Title
Measure preserving diffeomorphisms of the torus are unclassifiable

Journal
Journal of the European Mathematical Society, 24(8)

ISSN
1435-9855

Authors
Foreman, Matthew
Weiss, Benjamin

Publication Date
2021-12-08

DOI
10.4171/jems/1151

J. Eur. Math. Soc. 24, 2605–2690 (2022) DOI 10.4171/JEMS/1151

Matthew Foreman · Benjamin Weiss

Measure preserving diffeomorphisms of the torus are unclassifiable

Received May 22, 2019; revised November 5, 2020

Abstract. In 1932 von Neumann proposed classifying the statistical behavior of differentiable systems. In modern language this is interpreted as classifying diffeomorphisms of compact manifolds up to measure isomorphism. This paper proves that this is impossible in a rigorous sense.

Keywords. Classification of ergodic transformations, odometers, circular systems, distal transformations, Anosov–Katok method

Contents

Matthew Foreman: Mathematics Department, University of California, Irvine, Irvine, CA 92697, USA; mforeman@math.uci.edu

Benjamin Weiss: Einstein Institute of Mathematics, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel; weiss@math.huji.ac.il

Mathematics Subject Classification (2020): Primary 37A05; Secondary 37A35

1. Introduction

The isomorphism problem in ergodic theory was formulated by von Neumann in 1932 in his pioneering paper [23].[1] The problem has been solved for some classes of transformations that have special properties. Halmos and von Neumann [15] used the unitary operators defined by Koopman to completely characterize ergodic measure preserving transformations with pure point spectrum. They showed that these are exactly the transformations that can be realized as translations on compact groups. Another notable success in solving this problem was the classification of Bernoulli shifts using the notion of entropy introduced by Kolmogorov.

Starting in the late 1990s a different type of result began to appear: *anti-classification* results that demonstrate in a rigorous way that classification is not possible. This type of theorem requires a precise definition of what a classification is. Informally, a classification is a method of determining isomorphism between transformations by computing (in a liberal sense) other invariants for which equivalence is easy to determine.

The key words here are *method* and *computing*. For negative theorems, the more liberal a notion one takes for these words, the stronger the theorem. One natural way of what a computation is uses the Borel/non-Borel distinction. Saying a set X or function f is Borel is a loose way of saying that membership in X or the computation of f can be done using a countable (possibly transfinite) protocol whose basic input is membership in open sets. Saying that X or f is *not* Borel is saying that determining membership in X or computing f cannot be done with any countable amount of resources. (See [6] for an elementary discussion and a comparison with the more strict notion of *recursive computation*, which requires inherently finite resources.)

In the context of classification problems, saying that an equivalence relation E on a space X is *not* Borel is saying that there is no countable amount of initial information and no countable, potentially transfinite, protocol based on this information for determining, for arbitrary $x, y \in X$ whether xEy. Any such method must inherently use uncountable resources.[2]

An example of a positive theorem in the context of ergodic theory is due to Halmos ([14]) who showed that the collection of ergodic measure preserving transformations is a dense $\mathcal{G}_\delta$ set in the space of all measure preserving transformations of $([0, 1], \lambda)$ endowed with the weak topology. Moreover, he showed that the set of weakly mixing transformations is also a dense $\mathcal{G}_\delta$.[3]

[1] Two measure preserving transformations (abbreviated to 'MPTs' in the paper) T and S are isomorphic if there is an invertible measurable mapping between the corresponding measure spaces which commutes with the actions of T and S.

[2] Many well known classification theorems have as immediate corollaries that the resulting equivalence relation is Borel. An example of this is the Spectral Theorem, which has a consequence that the relation of Unitary Conjugacy for normal operators is a Borel equivalence relation.

[3] Relatively straightforward arguments show that the set of strongly mixing transformation is a first category Π^3_0 set. See [5].

The first anti-classification result in the area is due to Beleznay and Foreman [3] who showed that the class of *measure distal* transformations used in early ergodic theoretic proofs of Szemeredi's theorem is not a Borel set. Later Hjorth [16] introduced the notion of *turbulence* and showed that there is no Borel way of attaching algebraic invariants to ergodic transformations that completely determine isomorphism. Foreman and Weiss [10] improved this result by showing that the conjugacy action of the measure preserving transformations is turbulent – hence no generic class can have a complete set of algebraic invariants.

In considering the isomorphism relation as a collection $\mathcal{I}$ of pairs (S, T) of measure preserving transformations, Hjorth ([17]) showed that $\mathcal{I}$ is not a Borel set. However the pairs of transformations he used to demonstrate this were inherently non-ergodic, leaving open the essential problem:

Question. Is isomorphism of ergodic measure preserving transformations Borel?

This question was answered in the negative by Foreman, Rudolph and Weiss in [8]. This answer can be interpreted as saying that determining isomorphism between ergodic transformations is inaccessible to countable methods that use countable amounts of information.

In the same foundational paper from 1932 von Neumann expressed the likelihood that any abstract MPT is isomorphic to a continuous MPT and perhaps even to a differentiable one. This brief remark eventually gave rise to one of the yet outstanding problems in smooth dynamics, namely:

Question. Does every ergodic MPT with finite entropy have a smooth model?[4]

By a smooth model it is meant an isomorphic copy of the MPT which is given by smooth diffeomorphism of a compact manifold preserving a measure equivalent to the volume element. Soon after entropy was introduced, A. G. Kushnirenko showed that such a diffeomorphism must have finite entropy, and up to now this is the only restriction that is known. The current paper is the culmination of a series whose purpose is to show that the variety of ergodic transformations that have smooth models is rich enough so that the abstract isomorphism relation, when restricted to these smooth systems, is as complicated as the general isomorphism problem for ergodic measure preserving systems. We show that even when restricting to diffeomorphisms of the 2-torus that preserve Lebesgue measure this is the case. The formal statement of our solution to the isomorphism problem is:

Theorem 1. *If M is either the torus $\mathbb{T}^2$, the disk D or the annulus then the measure-isomorphism relation among pairs (S, T) of measure preserving C^∞-diffeomorphisms of M is not a Borel set with respect to the C^∞-topology.*

[4]In [23] on page 590, "Vermutlich kann sogar zu jeder allgemeinen Strömung eine isomorphe stetige Strömung gefunden werden [footnote 13], vielleicht sogar eine stetig-differentiierbare, oder gar eine mechanische. Footnote 13: Der Verfasser hofft, hierfür demnächst einen Beweis anzugeben."

Thus the isomorphism problem is impossible even for diffeomorphisms of compact surfaces.

How does one prove a result such as Theorem 1? The main tool is the idea of a *reduction* (see [6] and Section 4.6). A function $f : X \to Y$ reduces A to B if and only if for all $x \in X$:

$$x \in A \text{ if and only if } f(x) \in B.$$

If X and Y are completely metrizable spaces and f is a Borel function, then f is a method of reducing the question of membership in A to membership in B. Thus if A is not Borel then B cannot be either.

In the current context, the C^∞-topology on the smooth transformations refines the weak topology. Thus, by Halmos' result quoted earlier, on the torus (disk, etc.), the ergodic transformations are still a $\mathcal{G}_\delta$-set. (However the famous KAM theory shows that the ergodic transformations are no longer dense.) In particular, the C^∞-topology induces a metrizable complete and perfect topology on the measure preserving diffeomorphisms of $\mathbb{T}^2$. If M is a manifold with supporting a measure μ, we denote the space of C^∞, μ-measure preserving diffeomorphisms of M with the notation $\mathrm{Diff}^\infty(M, \mu)$. Elements of $\mathrm{Diff}^\infty(M, \mu)$ are also members of the group MPT of μ-measure preserving transformations. For $T \in \mathrm{Diff}^\infty(M, \mu)$ the centralizer of T in MPT is denoted $C(T)$.

If X is perfect and completely metrizable, a set $A \subseteq X$ is *analytic* if and only if A is the continuous image of a Borel set. A is *complete analytic* if and only if every analytic set can be reduced to A. It is a classical fact that complete analytic sets are not Borel.

The proof of Theorem 1 uses a well-known example of a complete analytic set. The underlying space X is the space $\mathcal{Trees}$ and A is the collection of ill-founded trees; those that have infinite branches. A precise statement of the main result of the paper:

Theorem 2. *There is a continuous function $F^s : \mathcal{Trees} \to \mathrm{Diff}^\infty(\mathbb{T}^2, \lambda)$, taking values among the ergodic transformations, such that for $\mathcal{T} \in \mathcal{Trees}$, if $T = F^s(\mathcal{T})$:*

(1) *$\mathcal{T}$ has an infinite branch if and only if $T \cong T^{-1}$, and*

(2) *$\mathcal{T}$ has two distinct infinite branches if and only if*

$$C(T) \neq \overline{\{T^n : n \in \mathbb{Z}\}}.$$

Corollary 3. *The following statements hold:*

- $\{T \in \mathrm{Diff}^\infty(\mathbb{T}^2, \lambda) : T$ *is ergodic and* $T \cong T^{-1}\}$ *is complete analytic.*
- $\{T \in \mathrm{Diff}^\infty(\mathbb{T}^2, \lambda) : T$ *is ergodic and* $C(T) \neq \overline{\{T^n : n \in \mathbb{Z}\}}$ *is complete analytic.*

Since the map

$$\iota(T) = (T, T^{-1})$$

is a continuous mapping of $\mathrm{Diff}^\infty(\mathbb{T}^2, \lambda)$ to $\mathrm{Diff}^\infty(\mathbb{T}^2, \lambda) \times \mathrm{Diff}^\infty(\mathbb{T}^2, \lambda)$ and reduces $\{T : T \cong T^{-1}\}$ to $\{(S, T) : S \cong T\}$, it follows that:

Corollary 4. *The set*

$$\{(S, T) : S \text{ and } T \text{ are ergodic diffeomorphisms of } \mathbb{T}^2 \text{ and are isomorphic}\}$$

is complete analytic and hence not Borel.

We note that the problem of finding even one measure preserving transformation not isomorphic to its inverse is difficult. This was not done until Anzai in [2]. In Math Review MR0047742, Halmos said, "By constructing an example of the type described in the title the author solves (negatively) a problem proposed by the reviewer and von Neumann [Ann. of Math. (2) 43, 332?350 (1942): MR0006617]".

More fine-grained information is now known and will be published elsewhere. For example, Foreman, in unpublished work, showed that the problem of "isomorphism of countable graphs" is Borel reducible to the isomorphism problem for ergodic measure preserving transformations.

The techniques of this paper also have foundational interest. A close analysis of our construction shows that the problem of whether T is isomorphic to its inverse is "Π^0_1-hard." (See [7]). This enables one to prove that truth or falsity of various open problems like the Riemann hypothesis is equivalent to the question of is T_{RH} isomorphic or not to its inverse for a specific measure preserving diffeomorphism T_{RH} of the torus given by our construction. Another consequence is the existence of a different diffeomorphism T_{ZFC} such that the question of whether T_{ZFC} is isomorphic to its inverse is independent of ZFC, the usual axioms for mathematics.

Here are two problems that remain open:

Problem 1. In contrast to [10], where the authors were able to show that the equivalence relation of isomorphism on abstract ergodic measure preserving transformations is *turbulent*, this remains open for ergodic diffeomorphisms of a compact manifold.

Problem 2. The problem of classifying diffeomorphisms of compact surfaces up to topological conjugacy remains largely open. Work of the first author with A. Gorodetski shows that the isomorphism relation itself is not Borel, but for a very specific type of diffeomorphisms of manifolds of dimension 5 and above. It is not know, for example for topologically minimal transformations.

We owe a substantial debt to everyone who has helped us with this project. Jean-Paul Thouvenot brought the Anosov–Katok technique to our attention and suggested using it to solve the von Neumann problem. Philipp Kunde aided us by reading the paper and providing comments and corrections. Others include Eli Glasner, Anton Gorodetski, Alekos Kechris, and Anatole Katok.

We particularly want to acknowledge the contribution of the late Dan Rudolph, who helped pioneer these ideas and was a co-author in [8], contributing techniques fundamental to this paper.

2. An outline of the argument

This section gives an outline of the argument for Theorem 2. It uses the main results from our earlier papers: *A symbolic representation of Anosov–Katok systems* ([11]) and *From odometers to circular systems: A global structure theorem* ([12]) which we briefly summarize. In [11], the Anosov–Katok technique of Approximation by Conjugacy is used to give a new symbolic representation for a class of measure preserving diffeomorphisms that are extensions of the rotations by certain Liouvillean α. These are called strongly uniform **Circular Systems**.[5]

In [12] two classes of symbolic systems are defined. The first, called *Odometer Based* systems, contains representatives of every finite entropy measure preserving transformation with an odometer factor. The second class is the collection of *Circular Systems*. These classes are made into categories by taking as morphisms synchronous and anti-synchronous factor maps. The main result is that there is a functorial isomorphism between $\mathcal{F}$ between these categories that takes strongly uniform systems to strongly uniform systems.

Since the main construction in [8] uses Odometer Based systems this map enables us to adapt that construction to the smooth setting. However in order to prove our main result we still have to take into account potential isomorphisms of Circular Systems that are neither synchronous nor anti-synchronous. It is to deal with this difficulty that we analyze what we call the *displacement function*.

To each α arising as a rotation factor of a circular system T one can associate a *displacement function* (Section 7.1) and use it to associate the set of *central values*, a subgroup of the unit circle. Its significance is the following:

(1) (Theorem 84) If β is central, then there is an $\phi \in \overline{\{T^n : n \in \mathbb{Z}\}}$ such that the rotation factor of ϕ is rotation by β.

(2) (Theorem 90) If T is built from sufficiently random words,[6] and $\phi \in C(T)$, then the canonical rotation factor of ϕ is rotation by a central value.

(3) It follows that if there is a $\phi \in C(T)$ and $\phi \notin \overline{\{T^n : n \in \mathbb{Z}\}}$, then there is a synchronous $\psi \in C(T)$ such that $\psi \notin \overline{\{T^n : n \in \mathbb{Z}\}}$.

(4) (Theorem 92) The analogous results relating isomorphisms ϕ between T and T^{-1} with central values is proved, allowing us to conclude that if T is isomorphic to T^{-1}, then there is an anti-synchronous isomorphism between T and T^{-1}.

(5) The previous two items are the content of Theorem 93, which says that for T satisfying the Timing Assumptions, to decide whether $T \cong T^{-1}$ or $C(T) \neq \overline{\{T^n : n \in \mathbb{Z}\}}$ it suffices to consider anti-synchronous and synchronous isomorphisms.

In [8] a continuous function F from the space of Trees to the strongly uniform odometer based transformations is constructed that:

- reduces the set of ill-founded trees to the transformations T that are isomorphic to their inverses (and *if* $T \cong T^{-1}$, then this is witnessed by an anti-synchronous isomorphism) and
- reduces the set of trees with two infinite branches to the transformations T whose centralizer is different from the powers of T (and if the centralizer contains an exotic element, it contains a synchronous exotic element).

[5] In a forthcoming paper we show how to drop the "strongly uniform" assumption.

[6] That is, T satisfies the *Timing Assumptions*.

Moreover, in the second case, there is a synchronous element of the centralizer with a specific piece of evidence that it is not the identity (it moves a $\mathcal{Q}_1^1$-equivalence class).

Composing one concludes that $\mathcal{F} \circ F$:

- reduces the set of ill-founded trees to collection of circular systems that are isomorphic to their inverses and
- reduces the set of trees with two infinite branches to the circular systems whose centralizer is different from the closure of the powers of T.

Continuously realizing the circular systems by R (as in [11]) completes the proof that:

- The collection of ergodic measure preserving diffeomorphisms T of the torus that are isomorphic to their inverses is complete analytic. Consequently, the set of pairs (S, T) of ergodic conjugate measure preserving diffeomorphisms is a complete analytic set.
- The collection of ergodic measure preserving diffeomorphisms T whose centralizer is different from the closure of the powers of T is complete analytic.

Figure 1 illustrates $F^s = R \circ \mathcal{F} \circ G$.

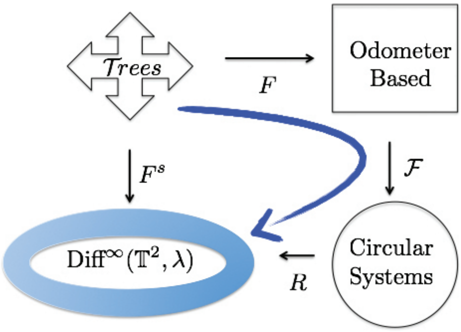

Fig. 1. The reduction F^s.

The next two sections review basic facts in ergodic theory and descriptive set theory, define *odometer based* and *circular* systems and review their properties and the facts shown in [11] and [12].

The analysis of the displacement function and the associated central values, which are a subgroup of the circle canonically associated to the Liouvillean α, is carried out in Sections 5–7. Finally, the proof of the main theorems are given in Section 8 modulo certain properties which impose some additional conditions on the parameters of the construction in [8]. These are verified in Section 9 and in Section 10 we spell out the dependencies between the various parameters and show that they can be realized.

3. Numerical requirements

The proof of Theorem 2 uses a construction with many interconnecting pieces, most of which are built by taking limits. This results in a large number of related sequences of variables, each having their own requirements and the estimates for the different pieces must be compatible.

The least interesting part of this paper is verifying the consistency of the numerical requirements. Sorting these requirements out is completely independent of the rest of the paper. For this reason, we list the numerical requirements in Section 11.1, and then give an argument for their consistency. We also note the specific requirement by number in the text as they are posited and used.

Contributing to the complexity of the situation is that many of the relationships between the variables come from internal arguments of the general form "taking δ small enough you can guarantee that $x < \epsilon$", with various variables in place of ϵ, δ and x. The exact relationship between ϵ and δ is not clear from the argument, but there is a requirement of the form *"δ is small as a function of ϵ."* A typical example of this is Sublemma 99 which says that, as a function of Q_1^n, if ϵ_n is take sufficiently small then an involved inequality involving I^*, u_i', v_i' and Q_1^n holds.

Complicating this task further is the fact that the construction in this paper depends on the construction in [8], which has its own numerical requirements. For a reader tracking the correspondence, in the appendix, we include a table for translating between the notation in this paper and the notation in [8].

The variables. Here is a list of variable sequences that have to be chosen during the construction:

$$k_n, l_n, q_n, s_n, e(n), p_n, q_n, \alpha_n, \epsilon_n, \varepsilon_n, \mu_n, Q_1^n.$$

Some of these variables have clear relationships that are externally determined. The main construction is of a function that has a tree as in input. That tree *directly* determines a sequence of parameters, such as G_1^n and $\langle M(s) : s \leq n\rangle$ that are not *chosen* during the construction. (In Section 11, we call these *exogenous* variables.) These parameters determine some of the numerical requirements.

Example 5. The words in the collection $\mathcal{W}_{n+1}$ are built by a sequence of M substitutions into equivalence classes of the relations $\mathcal{Q}_i^{n+1}$, where $M = \sup_S M(s)$ for S the collection of heights on nodes in the given tree at stage n. These substitution instances are closed under a sequence of $\mathbb{Z}_2$ actions of the groups $\langle G_i^n : i \leq M\rangle$. The number M and the dimensions of the $\mathbb{Z}_2$ actions are also determined by the tree. Thus s_{n+1} is determined by the exogenous variables G_i^n, $M(s)$, and the internally chosen variable $e(n+1)$. In this particular example, It is possible to give a completely explicit formula for $e(n+1)$ in terms s_{n+1} and vice versa.[7]

However that would be uninformative. What we need to see is that if $e(n+1)$ is large, then s_{n+1} is and vice versa and that each determines the other. This is the only relevant information for determining the consistency of the numerical requirements. We have thus eliminated one variable.

It would perhaps be more conventional to define all of the variables in advance, write down the list of inequalities and then show they are consistent. However the examples

[7] We have $s_{n+1} = (2^{Me(n+1)})G$ for numbers M and G determined exogenously.

above illustrate the difficulties with this. The inequalities are intimately intertwined with the details of the construction and are completely enigmatic without that context. For this reason we note the numerical requirements one by one as they accumulate and collect them in Section 11.1. We then proceed to show that they are consistent by the method we describe next. A reader with a preference for the conventional presentation is advised to skip directly to Section 11, read the reconciliation and then return to read the rest of the paper.

What could possibly go wrong? The only potential issue is that there may be a situation where the requirements are circular: for example, δ might have to be small as a function of ϵ, ϵ small as a function of μ and μ small as a function of δ. In symbols

$$\epsilon \to \delta \to \mu \to \epsilon.$$

So if you choose ϵ first, then δ then μ, you might find that your choice of ϵ was inadequate. Indeed, because there is a cycle in the dependency diagram there is no variable you can choose first and be certain of consistency.

Method for showing consistency. In Section 11 we analyze the dependencies and draw a dependency diagram giving the order of choice. Since that diagram is cycle free, all of the variables can be chosen to satisfy the accumulated requirements.

4. Preliminaries

The reader is referred to standard texts such as [22], [24] or [21]. Facts that are not standard and are simply cited here are proved in [12], [11] and [8].

4.1. Measure spaces

We will call separable non-atomic probability spaces *standard measure spaces* and denote them $(X, \mathcal{B}, \mu)$, where $\mathcal{B}$ is the Boolean algebra of measurable subsets of X and μ is a countably additive, non-atomic measure defined on $\mathcal{B}$. Maharam and von Neumann proved that every standard measure space is isomorphic to $([0,1], \mathcal{B}, \lambda)$, where λ is Lebesgue measure and $\mathcal{B}$ is the algebra of Lebesgue measurable sets.

If $(X, \mathcal{B}, \mu)$ and $(Y, \mathcal{C}, \nu)$ are measure spaces, an isomorphism between X and Y is a bijection $\phi : X \to Y$ such that ϕ is measure preserving and both ϕ and ϕ^{-1} are measurable. We will ignore sets of measure zero when discussing isomorphisms; i.e. we allow the domain and range of ϕ to be subsets of X and Y of measure one.

A measure preserving system is an object $(X, \mathcal{B}, \mu, T)$, where $T : X \to X$ is a measure isomorphism. A *factor map* between two measure preserving systems $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ is a measurable, measure preserving function $\phi : X \to Y$ such that

$$S \circ \phi = \phi \circ T.$$

A factor map is an isomorphism between systems iff ϕ is a measure isomorphism.

Let $T : (X, \mathcal{B}, \mu, T) \to (X, \mathcal{B}, \mu, T)$ be measure preserving, let $(Y, \mathcal{C})$ be a measurable space, $S : Y \to Y$ a measurable map and $\phi : X \to Y$ a measurable map such that $\phi T = S\phi$. Then we can define a measure $\nu = \phi^*\mu$ by setting $\nu(A) = \mu(\phi^{-1}(A))$. This measure makes ϕ a factor map from $(X, \mathcal{B}, \mu, T)$ to $(Y, \mathcal{C}, \nu, S)$.

4.2. *Presentations of measure preserving systems*

Measure preserving systems occur naturally in many guises with diverse topologies. As far as is known, the Borel/non-Borel distinction for dynamical properties is the same in each of these presentations and many of the presentations have the same generic classes. (See the forthcoming paper [9] which gives a precise condition for this.)

Here is a review the properties of the types of presentations relevant to this paper, which are: abstract invertible preserving systems, smooth transformations preserving volume elements and symbolic systems.

4.2.1. Abstract measure preserving systems. Since every standard measure space is isomorphic to the unit interval with Lebesgue measure, every invertible measure preserving transformation of a standard measure space is isomorphic to an invertible Lebesgue measure preserving transformation on the unit interval.

In accordance with the conventions of [5] we denote the group of measure preserving transformations of $[0, 1)$ by **MPT**.[8] Two measure preserving transformations are identified if they are equal on sets of full measure.

Two measure preserving transformations are isomorphic if and only if they are conjugate in the group **MPT** and we will use *isomorphic* and *conjugate* as synonyms. However some caution is order. If (M, μ) is a manifold, $T : M \to M$ is a smooth measure preserving transformation and ϕ is an arbitrary measure preserving transformation from M to M, then $\phi T \phi^{-1}$ is unlikely to be smooth. Thus, the equivalence relation of isomorphism of diffeomorphisms is not given by an action of the group of measure preserving transformations in an obvious way.

Given a measure space (X, μ) and a measure preserving transformation $T : X \to X$, define the *centralizer of* T to be the collection of measure preserving $S : X \to X$ such that $ST = TS$. This group is denoted $C(T)$. Note that this is the centralizer *in the group of measure preserving transformations*. In the case that X is a manifold and T is a diffeomorphism, $C(T)$ differs from the centralizer of T inside the group of diffeomorphisms.

To each invertible measure preserving transformation $T \in$ **MPT**, associate a unitary operator $U_T : L^2([0, 1]) \to L^2([0, 1])$ by defining $U(f) = f \circ T$. In this way **MPT** can be identified with a closed subgroup of the unitary operators on $L^2([0, 1])$ with respect to the weak operator topology[9] on the space of unitary transformations. This makes **MPT** into a Polish group. We will call this the *weak topology* on **MPT**. Halmos ([14]) showed that the ergodic transformations, which we denote $\mathcal{E}$, is a dense $\mathcal{G}_\delta$ set in **MPT**. In particular, the weak topology makes $\mathcal{E}$ into a Polish subspace of **MPT**.

[8] Recently several authors have adopted the notation $\mathrm{Aut}(\mu)$ for the same space.

[9] Which coincides with the strong operator topology in this case.

There is another topology on the collection of measure preserving transformations of X to Y for measure spaces X and Y. If $S, T : X \to Y$ are measure preserving transformations, the *uniform distance* between S and T is defined to be

$$d_U(S,T) = \mu\{x : Sx \neq Tx\}.$$

This topology refines the weak topology and is a complete, but not a separable topology.

4.2.2. Diffeomorphisms. Let M be a C^m-smooth compact finite-dimensional manifold and let μ be a standard measure on M determined by a smooth volume element. For each $k \leq m$ there is a Polish topology on the k-times differentiable homeomorphisms of M, the C^k-topology. If M is C^∞, then the C^∞-topology is the coarsest topology refining the C^k-topology for each $k \in \mathbb{N}$. It is also a Polish topology and a sequence of C^∞-diffeomorphisms converges in the C^∞-topology if and only if it converges in the C^k-topology for each $k \in \mathbb{N}$.

The collection of μ-preserving diffeomorphisms forms a closed nowhere dense set in the C^k-topology on the C^k-diffeomorphisms, and as such, inherits a Polish topology.[10] We will denote this space by $\mathrm{Diff}^k(M,\mu)$.

Viewing M as an abstract measure space one can also consider the space of abstract μ-preserving transformations on M with the weak topology. In [4] it is shown that the collection of a.e.-equivalence classes of smooth transformations form a Π^0_3-set in $\mathbf{MPT}(M)$, and hence the collection has the Property of Baire.

4.2.3. Symbolic systems. Let Σ be a countable or finite alphabet endowed with the discrete topology. Then $\Sigma^{\mathbb{Z}}$ can be given the product topology, which makes it into a separable, totally disconnected space that is compact if Σ is finite.

Notation. If $u = \langle \sigma_0, \ldots, \sigma_{n-1}\rangle \in \Sigma^{<\infty}$ is a finite sequence of elements of Σ, then we denote the cylinder set based at k in $\Sigma^{\mathbb{Z}}$ by writing $\langle u\rangle_k$. If $k = 0$, we abbreviate this and write $\langle u\rangle$. Explicitly: $\langle u\rangle_k = \{f \in \Sigma^{\mathbb{Z}} : f \upharpoonright [k, k+n) = u\}$. The collection of cylinder sets form a base for the product topology on $\Sigma^{\mathbb{Z}}$.

Let u, v be finite sequences of elements of Σ having length q. Given intervals I and J in $\mathbb{Z}$ of length q, we can view u and v as functions having domain I and J, respectively. We will say that u and v are located at I and J. We will say that u is *shifted by k* relative to v iff I is the shift of the interval J by k. We say that u is the *k-shift* of v iff u and v are the same words and I is the shift of the interval j by k.

The shift map

$$\mathrm{sh} : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$$

defined by setting

$$\mathrm{sh}(f)(n) = f(n+1)$$

[10] One can also consider the space of measure preserving homeomorphisms with the $\|\cdot\|_\infty$ topology, which behaves in some ways similarly.

is a homeomorphism. If μ is a shift-invariant Borel measure, then the resulting measure preserving system $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, \mathrm{sh})$ is called a *symbolic system*. The closed support of μ is a shift-invariant closed subset of $\Sigma^{\mathbb{Z}}$ called a *symbolic shift* or *sub-shift*.

Symbolic shifts are often described intrinsically by giving a collection of words that constitute a clopen basis for the support of an invariant measure. Fix a language Σ, and a sequence of collections of words $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ with the properties that:

(1) for each n all of the words in $\mathcal{W}_n$ have the same length q_n,

(2) each $w \in \mathcal{W}_n$ occurs at least once as a subword of every $w' \in \mathcal{W}_{n+1}$,

(3) 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

$$u_0 w_0 u_1 w_1 \ldots w_l u_{l+1} \tag{4.1}$$

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

$$\frac{\sum_i |u_i|}{q_{n+1}} < \epsilon_{n+1}. \tag{4.2}$$

The segments u_i in condition 4.1 are called the *spacer* or *boundary* portions of w.

Definition 6. A sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ satisfying properties (1)–(3) will be called a *construction sequence*.

If $\mathcal{W}$ is a collection of words in an alphabet Σ, we will say that $\mathcal{W}$ is *uniquely readable* if and only if whenever $u, v, w \in \mathcal{W}$ and $uv = pws$ then either:

- $p = \emptyset$ and $u = w$ or
- $s = \emptyset$ and $v = w$.

Equation (4.1) of clause (3) implies that each $\mathcal{W}_n$ is uniquely readable. We will need unique readability to parse elements of $\mathbb{K}$, the symbolic shift associated with the construction sequence.

Definition 7. Let $\mathbb{K}$ be the collection of $x \in \Sigma^{\mathbb{Z}}$ such that every finite contiguous subword of x occurs inside some w belonging to some $\mathcal{W}_n$. Then $\mathbb{K}$ is a closed shift-invariant subset of $\Sigma^{\mathbb{Z}}$ that is compact if Σ is finite.

The symbolic shifts built from construction sequences coincide with transformations built by *cut-and-stack* constructions.

Notation. For a word $w \in \Sigma^{<\mathbb{N}}$ we will write $|w|$ for the length of w.

Here is a natural set of measure one for the relevant measures:

Definition 8. Suppose that $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ is a construction sequence for a symbolic system $\mathbb{K}$ with each $\mathcal{W}_n$ uniquely readable. Let S be the collection $x \in \mathbb{K}$ 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, $x \restriction [-a_m, b_m) \in \mathcal{W}_n$.

Note that S is a dense shift-invariant $\mathcal{G}_\delta$ set.

Lemma 9 ([11]). *Fix a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ for a symbolic system $\mathbb{K}$ in a finite language. Then:*

(1) *$\mathbb{K}$ is the smallest shift-invariant closed subset of $\Sigma^{\mathbb{Z}}$ such that for all n, and $w \in \mathcal{W}_n$, $\mathbb{K}$ has non-empty intersection with the basic open interval $\langle w\rangle \subset \Sigma^{\mathbb{Z}}$.*

(2) *Suppose that there is a unique invariant measure ν on $S \subseteq \mathbb{K}$, then ν is ergodic.*

(3) *(See* [12]*.) If ν is an invariant measure on $\mathbb{K}$ concentrating on S, then for ν-almost every s there is an N for all $n > N$, there are $a_n \leq 0 < b_n$ such that $s \restriction [a_n, b_n) \in \mathcal{W}_n$.*

Example 10. 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 is a summable sequence of positive numbers $\langle \epsilon_n : n \in \mathbb{N}\rangle$ and $\langle d_n : n \in \mathbb{N}\rangle$, where $d_n : \mathcal{W}_n \to (0, 1)$ 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 $w = w_i$

$$\left|\frac{f(w, w')}{q_{n+1}/q_n} - d_n(w)\right| < \frac{\epsilon_{n+1}}{q_n}. \tag{4.3}$$

It is shown in [11] that uniform construction sequences are uniquely ergodic. A special case of uniformity is *strong uniformity:* when each $w \in \mathcal{W}_n$ occurs exactly the same number of times in each $w' \in \mathcal{W}_{n+1}$. This property holds for the circular systems considered in [11] and that are used for the proof of the main theorem of this paper (Theorem 2).

4.2.4. Locations. Let $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ be a uniquely readable construction sequence and let ν be a shift invariant measure on S. For $s \in S$ and each n either $s(0)$ lies in a well-defined subword of s belonging to $\mathcal{W}_n$ or in a spacer of a subword of s belonging to some $\mathcal{W}_{n+k}$. By Lemma 9 for ν-almost all x and for all large enough n there is a unique k with $0 \leq k < q_n$ such that $s \restriction [-k, q_n - k) \in \mathcal{W}_n$.

Definition 11. Let $s \in S$ and suppose that for some $0 \leq k < q_n$, $s \restriction [-k, q_n - k) \in \mathcal{W}_n$. Define $r_n(s)$ to be the unique k with this property. We will call the interval $[-k, q_n - k)$ the *principal n-block* of s, and $s \restriction [-k, q_n - k)$ its *principal n-subword*. The sequence of r_n will be called the *location sequence of s*.

Thus $r_n(s) = k$ is saying that $s(0)$ *is the k-th symbol in the principal n-subword of s containing* 0. We can view the principal n-subword of s as being located on an interval I inside the principal $n + 1$-subword. Counting from the beginning of the principal $n + 1$-subword, the $r_{n+1}(s)$ position is located at the $r_n(s)$ position in I.

Remark 12. It follows immediately from the definitions that if $r_n(s)$ is well-defined and $n \leq m$, the $r_m(s)$-th position of the word occurring in the principal m-block of s is in the $r_n(s)$-th position inside the principal n-block of s.

Lemma 13. [12] *Suppose that $s, s' \in S$ and $\langle r_n(s) : n \geq N\rangle = \langle r_n(s') : n \geq N\rangle$ and for all $n \geq N$, s and s' have the same principal n-subwords. Then $s = s'$.*

Thus an element of s is determined by knowing any tail of the sequence

$$\langle r_n(s) : n \geq N\rangle$$

together with a tail of the principal subwords of s.

Remark 14. Here are some consequences of Lemma 13:

(1) Given a sequence $\langle u_n : M \leq n\rangle$ with $u_n \in \mathcal{W}_n$, if we specify which occurrence of u_n in u_{n+1} is the principal occurrence, then $\langle u_n : M \leq n\rangle$ determines an $s \in \mathbb{K}$ completely up to a shift k with $|k| \leq q_M$.

(2) A sequence $\langle r_n : N \leq n\rangle$ and sequence of words $w_n \in \mathcal{W}_n$ comes from an infinite word $s \in S$ if both r_n and $q_n - r_n$ go to infinity and that the r_{n+1} position in w_{n+1} is in the r_n position in a subword of w_{n+1} identical to w_n. (*Caveat*: just because $\langle r_n : N \leq n\rangle$ is the location sequence of some $s \in S$ and $\langle w_n : N \leq n\rangle$ is the sequence of principal subwords of some $s' \in S$, it does not follow that there is an $x \in S$ with location sequence $\langle r_n : N \leq n\rangle$ and sequence of subwords $\langle w_n : N \leq n\rangle$.)

