UC Irvine

Title

From odometers to circular systems: A global structure theorem

Permalink

Journal

ISSN

Authors

Foreman, Matthew

Weiss, Benjamin

Publication Date

2019

DOI

Copyright Information

Peer reviewed

From Odometers to Circular Systems:
A Global Structure Theorem

Matthew Foreman, Benjamin Weiss

March 22, 2017

Abstract

The main result of this paper is that two large collections of ergodic measure preserving systems, the *Odometer Based* and the *Circular Systems* have the same global structure with respect to joinings. The classes are canonically isomorphic by a continuous map that takes factor maps to factor maps, measure-isomorphisms to measure-isomorphisms, weakly mixing extensions to weakly mixing extensions and compact extensions to compact extensions. The first class includes all finite entropy ergodic transformations with an odometer factor. By results in [5], the second class contains all transformations realizable as diffeomorphisms using the untwisted Anosov-Katok method. An application of the main result will appear in a forthcoming paper that shows that the diffeomorphisms of the torus are inherently unclassifiable up to measure-isomorphism. Other consequences include the existence measure distal diffeomorphisms of arbitrary countable distal height.

Contents

1

1 Introduction

The isomorphism problem in ergodic theory was formulated by von Neumann in 1932 in his pioneering paper [19]. Simply put it asks to determine when two measure preserving actions are isomorphic, in the sense that there is a measure isomorphism between the underlying measure space that intertwines the actions. It has been solved completely only for some special classes of transformations. Halmos and von Neumann [13] used the unitary operators defined by Koopman to completely characterize ergodic measure preserving transformations with pure point spectrum, these transformations can be concretely realized (in a Borel way) as translations on compact groups. Another notable success was the use of the Kolmogorov entropy to distinguish between measure preserving systems. Ornstein's work showed that entropy completely classifies a large class of highly random systems, such as independent processes, mixing Markov chains and certain smooth systems such as geodesic flows on surfaces of negative curvature.

Closely related to the isomorphism problem is the study of structural properties of measure preserving systems. These including mixing properties and compactness. A famous example is the Furstenberg-Zimmer structure theorem for ergodic measure preserving transformations, which characterizes every ergodic transformation as an inverse limit system of compact extensions followed by a weakly mixing extension. This result is fundamental for studying recurrence properties of measure preserving systems and the related proofs of Szemeredi-type combinatorial theorems ([9]).

In this paper we present a new phenomenon, *Global Structure Theory.* Most structure theorems in ergodic theory consider a single transformation *in vitro*. The approach here is study whole, intact ecosystems of transformations with their inherent relationships.

Our main result shows that two large collections of measure preserving transformations have exactly the same structure with respect to factors and isomorphisms (and more generally, joinings). More concretely, define the *odometer based* transformations to be those finite entropy transformations that contain a non-trivial odometer factor. Spectrally, this is equivalent to the associated unitary operator having infinitely many finite period eigenvalues. To each odometer, we can associate a class of symbolic systems, the

3

circular systems. In [5], it is shown that the circular systems coincide exactly with the ergodic transformations realizable as diffeomorphisms of the torus using the untwisted method of Approximation-by-Conjugacy, due to Anosov-Katok ([1]).

We can make two categories by taking the objects to be these two classes of systems and by taking morphisms to be factor maps (or more generally joinings) that preserve the underlying timing structure. The main theorem of this paper says that these two categories are isomorphic by a map that takes measure-isomorphisms to measure-isomorphisms, weakly mixing extensions to weakly mixing extensions and compact extensions to compact extensions. It follows that it takes distal towers to distal towers. Moreover the map preserves the simplex of non-atomic invariant measures, takes rank one transformations to rank one transformations and much more. (This will be discussed further in the forthcoming [8].) In other words the global structure of these two categories is identical.

We can get more detail by considering systems based on a fixed odometer map and circular systems based on that odometer map and an arbitrary fast growing coefficient sequence. Doing so gives us collections of pairwise isomorphic categories that can be amalgamated to yield the statement above. The main theorem is framed in this more granular setting.

Our result might be a mere curiosity, were it not for an application which we now describe.

Foreshadowed by a remarkable early result by Feldman [4], in the late 1990's 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 perhaps 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 the stronger the theorem. One natural notion is 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. Say that X or f is *not* Borel is saying that determining membership in X or computing f cannot be done with any amount of countable resources.

4

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 information and no countable transfinite protocol for determining, for arbitrary $x, y \in X$ whether xEy. *Any* such method must inherently use uncountable resources.[1]

In considering the isomorphism relation as a collection $\mathcal{I}$ of pairs (S, T) of measure preserving transformations, Hjorth showed that $\mathcal{I}$ is not a Borel set. However the pairs of transformations he used to demonstrate this were inherently non-ergodic[2], leaving open the essential problem:

Is isomorphism of ergodic measure preserving transformations Borel?

This question was answered by Foreman, Rudolph and Weiss in [6], where they gave a negative answer. 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 where von Neumann formulated the isomorphism problem he expressed the likelihood that any abstract measure preserving transformation is isomorphic to a continuous measure preserving transformation and perhaps even to a differentiable one. This brief remark eventually gave rise to one of the outstanding problems in smooth dynamics, namely:

Does every ergodic MPT have a smooth model?

By a smooth model is meant an isomorphic copy of the transformation 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.

This paper is the second in a series of papers whose original purpose was 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 it is in general. We show

---

[1]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.

[2]The ergodic components of the pairs were rotations of the circle.

this to be the case even when restricting to diffeomorphisms of the 2-torus that preserve Lebesgue measure this is the case. In the third paper we will complete the proof of the following theorem:

Theorem (Anti-classification of Diffeomorphisms). *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.*

It was natural for us to try to adapt our earlier work to establish this result. However we were faced at first with the following difficulty. The transformations built in [6] were based on odometers (in the sense that the Kronecker factor was an odometer). It is a well known open problem whether it is possible to have any smooth transformation on a compact manifold that has a non-trivial odometer factor. Thus proving the anti-classification theorem in the smooth context required constructing a different collection of hard-to-classify transformations and then showing that this collection could be realized smoothly. This is our application of the main result of this paper.

The paper ([5]) constructed a new collection of systems, the *Circular Systems*, which are defined as symbolic systems constructed using the *Circular Operator*, a formal operation on words. The main result in [5] has as a consequence that uniform circular systems can be realized as smooth models using the method developed by Anosov and Katok.

The primary theorem of this paper allows us to transfer the general isomorphism structure for odometer based systems to the isomorphism structure for circular systems, at least up to automorphisms of the underlying odometer or rotation. Namely there remains the issue of preserving the timing mechanism. In the forthcoming [7] it is shown how to construct odometers so that for the resulting circular systems, up to a small correction factor, all isomorphisms preserve the underlying timing structure. This allows us to conclude the proof of the anti-classification theorem for diffeomorphisms.

Here is a more concrete description of the results in the paper. In the present paper we are concerned with the entire class $\mathcal{OB}$ of systems based on a fixed odometer and the relations between them. The odometer is determined by a sequence of positive integers greater than one, $\langle k_n : n \in \mathbb{N} \rangle$. The the circular operator is determined by an additional sequence of integers $\langle l_n : n \in \mathbb{N} \rangle$. For this paper, the sequence of l_n's can be arbitrary subject to

the requirement that $\sum 1/l_n < \infty$. However for realizing circular systems as diffeomorphisms there is a fixed growth rate, determined by the size of the alphabet of the odometer based system and $\langle k_n : n \in \mathbb{N} \rangle$, that the sequence of l_n's must eventually exceed.

We describe $\mathcal{OB}$ symbolically here, but show in a forthcoming paper that $\mathcal{OB}$ consists of representations of arbitrary ergodic systems with finite entropy that have the specific odometer as a factor. In the language of "cutting and stacking" constructions these are those constructions where no spacers are introduced. We fix $\langle l_n : n \in \mathbb{N} \rangle$, and hence a sequence of circular operators. Applying these to each of the elements of $\mathcal{OB}$ we obtain a second class, $\mathcal{CB}$, of circular systems. This class consists of some of the extensions of a fixed irrational rotation which is determined by the circular operator. As remarked above, for suitably chosen coefficient sequences, this class can be characterized as those transformations realizable as diffeomorphisms using the Anosov-Katok technique. We consider the two classes as categories where the morphisms are graph joinings which are either the identity of the base or reverse it. These are called *synchronous* and *anti-synchronous* joinings respectively. Our main theorem then takes the form:

Theorem 1. *For a fixed circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N} \rangle$ the categories $\mathcal{OB}$ and $\mathcal{CB}$ are isomorphic by a functor $\mathcal{F}$ 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.*[3]

It is natural to extend the collections of morphisms of $\mathcal{OB}$ and $\mathcal{CB}$ to general synchronous and non-synchronous joinings. Because the ergodic joinings are not closed under composition, in extending Theorem 1 one is forced to consider at least *some* non-ergodic joinings. At the end of the paper we discuss how to extend Theorem 1 to expanded categories that have as morphisms arbitrary synchronous and anti-synchronous joinings. This involve expanding our analysis of generic sequences to non-ergodic joinings. We also describe some detailed analysis of the combinatorics behind the isomorphism $\mathcal{F}$.

We have provided a detailed table of contents which enumerates the contents of the paper. Here is a brief summary. Much of the section following

---

[3]E. Glasner showed that the functor takes compact extensions to compact extensions.

this one is standard, with the exception §2.6, which is exposes generic sequences for transformations and extends that notion to joinings. In §3, the reader will find an explanation of our two categories and a proof that circular systems contain a canonical rotation factor. Section 4 is primarily concerned with defining a map ♮ that is a symbolic analogue of complex conjugation on the unit circle. In sections 5 and 6 the mapping $\mathcal{F}$ is defined on morphisms, while §7 contains the proof of the main result. In §8 there is a more detailed analysis of of the dynamical properties of our mapping $\mathcal{F}$ which may prove useful in the future, and in the final section we collect some problems that are left open.

1.1 Acknowledgements

This work was inspired by the pioneering work of our co-author Dan Rudolph, who passed away before this portion of the grand project was undertaken. We owe an inestimable debt to J.P. Thouvenot who suggested using the Anosov-Katok technique to produce our badly behaved transformations rather than directly attacking the "odometer obstacle." We would like to thank E. Glasner for showing that $\mathcal{F}$ preserves compact extensions. Finally the first author would like to thank Christian Rosendal for asking very useful questions about how general our results were.

2 Preliminaries

This section establishes some of the conventions we follow in this paper. There are many sources of background information on this including any standard text or [20], [15]. A small portion of the material in this section was presented in [5], but is repeated here in an attempt to be self-contained. The reader is referred to [5] for any missing definitions.

2.1 Measure Spaces

We will call separable non-atomic probability spaces *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 $\mathcal{B}$.[4] We

---

[4]We will occasionally make an exception to this by calling discrete probability measures on a finite set *measures*; we hope that context makes the difference clear.

will often identify two members of $\mathcal{B}$ that differ by a set of μ-measure 0 and seldom distinguish between $\mathcal{B}$ and the σ-algebra of classes of measurable sets modulo measure zero unless we are making a pointwise definition and need to claim it is well defined on equivalence classes.

We will frequently use without explicit mention the Maharam-von Neumann result 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 (resp.) 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* or *conjugacy* between systems iff ϕ is a measure isomorphism. Following common practice, we will use the word *conjugacy* interchangeably with *isomorphism* in this context.

For a fixed measure space (X, μ) we can consider the collection of measure preserving transformations $T : X \to X$. These form a group that can be endowed with a Polish topology that has basic open sets described as follows. We fix a finite measurable partition $\mathcal{A}$ of X and an $\epsilon > 0$ and take as a neighborhood of T

$$\mathcal{N}(T, \mathcal{A}, \epsilon) =_{def} \{S : \sum_{a \in \mathcal{A}} \mu(Ta \Delta Sa) < \epsilon\}.$$

Details about this topology can be found in many sources including [12], [20].

2.2 Joinings

We remind the readers of the definitions. Extensive treatments of joinings can be found in [11] or [16]. All of the definitions and basic results about joinings necessary for this paper occur in Chapter 6 of the latter reference.

Definition 2. *A joining between two measure preserving systems $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ is a measure ρ on $X \times Y$ defined on the product σ-algebra $\mathcal{B} \otimes \mathcal{C}$ such that*

1. ρ is $T \times S$ invariant,

2. for each set $B \in \mathcal{B}$, $\rho(B \times Y) = \mu(B)$,

3. for each set $C \in \mathcal{C}$, $\rho(X \times C) = \nu(C)$.

The graphs of factor maps provide natural examples of joinings. We characterize these with a definition.

Definition 3. *A joining ρ is a* graph joining *between X and Y if and only if for all $C \in \mathcal{C}$ and all $\epsilon > 0$, there is a $B \in \mathcal{B}$ such that*

$$\rho((B \times Y) \Delta (X \times C)) < \epsilon.$$

A joining ρ between $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ is an invertible graph joining *if and only for all $B \in \mathcal{B}$ there is a $C \in \mathcal{C}$ such that*

$$\rho((B \times Y) \Delta (X \times C)) = 0 \tag{1}$$

and vice versa: for all $C \in \mathcal{C}$, there is a $B \in \mathcal{B}$ such that equation 1 holds.

Here are some standard facts (see [11]):

Proposition 4. *Let $\mathbb{X} = (X, \mathcal{B}, \mu, T)$ and $\mathbb{Y} = (Y, \mathcal{C}, \nu, S)$. Then*

1. *There is a canonical one-to-one correspondence between the collection of graph joinings of $\mathbb{X}$ and $\mathbb{Y}$ and the collection of factor maps from X to Y. A graph joining concentrates on the graph of the factor map. We can represent the graph joining corresponding to a measure preserving map $\phi : X \to Y$ by*

$$\rho_\phi = \int (\delta_x \times \delta_{\phi(x)}) d\mu(x).$$

2. *There is a canonical one-to-one correspondence between the collection of invertible graph joinings of $\mathbb{X}$ and $\mathbb{Y}$ and the collection of conjugacies between $\mathbb{X}$ and $\mathbb{Y}$.*

3. *Suppose that $\mathcal{B}' \subseteq \mathcal{B}$ and $\mathcal{C}' \subseteq \mathcal{C}$ are Boolean algebras that generate $\mathcal{B}$ and $\mathcal{C}$ respectively as σ-algebras. Let ρ be a joining of $\mathbb{X}$ with $\mathbb{Y}$ such that for all $\epsilon > 0$ and all $C \in \mathcal{C}'$ there are $B_1, \ldots B_n \in \mathcal{B}'$ such that we have $\rho(\bigcup_i (B_i \times Y) \Delta (X \times C)) < \epsilon$, then ρ is a graph joining.*

10

We note that perhaps a more proper term for an invertible graph joining is the earlier usage *diagonal joining*. In view of the results of this section we will often be careless and say that *ρ is a factor map* or *ρ is a conjugacy/isomorphism* to mean that ρ is a graph joining or ρ is an invertible graph joining.

To each joining ρ of $\mathbb{X}$ and $\mathbb{Y}$ we can associate its adjoint ρ^*, the joining of $\mathbb{Y}$ with $\mathbb{X}$ defined for $B \in \mathcal{B}$ and $C \in \mathcal{C}$ as:

$$\rho^*(C \times B) = \rho(B \times C).$$

If ρ is a graph joining corresponding to a factor map $\pi : X \to Y$, then ρ^* concentrates on $\{(y, x) : \pi(x) = y\}$.

The following is immediate:

Proposition 5. *ρ is an invertible graph joining if and only if both ρ and ρ^* are graph joinings.*

Thus we can apply Proposition 4, item 3 to both ρ and ρ^* to get a criterion for being the joining associated with a conjugacy.

A potential source of confusion. Proposition 4 allows us to identify graph joinings with factor maps and invertible graph joinings with conjugacies. These joinings are always *ergodic* as joinings. However, there are non-ergodic conjugacies between ergodic measure preserving transformations. More explicitly: there are ergodic systems (X, T) and (X, S) and non-ergodic isomorphisms $\phi : (X, T) \to (X, S)$.[5] The associated joining ρ_ϕ is, however, ergodic as a $T \times S$-invariant measure.

Let $(X, \mu), (Y, \nu)$ and $(Z, \tilde{\mu})$ be measure spaces and $\pi_X : X \to Y$ and $\pi_Z : Z \to Y$ be factor maps. We can define a canonical joining of $\mathbb{X}$ and $\mathbb{Z}$ that reflects the factor structure as follows. We let $\{\mu_y : y \in Y\}$ and $\{\tilde{\mu}_y : y \in Y\}$ be the disintegrations of $\mathbb{X}$ and $\mathbb{Z}$ over $\mathbb{Y}$ respectively. The *relatively independent joining* of $\mathbb{X}$ and $\mathbb{Z}$ over $\mathbb{Y}$ is the joining ρ:

$$\rho = \int (\mu_y \times \tilde{\mu}_y) d\nu(y).$$

We will sometimes write this as $X \times_Y Z$.

---

[5]The second author has given examples of of isomorphic ergodic transformations where every conjugacy is non-ergodic.

We will be concerned about categories of measure preserving systems where the morphisms are joinings. For this we must describe the composition operation. Suppose we are given joinings ρ_{XY} between X and Y and ρ_{YZ} between Y and Z. Then (Y, ν) is a common factor of both $(X \times Y, \rho_{XY})$ and $(Y \times Z, \rho_{YZ})$ and we can consider the relatively independent joining $\rho_{XY} \times_Y \rho_{YZ}$.

We define the composition of ρ_{XY} and ρ_{YZ} to be the projection of the relatively independent joining of ρ_{XY} and ρ_{YZ} to a measure on $X \times Z$. Formally, if $A \subseteq X \times Z$ and ρ is the relatively independent joining, then:

$$\rho_{XY} \circ \rho_{YZ}(A) = \rho(\{(x, y, z) : x, z \in A\}).$$

Example 6. *Suppose that $\pi_0 : X \to Y$ and $\pi_1 : Y \to Z$ are factor maps. If ρ_{XY} is the joining associated with π_0 and ρ_{YZ} is the joining associated with π_1, then $(\rho_{YZ}^* \circ \rho_{XY}^*)^*$ is the joining associated with the factor map $\pi_1 \circ \pi_0 : X \to Z$.[6]*

The following are standard facts (e.g. in §6.2 of [11]):

Proposition 7. *1. The operation of composition of joinings is associative: if ρ_1, ρ_2 and ρ_3 are joinings, then*

$$(\rho_1 \circ \rho_2) \circ \rho_3 = \rho_1 \circ (\rho_2 \circ \rho_3).$$

2. Suppose that $\pi^X : X \to X'$ and $\pi^Z : Z \to Z'$ are factor maps Let ρ_1 and ρ_2 be joinings of X, Y and Y, Z respectively. Let ρ_1^π be the projection of ρ_1 to a joining of X' and Y via $\pi^X \times id$ and ρ_2^π be defined similarly. Finally let $(\rho_1 \circ \rho_2)^\pi$ be the projection of the composition of ρ_1 and ρ_2 to a joining of X with Z. Then:

$$\rho_1^\pi \circ \rho_2^\pi = (\rho_1 \circ \rho_2)^\pi.$$

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.

---

[6]In the following, in the context of factor maps $\pi : X \to Y$ we will be sloppy about whether this is associated with a joining of X with Y or a joining of Y with X.

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: $\langle u \rangle_k = \{f \in \Sigma^{\mathbb{Z}} : f \restriction [k, k+n) = u\}$. The collection of cylinder sets form a base for the product topology on $\Sigma^{\mathbb{Z}}$.

Notation: For a word $w \in \Sigma^{<\mathbb{N}}$ we will write $|w|$ for the length of w. We will write $1_{\langle w \rangle}$ for the characteristic function of the interval $\langle w \rangle_0$ in $\Sigma^{\mathbb{Z}}$.

The shift map:
$$sh : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$$
defined by setting $sh(f)(n) = f(n+1)$ is a homeomorphism. If μ is a shift invariant Borel measure then the resulting measure preserving system $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, sh)$ 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 each $w' \in \mathcal{W}_{n+1}$,

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
$$u_0 w_0 u_1 w_1 \ldots w_l u_{l+1} \tag{2}$$
such that each $w_i \in \mathcal{W}_n$, $u_i \in \Sigma^{<\mathbb{N}}$ and for this parsing
$$\frac{\sum_i |u_i|}{q_{n+1}} < \epsilon_{n+1}.$$

The segments u_i in condition 2 are called the *spacer* or *boundary* portions of w.

Definition 8. *A sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ satisfying properties 1.)-3.) will be called a* construction sequence.

13

Associated with a construction sequence is a symbolic shift defined as follows. Let $\mathbb{K}$ be the collection of $x \in \Sigma^{\mathbb{Z}}$ such that every finite contiguous subword of x occurs inside some $w \in \mathcal{W}_n$. Then $\mathbb{K}$ is a closed shift invariant subset of $\Sigma^{\mathbb{Z}}$ that is compact if Σ is finite.[7]

Formally, we have constructed a symbolic shift. To get a measure preserving system we find a shift invariant measure μ concentrating on $\mathbb{K}$ and write $(\mathbb{K}, \mu)$. In [5] we define the notion of a *uniform* construction sequence and show that the resulting $\mathbb{K}$ are uniquely ergodic.

We want to be able to unambiguously parse elements of $\mathbb{K}$. For this we will use construction sequences consisting of uniquely readable words.

Definition 9. *Let Σ be a language and $\mathcal{W}$ be a collection of finite words in Σ. Then $\mathcal{W}$ is* uniquely readable *iff whenever $u, v, w \in \mathcal{W}$ and $uv = pws$ then either p or s is the empty word.*

In our constructions we will restrict our measures to a natural set:

Definition 10. *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, $x \upharpoonright [-a_m, b_m) \in \mathcal{W}_n$.*

Note that S is a dense shift invariant $\mathcal{G}_\delta$ set. The following lemma is routine:

Lemma 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.*

⊢   Item 1 is clear from the definitions. If X is a Polish space, $T : X \to X$ is a Borel automorphism and D is a T-invariant Borel set with a unique T-invariant measure on D, then that measure must be ergodic.   ⊣

Let $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ be a uniquely readable construction sequence, and $s \in S$. By the unique readability, for each n either $s(0)$ lies in a well-defined

---

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

subword of s belonging to $\mathcal{W}_n$ or in a spacer of a subword of s belonging to some $\mathcal{W}_{n+k}$.

Lemma 12. *Suppose that $\mathbb{K}$ is built from $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ and ν is a shift 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 $s \restriction [a_n, b_n) \in \mathcal{W}_n$.*

$\vdash$ Let B_n be the collection of $s \in S$ such that for some $a_n \leq 0 < b_n$, $s \restriction [a_n, b_n) \in \mathcal{W}_n$ but $s(0)$ is in a boundary portion of $s \restriction [a_n, b_n)$. By the Ergodic Theorem and clause 3.) of the definition of a construction sequence $\sum \nu(B_n) < \infty$.

It follows from the Borel-Cantelli Lemma that for almost all s there is an N such that for all $n \geq N$, $s \notin B_n$. Fix an $s \in S$ and such an N. From the definition of S there are arbitrarily large $n^* > N$ and $a_{n^*} \leq 0 < b_{n^*}$ such that $s \restriction [a_{n^*}, b_{n^*}) \in \mathcal{W}_{n^*}$. Using backwards induction from n^* to N and the definition of B_n, this also holds for all $n \in [N, n^*)$. $\dashv$

2.4   Locations

By Lemma 12 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 13. *Let $s \in S$ and suppose that for some $0 \leq k < q_n$, $s \restriction [-k, q_n - k) \in \mathcal{W}_n$. We define $r_n(s)$ to be the unique k with 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's will be called the* location sequence *of s.*

We interpret $r_n(s) = k$ as saying that $s(0)$ *is the kth 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 14. *Suppose that $s \in S$ has a principal n-block for all $n \geq N$. Let $N \leq n < m$. It follows immediately from the definitions that $r_n(s)$ and $r_m(s)$ are well defined and the $r_m(s)^{th}$ position of the principal m-block of s is in the $r_n(s)^{th}$ position inside the principal n-block of s.*

The next lemma tells us that 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.

Lemma 15. *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 $s = s'$.*

$\vdash$ Since $s, s' \in S$ there are sequences $\langle a_n, a'_n, b_n, b'_n : n \geq N \rangle$ tending to infinity such that $s \upharpoonright [-a_n, b_n) \in \mathcal{W}_n$ and $s' \upharpoonright [a'_n, b'_n) \in \mathcal{W}_n$. Since $r_n(s) = r_n(s')$ we know that $a_n = a'_n$ and $b_n = b'_n$. Since s and s' have the same principal subwords, $s \upharpoonright [a_n, b_n) = s' \upharpoonright [a'_n, b'_n)$. The lemma follows. $\dashv$

Remark 16. *We record some consequences of Lemma 15:*

1. *Suppose that we are 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, and the distances of the principle occurrence to the beginning of u_{n+1} go to infinity, then $\langle u_n : M \leq n \rangle$ determines an $s \in S \subseteq \mathbb{K}$ 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 $y = sh(x)$.*

2.5   A note on inverses of symbolic shifts

We define operators we label rev(), and apply them in several contexts

Definition 17. *If x is in $\mathbb{K}$, we define the reverse of x by setting $\mathrm{rev}(x)(k) = x(-k)$. For $A \subseteq \mathbb{K}$, define:*

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

16

If w is a word, we define $\mathrm{rev}(w)$ to be the reverse of w. If we are viewing w as sitting on an interval, we take $\mathrm{rev}(w)$ to sit on the same interval. Similarly, if $\mathcal{W}$ is a collection of words, $\mathrm{rev}(\mathcal{W})$ is the collection of reverses of the words in $\mathcal{W}$.

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

Proposition 18. *The map ϕ sending x to $\mathrm{rev}(x)$ is a canonical isomorphism between $(\mathbb{K}, sh^{-1})$ and $(\mathrm{rev}(\mathbb{K}), sh)$.*

We will use the notation $\mathbb{L}^{-1}$ for the system $(\mathbb{L}, sh^{-1})$ and $\mathrm{rev}(\mathbb{L})$ for the system $(\mathrm{rev}(\mathbb{L}), sh)$.

We can say more. For a fixed symbolic shift $\mathbb{K}$, the canonical isomorphism $\phi : \mathbb{L}^{-1} \to \mathrm{rev}(\mathbb{L})$ gives rise to a canonical correspondence

$$\rho \leftrightarrow \rho'$$

between joinings ρ of $(\mathbb{K}, sh)$ with $(\mathbb{L}, sh^{-1})$ and joinings ρ' of $(\mathbb{K}, sh)$ with $(\mathrm{rev}(\mathbb{L}), sh)$.

We will also use the following remark.

Remark 19. *Assume that there is a unique non-atomic measure on a shift invariant set $S \subseteq \mathbb{K}$. Then there is also a unique non-atomic shift invariant measure on $\mathrm{rev}(S)$ and for this measure, which we denote ν^{-1}, we have $\nu(\langle w \rangle) = \nu^{-1}(\langle \mathrm{rev}(w) \rangle)$.*

2.6   Generic points and sequences

Let T be a measure preserving transformation from (X, μ) to (X, μ), where X is a compact metric space. Let $C(X)$ be the space of all real valued complex functions. Then a point $x \in X$ is *generic* for T if and only if for all $f \in C(X)$,

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

The Ergodic Theorem tells us that for a given f and ergodic T equation above holds for a set of μ-measure one. Intersecting over a countable dense set of f gives a set of μ-measure one of generic points. For symbolic systems $\mathbb{K} \subseteq \Sigma^{\mathbb{Z}}$ we can describe generic points x as being those x such that the

17

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.

The symbolic systems we consider will be built from construction sequences and are characterized by the limiting properties of finite information. We now describe how this works in greater detail. A more complete discussion of this can be found in [21].

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

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

Thus $\hat{\mu}_n(w)$ is the relative measure of $\langle w \rangle$ among all $\langle w' \rangle, w' \in \mathcal{W}_n$. The denominator is a normalizing constant to account for spacers at stages $m > n$ and for shifts of size less than q_n.

Explicitly, if $A_n = \{s \in \mathbb{K} : s(0) \text{ is the start of a word in } \mathcal{W}_n\}$, then the sets $\{sh^j(A_n)\}_{j=0}^{q_n-1}$ are disjoint and their union has a measure that tends to one as n grows to infinity. The set A_n is partitioned into $|\mathcal{W}_n|$ many sets by the words $w \in \mathcal{W}_n$ and $\hat{\mu}_n$ gives their relative size in A_n. Since the measure of an arbitrary finite cylinder set can be calculated along the individual columns represented by a fixed w, it is clear that the $\hat{\mu}_n(w)$ determine uniquely the measure μ.

Using the unique readability of words in $\mathcal{W}_k$ a word w in Σ^{q_k+l} determines a unique sequence of words w_j in $\mathcal{W}_k$ such that ,

$$w = u_0 w_0 u_1 w_1 \ldots w_J u_{J+1}.$$

When $w \in \mathcal{W}_{k+l}$, each u_j is in the region of spacers added in $\mathcal{W}_{k+l'}$, $l' \leq l$. We will denote the *empirical distribution* of $\mathcal{W}_k$-words in w by $\text{EmpDist}_k(w)$. Formally:

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

Then *EmpDist* extends to a measure on $\mathcal{P}(\mathcal{W}_k)$ in the obvious way.

To finitize the idea of a generic point in $\mathbb{K}$ we introduce the notion of a generic sequence of words.

Definition 20. *A sequence $\langle v_n \in \mathcal{W}_n : n \in \mathbb{N} \rangle$ is a* generic sequence of words *if and only if for all k and $\epsilon > 0$ there is an N for all $m, n > N$,*

$$\|EmpDist_k(v_m) - EmpDist_k(v_n)\|_{var} < \epsilon.$$

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

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

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

It follows that if $\langle v_n : n \in \mathbb{N} \rangle$ is a generic sequence of words then it is generic for a unique measure μ. Even though Definition 20 involves only the measures $\hat{\mu}_k$ it is easy to see (using the Ergodic Theorem) that for any $u \in \Sigma^k$, if $\langle v_n : n \in \mathbb{N} \rangle$ is generic then the density of the occurrences of u in the v_n will converge to $\mu(\langle u \rangle)$.

We can summarize the exact relationship between the empirical distributions and the μ_{q_k} by saying that the empirical distribution is the proportion of occurrences of $w' \in \mathcal{W}_k$ among the k-words that appear in v_n, whereas μ_{q_k} is approximately the density of the locations of the start of k-words in v_n. Letting $u \in \mathcal{W}_k$, d be the density of the positions where an occurrence of u begins in v_n, and d_s be the density of locations of letters in some spacer u_i we see that these are related by:

$$d \ = \ \left(\frac{\text{EmpDist}(v_n)(u)}{q_k} \right) (1 - d_s)$$

We record the following consequence of the Ergodic Theorem for future reference:

Proposition 21. *Let $\mathbb{K}$ be an ergodic symbolic system with construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ and measure μ. Then for any generic s the sequence of principal subwords of s, $\langle w_n : n \in \mathbb{N} \rangle$, is generic for μ. In particular, generic sequences for μ exist.*

We will need a characterization of when a generic sequence of words $\langle w_n : n \in \mathbb{N} \rangle$ determines an ergodic measure.

Definition 22. *A sequence $\langle v_n : n \in \mathbb{N} \rangle$ with $v_n \in \mathcal{W}_n$ is an* ergodic sequence *if for any k and $\epsilon > 0$ there are $n_0 > k$, and m_0 such that for all $m \geq m_0$, if*

$$v_m = u_0 w_0 u_1 w_1 u_2 \ldots u_J w_J u_{J+1}$$

is the parsing of v_m into $\mathcal{W}_{n_0}$ words and spacers u_i then there is a subset $I \subseteq \{0, 1, 2 \ldots J\}$ with $|I|/J > 1 - \epsilon$ and for all $j, j' \in I$

$$\|EmpDist_k(w_j) - EmpDist_k(w_{j'})\|_{var} < \epsilon. \tag{3}$$

Notice that in the definition of an ergodic sequence $\langle v_n \rangle$ we are not assuming that it is a generic sequence for a measure. This follow easily (see Lemma 24), but we have not made it part of the definition to emphasize its finitary nature. In the next lemma we use the fact that the language Σ is finite.

