w)$ is simply the frequency of occurrences of w in w', which is given by

$$\mathrm{Freq}(w, w') = \frac{\text{number of occurrences of } w \text{ in } w'}{K_m/K_n}.$$

We use the following remark in [9].

Remark 39: For $n < m$, clause 3 of the definition of a construction sequence (Definition 9) together with Remark 34 implies that the frequency of each word $w \in \mathcal{W}_n$ inside each $w' \in \mathcal{W}_m$ is at least $1/k_n$.

Definition 38 can be restated for odometer based construction sequences as saying that if

$$f_n = \sup\{\mathrm{Freq}(w, w') : w \in \mathcal{W}_n, w' \in \mathcal{W}_{n+1}\}$$

is the supremum of the frequencies of the n-words in $n+1$-words, then the sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ has the **small word property** with respect to a sequence $\langle \delta_n : n \in \mathbb{N}\rangle$ if and only if for all n, $f_n < \delta_n$.

Since odometer based construction sequences have no spacers, the hypothesis about spacers mentioned in Lemma 34 does not arise. Restating the lemma in the language of frequencies, if $w \in \mathcal{W}_k$ and for all $w' \in \mathcal{W}_{k+1}$, $\eta_0 < \mathrm{Freq}(w, w') < \eta_1$, then for $k + l > k$, $w' \in \mathcal{W}_{k+l}$ we have

$$\eta_0 < \mathrm{Freq}(w, w') < \eta_1.$$

We preserve the following proposition for future use:

Proposition 40: *Let $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ be an odometer based sequence. Suppose that $f_n < \delta$. Then for all $w \in \mathcal{W}_n$ and all shift invariant measures μ on the limit $\mathbb{K}$*

$$\hat{\mu}_n(w) < \delta,$$
$$\mu(\langle w\rangle) < \delta/K_n.$$

The next lemma gives upper and lower bounds on measures of basic open intervals.

LEMMA 41: *Let $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ be an odometer based construction sequence for the system $\mathbb{K}$, and ρ be a shift-invariant measure on $\mathbb{K}$. Then for all words $w \in \mathcal{W}_n$*

$$\frac{1}{K_{n+1}} \leq \rho(\langle w\rangle) \leq \frac{f_n}{K_n}.$$

Proof. By the ergodic theorem, it suffices to show that for all $n < m$ and all $w \in \mathcal{W}_n, w' \in \mathcal{W}_m$,

$$\frac{1}{K_{n+1}} \leq \frac{|\{k \leq K_m : w' \restriction [k, k+k_n) = w\}|}{K_m} \leq \frac{f_n}{K_n}.$$

Write $w' = w_0 w_1 \cdots w_{K_m/K_{n+1}}$ where each $w_i \in \mathcal{W}_{n+1}$. Because w occurs at least once inside each w_i, the density of

$$D = \{k : w_i \restriction [k, k+k_n)\}$$

is at least $1/K_{n+1}$. The first inequality follows.

To see the second, for all i, note that for all w_i

$$f_n \geq \operatorname{Freq}(w, w_i) = \frac{\text{number of occurrences of } w \text{ in } w_i}{K_{n+1}/K_n}.$$

Hence the number of occurrences of w in w' is bounded by

$$f_n\Big(\frac{K_{n+1}}{K_n}\Big)\Big(\frac{K_m}{K_{n+1}}\Big) = f_n\Big(\frac{K_m}{K_n}\Big).$$

It follows that the density of D is bounded by

$$f_n\Big(\frac{K_m}{K_n}\Big)\Big(\frac{1}{K_m}\Big) = \frac{f_n}{K_n}. \qquad \blacksquare$$

Thus if $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ has the small word property with respect to $\langle \delta_n : n \in \mathbb{N}\rangle$ with $\delta_n < 1$, then for all $w \in \mathcal{W}_n, w' \in \mathcal{W}_{n+1}$ and all invariant measures ρ

$$\rho(\langle w'\rangle) < \frac{\delta_{n+1}}{K_{n+1}} < \rho(\langle w\rangle). \tag{10}$$

Our next step is to show that if $\mathfrak{O}$ is an odometer transformation, then $\mathfrak{O}$ has a presentation as an odometer based system with the small word property for some sequence $\langle \delta_n : n \in \mathbb{N}\rangle$ tending to 0. We do this by modifying Lemma 27.