(3) If $x, y \in S$ have the same principal n-subwords and $r_n(y) = r_n(x) + 1$ for all large enough n, then $y = \mathrm{sh}(x)$.

4.2.5. A note on inverses of symbolic shifts. We define operators we label $\mathrm{rev}(\cdot)$, and apply them in several contexts.

Definition 15. If x is in $\mathbb{K}$, define the reverse of x by setting $\mathrm{rev}(x)(k) = x(-k)$. For $A \subseteq \mathbb{K}$, define

$$\mathrm{rev}(A) = \{\mathrm{rev}(x) : x \in A\}.$$

If w is a word, let $\mathrm{rev}(w)$ to be the reverse of w sitting on the same interval. Explicitly, if $w : [a_n, b_n) \to \Sigma$ is the word, then $\mathrm{rev}(w) : [a_n, b_n) \to \Sigma$ and

$$\mathrm{rev}(w)(i) = w((a_n + b_n) - (i + 1)).$$

If $\mathcal{W}$ is a collection of words, $\mathrm{rev}(\mathcal{W})$ is the collection of reverses of the words in $\mathcal{W}$.

If $(\mathbb{K}, \mathrm{sh})$ is an arbitrary symbolic shift, then its inverse is $(\mathbb{K}, \mathrm{sh}^{-1})$. It will be convenient to have all of the shifts go in the same direction, thus:

Proposition 16. *The map ϕ sending x to $\mathrm{rev}(x)$ is a canonical isomorphism between $(\mathbb{K}, \mathrm{sh}^{-1})$ and $(\mathrm{rev}(\mathbb{K}), \mathrm{sh})$.*

Note that the notation $\mathbb{L}^{-1}$ stands for the system $(\mathbb{L}, \mathrm{sh}^{-1})$ and $\mathrm{rev}(\mathbb{L})$ for the system $(\mathrm{rev}(\mathbb{L}), \mathrm{sh})$.

4.3. Generic points

Let T be a measure preserving transformation from (X, τ, μ) to (X, τ, μ), where τ is a compact separable topology, and μ is a standard measure. Then a point $x \in X$ is *generic* for T if and only if for all $f \in C(X)$,

$$\lim_{N\to\infty} \left(\frac{1}{N}\right) \sum_{0}^{N-1} f(T^n(x)) = \int_X f(x)d\mu(x). \tag{4.4}$$

The Ergodic Theorem tells us that for a given f and ergodic T equation (4.4) holds for a set of μ-measure one. Intersecting over a countable dense set of $f \in C(X)$ gives a set of μ-measure one of generic points. For symbolic systems $\mathbb{K} \subseteq \Sigma^{\mathbb{Z}}$ the generic points are those x such that the μ-measure of all basic open intervals $\langle u \rangle_0$ is equal to the density of k such that u occurs in x at k.

4.4. Stationary codes and $\bar{d}$-distance

In this subection we briefly review a standard idea, that of a *stationary code*. A reader unfamiliar with this material who is interested in the proofs of the facts cited here should see [22].

Definition 17. Suppose that Σ is a countable language. A *code* of length $2N+1$ is a function $\Lambda : \Sigma^{[-N,N]} \to \Sigma$ (where $[-N, N]$ is the interval of integers starting at $-N$ and ending at N).

Given a code Λ, the *stationary code* determined by Λ is the function $\bar{\Lambda} : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$, where, given s

$$\bar{\Lambda}(s)(k) = \Lambda(s \upharpoonright [k - N, k + N]).$$

Let $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, \mathrm{sh})$ be a symbolic system. Given two codes Λ_0 and Λ_1 (not necessarily of the same length), define

$$D = \{s \in \Sigma^{\mathbb{Z}} : \overline{\Lambda}_0(s)(0) \neq \bar{\Lambda}_1(s)(0)\} \quad \text{and} \quad d(\Lambda_0, \Lambda_1) = \nu(D).$$

Then d is a semi-metric on the collection of codes. The following is a consequence of the Borel–Cantelli lemma.

Lemma 18. *Suppose that $\langle \Lambda_i : i \in \mathbb{N} \rangle$ is a sequence of codes such that*

$$\sum_i d(\Lambda_i, \Lambda_{i+1}) < \infty.$$

Then there is a shift-invariant Borel map $S : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ such that for ν-almost all s, $\lim_{i \to \infty} \overline{\Lambda}_i(s) = S(s)$.

A shift-invariant Borel map $S : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$, determines a factor $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, \mathrm{sh})$ of $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, \mathrm{sh})$ by setting $\mu = S^*\nu$. Hence a convergent sequence of stationary codes determines a factor of $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, \mathrm{sh})$.

Let Λ_0 and Λ_1 be codes. Define $\bar{d}(\bar{\Lambda}_0(s), \bar{\Lambda}_1(s))$ to be

$$\lim_{N \to \infty} \frac{|\{k \in [-N, N] : \bar{\Lambda}_0(s)(k) \neq \bar{\Lambda}_1(s)(k)\}|}{2N+1}.$$

More generally define the $\bar{d}$ metric on $\Sigma^{[a,b]}$ by setting

$$\bar{d}_{[a,b]}(x, y) = \frac{|\{k \in [a, b) : x(k) \neq y(k)\}|}{b - a}.$$

For $x, y \in \Sigma^{\mathbb{Z}}$, we set

$$\bar{d}(x, y) = \lim_{N \to \infty} \bar{d}_{[-N,N]}(x \upharpoonright [-N, N], y \upharpoonright [-N, N]),$$

provided this limit exists.

To compute distances between codes we will use the following application of the Ergodic Theorem.

Lemma 19. *Suppose that ν is ergodic. Let Λ_0 and Λ_1 be codes. Then for almost all $s \in \Sigma^{\mathbb{Z}}$,*

$$d(\Lambda_0, \Lambda_1) = \bar{d}(\bar{\Lambda}_0(s), \bar{\Lambda}_1(s)).$$

The next proposition is used to study alleged isomorphisms between measure preserving transformations. We again refer the reader to [22] for a proof.

Proposition 20. *Suppose that $\mathbb{K}$ and $\mathbb{L}$ are symbolic systems and $\phi : \mathbb{K} \to \mathbb{L}$ is a factor map. Let $\epsilon > 0$. Then there is a code Λ such that for almost all $s \in \mathbb{K}$,*

$$\bar{d}(\bar{\Lambda}(s), \phi(s)) < \epsilon. \tag{4.5}$$

To show that equation (4.5) cannot hold (and hence show that $\mathbb{L}$ is not a factor of $\mathbb{K}$), we will want to view $\bar{\Lambda}(s)$ as limits of Λ-images of large blocks of the form $s \upharpoonright [a, b]$ with $a < 0 < b$. There is an ambiguity in doing this: if the code Λ has length $2N + 1$, it does not make sense to apply it to $s \upharpoonright [k - N, k + N]$ for $k \in [a, a + 2N]$ or $k \in [b - 2N, b]$. However if $b - a$ is quite large with respect to N, then filling in the values for $\Lambda(s \upharpoonright [k - N, k + N])$ arbitrarily as k ranges over these initial and final intervals makes a negligible difference to the $\bar{d}$-distances of the result. In particular, if $\bar{d}(\bar{\Lambda}(s), \phi(s)) < \epsilon$, then for all large enough $a, b \in \mathbb{N}$, we have

$$\bar{d}_{[-a,b]}(\bar{\Lambda}(s \upharpoonright [-a, b]), \phi(s) \upharpoonright [-a, b]) < \epsilon,$$

no matter how we fill in the ambiguous portion.

The general phenomenon of ambiguity or disagreement at the beginning and end of large intervals is referred to by the phrase *end effects*. Because the end effects are usually negligible on large intervals we will often neglect them when computing $\bar{d}$ distances.

The next proposition is standard:

Proposition 21. *Suppose that $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, \mathrm{sh})$ is an ergodic symbolic system and that $\langle T_n : n \in \mathbb{N}\rangle$ is a sequence of functions from $\Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ that commute with the shift. Then the following are equivalent:*

(1) *The sequence $\langle T_n\rangle$ converges to S in the weak topology.*

(2) $\nu(\{s : T_n(s)(0) \neq S(s)(0)\}) \to 0$.

(3) *For ν-almost all s, $\bar{d}(T_n(s), S(s)) \to 0$.*

(4) *For some ν-generic s, for all $\gamma > 0$ we can find an N for all $n \geq N$, for all large enough a, b, the distance $\bar{d}(T_n(s) \upharpoonright [-a, b), S(s) \upharpoonright [-a, b)) < \gamma$.*

We finish with a remark that we will use in several places:

Remark 22. If w_1 and w_2 are words in a language Σ defined on an interval I and $J \subset I$ with $\frac{|J|}{|I|} \geq \delta$, then $\bar{d}_I(w_1, w_2) \geq \delta \bar{d}_J(w_1, w_2)$.

4.5. Rotations of the circle

Many of the arguments in this paper are based on an understanding of rational approximations to rotations of the circle. It is usually convenient to adopt additive notation and work on the unit interval $[0, 1)$, but this introduces ambiguities. Fix an $\alpha \in \mathbb{R}$. We use the symbol $\mathcal{R}_\alpha$ in two ways. The first way is that

$$\mathcal{R}_\alpha : S^1 \to S^1$$

by rotating the circle by $\alpha * 2\pi$ radians. The second, equivalent, way is that

$$\mathcal{R}_\alpha : [0, 1) \to [0, 1)$$

and is given by the formula

$$x \mapsto x + \alpha \bmod 1.$$

We note in both cases that we are really concerned with $[\alpha]$ (mod 1).

4.6. Descriptive set theory basics

Let X and Y be Polish spaces and $A \subseteq X, B \subseteq Y$.[11] A function $f : X \to Y$ *reduces* A to B if and only if for all $x \in X$,

$$x \in A \text{ if and only if } f(x) \in B.$$

For this definition to have content there must be some definability restriction on f. The relevant restrictions for this paper are either that *f is a Borel function* (i.e. the inverse image of an open set is Borel) or that *f is a continuous function* (i.e. the inverse image of an open set is open). The latter is clearly a stronger condition. If B is Borel and f is a Borel reduction, then A is clearly Borel. Taking the contrapositive, if A is *not* Borel, then B is not. If A is Borel (resp. continuously) reducible to B, we will write $A \preceq_B B$ (resp. $A \preceq_c B$). Both $\preceq_B$ and $\preceq_c$ are clearly pre-partial-orderings.[12]

If $\mathcal{S}$ is a collection of pairs (A, X) and $(B, Y) \in \mathcal{S}$, then B is *$\mathcal{S}$-complete* for Borel reductions (resp. continuous reductions) if and only if every $(A, X) \in \mathcal{S}$ is Borel reducible (resp. continuously reducible) to (B, Y). Being complete is interpreted as being at least as complicated as each set in $\mathcal{S}$.

For this to be useful there must be examples of sets that are not Borel. If X is a Polish space and $B \subseteq X$, then B is *analytic* (Σ^1_1) if and only if it the continuous image of a Borel subset of a Polish space. This is equivalent to there being a Polish space Y and a Borel set $C \subseteq X \times Y$ such that B is the projection to the X-axis of C.

Correcting a famous mistake of Lebesgue, Suslin proved that there are analytic sets that are not Borel. It follows immediately that complete analytic sets are not Borel. This paper uses a canonical example of such a set.

Let $\langle \sigma_n : n \in \mathbb{N} \rangle$ be an enumeration of $\mathbb{N}^{<\mathbb{N}}$, the finite sequences of natural numbers. Using this enumeration subsets $S \subseteq \mathbb{N}^{<\mathbb{N}}$ can be identified with functions

$$\chi_S : \mathbb{N} \to \{0, 1\}.$$

[11] The ideas in this subsection are just summaries, they are exposited in [5] and [19].

[12] The reader should be aware that this is a different notion than the notion of a reduction of *equivalence relations*.

A *tree* is a set $\mathcal{T} \subseteq \mathbb{N}^{<\mathbb{N}}$ such that if $\tau \in \mathcal{T}$ and σ is an initial segment of τ, then $\sigma \in \mathcal{T}$. The set $\{\chi_{\mathcal{T}} : \mathcal{T} \text{ is a tree}\}$ is a closed subset of $\{0,1\}^{\mathbb{N}}$, hence a Polish space with the induced topology. We call the resulting space $\mathcal{T}rees$. (In the sequel we will not always distinguish between $\mathcal{T}$ and $\chi_{\mathcal{T}}$.)

Because the topology on the space of trees is the "finite information" topology, inherited from the product topology on $\{0,1\}^{\mathbb{N}}$, the following characterizes continuous maps defined on $\mathcal{T}rees$.

Proposition 23. *Let Y be a topological space and $f : \mathcal{T}rees \to Y$. Then f is continuous if and only if for all open $O \subseteq Y$ and all $\mathcal{T}$ with $f(\mathcal{T}) \in O$ there is an $M \in \mathbb{N}$ for all $\mathcal{T}' \in \mathcal{T}rees$:*

$$\text{if } \mathcal{T} \cap \{\sigma_n : n \le M\} = \mathcal{T}' \cap \{\sigma_n : n \le M\}, \text{ then } f(\mathcal{T}') \in O.$$

An infinite branch through T is a function $f : \mathbb{N} \to \mathbb{N}$ such that for all $n \in \mathbb{N}$, $f \upharpoonright \{0, 1, 2, \ldots, n-1\} \in T$. A tree T is *ill-founded* if and only if it has an infinite branch.

The following theorem is classical; proofs can be found in [19] and [20].

Fact 24. Let $\mathcal{T}rees$ be the space of trees. Then:

(1) The collection of ill-founded trees is a complete analytic subset of $\mathcal{T}rees$.

(2) The collection of trees that have at least two distinct infinite branches is a complete analytic subset of $\mathcal{T}rees$.

The main results of this paper (Theorem 2 and Corollary 3) are proved by reducing the sets mentioned in Fact 24 to conjugate pairs of diffeomorphisms and concluding that the sets of conjugate pairs is complete analytic – so not Borel.

5. Odometer and circular systems

Two types of symbolic shifts play central roles for the proofs of the main theorem, the *odometer based* and the *circular* systems. Most of the material in this section appears in [12] in more detail and is reviewed here without proof.

5.1. Odometer based systems

We now define the class of *odometer based systems*. In a sequel to this paper ([13]), we prove that these are exactly the finite entropy transformations that have non-trivial odometer factors. We recall the definition of an odometer transformation. Let $\langle k_n : n \in \mathbb{N}\rangle$ be a sequence of natural numbers greater than or equal to 2. Let

$$O = \prod_{n=0}^{\infty} \mathbb{Z}/k_n\mathbb{Z}$$

be the $\langle k_n\rangle$-adic integers. Then O naturally has a compact abelian group structure and hence carries a Haar measure μ. The set O becomes a measure preserving system $\mathcal{O}$ by defining $T : O \to O$ to be addition by 1 in the $\langle k_n \rangle$-adic integers. Concretely, this is the map that "adds one to $\mathbb{Z}/k_0\mathbb{Z}$ and carries right". Then T is an invertible transformation that preserves the Haar measure μ on $\mathcal{O}$. Let $K_n = k_0 * k_1 * k_2 * \cdots * k_{n-1}$.

The following results are standard:

Lemma 25. *Let $\mathcal{O}$ be an odometer system. Then:*

(1) *$\mathcal{O}$ is ergodic.*

(2) *The map $x \mapsto -x$ is an isomorphism between $(O, \mathcal{B}, \mu, T)$ and $(O, \mathcal{B}, \mu, T^{-1})$.*

(3) *Odometer maps are transformations with discrete spectrum and the eigenvalues of the associated linear operator are the K_n-th roots of unity ($n > 0$).*

Any natural number $a < K_j$ can be uniquely written as

$$a = a_0 + a_1 k_0 + a_2(k_0 k_1) + \cdots + a_j(k_0 k_1 k_2 \ldots k_{j-1})$$

for some sequence of natural numbers $a_0, a_1, \ldots, a_j$ with $0 \le a_j < k_j$.

Lemma 26. *Suppose that $\langle r_n : n \in \mathbb{N} \rangle$ is a sequence of natural numbers with $0 \le r_n < k_0 k_1 \ldots k_n$ and $r_n \equiv r_{n+1} \bmod (k_0 k_1 \ldots k_n)$. Then there is a unique element $x \in O$ such that $r_n = x(0) + x(1)k_0 + \cdots + x(n)(k_0 k_1 \ldots k_{n-1})$ for each n.*

We now define the collection of symbolic systems that have odometer maps as their timing mechanism. This timing mechanism can be used to parse typical elements of the symbolic system.

Definition 27. Let $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ be a uniquely readable construction sequence with the properties that $\mathcal{W}_0 = \Sigma$ and for all n, $\mathcal{W}_{n+1} \subseteq (\mathcal{W}_n)^{k_n}$ for some k_n. The associated symbolic system will be called an *odometer based system*.

Thus odometer based systems are those built from construction sequences

$$\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$$

such that the words in $\mathcal{W}_{n+1}$ are concatenations of words in $\mathcal{W}_n$ of a fixed length k_n. The words in $\mathcal{W}_n$ all have length K_n and the words u_i in equation (4.1) are all the empty words.

Equivalently, an odometer based transformation is one that can be built by a cut-and-stack construction using no spacers. An easy consequence of the definition is that for odometer based systems, for all $s \in S$ and for all $n \in \mathbb{N}$, $r_n(s)$ exists.[13]

The next lemma justifies the terminology.

Lemma 28. *Let $\mathbb{K}$ be an odometer based system with each $\mathcal{W}_{n+1} \subseteq (\mathcal{W}_n)^{k_n}$. Then there is a canonical factor map*

$$\pi : S \to \mathcal{O},$$

where $\mathcal{O}$ is the odometer system determined by $\langle k_n : n \in \mathbb{N} \rangle$.

[13] S is defined in Definition 8.

Proof. For each $s \in S$, for all n, $r_n(s)$ is defined and both r_n and $k_n - r_n$ go to infinity. By Lemma 26, the sequence $\langle r_n(s) : n \in \mathbb{N}\rangle$ defines a unique element $\pi(s)$ in $\mathcal{O}$. It is easily checked that π intertwines sh and T. ■

Heuristically, the odometer transformation $\mathcal{O}$ parses the sequences s in $S \subseteq \mathbb{K}$ by indicating where the words constituting s begin and end. Shifting s by one unit shifts this parsing by one. We can understand elements of S as being an element of the odometer with words in $\mathcal{W}_n$ filled in inductively.

The following remark is useful when studying the canonical factor of the inverse of an odometer based system.

Remark 29. If $\pi : \mathbb{L} \to \mathcal{O}$ is the canonical factor map, then the function $\pi : L \to O$ is also factor map from $(\mathbb{L}, \mathrm{sh}^{-1})$ to $\mathcal{O}^{-1}$ (i.e. O with the operation "-1"). If $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ is the construction sequence for $\mathbb{L}$, then $\langle \mathrm{rev}(\mathcal{W}_n) : n \in \mathbb{N}\rangle$ is a construction sequence for $\mathrm{rev}(\mathbb{L})$. If $\phi : \mathbb{L}^{-1} \to \mathrm{rev}(\mathbb{L})$ is the canonical isomorphism given by Proposition 16, then Lemma 25 tells us that the projection of ϕ to a map $\phi^\pi : \mathcal{O} \to \mathcal{O}$ is given by $x \mapsto -x$.

The following is proved in [12]:

Proposition 30. *Let $\mathbb{K}$ be an odometer based system and suppose that ν is a shift invariant measure. Then ν concentrates on S.*

5.2. *Circular systems*

We now define *circular systems*. In [11] it is shown that the strongly uniform circular systems give symbolic characterizations of certain smooth diffeomorphisms defined by the Anosov–Katok method of conjugacies.

These systems are called *circular* because they are related to the behavior of rotations by a convergent sequence of rationals $\alpha_n = p_n/q_n$. The rational rotation by p/q permutes the $1/q$ intervals of the circle cyclically in a manner that the interval $[i/q, (i+1)/q)$ occurs in position[14]

$$j_i =_{\mathrm{def}} p^{-1} i \pmod{q}.$$

The operation $\mathcal{C}$ which we are about to describe models the relationship between rotations by p/q and p'/q' when p'/q' is very close to p/q.

Let k, l, p, q be positive natural numbers with $p < q$ relatively prime. For $0 \le i < q$, setting

$$j_i \equiv_q (p)^{-1} i \tag{5.1}$$

with $j_i < q$, it is easy to verify that

$$q - j_i = j_{q-i}. \tag{5.2}$$

For notational convenience later we set $j_q = q$.

[14]We assume that p and q are relatively prime and the exponent -1 indicates the multiplicative inverse modulo q.

Let Σ be a non-empty set such that neither b nor e belongs to Σ and let $w_0, \ldots, w_{k-1}$ be words in $\Sigma \cup \{b, e\}$. Define

$$\mathcal{C}(w_0, w_1, w_2, \ldots, w_{k-1}) = \prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (b^{q-j_i} w_j^{l-1} e^{j_i}). \tag{5.3}$$

We note that the product symbol Π is repeated concatenation as is the exponent. If w is a word, then w^0 is the empty string, $w^1 = w$, $w^2 = ww$ and so forth. The formula in equation (5.3) is a concatenation of q words, each of which is itself, a concatenation of k words. The words inside the parenthesis in equation (5.3) start with $q - j_i$ letters b, followed by concatenating $l - 1$ many words w, followed by concatenating j_i many letters e. Written with parenthesis

$$\mathcal{C}(w_0, w_1, w_2, \ldots, w_{k-1}) = \prod_{i=0}^{q-1} \left(\prod_{j=0}^{k-1} ((b^{q-j_i})(w_j^{l-1})(e^{j_i})) \right). \tag{5.4}$$

Informally, the i-th term, $\prod_{j=0}^{k-1} (b^{q-j_i} w_j^{l-1} e^{j_i})$ can be written as a block of $q - j_i$ letters b followed by w_0 concatenated with itself $l - 1$ times, followed by a block of j_i many letters e, followed by a block of $q - j_i$ letters b followed by w_1 concatenated with itself $l - 1$ times followed by a block of j_i letters e and so forth, ending with a block of w_{k-1} repeated $l - 1$ times followed by e repeated j_i many times

$$(bbb \ldots)(w_0 w_0 \ldots)(ee \ldots e)(bb \ldots b)(w_1 w_1 \ldots w_1)(ee \ldots e) \ldots$$
$$\ldots (bb \ldots b)(w_{k-1} w_{k-1} w_{k-1} \ldots w_{k-1})(ee \ldots e).$$

Remark 31. We make the following observations.

- Suppose that each w_i has length q. Then the length of $\mathcal{C}(w_0, w_1, \ldots, w_{k-1})$ is klq^2.
- For each occurrence of an e in $\mathcal{C}(w_0, \ldots, w_{k-1})$ there is an occurrence of b to the left of it.
- Suppose that $n < m$ and b occurs at n and e occurs at m and neither occurrence is in a w_i. Then there must be some w_i occurring between n and m.
- Words constructed with $\mathcal{C}$ are uniquely readable.

The $\mathcal{C}$ operation is used to build a collection of symbolic shifts. *Circular systems* will be defined using a sequence of natural number parameters k_n and l_n that is fundamental to the version of the Anosov–Katok construction presented in [18].

Fix an arbitrary sequence of positive natural numbers $\langle k_n : n \in \mathbb{N} \rangle$. Let $\langle l_n : n \in \mathbb{N} \rangle$ be an increasing sequence of natural numbers such that

Numerical Requirement 1. One has $l_0 > 20$ and

$$\sum_{k \geq n} \frac{1}{l_k} < \frac{1}{l_{n-1}}.$$

From the k_n and l_n we define sequences of numbers: $\langle p_n, q_n, \alpha_n : n \in \mathbb{N}\rangle$. Begin by letting $p_0 = 0$ and $q_0 = 1$ and inductively set

$$q_{n+1} = k_n l_n {q_n}^2 \tag{5.5}$$

(thus $q_1 = k_0 l_0$) and take

$$p_{n+1} = p_n q_n k_n l_n + 1. \tag{5.6}$$

Then clearly p_{n+1} is relatively prime to q_{n+1}.[15]

By setting $\alpha_n = p_n/q_n$, it is easy to check that there is an irrational α such that the sequence α_n converges rapidly to α.

Definition 32. A sequence of integers $\langle k_n, l_n : n \in \mathbb{N}\rangle$ such that $k_n \geq 2$, $\sum 1/l_n < \infty$ will be called a *circular coefficient sequence*.

Let Σ be a non-empty finite or countable alphabet. Build collections of words $\mathcal{W}_n$ in $\Sigma \cup \{b, e\}$ by induction as follows:

- Fix a circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$.
- Set $\mathcal{W}_0 = \Sigma \cup \{b, e\}$.
- Having built $\mathcal{W}_n$ choose a set $P_{n+1} \subseteq (\mathcal{W}_n)^{k_n}$ and form $\mathcal{W}_{n+1}$ by taking all words of the form $\mathcal{C}(w_0, w_1, \ldots, w_{k_n-1})$ with $(w_0, \ldots, w_{k_n-1}) \in P_{n+1}$.[16]

We call the elements of P_{n+1} *prewords*. The $\mathcal{C}$ operator automatically creates uniquely readable words, however we will need a stronger unique readability assumption for our definition of circular systems.

Strong Unique Readability Assumption. Let $n \in \mathbb{N}$, and view $\mathcal{W}_n$ as a collection Λ_n of letters. Then each element of P_{n+1} can be viewed as a word with letters in Λ_n. In the alphabet Λ_n, each $w \in P_{n+1}$ is uniquely readable.

Definition 33. A construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ will be called *circular* if it is built in this manner using the $\mathcal{C}$-operators, a circular coefficient sequence and each P_{n+1} satisfies the strong unique readability assumption.

Definition 34. A symbolic shift $\mathbb{K}$ built from a circular construction sequence will be called a *circular system*.

Notation. We will often write $\mathbb{K}^c$ and $\langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$ to emphasize that we are building circular systems and circular construction sequences. Circular words will often be denoted w^c for emphasis.

Definition 35. Suppose that $w = \mathcal{C}(w_0, w_1, \ldots, w_{k-1})$. Then w consists of blocks of w_i repeated $l - 1$ times, together with some letters b and e that are not in the words w_i. The

[15] p_n and q_n being relatively prime for $n \geq 1$, allows us to define the integer j_i in equation (5.1). For $q_0 = 1$, $\mathbb{Z}/q_0\mathbb{Z}$ has one element, [0], so we set ${p_0}^{-1} = p_0 = 0$.

[16] Passing from $\mathcal{W}_n$ to $\mathcal{W}_{n+1}$, use $\mathcal{C}$ with parameters $k = k_n$, $l = l_n$, $p = p_n$ and $q = q_n$ and take $j_i = (p_n)^{-1} i$ modulo q_n. By Remark 31, the length of each of the words in $\mathcal{W}_{n+1}$ is q_{n+1}.

interior of w is the portion of w in the words w_i. The remainder of w consists of blocks of the form b^{q-j_i} and e^{j_i}. We call this portion the *boundary* of w.

In a block of the form w_j^{l-1} the first and last occurrences of w_j will be called the *boundary* occurrences of the block w_j^{l-1}. The other occurrences will be the *interior* occurrences.

While the boundary consists of sections of w made up of letters b and e, not all letters b and e occurring in w are in the boundary, as they may be part of a power w_i^{l-1}.

The boundary of w constitutes a small portion of the word:

Lemma 36. *Suppose that $w = \mathcal{C}(w_0, w_1, \ldots, w_{k-1})$ and each w_i has length q. Then the proportion of the word w that belongs to its boundary is $1/l$. Moreover, the proportion of the word that is within q letters of boundary of w is $3/l$.*

Proof. The length of w is klq^2. The boundary portions are $q * k * q$ long. The number of letters within q letters of the boundary is $q * k * 3 * q$. ■

Remark 37. Let $v_0, \ldots, v_{k-1}$ and $w_0, \ldots, w_{k-1}$ be sequences of words of length q. The boundary portions of $\mathcal{C}(v_0, \ldots, v_{k-1})$ and $\mathcal{C}(w_0, \ldots, w_{k-1})$ occur in the same positions and by Lemma 36 have proportion $1/l$ of the length. Since all of the words v_i and w_i have the same length and the same multiplicity in the circular words, we see that

$$\begin{aligned}&\bar{d}(\mathcal{C}(v_0, \ldots, v_{k-1}), \mathcal{C}(w_0, \ldots, w_{k-1}))\\ &\quad \geq \left(1 - \frac{1}{l}\right)\bar{d}(v_0 v_1 v_2 \ldots v_{k-1}, w_0 w_1 \ldots w_{k-1}),\end{aligned}$$

where $v_0 v_1 v_2 \ldots v_{k-1}$ and $w_0 w_1 \ldots w_{k-1}$ are the concatenations of the various words.[17]

For proofs of the next lemma see [11, Lemma 20] and [12].

Lemma 38. *Let $\mathbb{K}^c$ be a circular system and let ν be a shift-invariant measure on $\mathbb{K}^c$. Then the following are equivalent:*

(1) *ν has no atoms.*

(2) *ν concentrates on the collection of $s \in \mathbb{K}^c$ such that $\{i : s(i) \notin \{b, e\}\}$ is unbounded in both $\mathbb{Z}^-$ and $\mathbb{Z}^+$.*

(3) *ν concentrates on S.*

If $\mathbb{K}^c$ is a uniform circular system (Example 10), then there is a unique invariant measure concentrating on S.

Moreover, there are only two ergodic invariant measures with atoms: the one concentrating on the constant sequence $\vec{b}$ and the one concentrating on $\vec{e}$.

Remark 39. If $\mathbb{K}^c$ is circular and $s \in \mathbb{K}^c$ has a principal n-subword and $m > n$, then s has a principal m-subword.

[17] Equality holds, a fact we will not use.

5.3. *An explicit description of* rev($\mathbb{K}^c$)

The symbolic system $\mathbb{K}^c$ is built by an operation $\mathcal{C}$ applied to collections of words. The system rev($\mathbb{K}^c$) is built by a similar operation applied to the reverse collections of words. In analogy to equation (5.3), we define $\mathcal{C}^r$ as follows.

Definition 40. Suppose that $w_0, w_1, \ldots, w_{k-1}$ are words in a language Σ. Given coefficients p, q, k, l with p and q relatively prime, let $j_i \equiv_q (p^{-1})i$ with $0 \le j_i < q$. Define

$$\mathcal{C}^r(w_0, w_1, w_2, \ldots, w_{k-1}) = \prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (e^{q-j_{i+1}}({w_{k-j-1}}^{l-1})b^{j_{i+1}}). \tag{5.7}$$

From equation (5.3), a $w \in \mathcal{W}^c_{n+1}$ is of the form $\mathcal{C}(w_0, \ldots, w_{k_n-1})$:

$$w = \prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (b^{q-j_i} w_j^{l-1} e^{j_i}), \tag{5.8}$$

where $q = q_n, k = k_n, l = l_n$ and $j_i \equiv_{q_n} (p_n)^{-1} i$ with $0 \le j_i < q_n$. By examining this formula, we see that

$$\mathrm{rev}(w) = \prod_{i=1}^{q} \prod_{j=1}^{k} e^{j_{q-i}} \,\mathrm{rev}(w_{k-j})^{l-1} b^{q-j_{q-i}}.$$

Applying the identity in formula (5.2) and recalling that we take $j_q = q$, so $q - j_q = 0$, we see that this can be rewritten as

$$\mathrm{rev}(w) = \prod_{i=1}^{q} \prod_{j=1}^{k} (e^{q-j_i} \,\mathrm{rev}(w_{k-j})^{l-1} b^{j_i}). \tag{5.9}$$

Thus

$$\mathrm{rev}(w) = \mathcal{C}^r(\mathrm{rev}(w_0), \mathrm{rev}(w_1), \ldots, \mathrm{rev}(w_{k-1})). \tag{5.10}$$

In particular, if $\langle \mathcal{W}^c_n : n \in \mathbb{N} \rangle$ is a construction sequence of a circular system $\mathbb{K}^c$, then rev($\mathcal{W}^c_{n+1}$) is the collection

$$\{\mathcal{C}^r(\mathrm{rev}(w_0), \mathrm{rev}(w_1), \ldots, \mathrm{rev}(w_{k_n-1})) : w_0 w_1 \ldots w_{k_n-1} \in P_n\}$$

and $\langle \mathrm{rev}(\mathcal{W}^c_n) : n \in \mathbb{N} \rangle$ is a construction sequence for rev($\mathbb{K}^c$).

5.4. *Understanding the words*

The words used to form circular transformations have quite specific combinatorial properties. Fix a sequence $\langle \mathcal{W}^c_n : n \in \mathbb{N} \rangle$ defining a circular system. Each $u \in \mathcal{W}^c_{n+1}$ has three *subscales*.

- **Subscale 0,** the scale of the individual powers of $w \in \mathcal{W}_n^c$ of the form w^{l-1}. We call each such occurrence of a w^{l-1} a 0-*subsection*.
- **Subscale 1,** the scale of each term in the product $\prod_{j=0}^{k-1}(b^{q-j_i} w_j^{l-1} e^{j_i})$ that has the form $(b^{q-j_i} w_j^{l-1} e^{j_i})$. We call these terms 1-*subsections*.
- **Subscale 2,** the scale of each term of $\prod_{i=0}^{q-1}(\prod_{j=0}^{k-1}(b^{q-j_i} w_j^{l-1} e^{j_i}))$ that has the form $\prod_{j=0}^{k-1}(b^{q-j_i} w_j^{l-1} e^{j_i})$. We call these terms 2-*subsections*.

Summary. We have

Whole word:	$\prod_{i=0}^{q-1}\prod_{j=0}^{k-1}(b^{q-j_i} w_j^{l-1} e^{j_i})$,
2-subsection:	$\prod_{j=0}^{k-1}(b^{q-j_i} w_j^{l-1} e^{j_i})$,
1-subsection:	$(b^{q-j_i} w_j^{l-1} e^{j_i})$,
0-subsection:	w_j^{l-1}.

For $m \leq n$, we will discuss "m-subwords" of a word w. These will be subwords that lie in $\mathcal{W}_m^c$, the m-th stage of the construction sequence. We will use "m-block" to mean the location of the m-subword.

Lemma 41. *Let $w = \mathcal{C}(w_0, \ldots, w_{k_n-1})$ for some n and let $q = q_n$, $k = k_n$, $l = l_n$. View $w : \{0, 1, 2 \ldots, klq^2 - 1\} \to \Sigma \cup \{b, e\}$.*

(1) *If m_0 and m_1 are such that $w(m_0)$ and $w(m_1)$ are at the beginning of n-subwords in the same 2-subsection, then $m_0 \equiv_q m_1$.*

(2) *If m_0 and m_1 are such that $w(m_0)$ is the beginning of an n-subword occurring in a 2-subsection $\prod_{j=0}^{k-1}(b^{q-j_i} w_j^{l-1} e^{j_i})$ and $w(m_1)$ is the beginning of an n-subword occurring in the next 2-subsection $\prod_{j=0}^{k-1}(b^{q-j_{i+1}} w_j^{l-1} e^{j_{i+1}})$, then $m_1 - m_0 \equiv_q -j_1$.*

Proof. To see the first point, the indices of the beginnings of n-subwords in the same 2-subsection differ by multiples of q coming from powers of a w_j and intervals of w of the form $b^{q-j_i} e^{j_i}$.