Lemma 23. *Any generic sequence $\langle v_n : n \in \mathbb{N} \rangle$ for an ergodic measure μ is an ergodic sequence.*

$\vdash$ Suppose we are given k and $\epsilon > 0$. For all $\delta > 0$ we can apply the Ergodic Theorem to find an N much bigger than q_k and a set B with $\mu(B) > 1 - \delta$ such that for all $s \in B$ and all $w \in \mathcal{W}_k$:

$$\left| \frac{1}{N} \sum_0^{N-1} 1_{\langle w \rangle}(T^i s) - \mu_{q_k}(\langle w \rangle) \right| < \delta.$$

Fix a generic point s for μ. Let $I = \{i \geq 0 : T^i s \in B\}$, and define an infinite sequence of disjoint intervals of length N that cover I by inductively letting $i_0 = min(I)$, and $i_{j+1} = min(\{i \in I : i \geq i_j + N\})$. We take the intervals to be the sequence

$$[i_0, i_0 + N - 1], [i_1, i_1 + N - 1], [i_2, i_2 + N - 1], \ldots$$

Notice that the complement of these intervals in $\mathbb{Z}^+$ has density less than δ since their union clearly covers I.

Though this is an infinite sequence of intervals, the fact our language is finite implies that only finitely many distinct words of length N occur as subwords of s on these intervals. For each such word w^*, the density of those

i in the domain of w^* such that an occurrence of a $w \in \mathcal{W}_k$ starts at i is within δ of $\mu_{q_k}(\langle w \rangle)$.[8]

Next take n_0 large enough that $N/q_{n_0} < \delta$, and parse s into words from $\mathcal{W}_{n_0}$ and the sections of s corresponding to spacers in words in W_j for some $j \geq n_0 + 1$. By taking n_0 large enough we can take the density of locations in s occurring in spacers to be arbitrarily small. Let δ' be this density.

The words from $\mathcal{W}_{n_0}$ have length much larger than N, and we can collect all those words $w \in \mathcal{W}_{n_0}$ that are $(1 - \sqrt{\delta})$-covered by the N-intervals we chose above into a set $A \subseteq \mathcal{W}_{n_0}$.

The proportion of $s \restriction \mathbb{Z}^+$ not covered by words in A can be split into the spacer section and the portion inside words w in $B = \mathcal{W}_{n_0} \setminus A$. For $w \in B$ the complement of the N-intervals has density at least $\sqrt{\delta}$. It follows that the density of sections of s covered by elements of B is less than $\sqrt{\delta}$.

Thus the fraction of s not covered by words in A is at most $\sqrt{\delta} + \delta'$. It is now clear that if δ, δ' are chosen to be sufficiently small then

$$\sum_{w \in A} \hat{\mu}_{n_0}(w) > 1 - \epsilon \tag{4}$$

and all $w \in A$ will have the property that

$$\|\text{EmpDist}_k(w) - \hat{\mu}_k\|_{var} < \epsilon/2$$

which implies inequality 3 for pairs of words in A. Using inequality 4 and the fact that $\langle v_n \rangle$ is generic for μ gives an m_0 so that for all $m \geq m_0$ when v_m is parsed into n_0 words a $(1 - \epsilon)$-fraction will lie in A and this concludes the proof. $\dashv$

We will also need the converse to Lemma 23, namely that the limiting measure defined by an ergodic sequence is, in fact, ergodic.

Lemma 24. *An ergodic sequence is generic and the measure μ defined by an ergodic sequence $\langle v_n : n \in \mathbb{N} \rangle$ is ergodic.*

$\vdash$ Inequality 3 implies that for each k and $w \in \mathcal{W}_k$, the limit of the density of occurrences of w in v_n exists as n goes to infinity. It follows (since $\mathcal{W}_k$ is finite) that $\langle v_n : n \in \mathbb{N} \rangle$ is a generic sequence and hence it defines a unique measure μ.

---

[8]By taking $N \gg q_k$, we can account for negligible "end effects" so that $\left| \frac{1}{N} \sum_0^{N - q_k - 1} 1_{\langle w \rangle}(T^i s) - \mu_{q_k}(\langle w \rangle) \right| < \delta$. We ignore end effects in the rest of the proof.

The ergodicity of μ is equivalent to the fact that the ergodic averages of all L^2 functions converge almost everywhere to a constant. Functions of the form $1_{\langle w \rangle}$ where $w \in \bigcup_n \mathcal{W}_n$ and their shifts linearly span a dense set in L^2 from which it easily follows that if μ were not ergodic there would be some k, and $w \in \mathcal{W}_k$ with $(1/N) \sum_0^{N-1} 1_{\langle w \rangle}(T^i x)$ converging μ-a.e. to a non-constant function. This means that there is a $\gamma > 0$ and disjoint sets B_0, B_1 of positive measure in $\mathbb{K}$ such that for all large enough N for all $x_0 \in B_0, x_1 \in B_1$

$$\left| \frac{1}{N} \sum_0^{N-1} 1_{\langle w \rangle}(T^i x_0) - \frac{1}{N} \sum_0^{N-1} 1_{\langle w \rangle}(T^i x_1) \right| \geq \gamma. \tag{5}$$

Take ϵ small compared to γ and $\mu(B_0), \mu(B_1)$. Find n_0, m_0 as in the definition of ergodic sequence for this k and ϵ. Choose N large enough that inequality 5 holds and so that q_{n_0}/N is negligible. Finally take $m \geq n_0$ so that N/q_m is negligible.

The inequality 5 depends only on the initial $(N + q_k)$-block of x_0 and x_1. Thus for large enough m we can compute $\mu(B_0)$ and $\mu(B_1)$ by the empirical distributions of the $(N + q_k)$-blocks in v_m.

Since N is large compared to q_{n_0} the frequency of occurrence of w in a block of length $N + q_k$ is determined by its frequencies in the words in $\mathcal{W}_{n_0}$ in the n_0-parsing of v_m. We now get a contradiction to inequality 5, since except for an ϵ-fraction, these w_{n_0}-words have their k-words distributed very close to $\hat{\mu}_k(w)$. $\qquad \dashv$

If S and T are symbolic systems then a joining ρ of S and T will be a symbolic system, but may not have well-defined construction sequence, even if S and T do.[9] Accordingly we must generalize our definition of *empirical distribution* to take into account the relative locations of words in typical $(s, t) \in \mathbb{K} \times \mathbb{L}$. We express this by shifting one of the basic open sets and considering words $(w, sh^s(v))$, which we view as starting at the locations $(0, s)$.

Let $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ and $\langle \mathcal{V}_n : n \in \mathbb{N} \rangle$ be uniquely readable construction sequences for $\mathbb{K}$ and $\mathbb{L}$ in the languages Σ, Λ respectively. Assume for simplicity that all words in $\mathcal{W}_n$ and $\mathcal{V}_n$ have the same length.

---

[9]We run into this problem when considering joinings of circular systems and their inverses that project to the $\natural$-map on the canonical factors; these notions are defined in future sections.

Let $n \leq n' < n + l$. Then we can uniquely parse a word $w \in \mathcal{W}_{n+l}$ as

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

where each $w_j \in \mathcal{W}_n$ and each u_j is in the region of spacers for words in $\mathcal{W}_{n+l'}$, $l' < l$. The similar statement holds for $v'_k \in \mathcal{V}_{n'}$, and $v \in \mathcal{V}_{n+l}$:

$$v = u'_0 v'_0 u'_1 v'_1 \ldots v'_K u'_{K+1}.$$

The definition must take into account the relative shifts of w and v, the shifts of (w_j, v_k) allow spacers to occur in different places and for the possibility that $J \neq K$.

Let $n \leq n' < n + l$ be natural numbers, $s, s' \in \mathbb{Z}$, and $(w', v') \in \mathcal{W}_n \times \mathcal{V}_{n'}$ and $(w, v) \in \mathcal{W}_{n+l} \times \mathcal{V}_{n+l}$. Write w and v in terms of n and n'-words as above. For s, s', define an *occurrence* of $(w', sh^{s'}(v'))$ in $(w, sh^s(v))$ to be a $j \leq J$ such that $w_j = w'$ and if k is the location of w_j in w, then v' occurs at $k + s'$ in $sh^s(v)$. We note the bijection between occurrences of $(w', sh^{s'}(v'))$ in $(w, sh^s(v))$ and occurrences of $(v', sh^{-s'}(w'))$ in $(v, sh^{-s}(w))$.

In defining empirical distributions for joinings we generalize Definition 20. The empirical distribution of a shifted pair is defined to be the proportion of times it occurs, relative to the proportion of times arbitrary pairs with the same shift occur.

Definition 25. *Fix w, s, v Let A be the collection*

$$\{j : \text{for some } (w^*, v^*) \in \mathcal{W}_n \times \mathcal{V}_{n'}, (w^*, sh^{s'}(v^*)) \text{ occurs at } j \text{ in } (w, sh^s(v))\}.$$

Assume that $A \neq \emptyset$. For $w' \in \mathcal{W}_n$ and $v' \in \mathcal{V}_{n'}$, we define:

$$EmpDist_{n,n',s'}(w, sh^s(v))(w', v') = \frac{|\{0 \leq j \leq J : (w', sh^{s'}(v')) \text{ occurs at } j\}|}{|A|}.$$

As before, $EmpDist_{n,n',s'}(w, sh^s(v))$ extends uniquely to a probability measure on $\mathcal{P}(\mathcal{W}_n \times \mathcal{V}_{n'})$. Definition 25 facilitates a notion of a generic sequence for a joining.

Definition 26. *A sequence of $\langle (w_n, v_n, s_n) \in \mathcal{W}_n \times \mathcal{V}_n \times \mathbb{Z} : n \in \mathbb{N} \rangle$ is called generic iff*

1. $\sum \frac{|s_n|}{q_n} < \infty$ *and*

2. for all n, n', s' and $\epsilon > 0$ there is an N for all $m, m' > N$,

$$\|EmpDist_{n,n',s'}(w_m, sh^{s_m}(v_n)) - EmpDist_{n,n',s'}(w_{m'}, sh^{s_{m'}}(v_{m'}))\|_{var} < \epsilon.$$

The definition of an ergodic sequence of pairs *is done analogously.*

It is easy to check that $\langle (w_n, v_n, s_n) : n \in \mathbb{N} \rangle$ is generic/ergodic if and only if $\langle (v_n, w_n, -s_n) : n \in \mathbb{N} \rangle$ is generic/ergodic. For ergodic joinings the analogues of Proposition 21, and Lemmas 23 and 24 hold and are proved in exactly the same way.

We have given these definitions in the case of a product of two symbolic shifts, but they generalize immediately to products of three or more shifts. For example, to consider three shifts with construction sequences $\langle \mathcal{U}_n \rangle_n, \langle \mathcal{V}_n \rangle_n, \langle \mathcal{W}_n \rangle_n$, we would consider a sequence of the form:

$$\langle (u_n, v_n, w_n, s_n, t_n) : n \in \mathbb{N} \rangle,$$

where the words belong to the respective construction sequences and the s_n's and t_n's give the shifts relative to the first coordinate.

We will be concerned with compositions of joinings, which involves products of three shifts. To prepare for this we need the notion of a conditional empirical distribution.

Definition 27. *Let $n, n' < n+l$. Given a fixed $w^* \in \mathcal{W}_{n'}$ and a pair $(w, v) \in \mathcal{W}_{n+l} \times \mathcal{V}_{n+l}$ and (s, s') we define the* conditional empirical distribution *to be:*

$$EmpDist_{n,s'}((w, sh^s(v)|w^*)(v') =$$

$$\frac{|\{0 \le j \le J : (w^*, sh^{s'}(v')) \ occurs \ at \ j\}|}{|\{j \le J : for \ some \ v^* \in \mathcal{V}_n, (w^*, sh^{s'}(v^*)) \ occurs \ at \ j\}|}$$

for $v' \in \mathcal{W}_n$.

Using the same ideas we can define the empirical distribution conditioned on a $v^* \in \mathcal{V}_k$ by looking at $(sh^{-s}(w), v)$ and counting occurrences of $(sh^{-s'}(w'), v^*)$ for the $w' \in \mathcal{W}_k$.

This definition generalizes to products of three or more systems. When working in three or more systems, there will be multiple s's playing the role of s' in Definition 27. They will refer to the position of the sequences being counted, *relative to the conditioning sequence.* So for example, if $\mathbb{K}, \mathbb{L}, \mathbb{M}$ have

construction sequences $\langle \mathcal{U}_n \rangle_n, \langle \mathcal{V}_n \rangle_n, \langle \mathcal{W}_n \rangle_n$ and $\langle (u_n, v_n, w_n, s_n, t_n) : n \in \mathbb{N} \rangle$ is a generic sequence for a joining ρ of $\mathbb{K}, \mathbb{L}$ and $\mathbb{M}$, then

$$EmpDists_{k,k',s,s'}(u_n, sh^{s_n}(v_n), sh^{t_n}(w_n)|v)$$

counts pairs $(sh^s(u), sh^{s'}(w))$, where $(u, w) \in \mathcal{U}_k \times \mathcal{W}_{k'}$ have been shifted by s and s' *relative to* v.

Let ρ_1 be a $T_1 \times T_2$-invariant measure on $X \times Y$ and ρ_2 a $T_2 \times T_3$-invariant measure on $Y \times Z$. Recall from Section 2.2 that the composition of ρ_1 and ρ_2 is defined to be projection of the relative independent joining of ρ_1 and ρ_2 over the common factor Y to a measure on $X \times Z$. We now describe a method for detecting generic sequences for relatively independent joinings.

Suppose that systems X and Z have a common factor Y.

$$(X, \mathcal{B}, \mu, T) \qquad\qquad (Z, \mathcal{D}, \tilde{\mu}, \tilde{T})$$

$$(Y, \mathcal{C}, \nu, S)$$
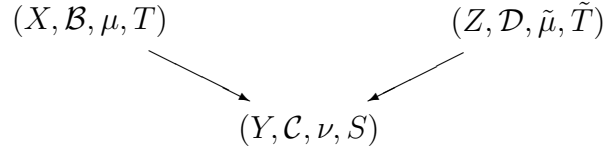
Let $\rho = X \times_Y Z$ be the relatively independent joining of X and Y. Let $\mu_y, \tilde{\mu}_y, \rho_y$ be the distintegrations of $\mu, \tilde{\mu}$ and ρ respectively. Then the relatively independent joining ρ is characterized by the fact that for ν-a.e y,

$$\rho_y = \mu_y \times \tilde{\mu}_y. \tag{6}$$

Let $\langle \mathcal{A}_n, \tilde{\mathcal{A}}_n, \mathcal{A}'_n : n \in \mathbb{N} \rangle$ be sequences of refining partitions that generate $\mathcal{B}, \mathcal{D}$ and $\mathcal{C}$ respectively. Since the sequence of partitions $\mathcal{A}_n \times \tilde{\mathcal{A}}'_n$ generates $\mathcal{B} \otimes \mathcal{D}$, equation 6 is equivalent to the property that for all $A_k \in \mathcal{A}_k, \tilde{A}_k \in \tilde{\mathcal{A}}_k$ and ν-a.e. y,

$$\mu_y(A_k) \times \tilde{\mu}_y(\tilde{A}_k) = \rho_y(A_k \times \tilde{A}_k) \tag{7}$$

To finitize this we approximate $\mu_y(A_k)$ by $\mu(A_k | A'_m(y))$ for large m, where $A'_m(y)$ is the atom of $\mathcal{A}'_m$ to which y belongs. We let $\mu_y(\mathcal{A}_k)$ be shorthand for the distribution $\langle \mu_y(A_k) : A_k \in \mathcal{A}_k \rangle$, and $\mu(\mathcal{A}_k | \mathcal{A}'_m)(y)$ stands for the conditional distribution $\mu(A_k | A'_m(y)), A_k \in \mathcal{A}_k$. (We use similar notation in Lemma 28 for the conditional distribution given by ρ, μ and $\tilde{\mu}$ on various partitions.)

By Martingale convergence,[10] for $\epsilon > 0$ and fixed k if m sufficiently large, then for $(1 - \epsilon)$ proportion of the y' in the same atom as y:

$$\|\mu_{y'}(\mathcal{A}_k) - \mu(\mathcal{A}_k|\mathcal{A}'_m)(y)\|_{var} < \epsilon$$

but for a collection of A'_m of whose union has ν-measure less than ϵ.

One can deal similarly with $\tilde{\mu}_n$ and ρ_y. We have shown:

Lemma 28. *In the notation above, ρ is the relatively independent joining of μ and $\tilde{\mu}$ if and only if for all $k, \epsilon > 0$, for all large enough m, there is a collection of atoms $A_m \in \mathcal{A}'_m$ of total measure at least $1 - \epsilon$ for which:*

$$\|\rho(\mathcal{A}_k \times \tilde{\mathcal{A}}_k|A_m) - \mu(\mathcal{A}_k|A_m) \times \tilde{\mu}(\tilde{\mathcal{A}}_k|A_m)\|_{var} < \epsilon. \tag{8}$$

We now express Lemma 28 in terms of sequences of finite words. Suppose that $\langle \mathcal{U}_n \rangle, \langle \mathcal{V}_n \rangle$, and $\langle \mathcal{W}_n \rangle$ are the uniquely readable construction sequences for X, Y and Z.

Proposition 29. *Let $\langle (u_n, v_n, w_n, s_n, t_n) \in \mathcal{U}_n \times \mathcal{V}_n \times \mathcal{W}_n \times \mathbb{Z}^2 : n \in \mathbb{N} \rangle$ be a sequence of words. Suppose that:*

1. *$\langle (u_n, v_n, s_n) \rangle_n$ is generic for ρ_1,*

2. *$\langle (v_n, w_n, t_n) \rangle_n$ is generic for ρ_2.*

3. *for all $\epsilon > 0, k$ and s^* for all sufficiently large k' there is an N and a set $G_{k'} \subset \mathcal{V}_{k'}$ and for each $v \in G_{k'}$ a set of indices $I_v \subseteq [0, q_{k'})$ that satisfies $|I_v| > (1 - \epsilon)q_{k'}$ such that for all $n > N$:*

   (a) *$\sum_{v \in G_{k'}} EmpDist(v_n)(v) > 1 - \epsilon$*
   and

   (b) *for all $v \in G_{k'}$ and $s \in I_v$,*

   $\|EmpDist_{k,k,s,s+s^*}(u_n, sh^{s_n}(v_n), sh^{t_n}(w_n)|v) \quad -$
   $EmpDist_{k,s}(u_n, sh^{s_n}(v_n)|v) * EmpDist_{k,s+s^*}(v_n, sh^{t_n-s_n}(w_n)|v)\|_{var}$

   is less than ϵ.

If ρ is the relatively independent joining of ρ_1, ρ_2, then $\langle (u_n, v_n, w_n, s_n, t_n) : n \in \mathbb{N} \rangle$ is a generic sequence for ρ.

---

[10]See (e.g.) [11], Theorem 14.26, page 261.

⊢   Observe that the hypothesis 3b implies a similar equation for any $k_1 < k$ while the other parameters are fixed. Now use hypothesis 3a with a summable sequence of ϵ's and we can conclude by the Borel-Cantelli lemma that for ν-almost every $y \in Y$ for k' sufficiently large, if $v_{k'}(y)$ is the principal k'-block of y with location $r_{k'}$, then the inequality in 3b will hold for $s = r_{k'}$ and $v = v_{k'}(y)$.

Now by hypotheses 1 and 2, the single empirical distributions are converging to $(\rho_1)_y$ and $(\rho_2)_y$ respectively (where $(\rho_i)_y$ is the disintegration of ρ_i over y).

It then follows by integration that the sequence of $(u_n, v_n, w_n, s_n, t_n)$'s is generic for a measure ρ on $X \times Y \times Z$, which is the relatively independent joining.   ⊣

Remark 30. *It follows immediately from hypothesis 3 of Proposition 29 that if we are given a finite set F of natural numbers then for all sufficiently large k' we can find an N, $G_{k'}$ and I_v as in hypothesis 3 so that (a) and (b) hold simultaneously for all $s^* \in F$.*

An immediate corollary of this is:

Corollary 31. *Suppose that $\langle (u_n, v_n, w_n, s_n, t_n) : n \in \mathbb{N} \rangle$ satisfies the hypotheses of Proposition 29. Then $\langle (u_n, sh^{t_n}(w_n)) : n \in \mathbb{N} \rangle$ is generic for $\rho_1 \circ \rho_2$.*

There is a converse to Proposition 29, namely that a generic sequence for the relatively independent joining of two odometer based system satisfies the conditions 1-3 of the Proposition. The first two are immediate while the third simply expresses the fact that the generic sequence sequence is actually representing the relatively independent joining. For later use we record this as:

Lemma 32. *Given joinings ρ_1 of $X \times Y$ and ρ_2 of $Y \times Z$ if $\langle (u_n, v_n, w_n, s_n, t_n) : n \in \mathbb{N} \rangle$ is generic for the relatively independent joining ρ then it satisfies the hypotheses of Proposition 29.*

2.7   Unitary Operators

We will use spectral tools introduced by Koopman and studied by Halmos and von Neumann. We reprise the basic facts we will use. Readers unfamiliar

with this material can find it in [20] or [11]. Let $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ be measure preserving systems.

If $T : X \to Y$ is a measure preserving transformation then T induces a unitary isometry $U_T : L^2(Y) \to L^2(X)$ by setting

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

If T is an isomorphism then U_T is invertible. Moreover if $U : L^2(Y) \to L^2(X)$ is multiplicative on bounded functions then there is a measure preserving transformation $T : X \to Y$ such that $U = U_T$.

If $\pi : X \to Y$ is a factor map, then the map $f \mapsto f \circ \pi$ gives an injection of $L^2(Y)$ into $L^2(X)$, whose range is a closed U_T invariant subspace. Conversely if $M \subseteq L^2(X)$ is a closed U_T invariant subspace containing 1 that is closed under taking complex conjugates, truncation and multiplication by elements of $M \cap L^\infty(X)$, then there is a factor $Y \subseteq X$ such that $M = L^2(Y)$.

For the rest of this discussion assume that T is ergodic. Then the eigenvalues of U_T all have multiplicity one and form a subgroup $G_T \subseteq \mathbb{T}$. The group G_T is an isomorphism invariant.

The collection of eigenfunctions generate a closed subspace of $L^2(X)$ corresponding to a factor K of X. This factor is called the *Kronecker factor*. If H is any subgroup of G_T then there is a further factor K_H of K that is canonically determined by the eigenfunctions coming from eigenvalues in H.

Assume that ϕ is an isomorphism from (X, T) to (Y, S). Then $G_T = G_S$ and if K_H^X, K_H^Y are the factors of X and Y determined by $H \subseteq G_T$ then U_ϕ determines an unique isomorphism between K_H^X and K_H^Y.

It follows from this that if $\alpha \in \mathbb{T}$ is an eigenvalue of U_T then there are factors of X and Y isomorphic to rotation $\mathcal{R}_\alpha$ of $\mathbb{T}$ by α. Moreover there is a unique isomorphism $U_\phi^\pi : (\mathbb{T}, \mathcal{B}, \lambda, \mathcal{R}_\alpha) \to (\mathbb{T}, \mathcal{B}, \lambda, \mathcal{R}_\alpha)$ that intertwines U_ϕ and the projection maps of X and Y to $(\mathbb{T}, \mathcal{B}, \lambda, \mathcal{R}_\alpha)$.

The analogous statement holds for odometers. If G_T consists of finite order eigenvalues and $\mathcal{O}$ is the corresponding odometer transformation, then there is a unique isomorphism $U_\phi^\pi : \mathcal{O} \to \mathcal{O}$ that intertwines U_ϕ and the projection maps of X and Y to $\mathcal{O}$.

2.8   Stationary Codes and $\bar{d}$-Distance

In this section we briefly describe a standard idea, that of a *stationary code* that we will use to understand the existence of factor maps and isomorphisms.

We review some standard facts here. A reader unfamiliar with this material who wants to see proofs should see [17].

Definition 33. *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 and an $s \in \Sigma^{\mathbb{Z}}$ we define the stationary code *determined by Λ to be $\bar{\Lambda}(s)$ where:*

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

Let $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, sh)$ be a symbolic system. Suppose we have two codes Λ_0 and Λ_1 that are not necessarily of the same length. Define $D = \{s \in \Sigma^{\mathbb{Z}} : \bar{\Lambda}_0(s)(0) \neq \bar{\Lambda}_1(s)(0)\}$ and $d(\Lambda_0, \Lambda_1) = \nu(D)$. Then d is a semi-metric on the collection of codes. The following is a consequence of the Borel-Cantelli lemma.

Lemma 34. *Let 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, $\lim_{i\to\infty} \bar{\Lambda}_i(s) = S(s)$*

A shift invariant Borel map $S : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$, determines a factor $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, sh)$ of $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, sh)$ by setting $\mu = S^*\nu$ (i.e. $\mu(A) = \nu \circ S^{-1}(A)$). Hence a convergent sequence of stationary codes determines a factor of $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, sh)$.

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

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

More generally we can 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

$$\bar{d}(x, y) = \overline{\lim}_{N\to\infty} \bar{d}_{[-N,N]}(x \restriction [-N, N], y \restriction [-N, N]),$$

provided this limit exists.

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

Lemma 35. *Suppose that $(\Sigma^{\mathbb{Z}}, sh, \nu)$ is ergodic and that Λ_0 and Λ_1 be codes. Then for almost all $s \in S$:*

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

We finish with a useful remark:

Remark 36. *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 $\bar{d}_I(w_1, w_2) \geq \delta \bar{d}_J(w_1, w_2)$.*

3 Odometer based and Circular Symbolic Systems

Two types of symbolic shifts play central roles for the proofs of our main theorem. We dub them *odometer based* and *circular* systems. In this section we give some general facts about symbolic systems with uniquely readable construction sequences, define odometer and circular systems, and show that every circular system has a canonical rotation factor.

3.1 Odometer Based Systems

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. We make O into 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 $K_n = k_0 * k_1 * k_2 \ldots k_{n-1}$.

The following results are standard:

Lemma 37. *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 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 $0 \leq a_j < k_j$.

Lemma 38. *Suppose that $\langle r_n : n \in \mathbb{N} \rangle$ is a sequence of natural numbers with $0 \leq r_n < k_0 k_1 \ldots k_{n-1}$ and $r_n \equiv r_{n+1} \mod (K_n)$. Then there is a unique element $x \in O$ such that $r_n = x(0) + x(1)k_0 + \ldots x(n)(k_0 k_1 \ldots k_{n-1})$ 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 39. *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 2 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 $\mathbb{K}$, for all $s \in \mathbb{K}$ and for all $n \in \mathbb{N}$, $r_n(s)$ exists.

Proposition 40. *Let $\mathbb{K}$ be an odometer based system and suppose that ν is a shift invariant measure. Then ν concentrates on S.*

$\vdash$ Let $B = \mathbb{K} \setminus S$. Then B is shift invariant. Suppose that ν gives B positive measure. For $s \in B$ let $a_n(s) \leq 0 \leq b_n(s)$ be the left and right endpoints of the principal n-block of s. Then for all $s \in B$ there is an $N \in \mathbb{N}$ such that:

1. for all $n, -N \leq a_n$ or

2. for all $n, b_n \leq N$.

We assume that ν gives the collection B^* of s such that there is an $N \in \mathbb{N}$ for all $n, -N \leq a_n$ positive measure, the other case is similar.

Define $f : B^* \to \mathbb{N}$ by setting $f(s) = $ least N satisfying item 1. Then f is a Borel function. Let $B_n = f^{-1}(n)$. Then the B_n's are disjoint, $B^* = \bigcup_{n \geq 0} B_n$ and $sh^{-1}(B_n) = B_{n+1}$. Hence for all $n, m, \nu(B_n) = \nu(B_m)$, a contradiction. $\dashv$

The next lemma justifies our terminology.

Lemma 41. *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 $\langle k_n : n \in \mathbb{N} \rangle$.

$\vdash$ For each $s \in S$, we know that for all $n, r_n(s)$ is defined and both r_n and $k_n - r_n$ go to infinity. By Lemma 38, 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. $\dashv$

In the forthcoming paper [8] we show a strong converse to this result: if T has finite entropy and an odometer factor then T can be presented by an odometer based system.

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.

We will use the following remark about the canonical factor of the inverse of an odometer based system.

Remark 42. *If $\pi : \mathbb{L} \to \mathcal{O}$ is the canonical factor map, then the function $\pi : \mathbb{L} \to O$ is also factor map from $(\mathbb{L}, sh^{-1})$ 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 18, then Lemma 37 tells us that the projection of ϕ to a map $\phi^\pi : \mathcal{O} \to \mathcal{O}$ is given by $x \mapsto -x$.*

From this remark we immediately see:

Lemma 43. *Let $\rho \leftrightarrow \rho'$ be the canonical correspondence between joinings of $(\mathbb{K}, sh)$ and $(\mathbb{L}, sh^{-1})$ and joinings of $(\mathbb{K}, sh)$ and $(\mathrm{rev}(\mathbb{L}), sh)$ given after Proposition 18. Then the joining ρ concentrates on the set of pairs (s, t) such that $\pi^{\mathbb{K}}(t) = -\pi^{\mathbb{L}}(s)$ if and only if ρ' concentrates on the collection of (s, t) such that $\pi^{\mathbb{K}}(s) = \pi^{\mathbb{L}^{-1}}(t)$.*

3.2 Circular systems

We now define and discuss circular systems. The paper [5] showed that the circular systems give symbolic characterizations of the smooth diffeomorphisms defined by the Anosov-Katok method of conjugacies. The construction sequences of circular systems have quite specific combinatorial properties that will be important to our understanding of the Anosov-Katok systems and their centralizers in the third paper in this series.

We call these systems *circular* because they are closely tied 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 along a sequence determined by some numbers $j_i =_{def} p^{-1}i \pmod q$: the interval $[i/q, (i+1)/q)$ is the j_i^{th} interval in the sequence.[11] The operation $\mathcal{C}$ which we are about to describe models the relationship between rotations by p/q and p'/q' when q' is very close to q.

Let k, l, p, q be positive natural numbers with $p < q$ relatively prime. Set

$$j_i \equiv_q (p)^{-1}i \tag{9}$$

with $j_i < q$. It is easy to verify that:

$$q - j_i = j_{q-i} \tag{10}$$

Let Σ be a non-empty set. We define an operation $\mathcal{C}$, which depends on p, q, an integer $l > 1$, and on sequences $w_0, \ldots w_{k-1}$ of words in a language $\Sigma \cup \{b, e\}$ by setting:[12]

$$\mathcal{C}(w_0, w_1, w_2, \ldots w_{k-1}) = \prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (b^{q-j_i} w_j^{l-1} e^{j_i}). \tag{11}$$

---

[11] We assume that p and q are relatively prime and the exponent -1 is the multiplicative inverse of p mod q.

[12] We use $\prod$ for repeated concatenation of words.

To start our construction we frequently take $p_0 = 0$ and $q_0 = 1$. In this case we adopt the convention that $j_0 = 0$. Hence

$$\begin{aligned} \mathcal{C}(w_0, w_1, \ldots w_{k-1}) &= \prod_{j<k} b^q w_j^{l-1} \\ &= \prod_{j<k} bw^{l-1}. \end{aligned}$$

Remark 44. *We remark:*

- *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.*

- *Every occurrence of an e in $\mathcal{C}(w_0, \ldots w_{k-1})$ has an occurrence of a b to the left of it. If $p \neq 0$ then every occurrence of a b has an e to the right of it.*

- *Suppose that $n < m$ and b occurs at position n in $\mathcal{C}(w_0, w_1, \ldots w_{k-1})$ 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.*

The $\mathcal{C}$ operator automatically creates uniquely readable words, as the next lemma shows, however we will need a stronger unique readability assumption for our definition of circular systems.