LEMMA 42: *Let $\mathfrak{O} = \prod_{n\in\mathbb{N}} \mathbb{Z}_{k_n}$ be an odometer system with invariant measure μ. Then $\mathfrak{O}$ is isomorphic to $(\mathbb{K}, \nu)$ where $\mathbb{K}$ is the limit of an odometer based construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ with f_n tending monotonically to zero exponentially fast; in particular, $\sum f_n < \infty$.*

Proof. Let $\mathfrak{O}$ be an odometer based on $\langle k_n : n \in \mathbb{N}\rangle$. Let n_i be a monotone strictly increasing sequence and define

$$l_i = \prod_{n_{i-1} \leq n < n_i} k_n.$$

By Proposition 8, $\mathfrak{O}$ is isomorphic to the odometer based on $\langle l_i : i \in \mathbb{N}\rangle$. Thus by passing to a subsequence we can assume that

$$k_n > 3s_n(2^n + 1)K_n$$

(recall that s_n is the number of words in $\mathcal{W}_n$). To begin, let $\mathcal{W}_0 = \Sigma = \{a, b, c\}$.[6]

Suppose that we have constructed $\mathcal{W}_n$ and it is enumerated in lexicographical order as $\{w_i^n : 1 \leq i \leq s_n\rangle$. For each non-identity permutation σ of $\{1, 2, 3, \ldots, s_n\}$, let w_σ be the three-fold concatenation of the words in $\mathcal{W}_n$ in the order given by σ:

$$w_\sigma = \left(\prod_{i=1}^{s_n} w_{\sigma(i)}^n\right)^3.$$

Note that the length of w_σ is $3s_nK_n$.

Write $k_n = s_n(c_n + 3) + d_n$ where $c_n \in \mathbb{N}$, $0 \leq d_n < s_n$ and let

$$\vec{t} = \left(\prod_{i=1}^{s_n} w_i^n\right)^{c_n} * \prod_{i=1}^{d_n} w_i^n.$$

Note that the final segment, $\prod_{i=1}^{d_n} w_i^n$, has length d_nK_n. Finally we let

$$\mathcal{W}_{n+1} = \{w_\sigma{}^\frown\vec{t} : \sigma \text{ is a non-trivial permutation of } \{1, 2, \ldots, s_n\}\}.$$

In words: we begin by making $s_n! - 1$ prefixes w_σ by concatenating the words in $\mathcal{W}_n$ in all possible orders. We then use a single, much longer, suffix to complete each word. Note that $\mathcal{W}_{n+1}$ has at least $s_n! - 1$ many words in it.

Since each prefix is uniquely readable and comes from a non-trivial permutation σ, the words in $\mathcal{W}_{n+1}$ are uniquely readable. Moreover, any two words in $\mathcal{W}_n$ occur with approximately the same frequency in each word in $\mathcal{W}_{n+1}$. This precision gets better in a summable way as n increases to ∞. The words in $\mathcal{W}_{n+1}$ are clearly concatenations of words in $\mathcal{W}_n$. Let $\mathbb{K}$ be the limit of $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$.

[6] This construction can be easily modified to work in a 2-letter alphabet, by changing $\mathcal{W}_1$ in an ad hoc way.

By assumption on k_n the prefix makes up less than 2^{-n} portion of a word in $\mathcal{W}_{n+1}$. Let

$$G = \{x \in \boldsymbol{O} : \text{for large enough } m, 3s_m < x(m) < k_m - 10\}.$$

For $x \in G$, for some k and all $m \geq k$ and $w \in \mathcal{W}_{m+1}$, $x(m)K_m$ is bigger than the length of the prefix w_σ and less than $K_{m+1} - (d_m + 10)k_m$.