To see the second point, let u and v be consecutive 2-subsections. In view of the first point it suffices to consider the last n-subword of u and the first n-subword of v. These sit on either side of an interval of the form $e^{j_i} b^{q-j_{i+1}}$. Since

$$j_i + q - j_{i+1} \equiv_q (p)^{-1} i - p^{-1}(i+1) \equiv_q -p^{-1} \equiv_q -j_1,$$

we see that

$$m_0 - m_1 = q + j_i + q - j_{i+1} \equiv_q -j_1. \qquad \blacksquare$$

Assume that $u \in \mathcal{W}_{n+1}^c$ and $v \in \mathcal{W}_{n+1}^c \cup \mathrm{rev}(\mathcal{W}_{n+1}^c)$ and v is shifted with respect to u. On the overlap of u and v, the 2-subsections of u split each 2-subsection of v into either one or two pieces. Since the 2-subsections all have the same length, the number of pieces in the splitting and the size of each piece is constant across the overlap except perhaps at the two ends of the overlap. If u splits a 2-subsection of v into two pieces, then we call the leftmost piece of the pair the even piece and the rightmost the odd piece.

If v is shifted only slightly, it can happen that either the even piece or the odd piece does not contain even one entire 1-subsection. In this case we will say that the split is *trivial on the left* or *trivial on the right*.

Lemma 42. *Assume that $u \in \mathcal{W}^c_{n+1}$ and $v \in \mathcal{W}^c_{n+1} \cup \mathrm{rev}(\mathcal{W}^c_{n+1})$ and v is shifted with respect to u. Suppose that the 2-subsections of u divide the 2-subsections of v into two non-trivial pieces. Then:*

(1) *The boundary portion of u occurring between each consecutive pair of 2-subsections of u completely overlaps at most one n-subword of v.*

(2) *There are two numbers s and t such that the positions of the 0-subsections of v in even pieces are shifted relative to the 0-subsections of u by s and the positions of the 0-subsections of v in odd pieces are shifted relative to the 0-subsections of u by t. Moreover, $s \equiv_q t - j_1$.*

Proof. This follows easily from Lemma 41. ■

In the case where the split is trivial Lemma 42 holds with just one coefficient, s or t. A special case of Lemma 42 that we will use is:

Lemma 43. *Assume that $u \in \mathcal{W}^c_{n+1}$ and $v \in \mathcal{W}^c_{n+1} \cup \mathrm{rev}(\mathcal{W}^c_{n+1})$ and v is shifted with respect to u. Suppose that the 2-subsections of u divide the 2-subsections of v into two pieces and that for some occurrence of a n-subword in an even (resp. odd) piece is lined up with an occurrence of some n-subword in u. Then every occurrence of a n-subword in an even (resp. odd) piece of v is either*

(a) *lined up with some n-subword of u or*

(b) *lined up with a section of a 2-subsection that has the form $e^{j_i} b^{q-j_i}$.*

Moreover, no n-subword in an odd (resp. even) piece of v is lined up with a n-subword in u.

5.5. Full measure sets for circular systems

Fix a sequence $\langle \varepsilon_n : n \in \mathbb{N} \rangle$ such that the following hold:

Numerical Requirement 2. $\langle \varepsilon_n : n \in \mathbb{N} \rangle$ is a decreasing sequence of numbers in $[0, 1)$ such that $6 \sum_{n>N} \varepsilon_n < \varepsilon_N$.

From Lemma 36, the boundary of a word $w_n \in \mathcal{W}_n$ has proportion $1/l_n$. Hence Numerical Requirement 1 implies that for all choices $\langle w_n : n \in \mathbb{N} \rangle$ with $w_n \in \mathcal{W}_n$, the sum of the proportion of the boundary sections of the words w_n is finite.

Definition 44. Let:

(1) E_n be the collection of $s \in S$ such that either s does not have a principal n-block or $s(0)$ is in the boundary of the principal n-block of s,

(2) $E^0_n = \{s : s(0)$ is in the first or last $\varepsilon_n l_n$ copies of w in a power of the form $w^{l_n - 1}$, where $w \in \mathcal{W}^c_n\}$,

(3) $E_n^1 = \{s : s(0)$ is in the first or last $\varepsilon_n k_n$ 1-subsections of the 2-subsection in which $s(0)$ is located$\}$,

(4) $E_n^2 = \{s : s(0)$ is in the first or last $\varepsilon_n q_n$ 2-subsections of its principal $n+1$-block$\}$.

Lemma 45. *Assume Numerical Requirements 1 and 2. Let ν be a shift-invariant measure on $S \subseteq \mathbb{K}^c$, where $\mathbb{K}^c$ is a circular system. Then:*

(1) *One has*

$$\sum_n \nu(E_n) < \infty.$$

(2) *For $i = 0, 1, 2$,*

$$\sum_n \nu(E_n^i) < \infty.$$

Proof. By the Ergodic Theorem we have $\nu(E_n) < 1/l_n$, and for $i = 0, 1, 2$, $\nu(E_n^i) < \varepsilon_n$. The result then follows by the summability of $1/l_n$ and $1/\varepsilon_n$. ∎

In particular, we see:

Corollary 46. *For ν-almost all s there is an $N = N(s)$ such that for all $n > N$,*

(1) *$s(0)$ is in the interior of its principal n-block,*

(2) *for $i = 0, 1, 2$, $s \notin E_n^i$.*

In particular, for almost all s and all large enough n,

(3) *if $s \restriction [-r_n(s), -r_n(s) + q_n) = w$, then*

$$s \restriction [-r_n(s) - q_n, -r_n(s)) = s \restriction [-r_n(s) + q_n, -r_n + 2q_n) = w,$$

(4) *$s(0)$ is not in a string of the form $w_0^{l_n-1}$ or $w_{k_n-1}^{l_n-1}$.*

Proof. Apply the Borel–Cantelli lemma using the previous lemma. ∎

The elements s of S such that some shift $\mathrm{sh}^k(s)$ fails one of conclusions (1)–(4) of Corollary 46 form a measure zero set. Consequently, we work on those elements of S whose whole orbit satisfies the conclusions of Corollary 46. Note however that for $t = \mathrm{sh}^k(s)$, the $N(t)$ in Corollary 46, depends on k.

Definition 47. We will call n *mature* for s (or say that s *is mature at stage* n) iff n is so large that $s \notin E_m \cup \bigcup_{0 \le i \le 2} E_m^i$ for all $m \ge n$.

If s is mature at stage n, then s is mature at stage $n+1$. Moreover, if $\mathrm{sh}^k(s)$ has the same principal n-block as s does, then $\mathrm{sh}^k(s)$ is mature if and only if $s(k)$ is not in the boundary portion of the principal n-block.

Numerical Requirement 3. The following hold:

$$\varepsilon_n k_n \to \infty,$$
$$\varepsilon_n l_n \to \infty,$$
$$\varepsilon_n q_n \to \infty.$$

Definition 48. We will use the symbol ∂_n in multiple equivalent ways. If $s \in S$ or $s \in \mathcal{W}_m^c$, define $\partial_n = \partial_n(s) \subseteq \mathbb{Z}$ to be the collection of $i \in \mathbb{Z}$ such that $\mathrm{sh}^i(s)(0)$ is in the boundary portion of an n-subword of s. In the spatial context define $s \in \partial_n \subseteq \mathbb{K}^c$ by putting $s \in \partial_n$ if $s(0)$ is the boundary of an n-subword of s.

For $s \in S$,

$$\partial_n(s) \subseteq \bigcup\{[l, l+q_n) : s \upharpoonright [l, l+q_n) \in \mathcal{W}_n^c\}.$$

The relationship between $\partial_n(s) \subseteq \mathbb{Z}$ and $\partial_n \subseteq \mathbb{K}^c$ is that for $s \in \mathbb{K}^c$,

$$i \in \partial_n(s) \subseteq \mathbb{Z} \text{ iff } \mathrm{sh}^i(s) \in \partial_n \subseteq \mathbb{K}^c.$$

The next lemma says that if s is mature at stage n, then we can detect locally those i for which the i-shifts of s are mature.

Lemma 49. *Suppose that $s \in S$, n is mature for s and $n < m$.*

(1) *Assume the first three numerical requirements. Suppose that $i \in [-r_m(s), q_m - r_m(s))$. Then n is mature for $\mathrm{sh}^i(s)$ iff*
   (a) $i \notin \bigcup_{n \leq k \leq m} \partial_k(s)$ *and*
   (b) $\mathrm{sh}^i(s) \notin \bigcup_{n \leq k < m}(E_k^0 \cup E_k^1 \cup E_k^2)$.

(2) *For all but at most $(\sum_{n<k\leq m} 1/l_k) + (\sum_{n\leq k<m} 6\varepsilon_k)$ proportion of the indices $i \in [-r_m(s), q_m - r_m(s))$, the point $\mathrm{sh}^i(s)$ is mature for n.*

Hence by Numerical Requirement 2, the proportion of $i \in [-r_m(s), q_m - r_m(s))$ for which the i-shift of s is not mature for n is less than $1/l_{n-1} + \varepsilon_{n-1}$.

Proof. The first item is immediate from the definition of *mature*. For the second item, first note that

$$\bigcup_{n\leq k\leq m} \partial_k(s) \cup \bigcup_{n\leq k<m} (E_k^0 \cup E_k^1 \cup E_k^2) = \partial_m(s) \cup \bigcup_{n\leq k<m} (\partial_k(s) \cup E_k^0 \cup E_k^1 \cup E_k^2).$$

Let $I = [-r_m(s), q_m - r_m(s))$. Since ∂_m has proportion $1/l_m$ of I, it suffices to show that for a fixed $k \in [n, m)$, the proportion of $i \in I$ such that $\mathrm{sh}^i(s) \in \partial_k \cup E_k^0 \cup E_k^1 \cup E_k^2$ is less than $1/l_k + 6\epsilon_k$.

There are at most q_m/q_k k-words appearing in $s \upharpoonright I$. There are at most $1/l_k$ many i in the boundary of each of these k-words. So total number of i in $\partial_k(s) \cap I$ is less than or equal to $(\frac{q_m}{q_k})(q_k/l_k)$, hence has proportion less than or equal to $1/l_k$ of I.

Similarly for $j = 0, 1, 2$ the number of i with $\mathrm{sh}^i(s) \in E_k^j$ and i is in the block corresponding to a k-subword of $s \upharpoonright I$ is at most $(q_m/q_k)2\varepsilon_k q_k$, and hence those i have proportion bounded by

$$\left(\frac{(q_m/q_k)2\varepsilon_k q_k}{q_m}\right) = 2\varepsilon_k$$

in I. It follows that the collection of $i \in I$ such that $\mathrm{sh}^i(s) \in E_k^0 \cup E_k^1 \cup E_k^2$ is bounded by $3 * 2\varepsilon_k$.

Numerical Requirements 1 and 2 imply that the sum in item (2) of the lemma is bounded by $1/l_{n-1} + \varepsilon_{n-1}$. ■

A very similar statement is the following:

Lemma 50. *Suppose that $s \in S$ and s has a principal n-block. Then n is mature provided that $s \notin \bigcup_{n \leq m} E^0_m \cup E^1_m \cup E^2_m$. In particular, if n is mature for s and s is not in a boundary portion of its principal $n-1$-block or in $E^0_{n-1} \cup E^1_{n-1} \cup E^2_{n-1}$, then $n-1$ is mature for s.*

5.6. *The circle factor*

Let $\langle k_n, l_n : n \in \mathbb{N}\rangle$ be a circular coefficient sequence and let $\langle p_n, q_n : n \in \mathbb{N}\rangle$ be the associated sequence defined by formulas 5.5 and 5.6. Let $\alpha_n = p_n/q_n$ and $\alpha = \lim \alpha_n$.

For a natural number $q \geq 1$, let $\mathcal{I}_q$ be the partition of the interval $[0, 1)$ with atoms $\langle [i/q, (i+1)/q) : 0 \leq i < q\rangle$, and refer to $[i/q, (i+1)/q)$ as I^q_i.[18] Since p_n and q_n are relatively prime, the rotation $\mathcal{R}_{\alpha_n}$ enumerates the partition $\mathcal{I}_{q_n}$ starting with $I^{q_n}_0$. Thus $\mathcal{I}_{q_n}$ has two natural orderings – the usual geometric ordering and the *dynamical* ordering given by the order that $\mathcal{R}_{\alpha_n}$ enumerates $\mathcal{I}_{q_n}$. Since $j_i = p^{-1}i \pmod q$, I^q_i is the j_i-th interval in the dynamical ordering.

Definition 51. For $x \in [0, 1)$ we will write $D_n(x) = j$ if x belongs to the j-th interval in the dynamical ordering of $\mathcal{I}_{q_n}$. Equivalently, $D_n(x) = j$ if $x \in I^{q_n}_{jp_n}$.

Informal description. Following [11], for each stage n, we have a periodic approximation τ_n to $\mathbb{K}^c$ consisting of towers $\mathcal{T}$ of height q_n whose levels correspond to subintervals of $[0, 1)$. This approximation refines the periodic permutation of $\mathcal{I}_{q_n}$ determined by $\mathcal{R}_{\alpha_n}$. If s is mature, then s lies is the $r^{th}_n(s)$ level of $\mathcal{I}_{q_n}$ in the dynamical ordering. Passing from τ_n to τ_{n+1} the mature points remain in the same levels of the n-towers as they are spread into the $n+1$-towers in τ_{n+1}. The towers of τ_{n+1} can be viewed as cut-and-stack constructions–filling in boundary points between cut n-towers. The fillers are taken from portions of the n-towers.

With this view each mature point remains in the same interval of $\mathcal{I}_{q_n}$ when viewed in τ_{n+1}. Moreover, if $s \in J \in \mathcal{I}_{q_{n+1}}$ and $J \subseteq I \in \mathcal{I}_{q_n}$, then $\mathcal{R}_{\alpha_{n+1}} J \subseteq \mathcal{R}_{\alpha_n} I$.

Thus the $n+1$-tower for $\mathcal{R}_{\alpha_{n+1}}$ has multiple contiguous sequences of levels of length q_n that are sublevels of the n-tower and the action of $\mathcal{R}_{\alpha_n}$ and $\mathcal{R}_{\alpha_{n+1}}$ agree on these levels.

Definition 52. Let $\Sigma_0 = \{*\}$. We define a circular construction sequence such that each $\mathcal{W}^c_n$ has a unique element as follows:

(1) $\mathcal{W}^c_0 = \{*\}$ and

(2) if $\mathcal{W}^c_n = \{w_n\}$, then $\mathcal{W}^c_{n+1} = \{\mathcal{C}(w_n, w_n, \ldots, w_n)\}$.

Let $\mathcal{K}$ be the resulting circular system.

It is easy to check that $\mathcal{K}$ has unique non-atomic measure since the unique n-word, w_n, occurs exactly $k_n(l_n - 1)q_n$ many times in w_{n+1}. This measure is ergodic.

[18] If $i > q$, then I^q_i refers to $I^q_{i'}$, where $i' < q$ and $i' \equiv i \mod q$.

Let $\mathbb{K}^c$ be an arbitrary circular system with coefficients $\langle k_n, l_n : n \in \mathbb{N}\rangle$. Then $\mathbb{K}^c$ has a canonical factor isomorphic to $\mathcal{K}$. This canonical factor plays a role for circular systems analogous to the role odometer transformations play for odometer based systems.

To see $\mathcal{K}$ is a factor of $\mathbb{K}^c$, define the following function:

$$\pi(x)(i) = \begin{cases} x(i) & \text{if } x(i) \in \{b, e\}, \\ * & \text{otherwise.} \end{cases} \tag{5.11}$$

Notation. Write w_n^α for the unique element of $\mathcal{W}_n^c$ in the construction sequence for $\mathcal{K}$. Then w_n^α lies in the principal n-block of the projection to $\mathcal{K}$ of any $s \in \mathbb{K}^c$ for which n is mature.

Theorem 53 ([11, Theorem 43]). *Let ν be the unique non-atomic shift-invariant measure on $\mathcal{K}$. Then*

$$(\mathcal{K}, \mathcal{B}, \nu, \mathrm{sh}) \cong (S^1, \mathcal{D}, \lambda, \mathcal{R}_\alpha),$$

where $\mathcal{R}_\alpha$ is the rotation of the unit circle by $\alpha * 2\pi$ radians and $\mathcal{B}, \mathcal{D}$ are the σ-algebras of measurable sets.

The isomorphism $\phi_0 : \mathcal{K} \to S^1$ asserted to exist in Theorem 53 is constructed as a limit of functions ρ_n, where ρ_n is defined by setting

$$\rho_n(s) = \frac{i}{q_n} \tag{5.12}$$

iff $I_i^{q_n}$ is the $r_n(s)$-th interval in the dynamical ordering.[19] Equivalently, since the r_n-th interval in the geometric ordering is $I^{q_n}_{p_n r_n(s)}$,

$$i \equiv p_n r_n(s) \bmod q_n. \tag{5.13}$$

The following follows from [11, Proposition 44].

Proposition 54. *Suppose that n is mature for s. Then*

$$r_n(s) = D_n(\phi_0(s)).$$

The proof of Theorem 2 requires understanding the correspondence between the geometric construction and its symbolic representation. The words in $\mathcal{W}_n$ correspond to cut-and-stack constructions, passing from stage n to $n+1$ via the $\mathcal{C}$ operator corresponds to basing the cut and stack construction on $\mathcal{R}_{\alpha_{n+1}}$ which agrees with the $\mathcal{R}_{\alpha_n}$ for most consecutive intervals of length q_n. A first step in understanding this correspondence is the next remark and lemma.

Remark 55. It will be helpful to understand ϕ_0^{-1} explicitly. To each point x in the range of ϕ_0, $s = \phi_0^{-1}(x)$ belongs to S. By Lemma 13, to determine s it suffices to know

[19] Thus r_n and ρ_n both have the same subset of S as their domain and contain the same information. They map to different places $r_n : S \to \mathbb{N}$, whereas $\rho_n : S \to [0, 1)$ and is the left endpoint of the r_n-th interval in the dynamical ordering.

$\langle r_n(s) : n \geq N\rangle$ for some N as well as the sequence $\langle w_n : n \geq N\rangle$ of principal subwords of s. Since we are working with $\mathcal{K}$, the only choice for w_n is w_n^α. For mature n, Proposition 54 tells us that $r_n(s) = D_n(x)$. Thus s is the unique element of S with the property that $\langle r_n(s) : n \in \mathbb{N}\rangle$ agrees with $\langle D_n(x) : n \in \mathbb{N}\rangle$ for all large n.

We isolate the following fact for later use:

Lemma 56. *Suppose that $\phi_0(s) = x$ and $n < m$ are mature for s. Then if I and J are the $D_n(x)$-th and $D_m(x)$-th intervals in the dynamical orderings of $\mathcal{I}^{q_n}$ and $\mathcal{I}^{q_m}$, then $J \subseteq I$.*

The natural way of representing the complex unit circle as an abelian group is multiplicatively: the rotation by $2\pi\alpha$ radians is multiplication by $e^{2\pi i\alpha}$. It is often convenient to identify the unit circle with $[0, 1)$. In doing so, multiplication by $e^{2\pi i\alpha}$ corresponds to "mod one" addition and the complex conjugate $\bar{z}$ corresponds to $-z$.

The following result is standard:

Proposition 57. *Let $\alpha \in [0, 1)$ be irrational. Suppose that $T : S^1 \to S^1$ is an invertible measure preserving transformation that commutes with $\mathcal{R}_\alpha$. Then for some β, $T = \mathcal{R}_\beta$ almost everywhere. Identifying S^1 with $[0, 1)$ there is a β such that for almost all $x \in S^1$,*

$$T(x) = x + \beta \bmod 1. \tag{5.14}$$

It follows that if T is an isomorphism between $\mathcal{R}_\alpha$ and $\mathcal{R}_\alpha^{-1}$, then $T(x) = -x + \beta \bmod 1$.

Definition 58. Using the identification of S^1 with $[0, 1)$ we view $\phi_0 : \mathcal{K} \to [0, 1)$. Given a rotation $\mathcal{R}_\beta$, we get a map $\mathcal{S}_\beta : \mathcal{K} \to \mathcal{K}$ such that

$$\mathcal{S}_\beta(s) = \phi_0^{-1}\mathcal{R}_\beta\phi_0(s).$$

We will occasionally abuse notation and write $s + \beta$ for $\mathcal{S}_\beta(s)$.

5.7. *Points of view*

Circular systems can be viewed from multiple perspectives: geometrically, as limits of periodic processes[20] and as symbolic shifts.

The n-th periodic process consists of a collection of s_n periodic towers with each tower having one level designated as a base. To pass from τ_n to τ_{n+1} the bulk of the τ_n-towers are repeated $q_n(k_n)(l_n - 1)$ many times in blocks of length $l_n - 1$ in each τ_{n+1}-tower. In between these blocks there are filler levels.

The words $w \in \mathcal{W}_n^c$ are in one-to-one correspondence with the towers in τ_n. The "$\mathcal{C}$" operation encodes the transition from τ_n to τ_{n+1}. The towers in τ_{n+1} correspond to words $\mathcal{C}(w_0, \ldots, w_{k_n-1})$. Each τ_n-tower T_j has a corresponding word $w_j \in \mathcal{W}_n$. Repeating stacking of T_j corresponds to the powers of w_j in $\mathcal{C}(w_0, \ldots, w_{k_n-1})$. The levels of a tower in τ_{n+1} are either contained in levels of τ_n-tower or are filler blocks labelled "b"

[20] See [11, Section 5] for the formal definition.

or “e.” The repetitions of each w_i in 0-subsections correspond to stacking parts of the levels of the corresponding tower in τ_n periodically $l_n - 1$ times.

The circle factor $\mathcal{K}_\alpha$ captures exactly the structure of the *levels* of the towers and how they interact as one moves from τ_n to τ_{n+1}. This is the idea behind for the construction of the isomorphism between $(\mathcal{K}_\alpha, \nu, \mathrm{sh})$ and $(S^1, \lambda, \mathcal{R}_\alpha)$ and made explicit in Proposition 54.

Given an $s \in \mathbb{K}^c$ that is mature for $n \leq m$ we can view its restriction to its principal m-subword as a particular tower in τ_m. Since s is mature for m, the principal subword is repeated many times on either side of $s(0)$. In particular, we see:

Remark 59. Suppose that n is mature for $s \in S \subseteq \mathbb{K}^c$, $n \leq m$ and $0 \leq d < q_m$. Then

$$r_n(\mathrm{sh}^d(s)) \equiv_{q_n} d + r_n(s). \tag{5.15}$$

The circle factor $\mathcal{K}_\alpha$ of $\mathbb{K}^c$ punctuates the elements of $S \subseteq \mathbb{K}^c$. Since there is only one word in each element of the construction sequence for $\mathcal{K}_\alpha$, we can view the levels of its tower as being of the form $[i/q_n, (i+1)/q_n)$ in the dynamical ordering. Then the cyclic permutation of these levels given by $\mathcal{R}_{p_n/q_n}$. This permutation preserves the dynamical ordering and, for s that are mature at stage n, reflect the behavior of $r_n(s)$.

5.8. *The natural map*

A specific isomorphism $\natural : (\mathcal{K}, \mathrm{sh}) \to (\mathrm{rev}(\mathcal{K}), \mathrm{sh})$ will serve as a benchmark for understanding of potential maps $\phi : \mathbb{K}^c \to \mathrm{rev}(\mathbb{K}^c)$. Viewing $\mathcal{R}_\alpha$ as a rotation of the unit circle by $\alpha * 2\pi$ radians one can view the transformation $\natural$ as a symbolic analogue of complex conjugation $z \mapsto \bar{z}$ on the unit circle, which is an isomorphism between $\mathcal{R}_\alpha$ and $\mathcal{R}_{-\alpha}$. Indeed, by Theorem 53, $\mathcal{K} \cong \mathcal{R}_\alpha$ and so $\mathrm{rev}(\mathcal{K}) \cong \mathcal{R}_{-\alpha}$. Copying $\natural$ over to a map on the unit circle will give an isomorphism ϕ between $\mathcal{R}_\alpha$ and $\mathcal{R}_{-\alpha}$. If we view z and α as elements of the unit interval and the rotation as addition modulo 1, Proposition 57 says that such an isomorphism must be of the form

$$\phi(z) = -z + \beta$$

for some β. It follows immediately from this characterization that $\natural$ is an involution.[21]

The map $\natural$ is defined as the limit of a sequence of codes $\langle \Lambda_n : n \in \mathbb{N} \rangle$ that converge to an isomorphism from $\mathcal{K}$ to $\mathrm{rev}(\mathcal{K})$ (see [12] for more details). The Λ_n will be shifting and reversing words. The amount of shift is determined by the Anosov–Katok coefficients p_n, q_n defined in equations (5.6) and (5.5).

Let $A_0 = 0$ and inductively

$$A_{n+1} = A_n - (p_n)^{-1}. \tag{5.16}$$

It is easy to check that

$$|A_{n+1}| < 2q_n. \tag{5.17}$$

[21] The particular β given by $\natural$ is determined by the specific variation of the definition one uses – indeed any *central* value can occur as a β. (See Section 8 for the definition and use of central values.)

Define a stationary code $\overline{\Lambda}_n$ with domain S that approximates elements of rev$(\mathcal{K})$ by defining

$$\Lambda_n(s) = \begin{cases} \mathrm{sh}^{A_n+2r_n(s)-(q_n-1)}(\mathrm{rev}(s))(0) & \text{if } r_n(s) \text{ is defined,} \\ b & \text{otherwise.} \end{cases} \tag{5.18}$$

The following result appears in [12]:

Theorem 60. *The sequence of stationary codes $\langle \overline{\Lambda}_n : n \in \mathbb{N}\rangle$ converges to a shift invariant function $\bar{\natural} : \mathcal{K} \to (\{*\} \cup \{b, e\})^{\mathbb{Z}}$ that induces an isomorphism $\natural$ from $\mathcal{K}$ to* rev$(\mathcal{K})$.

Remark 78 of [12] implies that the convergence is prompt: for a typical s and all large enough n, $\natural(s)$ agrees with $\bar{\Lambda}_n(s)$ on the principal n-block of s.

Caveat. Since $(\mathbb{K}^c)^{-1} = (\mathbb{K}^c, \mathrm{sh}^{-1})$ is trivially isomorphic to $(\mathrm{rev}(\mathbb{K}^c), \mathrm{sh})$, we often do not distinguish them. However, as in Definition 63 of the *synchronous* and *anti-synchronous* joinings, the notational distinction becomes important.

When viewing $(\mathbb{K}^c)^{-1}$ and $\mathbb{K}^c$ with the backwards shift and considering the action on the circle factor instead of using $\natural$, one must use

$$\mathrm{rev}(\,\cdot\,) \circ \natural \tag{5.19}$$

instead of simply $\natural$.

5.9. *Categories and the functor $\mathcal{F}$*

Fix a circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$. Let Σ be a language and $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ a construction sequence for an odometer based system with coefficients $\langle k_n : n \in \mathbb{N}\rangle$. Then for each n the operation $\mathcal{C}_n$ is well-defined. We define a construction sequence $\langle \mathcal{W}^c_n : n \in \mathbb{N}\rangle$ and bijections $c_n : \mathcal{W}_n \to \mathcal{W}^c_n$ by induction as follows:

(1) Let $\mathcal{W}^c_0 = \Sigma$ and c_0 be the identity map.

(2) Suppose that $\mathcal{W}_n$, $\mathcal{W}^c_n$ and c_n have already been defined.

$$\mathcal{W}^c_{n+1} = \{\mathcal{C}_n(c_n(w_0), c_n(w_1), \ldots, c_n(w_{k_n-1})) : w_i \in \mathcal{W}_n,\ w_0w_1\ldots w_{k_n-1} \in \mathcal{W}_{n+1}\}.$$

(Words in $\mathcal{W}_{n+1}$ are concatenations of k_n words in $\mathcal{W}_n$ and so can be written in the required form: as $w_0w_1\ldots w_{k_n-1}$ with $w_j \in \mathcal{W}_n$.) Define the map c_{n+1} by setting

$$c_{n+1}(w_0w_1\ldots w_{k_n-1}) = \mathcal{C}_n(c_n(w_0), c_n(w_1), \ldots, c_n(w_{k_n-1})).$$

Note in case 2 the prewords are

$$P_{n+1} = \{(c_n(w_0), c_n(w_1), \ldots, c_n(w_{k_n-1})) : w_0w_1\ldots w_{k_n-1} \in \mathcal{W}_{n+1}\}.$$

Remark 61. Some useful facts are:

- It follows from Lemma 36 and Numerical Requirement 1 that if $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ is an odometer based construction sequence, then $\langle \mathcal{W}^c_n : n \in \mathbb{N}\rangle$ is a construction sequence; i.e. the spacer proportions are summable.

- If each $w \in \mathcal{W}_n$ occurs exactly the same number of times in every element of $\mathcal{W}_{n+1}$, then $\langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$ is strongly uniform.
- Odometer words in $\mathcal{W}_n$ have length K_n. The length of the circular words in $\mathcal{W}_n^c$ is q_n.

Definition 62. Define a map $\mathcal{F}$ from the set of odometer based subshifts to circular subshifts as follows. Suppose that $\mathbb{K}$ is an odometer based shift built from a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$. Define

$$\mathcal{F}(\mathbb{K}) = \mathbb{K}^c,$$

where $\mathbb{K}^c$ has construction sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$.

The map $\mathcal{F}$ is one to one by the unique readability of words in $\mathcal{W}$. Suppose that $\mathbb{K}^c$ is a circular system with coefficients $\langle k_n, l_n : n \in \mathbb{N}\rangle$. We can recursively build functions ${c_n}^{-1}$ from words in $\Sigma \cup \{b, e\}$ to words in Σ. The result is a odometer based system $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ with coefficients $\langle k_n : n \in \mathbb{N}\rangle$. If $\mathbb{K}$ is the resulting odometer based system then $\mathcal{F}(\mathbb{K}) = \mathbb{K}^c$. Thus $\mathcal{F}$ is a bijection.

If $\mathbb{K}$ is an odometer based system, denote the odometer base by $\mathbb{K}^\pi$ and let $\pi : \mathbb{K} \to \mathbb{K}^\pi$ be the canonical factor map. If $\mathbb{K}^c$ is a circular system, let $(\mathbb{K}^c)^\pi$ be the rotation factor $\mathcal{K}$ and let $\pi : \mathbb{K}^c \to \mathcal{K}$ be the canonical factor map. For both odometer based and circular systems the underlying canonical factors serve as timing mechanisms. This motivates the following.

Definition 63. *Synchronous* and *anti-synchronous* joinings are defined as follows:[22]

(1) Let $\mathbb{K}$ and $\mathbb{L}$ be odometer based systems with the same coefficient sequence, and ρ a joining between $\mathbb{K}$ and $\mathbb{L}^{\pm 1}$. Then ρ is *synchronous* if ρ joins $\mathbb{K}$ and $\mathbb{L}$ and the projection of ρ to a joining on $\mathbb{K}^\pi \times \mathbb{L}^\pi$ is the graph joining determined by the identity map (the diagonal joining of the odometer factors); ρ is *anti-synchronous* if ρ is a joining of $\mathbb{K}$ with $\mathbb{L}^{-1}$ and its projection to $\mathbb{K}^\pi \times (\mathbb{L}^{-1})^\pi$ is the graph joining determined by the map $x \mapsto -x$.

(2) Let $\mathbb{K}^c$ and $\mathbb{L}^c$ be circular systems with the same coefficient sequence and ρ a joining between $\mathbb{K}^c$ and $(\mathbb{L}^c)^{\pm 1}$. Then ρ is *synchronous* if ρ joins $\mathbb{K}^c$ and $\mathbb{L}^c$ and the projection to a joining of $(\mathbb{K}^c)^\pi$ with $(\mathbb{L}^c)^\pi$ is the graph joining determined by the identity map of $\mathcal{K}$ with $\mathcal{L}$, the underlying rotations; ρ is *anti-synchronous* if it is a joining of $\mathbb{K}^c$ with $(\mathbb{L}^c)^{-1}$ and projects to the graph joining determined by $\mathrm{rev}(\cdot) \circ \natural$ on $\mathcal{K} \times \mathcal{L}^{-1}$.

The categories. Let $\mathcal{O}B$ be the category whose objects are ergodic odometer based systems with coefficients $\langle k_n : n \in \mathbb{N}\rangle$. The morphisms between objects $\mathbb{K}$ and $\mathbb{L}$ will be synchronous graph joinings of $\mathbb{K}$ and $\mathbb{L}$ or anti-synchronous graph joinings of $\mathbb{K}$ and $\mathbb{L}^{-1}$. We call this the *category of odometer based systems.*

[22] We use $\mathcal{L}$ for the notation for the rotation factor of a circular system $\mathbb{L}^c$. In this context, when taking inverses of symbolic systems we keep the same orientation for the symbolic system and use sh^{-1}.

Let $\mathcal{C}B$ be the category whose objects consists of all ergodic circular systems with coefficients $\langle k_n, l_n : n \in \mathbb{N}\rangle$. The morphisms between objects $\mathbb{K}^c$ and $\mathbb{L}^c$ will be synchronous graph joinings of $\mathbb{K}^c$ and $\mathbb{L}^c$ or anti-synchronous graph joinings of $\mathbb{K}^c$ and $(\mathbb{L}^c)^{-1}$. We call this the *category of circular systems.*

The main theorem of [12] is the following:

Theorem 64. *For a fixed circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$ the categories $\mathcal{O}B$ and $\mathcal{C}B$ are isomorphic by a function $\mathcal{F}$ that takes synchronous joinings to synchronous joinings, anti-synchronous joinings to anti-synchronous joinings, isomorphisms to isomorphisms and weakly mixing extensions to weakly mixing extensions.*[23]

It is also easy to verify that the map $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle \mapsto \langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$ takes uniform construction sequences to uniform construction sequences and strongly uniform construction sequences to strongly uniform construction sequences.

Remark 65. Were we to be completely precise we would take objects in $\mathcal{O}B$ to be *presentations* of odometer based systems by construction sequences $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$ without spacers and the objects in $\mathcal{C}B$ to be *presentations* by circular construction sequences. This subtlety does not cause problems in the sequel so we ignore it.

5.10. Propagating equivalence relations and actions

In [8], the number $M(s)$ is the first stage in the tree for which σ_m has length s. It is the first stage that the equivalence relation $\mathcal{Q}_s^m$ is defined.

The main result of [8] is the existence of a continuous function from the space of trees to odometer based transformations that reduces ill-founded trees to ergodic transformations isomorphic to their inverses. Components of the construction include equivalence relations $\langle \mathcal{Q}_s^n : M(s) \leq n, s \in \mathbb{N}\rangle$ and groups $\langle G_s^n : M(s) \leq n, s \in \mathbb{N}\rangle$. Some of their properties are:

(1) M is a monotone, strictly increasing function from $\mathbb{N}$ to $\mathbb{N}$.

(2) $\mathcal{Q}_0^0$ is the trivial equivalence relation with one equivalence class on $\mathcal{W}_0 = \Sigma$.