Lemma 45. *Suppose that Σ is a language, $b, e \notin \Sigma$, $0 < p < q$ and that $u_0, \ldots u_{k-1}, v_0, \ldots v_{k-1}$ and $w_0 \ldots w_{k-1}$, are words in the language $\Sigma \cup \{b, e\}$ of some fixed length $q < l/2$. Let*

$$\begin{aligned} u &= \mathcal{C}(u_0, u_1, \ldots u_{k-1}) \\ v &= \mathcal{C}(v_0, v_1, \ldots v_{k-1}) \\ w &= \mathcal{C}(w_0, w_1, \ldots w_{k-1}). \end{aligned}$$

Suppose that uv is written as pws where p and s are words in $\Sigma \cup \{b, e\}$. Then either p is the empty word and $u = w, v = s$ or s is the empty word and $u = p, v = w$.

⊢ The map $i \mapsto j_i$ is one-to-one. Hence each location in the word of length klq^2 is uniquely determined by the lengths of nearby sequences of b's and e's. ⊣

In fact something stronger is true: if $\sigma \in \Sigma$ occurs at place m in w then m is uniquely determined by the knowing the $w_0, w_1, \ldots w_{k-1}$ and the $kq^l/2 + 1$ letters on either side of σ.

We now describe how to use the $\mathcal{C}$ operation to build a collection of symbolic shifts. Our 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 [14].

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

$$q_{n+1} = k_n l_n q_n{}^2 \tag{12}$$

(thus $q_1 = k_0 l_0$) and take

$$p_{n+1} = p_n q_n k_n l_n + 1. \tag{13}$$

Then clearly p_{n+1} is relatively prime to q_{n+1}.[13]

Definition 46. *A sequence of integers $\langle k_n, l_n : n \in \mathbb{N} \rangle \rangle$ 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. We will construct the systems we study by building collections of words $\mathcal{W}_n$ in the alphabet $\Sigma \cup \{b, e\}$ by induction as follows:

- Fix a circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N} \rangle \rangle$.

- Set $\mathcal{W}_0 = \Sigma$.

- Having built $\mathcal{W}_n$ we 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 $(w_0, \ldots w_{k_n-1}) \in P_{n+1}$.[14]

---

[13]p_n and q_n being relatively prime for $n \geq 1$, allows us to define the integer j_i in equation 9. For $q_0 = 1$, $\mathbb{Z}/q_0\mathbb{Z}$ has one element, $[0]$, so we set $p_0{}^{-1} = p_0 = 0$.

[14]Passing from $\mathcal{W}_n$ to $\mathcal{W}_{n+1}$ we 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 44, the length of each of the words in $\mathcal{W}_{n+1}$ is q_{n+1}.

We will call the elements of P_{n+1} *prewords*.

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. We assume that in the alphabet Λ_n, each P_{n+1} is uniquely readable.

Definition 47. *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.*

It follows from Lemma 45 that each $\mathcal{W}_n$ in a circular construction sequence is uniquely readable.

Definition 48. *A symbolic shift $\mathbb{K}$ built from a circular construction sequence will be called a* circular system.

For emphasis we will often write circular construction sequences as $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$ and the associated circular shift $\mathbb{K}^c$. We sometimes write w^c to emphasize that a word is a circular word.

We will need to analyze the words constructed by $\mathcal{C}$ in detail. We start by describing the boundary and interior portions of the words.

Definition 49. *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 b's and e's that are not in the w_i's. The* interior *of w is the portion of w in the w_i's. 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 b's and e's, not all b's and e's 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 50. *The proportion of the word w written in equation 11 that belongs to its boundary is $1/l$. Moreover the proportion of the word that is within q letters of boundary of w is $3/l$.*

The next lemma was proved in [5] (Lemma 20).

Lemma 51. *Let $\mathbb{K}^c$ be a circular system and 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.*

Remark 52. *Let $\mathbb{K}^c$ be a circular system.*

1. *There are only two invariant atomic measures, one concentrates on the constant "b" sequence, the other on the constant "e" sequence.*

2. *for $\mathbb{K}^c$, Lemma 12 can be strengthened to say that for all $s \in S$ for all large enough n, the principal n-block of s exists.*

3. *The symbolic shift $\mathbb{K}^c$ has zero topological entropy.*

$\vdash$ A direct inspection reveals that the only periodic points in $\mathbb{K}^c$ are the two fixed points constant "b" and "e".

The second item follows because if s has a principal n-block at $[a_n, b_n)$ then it has a principal $n + 1$-block at some $[a_{n+1}, a_{n+1} + q_{n+1})$ for an a_{n+1} with $|a_{n+1}| \leq |a_n| + (q_{n+1} - q_n)$.

The fact that the topological entropy of $\mathbb{K}^c$ is zero follows easily from the fact that the l_n tend to infinity.

3.3 The structure of the words

The words used to form circular transformations have quite specific combinatorial properties. We begin with an important definition for our understanding of rotations; the three *subscales* at stage $n + 1$. Fix a sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$ defining a circular system. Using equation 11 we define the *subscales* of a word $w^* \in \mathcal{W}_{n+1}$:

> **Subscale 0** is the scale of the individual powers of $w_j \in \mathcal{W}_n^c$ of the form w_j^{l-1}; we call each such occurrence of a w_j^{l-1} a *0-subsection*

Subscale 1 is 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 is 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*.

<div align="center">

Summary

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}

</div>

By contrast we will discuss *n-subwords* of a word w. These will be subwords that lie in $\mathcal{W}_n^c$, the nth stage of the construction sequence. We will use *n-block* to mean the location of the *n*-subword.

3.4 The canonical circle factor $\mathcal{K}$

We now define a canonical factor $\mathcal{K}$ of a circular system and show that this factor is isomorphic to a rotation of the circle by α, where α is the limit of $\alpha_n = \frac{p_n}{q_n}$ as n goes to infinity.

Definition 53. *Let $\langle k_n, l_n : n \in \mathbb{N} \rangle\rangle$ be a circular coefficient sequence. Let $\Sigma_0 = \{*\}$. We define a circular construction sequence such that each $\mathcal{W}_n^c$ has a unique element as follows:*

1. *$\mathcal{W}_0 = \{*\}$ and*

2. *If $\mathcal{W}_n^c = \{w_n\}$ then $\mathcal{W}_{n+1}^c = \{\mathcal{C}(w_n, w_n, \ldots w_n)\}$.*

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

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

Let $\mathbb{K}^c$ be an arbitrary circular system with coefficients $\langle k_n, l_n \rangle$. 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 $\mathbb{K}^c$, we define the following function:

$$\pi(x)(i) = \begin{cases} x(i) & \text{if } x(i) \in \{b, e\} \\ * & \text{otherwise} \end{cases} \tag{14}$$

We record the following easy lemma that justifies the terminology of Definition 53:

Lemma 54. *Let π be defined by equation 14. Then:*

1. *$\pi : \mathbb{K}^c \to \mathcal{K}$ is a Lipshitz map,*

2. *$\pi(sh^{\pm 1}(x)) = sh^{\pm 1}(\pi(x))$ and thus*

3. *π is a factor map of $\mathbb{K}^c$ to $\mathcal{K}$ and $(\mathbb{K}^c)^{-1}$ to $\mathcal{K}^{-1}$*

A variant of item 3 is also true: π can be interpreted as a function from $\mathrm{rev}(\mathbb{K}^c)$ to $\mathrm{rev}(\mathcal{K})$. With this interpretation π is also a factor map. We will call $\mathcal{K}$ the *circle factor* of any circular system with construction coefficients $\langle k_n, l_n : n \in \mathbb{N} \rangle$.

Fix a circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N} \rangle$, and let $\mathcal{K}$ and $\langle \mathcal{W}_n^\alpha : n \in \mathbb{N} \rangle$ be given in definition 53. Let $\alpha_n = p_n/q_n$ and $\alpha = \lim \alpha_n$.

If $s \in S$, from $r_n(s)$ we can determine the locations of the beginnings and ends of the words w_n^α that contain $s(0)$. Since $|\mathcal{W}_n^\alpha| = 1$ for all n, for all $s \in S$ the sequence $\langle r_n(s) : n \in \mathbb{N} \rangle$ uniquely determines s.

Theorem 55. *Let ν be the unique non-atomic shift invariant measure on $\mathcal{K}$. Then*

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

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

$\vdash$ A more involved geometric proof of this fact is given in [5]. Here present a simple algebraic proof. As usual we identify the unit circle S^1 with $[0, 1)$ and use additive notation for the group operations.

By Lemma 12, the collection S' of $s \in S$ such that for all large enough n, the principal n-block of s exists, has measure one. We define a map $\phi_0 : S' \to [0, 1)$ by a limiting process. For s such that $r_n(s)$ exists, we let

$$\rho_n(s) = \frac{p}{q_n}$$

iff

$$p \equiv p_n r_n(s) \mod q_n$$

39

Claim 56. *If r_n is defined, then $|\rho_{n+1}(s) - \rho_n(s)| < 2/q_n$.*

⊢   From equation 11, we see that the position of $s(0)$ in an $n+1$-block is determined by the parameters $i \in [0, q_n - 1), j \in [0, k_n - 1), l^* \in [0, l - 1]$ and r_n, which determine its location among the 2-subsections, 1-subsections, 0-subsections and inside the n-words w_n respectively. Explicitly:

$$r_{n+1}(s) = i(k_n l_n q_n) + j(l_n q_n) + (q_n - j_i) + l^* q_n + r_n(s),$$

where $r_n(s)$ is the position of $s(0)$ in its principal w_n-word.

From the definition of ρ_{n+1}, and working *mod 1*:

$$
\begin{aligned}
\rho_{n+1} &= r_{n+1}(s)\left(\frac{p_{n+1}}{q_{n+1}}\right) \\
&= r_{n+1}(s)\left(\frac{p_n}{q_n} + \frac{1}{q_{n+1}}\right)
\end{aligned}
$$

Expanding this, using our formula for $r_{n+1}(s)$ and the fact that all but two terms of $r_{n+1}(s)$ are divisible by q_n, we get:

$$\rho_{n+1} = \left(-j_i\left(\frac{p_n}{q_n}\right) + r_n(s)\left(\frac{p_n}{q_n}\right)\right) + \left(\frac{i}{q_n} + \delta\right) \tag{15}$$

where

$$\delta = \frac{j}{k_n q_n} + \frac{1}{k_n l_n q_n} + \frac{l^*}{k_n l_n q_n} + \frac{r_n(s) - j_i}{k_n l_n q_n^2}.$$

The first and third terms of equation 15 cancel, thus:

$$\rho_{n+1} = \rho_n + \delta.$$

Since $\delta < 2/q_n$, the claim follows. ⊣

Since the sequence $1/q_n$ is summable, for almost all $s, \langle \rho_n(s) : n \in \omega \rangle$ is Cauchy. We define

$$\phi_0(s) = \lim_n \rho_n(s).$$

It is easy to check that ϕ_0 is one-to-one. By the unique ergodicity of the rotation $\mathcal{R}_\alpha$, Theorem 55 will be proved when we establish:

40

Claim 57. *The map $\phi_0 : S \to [0,1)$ satisfies:*

$$\phi_0(sh(s)) = \mathcal{R}_\alpha(\phi_0(s)).$$

In particular, if ν is the unique invariant measure on S

$$(\mathcal{K}, \mathcal{C}, \nu, sh) \cong ([0,1), \mathcal{B}, \lambda, \mathcal{R}_\alpha).$$

⊢ Suppose that $r_n(s)$ and $r_n(sh(s))$ both exist. Then $r_n(sh(s)) = r_n(s) + 1$. If follows that $\rho_n(sh(s)) = \rho_n(s) + p_n/q_n$. Taking limits we see that $\phi_0(sh(s)) = \phi_0(s) + \lim_n \alpha_n = \phi_0(s) + \alpha$. ⊣

This finishes the proof of Theorem 55. ⊣

3.5 Kronecker Factors

Both odometer transformations and irrational rotations of the circle are ergodic discrete spectrum transformations. Because the odometer transformation based on $\langle k_n : n \in \mathbb{N} \rangle$ is a factor of any odometer based system T and the rotation $\mathcal{R}_\alpha$ is a factor of any circular system S, both are factors of the respective Kronecker factors of T or S. In general it is not the whole Kronecker factor in either case.

We make the following lemma explicit in the case of odometer based transformations. In the case of systems with a circle factor the exactly analogous results hold.

Lemma 58. *Let $(\mathbb{K}, \mathcal{B}, \mu, T)$ and $(\mathbb{L}, \mathcal{C}, \nu, S)$ be measure preserving systems. Suppose that $\mathbb{K}$ has an odometer factor $\mathcal{O}$ and that $\phi : \mathbb{K} \to \mathbb{L}$ is an isomorphism. Then there is a unique odometer factor $\mathcal{O}^*$ of $\mathbb{L}$ with an isomorphism $\phi^\pi : \mathcal{O} \to \mathcal{O}^*$ such that the following diagram commutes:*

$$
\begin{array}{ccc}
\mathbb{K} & \xrightarrow{\ \phi\ } & \mathbb{L} \\
\Big\downarrow{\pi^{\mathbb{K}}} & & \Big\downarrow{\pi^{\mathbb{L}}} \\
\mathcal{O} & \xrightarrow{\ \phi^\pi\ } & \mathcal{O}^*
\end{array}
$$

If each finite order eigenvalue of $\mathbb{L}$ has multiplicity 1 (e.g. if $\mathbb{L}$ is ergodic), then $\mathcal{O}^*$ is the unique odometer factor of $\mathbb{L}$ isomorphic to $\mathcal{O}$.

⊢   Since the unitary operator $U_\phi : L^2(\mathbb{K}) \to L^2(\mathbb{K})$ takes eigenfunctions to eigenfunctions, we know that U_ϕ takes the subspaces of $L^2(\mathbb{K})$ corresponding to $\mathcal{O}$ to a subspace of $L^2(\mathbb{L})$ corresponding to an isomorphic copy of $\mathcal{O}$. The lemma follows.   ⊣

An immediate corollary of Lemma 58 is that if $\mathbb{K}$ and $\mathbb{L}$ are ergodic odometer based systems over the same odometer $\mathcal{O}$, with projections π_K and π_L, then ϕ^π is an isomorphism between the canonical odometer factors.

We record the following consequences for later use;

Proposition 59. *Suppose that $\mathbb{K}$ and $\mathbb{L}$ are both ergodic odometer based systems with coefficients $\langle k_n : n \in \mathbb{N} \rangle$. Then any isomorphism $\phi : \mathbb{K} \to \mathbb{L}$ takes the canonical odometer factor $\mathcal{O}^{\mathbb{K}}$ of $\mathbb{K}$ to the canonical odometer factor $\mathcal{O}^{\mathbb{L}}$ of $\mathbb{L}$.*

Similarly if $\mathbb{K}^c$ and $\mathbb{L}^c$ are both ergodic circular systems with the same coefficient sequences $\langle k_n, l_n : n \in \mathbb{N} \rangle \rangle$, then any isomorphism between $\mathbb{K}^c$ and $\mathbb{L}^c$ takes the canonical rotation $\mathcal{K}^{\mathbb{K}}$ to the canonical rotation factor $\mathcal{K}^{\mathbb{L}}$

⊢   In the first case there is a unique factor of $\mathbb{K}$ and $\mathbb{L}$ corresponding to the eigenvalues of $\mathcal{O}^{\mathbb{K}}$ and $\mathcal{O}^{\mathbb{L}}$. Any isomorphism must preserve the factor corresponding to these eigenvalues. The same argument works for $\mathcal{K}$, as it is isomorphic to the rotation by $\alpha = \lim_n p_n/q_n$.   ⊣

3.6   Uniform Systems

In [5] it is established that the *strongly uniform* circular systems with sufficiently fast growing $\langle l_n : n \in \mathbb{N} \rangle$, are realizable as measure preserving diffeomorphisms of the torus. Strongly uniform systems are those for which each word in $\mathcal{W}_n$ occurs the same number of times in each word in $\mathcal{W}_{n+1}$. These systems carry unique non-atomic invariant measures, simplifying much of what we do later in this paper. For example the correspondence between the measures ν on uniform odometer systems $\mathbb{K}$ and ν^c on their uniform circular system counterparts $\mathbb{K}^c$ given in equation 33, is automatic.

In the forthcoming [8] we show that arbitrary (i.e. non-uniform) circular systems are realizable as measure preserving diffeomorphisms of the torus, provided that the measures of the words in $\mathcal{W}_n$ go to zero.

4 Details of Circular Systems

This section examines the circular systems defined in section 3.2 in more detail. Initially we are given a circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N} \rangle$ and $\langle q_n : n \in \mathbb{N} \rangle$ where q_n satisfies the inductive definition in equation 12. When n is fixed, we again let $j_i = (p_n)^{-1}i$ modulo q_n and $0 \leq j_i < q_n$. Without significant loss of generality it is convenient to assume that $\sum 1/q_n < 1/10$.

To understand joinings of circular systems we will be comparing generic elements (s, t) of circular $\mathbb{K}^c$ and $\mathbb{L}^c$, and their parsings into subwords. We will use the following terminology:

Definition 60. *Let u, v be finite sequences of elements of $\Sigma \cup \{b, e\}$ 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 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.*

4.1 Understanding the words

We elaborate on the descriptions given in Section 3.3. Our first combinatorial lemma is the following:

Lemma 61. *Let $w = \mathcal{C}(w_0, \ldots w_{k_n-1})$ for some n and $q = q_n, k = k_n, l = l_n$. View w as a word in the alphabet $\Sigma \cup \{b, e\}$ lying on the interval of integers $[0, klq^2)$.*

1. *If m_0 and m_1 are the locations of the beginnings of 0-subsections in the same 2-subsection, then $m_0 \equiv_q m_1$.*

2. *If m_0 and m_1 are such that m_0 is the location of the beginning of a 0-subsection occurring in a 2-subsection $\prod_{j=0}^{k-1}(b^{q-j_i}w_j^{l-1}e^{j_i})$ and m_1 at the i beginning of a 0-subsection 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 $m_1 - m_0 \equiv_q -j_1$.*

$\vdash$ To see the first point, the indices of the beginnings of 0-subsections 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 0-subsection of u and the first 0-subsection of v. But 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 $m_0 - m_1 \equiv_q q + j_i + q - j_{i+1} \equiv_q -j_1$. $\dashv$

Assume that $u \in \mathcal{W}_{n+1}$ and $v \in \mathcal{W}_{n+1} \cup \mathrm{rev}(\mathcal{W}_{n+1})$ 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 all of the 2-subsections in both words 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 left piece of the pair the even piece and the right piece the odd piece.

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

Lemma 62. *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 0-subsection 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 subwords of u by t. Moreover $s \equiv_q t - j_1$.*

$\vdash$ This follows easily from Lemma 61 $\dashv$

In the case where the split is trivial we get Lemma 62 with just one coefficient, s or t.

A special case Lemma 62 that we will use is:

Lemma 63. *Suppose that the 2-subsections of u divide the 2-subsections of v into two pieces and that for some occurrence of an n-subword of v in an even (resp. odd) piece is lined up with an occurrence of some n-word in u. Then every occurrence of an n-word 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 portion 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.

4.2 Full measure sets for circular systems

Fix a summable sequence $\langle \varepsilon_n : n \in \mathbb{N} \rangle$ of numbers in $[0, 1)$ and a circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N} \rangle$. As we argued in the proof of Lemma 50, the proportion of boundaries that occur in words of $\mathcal{W}_n^c$ is always summable, independently of the way we build $\mathcal{W}_n^c$. Recall the set $S \subseteq \mathbb{K}^c$ given in Definition 10, where $\mathbb{K}^c$ is the symbolic shift defined from a construction sequence.

Definition 64. *We define some sets that a typical generic point for a circular system eventually avoids. Let:*

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

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

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 the principal $n + 1$-block of$s\}$.*

Lemma 65. *Assume that $\sum 1/l_n < \infty$. Let ν be a shift invariant measure on $S \subseteq \mathbb{K}^c$, where $\mathbb{K}^c$ is a circular system. Then:*

1.
$$\sum_n \nu(E_n) < \infty.$$

   Assume that $\langle \varepsilon_n \rangle$ is a summable sequence, then for $i = 0, 1, 2$:

2.
$$\sum_n \nu(E_n^i) < \infty.$$

⊢   This is an application of the Ergodic Theorem.   ⊣

In particular we see:

Corollary 66. *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. *$s \notin E_n^i$.*

   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 $w_{k_n - 1}^{l_n - 1}$.*

⊢   This follows from the Borel-Cantelli Lemma.   ⊣

The elements s of S such that some shift $sh^k(s)$ fails one of the conclusions 1.)-4.) of Corollary 66 form a measure zero set. Consequently we work on those elements of S whose whole orbit satisfies the conclusions of Corollary 66. Note, however that the $N(sh^k(s))$ depends on the shift k.

Definition 67. *We will call n* mature *for s (or say that s is mature at stage n) iff n is so large that $s \notin E_m \cup \bigcup_{0 \leq i \leq 2} E_m^i$ for all $m \geq n$.*

Thus if s is mature at stage n then for all $m > n$ the principal m-block of s exists and conclusions 1-4 of Corollary 66 hold.

Recall that in Section 3.2, we defined a canonical factor of a circular system which we called the circle factor. Since the notion of maturity only involves the punctuation of the words involved, it is an easy remark that for all $s \in S$, n is mature for s just in case n is mature for $\pi(s)$, where π is the canonical factor map.

For the following definition and lemma, we view $s \in S$ as a function with domain $\mathbb{Z}$, and $s \in \mathcal{W}_n$ as a function with domain $[0, q_n)$ or, sometimes, an interval $[k, k + q_n)$. In each of these cases we use *dom(s)* to mean the domain of s.

Definition 68. *We will use the symbol ∂_n in multiple equivalent ways. If $s \in S$ or $s \in \mathcal{W}_m^c$ we define $\partial_n = \partial_n(s)$ to be the collection of i such that $sh^i(s)(0)$ is in the boundary portion of an n-subword of s. This is well-defined by our unique readability lemma. In the spatial context we will say that $s \in \partial_n$ if $s(0)$ is the boundary of an n-subword of s.*

For $s \in S$

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

An integer, $i \in \partial_n(s) \subseteq \mathbb{Z}$ iff $sh^i(s)$, viewed as an element of $\mathbb{K}^c$, belongs to the n-boundary, ∂_n.

In what follows we will be considering a generic point s and all of its shifts. We will use the fact if s is mature at stage n, then we can detect locally those i for which the i-shifts of s are mature.

Lemma 69. *Suppose that $s \in S$, n is mature for s and $n < m$.*

1. *Suppose that $i \in [-r_m(s), q_m - r_m(s))$. Then n is mature for $sh^i(s)$ iff*

   (a) $i \notin \bigcup_{n \leq k \leq m} \partial_k$ and
   (b) $sh^i(s) \notin \bigcup_{n \leq k < m} (E_k^0 \cup E_k^1 \cup E_k^2)$.

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

In particular, if $\varepsilon_{n-1} > \sup_m (1/q_m) \sum_{k=n}^{m-1} 6\varepsilon_k q_{k+1}$, $1/l_{n-1} > \sum_{k=n}^{\infty} 1/l_k$ and n is mature for s, the upper density of those $i \in \mathbb{Z}$ for which the i-shift of s is not mature for n is less than $1/l_{n-1} + \varepsilon_{n-1}$.

Similarly:

Lemma 70. *Suppose that $s \in S$ and s has a principal n-block. Then n is mature provided that $s \notin \bigcup_{n \leq m} E_m^0 \cup E_m^1 \cup E_m^2$. In particular, if n is mature for s and s is not in a boundary portion of its principal $n - 1$-block or in $E_{n-1}^0 \cup E_{n-1}^1 \cup E_{n-1}^2$, then $n - 1$ is mature for s.*

4.3 The ♮ map

Proposition 59 implies that any isomorphism ϕ between an ergodic $(\mathbb{K}^c, sh)$ and $(\mathbb{K}^c, sh^{-1})$ induces an isomorphism ϕ^π between $(\mathcal{K}, sh)$ and $(\mathcal{K}, sh^{-1})$, where $\mathcal{K}$ is the canonical circle factor. Because $(\mathcal{K}, sh^{-1})$ is canonically isomorphic with $(\mathrm{rev}(\mathcal{K}), sh)$ (Proposition 18) and $(\mathcal{K}, sh)$ is isomorphic to the rotation $\mathcal{R}_\alpha$ of the circle, we see that $(\mathrm{rev}(\mathcal{K}), sh)$ is isomorphic to the rotation $\mathcal{R}_{-\alpha}$.

We use a specific isomorphism $\natural : (\mathcal{K}, sh) \to (\mathrm{rev}(\mathcal{K}), sh)$ as a benchmark for understanding of potential maps $\phi : \mathbb{K}^c \to \mathrm{rev}(\mathbb{K}^c)$. If we view $\mathcal{K}$ as a rotation $\mathcal{R}_\alpha$ of the unit circle by α 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 $\mathcal{R}_{-\alpha}$. Copying $\natural$ over to a map on the unit circle gives an isomorphism ϕ between $\mathcal{R}_\alpha$ and $\mathcal{R}_{-\alpha}$. Such an isomorphism must be of the form

$$\phi(z) = \bar{z} e^{2\pi i \beta}$$

for some β. It follows immediately from this characterization that $\natural$ is an involution, however for completeness we prove this directly (and symbolically) in Proposition 79.

As usual we find it more convenient to work on the unit interval $I = [0, 1)$ rather than the unit circle. The complex conjugacy map $z \mapsto \bar{z}$ corresponds to the map $x \mapsto -x$ on $[0, 1)$.

We begin by recalling from equation 11 the formula for a $w \in \mathcal{W}_{n+1}^c$ that is of the form $\mathcal{C}(w_0, \dots w_{k_n-1})$:

$$w = \prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (b^{q-j_i} w_j^{l-1} e^{j_i}) \tag{16}$$

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 10, we see that this can be rewritten as[15]

$$\text{rev}(w) = \prod_{i=1}^{q} \prod_{j=1}^{k} (e^{q-j_i} \text{rev}(w_{k-j})^{l-1} b^{j_i}).$$ (17)

We can reindex again and get another form of equation 17:

$$\text{rev}(w) = \prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (e^{q-j_{i+1}} \text{rev}(w_{k-j-1})^{l-1} b^{j_{i+1}}).$$ (18)

We can now state the basic lemma about the way w lines up with a shift of $\text{rev}(w)$.

Lemma 71. *Let $w \in \mathcal{W}_{n+1}^c$ and view w as sitting at location $[0, q_{n+1}) \subseteq \mathbb{Z}$. Let $q = q_n$ and $k = k_n$. Consider $sh^{-j_1}(\text{rev}(w))$ as being the word $\text{rev}(w)$ in location $[j_1, q_{n+1} + j_1)) \subseteq \mathbb{Z}$. For all but at most $2kq$ of the occurrences of an n-subword w_j of w starting in a location $r \in [0, q_{n+1})$, the reversed word $\text{rev}(w_{k-j-1})$ occurs in $sh^{-j_1}(\text{rev}(w))$ starting at r.*

⊢ The word w starts with a block of q b's and then a block of $l-1$ copies of w_0, whereas $\text{rev}(w)$ starts with a block of $q - j_1$ e's followed by $l-1$ copies of $\text{rev}(w_{k-1})$. Hence if we shift $\text{rev}(w)$ to the right by j_1 (to get $sh^{-j_1}(\text{rev}(w))$) the first copy of $\text{rev}(w_{k-1})$ is aligned with the first copy of w_0 in w. Hence all of the copies of $\text{rev}(w_{k-1})$ in the first 1-subsection are aligned with the copies of w_0 in the first 1-subsection of w. Because the consecutive blocks of b's and e's (or e's and b's) in the 2-subsections add up to q we see that every copy of $\text{rev}(w_{k-j-1})$ in the first 2-subsection of $sh^{-j_1}(\text{rev}(w))$ is aligned with with a copy of w_j.

We now argue as in Section 4.1. At the end of each 2-subsection, w has a block of e's of length j_i, followed at the beginning of the next 2-subsection, by a block of b's of length $q - j_{i+1}$. Together the e's and b's form a block of length $j_i + q - j_{i+1}$, which is equivalent $\text{mod}(q)$ to $-j_1$. Similarly the combined length of a block of b's and e's finishing and starting consecutive 2-subsections of $\text{rev}(w)$ is equal to $-j_1 \text{ mod}(q)$.

Both the beginning of the block of e's ending the kth 2-subsection and the end of the block of b's starting the $k + 1^{st}$ 2-subsection are of distance less than q from the location of the end of the kth 2-subsection. It follows

---

[15] We take $j_q = 0$.

from this and the comments in the previous paragraph, that if S_1 and S_2 are consecutive 2-subsections of w and S_1' and S_2' are the corresponding 2-subsections of $\text{rev}(w)$ then the beginning of the first occurrence of $\text{rev}(w_{k-1})$ in S_2' is within $2q$ of the first occurrence of w_0 is S_2 and their locations are equivalent $\text{mod}(q)$. Hence inside the first 1-subsection, the 0-subsections are lined up except for at most 2 copies of w_0. This pattern is continued through S_2, giving at most $2k$ locations of n-blocks that are not aligned in S_2.

Since there are less than q 2-subsections with potential misalignments, the Lemma is proved. ⊣

The next proposition gives a somewhat more detailed view into situation of Lemma 71.

Proposition 72. *Let* $w, w' \in \mathcal{W}_{n+1}^c$ *and suppose that*

$$w = \mathcal{C}(v_0, v_1, \ldots v_{k_n - 1}) \quad and \quad w' = \mathcal{C}(v_0', v_1', \ldots v_{k_n - 1}').$$

We look at the relative positions of n-words in w and $sh^{-j_1}(\text{rev}(w'))$.

1. *Each occurrence of v_i in w is either lined up with an occurrence of* $\text{rev}(v_{k_n - i - 1}')$ *or entirely lined up with a section of ∂_n inside $sh^{-j_1}(\text{rev}(w'))$.*

2. *There is a number C such that for all i the number of occurrences of* v_i *lined up with an occurrence of* $\text{rev}(v_{k_n - i - 1}')$ *is C.*

⊢ The first part is clear from the proof of Lemma 71. The second part follows because all of the 1-subsections in a given 2-subsection of w have the same alignment relative to $sh^{-j_1}(\text{rev}(w'))$. ⊣

Since the total number of occurrences of n-subwords in klq, the proportion of n-subwords lined up with ∂_n in $sh^{-j_i}(\text{rev}(w'))$ is at most $2/l$.

Suppose that $\mathcal{K}$ is given by the canonical construction sequence $\langle \mathcal{W}_n^\alpha : n \in \mathbb{N} \rangle$. We define a sequence of functions $\langle \Lambda_n : n \in \mathbb{N} \rangle$ and argue that they converge to an isomorphism from $\mathcal{K}$ to $\text{rev}(\mathcal{K})$.

We begin by defining an increasing sequence of natural numbers. Recall the definition of the Anosov-Katok coefficients p_n and q_n given in equations 13 and 12. Since p_n and q_n are relatively prime we can define $(p_n)^{-1}$ in $\mathbb{Z}/q_n\mathbb{Z}$. For the following definition we will view $(p_n)^{-1}$ as a *natural number* with $0 \leq (p_n)^{-1} < q_n$.[16]

---

[16] In the notation used to define $\mathcal{C}$, $(p_n)^{-1} = j_1$. However the notation j_1 is ambiguous (it depends on n), so we use $(p_n)^{-1}$ in this context.

We let $A_0 = 0$ and
$$A_{n+1} = A_n - (p_n)^{-1}. \tag{19}$$

Lemma 73. *If A_n is defined as above, then $|A_{n+1}| < 2q_n$.*

⊢ This is proved inductively using the fact that $q_{n+1} > 2q_n$. ⊣

Let $\mathcal{K}$ be the circular system in the language $\Sigma = \{*\}$, as given in Definition 53. We now define a stationary code $\overline{\Lambda}_n$ with domain S that approximates elements of $\text{rev}(\mathcal{K})$ by defining

$$\Lambda_n(s) = \begin{cases} sh^{A_n + 2r_n(s) - (q_n - 1)}(\text{rev}(s))(0) & \text{if } r_n(s) \text{ is defined} \\ b & \text{otherwise} \end{cases} \tag{20}$$

Since for all $s \in S$ and all large enough n, $r_n(s)$ is defined, the default value is only obtained for finitely many n.