As before we let

$$r_n = x(0) + x(1)K_1 + x(2)K_2 + \cdots + x(n)K_n.$$

Because $x(n) \geq 3s_m$ for all $n+1$ words $w \in \mathcal{W}_{n+1}$, the $x(n)^{th}$ letter is not in a prefix. It follows that for all $w \in \mathcal{W}_{n+1}$, the $x(n)^{th}$ subword is the same. Let w_m be this word, and let $\psi(x)$ be the element $s \in \mathbb{K}$ determined by $\langle r_n : n \geq k\rangle$ and $\langle w_n : n \geq k\rangle$.

As in Lemma 27, G is a measure one Borel set and the map

$$\psi : G \overset{1-1}{\longrightarrow} \mathbb{K}$$

is continuous and one-to-one.

It remains to show that the f_n are small. Since each word in $\mathcal{W}_n$ occurs very close to the same number of times in each $\mathcal{W}_{n+1}$, the frequencies of occurrences of each word is to $1/s_n$. Since s_n grows as an iterated factorial, f_n goes to zero exponentially. ∎

6.3. Arbitrary rates of descent. Fix an odometer based construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ with $f_n \leq b_n$ for some sequence $\langle b_n : n \in \mathbb{N}\rangle$ going to zero. Let $\langle \delta_i : i \in \mathbb{N}\rangle$ be a sequence of positive numbers less than one. Then there is a subsequence $\langle n_i : i \in \mathbb{N}\rangle$ such that $b_{n_i} < \delta_i$. Let $\mathcal{V}_i = \mathcal{W}_{n_i}$.

We claim that $\langle \mathcal{V}_i : i \in \mathbb{N}\rangle$ has the small word property with respect to $\langle \delta_i : i \in \mathbb{N}\rangle$. This follows because the frequency of each $w \in \mathcal{W}_{n_i}$ in each $w' \in \mathcal{W}_{n_i+1}$ is bounded above by $b_{n_i} < \delta_i$. Applying Lemma 34 we see that the frequency of each $w \in \mathcal{V}_i$ in each $w' \in \mathcal{W}_{n_{i+1}} = \mathcal{V}_{i+1}$ is bounded above by δ_i.

This subsequence can be chosen continuously in the parameters $\langle b_i, \delta_i\rangle$ and any tail of any sufficiently fast growing subsequence has the small word property with respect to $\langle \delta_n : n \in \mathbb{N}\rangle$. We elaborate on this after the next theorem.

We now note the following:

Theorem 43: *Let $\mathfrak{O}$ be an odometer system. Let $\langle \mathcal{W}_n^{\mathfrak{O}} : n \in \mathbb{N}\rangle$ be a construction sequence for $\mathfrak{O}$ that has the small word property for $\langle \delta_n : n \in \mathbb{N}\rangle$.*

- *If $T : (X, \mu) \to (X, \mu)$ is an ergodic transformation with finite entropy having $\mathfrak{O}$ as a factor, and $\langle \mathcal{W}_n^X : n \in \mathbb{N}\rangle$ is the presentation of X as a limit of the odometer based system $\langle \mathcal{W}_n^X : n \in \mathbb{N}\rangle$ constructed by Theorem 21 as modified in Remark 29, then $\langle \mathcal{W}_n^X : n \in \mathbb{N}\rangle$ has the small word property for $\langle \delta_n : n \in \mathbb{N}\rangle$.*
- *If x is a Toeplitz sequence with underlying odometer $\mathfrak{O}$, then the presentation of the orbit closure $\mathbb{L}$ of x as the limit $\mathbb{L}^*$ of an odometer based construction sequence given in Corollary 33 has the small word property with parameters $\langle \delta_n : n \in \mathbb{N}\rangle$.*

Proof. In both cases the words in the respective construction sequences were of the form (u, v), where v is in the construction sequence for a presentation of $\mathfrak{O}$. Since the construction sequence for $\mathfrak{O}$ has the small word property, the given construction sequence does as well. ∎

Theorem 43 reduces the problem of finding presentations of odometer based systems with the small word property to the problem of finding a presentation of the underlying odometer with the small word property. By Lemma 42, we can do this for a single sequence $\langle f_n\rangle$ tending to zero.