(3) $\mathcal{Q}_s^n$ is an equivalence relation on $\mathcal{W}_n$.

(4) For integers $n \geq M(s) + 1$, viewing elements of $\mathcal{W}_n$ as concatenations of words in $\mathcal{W}_{M(s)}$, $\mathcal{Q}_s^n$ is the product equivalence relation of $\mathcal{Q}_s^{M(s)}$. Hence we can view $\mathcal{W}_n/\mathcal{Q}_s^n$ as sequences of elements of $\mathcal{W}_{M(s)}/\mathcal{Q}_s^{M(s)}$ and similarly for $\mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$. These sequences have length K_n and are made of $K_n/K_{M(s)}$ many constant blocks of length $K_{M(s)}$.

(5) The groups $\langle G_s^n : M(s) \leq n, s \in \mathbb{N}\rangle$ are direct sums of copies of $\mathbb{Z}_2$ that have a designated canonical collection of free generators.[24] Each $G_s^{n+1} = G_s^n \oplus H$, where H is either $\mathbb{Z}/2\mathbb{Z}$ or H is trivial.

[23] Glasner showed that it takes compact extensions to compact extensions.

[24] These groups are described in detail in Section 10.2.

(6) Each group G_s^n acts freely on $\mathcal{W}_n/\mathcal{Q}_s^n \cup \mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$ in a manner that even parity group elements preserve the sets $\mathcal{W}_n/\mathcal{Q}_s^n$ and $\mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$ and the odd parity group elements send elements of $\mathcal{W}_n/\mathcal{Q}_s^n$ to $\mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$.

(7) The action of $G_s^n \subseteq G_s^{n+1}$ on $\mathcal{W}_{n+1} \cup \mathrm{rev}(\mathcal{W}_{n+1})$ is propagated from $\mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$ by the *skew-diagonal* action: if $g \in G_s^n$ is a canonical generator and if the word $w \in \mathcal{W}_{n+1} \cup \mathrm{rev}(\mathcal{W}_{n+1})$ is of the form $w_0 w_1 \dots w_{k_n-1}$, then

$$gw = gw_{k_n-1} \dots gw_1 gw_0.$$

We define corresponding equivalence relations and group actions on $\langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$. They will be used in Section 8.2.1 to state the timing assumptions and in Section 10.2 which gives the construction specifications from [8].[25]

An inductive understanding of $(\mathcal{Q}_s^n)^s$ and the G_s^n-actions is quite useful.

Inductive definition of $(\mathcal{Q}_s^n)^c$. Define

- $(\mathcal{Q}_0^n)^c$ to have exactly one class in each $\mathcal{W}_n^c$,
- for $w_0, w_1 \in \mathcal{W}_{M(s)}$ put $(c_{M(s)}(w_0), c_{M(s)}(w_1)) \in (\mathcal{Q}_s^{M(s)})^c$ iff $(w_0, w_1) \in \mathcal{Q}_s^{M(s)}$.

Suppose we are given $(\mathcal{Q}_s^n)^c$ on $\mathcal{W}_n^c$. Define an equivalence relation $\mathcal{Q}$ on $\mathcal{W}_{n+1}^c$ by setting $\mathcal{C}(w_0, \dots, w_{k_n-1})$ equivalent to $\mathcal{C}(w_0', \dots, w_{k_n-1}')$ if and only if for all i, w_i is $(\mathcal{Q}_s^n)^c$-equivalent to w_i'.

Rather than a full definition of the action of G_s^{n+1} on

$$\mathcal{W}_{n+1}^c/(\mathcal{Q}_s^{n+1})^c \cup \mathrm{rev}(\mathcal{W}_{n+1}^c/(\mathcal{Q}_s^{n+1}))^c,$$

we describe the how the action of G_s^n propagates: via the *circular skew diagonal action*:

Identify $\mathrm{rev}(\mathcal{W}_{n+1}^c/(\mathcal{Q}_s^{n+1})^c)$ with the collection of sequences of the form

$$\mathcal{C}^r(\mathrm{rev}([w_0]_{(\mathcal{Q}_s^n)^c}), \mathrm{rev}([w_1]_{(\mathcal{Q}_s^n)^c}), \dots, \mathrm{rev}([w_{k_n-1}]_{(\mathcal{Q}_s^n)^c}))$$

as $w_0 w_1 \dots w_{k_n-1}$ ranges over the elements of P_n.

To define the skew-diagonal action of G_s^n on classes of circular words, it suffices to specify it on the canonical generators, This is done by setting[26]

$$g\mathcal{C}([w_0], [w_1] \dots [w_{k-1}]) =_{\mathrm{def}} \mathcal{C}^r([gw_0], [gw_1], \dots, [gw_{k-1}])$$

whenever g is a canonical generator of G_s^n. We observe that the skew-diagonal action has the property that the canonical generators take elements of $\mathcal{W}_{n+1}^c/(\mathcal{Q}_s^{n+1})^c$ to elements of $\mathrm{rev}(\mathcal{W}_{n+1}^c/(\mathcal{Q}_s^{n+1})^c)$. It follows that the even parity elements of G leave the sets $\mathcal{W}_{n+1}^c/(\mathcal{Q}_s^{n+1})^c$ and $\mathrm{rev}(\mathcal{W}_{n+1}^c/(\mathcal{Q}_s^{n+1})^c)$ invariant and odd parity elements of G take $\mathcal{W}_{n+1}^c/(\mathcal{Q}_s^{n+1})^c$ to elements of $\mathrm{rev}(\mathcal{W}_{n+1}^c/(\mathcal{Q}_s^{n+1})^c)$ and vice versa.

As in [8] the equivalence relations $\langle \mathcal{Q}_s^n : n \in \mathbb{N}\rangle$ define factors $\mathbb{K}_s$ of $\mathbb{K}$ and similarly $\langle (\mathcal{Q}_s^n)^c : n \in \mathbb{N}\rangle$ define factors $\mathbb{K}_s$ of $\mathbb{K}^c$ The equivariant definitions given here imply that $\mathcal{F}$ takes each $\mathbb{K}_s$ to $\mathbb{K}_s^c$ and respects the actions of the G_s^n.

---

[25] If $\mathcal{Q}$ is an equivalence relation on $\mathcal{W}^c$ define $\mathrm{rev}(\mathcal{Q})$ by $(\mathrm{rev}(w_0), \mathrm{rev}(w_1)) \in \mathrm{rev}(\mathcal{Q})$ if and only if $(w_0, w_1) \in \mathcal{Q}$. In abuse of notation we will not distinguish between $(\mathcal{Q}_s^n)^c$ as a relation on $\mathcal{W}_n^c$, $(\mathcal{Q}_s^n)^c \cup \mathrm{rev}((\mathcal{Q}_s^n)^c)$ as a relation on $\mathcal{W}_n^c \cup \mathrm{rev}(\mathcal{W}_n^c)$ or $\mathcal{W}_n^c/(\mathcal{Q}_s^n)^c \cup \mathrm{rev}(\mathcal{W}_n^c/(\mathcal{Q}_s^n)^c)$.

[26] We use $[w_i]$ to denote $[w_i]/(\mathcal{Q}_s^n)^c$.

6. Understanding rotations

Let $\mathcal{K}$ be a rotation factor of a circular system with coefficient sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$. This section analyzes how automorphisms of $\mathcal{K}$ affect the parsing of elements of $\mathcal{K}$.

Let $(\mathbb{K}^c, \mu^c)$ and $(\mathbb{L}^c, \nu^c)$ be two circular systems that share a given circular coefficient sequence and let $\alpha = \lim \alpha_n$. Any isomorphism between $\mathbb{K}^c$ and $(\mathbb{L}^c)^{\pm 1}$ induces a unitary isomorphism U_ϕ from $L^2((\mathbb{L}^c)^{\pm 1})$ to $L^2(\mathbb{K}^c)$, and this isomorphism sends eigenfunctions for $n\alpha$ to eigenfunctions for $n\alpha$. Thus every isomorphism has to send the canonical factor $\mathcal{K}_\alpha$ of $\mathbb{K}^c$ to the canonical factor $\mathcal{K}_\alpha^{\pm 1}$ of $(\mathbb{L}^c)^{\pm 1}$. Explicitly: suppose that $\phi : \mathbb{K}^c \to (\mathbb{L}^c)^{\pm 1}$ is an isomorphism. Then $U_\phi : L^2((\mathbb{L}^c)^{\pm 1}) \to L^2(\mathbb{K}^c)$, and U_ϕ takes the space generated by eigenfunctions of U_{sh} in $L^2((\mathbb{L}^c)^{\pm 1})$ with eigenvalues $\{\alpha^n : n \in \mathbb{Z}\}$ to the space generated by corresponding eigenfunctions in $L^2(\mathbb{K}^c)$. Consequently, there is a measure preserving transformation ϕ^π making the following diagram commute:

$$\begin{tikzcd}
\mathbb{K}^c \arrow[r, "\phi"'] \arrow[d, "\pi"] & (\mathbb{L}^c)^{\pm 1} \arrow[d, "\pi"] \\
\mathcal{K}_\alpha \arrow[r, "\phi^\pi"] & \mathcal{K}_\alpha^{\pm 1}
\end{tikzcd} \tag{6.1}$$

By Theorem 53, $\mathcal{K}_\alpha$ is conjugate to the rotation $\mathcal{R}_\alpha$ of the unit circle by a map ϕ_0. Hence (using additive notation) ϕ^π must be conjugate to a transformation defined on the unit interval of the form $x \mapsto z + \beta$ for some $\beta \in [0, 1)$, where z is either x or $-x$, depending on whether ϕ^π maps to $\mathcal{K}_\alpha$ or $\mathcal{K}_\alpha^{-1}$. Since $\mathrm{rev}(\,\cdot\,) \circ \natural : \mathcal{K}_\alpha \to \mathcal{K}_\alpha^{-1}$ is an isomorphism, if ϕ maps to $(\mathbb{L}^c)^{-1}$, $\mathrm{rev}(\,\cdot\,) \circ \natural(x)$ can serve as an alternative to the benchmark to the map $x \mapsto -x$. Explicitly: the β associated to ϕ is the number making

$$\phi^\pi(s) = \mathrm{rev}(\,\cdot\,) \circ \natural(\mathcal{S}_\beta(s));$$

equivalently, $\mathrm{rev}(\,\cdot\,) \circ \natural^{-1} \circ \phi^\pi(s) = \mathcal{S}_\beta(s)$.[27]

Summarizing,

(A) If $\phi : \mathbb{K}^c \to \mathbb{L}^c$ is an isomorphism, then viewed as a map from $[0, 1)$ to $[0, 1)$, there is a unique $\beta \in [0, 1)$ such that for almost every x,

$$\phi^\pi(s) = \mathcal{S}_\beta(s).$$

(B) If $\phi : \mathbb{K}^c \to (\mathbb{L}^c)^{-1}$, then there is a unique β such that for almost every s,

$$\phi^\pi(x) = \mathrm{rev}(\,\cdot\,) \circ \natural(\mathcal{S}_\beta(s)).$$

Definition 66. In cases (A) and (B), we call the map $\mathcal{S}_\beta$ the *rotation associated with* ϕ.

We record the following facts.

[27]The reader is referred to the caveat at the end of Section 5.8, for the reason $\mathrm{rev}(\,\cdot\,) \circ \natural$ is used.

Lemma 67. *Let $\mathbb{K}^c$ be a circular system. Then*

(1) *The set of β associated with automorphisms of $\mathbb{K}^c$ form a group.*

(2) *If $\phi : \mathbb{K}^c \to (\mathbb{K}^c)^{-1}$ and $\psi : \mathbb{K}^c \to \mathbb{K}^c$ are isomorphisms, where $\phi^\pi = \mathrm{rev}(\cdot) \circ \natural \circ \mathcal{S}_\beta$ and $\psi^\pi = \mathcal{S}_\gamma$, then $(\phi \circ \psi)^\pi = \mathrm{rev}(\cdot) \circ \natural \circ \mathcal{S}_\delta$, where $\delta = \beta + \gamma$.*

Proof. It is easy to check that

- If ϕ, ψ are isomorphisms from $\mathbb{K}^c$ to $\mathbb{K}^c$ with $\phi^\pi = \mathcal{S}_\beta$ and $\psi^\pi = \mathcal{S}_\gamma$, then $(\phi \circ \psi)$ is also an isomorphism from $\mathbb{K}^c$ to $\mathbb{K}^c$ and $(\phi \circ \psi)^\pi = \mathcal{S}_\delta$, where $\delta = \beta + \gamma$.
- If ϕ is an isomorphism from $\mathbb{K}^c$ to $\mathbb{K}^c$, and $\phi^\pi = \mathcal{S}_\beta$, then $(\phi^{-1})^\pi = \mathcal{S}_{-\beta}$.

The second assertion is similar. ■

Given a rotation $\mathcal{R}_\beta$, set

$$S(\beta) = \bigcap_{n \in \mathbb{Z}} \mathcal{S}_\beta^n(S).$$

This can be described independently of $\mathcal{S}_\beta$ as

$$\{s \in S : \text{for all } n \in \mathbb{Z}, \phi_0(s) \in (\phi_0[S] + n\beta)\}.$$

It is clear that $\nu(S(\beta)) = 1$.

Define a sequence of functions $\langle d^n : n \in \mathbb{N} \rangle$. Each

$$d^n : S(\beta) \to \{0, 1, 2, \ldots, q_n - 1\}.$$

For $s \in S(\beta)$ and $t = \mathcal{S}_\beta(s)$ we have $t \in S(\beta)$ and $\phi_0(t) = \mathcal{R}_\beta \phi_0(s)$. All large enough n are mature for t, and t is determined by a tail segment of $\langle r_n(t) : n \in \mathbb{N} \rangle$.

Definition 68. If n is mature for both s and $t = \mathcal{S}_\beta(s)$, let

$$d^n(s) \equiv_{q_n} r_n(t) - r_n(s), \tag{6.2}$$

and $d^n(s) = 0$ otherwise. (We could have made a more general definition $d^n(s, t)$ for arbitrary t and take $t = \mathcal{S}_\beta(s)$ when we want to use $d^n(s)$.)

Explicitly: from the definition of r_n, $\phi_0(s) + \beta$ belongs to the $(r_n(s) + d^n(s))$-th interval in the dynamical ordering of $\mathcal{I}_{q_n}$.[28]

Fix an n and suppose that β is not a multiple of $1/q_n$. Then the interval $[\beta, \beta + 1/q_n)$ intersects two geometrically consecutive intervals of the form $[i/q_n, (i+1)/q_n)$.

Lemma 69. *Suppose that the integer n is mature for s and $\mathcal{S}_\beta(s)$. Then $d^n(s)$ belongs to $\{D_n(\beta), D_n(\beta + 1/q_n)\}$. Thus there are only two possible values for $d^n(s)$ and these values differ by j_1.*

[28]More accurately: if $j < q_n$ and $j \equiv_{q_n} r_n(s) + d^n(s)$, then $\phi_0(s) + \beta$ belongs to the j-th interval in the dynamical ordering of $\mathcal{I}_{q_n}$. Recall the relationship between symbolic shifts and the towers of intervals in the dynamical ordering given in Section 5.7.

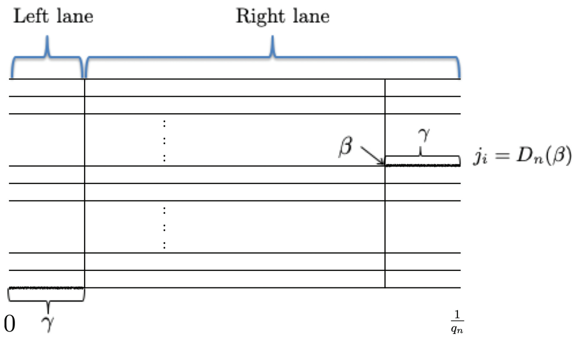

Fig. 2. Left lane and Right lane of the q_n-tower.

Proof. Suppose that $\beta \in [i/q_n, (i+1)/q_n)$ and $\gamma = (i+1)/q_n - \beta$. Then $D_n(\beta) = j_i$. We claim that, relative to those s for which n is mature for both s and $\mathcal{S}_\beta(s)$, d^n is constant on $\phi_0^{-1}(\bigcup_{j<q_n}[j/q_n, (j+1)/q_n - \gamma))$ and $\phi_0^{-1}(\bigcup_{j<q_n}[(j+1)/q_n - \gamma, (j+1)/q_n))$, where it takes values $D_n(\beta)$ and $D_n(\beta + 1/q_n)$, respectively (see Figure 2).

We show that d^n is constant on the first set. Suppose that n is mature for $s, \mathcal{S}_\beta(s)$ and $\phi_0(s) = x$ belongs to the interval $[0, \gamma)$. Then $x + \beta \in [i/q_n, (i+1)/q_n)$. Hence $r_n(\mathcal{S}_\beta(s)) = j_i = D_n(\beta)$. Since $r_n(s) = 0$, we know that $d^n(s) = j_i$. Now suppose that $s^* \in \phi_0^{-1}(\bigcup_{j<q_n}[j/q_n, (j+1)/q_n - \gamma))$ and that n is mature for s^* and $\mathcal{S}_\beta(s^*)$. Let $k = r_n(s^*)$. Then

$$\phi_0(t) = x + \frac{kp_n}{q_n}$$

for some $x \in [0, \gamma)$. So $\phi_0(s^* + \beta) \in [(i + 1 + kp_n)/q) - \gamma, (i + 1 + kp_n)/q)$. Hence

$$r_n(\mathcal{S}_\beta(s^*)) = (p_n)^{-1}(i + kp_n) = j_i + k.$$

Thus

$$d^n(s^*) = r_n(\mathcal{S}_\beta(s^*)) - r_n(s^*) = j_i + k - k = j_i.$$

If $s^* \in \phi_0^{-1}(\bigcup_{j<q_n}[(j+1)/q_n - \gamma, (j+1)/q_n))$, the proof is parallel.

Finally, β and $\beta + 1/q_n$ fall into consecutive intervals of $\mathcal{I}^{q_n}$ in the geometric ordering, and hence $D_n(\beta + 1/q_n) = D_n(\beta) + j_1$. ■

Define d_L^n and d_R^n by setting

$$d_L^n = D_n(\beta) \quad \text{and} \quad d_R^n = D_n\left(\beta + \frac{1}{q_n}\right).$$

Let

$$L_n = \{s : s \text{ is mature at stage } n \text{ and } r_n(s) + d_L^n \equiv_{q_n} r_n(\mathcal{S}_\beta(s))\}$$

and

$$R_n = \{s : s \text{ is mature at stage } n \text{ and } r_n(s) + d_R^n \equiv_{q_n} r_n(\mathcal{S}_\beta(s))\}.$$

We refer to L_n and R_n as the *left lane* and *right lane*, respectively.

Notation. Let β_n^L, β_n^R be the measures of the left and right lanes at stage n.

Lemma 70. *Consider* $(\mathcal{K}, \nu, \mathrm{sh})$ *and let* ι_n *be the measure of the collection* $\widetilde{M}_n$ *of* s *that are not mature at stage* n. *Then:*

(1) $\lceil q_n\beta\rceil - q_n\beta \geq \beta_n^L \geq \lceil q_n\beta\rceil - q_n\beta - \iota_n$,

(2) $q_n\beta - \lfloor q_n\beta\rfloor \geq \beta_n^R \geq q_n\beta - \lfloor q_n\beta\rfloor - \iota_n$,

(3) $\beta_n^L + \beta_n^R + \iota_n = 1$.

In particular, $\sum \beta_n^L < \infty$ *if and only if* $\sum(\lceil q_n\beta\rceil - \beta) < \infty$ *and* $\sum \beta_n^R < \infty$ *if and only if* $\sum(q_n\beta - \lfloor q_n\beta\rfloor) < \infty$.

Proof. Let M_n be the collection of S that are mature at stage n. In the proof of Lemma 69, we showed that L_n is

$$\phi_0^{-1}\left(\bigcup_{j<q_n}[j/q_n, (j+1)/q_n - \gamma)\right) \cap M_n$$

and R_n is

$$\phi_0^{-1}\left(\bigcup_{j<q_n}[(j+1)/q_n - \gamma, (j+1)/q_n)\right) \cap M_n,$$

where $\gamma = (i+1)/q_n - \beta$ and $\beta \in [i/q_n, (i+1)/q_n)$. Since there are q_n many levels and $q_n\gamma = \lceil q_n\beta\rceil - q_n\beta$ the inequalities in item (1) follow. Item (2) is similar. Item (3) follows since

$$S = \phi_0^{-1}\left(\bigcup_{j<q_n}[j/q_n, (j+1)/q_n - \gamma) \cup \bigcup_{j<q_n}[(j+1)/q_n - \gamma, (j+1)/q_n)\right) \cup \widetilde{M}_n.$$

The final assertion follows from Lemma 45. ∎

Restating the discussion:

Lemma 71. *For almost all* $s \in S \subseteq \mathbb{K}^c$ *that are mature at stage* n, $\mathcal{S}_\beta(s)(0) = s(i)$, *where* $i \equiv_{q_n} d_L^n$ *if* $s \in L_n$ *and* $i \equiv_{q_n} d_R^n$ *if* $s \in R_n$.

Proof. Assume that n is mature for s. Then on its principal n-block, the projection of s to $\mathcal{K}_\alpha$ agrees with w_n^α.[29] The values $s(0)$ and $\mathcal{S}_\beta(s)(0)$ are the $r_n(s)$-th and the $r_n(\mathcal{S}_\beta(s))$-th values of the word w_n^α. From equation (6.2),

$$r_n(\mathcal{S}_\beta(s)) = r_n(s) + d^n(s).$$

Hence

$$\mathcal{S}_\beta(s)(0) = s(d^n(s)),$$

and the lemma follows. ∎

[29] Recall w_n^α is the notation for the unique member of the n-th element $\mathcal{W}_n^c$ of the construction sequence for $\mathcal{K}_\alpha$.

The items in the following lemma are essentially Remark 12 and Lemma 56 in a different context.

Lemma 72. *For almost all s and for $n < m$ that are mature for s and $\mathcal{S}_\beta(s)$ the following hold:*

(1) *If $i \equiv_{q_n} r_n(s) + d^n(s)$ and $j \equiv_{q_m} r_m(s) + d^m(s)$, then the j-th place in the principal m-block of $\mathcal{S}_\beta(s)$ is in the i-th place of the principal n-block of $\mathcal{S}_\beta(s)$.*

(2) *Let I be the $r_n(s) + d^n(s)$-th interval of $\mathcal{I}^{q_n}$ and J the $r_m(s) + d^m(s)$-th interval of $\mathcal{I}^{q_m}$ in the dynamical orderings. Then $J \subseteq I$.*

Proof. This follows from Remark 55 and Lemma 56. To see this, note that

$$r_n(\mathcal{S}_\beta(s)) \equiv_{q_n} r_n(s) + d^n(s);$$

i.e. $\mathcal{S}_\beta(s)(0)$ is in the i-th place of the principal n-block of s, where

$$i \equiv_{q_n} r_n(s) + d^n(s). \qquad \blacksquare$$

Thus typical points in R_n and L_n are those in which the n-block of $\mathcal{S}_\beta(s)$ containing 0 is the shift of the block of s containing 0 by d^n_R and d^n_L, respectively.

We now describe how $d^n(\mathrm{sh}^k(s))$ changes. As k varies, $d^n(\mathrm{sh}^k(s))$ measures the shift between $\mathrm{sh}^k(s)(0)$ and $\mathcal{S}_\beta(\mathrm{sh}^k(s))(0)$. In regions where the principal n-subwords of both $\mathrm{sh}^s(s)$ and $\mathcal{S}_\beta(\mathrm{sh}^k(s))$ exist and are repeating $d^n(\mathrm{sh}^k(s))$ is constant. It is also constant as it crosses boundary regions of $\mathrm{sh}^k(s)$ and $\mathcal{S}_\beta(\mathrm{sh}^k(s))$ as long as those boundary regions have length q_n and are lined up with adjacent n-subwords. However for $m \geq n+1$, if the boundary section of an m-word of s or $\mathcal{S}_\beta(s)$ has length not divisible by q_n, the relative alignment between s and $\mathcal{S}_\beta(s)$ changes. This happens on regions of $\bigcup_{m \geq n+1} \partial_m(s) \cup \bigcup_{m \geq n+1} \partial(\mathcal{S}_\beta(s))$.

If n is mature for s, the principal n-word of s repeats on both sides of $s(0)$ and thus we see:

Lemma 73. *If s is mature at stage n, then $d^n(s)$ is constant on the principal n-block of s. Moreover, on $d^n(s)$ is constant on the even and odd overlaps of 2-subsections of $n+1$ subwords of s and $\mathcal{S}_\beta(s)$.*

The next lemma is used for the "nesting" arguments in Section 7.3. It says that the measure of the set of $s \in S$ with $d^n(s) = d^n_L$ or $d^n(s) = d^n_R$ can be closely computed as a density in every scale bigger than n.

Remark. The notation d^n_L and d^n_R are supposed to be suggestive of the left and right lanes. To a close approximation, if s is mature and in a left lane, then

$$d^n(s) = d^n_L$$

and similarly for the right lanes.

Lemma 74. *Let $n < m \in \mathbb{N}$ be natural numbers. Then*

$$\{0, 1, 2, \ldots, q_m - 1\} = P^n_L \cup U \cup P^n_R$$

such that for almost every s for which n is mature:[30]

(1) *if* $r_m(s) \in P^n_L$, *then* $s \in L_n$,

(2) *if* $r_m(s) \in P^n_R$, *then* $s \in R_n$,

(3) $|U| \leq 2q_n$,

(4) $||P^n_L|/q_m - \beta^L_m| < 2q_n/q_m$, *and*

(5) $||P^n_R|/q_m - \beta^R_m| < 2q_n/q_m$.

Proof. As in Lemma 69, let

$$\gamma = \frac{i+1}{q_n} - \beta,$$

where $i = p_n D_n(\beta)$ (see Figure 2). The partition $\mathcal{I}_{q_m}$ splits each interval $I \in \mathcal{I}_{q_n}$ into q_m/q_n subintervals. Let U be the indices of the $\mathcal{I}_{q_m}$ intervals that lie over or under γ and $\gamma + 1/q_m$. Explicitly: suppose that $\gamma \in I^m_{i_0}$ and $\gamma + 1/q_m \in I^m_{i_1}$. Let

$$U = \{i : \text{ for some } 0 \leq j < q_n,\ I^m_i = \mathcal{R}^j_{\alpha_n} I^m_{i_0}\}$$
$$\cup \{i : \text{ for some } 0 \leq j < q_n,\ I^m_i = \mathcal{R}^j_{\alpha_n} I^m_{i_1}\}.$$

Then $|U| = 2q_n$, and if $i \notin U$, then either

$$I^m_i \subseteq \bigcup_{j<q_n} [j/q_n,\ (j+1)/q_n - \gamma) \text{ or} \tag{6.3}$$

$$I^m_i \subseteq \bigcup_{j<q_n} [(j+1)/q_n - \gamma, (j+1)/q_n) \tag{6.4}$$

For $i \notin U$, put $i \in P^n_L$ if it satisfies equation (6.3) and $i \in P^n_R$ if it satisfies equation (6.4). It follows that for almost all s, if n is mature for s and $r_n(s) \in P^n_L$, then $d^n(s) = d^n_L$ and similarly for P^n_R. Since $P^n_R \cup P^n_L \cup U$ is a partition of q_m and $|U| \leq 2q_n$, the lemma follows. ∎

Lemma 75. *Let $f \in \{0,1\}^{\mathbb{N}}$ and let s be a typical member of $S(\beta)$.*

(1) *Let $\beta^*_n = p_n D_n(\beta) + f(i)/q_n$. Then the sequence $\langle \mathcal{R}_{\beta^*_n} : n \in \mathbb{N}\rangle$ converges to $\mathcal{R}_\beta$ in the C^∞-topology.*

As a result, in the language of symbolic systems:

(2) *Let $A_n = D_n(\beta + f(i)/q_n)$ and T the shift map on $\mathcal{K}_\alpha$. Then A_n is either d^n_L or d^n_R, depending on the value of f and for almost every $s \in S$, $\lim_{n\to\infty} T^{A_n}s = \mathcal{S}_\beta(s)$.*

(3) *With A_n as in item (2) and $\mathbb{K}^c$ an arbitrary circular system with the given coefficient sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$, define a_n and b_n to be the left and right endpoints of the principal n-block of $T^{A_n}(s)$. Then for almost all s, $\lim_{n\to\infty} a_n = -\infty$ and $\lim_{n\to\infty} b_n = \infty$.*

[30] Properly speaking the P^n_R and P^n_L notation should indicate m as well. Without any contextual indication of what m is we take $m = n + 1$.

Proof. The first item follows because $|\beta_n^* - \beta| < 2/q_n$. Hence β_n^* converges rapidly to β. The second item follows from the first via the isomorphism ϕ_0^{-1}. The third item follows because $\mathcal{S}_\beta(s) \in S$ and $T^{A_n}(s)$ converges to $\mathcal{S}_\beta(s)$ topologically. Hence for all n there is an N such that for all $m \geq N$, the principal n-block of $T^{A_m}(s)$ is the same as the principal n-block of $\mathcal{S}_\beta(s)$. Since the principal m-block of T^{A_m} contains the principal n-block of $\mathcal{S}_\beta(s)$ and $\mathcal{S}_\beta(s) \in S$, item (3) follows. ∎

If a_n and b_n are as in item (3), then

$$a_n = -r_n(s) + A_n \quad \text{and} \quad b_n = q_n - r_n(s) + A_n. \tag{6.5}$$

7. The displacement function

In this section we define a function Δ from S^1 to the extended positive real numbers that will eventually be shown to have the properties that

- $\Delta(\beta) < \infty$ implies that there is an element of the centralizer of $\mathbb{K}^c$ having $\mathcal{R}_\beta$ as its associated rotation.
- if $\mathbb{K}^c$ is built suitably randomly, then every element of the centralizer of $\mathbb{K}^c$, or isomorphism from $\mathbb{K}^c$ to $(\mathbb{K}^c)^{-1}$ has rotation factor β with $\Delta(\beta) < \infty$.

The idea behind the displacement function is simple: the number β determines $\mathcal{S}_\beta$ and hence a shift at each scale n. The words in $\mathcal{W}_{n+1}^c$ are of the form $\mathcal{C}(w_0, \ldots, w_{k_n-1})$. If the shift at stage n lines up most n-words with other n-words in the same argument of $\mathcal{C}$, then it is possible to build an element of the centralizer of any $\mathbb{K}^c$ having rotation factor β. If not, and we build $\mathbb{K}^c$ suitably randomly, then we can arrange that β is not a central value.

Fix β for the rest of this section, and let $T : \mathcal{K}_\alpha \to \mathcal{K}_\alpha$ be the shift map. The next lemma says that the principal n-blocks of $T^{d^n(s)}(s)$ and $\mathcal{S}_\beta(s)$ are exactly aligned.

Lemma 76. *Let $s, s^* \in \mathcal{K}_\alpha$ be typical and $n < m$ be mature for both. Define*

$$t^* = T^{d^m(s) - d^n(s)}(s^*).$$

Then $t^*(0)$ is in the same position of its principal n-block as $s^*(0)$ is in the principal n-block of s^*. In particular, $T^{d^m(s)-d^n(s)}(s^*)$ has its zero in a position inside an n-word in the construction sequence for some copy of w_n^α.

Proof. Since the n-blocks of s^* repeat on either side of the principal n-block of s^*, and these have length q_n, it suffices to show $d^m(s) - d^n(s) \equiv_{q_n} 0$. Let $t = T^{d^m(s)-d^n(s)}(s)$ and consider the point $s' = T^{d^m(s)}(s)$. Then $s'(0)$ is in the $(r_m(s) + d^m(s))$-th place in its principal m-block. By Lemma 72, $s'(0)$ is in the $(r_n(s) + d^n(s))$-th place in its principal n-block. Since $t = T^{-d^n(s)}(s')$, the point t has its 0 in the $r_n(s)$-th place of its principal n-block. Hence $r_n(t) = r_n(s)$ and by so by Remark 59, $d^m(s) - d^n(s) \equiv_{q_n} 0$. ∎

At first glance Lemma 76 looks puzzling as we are not assuming that any of

$$d^m(s) = d^m(s^*), \quad d^n(s) = d^n(s^*), \quad \text{or} \quad r_n(s) = r_n(s^*).$$

However, the assertion is a statement about how the n-towers sit inside the $n+1$-towers. For mature s, s^* this nesting repeats on either side of the principal n-blocks and hence behaves as in the cyclical approximations. Thus it is independent of the value of $d^n(s^*)$, $d^m(s^*)$ or $r_n(s^*)$, and simply reflects the cyclical structure.

For a particular $s \in \mathcal{K}$, the sequence of shifts $T^{d^n(s)}(s)$ converges to $\mathcal{S}_\beta(s)$. Lemma 76 tells us that this happens promptly: for mature n, $T^{d^n(s)}(s)$ has its 0-th place in the same position of its principal n-block as $\mathcal{S}_\beta(s)$ does.

Consider the location of 0 in the principal $n+1$-block of the point $T^{d^{n+1}(s)-d^n(s)}(s)$ relative to the position of 0 in the principal $n+1$-block of s. For some j_0 and j_1 the principal n-block of $T^{d^{n+1}(s)-d^n(s)}(s)$ arises from the j_0-th argument of $\mathcal{C}(w_n^\alpha, \ldots, w_n^\alpha)$ and the principal n-block of $s(0)$ is in a position coming from the j_1-st argument.

Definition 77. Let $s \in \mathcal{K}$. With indices j_0 and j_1 as just described, the j_0-th argument of $\mathcal{C}(w_n^\alpha, \ldots, w_n^\alpha)$ *β-matches* the j_1-st argument. The point $s \in \mathcal{K}$ is *well-β-matched* at stage n if s is mature at n and $j_0 = j_1$. If n is mature for s and $j_0 \neq j_1$, then s is *ill-β-matched.*

Lemma 78. *Let $\mathbb{K}^c$ be a circular system and consider $S \subseteq \mathbb{K}^c$. Let $s, s^* \in S$ and suppose that n is mature for $\pi(s), \pi(s^*), \mathcal{S}_\beta(\pi(s))$ and $\mathcal{S}_\beta(\pi(s^*))$ and that $\pi(s)$ is well-β-matched at stage n. Let $A_n = d^n(s)$ and $A_{n+1} = d^{n+1}(s)$. Then:*

(1) *one has*
$$r_n(T^{A_n}s^*) = r_n(T^{A_{n+1}}s^*)$$
and

(2) *if I is the interval $[-r_n(T^{A_n}s^*), q_n - r_n(T^{A_n}s^*)) \subseteq \mathbb{Z}$, then*
$$(T^{A_n}s^* \upharpoonright I) = (T^{A_{n+1}}s^* \upharpoonright I). \tag{7.1}$$

Proof. Lemma 76 asserts that 0 is located in the same place in the principal n-block of $T^{A_{n+1}-A_n}(\pi(s^*))(0)$ as 0 is in the principal n-block of $\pi(s^*)$. Since n is mature for s^*, the principal n-block of s^* is repeated on either side of $s^*(0)$. Since n is mature for $\mathcal{S}_\beta(\pi(s^*))$, the principal n-block of $T^{A_{n+1}}s^*$ is repeated at least twice on either side of $T^{A_{n+1}}(s^*)(0)$. It follows that 0 is in the same place in the principal n-block of $T^{A_n}(T^{A_{n+1}-A_n}(s^*))$ as 0 is in the principal n-block of $T^{A_n}(s^*)(0)$. This proves the first assertion.