Lemma 74. Λ_n *is given by a finite code.*

⊢ To check whether $r_n(s)$ is defined one need only examine s on the interval $[-q_n, q_n] \subseteq \mathbb{Z}$. The relevant portion of $\text{rev}(s)$ necessary to compute $\Lambda_n(s)$ is contained in $s \restriction [-q_n - A_n, q_n + A_n]$. Hence Λ_n is determined by a finite code. ⊣

The formula in equation 20 can be understood as follows. Suppose that $s \in S$ and s has a principal n-block. Then the element s^* defined as $sh^{2r_n(s) - (q_n - 1)}(\text{rev}(s))$ belongs to $\text{rev}(\mathcal{K})$, has a principal n-block that is the reverse of the principal n-block of s and moreover, the principal n-block of s^* is exactly lined up with the principal n-block of s.

The reverse of the principal n-block of s begins with a block of $q_{n-1} - (p_{n-1})^{-1}$ many e's, and hence if $s' = sh^{(-(p_{n-1})^{-1}) + 2r_n(s) - (q_n - 1)}(\text{rev}(s))$ then the first $n - 1$-subword of the principal n-block of s' is lined up with the first $n - 1$-subword of the principal n-block of s. The rest of the terms used to define A_n (coming from A_{n-1}) are used for lower order adjustments inside this principal n-block.

Thus, a qualitative description of $\overline{\Lambda}_n(s)$ can be given as follows:

1. It first reverses the principal n-block of s leaving it exactly lined up.

2. It then adjusts the result by shifting so that the first occurrence of a reverse $n-1$-block lines up with the first $n-1$-subword of the principal n-block of s. (So far we have described $sh^{(-(p_{n-1})^{-1})+2r_n(s)-(q_n-1)}(\text{rev}(s))$.) By Lemma 71, we get a sequence where the principal n-block of $\Lambda_n(s)$ has the vast majority of its $n-1$-blocks lined up with the $n-1$-blocks of s: all of them except those that span a section of boundary at the juncture of two 2-subsections of the principal n-word of s.

3. Finally it shifts by A_{n-1} which is the cumulative adjustment at earlier stages.

The next lemma follows from this description:

Lemma 75. *Let $n < m$ and suppose that $s \in \mathcal{K}$ has a principal m-block. Let $s' = sh^{2r_m - q + A_m - A_n}(\text{rev}(s))$. Then at least*

$$\prod_n^{m-1}(1 - \frac{2}{(l_i - 1)})$$

proportion of the n-blocks in the principal m-block of s are lined up with n-blocks in s'.

$\vdash$ We first consider $m = n+1$. By Lemma 71, all but $2k_n q_n$ of the n-blocks in w are aligned with the n-blocks in $sh^{-j}(\text{rev}(w))$. This is proportion

$$1 - \frac{2k_n q_n}{k_n q_n(l_n - 1)} = 1 - \frac{2}{l_n - 1}.$$

The general result follows by induction. $\dashv$

Theorem 76. *Suppose that $\langle k_n, l_n : n \in \mathbb{N} \rangle$ is a circular coefficient sequence. Then the sequence of stationary codes $\langle \overline{\Lambda}_n : n \in \mathbb{N} \rangle$ converges to a shift invariant function $\natural : \mathcal{K} \to (\{*\} \cup \{b, e\})^{\mathbb{Z}}$ that induces an isomorphism $\natural$ from $\mathcal{K}$ to $\text{rev}(\mathcal{K})$.*

$\vdash$ We first show that the sequence $\langle \overline{\Lambda}_n : n \in \mathbb{N} \rangle$ converges, which will follow if we show that the code distances between the Λ_n and Λ_{n+1} are summable. For notational simplicity, let $q = q_n, k = k_n, l = l_n$ and $j \equiv_q (p_n)^{-1}$ with $0 \leq j < q$.

Claim: There is a summable sequence of positive numbers δ_n such that for almost all s, the $\bar{d}$-distance between $\bar{\Lambda}_n(s)$ and $\bar{\Lambda}_{n+1}(s)$ is bounded by δ_n, and $\bar{\Lambda}_n(s)$ and $\bar{\Lambda}_{n+1}(s)$ agree on all but at most δ_n proportion of the n-blocks of s.

We use Lemma 35, which tells us that for a typical $s \in S$, the code distance between Λ_n and Λ_{n+1} is $\bar{d}(\bar{\Lambda}_n(s), \bar{\Lambda}_{n+1}(s))$, which is defined to be the density of

$$D =_{def} \{k : \Lambda_n(sh^k(s))(0) \neq \Lambda_{n+1}(sh^k(s))(0)\}. \tag{21}$$

Because $|\mathcal{W}_n^\alpha = 1$ for each n, there is only one possible n-subword at any location of any element of $\mathrm{rev}(\mathcal{K})$. Thus to compute $\bar{d}$-distance, it suffices count positions where the $\bar{\Lambda}_m$'s disagree on the *locations* of the n-subwords.

By Lemma 69 for a typical $s \in S \subseteq \mathcal{K}$ and all n, $I_n =_{def} \{i : n$ is not mature for $sh^i(s)\}$ has density at most $1/l_{n-1} + \varepsilon_{n-1}$, hence we can neglect these i when computing the density of D.

This allows us to assume that $r_{n+1}(s)$ is defined. We compute the density of the difference between $\bar{\Lambda}_n$ and $\bar{\Lambda}_{n+1}$ as they pass across an $n+1$-block in s. If this number is d then the distance between Λ_n and Λ_{n+1} is bounded by the sum of d and the density of I_n.

As Λ_{n+1} crosses an $n+1$-block it produces the reverse $n+1$-block shifted by A_{n+1}. Explicitly, if w is the $n+1$-block of s, as Λ_{n+1} crosses w it produces $sh^{A_{n+1}}(\mathrm{rev}(w))$. As Λ_n passes across this same section, each time it crosses an n-block w' it produces $sh^{A_n}(\mathrm{rev}(w'))$. If w' starts at r then the beginning of this copy of $sh^{A_n}(\mathrm{rev}(w'))$ is $r - A_n$.

We begin by rewriting $sh^{A_{n+1}}(\mathrm{rev}(w))$ as $sh^{A_n}(sh^{-j}(\mathrm{rev}(w)))$ where $j = (p_n)^{-1}$. By Lemma 71, all but $2kq$ of the n-blocks in w are aligned with the n-blocks in $sh^{-j}(\mathrm{rev}(w))$. Hence, relative to the complement of I_n, the portion of the principal $n+1$-block w of s that lies in an n-block aligned with an n-block of $sh^{-j}(\mathrm{rev}(w))$ is

$$\frac{k(l-1)q^2 - 2kq}{k(l-1)q^2} = 1 - \frac{2}{(l-1)q} \tag{22}$$

Because there is only one possible n-word, whenever $sh^{A_n}(\mathrm{rev}(w'))$ is aligned with $sh^{A_n}(sh^{-j}(\mathrm{rev}(w)))$ they are equal.

Putting this altogether, we see that Λ_n and Λ_{n+1} agree on all of the n-subwords of the principal $n+1$-block of s that are aligned with $sh^{-j}(\mathrm{rev}(w))$.

The disagreements are limited to the n-subwords that are not aligned and the boundary. The total length of the disagreements is therefore bounded by

$$(2kq) * q + kq^2 = 3kq^2.$$

This has proportion $3kq^2/klq^2 = 3/l$.

Thus the distance between Λ_n and Λ_{n+1} is bounded by $1/l_{n-1} + \varepsilon_{n-1} + 3/l_n$. In particular the distances are summable and the sequence $\langle \bar{\Lambda}_n : n \in \mathbb{N} \rangle$ converges almost everywhere to a function $\natural : \mathcal{K} \to (\Sigma \cup \{b, e\})^{\mathbb{Z}}$.

We now show that $\natural$ is an isomorphism between $\mathcal{K}$ and $\mathrm{rev}(\mathcal{K})$. Since $\bar{\Lambda}_n$ takes an n-block to a shift of the reverse n-block, it makes sense to discuss the *principal n-block* of $\bar{\Lambda}(s)$. Since the r_n's cohere as in Remark 14, for $n < m$, $r_m(\bar{\Lambda}_m(s))$ is in the $r_n(\bar{\Lambda}_m(s))^{th}$ position of the principal n-block of $\bar{\Lambda}_m(s)$ (provided both r_n and r_m are defined). An application of the Ergodic Theorem shows that if D_n is defined to be the collection of s such that:

$r_n(\bar{\Lambda}_n(s))$ exists and the principal n-words of $\bar{\Lambda}_n(s)$ and $\bar{\Lambda}_{n+1}(s)$ disagree

then $\sum \nu(D_n) < \infty$. From the Borel-Cantelli Lemma, it follows that for almost every s for all large enough n the principal n-blocks of $\bar{\Lambda}_n(s)$ and $\bar{\Lambda}_{n+1}(s)$ are the same, and thus that for $s \in S, \natural(s) \in \mathrm{rev}(\mathcal{K})$.

We now argue that if s is typical and $s^* = \natural(s)$, then $s^* \in \mathrm{rev}(S)$. It suffices to show that $\lim_{n \to \infty} -r_n(s^*) = -\infty$ and $\lim_{n \to \infty} q_n - r_n(s^*) = \infty$.[17]

If n is mature for s and large enough that for $m > n, \bar{\Lambda}_m(s)$ and $\bar{\Lambda}_n(s)$ have the same principal n-blocks, then $r_n(s^*) = r_n(s) + A_n$ unless $r_n(s) \in [0, |A_n|)$. Assuming that $r_n(s) \geq |A_n|$, we know from Lemma 73 that

$$r_n(s) - 2q_{n-1} < r_n^*(s) < r_n(s).$$

Hence, $-r_n^*(s) \leq 2q_{n-1} - r_n(s)$ and $q_n - r_n^*(s) \geq q_n - r_n(s)$. Applying Lemma 69 (using the fact that $\sum n q_{n-1}/q_n < \infty$, and hence $\sum |A_n|/q_n < \infty$) we see that for large n, $r_n(s) > |A_n|$ and that $r_n(s) - 2q_n \to \infty$. Since $q_n - r_n(s) \to \infty$ we have shown that $s^* \in \mathrm{rev}(S)$.

As noted before Theorem 55, if $s \in S$ then s is determined by any tail of the sequence $\langle r_n(s) : n \in \mathbb{N} \rangle$. In particular, if we know a tail of $\langle r_n(s^*) :$

---

[17]We are adopting the convention that in defining $r_n(s^*)$ for $s^* \in \mathrm{rev}(S)$ we count r_n from the left end of an n-block. Thus the position r in a word $w \in \mathcal{W}_n^\alpha$ corresponds to the position $q - 1 - r$ in $\mathrm{rev}(w)$.

$n \in \mathbb{N}\rangle$ we can determine s^*. Since for large n, $r_n(s^*) = r_n(s) + A_n$, $\natural$ is one-to-one on a set of measure one.

We can now conclude that $\natural$ is an isomorphism. It is shift invariant since it is a limit of stationary codes, it maps from S to $\mathrm{rev}(S)$, and is one-to-one on a set of ν-measure one. If we define a measure μ on the Borel sets of $\mathrm{rev}(\mathcal{K})$ by setting $\mu(A) = \nu(\natural^{-1}(A))$, then μ is a shift invariant, non-atomic measure on $\mathrm{rev}(S)$. Since S is uniquely ergodic, $\mathrm{rev}(S)$ is as well and thus μ must be equal to the unique invariant measure ν. We have shown that $\natural$ is an isomorphism between $\mathcal{K}$ and $\mathrm{rev}(\mathcal{K})$. $\dashv$

Definition 77. *We denote the limit of $\langle \bar{\Lambda}_n : n \in \mathbb{N} \rangle$ by $\natural : \mathcal{K} \to \mathrm{rev}(\mathcal{K})$.*

We describe the qualitative behavior of $\natural$ in a remark that we will use later:

Remark 78. *There is a summable sequence $\langle \delta_n \rangle$ such that for all but $1 - \delta_n$ measure of $s \in S \subseteq \mathcal{K}$, there is an interval I containing 0 in $\bar{\Lambda}_n(s)$ such that $s \upharpoonright I \in \mathcal{W}_n^\alpha$, and moreover $\bar{\Lambda}_{n+1}(s)$ and $\bar{\Lambda}_n(s)$ agree on this interval. It follows from the Borel-Cantelli Lemma that for almost all s and large enough n, $\natural(s)$ agrees with $\bar{\Lambda}_n(s)$ on the principal n-block of s. Thus for a typical s and large enough n, the map $\natural$ reverses the principal n-block while keeping its location and then shifts it by A_n.*

As noted at the beginning of this section, the next proposition follows immediately from Theorem 55, however we include a symbolic proof for completeness.

Proposition 79. *The map $\natural$ is an involution.*

$\vdash$ It is immediate from the qualitative description of $\bar{\Lambda}_n$ given before Lemma 75, that each $\bar{\Lambda}_n$ is an involution. To see that $\natural^2$ is the identity, let $\epsilon > 0$. We can choose an m_0 large enough that for all $m \geq m_0$, $\bar{\Lambda}_m$ and $\natural$ agree with $\bar{\Lambda}_{m_0}$ on all but ϵ proportion of the m_0-blocks and $\bigcup_{m_0+1}^\infty \partial_k$ has measure $\epsilon * 10^-6$. Then $\natural \circ \bar{\Lambda}_{m_0}$ is equal to the identity on a set of density at least $1 - \epsilon$. Letting $\epsilon \to 0$ and $m_0 \to \infty$ completes the argument. $\dashv$

4.4 Synchronous and Anti-synchronous joinings

Every odometer based system has a built in metronome: its odometer factor defined in Lemma 41. Correspondingly circular systems can be timed by their canonical rotation factor defined in Lemma 54.

Joinings between odometer based and circular systems may induce non-trivial automorphisms of the underlying timing structure. To avoid this complication we restrict ourselves to synchronous and anti-synchronous joinings: those which preserve or exactly reverse the underlying timing. We now make this idea precise.

Both the odometer transformations and rotations of a circle have easily understood inverse transformations and the isomorphisms between transformations and their inverses are given by the maps $x \mapsto -x$ and $\mathrm{rev}() \circ \natural$ respectively. If $\mathbb{K}$ and $\mathbb{L}$ are either odometer based or circular systems let $\mathbb{K}^\pi$ and $\mathbb{L}^\pi$ be the corresponding odometer or rotation systems on which they are based.

Definition 80.
- 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 $x \mapsto -x$.

- 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 $\mathrm{rev}() \circ \natural$ on $\mathcal{K} \times \mathcal{L}^{-1}$.

There is always a synchronous joining of odometer systems with the same underlying timing factor $\mathcal{O}$:

Definition 81. *Suppose that $\mathbb{K}$ and $\mathbb{L}$ are based on $\mathcal{O}$. Then the relatively independent joining of $\mathbb{K}$ and $\mathbb{L}$ over $\mathcal{O}$ is a synchronous joining, which we will call the* synchronous product joining. *The relatively independent joining*

of $\mathbb{K}$ *and* $\mathbb{L}^{-1}$ *over the map* $x \mapsto -x$ *we will call the* anti-synchronous product joining. *We will use the same terminology for the independent joinings of circular systems over the identity and* $\text{rev}() \circ \natural$.

5 Building the Functor $\mathcal{F}$

The main result of this paper concerns two categories whose objects are odometer based systems and circular systems respectively. The morphisms in these categories will be graph joinings. We will show that there is a functor taking odometer systems to circular systems that preserves the factor and conjugacy structure. In this section we focus on defining the function from odometer based systems to circular systems that underlies the functorial isomorphism between these categories.

We begin by defining a function from the odometer based symbolic shifts $\mathbb{K}$ to the circular symbolic shifts $\mathbb{K}^c$. After having done so we define $\mathcal{F}$ on the pairs $(\mathbb{K}, \mu)$ where μ is an invariant measure on $\mathbb{K}$. Finally we define $\mathcal{F}$ on synchronous and anti-synchronous graph joinings.

We will use the notation that $K_n = \prod_{i<n} k_i$. Then the K_n's are the lengths of the odometer based words in $\mathcal{W}_n$ and the q_n's are the lengths of the circular words in $\mathcal{W}_n^c$.

Except where otherwise stated we will assume that we are working with a fixed circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N} \rangle$.

Let Σ be a language and $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ be 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 $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$ and bijections $c_n : \mathcal{W}_n \to \mathcal{W}_n^c$ by induction as follows:

1. Let $\mathcal{W}_0^c = \Sigma$ and c_0 be the identity map.

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

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

   Define the map c_{n+1} by setting

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

We note in case 2 the prewords are:

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

Definition 82. *Define a map $\mathcal{F}$ from the set of odometer based systems (viewed as subshifts) to circular systems (viewed as subshifts) as follows. Suppose that $\mathbb{K}$ is 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 $\langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$.

Suppose that $\mathbb{K}^c$ is a circular system with coefficients $\langle k_n, l_n : n \in \mathbb{N}\rangle$. We can recursively 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 $\langle k_n : n \in \mathbb{N}\rangle$.[18]

If $\mathbb{K}$ is the resulting odometer based system then $\mathcal{F}(\mathbb{K}) = \mathbb{K}^c$. Thus we see:

Proposition 83. *The map $\mathcal{F}$ is a bijection between odometer based symbolic systems with coefficients $\langle k_n : n \in \mathbb{N}\rangle$ and circular symbolic systems with coefficients $\langle k_n, l_n : n \in \mathbb{N}\rangle$.*

$\vdash$ That $\mathcal{F}$ is one-to-one follows from the unique readability of words occurring in the construction sequence $\langle W_n : n \in \mathbb{N}\rangle$. $\dashv$

Remark 84. *It is clear from Definition 82 that $\mathcal{F}$ preserves uniformity and strong uniformity (see [5] for these notions). In fact it preserves much more: the simplex of non-atomic invariant measures, rank one transformations and so on. We verify much of this in this paper and more in the forthcoming [8].*

To understand the correspondence between measures on $\mathbb{K}$ and $\mathbb{K}^c$ we will have to understand the structure of basic open intervals. Recall that we write $\langle u \rangle_L$ to mean the basic open interval of $\mathbb{K}$ determined by u sitting on the interval $[L, L + |u|) \subseteq \mathbb{Z}$. Without the subscript L, $\langle u \rangle$ is shorthand for $\langle u \rangle_0$. We adopt the same conventions for $\mathbb{K}^c$, that the subscripts correspond to the beginning of the sequence and without a subscript the sequence begins at zero.

---

[18]We are using the strong unique readability assumption on the P_n's to see the unique readability of the words in the sequence $\langle \mathcal{W}_n : n \in \mathbb{N}\rangle$.

5.1 Genetic Markers

To see that $\mathcal{F}$ can be extended to a map from invariant measures on odometer based systems to invariant measures on circular systems, we begin by recalling how to identify elements of a symbolic system. Suppose that $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ is a construction sequence for an odometer based transformation $\mathbb{K}$. Let $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$ be the corresponding circular construction sequence for $\mathbb{K}^c$. By Lemma 15 to specify a typical $s \in \mathbb{K}$ or $s^c \in \mathbb{K}^c$, it suffices to give a tail of the sequence of principal n-blocks $\langle w_n(s) : N \leq n \in \mathbb{N} \rangle$ or $\langle w_n^c(s^c) : N \leq n \in \mathbb{N} \rangle$ along with the locations $\langle r_n(s) : N \leq n \rangle$ or $\langle r_n(s^c) : N \leq n \rangle$.

Definition 85. *Suppose that u, v are words in $\mathcal{W}_n$ and $\mathcal{W}_{n+1}$ respectively and u occurs as an n-subword of v in a particular location. Viewing v as a concatenation $w_0 w_1 \ldots w_{n_k-1}$ of n-subwords, there is a j such that $u = w_j$. Let $j_n^* = j$ and call j_n^* the* genetic marker *of u in v.*

Suppose that $u \in \mathcal{W}_n$ and $v \in \mathcal{W}_{n+k}$ and u is an n-subword of v occurring at a particular location. Then there is a sequence of words $u_n = u, u_{n+1}, \ldots u_{n+k-1}, u_{n+k} = v$ such that u_i is a $n+i$-subword of v at a definite location and the location of u in v is inside u_i. Let j_{n+i}^* be the genetic marker of u_{n+i} inside u_{n+i+1}. We call the sequence $\vec{j}^* = \langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$ the genetic marker *of u in v. If $\vec{j}^*$ is the genetic marker of some n-word inside and m-word, we will call it an (n, m)-genetic marker.*

If u occurs as a subword of v then the genetic marker $\langle j_n^*, j_{n+1}^* \ldots j_{n+k-1}^* \rangle$ of that occurrence codes its location inside v.

Suppose that $s \in \mathbb{K}$ has principal n-blocks $\langle w_n : n \in \mathbb{N} \rangle$. Each w_{n+1} is a concatenation of words $v_0 v_1 \ldots v_{k_n-1}$. Let

$$j_n' =_{def} \frac{r_{n+1}(s) - r_n(s)}{K_n} \tag{23}$$

or equivalently

$$r_{n+1}(s) = r_n(s) + j_n' K_n. \tag{24}$$

Each w_{n+1} is a concatenation of words $v_0 v_1 \ldots v_{k_n-1}$, and we see that $s(0)$ belongs to $v_{j_n'}$. In particular, the genetic marker of w_n inside w_{n+k} is the sequence $\langle j_n', j_{n+1}', \ldots j_{n+k-1}' \rangle$.

Genetic markers for regions of words in $\mathcal{W}_{n+k}^c$: In circular words, genetic markers code regions rather than subwords. Given u and v as above, we

can consider the construction of $c_{n+k}(v)$ starting with the collection $\{c_n(u) : u$ is an n-subword of $v\}$. Each of the genetic markers $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$ of a subword u of v determines a *region* of n-subwords of $c_{n+k}(v)$. More explicitly, in the first step of the construction we put u into the $(j_n^*)^{th}$ argument of $\mathcal{C}_n$. At the next step we put the result into the j_{n+1}^* argument of $\mathcal{C}_{n+1}$ and so on. Thus we see that there are bijections between

1. sequences $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$ with $0 \leq j_m^* < k_m$,

2. n-subwords u of v,

3. the regions of v^c occupied by the occurrences of powers $(u^c)^{l_n - 1}$ where u^c is the element of $\mathcal{W}_n^c$ determined by $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$.

Thus genetic markers give the correspondence between the regions of $c_{n+k}(v)$ that are not in $\bigcup_{n < m \leq n+k} \partial_m$ and particular occurrences of an n-word u in v.

The next lemma computes the number of occurrences of a $c_n(u)$ with a given genetic marker $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$ in $c_{n+k}(v)$.

Lemma 86. *Suppose that u^c occurs in v^c with genetic marker $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$. Then the number of occurrences of u^c in v^c with the same genetic marker $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$ is*

$$\prod_n^{n+k-1} q_i(l_i - 1). \tag{25}$$

⊢   Fix m and $v^c \in \mathcal{W}_m^c$. We prove equation 25 for $n = m - k$ by induction on $k \geq 1$. If $k = 1$ then we have a single genetic marker j_{m-1}^*. By formula 11 for $\mathcal{C}_{m-1}$ we see that the j_{m-1}^* argument occurs in v^c exactly $q_n(l_n - 1)$ times.

Suppose now that we know that formula 25 holds for $k - 1$. We show it for k. Let $n = m - k$ and u^c be the n-subword of v^c with genetic marker $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$. Let w^c be the subword of v^c with genetic marker $\langle j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$. Then:

$$|\{\text{occurrences of } u^c \text{ in } v^c \text{ with marker } \langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle\}|$$

is equal to

$$|\{\text{occurrences of } u^c \text{ in } w^c \text{ with marker } j_n^*\}| \times$$

$$|\{\text{occurrences of } w^c \text{ in } v^c \text{ with marker } \langle j_{n+1}^*, \ldots j_{n+k-1}^* \rangle\}|$$

The lemma follows. $\dashv$

Since particular (n, m)-genetic markers $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$ correspond to powers of u^c's that occur with the same multiplicity in v^c, independently of the marker, we see that for a given u and v:

$$\frac{|\{\text{occurrences of } u \text{ in } v\}|}{|\{n\text{-subwords of } v\}|} = \frac{|\{\text{occurrences of } c_n(u) \text{ in } c_{n+k}(v)\}|}{|\{\text{circular } n\text{-subwords of } c_{n+k}(v)\}|} \quad (26)$$

We can restate equation 26 in the language of section 2.6. It says that

$$\text{EmpDist}(v)(u) = \text{EmpDist}(c_{n+k}(v))(c_n(u)). \quad (27)$$

In particular, if we fix a set S^* of genetic markers we can compare the number of occurrences of a word with genetic marker in S^* in $v \in \mathcal{W}_{n+k}$ with the number of occurrences in the corresponding $v^c \in \mathcal{W}_{n+k}^c$. Specifically, the number of occurrences of a word u^c in v^c at some genetic marker in S^* is $|S^*| * \prod_n^{n+k-1} q_i(l_i - 1)$. The proportion of n-words occurring with a genetic marker in S^* relative to all n-words occurring in v^c is the same as the proportion of n-words with genetic markers in S^* occurring in v relative to the total number of genetic markers. The number of (n, m)-genetic markers is $\prod_n^{n+k-1} k_i$ so this proportion is equal to

$$\frac{|S^*|}{\prod_n^{n+k-1} k_i}. \quad (28)$$

This is simply a restatement of our discussion involving empirical distributions in Section 2.6.

We introduce some notation that allows us to compare densities of various sets between odometer based and circular words. For sets $A \subseteq [0, K_m)$ and $A^c \subseteq [0, q_m)$ we denote their densities by:

$$\begin{aligned} d_m(A) &= |A|/K_m \\ d_m^c(A^c) &= |A^c|/q_m \end{aligned}$$

Then d_m and d_m^c can be viewed as discrete probability measures on the sets $[0, K_m)$ and $[0, q_m)$ respectively.

Lemma 87. *Let $n \leq m$, $w \in \mathcal{W}_m$ and $w^c =_{def} c_m(w) \in \mathcal{W}_m^c$. We view w as sitting on the interval $[0, K_m)$ and w^c as sitting on $[0, q_m)$ Let S^* be a collection of (n, m)-genetic markers, g the total number of (n, m)-genetic markers and $d = |S^*|/g$. If:*

- *$A = \{k \in [0, K_m) : \text{some } u \in \mathcal{W}_n \text{ with genetic marker in } S^* \text{ begins at } k \text{ in } w\}$*

- *$A^c = \{k \in [0, q_m) : \text{some } u^c \in \mathcal{W}_n^c \text{ with genetic marker in } S^* \text{ begins at } k \text{ in } w^c\}$,*

then the following equations hold:

$$d_m(A) = \frac{d}{K_n} \tag{29}$$

$$d_m^c(A^c) = \frac{d}{q_n} \prod_{p=n}^{m-1}(1 - 1/l_p) \tag{30}$$

$$d_m(A) = \left(\frac{d_m^c(A^c)}{\prod_{p=n}^{m-1}(1 - 1/l_p)} \right) \left(\frac{q_n}{K_n} \right) \tag{31}$$

$$d_m^c(A^c) = d_m(A) \left(\prod_{p=n}^{m-1}(1 - 1/l_p) \right) \left(\frac{K_n}{q_n} \right). \tag{32}$$

$\vdash$ We prove equation 30. Equation 29 is similar but easier. The other two equations follow algebraically.

The union of the boundary regions ∂_p for $p = n$ to $m-1$ consist exactly of the elements of $[0, q_m)$ that are not part of any n-word. We denote the complement of $\bigcup_{p=n}^{m-1} \partial_p$ by $(\bigcup_{p=n}^{m-1} \partial_p)^{\backprime}$. The various ∂_p are pairwise disjoint and for each n^*, $(\bigcup_{p=n^*}^{m-1} \partial_p)^{\backprime}$ consists of the locations of entire n^*-words. Starting with $p = m-1$, iteratively deleting boundary sections as p decreases to n, and using Lemma 50 we see that the d_m^c-measure of $(\bigcup_{p=n}^{m-1} \partial_p)^{\backprime}$ is $\prod_{p=n}^{m-1}(1 - 1/l_p)$.

Let $B = \{k \in [0, q_m) : k$ is at the beginning of an n-word$\}$. Then B consists of a $1/q_n$ portion of the regions made up of n-words; i.e. $(\bigcup_{p=n}^{m-1} \partial_p)^{\backprime}$. We note that $A^c \subseteq B$ and B is disjoint from $\bigcup_{p=n}^{m-1} \partial_p$.

By Lemma 86, the number C_1 of n-words occurring in w^c with a given genetic marker does not depend on the marker. Let C_2 be the total number

of n-words occurring in w^c. Then:

$$\begin{aligned}
\frac{|A^c|}{|B|} &= \frac{|\{n\text{-words with genetic marker in } S^*\}}{C_2} \\
&= \frac{|S^*| * C_1}{g * C_1} \\
&= d.
\end{aligned}$$

We compute conditional expectations to get equation 30:

$$\begin{aligned}
d^c_m(A^c) &= d^c_m(A^c \mid (\bigcup_{p=n}^{m-1} \partial_p \curlyvee)) \, d_m((\bigcup_{p=n}^{m-1} \partial_p \curlyvee)) \\
&= d^c_m(A^c \mid B, (\bigcup_{p=n}^{m-1} \partial_p \curlyvee)) \, d_m(B \mid (\bigcup_{p=n}^{m-1} \partial_p \curlyvee)) \, d_m((\bigcup_{p=n}^{m-1} \partial_p \curlyvee)) \\
&= d \left(\frac{1}{q_n}\right) \prod_{p=n}^{m-1}(1 - 1/l_p)
\end{aligned}$$

Equation 29 is similar and 31, 32 follow from the first two equations by substitution. $\dashv$

The following relationship between pairs of measures ν on $\mathbb{K}$ and ν^c on $\mathbb{K}^c$

$$\nu^c(\langle c_n(u) \rangle) = \left(\frac{K_n}{q_n}\right) \nu(\langle u \rangle) \left(1 - \sum_n^\infty \nu^c(\partial_m)\right)$$

is the limit of equation 32 as m goes to infinity. This relationship will hold for a correspondence between measures that we build in forthcoming sections.

We note that since ∂_m has a density that depends only on the circular coefficient sequence, the measures of ∂_m is the same for all invariant measures. If we set d^{∂_n} be this density, then we can rewrite the previous equation as:

$$\nu^c(\langle c_n(u) \rangle) = \left(\frac{K_n}{q_n}\right) \nu(\langle u \rangle) \left(1 - \sum_n^\infty d^{\partial_n}\right) \tag{33}$$

A consequence of equation 33 is that for all basic open sets u, $\nu(\langle u \rangle)$ determines $\nu^c(\langle c_n(u) \rangle)$ and vice versa.

For counting arguments the following inequalities will be helpful.

Lemma 88. *Let n be a number greater than 0. Then there are constants K_n^U, K_n^L between 0 and 1 such that for all $k > 0$ and $w^c \in \mathcal{W}_{n+k}^c$ and all collections S^* of $(n, n+k)$-genetic markers,*
   if

$$A^c = \{i : i \text{ is the location of a start of an } n\text{-subword of } w^c \text{ indexed in } S^*\}$$

then

$$K_n^L |S^*| \leq \left(\frac{|A^c|}{q_{n+k}} \right) \left(\prod_{m=0}^{n+k-1} k_m \right) \leq K_n^U |S^*| \tag{34}$$

⊢   By equation 25 there are

$$|A^c| = |S^*| * \prod_{m=0}^{k-1} q_{n+m}(l_{n+m} - 1)$$

many i that occur at the beginning of occurrences of n-subwords with genetic markers in S^*. Since

$$q_{n+k} = k_n l_n q_n^2 \left(\prod_{m=1}^{k-1} k_{n+m} l_{n+m} q_{n+m} \right)$$

we have:

$$\frac{|A^c|}{q_{n+k}} = |S^*| * \left(\frac{1}{q_n} \right) \left(\prod_{m=1}^{k-1} (1 - \frac{1}{l_{n+m}}) \right) \left(\frac{1}{\prod_{m=0}^{k-1} k_{n+m}} \right).$$

Since the $\langle 1/l_n \rangle$ is a summable sequence, $\prod_{m=1}^{k-1}(1 - \frac{1}{l_{n+m}})$ converges as k goes to ∞. The inequality 34 follows.   ⊣

   Since $K_{n+k} = \prod_{m=0}^{n+k-1} k_m$, inequality 34 can be rewritten as:

$$K_n^L \frac{|S^*|}{K_{n+k}} \leq \frac{|A^c|}{q_{n+k}} \leq K_n^U \frac{|S^*|}{K_{n+k}} \tag{35}$$

Infinite genetic markers: Suppose that we are given a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ for an odometer based or circular system $\mathbb{K}$, $s \in S$ and an occurrence of an n-word u in s. Then we can inductively define an infinite sequence of words $\langle u_m : n \leq m \in \mathbb{N} \rangle$, letting $u_n = u$, and u_{m+1} to be

64

the $m + 1$-subword of s that contains u_m. For each $n < m$ we get a genetic marker $\langle j_n^*, j_{n+1}^*, \ldots j_{m-1}^* \rangle$, and these cohere as m goes to infinity. We define the *infinite genetic marker* to be $\vec{j^*} = \langle j_m^* : n \leq m \in \mathbb{N} \rangle$.