The small word property can be arranged continuously. Fix an odometer construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$, let $n_0 = 0$ and consider the following game $\mathfrak{G}(\langle \mathcal{W}_n : n \in \mathbb{N}\rangle)$. Let $\langle b_n : n \in \mathbb{N}\rangle$ be a sequence with $b_n > f_n$ for all n. At round $k \geq 0$:

- Player I plays $\epsilon_k > 0$.
- Player II plays $n_{k+1} > n_k$.

Player II wins $\mathfrak{G}(\langle \mathcal{W}_n : n \in \mathbb{N}\rangle)$ if and only if $b_{n_{k+1}} < \epsilon_k$ for all k.

We record the following remark for applications in other contexts.

Remark 44: If b_n converges to 0 then player II has a winning strategy in $\mathfrak{G}(\langle \mathcal{W}_n : n \in \mathbb{N}\rangle)$. Moreover, by Theorem 43, if $\mathcal{S}$ is this strategy for an odometer based presentation $\langle \mathcal{W}_n^{\mathfrak{O}} : n \in \mathbb{N}\rangle$, then $\mathcal{S}$ is also a winning strategy for all odometer based presentations $\langle \mathcal{W}_n^X : n \in \mathbb{N}\rangle$ built over $\langle \mathcal{W}_n^{\mathfrak{O}} : n \in \mathbb{N}\rangle$.

In particular, we can choose the subsequence n_k continuously in the ϵ_k.

References

[1] T. Adams, S. Ferenczi and K. Petersen, *Constructive symbolic presentations of rank one measure-preserving systems*, Colloquium Mathematicum **150** (2017), 243–255.

[2] F. Beleznay and M. Foreman, *The complexity of the collection of measure-distal transformations*, Ergodic Theory and Dynamical Systems **16** (1996), 929–962.

[3] T. Downarowicz, *The Choquet simplex of invariant measures for minimal flows*, Israel Journal of Mathematics **74** (1991), 241–256.

[4] T. Downarowicz and Y. Lacroix, *Almost* 1-1 *extensions of Furstenberg–Weiss type and applications to Toeplitz flows*, Studia Mathematica **130** (1998), 149–170.

[5] S. Ferenczi, *Systems of finite rank*, Colloquium Mathematicum **73** (1997), 35–65.

[6] M. Foreman, D. J. Rudolph and B. Weiss, *The conjugacy problem in ergodic theory*, Annals of Mathematics **173** (2011), 1529–1586.

[7] M. Foreman and B. Weiss, *From odometers to circular systems: a global structure theorem*, Journal of Modern Dynamics **15** (2019), 345–423.

[8] M. Foreman and B. Weiss, *A symbolic representation for Anosov–Katok systems*, Journal d'Analyse Mathématique **137** (2019), 603–661.

[9] M. Foreman and B. Weiss, *Measure preserving diffeomorphisms are unclassifiable*, Journal of the European Mathematical Society, to appear.

[10] W. Krieger, *On entropy and generators of measure-preserving transformations*, Transactions of the American Mathematical Society **149** (1970), 453–464.

[11] B. Weiss, *Strictly ergodic models for dynamical systems*, Bulletin of the American Mathematical Society **13** (1985), 143–146.

[12] B. Weiss, *Single Orbit Dynamics*, CBMS Regional Conference Series in Mathematics, Vol. 95, American Mathematical Society, Providence, RI, 2000.

[13] S. Williams, *Toeplitz minimal flows which are not uniquely ergodic*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **67** (1984), 95–107.