A repetition of this argument shows the second assertion as well, using the fact that s is well-β-matched. Indeed the definition of *well-β-matched* implies that the principal n-words of $T^{A_{n+1}-A_n}s$ and s are identical. Applying T^{A_n} to both, and using the fact that the principal n-words repeat one sees that the principal n-words $T^{A_{n+1}}s$ and $T^{A_n}s$ are identical. Since the issue of alignment only involves $\pi(s)$, item (2) holds for all s^* with $\pi(s) = \pi(s^*)$. Moreover, arguing as in the last paragraph using the repetition of the principal n-blocks, shifting by an $l < q_n$ does not change this. ∎

Comment. The terminology in this definition extends easily to general circular systems by saying that j_0-th argument and j_1-st arguments are β-matched in $s \in \mathbb{K}^c$ if and only if this is true in s^π, where s^π is the projection of s to $\mathcal{K}$. Similarly we write $d^n(s)$ for $d^n(\pi(s))$.

7.1. The definition of Δ

Let $(X, \mathcal{B}, \mu, T) = (\mathbb{K}^c, \mathcal{B}, \nu, \mathrm{sh})$ be a circular system. Define

$$\Delta_n(\beta) = \nu(\{s : s \text{ is ill-}\beta\text{-matched at stage } n\}) \tag{7.2}$$

and set[31]

$$\Delta(\beta) = \sum_n \Delta_n(\beta). \tag{7.3}$$

Definition 79. The number $\beta \in S^1$ is a *central value* iff $\Delta(\beta) < \infty$.

Note that $\Delta(\beta)$ is defined using the block structure of the $\mathcal{W}_n^c$ and hence is determined by β together with the sequences $\langle k_n \rangle$ and $\langle l_n \rangle$. Thus for $\beta \in S^1$ the property of being *central* depends only on the circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N} \rangle$, rather than on the particular circular system $\mathbb{K}^c$.

In Section 8.1, we show that if $\Delta(\beta)$ is finite, then there is an element T^* in the weak closure of $\{T^n : n \in \mathbb{Z}\}$ such that $(T^*)^\pi = \mathcal{S}_\beta$. In particular, β is the rotation factor of an element of the centralizer. That result does not use the results of the rest of this section.

7.2. Deconstructing $\Delta(\beta)$

Fix a β. Recall that $\langle \varepsilon_n : n \in \mathbb{N} \rangle$ is the sequence satisfying Numerical Requirement 2: $\varepsilon_N > 6 \sum_{n>N} \varepsilon_n$.

Suppose that s is typical, n is mature and s is ill-β-matched. Then there are four possibilities:

(1) $d^n(s) = d_L^n$ or d_R^n and

(2) $d^{n+1}(s) = d_L^{n+1}$ or d_R^{n+1}

Call these possibilities $P_{LL}, P_{LR}, P_{RL}, P_{RR}$.

Lemma 80. *Let $n, m \in \mathbb{N}$ with $n + 1 < m$. There is a partition*

$$\{P^{n,m}_{hd_1,hd_2} : hd_1, hd_2 \in \{L, R\}\} \cup \{U\}$$

of the set $\{0, 1, \ldots, q_m - 1\}$ such that for $s \in S$, if n is mature for s, then

(1) $r_m(s) \in P^{n,m}_{hd_1,hd_2}$ *implies* $(d^n(s), d^{n+1}(s)) = (d^n_{hd_1}, d^{n+1}_{hd_2})$,

(2) $|U| \leq 2q_n + 2q_{n+1}$.

Proof. This follows immediately from Lemma 74 by holding m fixed and applying the lemma successively to n and $n + 1$. Except for a set $U =_{\mathrm{def}} U_n \cup U_{n+1}$ that has at most $2q_n + 2q_{n+1}$ elements, every point in $\{0, 1, \ldots, q_m - 1\}$ belongs to some $P_i^n \cap P_j^{n+1}$. ■

The levels of the q_m-tower reflect the construction of w_m^α from n-words with $n < m$. If s and $\mathcal{S}_\beta(s)$ are mature at stage $n < m$, the locations of $s(0)$ and $T^{d_{n+1}(s) - d_n(s)}(s)(0)$

[31] Since being well or ill-matched only depends on $\pi(s)$ in this section we will not carefully distinguish between s and $\pi(s)$.

in their principal m-block and the pair $(d^n(s), d^{n+1}(s))$ determine whether s is ill-β-matched or not. For particular choices of $hd_1, hd_2 \in \{L, R\}$ either all typical s in P_{hd_1,hd_2} with n mature for both s and $\mathcal{S}_\beta(s)$ are well-β-matched *or* none are.

In the next section we will fix a particular choice of hd_1 and hd_2. For now let n, hd_1 and hd_2 be such that all n-mature s in configuration P_{hd_1,hd_2} are ill-β-matched. We use the symbol $\not\Downarrow_n$ (In LaTeX: `\not\Downarrow`) to indicate the *misaligned* points at stage n. Let

$$\not\Downarrow_n = \{s : s \text{ is ill-}\beta\text{-matched at stage } n \text{ and in configuration } P_{hd_1,hd_2}\}. \tag{7.4}$$

We need to localize the sets $\not\Downarrow_n$. The next lemma tells us that they are uniformly close to open sets:

Proposition 81. *Let $n, m \in \mathbb{N}$ with $n + 1 < m$. Then there is a set*

$$d^{n,m} \subseteq \{0, 1, \ldots, q_m - 1\}$$

such that if $s \in S$, n is mature for s and $r_m(s) + k \in d^{n,m}$, then

(1) *n is mature for $\mathrm{sh}^k(s)$,*

(2) *$d^n(\mathrm{sh}^k(s)) = d^n_{hd_1}$ and $d^{n+1}(\mathrm{sh}^k(s)) = d^{n+1}_{hd_2}$,*

(3) *$\mathrm{sh}^k(s) \in \not\Downarrow_n$ and*

$$\left|\frac{|d^{n,m}|}{q_m} - \nu(\not\Downarrow_n)\right| < 2\left(\frac{q_n + q_{n+1}}{q_m}\right) + \frac{1}{l_{n-1}} + \varepsilon_{n-1}.$$

Proof. Let s be an arbitrary point in S that is mature for n. Take $d^{n,m}$ to be those numbers of the form $r_m(s) + k$ (where $k \in [-r_m(s), q_m - r_m(s))$) such that $\mathrm{sh}^k(s)$ has its zero point in the set $P^{n,m}_{hd_1,hd_2}$ and n is mature for $\mathrm{sh}^k(s)$. Then $d^{n,m}$ is independent of the choice of s. By Lemma 49, the collection of k such that $\mathrm{sh}^k(s)$ is not mature for n has density at most $\frac{1}{l_{n-1}} + \varepsilon_{n-1}$. ∎

7.3. *Red zones*

Suppose that β is not central, i.e. $\Delta(\beta) = \infty$. Then for some fixed choice of (hd_1, hd_2), with hd_i belonging to $\{L, R\}$,

$$\sum_n \nu(\{s : s \text{ is ill-}\beta\text{-matched at stage } n \text{ and in configuration } P_{hd_1,hd_2}\})$$

is infinite. Fix such an hd_1, hd_2. Then with this choice for all n, $\not\Downarrow_n$ is well-defined, and moreover there is a set $G \subseteq \mathbb{N}$ such that if $n < m$ belong to G, then $n + 2 < m$ and

$$\sum_{n \in G} \nu(\not\Downarrow_n) = \infty. \tag{7.5}$$

Let s be a point in $\mathcal{K}_\alpha$ such that all of the shifts of s and $\mathcal{S}_\beta(s)$ are generic with respect to basic open sets, the sets E^i_n, $\not\Downarrow_n$, $P^{n,m}_{hd_1,hd_2}$ and the sets L_n, R_n. For large enough M, we will describe how to use s and the union $\bigcup_{n \in G} \not\Downarrow_n$ to identify a subset of the interval $[-r_M(s), q_M - r_M(s))$ consisting of misaligned points and having density arbitrarily close to one.

Assume that $s \in \not\Downarrow_n$ and n is mature for s and $\mathcal{S}_\beta(s)$. In defining $\not\Downarrow_n$, the choice that $(d^n(s), d^{n+1}(s)) = (hd1, hd2)$ together with $s(0)$, give us the relative locations of the overlap of the principal $n+1$-blocks of s and $\mathcal{S}_\beta(s)$.

Let u be the principal $n+1$-block of s and v be the principal $n+1$-block of $\mathcal{S}_\beta(s)$. and assume that they are in the position determined by $d^{n+1}(s)$. By Lemma 42, on the overlap the 2-subsections of v split the 2-subsections of u into either one or two pieces, and the positions of all of the even pieces are shifted by the same amount relative to the 2-subsections of v and similarly for the odd pieces.

We analyze the case where $s(0)$ occurs in an $n+1$-block, where the 2-subsections are split into two pieces. If they are only split into one piece (i.e. they are not split) the analysis is similar and easier. Without loss of generality we will assume that $s(0)$ occurs in an even overlap.

Since neither $s(0)$, nor $\mathcal{S}_\beta(s)(0)$ occur in the first or last $\varepsilon_n k_n$ 1-subsections of the principal 2-subsection that contains them, we know that the overlaps of the principal 2-subsections of $s(0)$ and $\mathcal{S}_\beta(s)(0)$ contain at least $\varepsilon_n k_n$ 1-subsections. The 0-subsections of the form $w_j^{l_n-1}$ of each 1-subsection of s in this overlap are split into at most three pieces, powers of the form $w_i^{s_0^n}$, w_i^r and $w_i^{s_1^n}$, where $0 \le r \le 2$, $l_n - (s_0^n + s_1^n) \le 3$ and the middle power w_i^r crosses a boundary section of $\mathcal{S}_\beta(s)$. The powers s_0^n and s_1^n are constant on the overlap of the 2-subsections, constant in all of the even pieces of the overlap of the 2-subsections of the principal $n+1$-block, and are determined by (hd_1, hd_2). Moreover, $s_i^n > \varepsilon_n l_n$. Again, without loss of generality we assume that $s(0)$ is in the left overlap corresponding to the power s_0^n.

Observation. There is a number j_0 between 0 and $k_n - 1$ that is determined by the pair $(d^n(s), d^{n+1}(s))$ such that the even piece of a 2-subsection that contains $s(0)$ is of the form

$$\prod_{j<j_0} b^{q_n - j_i} w_j^{l-1} e^{j_i},$$

except that the last 1-subsection may be truncated. Moreover, since $d^{n+1}(\mathrm{sh}^k(s))$ is constant for k in the principal $n+1$-block of s, if

$$t = k_n - j_0, \tag{7.6}$$

then $t \neq 0$ and for all $j < j_0$ the powers $w_j^{s_0^n}$ are β-matched with $w_{j+t}^{s_0^n}$ except for portions of the first and last power.

In particular, if k is such that the 0 position of $\mathrm{sh}^k(s)$ lies in the interior of initial power $w_j^{s_0^n}$ in an even overlap and $j < j_0$, then $\mathrm{sh}^k(s) \in \not\Downarrow_n$ because it is lined up with w_{j+t}.

Lemma 82. *Let $s \in \mathcal{K}$ and suppose that s and $\mathcal{S}_\beta(s)$ are generic, and that s is mature at n. Suppose that $m > n+2$. Then there is a set $B_n \subseteq \{0, \dots, q_m - 1\}$ such that if $k \in [-r_m(s), q_m - r_m(s))$ and $r_m(s) + k \in B_n$, then:*

(1) $\mathrm{sh}^k(s)$ *has its zero located in B_n,*

(2) *n is mature for* $\mathrm{sh}^k(s)$*,*

(3) $\mathrm{sh}^k(s) \in \not\Downarrow_n$*,*

(4) *there exist a* $j_0 > \varepsilon_n k_n$ *and a* $t \neq 0$ *such that* B_n *is:*

   (a) *a union of sets, each of the form* $\bigcup_{j<j_0} U_j$,

   (b) *each set* $\bigcup_{j<j_0} U_j$ *is a subset of a position of an occurrence in* s *of an* $n+1$*-subword* $\mathcal{C}(u_0, u_1, \dots, u_{k_n-1})$ *of* w_m^α *(with* $u_i = w_n^\alpha$*),*

   (c) *each* U_j *is a collection of non-n-boundary positions in* $u_j^{s_0^n}$ *such that* $u_j^{s_0^n}$ *is* β*-matched with* $u_{j+t}^{s_0^n}$*, except perhaps for the first or last copy of* u_j *in* $u_j^{s_0^n}$*, and*

   (d) *each set* $\bigcup_{j<j_0} U_j$ *is the collection of all non-n-boundary positions in* $u_j^{s_0^n}$ *in a block of the form*

$$\prod_{j<j_0} b^{q_n - j_i} u_j^{l_n - 1} e^{j_i}.$$

and

$$\left| \frac{B_n}{q_m} - \nu(\mathbb{V}_n) \right| < 2\left(\frac{q_n + q_{n+1}}{q_m} \right) + \frac{1}{l_{n-1}} + \varepsilon_{n-1}.$$

Proof. The first statement is automatic since $B_n \subseteq \{0, 1, \dots, q_m - 1\}$. Let $d^{n,m}$ be as in Proposition 81. If $k \in d^{n,m}$, then, as in the discussion before the statement of Lemma 82, $\mathrm{sh}^k(s)(0)$ occurs in the position of a power $u^{s_0^n}$, where u is the principal n-block of $\mathrm{sh}^k(s)$ and $u^{s_0^n}$ occurs on the left overlap of 1-subsections of the principal $n+1$-block of $\mathrm{sh}^k(s)$.

As in the observation before this lemma, to each $k \in d^{n,m}$ we can associate a set $\bigcup_{j<j_0} U_j$ containing k by taking all of the positions of the powers $u_j^{s_0^n}$ in the even overlap determined $\mathrm{sh}^k(s)(0)$, where k is not in the boundary of a u_j. Let B_n be the union of all of the collections $\bigcup_{j<j_0} U_j$ as k ranges over $d^{n,m}$.

Assertion (4) (c) follows from the observation and the fact that d^n and d^{n+1} are constant (and equal to $d^n_{hd_1}$ and $d^{n+1}_{hd_2}$) on $d^{n,m}$.

We show that if $k' \in B_n$, then n is mature for $\mathrm{sh}^{k'}(s)$ and that $\mathrm{sh}^{k'}(s) \in \mathbb{V}_n$. The maturity of n follows immediately from the maturity of s and the fact that the location of 0 in $\mathrm{sh}^k(s)$ is in a non-boundary portion of an n-subword of its principal $n+1$-block. That $\mathrm{sh}^{k'}(s) \in \mathbb{V}_n$ follows from the fact that $u_j^{s_0^n}$ is β-matched with $u_{j+t}^{s_0^n}$, and $t \neq 0$.

To finish, note that

$$d^{n,m} \subseteq \bigcup \bigcup_{j<j_0} U_j \subseteq \mathbb{V}_n.$$

Hence

$$\frac{|d^{n,m}|}{q_n} \leq \frac{|\bigcup \bigcup_{j<j_0} U_j|}{q_n} \leq \nu(\mathbb{V}_n).$$

Thus Lemma 82 follows from Lemma 81. ∎

We now define the *red zones* corresponding to β. Recall that if $n < m \in G$, then

$$n + 2 < m \quad \text{and} \quad \sum_{n \in G} \nu(\mathbb{V}_n) = \infty.$$

For $n < m$ consecutive elements of G, define

$$\delta_n = 4\left(\frac{q_{n+1}}{q_m} \right) + \frac{1}{l_{n-1}} + \varepsilon_{n-1}.$$

Then we see that:

- $\sum_{n\in G} \delta_n < \infty$, so
- $\sum_{n\in G} (\nu(\nparallel_n) - \delta_n) = \infty$,

and if B_n is the set defined in Lemma 82, then $\nu(\nparallel_n) - \delta_n \leq \frac{|B_n|}{q_m} \leq \nu(\nparallel_n)$.

Lemma 83. *Let N be a natural number and $\delta > 0$. Suppose that s and $\mathcal{S}_\beta(s)$ are generic, and that s is mature at N. Then there is a sequence of natural numbers $\langle n_i : 1 \leq i \leq i^* \rangle$, an M and sets $R_i \subseteq \{0, 1, 2, \ldots, q_M - 1\}$, for $1 \leq i \leq i^*$, such that*

(1) *$N < n_1$ and $n_i + 2 < n_{i+1} < M$,*

(2) *R_i is disjoint from R_j for $i \neq j$,*

(3) *R_i is a union of blocks of the form B_{n_i} described in condition (4) in Lemma 82 inside n_{i+1}-subwords of w_M^α,*

(4) *if $k \in R_i$, then $\mathrm{sh}^k(s) \in \nparallel_{n_i}$,*

(5) *the density of $\bigcup_i R_i$ in $\{0, 1, \ldots, q_M - 1\}$ is at least $1 - \delta$.*

Proof. We can assume that N is so large that $\bigcup_{n\geq N} \partial_n$ has measure less than $\delta/100$ and $1/l_N + \varepsilon_N < \delta/100$. From the definition of G we can find a collection $\langle n_i : i \leq i^* \rangle$ of consecutive elements of G so that

$$\prod_{1\leq i\leq i^*} (1 - \nu(\nparallel_{n_i}) + \delta_{n_i}) < \frac{\delta}{100}.$$

Choose an $M > n_{i^*} + 2$, and for notation purposes set $n_{i^*+1} = M$.

Define sets R_i and I_i by reverse induction from $i = i^*$ to $i = 1$ with the following properties:

(i) $\{0, 1, \ldots, q_M - 1\} \setminus ((\bigcup_{i^*\geq j\geq i} I_j) \cup (\bigcup_{i^*\geq j\geq i} R_j))$ consists of entire locations of words $w_{n_i}^\alpha$ in w_M^α,

(ii) $R_i \subseteq \{0, 1, \ldots, q_M - 1\} \setminus ((\bigcup_{i^*\geq j> i} I_j) \cup (\bigcup_{i^*\geq j> i} R_j))$ and has relative density at least $\nu(\nparallel_{n_i}) - \delta_{n_i}$,

(iii) the set $I_i \subseteq \bigcup_{j=n_i+1}^{n_{i+1}} \partial_j \cap \{0, 1, \ldots, q_M - 1\}$ and hence,

(iv) I_i has density less than or equal to $1/l_{n_i}$ in $\{0, 1, \ldots, q_M - 1\}$

To start, apply Lemma 82 with $m = n_{i^*+1}$, to get a set $B_{n_{i^*}} \subseteq \{0, 1, \ldots, q_M - 1\}$ of density at least $\nu(\nparallel_{n_{i^*}}) - \delta_{n_{i^*}}$ satisfying conditions (3)–(4) of the lemma we are proving. Set $R_{i^*} = B_{n_{i^*}}$. Let

$$I_{i^*} = \bigcup_{j=n_{i^*}+1}^{M} \partial_j \cap \{0, 1, \ldots, q_M - 1\}.$$

Suppose that we have defined $\langle R_j : i^* \geq j > i \rangle$ and $\langle I_j : i^* \geq j > i \rangle$ satisfying the induction hypothesis (i)–(iv).

Apply Lemma 82 again to get a set $B = B_{n_i}$ a subset of $\{0, 1, \ldots, q_{n_{i+1}} - 1\}$. Inside each copy $\{k, k + 1, \ldots, k + q_{n_{i+1}} - 1\}$ corresponding to a location in w_M^α of a $w_{n_{i+1}}^\alpha$

in the complement of $((\bigcup_{i^*\geq j>i} I_j) \cup (\bigcup_{i^*\geq j>i} R_j))$, we have a translated copy of B, $k+B$. Let R_i be the union of the sets $k+B$, where k runs over the locations the words $w^\alpha_{n_{i+1}}$ in the complement of $((\bigcup_{i^*\geq j>i} I_j) \cup (\bigcup_{i^*\geq j>i} R_j))$.

Then the density of R_i relative to

$$\{0,1,\dots,q_M-1\} \setminus \left(\left(\bigcup_{i^*\geq j>i} I_j\right) \cup \left(\bigcup_{i^*\geq j>i} R_j\right)\right)$$

is at least $\nu(\not{\mathbb{K}}_{n_i}) - \delta_{n_i}$. It follows from conclusion (3) of Lemma 82 that R_i is a union of non-boundary portions of blocks of length $q_{n_i}^{s^0_{n_i}-1}$ corresponding to positions of $w^\alpha_{n_i}$ in w^α_M.

Since R_i consists of a union of the non-boundary portion of locations of words $w^\alpha_{n_i}$,

$$\{0,1,\dots,q_M-1\} \setminus \left(\left(\bigcup_{i^*\geq j>i} I_j\right) \cup \left(\bigcup_{i^*\geq j>i} R_j\right) \cup R_i\right)$$

consists of the entire blocks of locations of $w^\alpha_{n_i}$ together with elements of $\bigcup_{j=n_i}^{n_{i+1}} \partial_j$. The latter set has density less than or equal to $1/l_{n_i-1}$. Let

$$I_i = \left(\{0,1,\dots,q_M-1\} \cap \bigcup_{j=n_i}^{n_{i+1}} \partial_j\right) \setminus \left(\left(\bigcup_{i^*\geq j>i} I_j\right) \cup \left(\bigcup_{i^*\geq j>i} R_j\right) \cup R_i\right).$$

It remains is to calculate the density of $\bigcup_{1\leq i\leq i^*} R_i$. At each step in the induction, we remove a portion of density at least $\nu(\not{\mathbb{K}}_{n_i}) - \delta_{n_i}$ from

$$\{0,1,\dots,q_M-1\} \setminus \left(\left(\bigcup_{i^*\geq j>i} I_j\right) \cup \left(\bigcup_{i^*\geq j>i} R_j\right)\right).$$

Let $\partial = \bigcup_{1\leq i\leq M} \partial_{n_i}$. Then the density of the union of the sets R_i is at least

$$1 - \prod_{i^*\geq i\geq 1} (1 - \not{\mathbb{K}}_{n_i}) - \frac{|\partial|}{q_m},$$

which is at least $1-\delta$. ■

8. The centralizer and central values

In the first part of this section we show that every central value is a rotation factor of an element of the closure of the powers of T and hence an element of the centralizer.

The second part shows a converse: if $\mathbb{K}^c$ is built sufficiently randomly, then the rotation factor of every element of the centralizer is a rotation by a central value.

We note in passing that every circular system is *rigid*: if s is mature for n, then $T^{q_n(l_n-2)}(s)$ has the same principal n-block as s does. It follows that $\overline{\{T^n : n \in \mathbb{Z}\}}$ is a perfect Polish monothetic group.

8.1. Building elements of the centralizer

If $\Delta(\beta)$ is finite, then the Borel–Cantelli lemma implies that for ν-almost every s, there is an n_0 such that for all $n \geq n_0$, s is well-β-matched at stage n. As a consequence, certain sequences of translations converge. Precisely:

Theorem 84. *Suppose that $\mathbb{K}^c$ is a uniform circular system with coefficient sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$. Let T be the shift map on $\mathbb{K}^c$ and let $\beta \in [0, 1)$ be a number such that $\Delta(\beta) < \infty$. Then there is a sequence of integers $\langle A_n : n \in \mathbb{N}\rangle$ such that $\langle T^{A_n} : n \in \mathbb{N}\rangle$ converges pointwise almost everywhere to a $T^* \in C(T)$ with $(T^*)^\pi = \mathcal{S}_\beta$. In particular, there is a sequence $\langle A_n : n \in \mathbb{N}\rangle$ such that $\langle T^{A_n} : n \in \mathbb{N}\rangle$ converges in the weak topology to a T^* with $(T^*)^\pi = \mathcal{S}_\beta$.*

Corollary 85. *If β is central, then there is a $\phi \in \overline{\{T^n : n \in \mathbb{Z}\}}$ such that $\phi^\pi = \mathcal{S}_\beta$.*

Proof. Let $\mathcal{T}$ be the tree of finite sequences $\sigma \in \{L, R\}^{<\infty}$. Choose an n_0 such that

$$G = \{s : n_0 \text{ is mature for } s \text{ and for all } m \geq n_0, s \text{ is well-}\beta\text{-matched at stage } m\}$$

has positive measure. By the König Infinity Lemma there is a function

$$f : \{m : m \geq n_0\} \to \{L, R\}$$

such that for all $m \geq n_0$, $\{s \in G : d^n(s) = d^n_{f(n)}$ for all n with $n_0 \leq n \leq m\}$ has positive measure. Let $A_n = d^n_{f(n)}$.

By Lemma 75, item (3) it follows that for a typical s the left and right endpoints of the principal n-blocks of $T^{A_n}s$ go to negative and positive infinity, respectively. Let s^* be a typical element of S; e.g. $\pi(s^*)$ and $\mathcal{S}_\beta(\pi(s^*))$ both belong to S^π, large enough n are mature for s^* and for all large n, $\pi(s^*)$ is well-β-matched at stage n. Then for all large n, the left and right endpoints of the principal n-block of $T^{A_n}s$ and $T^{A_{n+1}}s$ are the same. If s^* is well-β-matched at stage n, then the words constituting principal n-block of $T^{A_n}s$ and $T^{A_{n+1}}s$ are the same. It follows that for typical $s^* \in S$, the sequence $T^{A_n}s^*$ converges in the product topology on $(\Sigma \cup \{b, e\})^{\mathbb{Z}}$.

We now show that the map

$$s \mapsto \lim T^{A_n}s$$

is one-to-one. If $s \neq s'$, then either $\pi(s) \neq \pi(s')$ or there is an N such that for all $n \geq N$ the principal n-blocks of s and s' differ. We can assume that this N is so large that n is mature and well-β-matched for $\pi(s), \pi(s')$.

If $\pi(s) \neq \pi(s')$, then $\mathcal{S}_\beta(\pi(s)) \neq \mathcal{S}_\beta(\pi(s'))$. Hence the limits of $T^{A_n}s$ and $T^{A_n}s'$ differ. So assume that $\pi(s) = \pi(s')$. Then, since T^{A_n} is a translation by at most $q_n - 1$ and n is mature for all parties (so the principal n-blocks of $T^{A_n}s$ and $T^{A_n}s'$ repeat), we know that the principal n-blocks of $T^{A_n}s$ and $T^{A_n}s'$ differ. But for all $m > n$, the principal n-blocks of $T^{A_m}s$ agree with the principal n-blocks of $T^{A_n}s$ (and similarly for s'). Hence for all $m > N$ the principal N-blocks of $T^{A_m}s$ and $T^{A_m}s'$ differ. It follows that the limit map is one-to-one.

We need to see that for almost all s, $\lim_{n\to\infty} T^{A_n}s$ belongs to $\mathbb{K}^c$. By the definition of $\mathbb{K}^c$ this is equivalent to showing that for almost all s if $I \subseteq \mathbb{Z}$ is an interval, then

$\lim_{n\to\infty} T^{A_n} s \upharpoonright I$ is a subword of some $w \in \mathcal{W}_m^c$ for some m. However, by Lemma 78, for almost all s we can find an n so large that:

(1) $I \subseteq [-r_n(s), q_n - r_n(s))$,

(2) $T^{A_n} s$ and $\lim_{n\to\infty} T^{A_n} s$ agree on the location of the principal n-block of containing I, and

(3) $T^{A_n} s$ and $\lim_{n\to\infty} T^{A_n} s$ agree on what word lies on the principal n-block.

Since the principal n-block of $T^{A_n} s$ belongs to $\mathcal{W}_n^c$, we are done.

Summarizing, if one has $T^* = \lim_{n\to\infty} T^{A_n} s$, then for almost all s, $T^* s$ is defined and belongs to S. Moreover, T^* is one-to-one and commutes with the shift map.

Define a measure ν^* on S by setting $\nu^*(A) = \nu((T^*)^{-1} A)$. Then ν^* is a non-atomic, shift invariant measure on S. By Lemma 38, we must have $\nu^* = \nu$. In particular, we have shown that $T^* : \mathbb{K}^c \to \mathbb{K}^c$ is an invertible measure preserving transformation belonging to $\overline{\{T^n : n \in \mathbb{Z}\}}$, with $(T^*)^\pi = \mathcal{S}_\beta$. ∎

We make the following remark without proof as it is not needed in the sequel:

Remark 86. Suppose that $\mathbb{K}^c$ satisfies the hypothesis of Theorem 84 and β is a central value. Then for any sequence of natural numbers $\langle A_n : n \in \mathbb{N}\rangle$ such that $A_n \alpha$ converges to β sufficiently fast, the sequence $\langle T^{A_n} : n \in \mathbb{N}\rangle$ converges to a $T^* \in C(T)$ with $(T^*)^\pi = \mathcal{S}_\beta$.

8.2. Characterizing central values

The main result of this subsection is a converse of Corollary 85. If $\mathbb{K}^c$ is a circular system built from sufficiently random collections of words and ϕ is an isomorphism between $\mathbb{K}^c$ and $\mathbb{K}^c$, then $\phi^\pi = \mathcal{S}_\beta$ for some central β. Moreover, if ϕ is an isomorphism between $\mathbb{K}^c$ and $(\mathbb{K}^c)^{-1}$, then ϕ^π is of the form $\mathrm{rev}(\cdot) \circ \natural \circ \mathcal{S}_\beta$ for some central β.

In this subsection we will return to considering $(\mathbb{K}^c)^{-1}$ as $(\mathrm{rev}(\mathbb{K}^c), \mathrm{sh})$ with the forward shift, and hence can use $\natural$ instead of $\mathrm{rev}(\cdot) \circ \natural$.

8.2.1. The timing assumptions. Randomness assumptions about the words in the sets $\mathcal{W}_n^c$ will allow us to assert that the rotations associated with elements of the centralizer of $\mathbb{K}^c$ or isomorphisms between $\mathbb{K}^c$ and $(\mathbb{K}^c)^{-1}$ arise from central elements β. The last part of the paper shows that these additional randomness assumptions are consistent with the randomness assumptions used in [8] and describes how to build words with both collections of specifications.

Recall from Definition 34 that to specify a circular system with coefficient sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$ it suffices to inductively specify collections of prewords $P_{n+1} \subseteq (\mathcal{W}_n^c)^{k_n}$, and define $\mathcal{W}_{n+1}^c$ as the collection of words

$$\{\mathcal{C}(w_0, \dots, w_{k_n-1}) : w_0 w_1 \dots w_{k_n-1} \in P_{n+1}\}.$$

In the construction, there will be an equivalence relation $\mathcal{Q}_1^1$ on $\mathcal{W}_1^c$ that is lifted from an analogous equivalence relation on the first step of the odometer construction $\mathcal{W}_1$. It is built in Section 10; we describe its properties here. Let $\langle \mathcal{Q}_1^n : n \in \mathbb{N}\rangle$ be the sequence of propagations of $\mathcal{Q}^1_1$. As the construction progresses there are groups G^n_1 acting freely on the set of $\mathcal{Q}^n_1$ equivalence classes of words in $\mathcal{W}^c_n$. Each G^n_1 is a finite sum of copies of $\mathbb{Z}_2$. Inductively, $G^{n+1}_1 = G^n_1$ or $G^{n+1}_1 = G^n_1 \oplus \mathbb{Z}_2$. The action of G^n_1 on $\mathcal{W}^c_{n+1}$ arising from the G^{n+1}_1 action via the inclusion map of G^n_1 into G^{n+1}_1 is the skew-diagonal action. We will write $[w]_1$ for the $\mathcal{Q}^n_1$-equivalence class of a $w \in \mathcal{W}^c_n$ and $G^n_1[w]_1$ for the orbit of $[w]_1$ under G^n_1. If $w \in \mathcal{W}^c_{n+1}$ and $C \in \mathcal{W}^c_n/\mathcal{Q}^n_1$, then we say that C *occurs at* t if there is a $v \in \mathcal{W}^c_n$ sitting on the interval $[t, t+q_n)$ inside w and $C = [v]_1$.

Numerical Requirement 4. One has

$$\sum \frac{|G^n_1|}{|\mathcal{Q}^n_1|} < \infty.$$

This can be satisfied by taking $\frac{|G^n_1|}{|\mathcal{Q}^n_1|} < 2^{-n}$.

We note that G^n_1 is determined directly by the first n-nodes in tree we are using in the domain of the reduction, and hence $|G^n_1|$ is determined by the tree. So this requirement on $|\mathcal{Q}^n_1|$ does not depend on any of the other variables being chosen during the construction. In what follows we call such requirements *absolute* requirements.

Notation. As an aid to tracking corresponding variables, script letters are used for sets and non-script Roman letters for the corresponding cardinalities. For example we will use $\mathcal{Q}_n$ for an equivalence relation and Q_n for the number of classes in that equivalence relation.

Here are the assumptions used to prove the converse to Corollary 85. The first three assumptions follow immediately from the definitions in Section 5.10.

(T1) The equivalence relation $\mathcal{Q}^{n+1}_1$ is the equivalence relation on $\mathcal{W}^c_{n+1}$ propagated from $\mathcal{Q}^n_1$.

(T2) The group G^n_1 acts freely on $\mathcal{W}_n/\mathcal{Q}^n_1 \cup \mathrm{rev}(\mathcal{W}_n/\mathcal{Q}^n_1)$

(T3) The canonical generators of the group G^n_1 send elements of $\mathcal{W}^c_n/\mathcal{Q}^n_1$ to elements of $\mathrm{rev}(\mathcal{W}^c_n/\mathcal{Q}^n_1)$ and vice versa.

The next axiom states that the $\mathcal{Q}^n_1$ classes are widely separated from each other.

(T4) There is a number γ such that $0 < \gamma < 1/4$ such that for each n and each pair $w_0, w_1 \in \mathcal{W}^c_n \cup \mathrm{rev}(\mathcal{W}^c_n)$ and each $j \geq q_n/2$ if $[w_0]_1 \neq [w_1]_1$, then

$$\bar{d}(w_0 \restriction [0, j), w_1 \restriction [0, j)) \geq \gamma,$$
$$\bar{d}(w_0 \restriction [q_n - j, q_n), w_1 \restriction [q_n - j, q_n)) \geq \gamma$$
$$\bar{d}(w_0 \restriction [0, j), w_1 \restriction [q_n - j, q_n)) \geq \gamma.$$

Remark 87. In axioms (T5)–(T7) we write $|x_n| \approx \frac{1}{y_n}$ to mean that $||x_n| - \frac{1}{y_n}| < \mu_n$, where $\mu_n \ll \min(\varepsilon_n, 1/Q^n_1)$.

Numerical Requirement 5. μ_n is chosen small relative to $\min(\varepsilon_n, 1/Q^n_1)$. Explicitly: if $t_n = \min(\varepsilon_n, 1/Q^n_1)$, then

$$0 < \mu_n < t_n \min_{k \leq n} 2^{-n-2} \frac{1}{t_k}.$$

In the next assumption we count the occurrences of particular n-word v that are lined up in an $n+1$-preword w_0 with the occurrences of a particular $\mathcal{Q}_1^n$-class in the shift of another $n+1$-preword w_1 or its reverse. The shift (by t n-subwords), must be non-zero and be such that there is a non-trivial overlap after the shift.