If an n-word u occurs inside an occurrence of an m-word v in s, then $v = u_m$. Thus their infinite genetic markers agree on the tail $\langle j_i^* : m \leq i \in \mathbb{N} \rangle$.

As in Remark 16, if we are given a sequence of words $\langle u_m : n \leq m \rangle$, with $u_m \in \mathcal{W}_m$, and an infinite sequence $\langle j_m : n \leq m \rangle$ such that the genetic marker j_m denotes an instance of u_m in u_{m+1} then we can find an $s \in \mathbb{K}$ with $\langle u_m : m \geq n \rangle$ as a tail of its principal subwords. If $\mathbb{K}$ is odometer then s is unique up to a shift of size less than or equal to K_m. A similar statement holds for circular systems.

5.2 TU and UT.

To understand the relationships between $\mathbb{K}$ and $\mathbb{K}^c$, we define maps $TU : S \to S^c$ and $UT : S^c \to S$ where $S \subseteq \mathbb{K}$ and $S^c \subseteq \mathbb{K}^c$ are as in definition 10. The map TU will be one-to-one but UT will not, in general it is continuum-to-one. Nevertheless $UT \circ TU$ will be the identity map.

We begin by considering a element $s \in S$. Let u_n be the principal n-subword of s. The sequence $\langle u_n : n \in \mathbb{N} \rangle$ determines a sequence of circular words $\langle u_n^c : n \in \mathbb{N} \rangle$ which we assemble to define $TU(s)$. Let $\vec{j} = \langle j_n : n \in \mathbb{N} \rangle$ be the infinite genetic marker of $s(0)$. To describe $TU(s)$ completely we need to define $\langle r_n^c : n \in \mathbb{N} \rangle$. Set $r_0^c = 0$, and inductively define r_{n+1}^c to be the $(r_n^c)^{th}$ position in the first occurrence of an n-word with genetic marker j_n in u_{n+1}^c. Set $TU(s)$ to be the element of $\mathbb{K}^c$ with principal subwords $\langle u_n^c : n \in \mathbb{N} \rangle$ and location sequence $\langle r_n^c : n \in \mathbb{N} \rangle$.

We define a map UT that associates an element of $\mathbb{K}$ to each element of S^c. Given such an $s^c \in S^c$, let $\langle u_n^c : n \geq N \rangle$ be its sequence of principal n-subwords. For each $n \geq N, u_n^c$ occurs as $u_{j_n^*}$ in the preword corresponding to u_{n+1}^c. Let $u_n = c_n^{-1}(u_n^c)$. Then the sequence of words $\langle u_n : n \in \mathbb{N} \rangle$ and genetic markers $\langle j_n^* : n \geq N \rangle$ determine an element of $s \in \mathbb{K}$ except for the location of 0 in the double ended sequence. (The sequence is double ended because $s \in S^c$.)

We determine this location arbitrarily in a manner that makes the sequence of u_n's the principal n-blocks of s $(n \geq N)$ and the j_n^* the sequence of genetic markers of these n-blocks. Let $\bar{0}$ be a sequence of zeros of length N. Then $\bar{0}^\frown \langle j_n^* : n \geq N \rangle$ is a well-defined member of the odometer $\mathcal{O}$ associated with $\mathbb{K}$. From equation 24, $\bar{0}^\frown \langle j_n^* : n \geq N \rangle$ determines a sequence

$\langle r_n : n \in \mathbb{N} \rangle$. Thus by Lemma 15, the pair $\langle u_n : n \geq N \rangle$ and $\bar{0}^\frown \langle j_n^* : n \geq N \rangle$ determines a unique element s of $\mathbb{K}$ which we will denote $UT(s^c)$. It is easy to check that $UT \circ TU = id$ and that for each $s \in S$, there is a perfect set of s^c with $UT(s^c) = s$.

We can get more precise information about correspondences between $\mathbb{K}$ and $\mathbb{K}^c$ by noting that if we are given a sequence $\langle u_n : n \in \mathbb{N} \rangle$ of principal subwords of an $s \in S$, the genetic markers $\langle j_n : n \in \mathbb{N} \rangle$ define an element s^c of $\mathbb{K}^c$ up to a choices $(s^c)^\pi \in \mathcal{K}$. Specifically, suppose that $s^* \in \mathcal{K}$ is such that the infinite genetic marker of $s^*(0)$ is $\langle j_n : n \in \mathbb{N} \rangle$. Then there is an $s^c \in \mathcal{K}^c$ that has a sequence of principal n-blocks $\langle u_n^c : n \in \mathbb{N} \rangle$.

The following lemma will be useful for understanding joinings.

Lemma 89. *Let $s \in S$. Then $\{TU(sh^k(s)) : k \in \mathbb{Z}\} \subseteq \{sh^k(TU(s)) : k \in \mathbb{Z}\}$. If $s \in S$, $s^c = TU(s)$ and $u \in \mathcal{W}_n$, then there is a canonical correspondence between occurrences of u in s and finite regions of s^c where u^c occurs. The occurrences of u^c in these finite regions have the same infinite genetic marker $\langle j_m : m > n \rangle$ in s^c as u does in s.*

$\vdash$ Given an $s \in S$ and a k, the shift $sh^k(s)$ and s have a tail of the principal n-blocks $\langle u_n : N \leq n \rangle$ in common. Moreover the genetic markers associated with this tail are the same for both s and $sh^k(s)$. It follows that $TU(sh^k(s))$ is a shift of $TU(s)$.

We can describe the correspondence as follows. If u occurs in s at k, then u is the principal n-word of $sh^k(s)$. Choose an N so large that some N-word u^* is the principal N-word of both s and $sh^k(s)$. Then $(u^*)^c$ is the principal N-block of s^c. Let $\vec{j}$ be the genetic marker of the occurrence of u (at k) in u^*. The region of s^c corresponding to this occurrence of u is the collection of occurrences of u^c with the genetic marker $\vec{j}$ in the principal N-block of s^c.$\dashv$

5.3 Transferring measures up and down, I

In this section we develop the tool we need for lifting measures on $\mathbb{K}$ to measures on $\mathbb{K}^c$. This will also allow us to establish a one-to-one correspondence between synchronous joinings on odometer systems and synchronous joinings on the corresponding circular systems. Throughout this section we will use π to denote either the projection of an odometer based system to its canonical odometer factor or a circular system to its canonical circular factor.

We begin with a proposition relating sequences of words in a construction sequence for an odometer based system to sequences of words in a construction sequence for a circular system.

Proposition 90. *Let $\langle v_n : n \in \mathbb{N} \rangle$ be a sequence with $v_n \in \mathcal{W}_n$. Let $v_n^c = c_n(v_n)$. Then:*

1. *$\langle v_n : n \in \mathbb{N} \rangle$ is an ergodic sequence iff $\langle v_n^c : n \in \mathbb{N} \rangle$ is an ergodic sequence.*

2. *$\langle v_n : n \in \mathbb{N} \rangle$ is a generic sequence for a measure ν iff $\langle v_n^c : n \in \mathbb{N} \rangle$ is a generic sequence for a measure ν^c. In case either sequence is generic, the measures ν and ν^c satisfy equation 33.*

$\vdash$ Both parts follow immediately from the definitions using equations 27 and 28 to relate the frequencies of k-words $w \in \mathcal{W}_k$ in n-words $u \in \mathcal{W}_n$, for $k < n$ to the frequencies of $c_k(w)$ in the corresponding $c_n(u)$. Equation 33 follows from the Ergodic Theorem and Lemma 87. $\dashv$

We endow that collection of invariant measures on a symbolic system $(\mathbb{K}, sh)$ with the weak* topology.

Theorem 91. *Let $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ be a uniquely readable construction sequence for an odometer based system $\mathbb{K}$ and $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$ be the associated circular construction sequence for $\mathbb{K}^c$. Then there is a canonical affine homeomorphism $\nu \mapsto \nu^c$ between shift invariant measures ν concentrating on $\mathbb{K}$ and non-atomic, shift invariant measures ν^c such that equation 33 holds between ν and ν^c.*

$\vdash$ By Proposition 40 and Lemma 51 we can assume that ν and ν^c concentrate on S and S^c respectively.

We begin by defining the correspondence for ergodic measures. Suppose that we are given an ergodic measure ν and we want to associate a measure ν^c. Let $s \in S$ be a generic point for $(\mathbb{K}, \nu)$. Let $\langle v_n : n \in \mathbb{N} \rangle$ be the sequence of principal n-blocks of s. By Proposition 21 this sequence is generic for ν. By Proposition 90, if we let $v_n^c = c_n(v_n)$, then $\langle v_n^c : n \in \mathbb{N} \rangle$ is an ergodic sequence. Let ν^c be the measure associated with $\langle v_n^c : n \in \mathbb{N} \rangle$. Then ν^c is ergodic and equation 33 holds by Proposition 90.

The other direction is similar, let $s^c \in S^c$ be generic for ν^c. Propositions 21 and 90 imply that if $\langle v_n^c : n \in \mathbb{N} \rangle$ is the sequence of principal n-blocks of

s^c and $v_n = c_n^{-1}(v_n^c)$, then $\langle v_n : n \in \mathbb{N} \rangle$ is ergodic and generic for a measure ν. Again equation 33 holds by Proposition 90.

Suppose now that ν is an arbitrary measure on $\mathbb{K}$. Write the ergodic decomposition of ν as:

$$\nu = \int \nu_i d\mu(i).$$

We define ν^c by

$$\nu^c = \int \nu_i^c d\mu(i)$$

which gives a corresponding measure on $\mathbb{K}^c$. Since equation 33 holds between corresponding ergodic components ν_i and ν_i^c, it holds between ν and ν^c.

By the ergodic decomposition theorem the map $\nu \mapsto \nu^c$ is a surjection. Since the map is invertible, it is a bijection. The map is affine by construction.

It remains to show that it is a homeomorphism. To see that $\nu \mapsto \nu^c$ is weak* continuous it suffices to show that for all $\epsilon > 0$ and $n \in \mathbb{N}$ there is a δ and an m such that for all invariant μ, ν, if for all $u \in \mathcal{W}_m$

$$|\mu(\langle u \rangle) - \nu(\langle u \rangle)| < \delta$$

we know that for all $v \in \mathcal{W}_n$ we have

$$|\mu^c(\langle v^c \rangle) - \nu^c(\langle v^c \rangle)| < \epsilon.$$

But the equation 33 easily implies this taking $m = n$ and

$$\delta < \left(\frac{K_n}{q_n} \right) \left(1 - \sum_{n}^{\infty} d^{\partial n} \right) * \epsilon/4.$$

The argument that the inverse is continuous is the same. $\dashv$

Definition 92. *We will call a pair (ν, ν^c) constructed as in Theorem 91* corresponding measures.

Remark 93. *It follows from Proposition 90 that if ν and ν^c are corresponding measures on $\mathbb{K}$ and $\mathbb{K}^c$ and $s \in \mathbb{K}$ is arbitrary then s is generic for ν iff $TU(s)$ is generic for ν^c. The point s is generic just in case its sequence of principal subwords is generic for ν. By item 2 of Proposition 90, this holds just in case the sequence of principal subwords of $TU(s)$ is generic; i.e. $TU(s)$ is generic.*

We can use Theorem 91 to characterize the possible simplexes of invariant measures for circular systems. By a theorem of Downarowicz ([3], Theorem 5), every non-empty compact metrizable Choquet simplex is affinely homeomorphic to the simplex of invariant probability measures for a dyadic Toeplitz flow. Note that the space of invariant probability measures is always a compact Choquet simplex, hence this theorem is optimal.

Since Toeplitz flows are special cases of odometer based systems it follows immediately that every non-empty compact metrizable Choquet simplex is affinely homeomorphic to the simplex of invariant measures of a 2-symbol odometer based system.

Let K be a compact Choquet simplex and $\mathbb{K}$ an odometer based system having its simplex of invariant probability measures affinely homeomorphic to K. Let $\mathbb{K}^c$ be a circular system corresponding to an odometer based system $\mathbb{K}$. Then the non-atomic measures on $\mathbb{K}^c$ are a Choquet simplex isomorphic to K. There are two additional ergodic measures, the atomic measures concentrating on the constant "b" sequence and on the constant "e" sequence. These two atomic measures are isolated among the ergodic measures.

In the forthcoming [8] we discuss the question of invariant measures further and show that $\mathcal{F}$ preserves several other properties, such as being *rank one*.

6 $\mathbf{P}^-, \mathbf{P}^\natural$, genetic markers and the $\natural$-map

Our goal is to understand the structure of synchronous and anti-synchronous joinings between pairs of ergodic systems $(\mathbb{K}, \mathbb{L}^{\pm 1})$. We will use Theorem 91 to define a bijection between synchronous joinings of odometer based systems and synchronous joinings of circular systems. This is relatively easy: to a joining of $\mathbb{K}$ with $\mathbb{L}$ that projects to the identity we can directly associate an odometer system $(\mathbb{K}, \mathbb{L})^\times$ with a measure ν such that the corresponding measure ν^c on $((\mathbb{K}, \mathbb{L})^\times)^c$ can be identified with a measure on $\mathbb{K}^c \times \mathbb{L}^c$ that projects to the identity. We carry this construction out in detail in section 7 and show that the map $\nu \mapsto \nu^c$ given by Theorem 91 gives a bijection between synchronous joinings of the two kinds of systems.

The situation for anti-synchronous joinings of $\mathbb{K}$ and $\mathbb{L}^{-1}$ is more complicated. In Lemma 43, we remarked that the anti-synchronous joinings of $\mathbb{K}$ and $\mathbb{L}^{-1}$ can be identified with joinings of $\mathbb{K}$ and $\mathrm{rev}(\mathbb{L})$ that concentrate

on $\{(s,t) : \pi s = \pi t\}$. Similarly we can identify the anti-synchronous join-ings of $\mathbb{K}^c$ and $(\mathbb{L}^c)^{-1}$ with joinings of $\mathbb{K}^c$ with $\mathrm{rev}(\mathbb{L}^c)$ that concentrate on $\{(s^c, t^c) : \pi t^c = \natural(\pi s^c)\}$. We give notation for these sets:

1. Let $\mathbf{P}^-$ be the collection of anti-synchronous joinings ρ of $\mathbb{K}$ and $\mathbb{L}^{-1}$.

2. Let $\mathbf{P}^\natural$ be the collection of anti-synchronous joinings ρ^c of $\mathbb{K}^c$ and $(\mathbb{L}^c)^{-1}$.

To understand the relationship between $\mathbf{P}^-$ and $\mathbf{P}^\natural$ we need an analogue of Lemma 87, and the corresponding analogue of equation 33. We now describe the tools we use to do this.

Fix construction sequences for $\langle \mathcal{U}_n : n \in \mathbb{N} \rangle$ and $\langle \mathcal{V}_n : n \in \mathbb{N} \rangle$ for $\mathbb{K}$ and $\mathbb{L}$ respectively based on $\langle k_n : n \in \mathbb{N} \rangle$ and $\mathbb{K}^c, \mathbb{L}^c$ the corresponding circular systems based on $\langle k_n, l_n : n \in \mathbb{N} \rangle$.

Let (s, t) be an arbitrary point in $\mathbb{K} \times \mathbb{L}$ with $\pi t = -\pi s$ and $s \in S^\mathbb{K}, t \in S^\mathbb{L}$. Let $\langle u_n : n \in \mathbb{N} \rangle$ and $\langle v_n : n \in \mathbb{N} \rangle$ be the sequence of principal subwords of s and t respectively. If $s^c = TU(s)$ and $t^c = TU(t)$, then $\langle u_n^c : n \in \mathbb{N} \rangle$ and $\langle v_n^c : n \in \mathbb{N} \rangle$ are the sequences of principal subwords of s^c and t^c.

Let $x = \natural(\pi s^c)$. Then $x \in \mathrm{rev}(\mathcal{K})$ and set $r_n = r_n(x)$.

Definition 94. *Define $\hat{t} \in \mathrm{rev}(\mathbb{L}^c)$ by taking $\langle \mathrm{rev}(v_n^c) : n \in \mathbb{N} \rangle$ as its principal n-subword sequence and $\langle r_n : n \in \mathbb{N} \rangle$ as its location sequence.*

We will study the relationship between $\mathbf{P}^-$ and $\mathbf{P}^\natural$ via the function taking (s, t) to $(s^c, \hat{t})$.

6.1   Genetic Markers revisited

To understand the relationship between joinings ρ in $\mathbf{P}^-$ and ρ^c in $\mathbf{P}^\natural$ we need to take into account the manner that $\natural$ shifts the reverse of the second coordinate of a the image of a generic pair (s, t) for $\mathbb{K} \times \mathbb{L}^{-1}$ and the in-terplay between the map $\natural$ and genetic markers. Let $n < m$. Suppose that $(u', \mathrm{rev}(v'))$ is a pair of n-words coming from $\mathcal{U}_n \times \mathrm{rev}(\mathcal{V}_n)$ that occur aligned inside m-words $(u, \mathrm{rev}(v)) \in \mathcal{U}_m \times \mathrm{rev}(\mathcal{V}_m)$. If u' and $\mathrm{rev}(v')$ occur at the same location in $(u, \mathrm{rev}(v))$, then $\vec{j}_{u'}$ determines $\vec{j}_{v'}$ in the following way:

for $n \leq r < m$ we must have

$$(j_{v'})_r = k_r - (j_{u'})_r - 1 \tag{36}$$

(where $\vec{j}_{u'} = (j_n, j_{n+1}, \dots j_{m-1})$).

Definition 95. *Let $(u', v') \in \mathcal{W}_n$ and $(u, v) \in \mathcal{W}_m$. Define the (n, m)-genetic marker of an occurrence of the pair $(u', \mathrm{rev}(v'))$ in $(u, \mathrm{rev}(v))$ to be $(\vec{j}_{u'}, \vec{j}_{v'})$ where $\vec{j}_{u'}$ is the genetic marker of u' in u and $\vec{j}_{v'}$ is the genetic marker of v' in v.[19] We call $\vec{j}_{u'}$ and $\vec{j}_{v'}$ a* conjugate pair.

Being a conjugate pair is equivalent to satisfying the numerical relationship given in equation 36 and thus either element of a conjugate pair determines the other. Hence for purposes of counting conjugate pairs we need only use the first coordinates, $j_{u'}$.

Let $(u, \mathrm{rev}(v)) \in \mathcal{U}_m \times \mathrm{rev}(\mathcal{V}_m)$ be words that occur in a pair $(s, \mathrm{rev}(t)) \in \mathbb{K} \times \mathrm{rev}(\mathbb{L})$. Then the relative alignment of u^c and $\mathrm{rev}(v^c)$ in $(s^c, \hat{t})$ is determined by the $\natural$-map. This is approximated with a high degree of accuracy by where the code Λ_m sends intervals. Accordingly:

Definition 96. *Define the pair $(u, \mathrm{rev}(v))^c$ to be $(u^c, sh^{A_m}(\mathrm{rev}(v^c)))$.*

Thus $\langle (u, \mathrm{rev}(v))^c \rangle_l$ determines a basic open interval in $\mathbb{K}^c \times \mathrm{rev}(\mathbb{L}^c)$ which we might also write as $(\langle u^c \rangle_l \times \mathrm{rev}(\mathbb{L}^c)) \cap (\mathbb{K}^c \times \langle \mathrm{rev}(v^c) \rangle_{l+A_m})$. Alternatively we could write this as:

$$\{(f, g) \in \mathbb{K}^c \times \mathrm{rev}(\mathbb{L}^c) : f \restriction [l, l + q_m) = u^c \text{ and}$$
$$g \restriction [l + A_m, l + A_m + q_m) = \mathrm{rev}(v^c)\}.$$

We now have a lemma extending Lemma 72 which says that if u and v belong to $\mathcal{U}_{n+1}$ and $\mathcal{V}_{n+1}$ then, relative to $sh^{-j_1}(\mathrm{rev}(v))$, all occurrences of $(u')^c \in \mathcal{U}_n^c$ in u^c are either lined up with an occurrence of a $\mathrm{rev}((v')^c)$ for some $(v')^c \in \mathcal{V}_n^c$ or a boundary section of $sh^{-j_1}(\mathrm{rev}(v^c))$. The lemma also says that if $(u')^c, \mathrm{rev}((v')^c)$ are lined up then $\vec{j}_{u'}$ and $\vec{j}_{v'}$ form a conjugate pair.[20]

Proposition 97. *Let $n < m$ and $u \in \mathcal{U}_m, v \in \mathcal{V}_m$. Then for $u' \in \mathcal{U}_n, v' \in \mathcal{V}_n$ we consider occurrences of $(u', \mathrm{rev}(v'))^c$ in $(u, \mathrm{rev}(v))^c$.[21]*

1. *If $(u', \mathrm{rev}(v'))^c$ occurs in $(u, \mathrm{rev}(v))^c$, then $\vec{j}_{u'}$ and $\vec{j}_{v'}$ form a conjugate pair.*

---

[19]Note that the genetic marker $\vec{j}_{u'}$ denotes a different position inside $\mathrm{rev}(v)$ then it does in u.

[20]In this case both $\vec{j}_{u'}$ and $\vec{j}_{v'}$ are of length one.

[21]Since $A_m \neq A_n$ we are considering different shifts in $(u, \mathrm{rev}(v))^c$ and $(u', \mathrm{rev}(v'))^c$.

2. *There is a constant $C = C(n, m)$ such that all conjugate pairs occur C times.*

3. *Fix a conjugate pair $(j_{u'}, j_{v'})$ of genetic markers of $(u', \mathrm{rev}(v'))$. If k is a location of an occurrence of $(u')^c$ in u^c with genetic marker $\vec{j}_{u'}$, but not a location of $(u', \mathrm{rev}(v'))^c$, then the section of $sh^{A_m}(\mathrm{rev}(v^c))$ in the interval $[k + A_n, k + A_n + q_n)$ is contained in $\bigcup_{i=n+1}^{m} \partial_i$.*

⊢ Item 1 is immediate from the definitions.

The latter items are asking about pairs of the form $((u')^c, sh^{A_n}(\mathrm{rev}((v')^c)))$ occurring in $(u^c, sh^{A_m}(\mathrm{rev}(v^c))$. Such a pair occurs at k if and only if the pair $((u')^c, \mathrm{rev}((v')^c))$ occurs aligned in $(u, sh^{A_m - A_n}(\mathrm{rev}(v^c)))$ at k. Item 3 is equivalent to saying that $(u')^c$ is lined up with a portion of $sh^{A_m - A_n}(\mathrm{rev}(v^c))$ contained in $\bigcup_{i=n+1}^{m} \partial_i$.

We fix m and prove 2 and 3 by induction on $m - n$. The case that $m = n + 1$ is the content of Lemma 72. Suppose that the proposition is true for m and $n + 1$, we prove it for m and n.

A pair of $n + 1$-circular words $(w_0, w_1)^c$ lined up in the shifted pair $(u^c, sh^{A_m - A_{n+1}}(\mathrm{rev}(v^c)))$ must have conjugate genetic markers. Moreover any there is a number C_0 such that any pair with conjugate genetic markers occurs lined up C_0 many times.

Fix an occurrence k of an $n + 1$-word w_0 so that no word in $sh^{A_m}(\mathrm{rev}(v^c))$ occurs at $[k + A_{n+1}, k + A_{n+1} + q_{n+1})$, i.e w_0 is not lined up with the reverse of an $n + 1$-word in $sh^{A_m - A_{n+1}}(\mathrm{rev}(v^c))$. Then w_0 is lined up with a segment of $sh^{A_m - A_{n+1}}(\mathrm{rev}(v^c))$ that is a subset of in $\bigcup_{n+2}^{m} \partial_i$. To pass from $A_m - A_{n+1}$ to $A_m - A_n$ we shift by $-j_1$, where $j_1 = p_n^{-1} \mod q_n$. Noting that each reversed $n + 1$-word ends with a string of b's of length q_n, we see that after the additional shift there can be no n-subwords inside w_0 lined up with anything besides a portion of $sh^{A_m - A_n}(\mathrm{rev}(v^c))$ contained in $\bigcup_{n+1}^{m} \partial_i$.

Suppose that u' and v' are n-words and we have an occurrence of $(u')^c$ and $\mathrm{rev}((v')^c)$ lined up in the pair $(u^c, sh^{A_m - A_n}(\mathrm{rev}(v^c)))$. If $\vec{j}_{u'} = k_0 \frown \vec{j}_{u'}^{*}$ and $\vec{j}_{v'} = k_1 \frown \vec{j}_{v'}^{*}$, we let (w_0, w_1) be the occurrence of $n + 1$-subwords of (u, v) with genetic markers $\vec{j}_{u'}^{*}$ and $\vec{j}_{v'}^{*}$ that contain u' and v'. It follows from the previous paragraph that the genetic markers of w_0 and w_1 are conjugate and $w_0^c, \mathrm{rev}(w_1^c)$ are aligned in $(u^c, sh^{A_m - A_{n+1}}(\mathrm{rev}(v^c)))$. By Lemma 72, k_0 and k_1 are conjugate and thus $\vec{j}_{u'}$ and $\vec{j}_{v'}$ are conjugate.

Further each conjugate pair occurs aligned the same number C_1 of times in the pair $(w_0^c, sh^{-j_1}(\mathrm{rev}(w_1^c)))$. The number C_1 is independent of w_0, w_1

72

and k_0 and k_1. It follows now that given a conjugate pair of genetic markers $(\vec{j}_{u'}, \vec{j}_{v'})$, the number of occurrences of a pair of circular n-words with genetic marker $\vec{j}_{u'}$ in u^c aligned with an occurrence of a circular word with genetic marker $\vec{j}_{v'}$ is in v^c is $C_0 * C_1$.

To finish we note that the unaligned n-words are in two categories, those that are not aligned because the $n+1$-words that contain them are not aligned, or those that are not aligned by the final shift $-j_1$. In each case, the unaligned n-words in u occur across from boundary sections in the word $sh^{A_m - A_n}(\text{rev}(v^c))$. ⊣

Thus, using the backwards $\mathcal{C}$-operation to wrap words around the circle in opposite directions introduces some *slippage*, but the slippage is uniform and predictable.

Definition 98. *Suppose that $\vec{j}$ and $\vec{j}'$ are a conjugate pair of (n,m)-genetic markers and $u^c \in \mathcal{U}_m^c, v^c \in \mathcal{V}_m^c$. Let $(u')^c$ and $(v')^c$ have genetic markers $\vec{j}$ and $\vec{j}'$ in u^c, v^c respectively. Then the set of locations k such that $(u')^c$ occurs in u^c starting at k with genetic marker $\vec{j}$ but $\text{rev}((v')^c)$ does not occur starting at $k + A_n$ in $sh^{A_m}(\text{rev}(v^c))$ is called the (n,m)-slippage of $\vec{j}$.*

A location k can belong to the slippage of $\vec{j}$ for two mutually exclusive reasons. Either, for some proper tail segment $\vec{j}^*$ of $\vec{j}$, k is part of the slippage of the subword of u^c with genetic marker $\vec{j}^*$ or k is part of the slippage of the j_n inside the $n+1$ word containing u caused by sh^{-j_1}.

Let $SL_{n,m}$ stand for the (n,m)-slippage of n-subwords of u^c; i.e. the locations k in u^c of some n-word $(u')^c$ such that there there is no n-word $\text{rev}((v')^c)$ at position $k + A_m$. Inside an m-word u^c we find multiple copies of $SL_{n,n+1}$ corresponding the location of each $n+1$ word in u^c. Denote the union of these copies as $SL_{n,n+1}^m$. Then it follows that:

$$SL_{n,m} = \bigcup_{k=n}^{m-1} SL_{k,k+1}^m \cap \{\text{locations of } n\text{-words}\} \tag{37}$$

and moreover the union is disjoint.

The slippage is the portion of of the words that we have no control over when counting, so we want to be able to estimate the proportion of words in the slippage. Let

$$\varpi_n^m = \frac{|SL_{n,m}|}{|\{n - \text{subwords of } u^c\}} \tag{38}$$

The next proposition allows us to control the (n, m)-slippage by controlling the successive $(n, n+1)$-slippages.

Proposition 99.

$$1 - \varpi_n^m = \prod_n^{m-1} (1 - \varpi_i^{i+1}). \tag{39}$$

⊢ We begin by noting that for n^* between n and m, all pairs $(u^*, \text{rev}(v^*))$ of n^*-words have the same proportion of slippage of n-words in $(u^*, \text{rev}(v^*))^c$. Thus $\varpi_n^{n^*}$ is equal to the proportion of slippage of all of the n-words occuring in pairs $(u^*, \text{rev}(v^*))^c$ of n^*-subwords of $(u, \text{rev}(v))^c$.

The argument is similar to Lemma 87. Starting with $n^* = m - 2$ and decreasing until $n^* = n + 1$, using that fact that the union in equation 37 is disjoint, one inductively demonstrates that:

$$(1 - \varpi_n^m) = (1 - \varpi_n^{n^*}) \prod_{n^*}^{m-1} (1 - \varpi_i^{i+1}).$$

⊣

We can combine item 3 of Lemma 97 with equation 39 to see that if k is in $SL_{n,m}$, then $[k + A_n, k + A_n + q_n)$ is a subset of $\bigcup_{i=n+1}^m \partial_i^{v^c}$. It thus follows from Lemma 75 that:

$$1 - \varpi_n^m \geq \prod_n^{m-1} (1 - \frac{2}{(l_i - 1)}). \tag{40}$$

Because the definition of ϖ_n^m was made entirely in terms of genetic markers, the whole discussion could have been carried out simply by considering $\mathcal{K}^c \times \text{rev}(\mathcal{K}^c)$. The numerics depend only on the circular coefficient sequence, not on particular construction sequences $\langle \mathcal{U}_n, \mathcal{V}_n : n \in \mathbb{N} \rangle$.

Viewing the operator $\natural$ as the limit of the codes Λ_m, we can pass to infinity and define SL_n^∞ similarly and let ϖ_n^∞ be the proportion of locations k of n-subwords of a typical $s \in \mathbb{K}^c$ such that no n-subword of $\natural(\text{rev}(\pi(s)))$ occurs at $k + A_n$.

Then:

$$
\begin{aligned}
(1 - \varpi_n^\infty) &= \prod_n^\infty (1 - \varpi_i^{i+i}) \\
&\geq \prod_n^\infty (1 - 2/l_i) \qquad (41) \\
&> 0.
\end{aligned}
$$

It follows that $\sum_1^\infty \varpi_i^{i+1} < \infty$.

We now formulate and prove the version of Lemma 87 involving the $\natural$ map. One might expect that would require considering arbitrary pairs of genetic markers $\vec{j}$ and $\vec{j'}$. However, by Proposition 97, if u' occurs in u with (n, m)-genetic marker $\vec{j}$, then the only genetic marker it can occur lined up with in $\mathrm{rev}(v)$ is its conjugate pair. Similarly either of the genetic markers of aligned words $(u')^c$ occurring in u^c and $sh^{A_n}(\mathrm{rev}((v')^c))$ occurring in $sh^{A_m}(\mathrm{rev}(v))$ determine the other member of the conjugate pair.

It follows that we need only consider pairs $(u', \mathrm{rev}(v'))$ whose genetic markers are conjugate in $(u, \mathrm{rev}(v))$. Since the map $\vec{j}$ to $\vec{j'}$ is a bijection we will refer to either of $\vec{j}$ or $\vec{j'}$ as the genetic marker of a pair $(u', \mathrm{rev}(v'))$ or equivalently $(u', \mathrm{rev}(v'))^c$.