(T5) Let w_0, w_1 be prewords in P_{n+1}, and w_1' be either w_1 or $\mathrm{rev}(w_1)$. Write

$$w_0 = v_0 v_1 \dots v_{k_n-1} \quad \text{and} \quad w_1' = u_0 u_1 \dots u_{k_n-1},$$

with $u_i, v_j \in \mathcal{W}_n^c \cup \mathrm{rev}(\mathcal{W}_n^c)$. Let $\mathcal{C} \in \mathcal{W}_n^c/\mathcal{Q}_1^n$ or $\mathcal{C} \in \mathrm{rev}(\mathcal{W}_n^c)/\mathcal{Q}_1^n$ according to whether $w_1' = w_1$ or $w_1' = \mathrm{rev}(w_1)$. For all integers t with $1 \le t \le (1-\varepsilon_n)(k_n)$, $v \in \mathcal{W}_n^c$, we have:

(a) (This is comparing w_0 with $\mathrm{sh}^{tq_n}(w_1')$.) Let

$$J(v) = \{k < k_n - t : v = v_k\}.$$

Then

$$\frac{|\{k \in J(v) : u_{t+k} \in \mathcal{C}\}|}{|J(v)|} \approx \frac{1}{Q_1^n}.$$

(b) (This is comparing $\mathrm{sh}^{tq_n}(w_0)$ with w_1'.) Let

$$J(v) = \{k : t \le k \le k_n - 1 \text{ and } v = v_k\}.$$

Then

$$\frac{|\{k \in J(v) : u_{t-k} \in \mathcal{C}\}|}{|J(v)|} \approx \frac{1}{Q_1^n}.$$

(T6) Suppose that $w_0 w_1 \dots w_{k_n-1}, w_0' w_1' \dots w_{k_n-1}' \in P_{n+1}$ are prewords, and suppose that $1 \le t \le (1-\varepsilon_n)k_n$ and $\varepsilon_n k_n \le j_0 \le k_n - t$. Let

$$S = \{k < j_0 : \text{for some } g \in G_1^n, g[w_k]_1 = [w_{k+t}']_1\}.$$

Then

$$\frac{|S|}{j_0} \approx \frac{|G_1^n|}{Q_1^n}.$$

(T7) Let w_0, w_1 be prewords in P_{n+1}, and let w_1' be either w_1 or $\mathrm{rev}(w_1)$. Suppose that $[w_1']_1 \notin G_1^n[w_0]_1$. Write

$$w_0 = v_0 v_1 \dots v_{k_n-1} \quad \text{and} \quad w_1' = u_0 u_1 \dots u_{k_n-1},$$

with $u_i, v_j \in \mathcal{W}_n^c \cup \mathrm{rev}(\mathcal{W}_n^c)$. Let $\mathcal{C} \in \mathcal{W}_n^c/\mathcal{Q}_1^n$ or $\mathcal{C} \in \mathrm{rev}(\mathcal{W}_n^c)/\mathcal{Q}_1^n$ according to whether $w_1' = w_1$ or $w_1 = \mathrm{rev}(w_1)$. Then for all $v \in \mathcal{W}_n^c$ if

$$J(v) = \{t : v_t = v\},$$

then

$$\frac{|\{t \in J(v) : u_t \in \mathcal{C}\}|}{|J(v)|} \approx \frac{1}{Q_1^n}. \tag{8.1}$$

Definition 88. We will call the collection of axioms (T1)–(T7) the *timing assumptions* for a construction sequence and an equivalence relation $\mathcal{Q}_1^1$.

8.2.2. Codes and $\bar{d}$-distance. We now prove some lemmas about $\bar{d}$.[32]

Lemma 89. *Let $w_0 \in \mathcal{W}^c_{n+1}$ and $w_1 \in \mathcal{W}^c_{n+1} \cup \mathrm{rev}(\mathcal{W}^c_{n+1})$ such that $[w_0]_1 \notin G^n_1[w_1]_1$. Let $r > 1000$ and let J_0, J_1 be intervals in $\mathbb{Z}$ of length $r * q_{n+1}$. Let I be the intersection of the two intervals. Put w_0^r on J_0 and w_1^r on J_1 and suppose that all but (possibly) the first or last copies of w_0 are included in I. Let $\bar{\Lambda}$ be a stationary code such that the length of Λ is less than $q_n/10000$. Then:*

$$\bar{d}(\bar{\Lambda}[w_0^r \upharpoonright I], w_1^r \upharpoonright I) > \frac{1}{50}\left(1 - \frac{1}{Q^n_1}\right)\gamma. \tag{8.2}$$

Proof. Since the length of the code Λ is much smaller than q_n and $r > 10000$, the end effects of Λ are limited to the first and last copies of w_0 and thus affect at most $(1/5000)$ proportion of $\bar{d}(\bar{\Lambda}[w_0^r \upharpoonright I], w_1^r \upharpoonright I)$. Removing the portion of I across from the first or last copy of w_0 leaves a segment of I of proportion at least $4999/5000$.

For all of the copies of w_0, except perhaps at most one at the end of J_0, there is a corresponding copy of w_1 that overlaps w_0 in a section of at least $q_{n+1}/2$. Discard the portions of I arising from copies of w_0 not overlapping the corresponding copies of w_1. After the first two removals we have a portion of I of proportion at least $(1/2)(4999/5000)$.

Because w_0 and w_1 have the same lengths, the relative alignment between any two corresponding copies of w_0 and w_1 in the powers w_0^r and w_1^r are the same. In particular, the "even overlaps" and "odd overlaps" are the same in each remaining portion of the corresponding copies of w_0 and w_1.

By Lemma 42, there are $s, t < q_n$ such that on the even overlaps all of the n-subwords of $\mathrm{sh}^s(w_0^r)$ are either lined up with an n-subword of w_1^r or with a boundary section of w_1, and all of the n-subwords of w_0 in an odd overlap are lined up with an n-subword or a boundary section of w_1^r by $\mathrm{sh}^t(w_0^r)$.

Either the even overlaps or the odd overlaps contain at least $1/2$ of the n-subwords that are not across from boundary portions of w_1. Assume that $1/2$ of the n-subwords lie in even overlaps and discard the portion of I on the odd overlaps. (If more than $1/2$ of the n-subwords are in odd overlaps, we would focus on those.)

Let $(w_0^*)^r = \mathrm{sh}^s(w_0^r)$ on the even overlaps. Denote any particular copy of w_0 in $(w_0^*)^r$ as w_0^*. Then, except for $\mathcal{W}^c_n$-words that get lined up with a boundary section of w_1, every n-subword of $(w_0^*)^r$ coming from an even overlap of $(w_0)^r$ gets lined up with an n-subword of $(w_1)^r$. Write $w_0 = \mathcal{C}(v_1, v_2, \ldots, v_{k_n-1})$ and $w_1 = \mathcal{C}(u_1, u_2, \ldots, u_{k_n-1})$ (or, respectively, $w_1 = \mathcal{C}^r(\mathrm{rev}(u_1), \mathrm{rev}(u_2), \ldots, \mathrm{rev}(u_{k_n-1}))$). Then each n-subword of w_0^* coming from an even overlap is of the form v_i for some i. There is a t such that for all i if v_i occurs in any copy of w_0^* and comes from an even overlap, then either:

(a) v_i is lined up with u_{i+t} (respectively $\mathrm{rev}(u_{k_n-(i+t)-1})$) or

(b) v_i is lined up with a boundary portion of w_1 or

(c) v_i is lined up with u_{i+t+1} (respectively $\mathrm{rev}(u_{k_n-(i+t+1)-1})$).

[32] Basic notation and facts about stationary codes are reviewed in Section 4.4.

On copies of v_i coming from even overlaps of 2-subsections the powers of v_i in alternatives a.) and c.) are constant. Since the even overlaps of the 2-subwords has size at least half of the lengths of the 2-subwords, it follows that $0 \leq t \leq k_n/2$.

Since all of $v_i^{l_n-1}$ satisfies (a), (b), or (c), after discarding the words v_i in case (b) half of the remaining words v_i satisfy (a) or (c). Keep the larger alternative and discard the other. What is left after all of the trimming has size at least

$$(4999/5000)(1/2)(1/2)(1-2|\partial_{n+1}|) > 1/10$$

proportion of I.

For some t what remains consists of n-subwords v_i in even overlaps of $(w_0)^r$ that, after being shifted by s to be subwords of $(w_0^*)^r$, are aligned with occurrences of n-subwords of $(w_1)^r$ of the form u_{i+t} ($\mathrm{rev}(u_{k_n-(i+t)-1})$ respectively). For the rest of this proof of Lemma 89 we will call these the *good occurrences* of n-subwords.

Claim. *Suppose that $v \in \mathcal{W}_n^c$ and let*

$$J^*(v) = \{y \in I : y \textit{ is at the beginning of a good occurrence of } v \textit{ in } (w_0^*)^r\}.$$

Furthermore, let $\mathcal{C} \in \mathcal{W}_n^c/\mathcal{Q}_1^n$ or $\mathcal{C} \in \mathrm{rev}(\mathcal{W}_n^c)/\mathcal{Q}_1^n$ depending on whether $w_1 \in \mathcal{W}_{n+1}^c$ or $w_1 \in \mathrm{rev}(\mathcal{W}_{n+1}^c)$. Then

$$\left| \frac{|\{y \in J^*(v) : \textit{some element of } \mathcal{C} \textit{ occurs at } y \textit{ in } w_1\}|}{|J^*(v)|} - \frac{1}{Q_1^n} \right| \tag{8.3}$$

is bounded by $2/q_n + 2/l_n + \mu_n$.

We prove the claim. We have two cases:

Case 1: $t = 0$. In this case we have a *trivial split* in the language of Section 5.4. The overlap of the 2-subsections contains the whole of the two subsections except for a portion of one 1-subsection. Since $[w_0]_1 \notin G_1^n[w_1]_1$, we can apply axiom (T7) to the words w_0 and w_1. The claim follows from inequality (8.1), which is the preword version of formula (8.3), after taking into account the boundary and the words at the ends of the blocks of $(w_0^*)^r$ and the truncated 1-subsections.

Case 2: $t \neq 0$. In this case the split is non-trivial. Because the even overlaps are at least as big as the odd overlaps of 2-subsections, the even overlap looks like

$$\prod_{j=0}^{t^*} (b^{q-j_i} v_j^{l-1} e^{j_i})$$

but with a portion of its last 1-subsection possibly truncated. In particular, it has an initial segment of the form

$$\prod_{j=0}^{t^*-1} (b^{q-j_i} v_j^{l-1} e^{j_i}),$$

where $t^* \geq k_n/2$.

It follows from the timing assumption (T5) that if $J' = \{y \in J(v) :$ some element of $\mathcal{C}$ occurs across from a word starting at y in the first $t^* - 1$ 1-subsections$\}$, then

$$\left| \frac{|J'|}{|J(v)|} - \frac{1}{Q_1^n} \right| < \mu_n.$$

Any variation between the quantity in formula (8.3) and the estimate in (T5) is due to the portion of the last 1-subsection of the even overlaps. This takes up a proportion of the remaining even overlap less than or equal to $1/t^* \leq 2/q_n$. This proves the Claim.[33]

We now shift $(w_0^*)^r$ back to be w_0^r and consider s. There is an $l' \geq l/2 - 1 \geq l/3$ such that all of the good occurrences of a $v \in \mathcal{W}_n^c$ in $(w_0^*)^r$ are in a power $v^{l'}$. Depending on whether $s \leq q_n/2$ or $s > q_n/2$, for each good occurrence of a v_j in $(w_0^*)^r$ either:

(a) there are at least $l' - 1$ powers of v_j in the corresponding occurrence in w_0 such that their left overlap with u_{j+t} has length at least $q_n/2$ or

(b) there are at least $l' - 1$ powers of v_j in the corresponding occurrence in w_0 such that their right overlap with u_{j+t} has length at least $q_n/2$

Without loss of generality we assume alternative (a). Suppose that the overlap has length o in all of the good occurrences. Then the left side of v_j overlaps the right side of u_{j+t} by at least $q_n/2$.

By axiom (T4), if $v \in \mathcal{W}_n^c$,

$$\begin{aligned} \bar{d}(\bar{\Lambda}[(v \upharpoonright [0,o)], u_{j+t} \upharpoonright [q_n - o - 1, q_n)) < \gamma/2, \\ \bar{d}(\bar{\Lambda}(v \upharpoonright [0,o), u_{j'+t} \upharpoonright [q_n - o - 1, q_n)) < \gamma/2, \end{aligned}$$

then $[u_{j+t}]_1 = [u_{j'+t}]_1$. It follows that if we fix a $v \in \mathcal{W}_n^c$ and let

$$J(v) = \{j : v_j = v\},$$

then

$$\frac{|\{j \in J(v) : \bar{d}(c(v_j \upharpoonright [0,o), u_{j+t} \upharpoonright [q_n - o - 1, q_n - 1)) < \gamma/2\}|}{|J(v)|} < \frac{1}{Q_1^n} + \mu_n.$$

Since at least $1/20$ proportion of I consists of left halves of good occurrences of the various v's belonging to $\mathcal{W}_n^c$, it follows that

$$\bar{d}(\bar{\Lambda}[w_0^r \upharpoonright I], w_1^r) \geq \frac{1}{20}\left(1 - \frac{1}{Q_1^n} - \mu_n\right)(\gamma/2). \tag{8.4}$$

The lemma follows. ∎

8.2.3. Elements of the centralizer. In this subsection we prove the theorem linking central values to elements of the centralizer of $\mathbb{K}^c$.

---

[33] Axiom (T5b) takes care of the case where the relevant overlaps is odd.

Theorem 90. *Suppose that* $(\mathbb{K}^c, \mathcal{B}, \nu, \mathrm{sh})$ *is a circular system built from a circular construction sequence satisfying the timing assumptions. Let* $\phi : \mathbb{K}^c \to \mathbb{K}^c$ *be an automorphism of* $(\mathbb{K}^c, \mathcal{B}, \nu, \mathrm{sh})$. *Then* $\phi^\pi = \mathcal{S}_\beta$ *for some central value* β.

Proof. Fix a ϕ and suppose that $\phi^\pi = \mathcal{S}_\beta$. We must show that β is central. Suppose not. The idea of the proof is to choose a stationary code $\overline{\Lambda^*}$ well approximating ϕ and an N such for all $M > N$, passing over the principal M-block of most $s \in \mathbb{K}^c$ with $\overline{\Lambda^*}$ gives a string very close to $\phi(s)$ in $\bar{d}$-distance. Consider an s where $\overline{\Lambda^*}$ codes well on this principal M-block.

Use Lemma 83 to build a red zone corresponding to M. Lemma 89 implies that $\overline{\Lambda^*}$ cannot code well on the red zone. Since the red zone takes up the vast majority of the principal M-block, $\overline{\Lambda^*}$ cannot code well on the principal M-block, yielding a contradiction. In more detail:

Let γ be as in axiom (T4). By Proposition 20 there is an code Λ^* such that for almost all $s \in \mathbb{K}^c$,

$$\bar{d}(\bar{\Lambda}^*(s), \phi(s)) < 10^{-9}\gamma.$$

By the Ergodic Theorem there is an N_0 so large that for a set $E \subseteq \mathbb{K}^c$ of measure $7/8$ for all $s \in E$ and all $N > N_0$, s is mature for N and if B is the principal N-block of s, then

$$\bar{d}(\overline{\Lambda^*}(s \upharpoonright B), \phi(s) \upharpoonright B) < 10^{-9}\gamma. \tag{8.5}$$

Let $s \in E$. Choose an $N > N_0$ such that the code length of Λ^* is much smaller than q_N, $1/Q_1^N < 10^{-9}$ and $l_N > 10^{12}$. Apply Lemma 83, with $\delta = 10^{-9}$ to find an integer M and $\langle R_i : i < i^* \rangle$ satisfying the conclusions of Lemma 83. Since $\bigcup_{i<i^*} R_i \subseteq q_M$, we view $\bigcup_{i<i^*} R_i$ as a subset of the principal M-block of s.

Each R_i is a union of collections of locations of the form

$$\bigcup_{j<j_0} U_j,$$

with each U_j consisting of the locations of $u_j^{s_0^{n_i}}$ for $j \in [0, j_0)$ (for some j_0).[34] Moreover, there is a t such that each power

$$u_j^{s_0^{n_i}}$$

is β-matched with a $v_{j+t}^{s_0^{n_i}}$ in $\phi(s)$ for some $t \neq 0$.

Because $j_0 > \varepsilon_n k_n$, axiom (T6) applies and thus for at least

$$\left(1 - \frac{|G_1^{n_i}|}{Q_1^{n_i}} + \mu_{n_i}\right)$$

proportion of $\{u_0, u_1, \ldots, u_{j_0-1}\}$, u_j and v_{j+t} are in different $G_1^{n_i}$-orbits. In Lemma 89, inequality (8.2) implies that if u_i and v_{i+t} are in different $G_1^{n_i}$ orbits, then, restricted to the overlaps of the locations of all of the

$$u_j^{s_0^{n_i}} \quad \text{and} \quad v_{j+t}^{s_0^{n_i}},$$

[34]Note that $s_0^{n_i}$ is as in condition 4 (c) of Lemma 82.

the $\bar{d}$ distance between $\bar{\Lambda}^*(s) \upharpoonright U_j$ and $\phi(s) \upharpoonright U_j$ is at least

$$\frac{1}{50}\left(1 - \frac{1}{Q_1^{n_i}}\right)\gamma.$$

Since the first and last powers of u_j in $u_j^{s_0^{n_i}}$'s take up $2/s_0^{n_i}$ of $u_j^{s_0^{n_i}}$ and $s_0^{n_i} \geq l_{n_i}/2 - 2$, we know that

$$\bar{d}(\overline{\Lambda^*}(s) \upharpoonright U_j, \phi(s) \upharpoonright U_j) \geq (1 - 10^{-11})\frac{1}{50}\left(1 - \frac{1}{Q_1^{n_i}}\right)\gamma.$$

Because the proportion of indices j for which u_j and v_{j+t} are in different $G_1^{n_i}$-orbits is at least

$$1 - \frac{|G_1^{n_i}|}{Q_1^{n_i}} + \mu_{n_i},$$

it follows that

$$\bar{d}\left(\bar{\Lambda}^*(s) \upharpoonright \bigcup_{j<j_0} U_j, \phi(s) \upharpoonright \bigcup_{j<j_0} U_j\right)$$

is at least

$$\left(1 - \frac{|G_1^{n_i}|}{Q_1^{n_i}} + \mu_{n_i}\right)(1 - 10^{-11})\frac{1}{500}\left(1 - \frac{1}{Q_1^{n_i}}\right)\gamma.$$

This in turn is at least $\gamma/1000$. Since R_i is a union of sets of the form $\bigcup_{j<j_0} U_j$, we have

$$\bar{d}(\overline{\Lambda^*}(s) \upharpoonright R_i, \phi(s) \upharpoonright R_i) \geq \frac{\gamma}{1000}.$$

Since $\bigcup_{i<i^*} R_i$ has density at least $1 - 10^{-9}$ if B is the principal M-block of s, it follows that

$$\bar{d}(\overline{\Lambda^*}(s \upharpoonright B), \phi(s) \upharpoonright B) > \frac{\gamma}{10^4}.$$

However, this contradicts inequality (8.5). ∎

Corollary 91. *Let $\mathbb{K}^c$ be a circular system built from a circular construction sequence satisfying the timing assumptions. Then β is a central value if and only if there is an element $\phi \in \overline{\{T^n : n \in \mathbb{N}\}}$ with $\phi^\pi = \mathcal{S}_\beta$. It follows that for each construction sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$ satisfying the numerical requirements collected in Section 11, the central values form a subgroup of the unit circle.*

Proof. Theorem 84 says that if β is central, there is a $\phi \in \overline{\{T^n : n \in \mathbb{N}\}}$ with $\phi^\pi = \mathcal{S}_\beta$. Theorem 90 is the converse. To see the last statement, we prove in Section 10 that for every coefficient sequence satisfying the numerical requirements, we can find a circular construction sequence satisfying the timing assumptions. ∎

8.2.4. Isomorphisms between $\mathbb{K}^c$ and $(\mathbb{K}^c)^{-1}$. We now prove a theorem closely related to Theorem 90.

Theorem 92. *Suppose that* $(\mathbb{K}^c, \mathcal{B}, \nu, \mathrm{sh})$ *is a circular system built from a circular construction sequence satisfying the timing assumptions. Suppose that*

$$\phi : (\mathbb{K}^c, \mathcal{B}, \nu, \mathrm{sh}) \to ((\mathbb{K}^c)^{-1}, \mathcal{B}, \nu, \mathrm{sh})$$

is an isomorphism. Then $\phi^\pi = \natural \circ \mathcal{S}_\beta$ *for some central value* β.

Proof. We concentrate here on the differences with the proof of Theorem 90. The general outline is the same: Fix a ϕ. Then there is a unique β such that $\phi^\pi = \natural \circ \mathcal{S}_\beta$. Suppose that β is not central. Choose a stationary code $\overline{\Lambda^*}$ that well approximates ϕ in terms of $\bar{d}$ distance (say within $\gamma/10^{10}$), and derive a contradiction by choosing a large M and getting lower bounds for $\bar{d}$ distance along the principal M-block of a generic s.

This is done by first comparing a typical s with $\mathcal{S}_\beta(s)$. As in Theorem 90, a definite fraction of a large principal M-block of s is misaligned with $\mathcal{S}_\beta(s)$. But most of the n-blocks of $\mathcal{S}_\beta(s)$ are aligned with reversed n-blocks of $\natural(\mathcal{S}_\beta(s))$ that have been shifted by a very small amount. This can be quantified by looking at the codes $\bar{\Lambda}_n$ for large n, which agree with $\natural$ on the M-block of $\mathcal{S}_\beta(s^\pi)$.

Here are more details. Recall $\natural$ is the limit of a particular sequence of stationary codes $\langle \bar{\Lambda}_n : n \in \mathbb{N}\rangle$. The proof of Theorem 60 showed that for almost all $s^\pi \in \mathcal{K}$ for all large enough n the principal n-blocks of $\bar{\Lambda}_n(s^\pi)$ and $\bar{\Lambda}_{n+1}(s^\pi)$ agree. Fix a generic $s \in \mathbb{K}^c$ and a large N such that:

(1) the code $\overline{\Lambda^*}$ codes ϕ well on the principal n-block of s for all $n \geq N$,

(2) for all $n \geq N$ the principal n-blocks of $\bar{\Lambda}_n(\mathcal{S}_\beta(\pi(s)))$ and $\bar{\Lambda}_{n+1}(\mathcal{S}_\beta(\pi(s)))$ agree,

(3) s is mature at N,

(4) the length of Λ is very small relative to N, and

(5) l_N is very large.

Comparing $\pi(s)$ and $\mathcal{S}_\beta(\pi(s))$, Lemma 83 gives us an $M > N$ and a red zone in the principal M-block s. We assume that the red zones take up at least $1 - 10^{-9}$ proportion of the principal M-block and have the form given in Lemma 83.

We will derive a contradiction by showing that $\overline{\Lambda^*}$ cannot code well. This is done by considering the blocks of $\phi(s)$ that are lined up with the red zones of the principal M-block of s and using Lemma 89 to see that $\overline{\Lambda^*}$ cannot code well on these sections. This is possible because the mismatched n-blocks of $\mathcal{S}_\beta(\pi(s))$ are lined up closely with the n-blocks of $\natural(\mathcal{S}_\beta(\pi(s))) = \phi^\pi(s)$. Explicitly: Use Lemma 83 to choose red zones $\langle R_i : i < i^*\rangle$ that take up a $1 - 10^{-9}$ proportion of the principal M-block of s.[35]

The boundary portions of n-words with $n < M + 1$ take up at most $2/l_M$ proportion of the overlap of the principal M-blocks of s and $\phi(s)$. Since this proportion is so small, Remark 22 allows us to completely ignore blocks corresponding to n_i-words in s that are lined up with boundary in $\phi(s)$ and vice versa.

We now examine the how $\natural(\mathcal{S}_\beta(\pi(s)))$ compares with $\mathcal{S}_\beta(\pi(s))$. Temporarily denote $\mathcal{S}_\beta(\pi(s))$ by s'. By the choice of s, for all $n \in [N, M]$ the alignments of the principal n-blocks of $\bar{\Lambda}_n(s')$ and $\bar{\Lambda}_M(s')$ agree.

[35] We use the notation in Lemma 83 and Theorem 60.

The red zones of s^π line up blocks of the form $u_j^{s_0^{n_i}}$ with blocks of the form $v_{j+t}^{s_0^{n_i}}$ occurring in s' that are shifted by $d^{n_i}(s)$ (so $t \neq 0$). Except for those blocks that line up with boundary portions of $\natural(s')$ these blocks are lined up with blocks of the form $\mathrm{sh}^{A_{n_i}}(\mathrm{rev}(v_{k_{n_i}-(j+t)-1}))$ in $\natural(s')$.[36] Inequality (5.17), says that $A_{n_i} < 2q_{n_i-1}$. In particular, the blocks of powers of v_{j+t} are lined up with a very small shift of $\mathrm{rev}(v_{k_{n_i}-(j+t)-1})$ in $\natural(s')$.

Thus vast majority of blocks U_j that are positions of $u_j^{s_0^{n_i}}$ in s^π are lined up with a shift by less than q_{n_i} of a block of $\natural(s^\pi)$ in a position of

$$v_{k_{n_i}-(j+t)-1}^{s_0^{n_i}}$$

in $\natural(s')$. Consider s and $\phi(s)$. Suppose that u_j are the n_i-words of s corresponding to the U_j and $v_{k_{n_i}-(j+t)-1}$ are the n_i-words of $\phi(s)$ across from them. By axiom (T5a), at most

$$\frac{1}{Q_1^{n_i}} + \mu_{n_i}$$

of the $j < j_0$ happen to have $[u_j]_1 \in G_1^{n_i}[v_{k_{n_i}-(j+t)-1}]_1$. At least

$$1 - \frac{1}{Q_1^{n_i}} + \mu_{n_i}$$

proportion of the powers of u_j the $\bar{d}$-distance between $\overline{\Lambda^*}$ and ϕ is at least

$$\frac{1}{500}\left(1 - \frac{1}{Q_1^n}\right)\gamma.$$

It follows that on R_i the $\bar{d}$-distance is at least $\gamma/1000$. If we choose $\bigcup_{i<i^*} R_i$ to have density at least $1 - 10^{-9}$ and let B be the principal M-block of s, then (as in Theorem 90)

$$\bar{d}(\overline{\Lambda^*}(s \restriction B), \phi(s) \restriction B) > \gamma/10^4,$$

a contradiction. ∎

8.3. Synchronous and anti-synchronous isomorphisms

View a circular system $(\mathbb{K}^c, \mathcal{B}, \nu, \mathrm{sh})$ as an element T of the space **MPT** endowed with the weak topology.

Theorem 93. *Suppose that $\mathbb{K}^c$ is a circular system satisfying the timing assumptions. Then:*

(1) *If there is an isomorphism $\phi : \mathbb{K}^c \to \mathbb{K}^c$ such that $\phi \notin \overline{\{T^n : n \in \mathbb{Z}\}}$, there is an isomorphism $\psi : \mathbb{K}^c \to \mathbb{K}^c$ such that $\psi \notin \overline{\{T^n : n \in \mathbb{Z}\}}$ and ψ^π is the identity map.*

(2) *If there exists an isomorphism $\phi : \mathbb{K}^c \to (\mathbb{K}^c)^{-1}$, then there exists an isomorphism $\psi : \mathbb{K}^c \to (\mathbb{K}^c)^{-1}$ such that $\psi^\pi = \natural$.*

[36] See the qualitative discussion of $\natural$ that occurs after its definition in [12].

Proof. To see assertion (1), let $\phi : \mathbb{K}^c \to \mathbb{K}^c$ be an isomorphism with $\phi \notin \overline{\{T^n : n \in \mathbb{Z}\}}$. Then by Theorem 90, $\phi^\pi = \mathcal{S}_\beta$ for a central β. Corollary 91 implies that there exists a $\theta \in \overline{\{T^n : n \in \mathbb{N}\}}$ such that $\theta^\pi = \mathcal{S}_{-\beta}$. Then $\phi \circ \theta : \mathbb{K}^c \to \mathbb{K}^c$ is an isomorphism such that $(\phi \circ \theta)^\pi$ is the identity map. Since $\overline{\{T^n : n \in \mathbb{N}\}}$ is a group, $\phi \circ \theta \notin \overline{\{T^n : n \in \mathbb{N}\}}$.

The proof of assertion (2) is very similar. Suppose that $\phi : \mathbb{K}^c \to (\mathbb{K}^c)^{-1}$ is an isomorphism. Then, by Theorem 92, $\phi^\pi = \natural \circ \mathcal{S}_\beta$ for a central β. Let $\theta \in \overline{\{T^n : n \in \mathbb{N}\}}$ be such that $\theta^\pi = \mathcal{S}_{-\beta}$. Then $\phi \circ \theta$ is an isomorphism between $\mathbb{K}^c$ and $(\mathbb{K}^c)^{-1}$ with $(\phi \circ \theta)^\pi = \natural$. ∎

9. The proof of the main theorem

In this section we prove the main theorem of this paper, Theorem 2. By Fact 24, it suffices to prove the following:

Theorem 94. *There is a continuous function $F^s : \mathcal{T}rees \to \mathrm{Diff}^\infty(\mathbb{T}^2, \lambda)$ such that for $\mathcal{T} \in \mathcal{T}rees$, if $T = F^s(\mathcal{T})$,*

(1) *$\mathcal{T}$ has an infinite branch if and only if $T \cong T^{-1}$,*

(2) *$\mathcal{T}$ has two distinct infinite branches if and only if*

$$C(T) \neq \overline{\{T^n : n \in \mathbb{Z}\}}.$$

We split the proof of this theorem into three parts. In the first we assume the timing assumptions hold, define F^s and show that it is a reduction. In the second part we show that F^s is continuous.

The third part of the proof augments the specifications of [8] with two additional randomness properties, shows that the additional properties imply the timing assumptions and describes how to perform the word construction from [8] with these additional requirements. We present the third part of the proof separately in Section 10.

We begin by defining F^s. The main result of [8] relied on the construction of a continuous function $F : \mathcal{T}rees \to$ **MPT** such that for all $\mathcal{T} \in \mathcal{T}rees$, if $S = F(\mathcal{T})$, then:

Fact 1. The tree $\mathcal{T}$ has an infinite branch if and only if $S \cong S^{-1}$.

Fact 2. The tree $\mathcal{T}$ has two distinct infinite branches if and only if

$$C(S) \neq \{S^n : n \in \mathbb{Z}\}.$$

Fact 3. The function F took values in the strongly uniform odometer based transformations and for S in the range of F, $S \cong S^{-1}$ if and only if there is an anti-synchronous isomorphism ϕ between S and S^{-1}.

Fact 4 ([8, Corollary 40, p. 1565]). If S is in the range of F and $C(S) \neq \{S^n : n \in \mathbb{Z}\}$, then there is a synchronous $\phi \in C(S)$ such that for some n, non-identity element $g \in G_1^n$ and all generic $s \in \mathbb{K}$ and all large enough m, if u and v are the principal m-subwords of s and $\phi(s)$ respectively, then

$$[v]_1 = g[u]_1.$$

Fact 5 ([8, Equations 1 and 2 on p. 1546 and p. 1547]). For all n_0 there is an M such that if $\mathcal{T}$ and $\mathcal{T}'$ are trees and[37]

$$\mathcal{T} \cap \{\sigma_n : n \le M\} = \mathcal{T}' \cap \{\sigma_n : n \le M\},$$

then the first n_0-steps of the construction sequences for $F(\mathcal{T})$ are equal to the first n_0-steps of the construction sequence for $F(\mathcal{T}')$; i.e. $\langle \mathcal{W}_k(\mathcal{T}) : k < n_0 \rangle = \langle \mathcal{W}_k(\mathcal{T}') : k < n_0 \rangle$.

Fact 6. The construction sequence for $F(\mathcal{T})$ satisfies the *specifications* given in [8]. In Section 10.2, these specifications are augmented by the addition of (J10.1) and (J11.1). In Section 10.3 we argue that if $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ is a construction sequence for an odometer based system that satisfies the augmented specifications, then the associated circular construction sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$ satisfies the timing assumptions.

Moreover, the construction sequence for $F(\mathcal{T})$ is strongly uniform and hence the construction sequence for $\mathcal{F} \circ F(\mathcal{T})$ is strongly uniform.

Fact 7. Construction sequences satisfying the augmented specifications are easily built using the techniques of [8] with no essential changes; consequently we can assume that the construction sequences for $F(\mathcal{T})$ satisfy the augmented specifications.

In [11, Theorem 60] it is shown that if $\langle W_n^c : n \in \mathbb{N} \rangle$ is a strongly uniform circular construction sequence with coefficients $\langle k_n, l_n : n \in \mathbb{N} \rangle$, where $\langle l_n : n \in \mathbb{N} \rangle$ grows fast enough and $|\mathcal{W}_n^c|$ goes to infinity then there is a smooth measure preserving diffeomorphism $T \in \text{Diff}^\infty(\mathbb{T}^2, \lambda)$ measure theoretically isomorphic to $\mathbb{K}^c$. This gives a map R from circular systems with fast growing coefficients to $\text{Diff}^\infty(\mathbb{T}^2, \lambda)$.

If $\mathcal{F}$ is the canonical functor from odometer systems to circular systems, we define

$$F^s = R \circ \mathcal{F} \circ F$$

(see Figure 3).

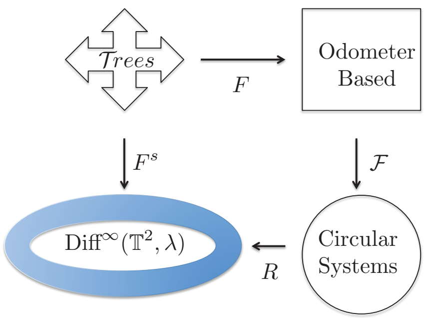

Fig. 3. The definition of F^s.

[37] See Section 4.6 for notation.

9.1. F^s is a reduction

Because R preserves isomorphism, to show that $F^s = R \circ \mathcal{F} \circ F$ is a reduction, it is suffices to show that $\mathcal{F} \circ F$ is a reduction. Let S be the transformation corresponding to the system $\mathbb{K} = F(\mathcal{T})$ and T the transformation corresponding to $\mathbb{K}^c = \mathcal{F} \circ F(\mathcal{T})$.

Item (1) *of Theorem* 94. Suppose that $\mathcal{T}$ is a tree and $\mathcal{T}$ has an infinite branch. By Facts 1 and 3, there is an anti-synchronous isomorphism $\phi : \mathbb{K} \to \mathbb{K}^{-1}$. By [12, Theorem 105], if $\mathbb{K}^c = \mathcal{F}(\mathbb{K})$, there is an isomorphism $\phi^c : \mathbb{K}^c \to (\mathbb{K}^c)^{-1}$.

Now suppose that $F^s(\mathcal{T}) \cong (F^s(\mathcal{T}))^{-1}$. Then we have $\mathbb{K}^c \cong (\mathbb{K}^c)^{-1}$. By Fact 6, the construction sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$ for $\mathcal{F}^s(\mathcal{T})$ satisfies the timing assumptions. By Theorem 93, there is an anti-synchronous isomorphism $\phi^c : \mathbb{K}^c \to (\mathbb{K}^c)^{-1}$. Again by [12, Theorem 105], there is an isomorphism between $\mathbb{K}$ and $\mathbb{K}^{-1}$. By [8], $\mathcal{T}$ has an infinite branch.

Item (2) *of Theorem* 94. Suppose that $\mathcal{T}$ has at least two infinite branches. Then the centralizer of $S = F(\mathcal{T})$ is not equal to the powers of S. By Fact 4, we can find a synchronous $\phi \in \overline{C(S) \setminus \{S^n : n \in \mathbb{Z}\}}$. Let $\psi = \mathcal{F}(\phi)$; then ψ is synchronous. We claim that $\psi \notin \overline{\{T^n : n \in \mathbb{Z}\}}$. Using Fact 4, and lifting the group action of G_1^n and the equivalence relation $\mathcal{Q}_1^n$, we see that for all generic $s^c \in \mathbb{K}^c$, and all large enough m, if u^c and v^c are the principal m-subwords of s^c and $\psi(s^c)$, respectively, then

$$[v^c]_1 = g[u^c]_1$$

for some $g \neq e$. In particular, $[v^c]_1 \neq [u^c]_1$.

By the timing assumption (T4), there is a $\gamma > 0$ such that for all large m and all shifts A with $|A|$ of size less than $q_m/2$, we have

$$\bar{d}(T^A(u^c), v^c) > \gamma. \tag{9.1}$$

Suppose that $\psi \in \overline{\{T^n : n \in \mathbb{Z}\}}$. Then, by Proposition 21, we can find an $A \in \mathbb{Z}$ and a generic s^c such that

$$\bar{d}(T^A(s^c), \psi(s^c)) < \gamma/2. \tag{9.2}$$

But inequality (9.2) and the Ergodic Theorem imply that for large enough $m \gg A$ if u^c and v^c are the principal m-blocks of s^c and $\psi(s^c)$, then

$$\bar{d}(T^A(u^c), v^c) < \gamma,$$

contradicting inequality (9.1).

Now suppose that there is a $\psi \in C(T)$ such that $\psi \notin \overline{\{T^n : n \in \mathbb{Z}\}}$. Then by Theorem 93, there is such a ψ that is synchronous. In particular, for all n, $\psi \neq T^n$. Thus if S is the transformation corresponding to $F(\mathcal{T})$, then $\mathcal{F}^{-1}(\psi)$ belongs to the centralizer of S and is not a power of S.

9.2. F^s is continuous

Fix a metric d on $\mathrm{Diff}^\infty(\mathbb{T}^2, \lambda)$ yielding the C^∞-topology. For each circular system T, let $\langle P_n^T : n \in \mathbb{N}\rangle$ be the sequence of collections of prewords used to construct T. By

[11, Proposition 61], given $T = F^s(\mathcal{T})$ and a C^∞-neighborhood B of T, there is a large enough M, for all $S \in \mathrm{range}(R)$ if $\langle P_n^S : n \leq M\rangle = \langle P_n^T : n \leq M\rangle$, then $S \in B$. For all odometer based transformations, the sequence $\langle \mathcal{W}_n : n \leq M\rangle$ determines $\langle P_n : n \leq M\rangle$. Hence for all $\mathcal{T}'$, if the first M members of the construction sequence for $F(\mathcal{T}')$ are the same as the first M members of the construction sequence for $F(\mathcal{T})$, then $F(\mathcal{T}') \in U$. By Fact 5, there is a basic open interval $V \subseteq \mathcal{T}rees$ that contains $\mathcal{T}$ and is such that the first M members of the construction sequence are the same for all $\mathcal{T}' \in V$. It follows that for all $\mathcal{T}' \in V$, $F^s(\mathcal{T}') \in U$.

9.3. Numerical requirements arising from smooth realizations

The construction of R depends on various estimates that put lower bounds on the growth of the coefficient sequences. We now list these numerical requirements. The claims in this subsection presuppose a knowledge of [11].

The map R depends on various smoothed versions h_n^s of the permutations h_n of the unit interval arising from $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$. To solve this problem, we fix in advance such approximations, making sure that each approximation h_n^s agrees sufficiently well with h_n as to not disturb the other estimates.

This introduces various numerical constraints on the growth of the coefficients l_n. The diffeomorphism T is built as a limit of periodic approximations T_n. To make the sequence of T_n converge at each stage, l_n must be chosen sufficiently large. Thus the growth rate of l_n depends on $\langle k_m, s_m, h_m : m \leq n\rangle$, $\langle l_m : m < n\rangle$, s_{n+1}, h_{n+1}. Since there are only finitely many possibilities for sequences $\langle h_m : m \leq n\rangle$ corresponding to a given sequence $\langle k_m : m \leq n\rangle$, $\langle s_m : m \leq n+1\rangle$, we can find one growth rate that is sufficiently fast to work for all choices of permutations h_m. This is discussed in detail in [11, p. 34], where the lower bound is called l_n^*.

Numerical Requirement 6. The coefficient l_n is big enough relative to a lower bound determined by $\langle k_m, s_m : m \leq n\rangle$, $\langle l_m : m < n\rangle$ and s_{n+1} to make the periodic approximations to the diffeomorphism converge. Moreover, $k_n \leq l_n$.

Remark 95. Choosing α_{n+1} close to α_n is a fundamental idea of the method of Approximation by Conjugacy, due to Anosov and Katok. By equations (5.5) and (5.6), this is equivalent to taking l_n large. The magnitude of l_n is not calculated, but instead it shown that as l_n increases a sequence of periodic diffeomorphisms well approximates a given periodic diffeomorphism. Then in the original sources [1] and [18], one simply takes l_n *sufficiently large*. This is what Numerical Requirement 6 is repeating.

The argument for the ergodicity of the diffeomorphism formally required that:

Numerical Requirement 7. We have $s_n \to \infty$ as $n \to \infty$, s_{n+1} is a multiple of s_n.

The reader is referred to Example 5 for a discussion of $s(n)$ and its growth.

The next requirement makes it possible to choose s_{n+1} and then, by making k_n sufficiently large, construct s_{n+1} sufficiently random words using elements of $\mathcal{W}_n$.

Numerical Requirement 8. We have $s_{n+1} \leq s_n^{k_n}$.

10. The specifications

In this section we describe how the timing assumptions are related to the specifications given in [8], show that they are compatible and indicate how to construct odometer words so that both sets of assumptions hold. This completes the proof of Theorem 94, subject to the verification that all of the numerical requirements we have introduced are consistent with the numerical requirements of [8]. We take this up in Section 11. We will assume that the reader is familiar with [8, Sections 7 and 8].

10.1. Corresponding specifications

Table 1 links the timing assumptions we use in this paper to the corresponding specification in [8]. (We remind the reader that Appendix A has a table giving corresponding notation between [8] and this paper.)

Timing assumption	Specification
(T1)	Q5
(T2)	Q7
(T3)	A8
(T4)	New
(T5)	J10
(T6)	J10
(T7)	J11

Tab. 1. The specifications in [8] related to the timing assumptions in this paper.

Specification (T4) does not directly correspond to one of the specifications, but (as we will show) holds naturally in the circular words lifted from an odometer construction satisfying the specifications.

Numerical Requirement 9. In the current construction we have two summable sequences: $\langle \epsilon_n : n \in \mathbb{N}\rangle$ and $\langle \varepsilon_n : n \in \mathbb{N}\rangle$. We use the lunate "$\epsilon_n$" notation for the specifications from [8] and the classical "ε_n" notation ("varepsilon" in LaTeX) for the numerical requirements relating to circular systems and their realizations as diffeomorphisms. A requirement for the construction is that

$$\epsilon_n < \varepsilon_n.$$

We also assume that the ϵ_n are decreasing and $\epsilon_0 < 1/40$.

10.2. Augmenting the specifications from [8]

The paper [8] constructs a reduction F from the space of trees to the odometer based systems. The system $\mathbb{K} = F(\mathcal{T})$ was built according to a list of specifications which we reproduce here in order to show how to strengthen them to imply the timing assumptions used in the proofs of Theorems 93 and 94 and to verify that the strengthened assumptions are consistent. The specifications directly relevant to the timing assumptions are (J10) and (J11). The others, which describe the scaffolding for the construction, are only relevant in that they set the stage for the application of the functor $\mathcal{F}$ defined in Section 5.

Here are some definitions from [8] that are used in the specifications. We advise the reader that a table giving the notational changes between [8] and this paper is in Appendix A.

Fix an enumeration of the finite sequences of natural numbers, $\langle \sigma_n : n \in \mathbb{N}\rangle$, with the property that if σ is an initial segment of τ, then σ is enumerated before τ. Let $\mathcal{T}$ be a tree whose elements are $\langle \sigma_{n_i} : i \in \mathbb{N}\rangle$. Here are the specifications for the construction sequence $\mathcal{W} = \mathcal{W}(\mathcal{T})$ used to build $\mathcal{F}(\mathcal{T})$.

There is a sequence of groups G_s^n built as follows. For all n, G_0^n is the trivial group (e) and if we let

$$X_s^n = \{\sigma_{n_i} : i \leq n \text{ and } \sigma_{n_i} \text{ has length } s\},$$

then

$$G_s^n = \sum_{\sigma \in X_s^n} (\mathbb{Z}_2)_\sigma,$$

i.e. G_s^n is a direct sum of copies of $\mathbb{Z}_2$ indexed by elements of X_s^n. There are canonical homomorphisms from G_{s+1}^n to G_s^n that send a generator of G_{s+1}^n corresponding to a sequence of the form $\tau^\frown j$ to the generator of G_s^n corresponding to τ.

The sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$, equivalence relations $\mathcal{Q}_s^n$ and the group actions of G_s^n are constructed inductively. The words in $\mathcal{W}_n$ are sequences of elements of $\Sigma = \{0, 1\}$. To start, $\mathcal{W}_0 = \{0, 1\}$ and $\mathcal{Q}_0^0$ is the trivial equivalence relation with one class. The collection of words $\mathcal{W}_n$ is built when the n-th element of $\mathcal{T}$ is considered. We will say that words in $\mathcal{W}_n$ have *even* parity and words in $\text{rev}(\mathcal{W}_n)$ have *odd* parity.

We begin by restating the specifications from [8] using the indexing conventions in this paper ($n \mapsto n+1$ vs $m \mapsto n$). (E1)–(A9) are exactly the same, however we modify the joining specifications (J10), (J11) slightly for the needs of this paper.

(E1) Any pair w_1, w_2 of words in $\mathcal{W}_n$ have the same length.

(E2) Every word in $\mathcal{W}_{n+1}$ is built by concatenating words in $\mathcal{W}_n$. Every word in $\mathcal{W}_n$ occurs in each word of $\mathcal{W}_{n+1}$ exactly p_n^2 times, where p_n is a large prime number chosen when the n-th element of $\mathcal{T}$ is considered.

(E3) (Unique readability) If $w \in \mathcal{W}_{n+1}$ and

$$w = p w_1 \dots w_k s$$

where each $w_i \in \mathcal{W}_n$ and p or s are sequences of 0's and 1's that have length less than that of any word in $\mathcal{W}_n$, then both p and e are the empty word. If $w, w' \in \mathcal{W}_{n+1}$ are such that $w = w_1 w_2 \dots w_{k_n}$ and $w' = w'_1 w'_2 \dots w'_{k_n}$ with $w_i, w'_i \in \mathcal{W}_n$, and $k = [k_n/2] + 1$, then we have $w_k w_{k+1} \dots w_{k_n} \neq w'_1 w'_2 \dots w'_{k_n - [k] - 1}$, i.e. the first half of w' is not equal to the second half of w.

Let $s(n)$ be the length of the longest sequence among the first n sequences in $\mathcal{T}$ and if $\mathcal{T} = \langle \sigma_{n_i} : i \in \mathbb{N}\rangle$, then $M(s)$ is the least i such that σ_{n_i} has length s.

The equivalence relations $\mathcal{Q}_s^n$ on $\mathcal{W}_n$ are defined for all $s \le s(n)$. The equivalence relation $\mathcal{Q}_0^0$ on $\mathcal{W}_0$ is the trivial equivalence relation with one class.

(Q4) Suppose that $n = M(s)$. Then any two words in the same $\mathcal{Q}_s^n$ equivalence class agree on an initial segment of proportion least $(1 - \epsilon_n)$.

(Q5) For $n \ge M(s) + 1$, $\mathcal{Q}_s^n$ is the product equivalence relation of $\mathcal{Q}_s^{M(s)}$. Hence we can view $\mathcal{W}_n/\mathcal{Q}_s^n$ as sequences of elements of $\mathcal{W}_{M(s)}/\mathcal{Q}_s^{M(s)}$ and similarly for $\mathrm{rev}(\mathcal{W}_n)/\mathcal{Q}_s^n$.

(Q6) $\mathcal{Q}_{s+1}^n$ refines $\mathcal{Q}_s^n$ and each $\mathcal{Q}_s^n$ class contains $2^{e(n)}$ many $\mathcal{Q}_{s+1}^n$ classes, where e is a strictly increasing function. The speed of growth of e is discussed in Section 11.

(A7) G_s^n acts freely on $\mathcal{W}_n/\mathcal{Q}_s^n \cup \mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$ and the G_s^n action is subordinate to the G_{s-1}^n action via the natural homomorphism $\rho_{s,s-1}$ from G_s^n to G_{s-1}^n.

(A8) The canonical generators of $G_s^{M(s)}$ send elements of $\mathcal{W}_{M(s)}/\mathcal{Q}_s^{M(s)}$ to elements of $\mathrm{rev}(\mathcal{W}_{M(s)})/\mathcal{Q}_s^{M(s)}$ and vice versa.

(A9) If $M(s) \le n$ and we view
$$G_s^{n+1} = G_s^n \oplus H,$$
the action of the group G_s^n on $\mathcal{W}_n/\mathcal{Q}_s^n \cup \mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$ is extended to an action on $\mathcal{W}_{n+1}/\mathcal{Q}_s^{n+1} \cup \mathrm{rev}(\mathcal{W}_{n+1}/\mathcal{Q}_s^{n+1})$ by the skew diagonal action. If H is non-trivial, then $H = \mathbb{Z}_2$ and its canonical generator maps $\mathcal{W}_{n+1}/\mathcal{Q}_s^{n+1}$ to $\mathrm{rev}(\mathcal{W}_{n+1}/\mathcal{Q}_s^{n+1})$.

Note. While it is not explicitly given as a specification in [8], the construction sequence has the property that if $g \in G_s^n$ is a canonical generator, then for $m > n$, $\mathcal{W}_m/\mathcal{Q}_s^m$ is closed under the skew diagonal action of g.

Suppose that u and v are elements of $\mathcal{W}_{n+1} \cup \mathrm{rev}(\mathcal{W}_{n+1})$ and (u', v') an ordered pair from $\mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$. Suppose that u and v are in positions shifted relative to each other by t units. Then an *occurrence* of (u', v') in $(\mathrm{sh}^t(u), v)$ is a t' such that u' occurs in u starting at $t + t'$ and in v starting at t'. Let Q_s^n be the number of classes of $\mathcal{Q}_s^n$ and let C_s^n be the number of elements of each Q_s^n class.[38]

To prove the timing assumptions, we need to strengthen specifications (J10) and (J11) to deal with $\bar{d}$-distance on initial and tail segments and on words that are shifted. The spirit of specification (J10) is that pairs of n-words (u', v') occur randomly in the overlap of u and v when u is shifted by a suitable multiple t of the lengths of n-words. Specification (J10.1) says the same thing relative to non-trivial initial segments of the overlap of the shift of u and v.

Specification (J11) says that if $[u]_s$ is in the G_s^n-orbit of $[v]_s$ and s is maximal with this property, then the occurrences of (u', v') are approximately conditionally random. More explicitly, suppose that $g[u]_s = [v]_s$, and we are given $u' \in \mathcal{W}_n$. Then there are Q_s^n many pairs of $\mathcal{Q}_s^n$-classes $([u^*]_s, [v^*]_s)$ with $g[u^*]_s = [v^*]_s$, and so $([u']_s, [v']_s)$ should occur randomly $1/Q_s^n$ proportion of the time. There are C_s^n many elements of $\mathcal{W}_n$ in

[38] We have changed the variables used in the statement of J10 in [8] to conform to the notation described in Appendix A.

the $\mathcal{Q}^n_s$-classes, and conditional on $g[u']_n = [v']_n$, the chances of such a pair (u', v') randomly matching is $1/(C^n_s)^2$. Specification (J11.1) strengthens this (but only for $\mathcal{Q}^n_0$, which is the trivial equivalence relation and $G^n_0 = \langle e \rangle$) by asking that this holds over any non-trivial interval of length $j_0 K_n$ at the beginning or end of an $n+1$-word.

Here are the joining specifications as given in [8]:

(J10) Let u and v be elements of $\mathcal{W}_{n+1} \cup \mathrm{rev}(\mathcal{W}_{n+1})$. Let $1 \leq t < (1-\epsilon_n)(k_n)$ be an integer. Then for each pair $u', v' \in \mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$ such that u' has the same parity as u and v' has the same parity as v, let $r(u', v')$ be the number of occurrences of (u', v') in $(\mathrm{sh}^{tK_n}(u), v)$ on their overlap. Then

$$\left| \frac{r(u', v')}{k_n - t} - \frac{1}{s_n^2} \right| < \epsilon_n.$$

(J11) Suppose that $u \in \mathcal{W}_{n+1}$ and $v \in \mathcal{W}_{n+1} \cup \mathrm{rev}(\mathcal{W}_{n+1})$. We let $s = s(u, v)$ be the maximal i such that there is a $g \in G^n_i$ such that $g[u]_i = [v]_i$. Let $g = g(u, v)$ be the unique g with this property and $(u', v') \in \mathcal{W}_n \times (\mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n))$ be such that $g[u']_s = [v']_s$. Let $r(u', v')$ be the number of occurrences of (u', v') in (u, v). Then

$$\left| \frac{r(u', v')}{k_n} - \frac{1}{Q^n_s}\left(\frac{1}{C^n_s}\right)^2 \right| < \epsilon_n.$$

The strengthening of (J10) is:

(J10.1) Let u and v be elements of $\mathcal{W}_{n+1} \cup \mathrm{rev}(\mathcal{W}_{n+1})$. Let $1 \leq t < (1-\epsilon_n)(k_n)$. Let j_0 be a number between $\epsilon_n k_n$ and $k_n - t$. Then for each pair $u', v' \in \mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$ such that u' has the same parity as u and v' has the same parity as v, let $r(u', v')$ be the number of $j < j_0$ such that (u', v') occurs in $(\mathrm{sh}^{tK_n}(u), v)$ in the (jK_n)-th position in their overlap. Then

$$\left| \frac{r(u', v')}{j_0} - \frac{1}{s_n^2} \right| < \epsilon_n.$$

The next assumption is a strengthening of a special case of (J11).

(J11.1) Suppose that $u \in \mathcal{W}_{n+1}$ and $v \in \mathcal{W}_{n+1} \cup \mathrm{rev}(\mathcal{W}_{n+1})$ and $[u]_1 \notin G^n_1[v]_1$.[39] Let j_0 be a number between $\epsilon_n k_n$ and k_n. Suppose that I is either an initial or a tail segment of the interval $\{0, 1, \ldots, K_{n+1} - 1\}$ having length $j_0 K_n$. Then for each pair $u', v' \in \mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$ such that u' has the same parity as u and v' has the same parity as v, let $r(u', v')$ be the number of occurrences of (u', v') in $(u \upharpoonright I, v \upharpoonright I)$. Then

$$\left| \frac{r(u', v')}{j_0} - \frac{1}{s_n^2} \right| < \epsilon_n.$$

We have augmented the specifications in [8] with (J10.1) and (J11.1). Formally, we must argue that it is possible to build construction sequences satisfying the additional specifications. This means constructing s_{n+1} many pseudo-random words. This is done using

[39] In the language of (J11): $s(u, v) = 0$, $Q^n_0 = 1$ and $C^n_0 = s_n$.

a routine modification of the techniques of [8], where the collections of words $\mathcal{W}_n$ are built probabilistically. For $n \geq 1$ the words in $\mathcal{W}_{n+1}$ are built by iteratively substituting words into $K_{n+1}/K_{M(i)}$-sequences of classes $\mathcal{Q}_i^n$, by induction on $i \leq i^*$, where i^* is maximal with $M(i^*) \leq n$. The classes of words $\mathcal{W}_{n+1}/\mathcal{Q}_{n+1}^i$ are built by induction on i. A word $w \in \mathcal{W}_{n+1}/\mathcal{Q}_{i+1}^{n+1}$ (or in $\mathcal{W}_{n+1}$ if $i = i^*$) can be viewed as a result of substituting elements of $\mathcal{W}_n/\mathcal{Q}_{i+1}^n$ (or $\mathcal{W}_n$) into a word in $\mathcal{W}_{n+1}/\mathcal{Q}_i^{n+1}$.

Suppose that $[w]_i \in \mathcal{W}_{n+1}/\mathcal{Q}_i^{n+1}$ has been built and is given by $K_{n+1}/K_{M(i)}$ many consecutive classes $C_1C_2\dots C_{K_{n+1}/K_{M(i)}}$. Then $[w]_{i+1} \in \prod_{j<K_{n+1}/K_{M(i)}} C_j$. Viewing these as independent trials and taking k_n large enough (so that $K_{n+1}/K_{M(i)}$ is very large) the finitary *Law of Large Numbers* shows that the vast majority of choices of $2^{e(n)}$ words satisfy (J10), (J10.1), (J11) and (J11.1):

Remark 96. As noted in Example 5, given the number of substitutions to be made (which is one more than the maximal s such that $\mathcal{Q}_s^n$ is defined) and the size of the groups G_s^n one can give an explicit formula relating the sizes of $e(n+1)$ and s_{n+1}. Given one of the two, one can solve for the other. Moreover, when one goes up the other does as well. This co-determination means that the requirements can be stated in terms of either variable. We state the requirements in terms of the s_n.

In the construction, getting the additional .1 for (J10) and (J11) only involves taking k_n larger than was necessary in [8]. This is described in this notation in [7].

This leads to a numerical requirement:

Numerical Requirement 10. The number k_n is chosen sufficiently large relative to a lower bound determined by s_{n+1} for the Law of Large Numbers arguments to work.

10.3. Verifying the timing assumptions

In this subsection we prove that the augmented specifications (E1)–(J11.1) imply the timing assumptions, introduced in Section 8.2.1. The first three timing assumptions, i.e. (T1)–(T3), follow easily from the results in Section 5.10 together with specifications (Q5), (A7) and (A8).

The following remark is easy and illustrates the idea behind the demonstrations of (T4)–(T7).

Remark 97. Suppose that $\mathcal{L}$ is an alphabet with s symbols in it and $\mathcal{C} \subset \mathcal{L}$ with $|\mathcal{C}| = C$. For u, v words in $\mathcal{L}$ of the same length and $x, y \in \mathcal{L}$, set $r(x, y)$ to be the number of occurrences of (x, y) in (u, v), $r(x, \mathcal{C})$ to be the number of occurrences of some element of $\mathcal{C}$ opposite an occurrence of x in u and $f(x)$ to be the number of occurrences of x in u. Then for all $\mu > 0$ there is an $\epsilon = \epsilon(\mu, s)$ such that whenever u, v are two words in $\mathcal{L}$ of the same length ℓ, if for all $x, y \in \mathcal{L}$,

$$\left|\frac{r(x,y)}{\ell} - \frac{1}{s^2}\right| < \epsilon,$$

then for all x,

$$\left|\frac{r(x,\mathcal{C})}{f(x)} - \frac{C}{s}\right| < \mu.$$

Proof. Because $f(x) = \sum_y r(x, y)$, by taking ϵ sufficiently small we can arrange that

$$\frac{f(x)}{\ell} \approx \frac{1}{s},$$

and the approximation improves as ϵ gets smaller. Simplemindedly,

$$\frac{r(x,y)}{f(x)} = \frac{r(x,y)}{\ell}\frac{\ell}{f(x)} \approx \frac{1}{s^2}s \approx \frac{1}{s}.$$

Since $r(x, \mathcal{C}) = \sum_{y\in\mathcal{C}} r(x, y)$, we see that

$$\frac{r(x,\mathcal{C})}{f(x)} \approx \frac{C}{s}.$$

As we take ϵ smaller, the final approximation improves. ■

We now establish the timing assumptions (T4)–(T7). Recall that in the context of the timing assumptions the notation $a \approx b$ means that $|a - b| < \mu_n$.

Assumption (T5). Assume that specification (J10) holds for sufficiently small ϵ_n. To use Remark 97 to see (T5), take $\mathcal{L} = \mathcal{W}_n$, the number $f(x)$ to be $|J(v)|$ and C to be the cardinality of any equivalence class of $\mathcal{Q}_1^n$ and $s = s_n$. Since each class of $\mathcal{Q}_1^n$ has the same number of elements, $\frac{s}{C}$ is equal to the number of classes: $\frac{s}{C} = Q_1^n$. Thus $\frac{C}{s} = \frac{1}{Q_1^n}$ and (T5) follows.

Assumption (T6). We can write the set S as

$$S = \bigcup_{v\in\mathcal{W}_n^c}\bigcup_{g\in G_1^n}\{k < j_0 : v = w_k \text{ and } w'_{k+t} \in g[v]_1\}.$$

which can be written in turn as

$$S = \bigcup_{v\in\mathcal{W}_n^c}\bigcup_{g\in G_1^n}\bigcup_{v'\in g[v]_1}\{k < j_0 : v = w_k \text{ and } w'_{k+t} = v'\}.$$

Thus, using (J10.1), we can estimate the size of S as

$$|S| \approx s_n|G_1^n|C_1^n\left(\frac{j_0}{s_n^2}\right).$$

Since $C_1^n = \frac{s_n}{Q_1^n}$, we can simplify this to $\frac{|G_1^n|}{Q_1^n}j_0$. Assumption (T6) follows.

Assumption (T7). Under the assumption that $[w_1']_1 \notin G_1^n[w_0]_1$, $s = 0$ and $\mathcal{Q}_0^n$ is the trivial equivalence relation. The estimate in (J11) simplifies to

$$\left|\frac{r(u',v')}{k_n} - \frac{1}{s_n^2}\right| < \epsilon_n. \tag{10.1}$$

To apply Remark 97, we again set $\mathcal{L} = W_n$ and $x = v$ and $|J(v)| = f(x)$, in the language of the remark. With this notation, $l = k_n$ and equation (10.1) is the hypothesis of

Remark 97. The conclusion of the remark is that

$$\frac{|\{t \in J(v) : \mathcal{C} \text{ occurs at } t \text{ in } [u'_1]_1[u'_2]_1 \dots [u'_{k_n-1}]_1\}|}{|J(v)|} \approx \frac{C_n^1}{s_n}. \tag{10.2}$$

Since $\frac{C_n^1}{s_n} = \frac{1}{Q_n^1}$, assumption (T7) follows.

Note that the verification of (T5)–(T7) uses Remark 97 for a small enough $\epsilon(\mu_n, s_n)$. We make this a requirement on ϵ_n.

Numerical Requirement 11. The number ϵ_n is sufficiently small relative to μ_n that the timing assumptions (T5)–(T7) hold.

Assumption (T4). Note that (T4) is the hardest timing assumption to verify. We motivate the proof by remarking that if u, v are long mutually random words in a language $\mathcal{L}$ that has s letters, then $\bar{d}(u, v) \approx 1 - 1/s^2$. Thus u and v are far apart. Specifications (J10.1) and (J11.1) imply that most (u, v) and their relative shifts are nearly mutually random. We use this to establish that w_0 and w_1 are distant in $\bar{d}$.

Numerical Requirement 12. One has $\epsilon_0 k_0 > 20$, the $\epsilon_n k_n$ are increasing and $\sum 1/\epsilon_n k_n$ is finite.

Let

$$\gamma_1 = (1 - 1/4 - \epsilon_0)(1 - 1/\epsilon_0 k_0)(1 - 1/l_0).$$

For $n \geq 2$, set

$$\gamma_n = \gamma_1 \prod_{0<m<n} (1 - 10(1/k_m\epsilon_m + 1/q_m + 1/l_m + 1/Q_1^m + \epsilon_{m-1}))$$

and finally

$$\gamma = \gamma_1 \prod_{0<m} (1 - 10(1/k_m\epsilon_m + 1/q_m + 1/l_m + 1/Q_1^m + \epsilon_{m-1})).$$

Assumption (T4) says that if $w_0, w_1 \in \mathcal{W}_n^c \cup \mathrm{rev}(\mathcal{W}_n^c)$ are not $\mathcal{Q}_1^n$-equivalent, then the overlaps of sufficiently long initial segments, or sufficiently long tail segments or of a sufficiently long initial segment with a tail segment of w_0 and w_1 are at least γ distant in $\bar{d}$. In (T4) *sufficiently long* means at least half of the length of the word. We prove something stronger by induction on n:

Proposition 98. *Let $n \geq 0$ and $w_0, w_1 \in \mathcal{W}_{n+1}^c \cup \mathrm{rev}(\mathcal{W}_{n+1}^c)$ with $[w_0]_1 \neq [w_1]_1$. Let I be an initial segment and let T be a tail segment of $\{0, 1, \dots, q_{n+1} - 1\}$ of the same length $\ell > \epsilon_n q_{n+1}$. Then we have*

$$\bar{d}(w_0 \upharpoonright I, w_1 \upharpoonright I) \geq \gamma_{n+1}, \tag{10.3}$$

$$\bar{d}(w_0 \upharpoonright T, w_1 \upharpoonright T) \geq \gamma_{n+1}, \tag{10.4}$$

$$\bar{d}(w_0 \upharpoonright I, w_1 \upharpoonright T) \geq \gamma_{n+1}. \tag{10.5}$$

Proof. We will consider the situation where $w_0, w_1 \in \mathcal{W}_{n+1}^c$. The situation where they both belong to $\mathrm{rev}(\mathcal{W}_{n+1}^c)$ follows, and the argument in the case where w_0, w_1 have different parities is a small variation of the basic argument.

The strategy for the proof is to consider $n+1$-words w_0 and w_1 and gradually eliminate small portions of I and T so that we are left with only segments of n-words that align in w_0 and w_1 in such a way that they have large $\bar{d}$-distance. The remaining portions of the w_0 and w_1 are far apart and they constitute most of the segments of each word. By Remark 22, we get an estimate on the distance of w_0 and w_1.

Suppose that

$$w_0 = \mathcal{C}(u_0, u_1, \ldots, u_{k_n-1}),$$
$$w_1 = \mathcal{C}(v_0, v_1, \ldots, v_{k_n-1}),$$

and let $u_i' = c_n^{-1}(u_i), v_i' = c_n^{-1}(v_i)$.

A general initial segment $w \restriction I$ of a word $w \in \mathcal{W}_{n+1}^c$ has the following form with $q = q_n, k = k_n, l = l_n$. For some $0 \le i_0 \le q_n, 0 \le j_0 \le k_n$,

$$\prod_{i<i_0}\left(\prod_{j<k} b^{q-j_i} w_j^{l-1} e^{j_i}\right) * \left(\prod_{j<j_0} b^{q-j_{i_0}} w_j^{l-1} e^{j_{i_0}}\right) * (b^{q^*} w_{j_0}^{l^*} w^* e^{j^*}),$$

where w^* is a possibly empty, possibly incomplete n-word, $0 \le j^* < j_{i_0}, 0 \le l^* \le l-1$, $0 \le q^* \le q - j_{i_0}$. This is a block of complete 2-subsections, followed by a block of complete 1-subsections, followed by a possibly empty, incomplete 1-subsection.

Similarly, a general tail segment $w \restriction T$ as the following form:

$$(b^{q^*} w^* w_{j_0}^{l^*} e^{j^*}) * \left(\prod_{j_0 \le j<k} b^{q-j_{i_0}} w_j^{l-1} e^{j_{i_0}}\right) * \prod_{i_0 \le i<q}\left(\prod_{j<k} b^{q-j_i} w_j^{l-1} e^{j_i}\right).$$

Initial segments. We now argue for inequality (10.3). To start, we take $n = 0$. In this case $q_0 = 1$ and $q_1 = k_0 l_0$. The initial segment $w_i \restriction I$ are of the form

$$\prod_{j<j_0} b w_j^{l_0-1} * u,$$

where u is a proper initial segment of a word of the form $b w_{j_0}^{l_0-1}$ that has length M, for some $M < l_0$.

If we throw away the tail segment u, we have thrown away proportion $M/\epsilon_0 k_0 l_0$. Since $M < l_0$, we have removed a portion of less than $\epsilon_0 k_0$ and the segment I_0 that is left has proportion at least $1 - (1/\epsilon_0 k_0)$ and is made up of a product of j_0 many 1-subsections.

We now consider $n > 0$. Since $\epsilon_n q_{n+1} = (\epsilon_n k_n l_n q_n) * q_n$, one of the following holds:

(1) There are no complete 2-subsections, in which case we must have $j_0 + 1 > \epsilon_n k_n q_n$.

(2) There is at least one complete 2-subsection and $j_0 \ge \epsilon_n k_n$.

(3) There is at least one complete 2-subsection and $j_0 < \epsilon_n k_n$.

In the first case, since $j_0 + 1 > \epsilon_n k_n q_n$, we know that $j_0 > \epsilon_n k_n$. Thus eliminating the partial 1-subsection at the end we are left with a concatenation of at least $\epsilon_n k_n$ complete 1-subsections and we have removed less than $1/\epsilon_n k_n$ portion of I. Similarly in the second case we can eliminate the incomplete 1-subsection at the end by removing proportion less than $1/\epsilon_n k_n$ of I. In the final case by removing both the final incomplete 1-subsection and $(\prod_{j<j_0} b^{q-j_{i_0}} w_j^{l-1} e^{j_{i_0}})$ we eliminate at most $1/q_n$ proportion of I.

In all three cases, we are left an I_0 such that $w_0 \restriction I_0$ and $w_1 \restriction I_0$ are made up of a possibly empty initial segment of complete 2-subsections followed either by no complete 1-subsections or at least $\epsilon_n k_n$ complete 1-subsections. We now delete the boundary portions of $w_0 \restriction I_0$, which are aligned with the boundary portions of $w_1 \restriction I_0$. These have proportion $1/l_n$ in each complete 1-subsection – hence proportion $1/l_n$ of I_0. Let I_1 be the remaining portion of I. Then I_1 contains proportion at least

$$(1 - 1/\epsilon_n k_n - 1/q_n)(1 - 1/l_n)$$

of I.

Case 1: $[w_0]_1 \notin G_1^n[w_1]_1$.[40] Let u' be the concatenation of $(u'_0, u'_1 \dots u'_{k_n-1})$, and v' similarly the concatenation of the v'_i. Then $u', v' \in \mathcal{W}_{n+1}$ and $[u']_1 \notin G_1^n[v']_1$. Let $u, v \in \mathcal{W}_n$ and I^* be an initial or final segment of $\{0, 1, \dots, k_n - 1\}$ of length at least $\epsilon_n k_n$.

Sublemma 99. *If ϵ_n is sufficiently small as a function of Q_1^n, then*

$$\frac{|\{i \in I^* : [u'_i]_1 = [v'_i]_1\}|}{|I^*|}$$

is within $\frac{1}{Q_1^n}$ of $\frac{1}{Q_1^n}$.