We are reduced to considering sets $S^* \subseteq \{(n, m)\text{-genetic markers}\}$ rather than sets of pairs of genetic markers. Let $n < m$ and let S^* be a set of (n, m)-genetic markers of pairs of n-words in $(u, \mathrm{rev}(v))$. Let

$$
\begin{aligned}
A = \quad &\{k \in [0, K_m) : \text{some } u' \text{ with with genetic marker} \\
&\text{in } S^* \text{ begins at } k \text{ in } u\}
\end{aligned}
$$

and

$$
\begin{aligned}
A^c = \quad &\{k \in [0, q_m) : \text{for some } u' \text{ with genetic marker in } S^*, \\
&\text{there is a } v' \text{ such that } (u')^c \text{ occurs beginning at k in } u^c \\
&\text{and } \mathrm{rev}((v')^c) \text{ occurs beginning at } k + A_n \text{ in } sh^{A_m}(\mathrm{rev}(v^c))\}
\end{aligned}
$$

and define

$$
\begin{aligned}
d_m(A) &= |A|/K_m \\
d_m^c(A^c) &= |A^c|/q_m
\end{aligned}
$$

75

If $(u')^c$ occurs at k in u and $\mathrm{rev}((v')^c)$ occurs at $k + A_n$ in $sh^{A_m}(v)$ then $(u', \mathrm{rev}(v'))^c$ occurs at k in $(u^c, sh^{A_m}(\mathrm{rev}(v^c)))$

Lemma 100. *Let $n < m$ and $(u, v) \in \mathcal{U}_m \times \mathcal{V}_m$. Let S^* be a collection of (n, m)-genetic markers, g the total number of (n, m)-genetic markers[22] and $d = |S^*|/g$. Then (in the notation above):*

$$d_m(A) = \frac{d}{K_n} \tag{42}$$

$$d_m^c(A^c) = \frac{d}{q_n} \prod_{p=n}^{m-1}(1 - 1/l_p)(\prod_{i=n}^{m-1}(1 - \varpi_i^{i+1}) \tag{43}$$

$$d_m(A) = \left(\frac{d_m^c(A^c)}{\prod_{p=n}^{m-1}(1 - 1/l_p)(\prod_{i=n}^{m-1}(1 - \varpi_i^{i+1})} \right) \left(\frac{q_n}{K_n} \right) \tag{44}$$

$$d_m^c(A^c) = d_m(A) \left(\prod_{p=n}^{m-1}(1 - 1/l_p) \right) \left(\prod_{i=n}^{m-1}(1 - \varpi_i^{i+1}) \right) \left(\frac{K_n}{q_n} \right). \tag{45}$$

⊢ The proof is essentially the same as the proof of Lemma 87, indeed the proof of equation 42 *is* the same. Because all genetic markers occur with the same frequency, after allowing for the portions u^c in boundary sections and in slippage (which are disjoint), d/q_n is the density of locations k of occurrences of words with genetic markers in S^*. Once again equations 44 and 45 follow from 42 and 43 by substitution. ⊣

The equation relating $\rho \in \mathbf{P}^-$ and $\rho^c \in \mathbf{P}^\natural$ that corresponds to equation 33 is:

$$\rho^c(\langle(u, \mathrm{rev}(v))^c\rangle) = \left(\frac{K_n}{q_n} \right) \rho(\langle(u, v)\rangle)(1 - \sum_n^\infty \rho^c(\partial_m))(1 - \varpi_n^\infty).$$

Once again $\rho^c(\partial_m)$ is independent of the choice of ρ^c. Setting $d_\rho^{\partial_m} = \rho^c(\partial_m)$, we can write the previous equation as:

$$\rho^c(\langle(u, \mathrm{rev}(v))^c\rangle) = \left(\frac{K_n}{q_n} \right) \rho(\langle(u, v)\rangle)(1 - \sum_n^\infty d_\rho^{\partial_m})(1 - \varpi_n^\infty). \tag{46}$$

---

[22] As before it is easy to check that $g = \prod_n^{m-1} k_i$.

Understanding empirical distributions of joinings along the natural map involves studying how the slippage affects each pair of n-words. Fix $u' \in \mathcal{U}_n, v' \in \mathcal{V}_n$ and $u \in \mathcal{U}_m, v \in \mathcal{V}_m$ where $n < m$. Let the conjugate pair $(\vec{j}, \vec{j'})$ be the genetic marker of $(u', \text{rev}(v'))$ in $(u, \text{rev}(v))$. Then, as remarked earlier $\vec{j'}$ is determined by $\vec{j}$, since they are a conjugate pair. Define $SL_{n,m}(u', \text{rev}(v'))$ to be the collection of locations $k \in SL_{n,m}$ of n-subwords of u^c that have genetic marker $\vec{j}$. Item 2 of Proposition 97 implies that $|SL_{n,m}(u', \text{rev}(v'))|$ is the same for all choices of $(u', \text{rev}(v'))$. Since $SL_{n,m}$ is the union over all possible pairs of $SL_{n,m}(u', \text{rev}(v'))$, we see that

$$
\begin{aligned}
\varpi_n^m \quad &=_{def} \quad \frac{|SL_{n,m}|}{|\{n - \text{subwords of } u^c\}|} \\
&= \quad \frac{|SL_{n,m}(u', \text{rev}(v'))|}{|\{\text{subwords of } u^c \text{ with genetic marker } \vec{j}\}|}
\end{aligned}
\tag{47}
$$

From the definition:

$$
EmpDist_{n,n,A_n}((u, \text{rev}(v))^c)((u')^c, (\text{rev}(v'))^c)
$$

is equal to

$$
\frac{|\{\text{occurrences of } (u', \text{rev}(v'))^c \text{ in } (u, \text{rev}(v))^c\}|}{|\text{for some } (u^*, v^*) \in \mathcal{W}_n \times \mathcal{V}_n, (u^*, \text{rev}(v^*))^c \text{ occurs in } (u, \text{rev}(v))^c\}|}
$$

This in turn is equal to:

$$
\frac{(1 - \varpi_n^m)|\{\text{subwords of } u^c \text{ with genetic marker} \vec{j}\}|}{(1 - \varpi_n^m)|\{n\text{-subwords of } u^c\}|}
$$

which in turn is equal to

$$
EmpDist_{n,n,0}(u, \text{rev}(v))(u', \text{rev}(v')).
$$

For notational convenience we write:

$$
EmpDist(u, \text{rev}(v))(u', \text{rev}(v')) =_{def} EmpDist_{n,n,0}(u, \text{rev}(v))(u', \text{rev}(v'))
$$

and

$$
\begin{aligned}
EmpDist((u, \text{rev}(v))^c)((u', \text{rev}(v'))^c) &=_{def} \\
EmpDist_{n,n,A_n}((u, \text{rev}(v))^c)((u')^c, (\text{rev}(v'))^c).
\end{aligned}
$$

Summarizing:

$$
EmpDist(u, \text{rev}(v))(u', \text{rev}(v')) = EmpDist((u, \text{rev}(v))^c)(u', \text{rev}(v'))^c
\tag{48}
$$

6.2 Transferring measures up and down, II

In this section we describe the correspondence between joinings in $\mathbf{P}^-$ and $\mathbf{P}^\natural$. We do this by considering generic points for the joinings and transferring them up or down.

For the reader's convenience we repeat a definition. Let (s, t) be an arbitrary point in $\mathbb{K} \times \mathbb{L}$ with $\pi t = -\pi s$ and $s \in S^{\mathbb{K}}, t \in S^{\mathbb{L}}$. Let $\langle u_n : n \in \mathbb{N} \rangle$ and $\langle v_n : n \in \mathbb{N} \rangle$ be the sequence of principal subwords of s and t respectively. Then $\langle u_n^c : n \in \mathbb{N} \rangle$ and $\langle v_n^c : n \in \mathbb{N} \rangle$ are the sequences of principal subwords of $s^c = TU(s)$ and $t^c = TU(t)$. If $x = \natural(\pi s^c)$, then $x \in \mathrm{rev}(\mathcal{K})$ and we can set $r_n = r_n(x)$. Recall that we defined $\hat{t} \in \mathrm{rev}(\mathbb{L}^c)$ by taking $\langle \mathrm{rev}(v_n^c) : n \in \mathbb{N} \rangle$ as its principal n-subword sequence and $\langle r_n : n \in \mathbb{N} \rangle$ as its location sequence.

The following follows immediately from equation 27:

Lemma 101. *The sequence t is generic for an invariant measure μ on $\mathbb{L}$ if and only if $\hat{t}$ is generic for an invariant measure μ^* on $\mathrm{rev}(\mathbb{L}^c)$.*

We will study the relationship between $\mathbf{P}^-$ and $\mathbf{P}^\natural$ via the function taking (s, t) to $(s^c, \hat{t})$. If $[a_n, b_n]$ is the location of the principal n-block of s^c, we define w_n^c to be the word $(u_n^c, \hat{t}_n)$ (in the language $\Sigma \times \Lambda$) where $\hat{t}_n = \hat{t} \restriction [A_n + a_n, A_n + b_n]$. Rephrasing this, if $(u_n, \mathrm{rev}(v_n))$ are the principal n-subwords of (s, t) then $w_n^c = (u_n, \mathrm{rev}(v_n))^c$.

Proposition 102. *The sequence $\langle (u_n, \mathrm{rev}(v_n)) : n \in \mathbb{N} \rangle$ is a generic sequence (resp. an ergodic sequence) if and only if $\langle w_n^c : n \in \mathbb{N} \rangle$ is a generic sequence (resp. an ergodic sequence).*

⊢   This follows immediately from equation 48.   ⊣

It is worth remarking that Proposition 102 can be restated in the language of Definition 26 as saying that $\langle (u_n, \mathrm{rev}(v_n), 0) : n \in \mathbb{N} \rangle$ is a generic sequence if and only if $\langle (u_n^c, \mathrm{rev}(v_n^c), A_n) : n \in \mathbb{N} \rangle$ is a generic sequence.

The next theorem is the analogue of Theorem 91 adapted to lifting joinings of $\mathbb{K}$ with $\mathbb{L}^{-1}$ to joining of $\mathbb{K}^c$ with $(\mathbb{L}^c)^{-1}$. In the theorem the notation (ν, ν^c) and (μ, μ^c) refer to pairs of corresponding measures. We assume that $\mathbb{K}$ is built in the language Σ and $\mathbb{L}$ is built in the language Λ.

Theorem 103. *Suppose that $\langle \mathcal{U}_n : n \in \mathbb{N} \rangle$ and $\langle \mathcal{V}_n : n \in \mathbb{N} \rangle$ are construction sequences for two ergodic odometer based systems $(\mathbb{K}, \nu)$ and $(\mathbb{L}, \mu)$ with the*

same sequence parameters $\langle k_n : n \in \mathbb{N}\rangle$. Let $(\mathbb{K}^c, \nu^c)$ and $(\mathbb{L}^c, \mu^c)$ be the associated ergodic circular systems built with a circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$. Then there is a canonical affine homeomorphism $\rho \mapsto \rho^c$ between the simplex of anti-synchronous joinings ρ of $(\mathbb{K}, \nu)$ and $(\mathbb{L}^{-1}, \mu)$ and the simplex of anti-synchronous joinings of $(\mathbb{K}^c, \nu^c)$ and $((\mathbb{L}^c)^{-1}, \mu^c)$ such that equation 46 holds between ρ and ρ^c.

⊢ Suppose that we are given an anti-synchronous ergodic joining ρ between $\mathbb{K}$ and $\mathbb{L}^{-1}$. Let (s, t) be generic for ρ. By lemma 23, the sequence of principal n-blocks, $\langle (u_n, \mathrm{rev}(v_n)) : n \in \mathbb{N}\rangle$ is ergodic. By Proposition 102 the sequence $\langle w_n^c : n \in \mathbb{N}\rangle$ define an ergodic measure ρ^c. Since the $\langle (u_n, \mathrm{rev}(v_n)) : n \in \mathbb{N}\rangle$ satisfy equation 45, the Ergodic Theorem implies that ρ^c and ρ satisfy equation 46. It is easy to check that the definition of ρ^c is independent of the choice of the generic pair (s, t).

For the other direction we can assume that we are given a generic pair $(s^c, \hat{t})$ for an ergodic measure ρ^c on $\mathbb{K}^c \times \mathrm{rev}(\mathbb{L}^c)$ that concentrates on pairs $(s^c, \mathrm{rev}(t^c)) \in \mathbb{K}^c \times \mathrm{rev}(\mathbb{L}^c)$ such that $\pi(\mathrm{rev}(t^c)) = \natural(\pi(s^c))$. Taking principal subwords gives us a generic sequence $\langle (u_n^c, \hat{t}_n) : n \in \mathbb{N}\rangle$. Each $\hat{t}_n$ is a well-defined word $\mathrm{rev}(v_n^c)$ in $\mathrm{rev}(\mathcal{V}_n^c)$.

As in the definition of UT the pair $(s^c, \mathrm{rev}(\hat{t}))$ gives a pair of sequences of genetic markers $(\langle j_n : n \geq N\rangle, \langle j'_n : n \geq N\rangle$ for some N. Letting $u_n = c_n^{-1}(u_n^c)$ and $v_n = c_n^{-1}(\mathrm{rev}(\hat{t}_n))$ the sequences $\langle u_n, j_n\rangle$ and $\langle v_n, j'_n\rangle$ determine a pair in $\mathbb{K} \times \mathbb{L}$ up to finite translations. These sequences are defined independently of the exactly location of the zero of $\hat{t}$; the small shifts used in the definition of $\natural$ do not change the two sequences.

If we let $(s, t) = (UT(s^c), UT(\mathrm{rev}(\hat{t})))$, making small adjustments if necessary to make (s, t) anti-synchronous, we get an element of $\mathbb{K} \times \mathbb{L}^{-1}$. Applying Proposition 102 again we see the theorem.

We can extend this correspondence to non-ergodic joinings ρ on $\mathbb{K} \times \mathbb{L}^{-1}$ and ρ^c on $\mathbb{K}^c \times \mathrm{rev}(\mathbb{L}^c)$, exactly as in Theorem 91; to go up we take an ergodic decomposition of ρ:

$$\rho = \int \rho_i d\mu(i)$$

and define

$$\rho^c = \int \rho_i^c d\mu(i).$$

To go down we use the ergodic decomposition theorem and the measure $\mu(i)$ to reverse this process.

Clearly the map $\rho \mapsto \rho^c$ is an affine bijection. It remains to show that it is continous. However, just as in Theorem 91, we see from equation 46, that for each n there is a constant C_n, independent of ρ such that for all $u \in \mathcal{U}_n, v \in \mathcal{V}_n$,

$$\rho^c(\langle (u, rev(v))^c \rangle) = C_n \rho(\langle (u, v) \rangle).$$

This clearly implies that the map $\rho \mapsto \rho^c$ is a weak* homeomorphism. $\dashv$

The proof of Theorem 103 shows that (s, t) is generic for ρ if and only if the pair $(s^c, \hat{t})$ is generic for ρ^c. Moreover, the proofs of Theorems 91 and 103 are quite robust. In particular the constructions of the corresponding measures are independent of the various choices of generic points s or s^c, (s, t) or $(s^c, \hat{t}\,)$.

7 The Main Result

We now turn to the main results of this paper. Fix an arbitrary circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N} \rangle$ for the rest of the section. Let $\mathcal{OB}$ be the category whose objects are ergodic odometer based systems with coefficients $\langle k_n : n \in \mathbb{N} \rangle$. The morphisms between objects $(\mathbb{K}, \mu)$ and $(\mathbb{L}, \nu)$ will be synchronous graph joinings of $(\mathbb{K}, \mu)$ and $(\mathbb{L}, \nu)$ or anti-synchronous graph joinings of $(\mathbb{K}, \mu)$ and $(\mathbb{L}^{-1}, \nu)$. We call this the *category of odometer based systems*.

Let $\mathcal{CB}$ 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 $(\mathbb{K}^c, \mu^c)$ and $(\mathbb{L}^c, \nu^c)$ will be synchronous graph joinings of $(\mathbb{K}^c, \mu^c)$ and $(\mathbb{L}^c, \nu^c)$ or anti-synchronous graph joinings of $(\mathbb{K}^c, \mu^c)$ and $((\mathbb{L}^c)^{-1}, \nu^c)$. We call this the *category of circular systems*.

Remark 104. *Were we to be completely precise we would take objects in* $\mathcal{OB}$ *to be* presentations *of odometer based systems by construction sequences* $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ *without spacers together with suitable generic sequences and the objects in* $\mathcal{CB}$ *to be* presentations *by circular construction sequences and their generic sequences. This subtlety does not cause problems in the applications so we ignore it.*

The main theorem of this paper is the following:

Theorem 105. *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.*

Elaborating on Example 6:

Corollary 106. *The map $\mathcal{F}$ preserves systems of factor maps (or alternatively extensions). Explicitly: let $\langle I, \leq_I \rangle$ be a partial ordering, $\langle X_i : i \in I\rangle$ be a family of odometer based systems and $\langle \pi_{i,j} : j \leq i\rangle$ is a commuting family of factor maps with $\pi_{i,j} : X_i \to X_j$. Then $\langle \mathcal{F}(\pi_{i,j}) : j \leq i\rangle$ is a commuting family of factor maps among $\langle \mathcal{F}(X_i) : i \in I\rangle$. Moreover the analogous statement holds for circular systems $\langle X_i^c : i \in I\rangle$, factor maps $\langle \pi_{i,j} : j \leq i\rangle$ and $\mathcal{F}^{-1}$.*

Theorem 105 can be interpreted as saying that the *whole* isomorphism and factor structure of systems based on the odometer $\langle k_n : n \in \mathbb{N}\rangle$ is canonically isomorphic to the isomorphism and factor structure of circular systems based on $\langle k_n, l_n : n \in \mathbb{N}\rangle$. We call this a *Global Structure Theorem*.

7.1 The proof of the main theorem

Before we prove theorem 105 we owe the following lemma:

Lemma 107. *Both $\mathcal{O}B$ and $\mathcal{C}B$ are categories, and the composition of synchronous joinings is synchronous, the composition of two anti-synchronous joinings is synchronous and the composition of a synchronous and an anti-synchronous joining (in either order) is anti-synchronous.*

$\vdash$ To see that $\mathcal{O}B$ and $\mathcal{C}B$ are categories we must see that the morphisms are closed under composition. This is equivalent to the statement that the composition of two synchronous or anti-synchronous joinings are synchronous or anti-synchronous. This, in turn follows from Proposition 7 (item 2) applied to joinings of odometers or rotations. $\dashv$

We now prove Theorem 105.
$\vdash$ By Proposition 83 the map $\mathcal{F}$ gives a bijection between the objects of $\mathcal{O}B$ and $\mathcal{C}B$ and hence it remains to define the functor on the morphisms (i.e. joinings between systems $(\mathbb{K}, \mu)$ and $(\mathbb{L}^{\pm 1}, \nu)$) and show that it preserves composition.

7.1.1   Defining $\mathcal{F}$ on morphisms

We split the definition of $\mathcal{F}(\rho)$ into two cases according to whether ρ is synchronous or anti-synchronous. In both cases we define $\mathcal{F}$ for arbitrary joinings even though the only joinings we use as morphisms in the categories are graph joinings; in particular the morphisms in each category are ergodic.

Case 1: ρ is synchronous:

Suppose that ρ a synchronous joining of odometer based systems $\mathbb{K}$ and $\mathbb{L}$ with coefficient sequence $\langle k_n : n \in \mathbb{N} \rangle$ that are constructed with symbols in Σ and Λ from construction sequences $\langle \mathcal{U}_n : n \in \mathbb{N} \rangle$ and $\langle \mathcal{V}_n : n \in \mathbb{N} \rangle$. We define a new construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ with the symbol set $\Sigma \times \Lambda$.

Given n, we put a sequence

$$\langle (\sigma_0, \lambda_0), (\sigma_1, \lambda_1) \ldots (\sigma_{K_n-1}, \lambda_{K_n-1}) \rangle$$

into $\mathcal{W}_n$ if and only there are words $u = (\sigma_0, \ldots \sigma_{K_n-1}) \in \mathcal{U}_n$ and $v = (\lambda_0, \ldots \lambda_{K_n-1}) \in \mathcal{V}_n$.

It is easy to check that $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ is an odometer based construction sequence with coefficients $\langle k_n : n \in \mathbb{N} \rangle$. Let $(\mathbb{K}, \mathbb{L})^{\times}$ be the associated odometer based system. Since ρ is synchronous, it concentrates on members of $\mathbb{K} \times \mathbb{L}$ that correspond to elements of $(\mathbb{K}, \mathbb{L})^{\times}$. We can canonically identify ρ with a shift invariant measure ν on $(\mathbb{K}, \mathbb{L})^{\times}$.

Let $((\mathbb{K}, \mathbb{L})^{\times})^c$ be the circular system associated with $(\mathbb{K}, \mathbb{L})^{\times}$. We can apply Theorem 91 to find shift invariant measure ν^c on $((\mathbb{K}, \mathbb{L})^{\times})^c$ associated with ν that is ergodic just in case ν is ergodic. Shift invariant measures on $((\mathbb{K}, \mathbb{L})^{\times})^c$ can be canonically identified with synchronous joinings on $\mathbb{K}^c \times \mathbb{L}^c$. Let ρ^c be the joining of $\mathbb{K}^c \times \mathbb{L}^c$ corresponding to ν^c. We let $\mathcal{F}(\rho) = \rho^c$.

Explicitly: A generic sequence $\langle (u_n, v_n, 0) : n \in \mathbb{N} \rangle$ for the joining ρ, can be viewed as a generic sequence $\langle (u_n, v_n) : n \in \mathbb{N} \rangle$ for $(\mathbb{K}, \mathbb{L})^{\times}$ and transformed into a generic sequence $\langle (u_n^c, v_n^c) : n \in \mathbb{N} \rangle$ for $((\mathbb{K}, \mathbb{L})^{\times})^c$. The latter corresponds to a generic sequence of the form $\langle (u_n^c, v_n^c, 0) : n \in \mathbb{N} \rangle$ for the joining ρ^c. This process is clearly reversible so $\mathcal{F}$ is a bijection between the synchronous joinings of $\mathcal{O}B$ and the synchronous joinings of $\mathcal{C}B$.

We must show that if ρ is a graph joining then so is ρ^c. Once this is established it follows by symmetry that if ρ is an isomorphism then ρ^c is an isomorphism. Namely if ρ^* is the adjoint joining of $\mathbb{L}$ with $\mathbb{K}$ defined as $\rho^*(A) = \rho(\{(s, t) : (t, s) \in A\})$, then $(\rho^*)^c = (\rho^c)^*$. Hence ρ^* is a graph joining iff $(\rho^c)^*$ is a graph joining.

Suppose that ρ is a graph joining. We apply Proposition 4, part 3. It suffices to show that for all basic open sets in $\mathbb{K}^c$ of the form $\langle u^c \rangle_0$ where $u^c \in \mathcal{U}_n^c$ and all $\epsilon > 0$, there are words $v_1^c, v_2^c \ldots v_{k*}^c$ that belong to $\bigcup_n \mathcal{V}_n^c$ and locations $l_1^c, \ldots l_{k*}^c$ such that:

$$\rho^c((\langle u^c \rangle_0 \times \mathbb{L}^c) \Delta (\mathbb{K}^c \times \bigcup \langle v_j^c \rangle_{l_j^c})) < \epsilon. \tag{49}$$

Consider u such that $c_n(u) = u^c$. Because ρ is a graph joining, for all $\delta > 0$ we can find words $v_1, \ldots v_{k'}$ and locations $l_1, \ldots l_{k'}$ such that

$$\rho((\langle u \rangle_0 \times \mathbb{L}) \Delta (\mathbb{K} \times \bigcup_{i \leq k'} \langle v_i \rangle_{l_i})) < \delta. \tag{50}$$

Without loss of generality we can assume that for some $m \geq n$ each v_i is an m-word and that each $l_i \leq 0$.

Let (s, t) be generic for ρ and considering the pair $s^c = TU(s)$, $t^c = TU(t)$. Then by Remark 93 (s^c, t^c) is generic for ρ^c. We will choose words v_j^c and locations l_j^c and compute the measure in inequality 49 by computing the density of locations representing points in the symmetric difference. Let

$$
\begin{aligned}
B_0 &= \{k : u \text{ occurs at } k \text{ in } s, \text{ but for no } i \text{ does } v_i \text{ occur in } t \\
&\qquad \text{at } l_i + k\} \\
B_1 &= \{k : \text{for some } i, v_i \text{ occurs in } t \text{ at } k + l_i \text{ but } u \text{ does} \\
&\qquad \text{not occur in } s \text{ at } k\}
\end{aligned}
$$

By inequality 50, $B_0 \cup B_1$ can be taken to have density less than δ.

Given words and locations $\{v_j^c, l_j^c : j \in J\}$ we can define two sets $B_0^c, B_1^c \subseteq \mathbb{Z}$, as follows:

$$
\begin{aligned}
B_0^c &= \{k : u^c \text{ occurs in } s^c \text{ at } k \text{ but for no } j \text{ does } v_j^c \text{ occurs in } t^c \\
&\qquad \text{at } l_j^c + k\} \\
B_1^c &= \{k : \text{for some } j \ v_j^c \text{ occurs in } t^c \text{ at } l_j^c + k \text{ but } u^c \text{ does} \\
&\qquad \text{not occur in } s^c \text{ at } k.\}
\end{aligned} \tag{51}
$$

We need to find the words and locations v_j^c, l_j^c so that the density of $B_0^c \cup B_1^c$ is less than ϵ.

For each i, if $-l_i$ is not the location of the beginning of an n-word in v_i then dropping $\langle v_i \rangle_{l_i}$ reduces the measure of the symmetric difference in inequality 50. Thus, without loss of generality we can assume that for all i, there is an (n, m)-genetic marker $\vec{j}(i)$ coding the location of the n-word in v_i that starts at $-l_i$. Since $B_0 \cup B_1$ has density less than δ, the density of k such that either:

1. u occurs at k but for each i, k is not the position of the beginning of an n-word with genetic marker $\vec{j}(i)$ in an occurrence of v_i or

2. for some i, k *is* the position of the beginning of an n-word with genetic marker $\vec{j}(i)$ in an occurrence of v_i, but u does not occur at k,

has density less than δ.

We are in a position to define the v_j^c and the l_j^c. For each i we define index sets J_i and a collection $\{l_j^c : j \in J_i\}$. We arrange the J_i's so that they are pairwise disjoint and for some k^*, $\bigcup_i J_i = \{j : 1 \leq j \leq k^*\}$. For $j \in J_i$, all of the v_j^c are the same and equal to $c_m(v_i)$. For a fixed i, let $\{-l_j^c : j \in J_i\}$ be the collection of locations of the beginnings of n-subwords of $c_m(v_i)$ that have genetic marker $\vec{j}(i)$.

To compute the density of $B_0^c \cup B_1^c$, it suffices to consider an extremely large M and compute the density of $B_0^c \cup B_1^c$ inside the principal M-subword (w_0^c, w_1^c) of (s^c, t^c). Let (w_0, w_1) be the principal M-subword of (s, t) and $c_M(w_0) = w_0^c$ and $c_M(w_1) = w_1^c$.

We now argue as in Lemma 87. Let d_0 be the density of $B_0 \cup B_1$ in (w_0, w_1) and d_0^c be the density of $B_0^c \cup B_1^c$ in (w_0^c, w_1^c). Among all n-words the proportion d_p that begin with an element of $B_0 \cup B_1$ is $d_0 * K_n$. The density of $k \in \mathbb{Z}$ that start n-words in (s, t) is $(1 - \mu(\bigcup_n^\infty \partial_i))/q_n$. Letting d^* be the density of $k \notin \bigcup_n^M \partial_i$, we see that d^* is bounded away from 0 and 1 independently of M. The proportion d_p^c of circular n-subwords of (w_0^c, w_1^c) that begin with a $k \in B_0^c \cup B_1^c$ is

$$\frac{d_0^c * q_n}{(1 - d^*)}.$$

Since ρ concentrates on $\{(s, t) : \pi(s) = \pi(t)\}$ and ρ^c concentrates on $\{(s^c, t^c) : \pi(s^c) = \pi(t^c)\}$, the n-words with a particular genetic marker in w_0 occupy the position of the same genetic marker in w_1 and similarly for w_0^c and w_1^c. The (n, M)-genetic markers set up a one-to-one correspondence between n

subwords u^* of w_0 and regions of w_0^c that consist of occurrences of $(u^*)^c$ that have the same genetic marker. Each of the regions of w_0^c with the same genetic marker have the same number of n-words in them.

Temporarily call an n-subword of (w_0^c, w_1^c) *bad* if it begins with a k in $B_0^c \cup B_1^c$ and similarly for n-subwords of (w_0, w_1) and $B_0 \cup B_1$. Then the property of being bad is determined by the (n, M)-genetic marker of the n-word: if k is the beginning of n-subword of w_0 with genetic marker $\vec{j}$, and k' is the beginning of an n-subword of w_0^c with the same genetic marker in w_0^c, then $k \in B_0 \cup B_1$ if and only iff $k' \in B_0^c \cup B_1^c$.

It follows the proportion of bad n-subwords of (w_0, w_1) is the same as the proportion of bad subwords of (w_0^c, w_1^c). In otherwords:

$$d_p = d_p^c.$$

It follows that

$$d_0 * K_n = \frac{d_0^c * q_n}{(1 - d^*)}.$$

Thus by taking δ small enough and M large enough we can make d_0 as small as we want, and thus arrange that $d_0^c \ll \epsilon$ as desired.

To finish showing that $\mathcal{F}$ is a bijection between graph joinings in each category and isomorphisms in each category we must also show that if ρ^c is a graph joining then so is ρ. But this is very similar. Given a $u^c \in \mathcal{U}_n^c$, and an $\epsilon > 0$ we can find $v_1^c, \ldots v_{k*}^c$ and locations $l_1^c, \ldots l_{k*}^c$ so that inequality 49 holds. Again we can assume that for some m, for all j, $v_j^c \in \mathcal{W}_m^c$. The numbers $|l_j^c|$ determine locations in v_j^c of beginnings of n-words. We can augment our collection of locations by adding more l_j^c's so that if l is the start of a location in v_j^c that has the same (n, m)-genetic marker as l_j^c, then for some j' we have $l_{j'}^c = -l$ and $v_{j'}^c = v_j^c$. In doing this we do not increase the density of $B_0^c \cup B_1^c$. Reversing the procedure above this gives words $v_j \in \bigcup_n \mathcal{V}_n$ and locations l_j such that the density of $B_0 \cup B_1$ is less than ϵ. (Note the lack of boundary in $\mathbb{K} \times \mathbb{L}$ makes the computation easier by reducing the density of $B_0 \cup B_1$.)

Case 2: ρ is anti-synchronous

On the anti-synchronous joinings we take $\mathcal{F}$ to be the bijection between anti-synchronous joinings of $(\mathbb{K}, \mu)$ with $(\mathbb{L}^{-1}, \nu)$ and of the circular systems $(\mathbb{K}^c, \mu^c)$ with $((\mathbb{L}^c)^{-1}, \nu^c)$ defined in Theorem 103. We show that $\mathcal{F}$ takes anti-synchronous graph joinings to anti-synchronous graph joinings and vice versa. Having done this it will follow by a symmetry argument that $\mathcal{F}$ sends anti-synchronous isomorphisms to anti-synchronous isomorphisms.

Suppose that ρ is an anti-synchronous graph joining; i.e. ρ is a graph joining of $\mathbb{K}$ with $\mathbb{L}^{-1}$ that concentrates on $\{(s,t) : \pi(t) = -\pi(s)\}$. The map $x \mapsto \mathrm{rev}(x)$ projects to the odometer map $\pi(x) \mapsto -\pi(x)$; in particular $\mathrm{rev}(\mathbb{L})$ is based on the same odometer that $\mathbb{L}$ is. By Lemma 43 we can view ρ as a graph joining of $\mathbb{K}$ with $\mathrm{rev}(\mathbb{L})$ that concentrates on $\{(s,t) : \pi(s) = \pi(t)\}$. Similarly we view ρ^c as concentrating on $\mathbb{K}^c \times \mathrm{rev}(\mathbb{L}^c)$.