Proof. Let (u^*, v^*) be the concatenations of $\{u'_i : i \in I^*\}$ and $\{v'_i : i \in I^*\}$. By (J11.1), we see that the number $r(u, v)$ of occurrences of (u, v) in (u^*, v^*) satisfies

$$\frac{r(u, v)}{|I^*|} \approx \left(\frac{1}{s_n}\right)^2. \tag{10.6}$$

Fix such an I^* and let $\mathcal{C}$ be a $\mathcal{Q}_1^n$-class. Then $\mathcal{C}$ has C_1^n elements. It follows from equation (10.6) that the number of occurrences of a pair (u, v) in (u^*, v^*) with $u, v \in \mathcal{C}$ takes proportion of $|I^*|$ approximately

$$\frac{(C_1^n)^2}{s_n^2} = \left(\frac{1}{Q_1^n}\right)^2$$

Since there are Q_1^n many classes $\mathcal{C}$ that need to be considered we see that the number of pairs u'_i and v'_i with $[u'_i]_1 = [v'_i]_1$ is approximately

$$(1/Q_1^n)|I^*|. \tag{10.7}$$

Hence for small enough ϵ_n, we can see the conclusion of the sublemma. ■

Numerical Requirement 13. The numbers ϵ_n should be small enough as a function of Q_1^n that estimate in the conclusion of Sublemma 99 hold:

$$\left| \frac{|\{i \in I^* : [u'_i]_1 = [v'_i]_1\}|}{|I^*|} - \frac{1}{Q_1^n} \right| < \frac{1}{Q_1^n}. \tag{10.8}$$

[40] We note that because $G_1^0 = \langle e \rangle$, if $n = 0$ we are in Case 1.

The locations in $w_0 \upharpoonright I_1$ are made up of powers u_i^{l-1}. These fall into two categories, those locations occurring in whole 2-subsections and those occurring in the final product of 1-subsections. Applying the previous reasoning separately to the whole 2-subsections and the either-empty-or-relatively-long product of 1-subsections at the end of I, we see that the proportion of u_i occurring in $w_0 \upharpoonright I_1$ across from a v_i in $w_1 \upharpoonright I_1$ that is $\mathcal{Q}_1^n$ equivalent is also extremely close to $1/Q_1^n$.

If $n = 0$, then specification (J11.1) implies that

$$\left|\bar{d}(u^*, v^*) - \frac{3}{4}\right| < \epsilon_0.$$

So $\bar{d}(w_0 \upharpoonright I_1, w_1 \upharpoonright I_1) > (1 - 1/4 - \epsilon_0)$ and hence

$$\bar{d}(w_0 \upharpoonright I, w_1 \upharpoonright I) > \gamma_1.$$

In general, the induction hypothesis yields that $\mathcal{Q}_1^n$-inequivalent words have $\bar{d}$-distance at least γ_n-apart. Thus on I_1,

$$\bar{d}(w_0 \upharpoonright I_1, w_1 \upharpoonright I_1) > (1 - 2/Q_1^n)\gamma_n. \tag{10.9}$$

Allowing for agreement on boundary portions and applying Remark 22 we see that

$$\bar{d}(w_0 \upharpoonright I, w_1 \upharpoonright I) \geq \left(1 - 2\left(\frac{1}{Q_1^n} + \frac{1}{\epsilon_n k_n} + \frac{1}{q_n} + \frac{1}{l_n}\right)\right)\gamma_n > \gamma_{n+1}.$$

Case 2: $[w_0]_1 \in G_1^n[w_1]_1$. In this case $n \neq 0$. Let $g \in G_1^n$ with $g[w_1]_1 = [w_0]_1$. Since $[w_0]_1 \neq [w_1]_1$, g is not the identity. Since G_1^n acts diagonally, for all i with u_i intersecting the interval I_1, we have $[u_i]_1 = g[v_i]_1$. In particular, $[u_i]_1 \neq [v_i]_1$.

Hence $\bar{d}(w_0 \upharpoonright I_1, w_1 \upharpoonright I_1) \geq \gamma_n$, and thus

$$\bar{d}(w_0 \upharpoonright I, w_1 \upharpoonright I) \geq \left(1 - 2\left(\frac{1}{\epsilon_n k_n} + \frac{1}{q_n} + \frac{1}{l_n}\right)\right)\gamma_n > \gamma_{n+1}.$$

Tail segments. The argument for tail segments (inequality (10.4)) follows the argument for initial segments, except that we delete small parts of the beginning of T, instead of the end of I.

Tail Segments compared to initial segments. To show inequality (10.5), we proceed by induction, considering $w_0, w_1 \in \mathcal{W}_{n+1}^c$. In the comparing two initial segments or two tail segments, not only did the 2 and 1-subsections line up, but the n-subwords did as well. When comparing initial segments with tail segments, the n-subwords may be shifted, causing additional complications. The proof proceeds as in the easier cases, eliminating small sections of I (or equivalently T) a bit at a time until we are left with n-words. The alignment of these n-words allows us to apply the induction hypothesis and conclude that the vast majority of I and T have $\bar{d}$-distance at least γ_n.

(a) Of the 2-subsections of w_0 that intersect I, at most one is not a subset of I (namely the last one), and similarly except for possibly the first 2-subsection intersecting $w_1 \upharpoonright T$, $w_1 \upharpoonright T$ is made up of whole 2-subsections.

(b) Each 2-subsection of $w_0 \restriction I$ overlaps one or two 2-subsections of $w_1 \restriction T$. An overlap of a 2-subsection of $w_0 \restriction I$ with a 2-subsection of $w_1 \restriction T$ that has proportion bigger than ϵ_n of the 2-subsection implies that the overlap contains at least $\epsilon_n k_n$ complete 1-subsections.

(1) Among the complete 2-subsections of $w_0 \restriction I$, delete overlaps of proportion less than ϵ_n.

(2) Delete the possible partial 2-subsection at the end of $w_0 \restriction I$ if it contains less than $\epsilon_n k_n$ complete 1-subsections.

The proportion of I that has been deleted is less than $2\epsilon_n$.

(c) It could be that some of the portions of the remaining 2-subsections start or end with incomplete 1-subsections; i.e. not a whole word of the form $b^{q_n - j_i} v_j^{l_n} e^{j_i}$. Delete these incomplete sections. This leaves initial or tail segments of 2-subsections of the form $\prod_{j<k_n} b^{q_n-j_i} v_j^{l_n-1} e^{j_i}$ that consist of at least $\epsilon_n k_n$ whole 1-subsections. This trimming removes at most $1/k_n\epsilon_n$ proportion of I.

(d) We also remove the boundary sections of $w_0 \restriction I$. This removes at most $1/l_n$ of what remains of I at this stage.

(e) We are left with a portion $I' \subset I$ such that $w_0 \restriction I'$ consisting entirely of 0-subsections. These are blocks of the form u_j^{l-1}, where $u_j \in \mathcal{W}_n^c$. Each individual n-word u_i can occur opposite a portion of $w_1 \restriction T$ in various ways. These are:

(i) u_i might occur exactly opposite a v_{i+t}[41] or

(ii) u_i might span portions of two copies of v_{i+t} in a power v_{i+t}^{l-1}. The two copies have the form $v_{i+t}v_{i+t}$, or

(iii) u_i might overlap a portion of the boundary of w_1. This can happen in two ways: boundary inside a 2-subsection (i.e. boundary of the form $e^{j_i}b^{q_n-j_i}$) and boundary between consecutive 2-subsections (i.e. boundary of the form $e^{j_i}b^{q_n-j_{i+1}}$). In each $u_i^{l_n-1}$ there are at most three copies of u_i overlapping boundary portions of w_1.

Hence by removing proportion at most $4/l_n$ we are left with a portion of $w_0 \restriction I$ consisting of powers of the words u_j that do not overlap any boundary in w_1.

(f) After the deletions described in (a)–(e) the remaining portions of $w_0 \restriction I$ consists of blocks of powers of u_i's in initial segments of 2-subsections:

$$u_0u_0\dots u_0 * u_0\dots u_0 \,\#\, u_1u_1\dots u_1 * u_1\dots u_1 \,\#\cdots\#\, u_k\dots u_k * u_k\dots u_k$$

and in tail segments of 2-subsections:

$$\begin{aligned} &u_ju_j\dots u_j * u_j\dots u_j \,\#\, u_{j+1}u_{j+1}\dots u_{j+1} * u_{j+1}\dots u_{j+1} \,\#\cdots \\ &\quad \#\, u_{k_n-1}\dots u_{k_n-1} * u_{k_n-1}\dots u_{k_n-1}, \end{aligned}$$

where $*$'s stand for u's deleted opposite boundary of w_1 and #'s stand for the boundary of w_0 that has been deleted. An important point for us is that in each block $k \geq \epsilon_n k_n$ and $k_n - j - 1 \geq \epsilon_n k_n$.

[41] This is what happens in the case that $n = 0$.

Consider the words u_j in situation described in item (e) (ii) above. The v_{i+t}'s split u_i into two pieces. By deleting a portion of the individual u_j's of size less than $\epsilon_{n-1}q_n$ we can assume that all of the overlap of u_j's is in sections of length at least $\epsilon_{n-1}q_n$. By doing this for all u_j's we remove a parts of the remaining elements of w_0 of proportion at most ϵ_{n-1}.

(g) We now look more carefully at the two types of blocks of words described in item (f). The case in item e.)i. is similar and easier than the case in item (e) (ii) so we omit it. Along the blocks described in (f) the initial portions of u_i are lined up with v_{i+t} and the second portions are lined up with v_{i+t+1}. Critically, the t is *constant* along the block.

According to whether $t = 0$ or not, we apply specifications (J11.1) (as in Case 1 of the *Initial segments* argument) and (J10) to see that at most proportion $2/Q_1^n$ of the words u_i in a segment of the forms in (f) are lined up with v_{i+t} are $\mathcal{Q}_1^n$-equivalent. Hence we can make a final deletion of proportion at most $2/Q_1^n$ to get a portion $I^* \subseteq I$ consisting of relatively long pieces of $\mathcal{W}_n^c$-words in $w_0 \restriction I'$ overlapping $\mathcal{W}_n^c$-words in $w_1 \restriction T$ that lie in different Q_1^n equivalence classes.

We now finish the argument using Remark 22. After all of the deletions we are left with I^* having at least $(1 - (2\epsilon_n + 1/\epsilon_n k_n + 5/l_n + \epsilon_{n-1} + 2/Q_1^n))$-proportion of I and $w_0 \restriction I^*$ consists of relatively long pieces of $\mathcal{W}_n^c$ words that are overlapping portions of $\mathcal{W}_n^c$ words in $w_1 \restriction T$ that lie in different $\mathcal{W}_1^n$-classes.

By the induction hypothesis each of the pieces of n-words in $w_0 \restriction I^*$ of $\bar{d}$-distance at least γ_n from the corresponding portion of w_1. Consequently,

$$\bar{d}(w_0 \restriction I, w_1 \restriction T) > \gamma_n(1 - (2\epsilon_n + 1/\epsilon_n k_n + 5/l_n + \epsilon_{n-1} + 2/Q_1^n)) > \gamma_{n+1},$$

thus finishing the proof of Proposition 98. ∎

Since assumption (T4) is an immediate corollary of Proposition 98, we have finished verifying the timing assumptions.

We note in passing that inequality (10.5) holds even if $w_0 = w_1$ provided that the choice of initial and tail segment misalign corresponding 1-subsections.

We have proved:

Theorem 100. *Suppose that $\mathbb{K}^c$ is a system in the range of F^s with construction sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$. Then $\langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$ satisfies the timing assumptions.*

11. The consistency of the numerical requirements

During the course of this construction we have accumulated numerical conditions about growth and decay rates of several sequences. The majority of the numerical constants are not inductively determined – they are given immediately by knowing a small portion of the tree $\mathcal{T}$. We call these *exogenous* requirements. Other sequences of numbers depend on previous choices for the numbers – hence are determined recursively. In this section we list the recursive requirements, explicate their interdependencies and resolve their consistency.

Some of the conditions are easy to satisfy, as they do not refer to other sequences. For example, Numerical Requirement 1 (that $\sum_n 1/l_n < \infty$) can be satisfied once and for all by assuming that $l_n > 20 * 2^n$. Others are trickier, in that they depend on the growth rates of the other sequences. For example, in defining the sequence of k_n's we require that k_n be large relative to the choice of s_{n+1}. We call the former type of conditions *Absolute* and the latter *Dependent*. The Dependent conditions introduce the risk of circular or inconsistent growth and decay rate conditions.

Our approach here is to gather all of the conditions arising in this paper and its predecessors and classify them as Absolute or Dependent. We label them A or D accordingly. This process allows us to make a diagram of the Dependent conditions to verify that there are no circularities. The lack of a cycle in the diagram gives a clear method of recursively satisfying all of the numerical conditions.

Due to an overabundance of numerical parameters we were forced into some awkward notational choices. As noted before we have two types of epsilons: the *lunate* ϵ_n, often used for set membership and the *classical* ε_n. They play similar but slightly different roles. The lunate epsilons come from construction requirements arising in [8] and their strengthenings. The classical epsilons come from requirements related to circular systems and realizing them as smooth systems. As is to be expected there is interaction between the two. This occurs via the intermediary numbers we called μ_n's in Numerical Requirements 5 and 11.

11.1. The numerical requirements collected

In this subsection we collect the relevant numerical requirements. Specifically, in constructing $F^s(\mathcal{T})$ we are presented with $\mathcal{T}$ as a subsequence $\langle \sigma_{n_i} : i \in \mathbb{N}\rangle$ of a fixed enumeration of $\mathbb{N}^{<\mathbb{N}}$.

In the formal statements of the specifications in [8] for the construction sequence corresponding to $\mathcal{T}$, $\mathcal{W}_n$ is built just in case $\sigma_n \in \mathcal{T}$. This leads to a construction sequence of the form $\langle \mathcal{W}_{n_i} : i \in \mathbb{N}\rangle$ with gaps corresponding to m's, where $\sigma_m \notin \mathcal{T}$. To simplify notation, we reindex $\langle \mathcal{W}_{n_i} : i \in \mathbb{N}\rangle$ as $\langle \mathcal{W}_i : i \in \mathbb{N}\rangle$, where $\langle \mathcal{W}_i : i < j\rangle$ is determined by $\langle \sigma_{n_i} : i < j\rangle$. In [8], the specifications discussed "successive" (or "consecutive") elements of $\mathcal{T}$. These are σ_m and σ_n that belong to $\mathcal{T}$, but have no $\sigma_j \in \mathcal{T}$ with $j \in (m, n)$. In our new notation successive elements σ_m and σ_n of $\mathcal{T}$ correspond to $\mathcal{W}_i$ and $\mathcal{W}_{i+1}$, where $m = n_i$. Having adopted this convention we do not distinguish between $\langle \mathcal{W}_i : i \in \mathbb{N}\rangle$ and $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$. To emphasize the dependence on $\mathcal{T}$, we will occasionally write $\langle \mathcal{W}_n(\mathcal{T}) : n \in \mathbb{N}\rangle$.

We begin with the requirements inherited from [8].

Inherited numerical requirements. We have changed the notation from [8] as described in Appendix A. The number of elements of $\mathcal{W}_m$ is denoted s_m; the numbers Q_s^m and C_s^m denote the number of classes and sizes of each class of $\mathcal{Q}_s^m$, respectively. In [8] we have sequences $\langle \epsilon_n : n \in \mathbb{N}\rangle$, $\langle s_n, k_n, e(n), p_n : n \in \mathbb{N}\rangle$

Inherited Requirement 1. The sequence $\langle \epsilon_n : n \in \mathbb{N}\rangle$ is summable.

Inherited Requirement 2. The number of $\mathcal{Q}^n_{s+1}$ classes inside each $\mathcal{Q}^n_s$ class is $2^{e(n)}$. The numbers $e(n)$ will be chosen to grow fast enough that

$$2^n 2^{-e(n+1)} < \epsilon_n. \tag{11.1}$$

If s is the maximal length of an element of $\mathcal{T} \cap \{\sigma_m : m \leq n\}$ and

$$|\mathcal{T} \cap \{\sigma_m : m \leq n\}| = i_0,$$

then we set

$$C_s^{i_0} = 2^{e(i_0)}$$

as well. This forces s_n, Q_s^n and C_s^n all to be powers of 2 that are determined by $e(n)$. In particular, let σ_m and σ_n be successive elements of $\mathcal{T}$. Then s_n is the number of words one gets by iteratively substituting $e(n)$ many elements into words in $\mathcal{W}_n^i/Q_i^n$ and closing under G_i^m are successive for $i = 0, 1, \ldots, s$.[42]

By Remark 96, s_n and $e(n)$ are monotonically co-determined. Hence we can state this requirement as saying:

s_{n+1} is large enough in terms of ϵ_n that inequality (11.1) holds.

Inherited Requirement 3. If $\mathcal{T} = \langle \sigma_{n_i} : i \in \mathbb{N} \rangle$, then

$$2\epsilon_i s_i^2 < \epsilon_{i-1}. \tag{11.2}$$

Inherited Requirement 4. We have

$$\epsilon_i k_i s_{i-1}^{-2} \to \infty \quad \text{as } i \to \infty. \tag{11.3}$$

Inherited Requirement 5. We have

$$\prod_{n \in \mathbb{N}} (1 - \epsilon_n) > 0. \tag{11.4}$$

Since this is equivalent to the summability of the ϵ_n-sequence, it is redundant and we will ignore it in the rest of this paper.

Inherited Requirement 6. There will be prime numbers p_n such that

$$K_n = p_n^2 s_{n-1} K_{n-1}$$

(i.e. $k_n = p_n^2 s_{n-1}$). The p_n's grow fast enough to allow the probabilistic arguments in [8] involving k_n to go through.

Inherited Requirement 7. The number s_n is a power of 2.

Inherited Requirement 8. The construction of $F(\mathcal{T})$ requires that if $\mathcal{T} = \langle \sigma_{i_n} : n \in \mathbb{N} \rangle$, then $\epsilon_n < 2^{-i_n}$.

[42] It is possible to give a closed form formula for this, but it is complicated and uninformative.

Numerical requirements introduced in this paper.

Numerical Requirement 1. One has $l_0 > 20$ and $\sum_{k=n} 1/l_k < 1/l_{n-1}$.

Numerical Requirement 2. $\langle \varepsilon_n : n \in \mathbb{N}\rangle$ is a sequence of numbers in $[0, 1)$ such that $6\sum_{n>N} \varepsilon_n < \varepsilon_N$.

Numerical Requirement 3. The numbers k_n, l_n and q_n grow fast enough that $\varepsilon_n k_n \to \infty$, $\varepsilon_n l_n \to \infty$, $\varepsilon_n q_n \to \infty$.

Numerical Requirement 4. One has

$$\sum \frac{|G_1^n|}{Q_1^n} < \infty,$$

which is satisfied if $\frac{|G_1^n|}{Q_1^n} < 2^{-n}$.

Numerical Requirement 5. μ_n is chosen small relative to $\min(\varepsilon_n, 1/Q_1^n)$.

Numerical Requirement 6. The number l_n is big enough relative to a lower bound determined by $\langle k_m, s_m : m \leq n\rangle$, $\langle l_m : m < n\rangle$ and s_{n+1} to make the periodic approximations to the diffeomorphism converge.[43] Moreover, $k_n \leq l_n$.

Numerical Requirement 7. We have $s_n \to \infty$ as $n \to \infty$ and s_{n+1} is a power of s_n.

Numerical Requirement 8. We have $s_{n+1} \leq s_n^{k_n}$.

Numerical Requirement 9. The ϵ_n are decreasing, $\epsilon_0 < 1/40$ and $\epsilon_n < \varepsilon_n$.

Numerical Requirement 10. The number k_n is chosen sufficiently large relative to a lower bound determined by s_{n+1}, ϵ_n so that the Law of Large Numbers argument from [8] works.

Numerical Requirement 11. The number ϵ_n is small relative to μ_n.

Numerical Requirement 12. One has $\epsilon_0 k_0 > 20$, the $\epsilon_n k_n$ are increasing and $\sum 1/\epsilon_n k_n$ is finite.

Numerical Requirement 13. The numbers ϵ_n should be small enough, as a function of Q_1^n, that estimate (10.8) holds.

11.2. Resolution

A list of parameters, their first appearances and their constraints. We classify the constraints on a given sequence according to whether they refer to other sequences or not. Requirements that inductively refer to the same sequence are straightforwardly consistent. Those that refer to other sequences risk the possibility of being circular and thus inconsistent. As noted above refer to the former as *Absolute* conditions and the latter as *Dependent* conditions.

[43] This is discussed in detail in [11, pp. 34–35], where the lower bound is called l_n^*.

(1) **The sequence $\langle k_n : n \in \mathbb{N}\rangle$.**
Absolute conditions: None for $\langle k_n : n \in \mathbb{N}\rangle$.
Dependent conditions:

(D1) Numerical Requirement 10, k_n depends on s_{n+1}, ϵ_n.

(D2) Inherited Requirement 6. We can satisfy Inherited Requirement 6 by taking k_n large enough to satisfy Numerical Requirement 10 and of the form

$$k_n = p_n^2 s_{n-1}.$$

(D3) From Inherited Requirement 4, equation (11.3) requires that $\epsilon_n k_n s_{n-1}^{-2}$ goes to ∞ as n goes to ∞. This can be satisfied by choosing k_n large enough as a function of ϵ_n, s_{n-1}. We note that equation (11.3) implies that $\sum 1/\epsilon_n k_n$ is finite.

(D4) Numerical Requirement 12 says that $\epsilon_0 k_0 > 20$ and the $\epsilon_n k_n$ are increasing and $\sum 1/\epsilon_n k_n$ is finite. As noted the last condition follows from D3. The other parts of Numerical Requirement 12 are satisfied by taking k_n large relative to ϵ_n.

(D5) Numerical Requirement 8 implies that k_n is large enough that $s_{n+1} \leq s_n^{k_n}$. This implies that k_n is large relative to s_{n+1}.

From (D1)–(D5), we see that k_n is dependent on the choices of $\langle k_m, l_m : m < n\rangle$, $\langle s_m : m \leq n+1\rangle$, and ϵ_n.

(2) **The sequence $\langle l_n : n \in \mathbb{N}\rangle$.**
Absolute conditions:

(A1) Numerical Requirement 1 says that $1/l_n > \sum_{k=n+1}^{\infty} 1/l_k$. We also require that $l_n > 20 * 2^n$, an exogenous requirement.

Dependent conditions:

(D6) By Numerical Requirement 6, the number l_n is bigger than a number determined by $\langle k_m, s_m : m \leq n\rangle$, $\langle l_m : m < n\rangle$ and s_{n+1}.

(D7) The sequence $\langle l_n : n \in \mathbb{N}\rangle$ must grow fast enough that $\varepsilon_{n+1} q_{n+1} \to \infty$. This can be arranged by making $\varepsilon_{n+1} q_{n+1} > n+1$. Since $q_{n+1} = k_n l_n q_n^2$, this puts lower bound on l_n dependent on ε_{n+1}.

Thus l_n depends on $\langle k_m, s_m : m \leq n\rangle$, $\langle l_n : m < n\rangle$, ε_{n+1} and s_{n+1}.

(3) **The sequences $\langle s_n : n \in \mathbb{N}\rangle$ and $\langle e(n) : n \in \mathbb{N}\rangle$.** We treat these sequences as equivalent since s_n is a power of 2 determined by $e(n)$ and the elements of the tree in the domain of the reduction. Moreover, increasing one increases the other and vice versa. Since they are co-determined, they are chosen at the same time.
Absolute conditions:

(A2) Inherited Requirement 7 says that s_n is a power of 2.

Numerical Requirement 7 says that:

(A3) The sequence s_n goes to infinity.

(A4) s_{n+1} is a multiple of s_n.

(A5) As $e(n)$ determines Q_1^n, Numerical Requirement 4 puts an exogenous sequence of lower bounds on $e(n)$, for example that

$$\frac{|G_1^n|}{Q_1^n} < 2^{-n}.$$

This requires that $e(n)$ be chosen large and, since $e(n)$ and s_n are inter-determined, can be satisfied by taking $s(n)$ large.

Dependent conditions:

(D8) Numerical Requirement 3 makes s_n depend on ϵ_{n-1}.

The result is that the number s_{n+1} depends on the first $n+1$ elements of the tree $\mathcal{T}$, $\langle k_m, s_m, l_m : m < n\rangle$, s_n, and ϵ_n.[44]

(4) **The sequence $\langle \epsilon_n : n \in \mathbb{N}\rangle$.**
Absolute conditions:

(A6) Numerical Requirement 9 and Inherited Requirement 1 require that $\langle \epsilon_n : n \in \mathbb{N}\rangle$ is decreasing and summable and $\epsilon_0 < 1/40$.

(A7) Inherited Requirement 8 says that if $\mathcal{T} = \langle \sigma_{i_n} : n \in \mathbb{N}\rangle$, then $\epsilon_n < 2^{-i_n}$

Dependent conditions:

(D9) Numerical Requirement 9 requires that $\epsilon_n < \varepsilon_n$.

(D10) Equation (11.2) of Inherited Requirement 3 says $2\epsilon_n s_n^2 < \epsilon_{n-1}$.

(D11) Numerical Requirement 11 says that ϵ_n must be small enough relative to μ_n.

(D12) Numerical Requirement 13 says that ϵ_n is small as a function of Q_1^n.

The result is that ϵ_n depends exogenously on the first n elements of $\mathcal{T}$, and on Q_1^n, s_n, ε_n, ϵ_{n-1} and μ_n.

(5) **The sequence $\langle \varepsilon_n : n \in \mathbb{N}\rangle$.**
Absolute conditions:

(A8) Numerical Requirement 2 says that $6\sum_{n>N} \varepsilon_n < \varepsilon_N$. This can be arranged by taking $\varepsilon_n < 12^{-n}\varepsilon_{n-1}$.

Dependent conditions: Numerical Requirement 3 imposes three Dependent conditions on ε_n: $\varepsilon_n k_n \to \infty$, $\varepsilon_n l_n \to \infty$, $\varepsilon_n q_n \to \infty$. We deal with these in turn.

(a) The requirement that $\langle \varepsilon_n k_n : n \in \mathbb{N}\rangle$ goes to infinity already follows from the fact that $\epsilon_n < \varepsilon_n$ and item (D4).

(b) $\langle \varepsilon_n l_n : n \in \mathbb{N}\rangle$ goes to infinity. This follows from $k_n \leq l_n$, which is covered in Dependent condition (D6).

(c) $\langle \varepsilon_n q_n : n \in \mathbb{N}\rangle$ goes to infinity. This follows from Dependent condition (D7).

Thus there are no new Dependent conditions.

(6) **The sequence $\langle Q_1^n : n \in \mathbb{N}\rangle$.**
Absolute conditions: There are no new Absolute conditions.

[44] It is important to observe that the choice of s_{n+1} does *not* depend on k_n or l_n.

Dependent conditions:

(D13) Numerical Requirement 4 says that

$$\frac{|G_1^n|}{Q_1^n} < 2^{-n}.$$

But since Q_1^n is determined by s_n and the first n-elements of the tree, Numerical requirement 4 is taken care of by (A5).

There are no new Dependent conditions.

(7) **The sequence $\langle \mu_n : n \in \mathbb{N} \rangle$.** This sequence gives the required pseudo-randomness in the timing assumptions.
Absolute conditions: There are no new Absolute conditions.
Dependent conditions:

(D14) Numerical Requirement 5 requires that μ_n be very small relative to ε_n and $\frac{1}{Q_1^n}$.

The number μ_n is dependent on ε_n and Q_1^n.

The recursive dependencies of the various coefficients are summarized in Figure 4, in which an arrow from a coefficient to another coefficient shows that the latter is dependent on the former. Here is the order the coefficients can be chosen consistently.

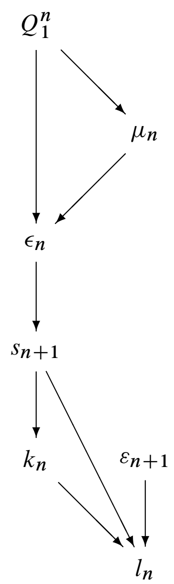

Fig. 4. Order of choice of Numerical parameters dependency diagram.

11.3. The inductive order of choices

We begin by setting $s_0 = 2, s_1 = 8, p_0 = 0, q_0 = k_0 = 1, l_0 = 21$; Q_1^0 is not defined, but Q_1^1 is determined by s_1; $\mu_0 = \epsilon_0 = k_0 = l_0 = 1, \varepsilon_0 = 1.1, \varepsilon_1 = \varepsilon_0/12$,

Assume:

The coefficient sequences $\langle k_m, l_m, Q_1^m, \mu_m, \epsilon_m : m < n\rangle$, $\langle \varepsilon_m : m \leq n\rangle$ and s_n have been chosen. The first $n + 1$ sequences on the tree $\mathcal{T}$ are known.

To do:

Choose $k_n, l_n, Q_1^n, \mu_n, \epsilon_n, \varepsilon_{n+1}$ and s_{n+1}. Each requirement is to choose the corresponding variable *large enough* or *small enough* where these adjectives are determined by the dependencies enumerated above.

Figure 4 gives an order to consistently choose the next elements on the sequences. Choose the successor coefficients in the following order:

$$Q_1^n, \varepsilon_{n+1}, \mu_n, \epsilon_n, s_{n+1}, k_n, l_n.$$

We note that Q_1^n is redundant in the diagram above since it is determined by s_n, but we include it as a bridge from stage $n - 1$.

Appendix A. Notation table

In this paper we have adopted the notation used in [1], which conflicts with the notation in [8], accordingly we provide a table for translating between the two. In the table, NEW means the notation used in this paper, OLD means the notation used in [8].

NEW	OLD	Description
s_n	W_n	s_n is the number of words in $\mathcal{W}_n^c$
k_n	l_{n+1}/l_n	the number of words concatenated to make $\mathcal{W}_{n+1}$ from $\mathcal{W}_n$
$e(n)$	$k(n)$	controls the number of $\mathcal{Q}_{s+1}$ classes in each $\mathcal{Q}_s$ class
γ	s_1	the separation between $\mathcal{Q}_1^n$ classes
K_n	l_n	K_n is this paper's notation for the lengths of the odometer based words in W_n, l_n was the notation for the lengths of the words in [8]
q_n	l_n	the lengths of the circular words in current paper vs. odometer based words in [8]; the new q_n refers to the lengths of the words in $\mathcal{W}_n^c$
l_n	no analogue	coefficient needed to grow fast for smooth transformations

An equivalent description of the numbers we are calling k_n in this paper is that they are the number of words in $\mathcal{W}_n^c$ concatenated to form elements of P_{n+1}. The number k_n is equal to the number K_{n+1}/K_n and l_{n+1}/l_n in the old notation of [8].

Funding. The first author would like to acknowledge partial support from the US National Science Foundation grant number 2100367.

References

[1] Anosov, D. V., Katok, A. B.: New examples in smooth ergodic theory. Ergodic diffeomorphisms. Trudy Moskov. Mat. Obšč. **23**, 3–36 (1970) Zbl 0218.58008 MR 0370662

[2] Anzai, H.: On an example of a measure preserving transformation which is not conjugate to its inverse. Proc. Japan Acad. **27**, 517–522 (1951) Zbl 0044.12502 MR 47742

[3] Beleznay, F., Foreman, M.: The complexity of the collection of measure-distal transformations. Ergodic Theory Dynam. Systems **16**, 929–962 (1996) Zbl 0869.58032 MR 1417768

[4] Bjork, A.-D.: Criteria for rank-one transformations to be weakly mixing, and the generating property. ProQuest LLC, Ann Arbor (2009); PhD Thesis, University of California, Irvine (2009) MR 2713155

[5] Foreman, M.: A descriptive view of ergodic theory. In: Descriptive Set Theory and Dynamical Systems (Marseille–Luminy, 1996), London Math. Soc. Lecture Note Ser. 277, Cambridge University Press, Cambridge, 87–171 (2000) Zbl 0962.03043 MR 1774425

[6] Foreman, M.: What is a Borel reduction? Notices Amer. Math. Soc. **65**, 1263–1268 (2018) Zbl 1402.03069 MR 3837073

[7] Foreman, M.: Gödel diffeomorphisms. arXiv:2009.06735 (2020)

[8] Foreman, M., Rudolph, D. J., Weiss, B.: The conjugacy problem in ergodic theory. Ann. of Math. (2) **173**, 1529–1586 (2011) Zbl 1243.37006 MR 2800720

[9] Foreman, M., Rudolph, D. J., Weiss, B.: Universal models for measure preserving transformations. To appear

[10] Foreman, M., Weiss, B.: An anti-classification theorem for ergodic measure preserving transformations. J. Eur. Math. Soc. (JEMS) **6**, 277–292 (2004) Zbl 1063.37004 MR 2060477

[11] Foreman, M., Weiss, B.: A symbolic representation for Anosov–Katok systems. J. Anal. Math. **137**, 603–661 (2019) Zbl 1429.37003 MR 3938015

[12] Foreman, M., Weiss, B.: From odometers to circular systems: A global structure theorem. J. Mod. Dyn. **15**, 345–423 (2019) Zbl 07350263 MR 4042165

[13] Foreman, M., Weiss, B.: Odometer based systems. arXiv:2009.10162 (2020)

[14] Halmos, P. R.: Lectures on Ergodic Theory. Chelsea Publishing, New York (1960) Zbl 0096.09004 MR 0111817

[15] Halmos, P. R., von Neumann, J.: Operator methods in classical mechanics. II. Ann. of Math. (2) **43**, 332–350 (1942) Zbl 0063.01888 MR 6617

[16] Hjorth, G.: Classification and Orbit Equivalence Relations. Math. Surveys Monogr. 75, American Mathematical Society, Providence (2000) Zbl 0942.03056 MR 1725642

[17] Hjorth, G.: On invariants for measure preserving transformations. Fund. Math. **169**, 51–84 (2001) Zbl 0990.03036 MR 1852353

[18] Katok, A.: Combinatorial Constructions in Ergodic Theory and Dynamics. Univ. Lecture Ser. 30, American Mathematical Society, Providence (2003) Zbl 1030.37001 MR 2008435

[19] Kechris, A. S.: Classical Descriptive Set Theory. Grad. Texts in Math. 156, Springer, New York (1995) Zbl 0819.04002 MR 1321597

[20] Marker, D.: Descriptive set theory. http://homepages.math.uic.edu/~marker/math512/dst.pdf (2002)

[21] Petersen, K.: Ergodic Theory. Cambridge Stud. Adv. Math. 2, Cambridge University Press, Cambridge (1989) Zbl 0676.28008 MR 1073173

[22] Shields, P. C.: The Ergodic Theory of Discrete Sample Paths. Grad. Stud. Math. 13, American Mathematical Society, Providence (1996) Zbl 0879.28031 MR 1400225

[23] von Neumann, J.: Zur Operatorenmethode in der klassischen Mechanik. Ann. of Math. (2) **33**, 587–642 (1932) Zbl 0005.12203 MR 1503078

[24] Walters, P.: An Introduction to Ergodic Theory. Grad. Texts in Math. 79, Springer, New York (1982) Zbl 0475.28009 MR 648108