We must show that for all basic open sets in $\mathbb{K}^c$ of the form $\langle u^c \rangle_0$ where $u^c \in \mathcal{U}_n^c$ and all $\epsilon > 0$, there are words $v_1^c, v_2^c \ldots v_{k^*}^c$ that belong to $\bigcup_n \mathcal{V}_n^c$ and locations $l_1^c, \ldots l_{k^*}^c$ such that:

$$\rho^c(((\langle u^c \rangle_0 \times \mathrm{rev}(\mathbb{L}^c))\Delta(\mathbb{K}^c \times \bigcup \langle \mathrm{rev}(v_j^c) \rangle_{l_j^c})) < \epsilon.$$

Consider u such that $c_n(u) = u^c$. Because ρ is a graph joining for all $\delta > 0$ and all large enough m we can find words $v_1, \ldots v_{k'} \in \mathcal{V}_m$ and locations $l_1, \ldots l_{k'}$ such that

$$\rho(((\langle u \rangle_0 \times \mathrm{rev}(\mathbb{L}))\Delta(\mathbb{K} \times \bigcup \langle \mathrm{rev}(v_i) \rangle_{l_i})) < \delta. \tag{52}$$

Without loss of generality we can assume that each $l_i \leq 0$. We will take m sufficiently large according to a restriction we define later.

Let (s,t) be generic for ρ and let $\hat{t}$ be as in Definition 94. Then $(s^c, \hat{t})$ is generic for ρ^c. We argue as before considering sets:

$$
\begin{aligned}
B_0 \;=\;& \{k : u \text{ occurs at } k \text{ in } s, \text{ but for no } i \text{ does } \mathrm{rev}(v_i) \text{ occur in } \mathrm{rev}(t) \\
& \text{at } l_i + k.\} \tag{53} \\
B_1 \;=\;& \{k : \text{for some } i, \mathrm{rev}(v_i) \text{ occurs in } \mathrm{rev}(t) \text{ at } k + l_i \text{ but } u \text{ does} \tag{54} \\
& \text{not occur in } s \text{ at } k.\}
\end{aligned}
$$

Then inequality 52, shows that $B_0 \cup B_1$ can be taken to have density less than any positive δ.

Given words and locations $\{v_j^c, l_j^c : j \in J\}$ we consider $B_0^c, B_1^c \subseteq \mathbb{Z}$, as follows:

$$
\begin{aligned}
B_0^c \;=\;& \{k : u^c \text{ occurs in } s^c \text{ at } k \text{ but for no } j \text{ does } v_j^c \text{ occurs in } \hat{t} \\
& \text{at } l_j^c + k\} \\
B_1^c \;=\;& \{k : \text{for some } j, v_j^c \text{ occurs in } \hat{t} \text{ at } l_j^c + k \text{ but } u^c \text{ does} \\
& \text{not occur in } s^c \text{ at } k.\}
\end{aligned}
$$

Given $\{(v_i, l_i) : 1 \leq i \leq k'\}$, we need to find the words and locations v_j^c, l_j^c so that the density of $B_0^c \cup B_1^c$ is less than ϵ. As in the synchronous case, for each i we build index sets J_i so that the J_i's to be disjoint and have union the interval $\{j : 1 \leq j \leq k^*\}$ for some k^*. For all $j \in J_i$ we take $v_j^c = c_m(v_i)$. We need to find a collection of locations $\{l_j : j \in J_i\}$.

Fix an $i \leq k'$. Without loss of generality we can assume that l_i is the beginning of a reversed n-block $\text{rev}(v')$ in $\text{rev}(v_i)$, since otherwise, discarding $\langle \text{rev}(v_i) \rangle_{l_i}$ makes inequality 52 sharper. If $(s_0, \text{rev}(t_0)) \in \mathbb{K} \times \text{rev}(\mathbb{L})$ is an arbitrary member of

$$(\langle u \rangle_0 \times \mathbb{L}) \cap (\mathbb{K} \times \langle \text{rev}(v_i) \rangle_{l_i})$$

with $\pi(s_0) = -\pi(t_0)$, then there is an m-word u^* such that $s_0 \in \langle u^* \rangle_{l_i}$. Let $\vec{j}(i)$ be the genetic marker of u in u^*. We note that $\vec{j}(i)$ does not depend on s_0, since it is determined entirely by the location of u in u^* and u^* must be aligned with $\text{rev}(v_i)$.

The genetic marker $\vec{j}(i)$ defines a region of n-words in $\mathcal{U}_n^c$ inside an m-word in $\mathcal{U}_m^c$. Let L_i be the collection of l that are at the beginning of an n-word in $\mathcal{U}_n^c$ with genetic marker $\vec{j}(i)$ in an m-word in $\mathcal{U}_m^c$ and set

$$\{l_j^c : j \in J_i\} = \{A_m - l : l \in L_i\}. \tag{55}$$

This determines the collection $\{v_j^c, l_j^c : 1 \leq j \leq k^*\}$.

We now compute the density of $B_0^c \cup B_1^c$ in terms of the density of $B_0 \cup B_1$. To do this it suffices to consider a large enough M that s^c has a principal M-block $[a_M, b_M)$ and compute densities inside this principal M-block. If this is sufficiently small we can deduce that the density of $B_0^c \cup B_1^c$ is small in $\mathbb{Z}$. By Remark 78, we can also assume that M is so large that $\natural$ restricted to this principal M-block is equal to $\bar{\Lambda}_M$ along this M-block; equivalently the principal M-block of $\hat{t}$ is $[a_M + A_m, b_M + A_M)$.

From Proposition 97, we know that if I is an m-sub-block of $s^c \restriction [a_M, b_M)$ then either:

1. the corresponding sub-block of $\hat{t}$ is at $sh^{A_m}(I)$ or

2. I is part of the (m, M)-slippage.

By item 2 of Proposition 97, the number of m-sublocks in each case that correspond to a given (n, M)-genetic marker does not depend on the genetic marker. Further in the second case $sh^{A_m}(I)$ is entirely part of $\bigcup_{m+1}^M \partial_i(\hat{t})$.

We compute the density d_0^c of elements of $B_0^c \cup B_1^c$ by separating them into these two sources. Explicity, we divide into:

Slippage: Those $k \in B_0^c \cup B_1^c$ that begin an n-subword of a location of an m-subword of s^c that is in the (m, M)-slippage.

Mistakes: those $k \in B_0^c \cup B_1^c$ such that k is the location of the beginning of a circular n-subword inside $s^c \restriction [a_M, b_M)$ and $[k + A_m, k + q_m + A_m)$ is the location of an m-word in $\hat{t}$.

We compute the density of the Mistakes and the Slippage separately. Again we will call n-subwords that begin with elements of $B_0 \cup B_1$ or $B_0^c \cup B_1^c$ *bad*.

Both the Mistakes and the Slippage occur at the beginning of n-subwords of $s^c \restriction [a_M, b_M)$. Define d_b to be density of $\bigcup_{n+1}^M \partial_i$ in $[a_M, b_M)$. Then proportion of $k \in [a_M, b_M)$ that begin n-subwords is:

$$\frac{1 - d_b}{q_n}.$$

Of these a proportion ϖ_m^M of the n-subwords are in the Slippage. Thus the collection of k that belong to the Slippage has density

$$\varpi_m^M \left(\frac{1 - d_b}{q_n} \right).$$

Since ϖ_m^M goes to zero as m goes to infinity we can make this term as small as desired by taking m large enough.

Let $[a_M', b_M')$ be the location of the principal M-block of s (and thus of $\mathrm{rev}(t)$). Let d_0 be the density of $B_0 \cup B_1$ in $[a_M', b_M')$.

Suppose now that k belongs to the Mistakes. Let $\vec{j}$ be the (n, M)-genetic marker of the word beginning with k in $s^c \restriction [a_M, b_m)$. Then there is a unique k' in $[a_M', b_M')$ that is at the beginning of an n-subword of $s \restriction [a_M', b_M')$ and has genetic marker $\vec{j}$. By construction, for k that are not in the Slippage:

$$k \in B_0^c \cup B_1^c \text{ iff } k' \in B_0 \cup B_1. \tag{56}$$

Let d_p be the proportion of m-subwords of $s \restriction [a_M', b_M')$ that begin with a $k \in B_0 \cup B_1$. Since every genetic marker is represented exactly the same number of times in the complement of the slippage (Proposition 97), the proportion of words that begin with k in the Mistakes is

$$d_p^c = d_p * (1 - \varpi_m^M). \tag{57}$$

If d_0 is the density of $B_0 \cup B_1$ in $[a'_M, b'_M)$ and d_0^c is the density of the Mistakes, then

$$d_0 = d_p/K_n \tag{58}$$

$$d_0^c = d_p^c \left(\frac{1 - d_b}{q_n} \right) \tag{59}$$

Putting together equations 57, 58 and 59, we see that if we make d_0 sufficiently small we can make d_0^c as small as desired.

Summarizing: By taking M large enough, the density of $B_0^c \cup B_1^c$ is well approximated by the density of $B_0^c \cup B_1^c$ inside $[a_m, b_m)$. This is the sum of the density of the (m, M) slippage and the density of the Mistakes. We can make the density of the Slippage arbitrarily small by taking m large enough and the density of the Mistakes arbitrarily small by taking δ_0 sufficiently small. This establishes the claim that if ρ is a graph joining then so is ρ^c.

We must show that if ρ^c is a graph joining then so is ρ. We suppose that we are given a $u \in \mathcal{U}_n$, we must find $\{v_i, l_i : i \le k'\}$ so that equation 52 holds. Let $u^c = c_n(u)$ and approximate $\langle u^c \rangle_0 \times \mathrm{rev}(\mathbb{L}^c)$ using $\{v_j^c, l_j^c : i \le k^*\}$. Again, we can assume that the collection of locations is saturated in the sense that if l is the start of a location in v_j^c that has the same (n, m)-genetic marker as l_j^c, then for some j' we have $l_{j'}^c = -l$ and $v_{j'}^c = v_j^c$. In doing this we do not increase the density of $B_0^c \cup B_1^c$. We can now use equations 57, 58 and 59 again to see that if d_0^c is made sufficiently small then so is d_0.

Our next claim is that ρ is an isomorphism if and only if ρ^c is an isomorphism. Recall from Proposition 5 that ρ is an isomorphism iff both ρ and ρ^* are graph joinings. Thus if ρ is an isomorphism, both ρ^c and $(\rho^*)^c$ are graph joinings. Since $\natural$ is an involution:

$$(\rho^*)^c = (\rho^c)^*.$$

Thus if ρ is an isomorphism, so is ρ^c.

Reversing this line of reasoning shows that if ρ^c is a graph joining then ρ is.

7.1.2 $\mathcal{F}$ preserves composition

To finish the proof that $\mathcal{F}$ is a functor we must show that $\mathcal{F}$ preserves composition. The argument splits into four natural cases: composing synchronous

joinings, composing a synchronous joining with an anti-synchronous joining on either side and composing two anti-synchronous joinings. We will carefully work out the case for compositions of synchronous embeddings, and discuss the appropriate modification in the cases involving at least one anti-synchronous embedding after Lemma 108.

The cases differ only that the shifts involved in the generic sequences have different forms. For ergodic synchronous joinings generic sequences can be taken to be of the form $\langle (u_n, v_n, 0) : n \in \mathbb{N} \rangle$, whereas for anti-synchronous joinings of $\mathbb{K}^c$ and $\mathrm{rev}(\mathbb{L}^c)$ a natural generic sequence is of the form $\langle (u_n^c, \mathrm{rev}(v_n^c), A_n) : n \in \mathbb{N} \rangle$.[23]

Preparatory Remarks

In the characterization of the relatively independent joining ρ of ρ_1 and ρ_2 given in Lemma 28 and Proposition 29, the partitions $\mathcal{A}_k, \mathcal{A}'_k$ and $\tilde{\mathcal{A}}_k$ are given by $\langle u_k \rangle_{s_1}, \langle v_k \rangle_{s_2}$ and $\langle w_k \rangle_{s_3}$ for $s_1, s_2, s_3 \in \mathbb{Z}$. Formally the partitions $\mathcal{A}_k \times \mathcal{A}'_k, \mathcal{A}_k \times \tilde{\mathcal{A}}_k$ and $\mathcal{A}'_k \times \tilde{\mathcal{A}}_k$ and $\mathcal{A}_k \times \mathcal{A}'_k \times \tilde{\mathcal{A}}_k$ consist of all possible products of these basic open sets. However, in the situation we are considering we have synchronous and anti-synchronous joinings. For synchronous joinings we can build a generating family for the relatively independent joining ρ of ρ_1 and ρ_2 by considering products of pairs of basic open intervals in the same locations; e.g. pairs of the form $\langle u_k \rangle_s \times \langle w_k \rangle_s$. As a consequence, for verifying the hypotheses of Proposition 29 we can restrict our attention to the case where $s^* = 0$.

In the case of anti-synchronous joinings we need to distinguish the odometer based from the circular systems. For anti-synchronous joinings of odometer based systems $\mathbb{K}$ with $\mathbb{M}^{-1}$ we can consider only intervals of the form $\langle u_k \rangle_s \times \langle \mathrm{rev}(w_k) \rangle_{s+s^*}$ where $s^* = 0$. For anti-synchronous joinings of the circular systems $\mathbb{K}^c$ with $\mathbb{M}^c$, asymptotically the Empirical Distances concentrate on words of the form $\langle u_k^c \rangle \times \langle \mathrm{rev}(w_k^c) \rangle_{A_k}$ (where A_k is the amount of shift for $\natural$ at scale k). Moreover, translations of sets of this form generate the measure algebra of the anti-synchronous joining.

Thus in the proof of the next lemma, to verify the hypothesis 3 of Proposition 29 we can take $s^* = 0$ or $s^* = A_k$ depending on whether $\rho_1 \circ \rho_2$ is synchronous or anti-synchronous.

Fix odometer based systems $\mathbb{K}, \mathbb{L}$ and $\mathbb{M}$ with construction sequences $\langle \mathcal{U}_n : n \in \mathbb{N} \rangle, \langle \mathcal{V}_n : n \in \mathbb{N} \rangle$ and $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ respectively. Let ρ_1 and ρ_2

---

[23]i.e. $\langle (u_n, \mathrm{rev}(v_n))^c : n \in \mathbb{N} \rangle$.

be synchronous graph joinings of $\mathbb{K}$ and $\mathbb{L}$, and $\mathbb{L}$ and $\mathbb{M}$ respectively and ρ their relatively independent joining over $\mathbb{L}$.

Since ρ_1 and ρ_2 are graph joinings so is their composition. Thus the relatively independent joining is ergodic. Hence by Lemma 32 we can find generic sequences for ρ_1, ρ_2 and ρ that satisfy the hypothesis of Proposition 29.

Lemma 108. *Let $\langle (u_n, v_n, w_n, 0, 0) : n \in \mathbb{N} \rangle$ be generic for ρ. Then the sequence $\langle (u_n^c, v_n^c, w_n^c, 0, 0) : n \in \mathbb{N} \rangle$ is generic for the relatively independent joining ρ^c of ρ_1^c with ρ_2^c.*

Assuming the lemma, we show that $\mathcal{F}$ preserves compositions. Corollary 31 shows that $\langle (u_n, w_n, 0) : n \in \mathbb{N} \rangle$ is generic for $\rho_1 \circ \rho_2$. From the way that $\mathcal{F}$ is constructed, if $\nu^c = \mathcal{F}(\rho_1 \circ \rho_2)$, then $\langle (u_n^c, w_n^c) : n \in \mathbb{N} \rangle$ is generic for ν^c (viewed as a measure on a circular system). From Lemma 108 and Corollary 31, we know that $\langle (u_n^c, w_n^c, 0) : n \in \mathbb{N} \rangle$ is generic for $\rho_1^c \circ \rho_2^c$. Hence $\mathcal{F}(\rho_1 \circ \rho_2) = \mathcal{F}(\rho_1) \circ \mathcal{F}(\rho_2)$ as desired.

It remains to prove Lemma 108.

$\vdash$ We claim that $\langle (u_n^c, v_n^c, w_n^c, 0, 0) : n \in \mathbb{N} \rangle$ satisfies the hypotheses of Proposition 29 for the joinings ρ_1^c and ρ_2^c.

The first two hypotheses follow immediately: ρ_1^c and ρ_2^c are constructed by taking the generic sequences $\langle (u_n^c, v_n^c, 0) : n \in \mathbb{N} \rangle$ and $\langle (v_n^c, w_n^c, 0) : n \in \mathbb{N} \rangle$ determined by $\langle (u_n, v_n, 0) : n \in \mathbb{N} \rangle$ and $\langle (v_n, w_n, 0) : n \in \mathbb{N} \rangle$ respectively, and the measures did not depend on the precise generic sequence taken. Hypothesis 3 remains to be shown.

We are given $\epsilon > 0$, k and s^* and need to find $(k')^c$, $G_{(k')^c}^c$ and the I_{v^c}'s so that inequalitites 3a and 3b hold. Since ρ_1^c and ρ_2^c are synchronous, so is the relatively independent joining. By the preparatory remarks can take s^*, the relative location of words in $\mathbb{K}$ and $\mathbb{M}$ to be 0. Since the sequence of $(u_n, v_n, w_n, 0, 0)$'s is generic for the relatively independent product of ρ_1 and ρ_2, we can find $k', N, G_{k'} \subseteq \mathcal{V}_{k'}$ and for each $v \in G_{k'}$ a set $I_v \subset [0, K_{k'})$ such that the conditions in hypothesis 3 hold in the odometer context.[24]

Choose k' so large that the density d_b of the boundary portions of circular k'-words is less than $\epsilon * 10^{-6}$ and so that for each $v \in G_{k'}$, there is an I_v with

$$|I_v| > \left(\frac{1 - (\epsilon * 10^{-6})}{1 - d_b} \right) * K_{k'}.$$

---

[24]For odometer systems, the length of the words in $\mathcal{U}_{k'}, \mathcal{V}_{k'}$ and $\mathcal{W}_{k'}$ is $K_{k'}$, for circular systems the words at stage k' have length $q_{k'}$.

Let $(k')^c = k'$, and $G_{k'}^c = \{v^c : v \in G_{k'}\}$. For each $v^c \in G_{k'}^c$ we define the set $I_{v^c} \subseteq [0, q_{k'})$. Each $I_v \subseteq [0, K_{k'})$ and each $s \in I_v$ has a genetic marker $\vec{j}_s$ in v. We let $I_{v^c} = \{s^c : s^c$ has the same genetic marker in v as some $s \in I_v$ does in $v\}$. Equation 26 implies that

$$\frac{|I_v|}{K_{k'}} = \frac{|I_{v^c}|}{q_{k'}}(1 - d_b)$$

and thus $|I_{v^c}| > (1 - \epsilon)q_{k'}$.

Equation 27 implies that for $v \in G_{k'}$ and all large n,

$$EmpDist(v_n)(v) = EmpDist(v_n^c)(v^c),$$

from which hypothesis 3a follows immediately.

Fix a $v_0^c \in G_{k'}^c$ and an $s^c \in I_{v_0^c}$. Let $v_0 \in G_{k'}$ correspond to v_0^c, and $s \in I_v$ correspond to s^c. Let $(u^c, w^c) \in \mathcal{U}_k^c \times \mathcal{W}_k^c$. To see hypothesis 3b, we need to compute the empirical distributions of $(u^c, w^c), u^c$ and w^c conditioned on v_0^c.

Let A^c be the collection of $((u')^c, v_0^c, (w')^c) \in \mathcal{U}_{k'}^c \times \mathcal{V}_{k'}^c \times \mathcal{W}_{k'}^c$ such that u^c occurs at s^c in $(u')^c$ and w^c occurs at s^c in $(w')^c$. Let B^c be the collection of all $((u')^c, v_0^c, (w')^c) \in \mathcal{U}_{k'}^c \times \mathcal{V}_{k'}^c \times \mathcal{W}_{k'}^c$. Then:

$$EmpDist_{k,k,s^c,s^c}(u_n^c, v_n^c, w_n^c | v_0^c)(u^c, w^c) = \frac{EmpDist_{k'}(u_n^c, v_n^c, w_n^c)(A^c)}{EmpDist_{k'}(u_n^c, v_n^c, w_n^c)(B^c)}. \quad (60)$$

As in the definition of $\mathcal{F}$ in Section 7.1.1, we can view the relatively independent joining ρ on $\mathbb{K} \times_{\mathbb{L}} \mathbb{M}$ as concentrating on a single odometer system $(\mathbb{K}, \mathbb{L}, \mathbb{M})^\times$ and ρ^c, the relatively independent joining of ρ_1^c, ρ_2^c as concentrating on $((\mathbb{K}, \mathbb{L}, \mathbb{M})^\times)^c$, which is canonically isomorphic to $\mathbb{K}^c \times_{\mathbb{L}^c} \mathbb{M}^c$.

In the odometer system $(\mathbb{K}, \mathbb{L}, \mathbb{M})^\times$, consider the set A consisting of those k'-words (u', v_0, w') such that u' and w' have u and v in position s. Then $A^c = \{((u')^c, v_0^c, (w')^c) : (u', v_0, w') \in A\}$. Similarly $B^c = \{((u')^c, v_0^c, (w')^c) : (u', v_0, w') \in B\}$. Equation 27 implies that

$$EmpDist(u_n, v_n, w_n)(A) = EmpDist(u_n^c, v_n^c, w_n^c)(A^c). \quad (61)$$

and

$$EmpDist(u_n, v_n, w_n)(B) = EmpDist(u_n^c, v_n^c, w_n^c)(B^c). \quad (62)$$

Finally noting that

$$EmpDist_{k,k,s,s}(u_n, v_n, w_n | v_0)(u, w) = \frac{EmpDist_{k'}(u_n, v_n, w_n)(A)}{EmpDist_{k'}(u_n, v_n, w_n)(B)}, \quad (63)$$

92

and using equations 60 and 61 we see that

$$EmpDist_{k,k,s^c,s^c}(u_n^c, v_n^c, w_n^c|v_0^c)(u^c, w^c) = \qquad (64)$$
$$EmpDist_{k,k,s,s}(u_n, v_n, w_n|v_0)(u, w).$$

Arguing in the same manner we see:

$$EmpDist_{k,s^c}(u_n^c, v_n^c|v_0^c)(u^c) = EmpDist_{k,s}(u_n, v_n|v_0)(u) \qquad (65)$$
$$EmpDist_{k,s^c}(v_n^c, w_n^c|v_0^c)(v^c) = EmpDist_{k,s}(v_n, w_n|v_0)(v) \qquad (66)$$

Since for large n,

$$\|EmpDist_{k,k,s,}(u_n, v_n, w_n|v_0) - EmpDist_{k,s}(u_n, v_n|v) * EmpDist_{k,s}(v_n, w_n)|v)\|$$
$$< \epsilon,$$

from equations 64, 65 and 66 we get the desired conclusion that

$$\|EmpDist_{k,k,s^c,s^c}(u_n^c, v_n^c, w_n^c|v_0^v)$$
$$-EmpDist_{k,s}(u_n^c, v_n^c|v_0^c) * EmpDist_{k,s}(v_n^c, w_n^c|v_0^c)\|$$

is less than ϵ. ⊣

Lemma 108 holds where one or both of the joinings ρ_1 and ρ_2 are anti-synchronous as well, however the shift coefficients for the circular systems are no longer all 0 but belong to $\{0, \pm A_n\}$ depending on which joinings are anti-synchronous. Similarly $s^* \in \{0, \pm A_k\}$. The argument follows the same path until it reaches equation 61. This equation relies, in turn on equation 27. The analogue of equation 27 for anti-synchronous joinings is equation 48, which in turn carries over to the relatively independent product. The upshot is that equations 64, 65 and 66 hold after applying the appropriate shifts of u_n^c and v_n^c relative to u_n^c.

This finishes the proof of Theorem 105. ⊣

7.2 Weakly-Mixing and Compact Extensions

We now show that $\mathcal{F}$ preserves weakly-mixing and compact extensions. The fact that compact extensions are preserved is due to E. Glasner and we reproduce the proof here with his kind permission.

Proposition 109. *Let $(\mathbb{K}, \mu)$ and $(\mathbb{L}, \nu)$ be ergodic and suppose that ρ and ρ^c are corresponding synchronous joinings determining factor maps*

$$
\begin{aligned}
\pi &: \mathbb{K} \to \mathbb{L} \\
\pi^c &: \mathbb{K}^c \to \mathbb{L}^c.
\end{aligned}
$$

Then $\mathbb{K}$ is a weakly mixing extension of $\mathbb{L}$ (via π) if and only if $\mathbb{K}^c$ is a weakly mixing extension of $\mathbb{L}^c$ (via π^c).

⊢ Recall that if $\pi : X \to Y$ is a factor map from $(X, \mathcal{B}, \mu, T)$ to $(Y, \mathcal{C}, \nu, S)$, then the extension is weakly-mixing if the relatively independent joining $X \times_Y X$ of X with itself over Y is ergodic relative to Y. In case Y is ergodic, this simply means that the relatively independent joining is ergodic.

Suppose that $\mathbb{K}$ and $\mathbb{L}$ are odometer based systems with construction sequences $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ and $\langle \mathcal{V}_n : n \in \mathbb{N} \rangle$ respectively. If ρ is a synchronous factor joining of $\mathbb{K}$ over $\mathbb{L}$, and the extension is weakly-mixing then we can find an ergodic sequence of words $\langle (u_n, v_n, w_n) \in \mathcal{W}_n \times \mathcal{V}_n \times \mathcal{W}_n : n \in \mathbb{N} \rangle$ that is generic for the relatively independent joining of ρ with itself over $\mathbb{L}$, i.e. $\rho \times_{\mathbb{L}} \rho$. This sequence will satisfy the hypotheses of Proposition 29. It follows that the sequence of (u_n^c, v_n^c, w_n^c)'s is also generic for an ergodic measure ν. As we argued in Lemma 108, the (u_n^c, v_n^c, w_n^c)'s also satisfy the hypothesis of Proposition 29. It follows that ν is the relatively independent joining $\rho^c \times_{\mathbb{L}} \rho^c$. Since ν is ergodic ρ^c is weakly mixing.

If, on the other hand the sequence of (u_n, v_n, w_n) is *not* ergodic, then the sequence (u_n^c, v_n^c, w_n^c) is also not ergodic. Hence if ρ^c is weakly-mixing, then ρ is weakly mixing. ⊣

It is immediate from the Furstenberg-Zimmer structure theorem ([11], Chapter 10, Proposition 10.14) that X is a relatively distal extension of Y if and only if there is no intermediate extension Z of Y, with X being a non-trivial weakly-mixing extension of Z. Thus $\mathcal{F}$ takes measure-distal extensions to measure-distal extensions.

What requires more effort to establish is the following:

Proposition 110. *(E. Glasner) The functor $\mathcal{F}$ takes compact extensions to compact extensions.*

⊢ Glasner's proof uses a result proved in the forthcoming [8]: If $(\mathbb{K}, \mu)$ is an ergodic odometer based system the X is a compact group extension of

$(\mathbb{K}, \mu)$ then there is a representation of X as an odometer based system with the same coefficients.

Since X is a compact extension of Y if and only if X is a factor of a compact group extension of Y,[25] it suffices to show that $\mathcal{F}$ takes compact group extensions to compact group extensions.

To prove that $\mathcal{F}$ takes compact group extensions to compact group extensions we use a remarkable theorem of Veech that characterizes group extensions $\pi : X \to Y$ of ergodic systems. The criteria is that every ergodic joining of X with itself that is the identity on Y (i.e. ρ, as a measure, concentrates on those pairs (x_1, x_2) such that $\pi(x_1) = \pi(x_2)$) comes from a graph joining which is an isomorphism of $(X, \mathcal{B}, \mu, T)$ that projects to the identity map on Y.[26]

Explicity, Theorem 6.18, on page 136 of [11] shows that if, in the ergodic decomposition of the relatively independent product $X \times_Y X$, only graph joinings appear, then X is a compact group extension. The converse follows from Proposition 6.15, part 2 in [11], that if X is a compact group extension of Y then every ergodic self-joining of X over Y which is the identity on Y is a graph joining.

The map $\mathcal{F}$ takes ergodic joinings to ergodic joinings, and all graph joinings to graph joinings, and the identity joining to the identity joining. Thus we see it preserves group extensions. $\dashv$

Furstenberg [9] and Zimmer [22] independently showed that for every ergodic system $\mathbb{X}$ there is an ordinal α and a tower of extensions $\langle X_\beta : \beta \leq \alpha \rangle$ such that X_0 is the trivial system, $X_\alpha = X$ and for all $\beta < \alpha$, $X_{\beta+1}$ is a compact extension of X_β, unless $\alpha = \beta + 1$ where X_α is either a compact or a weakly mixing extension of X_β. If there is no compact extension at the end of the tower, then X is *measure-distal* and $\langle X_\beta : \beta < \alpha \rangle$ is a *distal tower* approximating X. The least ordinal such that $\mathbb{X}$ can be represented this way is the *distal height* or *distal order* of $\mathbb{X}$.

Let $(\mathbb{K}, \mu)$ be an odometer based system and consider the odometer factor $\mathcal{O}$. Let $(\mathbb{K}', \mu')$ be the Kronecker factor of $(\mathbb{K}, \mu)$. Then we have

---

[25]See [10] for an explicit statement and proof.
[26]This first appears in [18].

$$(\mathbb{K}, \mu)$$

$$\downarrow \pi_1$$

$$(\mathbb{K}', \mu')$$

$$\downarrow \pi_2$$

$$\mathcal{O}$$

where π_2 may or may not be a trivial factor map. This tower is carried by $\mathcal{F}$ to

$$(\mathbb{K}^c, \mu^c)$$

$$\downarrow \pi_1$$

$$((\mathbb{K}')^c, (\mu')^c)$$

$$\downarrow \pi_2$$

$$\mathcal{R}_\alpha$$

If $\mathbb{K}'$ is a non-trivial extension of $\mathcal{O}$, then Glasner's result tells us that $(\mathbb{K}')^c$ is a compact extension of $\mathcal{R}_\alpha$, but is silent on the issue of whether $(\mathbb{K}')^c$ is discrete spectrum; i.e. we do not know whether $\mathcal{F}$ takes the Kronecker factor of $\mathbb{K}$ to the Kronecker factor of $\mathbb{K}^c$.

Suppose now that $\mathbb{K}$ is given by a finite tower of factors:

$$\mathcal{O} \longleftarrow \mathbb{K}_0 \longleftarrow \mathbb{K}_1 \longleftarrow \cdots \longleftarrow \mathbb{K}_{N-1} = \mathbb{K}$$

where $\mathbb{K}_0$ is the Kronecker factor of $\mathbb{K}$ and for all $i, \mathbb{K}_{i+1}$ is the maximal compact extension of $\mathbb{K}_i$ in $\mathbb{K}$. Then $\mathbb{K}$ is distal of height N. The map $\mathcal{F}$ carries this to a tower of compact extensions

$$\mathcal{R}_\alpha \longleftarrow \mathbb{K}_0^c \longleftarrow \mathbb{K}_1^c \longleftarrow \cdots \longleftarrow \mathbb{K}_{N-1}^c = \mathbb{K}^c$$

From this we see that the distal height of $\mathbb{K}^c$ is either N or $1 + N$.

We do not know an example whether the height of $\mathbb{K}^c$ can be $1 + N$. However the ordinary skew product construction applied to odometers gives examples of distal height n where $\mathcal{O}$ is the Kronecker factor. Hence from our analysis we see that *there are* ergodic circular systems with distal height N for all finite N.

In [2], Beleznay and Foreman proved that for all countable ordinals α there is an ergodic measure preserving transformation T of distal height α. In that construction there are no eigenvalues of the operator U_T of finite order. Hence if we let $\mathcal{O}$ be an odometer with coefficient sequence $\langle k_n : n \in \mathbb{N} \rangle$ going to infinity, $T \times \mathcal{O}$ is an ergodic transformation with distal height α and zero entropy. In the forthcoming [8] we see that this implies that $T \times \mathcal{O}$ can be presented as an odometer based transformation. By the analysis we just gave we see that $(T \times \mathcal{O})^c$ is a circular system with height $1 + \alpha$. In [8] we see that $(T \times \mathcal{O})^c$ can be realized as a smooth transformation. For infinite α, $1 + \alpha = \alpha$, hence we have:

Theorem 111. *Let N be a finite or countable ordinal. Then there is an ergodic measure distal diffeomorphism of $\mathbb{T}^2$ of distal height N.*

7.3   Continuity

Fix a measure space (X, μ). As noted in Section 2.3, we can identify symbolic shifts built from construction sequences with cut-and-stack constructions (whose levels generate X). By fixing a countable generating set in advance, we can make this association canonical. The levels in the cut-and-stack construction give the relationship with arbitrary partitions of X. In this way the usual weak topology on measure preserving transformation of X described in Section 2.1 determines a topology on the presentations of symbolic shifts as limits of construction sequences.

The finitary nature of the maps $\langle c_n : n \in \mathbb{N} \rangle$ that give bijections between words in $\mathcal{W}_n$ and words in $\mathcal{W}_n^c$ easily shows that the map $\mathcal{F}$ is a *continuous* map from the presentations of odometer based systems to presentations of circular systems. Thus we have:

Corollary 112. *The functor $\mathcal{F}$ is a homeomorphism from the objects in $\mathcal{O}B$ to $\mathcal{C}B$.*

For the purposes of the complexity of the isomorphism relation we note:

Corollary 113. *The map $\mathcal{F}$ is a continuous reduction of conjugacy between odometer based systems and circular systems.*

7.4 Extending the main result

In the main result we restricted the morphisms to graph joinings, largely because compositions of graph joinings are ergodic joinings. Unfortunately a composition of ergodic joinings is not necessarily ergodic, and non-ergodic joinings also arise naturally as relatively independent joinings of ergodic joinings. In this section we indicate how to extend our results to the broader categories that include non-ergodic joinings as morphisms. For convenience, we will continue to require that our objects are ergodic measure preserving systems.

Let $\mathcal{O}B^+$ and $\mathcal{C}B^+$ be the categories that have the same objects as $\mathcal{O}B$ and $\mathcal{C}B$, but where the collections of morphisms are expanded to include *all* synchronous and anti-synchronous joinings (rather than just graph joinings).

In Section 7.1.1, the definition of $\mathcal{F}$ included all such joinings ($\mathcal{F}(\rho)$ for a non-ergodic ρ was defined via an ergodic decomposition). Thus without modification we can view $\mathcal{F}$ as a map:

$$\mathcal{F} : \mathcal{O}B^+ \to \mathcal{C}B^+.$$

To show that $\mathcal{F}$ is a morphism between these categories, i.e. to show preserves composition for arbitrary morphisms, we develop a more combinatorial approach to lifting morphisms that coincides with the original definition.

We start by generalizing the notion of a generic sequence of words to include non-ergodic measures. Suppose $\mathbb{K}$ is a symbolic system with a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$. Let μ be a shift invariant measure which we assume is supported on the set $S \subseteq \mathbb{K}$ (where S is given in definition 10). The ergodic decomposition theorem gives a representation of μ as $\int \mu_p d\lambda(p)$, where each μ_p is a shift invariant ergodic measure and λ is a probability measure on a set P parameterizing the ergodic components. For each p, there is a generic sequence of words $\langle w_n^p : n \in \mathbb{N} \rangle$ for the measure μ_p. The main observation is that the set of probability measures on words of a fixed length is compact. Thus for any fixed k and $\epsilon > 0$, we can find a finite set $P_k \subseteq P$ of parameters so that for all p, there is some $p' \in P_k$ with[27]

$$\|\hat{\mu}_k^p - \hat{\mu}_k^{p'}\| < \epsilon. \tag{67}$$

---

[27]The notions of *EmpDist* and $\hat{\mu}_k$ are given in the beginning of Section 2.6.

This gives a partition of the parameter space into sets $\{E_p : p \in P_k\}$ such that inequality 67 holds for all $p' \in E_p$.

Now let n be sufficiently large such that for each $p \in P_k$, we can find an element $w_n^p \in \mathcal{W}_n$ with

$$\|EmpDist_k(w_n^p) - \hat{\mu}_k^p\| < \epsilon. \tag{68}$$

If we denote $\lambda(E_p)$ by $\alpha(p)$, then $\alpha(p) \geq 0$ and $\sum_{p \in P_k} \alpha_p = 1$. It is clear that one can obtain $\hat{\mu}_k$ up to a small error from the finite data $\{(w_n^p, \alpha(p)) : p \in P_k\}$, which is a weighted finite collection of words.

For the symbolic sequences that we are interested in, such as the circular systems, the measure of the spacers is independent of the invariant measure μ (see Section 5.1). This means that for all n, p, the sum $\sum_{w' \in \mathcal{W}_n} \mu_{q_n}^p(\langle w' \rangle)$ is the same. In this context using inequality 68 we can arrange the inequality:

$$\|(\sum_{p \in P_k} \alpha(p) EmpDist_k(w_n^p)) - \hat{\mu}_k\| < \epsilon.$$

The measure λ is defined on the extreme points of the simplex of shift invariant probability measures and if we choose the finite sets P_k to consist of points that lie in the closed support of λ then we an easily ensure that when we go from (k, ϵ) to a (k', ϵ') with $k' > k, \epsilon' < \epsilon$ that $P_{k'} \supseteq P_k$. Taking a sequence $k \to \infty$ and $\epsilon_k \to 0$ with $\sum \epsilon_k < \infty$, we get a set $\{\nu_1, \nu_2, \ldots\}$ of ergodic measures and finite sets $I_k \subseteq I_{k+1}$ of integers with probability measures α_k on I_k such that $(\sum_{i \in I_k} \alpha_k(i) \nu_i)$ converges to μ in the weak* topology.

Definition 114. *Let n_k go monotonically to infinity and $\{(w_{n_k}^i, \alpha_k(i))_k\}$ be a weighted sequence of words as above. Suppose that for each k and $i \in I_k$, $\|EmpDist_k(w_{n_k}^i) - \hat{\nu}_{i,k}\| < \epsilon_k$, then we call $\{(w_{n_k}^i, \alpha_k(i))\}$ a generic sequence for μ.*

We note that for a fixed i, as k varies $\{w_{n_k}^i\}$ is a generic sequence for ν_i–which is one of the ergodic measures in the support of λ.

In a manner exactly analogous to the analysis in Section 2.6, Definition 114 can be extended to products of symbolic systems, allowing for shifting of words in construction sequences.

Restricting our objects to ergodic systems $(X, \mathcal{B}, \mu, T), (Y, \mathcal{C}, \nu, S)$ and $(Z, \mathcal{D}, \tilde{\mu}, \tilde{T})$ allows us to deal with the non-ergodic analogue of the material discussed between Definition 25 and Lemma 32 in a relatively straightforward way which we now discuss.

For the analogue of Proposition 29 in the non-ergodic case let us make the following observation. Fix a non-ergodic joining ρ of X and Y that has ergodic decomposition $\rho = \int \rho^p d\lambda(p)$, where, by the ergodicity of X and Y, each ρ^p is also a joining of X with Y. Fix a k and an $\epsilon > 0$ and a cylinder set determined by a word $u \in \mathcal{W}_k^X$, at location s^* and let ϕ represent its indicator function. For k' large, by the Martingale convergence theorem, there is a subset G of Y of measure close to one such that when we look at the conditional expectation of ϕ with respect to the partition induced by the principal k'-words of $y \in G$, for $\mathcal{A}_{k'}$ and compare it to $\mathbb{E}(\phi|\mathcal{D})$, the error is small.

The element of that partition that contains y is given by a word $v_y \in \mathcal{W}_{k'}^Y$ and a location parameter s_y, and the conditional expectation is:

$$\frac{\rho(sh^{s^*}(\langle u \rangle) \cap sh^{s_y}(\langle v_y \rangle))}{\nu(\langle v_y \rangle)} \tag{69}$$

This easily gives a set $G_{k'} \subseteq \mathcal{W}_{k'}$ with $\hat{\nu}_k(G_{k'}) > 1 - \epsilon$ and a $J_v \subseteq [0, q_{k'})$ such that for $v \in G_{k'}, j \in J_v$, formula 69 gives a good approximation to $\rho_y(sh^{s^*}(\langle u \rangle))$ for most of the $y \in sh^{s_y}(\langle v_y \rangle)$.

If we have a generic sequence of weighted words for ρ, then we can use it to calculate the expression in 69. This observation makes it possible for us to formulate Proposition 29 for non-ergodic joinings.

We are given ergodic systems X, Y, Z and are given construction sequences $\langle \mathcal{U}_n, \mathcal{V}_n, \mathcal{W}_n : n \in \mathbb{N} \rangle$ such that for each n, the words in each $\mathcal{U}_n, \mathcal{V}_n, \mathcal{W}_n$ have the same length. Two joinings ρ_1 of X and Y and ρ_2 of Y and Z are given. The analogue of Proposition 29 is now:

Proposition 115. *Let*

$$\langle \{(u_{n_k}^i, v_{n_k}^i, w_{n_k}^i, s_{n_k}^i, t_{n_k}^i) : i \in I_k\}, \alpha_k \in Prob(I_k) : k \in \mathbb{N} \rangle \tag{70}$$

be a sequence of weighted words and $\sum \epsilon_k < \infty$. Suppose that the following hypothesis are satisfied:

1. *$\langle \{(u_{n_k}^i, v_{n_k}^i, s_{n_k}^i) : i \in I_k\}, \alpha_k) \rangle_k$ is generic for ρ_1,*

2. *$\langle \{(v_{n_k}^i, w_{n_k}^i, t_{n_k}^i) : i \in I_k\}, \alpha_k \rangle_k$ is generic for ρ_2,*

3. *For all ϵ, k, s^* there are k', N and a set $G_{k'} \subset \mathcal{W}_{k'}^Y$ and for each $v \in G_{k'}$ there is a set $J_v \subseteq [0, q_{k'})$ such that*

(a) $\sum_{v \in G_{k'}} EmpDist(v_{n_k})(v) > 1 - \epsilon$

(b) $|J_v| > (1 - \epsilon)q_{k'}$

(c) For all $v \in G_{k'}$ and $s \in J_v$, if $n_k > N$,

$$\| \sum_{i \in I_k} EmpDist_{k_0,k_0,s,s+s^*}(u_{n_k}^i, sh^{s_{k_n}^i}(v_{n_k}^i), sh^{t_{k_n}^i}(w_{n_k})|v)\alpha_k(i) \quad -$$
$$\sum_{i \in I_k} EmpDist_{k_0,s}(u_{n_k}^i, sh^{s_{k_n}^i}(v_{n_k}^i)|v)\alpha_k(i)*$$
$$\sum_{i \in I_k} EmpDist_{k_0,s+s^*}(v_{n_k}^i, sh^{t_{n_k}^i - s_{k_n}^i}(w_{n_k}^i)|v)\alpha_k(i) \| < \epsilon$$

Then the weighted sequence given in 70 is generic for the relatively independent joining $X \times_Y Z$.

The analogues of Corollary 31 and Lemma 32 are easily verified, giving us a characterization of compositions of non-ergodic joinings and the existence of generic sequences satisfying the hypothesis of Proposition 115.

Verifying that $\mathcal{F}$ preserves composition is now straightforward in the manner of Section 7.1.2: the $G_{k'}^c$ and J_{v^c} are constructed in exactly the same way. Checking the conditional distributions of short words relative to longer words (k vs. k') involves counting k'-words, and these are counted using Equation 27 for each component (u_k, v_k, w_k) separately. The weighted average is then preserved.

8 Lagend

In this section we explore the interplay of the geometric, arithmetic and combinatorial aspects of the manner in which $\mathcal{F}$ wraps the odometer based words around the circle. The map $\mathcal{F}$ does not preserve the dynamics of the odometer when transforming it into a rotation, indeed it can't. The shift sh^k of the odometer corresponds to a shift sh^{k^c} of the rotation. The relationship between k and k^c is characterized combinatorially as an optimal wrapping property. The latter is defined in terms of the notion of a *perfect match*. The results in this section can be used to give an alternate proof of the fact that if $(\mathbb{K}, \mu)$ is ergodic then so is $(\mathbb{K}^c, \mu^c)$ that does not use the notion of a generic sequence of words.

Central to our understanding circular systems is the manner in which an s^c had its n-words aligned with n-words in $sh^k(t^c)$. A word u occurs in s^c lined up with a word w in $sh^k(t^c)$ if and only if u occurs at some location l in s and w occurs at $k + l$ in t^c.

Definition 116. *Let $\vec{x}, \vec{y}$ be strings in the language $\Sigma \cup \{b, e\}$ and u, v be words of the same length. A k-match of u and v in $\vec{x}$ and $\vec{y}$ is a location l in the domain of $\vec{x}$ such that u occurs at l in $\vec{x}$ and v occurs at $l + k$ in $\vec{y}$.*

If w_0^c, w_1^c are circular m-words then a perfect match *of u^c, v^c in w_0^c, w_1^c is a k such that there are (n, m)-genetic markers $\vec{j}_u, \vec{j}_v$ such that u^c occurs in w_0^c and v^c occurs in w_1^c with genetic markers $\vec{j}_u$ and $\vec{j}_v$ respectively and k is a match between all occurrences of u^c and v^c with these genetic markers.*

Thus k is a perfect match of u and v if and only if the occurrences of $\vec{j}_u$ in w_0^c are exactly aligned with the occurrences of $\vec{j}_v$ is w_1^c.

We will say that k *is a match between u and v* if there is a location l such that such that k is a match between u and v at l, and that *every k-match is perfect* when k has the property that for every occurrence of a pair of words u^c, v^c in w_0^c, w_1^c, if k is a match between u^c, v^c then k is a perfect match between u^c, v^c. The astute reader will have already recognized that being a match or a perfect match only refers to the genetic markers and the underlying circular factor–thus the actual identities of u^c, v^c, w_0^c and w_1^c are not material–only the locations of the genetic markers.

The notion of a *perfect match* is vacuous for odometer words; for if u, v are odometer n-words and w_0, w_1 are odometer m-words then u, v are the unique pair with a genetic markers $\vec{j}_u$ and $\vec{j}_v$. Moreover, if k matches any pair of n-subwords, k matches every pair of corresponding n-subwords in the overlap of w_0 and $sh^k(w_1)$.

If $k > 0$, then the $n + 1$-subwords of w_1 in the overlap of w_0 and $sh^k(w_1)$ are split into two pieces by the $n + 1$-subwords of w_0; the left portion of each of the $n + 1$-subwords of w_0 in the overlap coincides with the right portion of the corresponding $n + 1$-subword of w_1. Call the matches in the left portion of w_0 *left-matches.*

Discussion. Let u^c have genetic marker $\vec{j}_{u^c} = (j_n, j_{n+1}, \ldots j_{m-1})$ in w_0^c and suppose that u^c sits inside the $n + 1$ word $(u')^c$ with genetic marker $(j_{n+1}, \ldots j_{m-1})$. Then words with genetic marker $\vec{j}_u$ sit inside every 2-subsection of u'. It follows that if $k^c > 0$ and k^c is a perfect match of u^c with v^c having genetic marker $\vec{j}_{v^c} = (j'_n, j'_{n+1}, \ldots j'_{m-1})$ win w_1^c, then $j_n \leq j'_n$. Thus

the relative position of v^c in the $n + 1$-subword of w_1^c with genetic marker $(j'_{n+1}, \ldots j'_{m-1})$ is to the right of the position of u^c in $(u')^c$; i.e. the relative shift is to the left to match u^v with v^c. For this reason, when $k^c > 0$ we need only consider left shifts.

It is also easy to see that perfect matches between n-words with genetic markers $\vec{j}$ and $\vec{j}'$ inside an m-words w_0^c, w_1^c are those k^c that match the first occurrence of an n-word with genetic marker $\vec{j}$ in w_0^c with the first occurrence of an n-word with genetic marker $\vec{j}'$ in w_1^c.

The next lemma says that perfect matches can be viewed as the locations of shifts of odometer based words wrapped around the circle.

Lemma 117. *Suppose that $w_0, w_1 \in \mathcal{W}_m$ and $w_i^c = c_m(w_i)$. Let $n < m$ and $0 \le k^c < q_m$ and suppose that k^c is a perfect match between some pair of n-subwords of w_0^c and w_1^c. Then there is a unique k such that for all genetic markers $\vec{j}, \vec{j}',*

- *k^c is a perfect match between the n-subwords of the w_i^c with genetic markers $\vec{j}$ and $\vec{j}'$ iff*

- *k is a left match between the n-subwords of w with genetic markers $\vec{j}$ and $\vec{j}'$.*

The Lemma has an obvious analogue for negative k^c and right matches.

$\vdash$ Suppose that k^c is a perfect match between $\vec{j}$ and $\vec{j}'$. Call the subwords of w_0, w_1 with genetic markers $\vec{j}$ and $\vec{j}'$ u and v. Then u^c, v^c are perfectly matched by k^c. Let k be the distance between the locations of u and v. Since $k^c \ge 0$ we have $k \ge 0$. From our discussion we seen that k is a left match of u, v. We claim that this k satisfies the lemma.

Let u', v' be the $n+1$-subwords of w_0, w_1 inside which u, v occur. Suppose that $u = u_0 u_1 \ldots u_{k_n-1}$ and $v = v_0 v_1 \ldots v_{k_n-1}$, so $(u')^c = \mathcal{C}((u_0)^c, \ldots, (u_{k_n-1})^c)$, $(v')^c = \mathcal{C}((v_0)^c, \ldots, (v_{k_n-1})^c)$. If u_i, v_j are left matched by k in u', v', then the first occurrences of $(u_i)^c$ and $(v_j)^c$ are matched by k^c, hence *inside* $(u')^c, (v')^c$, k^c is a perfect match of $(u_i)^c$ and $(v_j)^c$.

The relative position of $(u')^c$ and $(v')^c$ is duplicated over all $n + 1$-words with genetic markers $j_{u'}, j_{v'}$ in w_0 and $sh^k(w_1)$. It follows that k^c is a perfect match of u^c and v^c inside w_0^c, w_1^c.

From the uniformity of the relative positions of $n+1$-words it also follows that any two n-subwords of $n + 1$-subwords $(u^*)^c, (v^*)^c$ in positions i, j that

are k^c matched in $(u')^c, (v')^c$ are k^c-matched. Since these exactly coincide with the n-subwords of w_0, w_1 that are left-matched by k, we have proved the lemma. ⊣

Lemma 118. *Let $w_0 = w_1 = w$. Let $n \in \mathbb{N}$ and $k \in \mathbb{Z}$. Then:*

1. *Let $M(k)$ be the least M such that $k < q_M$. Let $u, v \in \mathcal{W}_n^c$. Then if k matches u, v inside $w \in \mathcal{W}_m^c$ with $m \geq M(k)$ then u, v occur inside the same $M(k)$-subword of w.*

2. *Let $m \geq M(k)$. Then k is a perfect match of occurrences of u, v inside an m-word iff k is a perfect match inside the $M(k)$ word in which they appear.*

⊢   Use Lemmas 61 and 62.   ⊣

Item 2 means that we usually don't have to refer to a long words when we are discussing perfect matches of u and v and fixes the scale of the potential perfect matches.

We can identify perfect matches numerically:

Lemma 119. *Let $A < q_N$, $w^c \in \mathcal{W}_N^c$. Then there is a (m, n)-genetic marker $\vec{j}$ such that A is the location of the first occurrence of some word genetic marker $\vec{j}$ if and only if*

$$A = c_{N-1}l_{N-1}q_{N-1} + c_{N-2}l_{N-1}q_{N-2} + \cdots + c_m l_m q_m \qquad (71)$$

where $0 \leq c_i < k_i$.

From Lemma 117, we see the correspondence between odometer translations and circular translation. We now address the question: given an arbitrary circular translation, how does one adjust it to get an odometer translation that gives the best fit among a given collection of n-words?

Theorem 120. *Let $n \in \mathbb{N}$ and $k \in \mathbb{Z}$. Then if $\{(u_i, v_i) : i \in I\} \subseteq \mathcal{W}_n^c \times \mathcal{W}_n^c$, $w \in \mathcal{W}_m^c$ $(m \geq M(k))$ then there is a k' such that $|k' - k| < q_m$ and:*

1. *all k'-matches of a (u_i, v_i) in w are perfect matches,*

2. *and*

$$\sum_i |k'\text{-matches of a } u_i \text{ with a } v_i| \geq \sum_i |k\text{-matches of a } u_i \text{ with a } v_i|$$

⊢ Without loss of generality $k \geq 0$ (otherwise we reverse the role of u and v). Words in $\mathcal{W}_m$ with $m > M(k)$ start with a block of b's of length at least $q_{M(k)}$. Hence if k matches n-words u, v inside $w \in \mathcal{W}_m^c$, they both must occur in some $M(k)$-subword of w.

To see item 1, we need to show how to improve k to a k' that is a perfect match. Changing k will involve sacrificing some of the matches of pairs in I, but this will be compensated by the additional multiplicity of the remaining matches.

We prove by induction on $d \geq 1$, that for all m, n with $m - n = d$ and all collections of pairs of m-words $\{(w_0^j, w_1^j) : j \in J\}$ and all k, all $m \geq M(k)$, all natural number weightings $\{\alpha_j : j \in J\}$ and all $\{(u_i, v_i) : i \in I\} \subseteq \mathcal{W}_n^c \times \mathcal{W}_n^c$ we can find a k' such that $|k' - k| < q_m$ such that (a) holds and

$$\sum_j \sum_i \alpha_j |\{k'\text{-matches of a } u_i \text{ and a } v_i \text{ in } (w_0^j, w_1^j)\}| \geq$$
$$\sum_j \sum_i \alpha_j |\{k\text{-matches of a } u_i \text{ and a } v_i \text{ in } (w_0^j, w_1^j)\}|$$

Suppose first that $d = 1$. Then successive 2-subsections of m-words are separated by boundary sections of size

$$j_i + (q - j_{i+1}) \equiv p_n^{-1} \pmod{q}.$$

Because q_n does not divide p_n, given a 2-subsection $\vec{s}$ of w_0^j there is a unique 2-subsection $\vec{t}$ of w_1^j within which k can match n-words. Moreover this does not depend on j, but rather the underlying locations of the words.

We start by lining up blocks of the form $u_i^{l_n - 1}$ with blocks of the form $v_i^{l_n - 1}$. To do this we classify the k-matches of a pair $(u, v) = (u_i, v_i)$ into *left block matches* if u and $sh^k(v)$ align as[28]:

```
.   .   u*  u*  u*   B   u   .   .   u   u    u    u   u   u   .   .
.   .   v   v   v    v   v   .   .   v   v    B    v*  v*  v*  .   .
```

and *right block matches* if u and $sh^k(v)$ align as:

```
.   .   u   u   u    u   u   .   .   u   u    B    u*  u*  u*  .   .
.   .   v*  v*  v*   B   v   .   .   v   v    v    v   v   v   .   .
```

---

[28] In both of these graphics the second row is a portion of $sk^k(w_1^j)$ and B represents a boundary section. These pictures are independent of j.

Note by taking k' to be $k + lq_n$ for some $l < l_n - 1$ we can turn all left block matches of all of the (u_i, v_i) into matches of entire $u_i^{l_n - 1}$ with $sh^k(v_i^{l_n - 1})$, but doing so destroys completely some of the right block matches. Similarly if we can shift to make all right block matches into matches of $u_i^{l_n - 1}$ with $sh^k(v_i^{l_n - 1})$ by destroying left block matches.

If we examine a particular left block match of a pair (u_i, v_i) in some w_0^j and a right block match of another pair (u_i', v_i') in w_1^j and we change k to k' to make $u_i^{l_n - 1}$ match with $sh^{k'}(v_i^{l_n - 1})$ then the sum of k'-matches between (u_i, v_i) and (u_i', v_i') goes up by one: we lose the right block matches but we gain left block matches and we gain one more match from the boundary section.

Suppose that

$$\sum_j \sum_j \alpha_j |\{\text{left block matches in } (w_0^j, w_1^j)| \geq$$
$$\sum_j \sum_j \alpha_j |\{\text{right block matches in } (w_0^j, w_1^j)|$$

Then from the previous paragraph that if we take $k' = k + lq_n$ for some $l < l_n - 1$ then we can make all left block matches have multiplicity $l_n - 1$ (while removing right block matches) and have:

$$\sum_j \sum_i \alpha_j |k'\text{-matches of a } u_i \text{ with a } v_i \text{ in some } (w_0^j, w_1^j)| \geq$$
$$\sum_j \sum_i \alpha_j |k\text{-matches of a } u_i \text{ with a } v_i \text{ in some } (w_0^j, w_1^j)|.$$

If, on the other hand, the weighted sum of the right block matches is greater than weighted sum of the left block matches, we shift the other direction to fix all right block matches and destroy all left block matches.

Thus we can assume that we have a k such that for all (u_i, v_i), sh^k matches $(l_n - 1)$-powers of u_i with $(l_n - 1)$-powers of v_i. This k would be a perfect match except that it matches n-words across 2-subsections. Writing each $w_s^j = \mathcal{C}_n(w_1, \ldots w_{k_n})$ then sh^k matches blocks of the form $w_s^{l_n - 1}$ in one 1-subsection of w_0^j with a block of the form $w_{s'}^{l_n - 1}$ in a (potentially) different 1-subsection of w_1^j. Moreover $s - s'$ is constant on all of these matches, since the differences between starts of $w_j^{l_n - 1}$-blocks are of length $l_n q_n$. Fix such a pair s, s'. By changing k so that it lines up $w_s^{l_n - 1}$ with $w_{s'}^{l_n - 1}$ in the first

1-subsection we create a perfect match of $n = m - 1$-words and increase the total number of matches of the form (u_i, v_i). This establishes the case where $d = 1$.

We now do the induction step. Let $d = m - n$ and assume the result holds for $d - 1$. Suppose that we are given $\{\alpha_j : j \in J\}$.

We can decompose a k-match between n-subwords of w_0^j and w_1^j as $k_1 + k^*$ where $k^* \in [-q_{m-1} + 1, q_{m-1} - 1]$ and k_1 is a match of $m - 1$ subwords of w_0^j and w_1^j.

Here is a picture of a pair $(u', v') \in \mathcal{W}_{m-1}^c \times \mathcal{W}_{m-1}^c$ comparing w_0 in the upper row with the k_1-shift of w_1 in the lower row.

. . . .	u'	u'	u'	B	u*	u*	u*	u*	. . . .
. . . .	B	v'	v'	v'	v'	v'	v'	v'	. . . .

Here is a picture after the $k = k_1 + k^*$ shift of w_1:

. . . .	u'	u'	u'	B	u*	u*	u*	u*	. . . .
. . . .	B	v'	v'	v'	v'	v'	v'	v'	. . . .

Let $\{(u', v')_{i'} : i' \in I'\}$ be the collection of pairs (u', v') from $\mathcal{W}_{m-1}^c$ sitting inside a pair (w_0^j, w_1^j) that contain k-matches of words (u_i, v_i). Arguing as in the case $d = 1$ we can adjust k_1 to a k_1' so that it is a perfect match of $m - 1$-words in I' and, summing over I and J, the weighted sum of $k_1' + k^*$-matches of pairs in I does not decrease.[29]

This is how the $m - 1$-words look after shifting by $k_1' + k^*$:

. . . .	B	u'	u'	...	u'	u'	u'	u'	B
. . . .	B	v'	v'	...	v'	v'	v'	B

The offset of the copies of u' and v' is k^*. Note that the boundary sections line up.

We now are in the position of having shifted by k_1' so that the powers of pairs $\{(u', v')_{i'} : i' \in I'\}$ are lined up. The additional shift k^* has absolute value less than q_{m-1}. Moreover all of the words $\{(u', v')_{i'} : i' \in I'\}$ are lined up the same way when shifted by k^*.

---

[29] We note that it is not enough to increase the weighted sum of the number of matches of pairs in I', because various I' matches may contain different number of I-matches. Nonetheless, arguing as in the case $d - 1$, one of the two possibilities for lining up the $m - 1$ subwords does not decrease the weighted sum of the number of $k_1' + k^*$-matches of I-words.
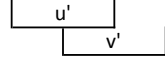
We call an occurrence of a $(u', v')_{i'}$ that is lined up in $(w_0^j, sh^{k_1'}(w_1^j))$ *good*. Let $\beta_{j,i'}$ be the number of good occurrences of $(u', v')_{i'}$ and

$$\alpha_{i'}' = \sum_j \alpha_j \beta_{j,i'}.$$

Note that

$$\sum_{i'} \sum_i \alpha_{i'}' |(k_1' + k^*)\text{-matches of a } u_i \text{ with a } v_i \text{ in some good occurrence of } (u', v')_{i'}|$$

$$= \sum_j \sum_i \alpha_j |(k_1' + k^*)\text{-matches of a } u_i \text{ with a } v_i \text{ in some } (w_0^j, w_1^j)|.$$

We now view the pairs $\{(u', v')_{i'} : i' \in I'\}$ as sitting on the intervals $[0, q_{m-1} - 1]$ and then shifting v' by k^*:

We are in a position to apply our induction hypothesis with I' playing the role of J, the $\alpha_{i'}'$'s being the α_j's, $d - 1 = (m - 1) - n$ and the shift being k^*.

The result is a k^{**} such that every k^{**}-match of a (u_i, v_i) in a $(u', v')_{i'}$ is perfect and

$$\sum_{i'} \sum_i \alpha_{i'}' |(k^{**})\text{-matches of a } u_i \text{ with a } v_i \text{ in some } (u', v')_{i'}| \geq$$

$$\sum_{i'} \sum_i \alpha_{i'}' |(k^*)\text{-matches of a } u_i \text{ with a } v_i \text{ in some } (u', v')_{i'}|.$$

We note that every $k_1' + k^{**}$-match of a (u_i, v_i) in a (w_0^j, w_1^j) is perfect. Since

$$\sum_{i'} \sum_i \alpha_{i'}' |(k_1' + k^{**})\text{-matches of a } u_i \text{ with a } v_i \text{ in some good occurrence of } (u', v')_{i'}|$$

$$= \sum_j \sum_i \alpha_j |(k_1' + k^{**})\text{-matches of a } u_i \text{ with a } v_i \text{ in some } (w_0^j, w_1^j)|.$$

we see that

$$\sum_j \sum_i \alpha_j |(k_1' + k^{**})\text{-matches of a } u_i \text{ with a } v_i \text{ in some } (w_0^j, w_1^j)| \geq$$

$$\sum_j \sum_i \alpha_j |k\text{-matches of a } u_i \text{ with a } v_i \text{ in some } (w_0^j, w_1^j)|.$$

This completes the proof of Lemma [120]. ⊣

9 Open Problems

We finish with two open problems that we find interesting and believe to be feasible. The first is to characterize the class of transformations isomorphic to circular systems in Ergodic-theoretic terms. All circular systems have common properties such that can be described in terms of rigidity sequences or zero entropy. The suggestions is to find a complete characterization in using this type of notion.

The second problem can be stated as follows. For the realization problem, the underlying rotation α of a circular system must be Liouvillian; however realization is not necessary for the results in this paper. Can an arbitrary irrational α be the underlying rotation of a circular system?

References

[1] D. V. Anosov and A. B. Katok. New examples in smooth ergodic theory. Ergodic diffeomorphisms. *Trudy Moskov. Mat. Obšč.*, 23:3–36, 1970.

[2] Ferenc Beleznay and Matthew Foreman. The complexity of the collection of measure-distal transformations. *Ergodic Theory Dynam. Systems*, 16(5):929–962, 1996.

[3] Tomasz Downarowicz. The Choquet simplex of invariant measures for minimal flows. *Israel J. Math.*, 74(2-3):241–256, 1991.

[4] Jacob Feldman. Borel structures and invariants for measurable transformations. *Proc. Amer. Math. Soc.*, 46:383–394, 1974.

[5] M. Foreman and B. Weiss. A symbolic representation of Anosov-Katok systems. *To appear in Journal d'Analyse Mathematique*, pages 1–63, 2015.

[6] Matthew Foreman, Daniel J. Rudolph, and Benjamin Weiss. The conjugacy problem in ergodic theory. *Ann. of Math. (2)*, 173(3):1529–1586, 2011.

[7] Matthew Foreman and Benjamin Weiss. Measure preserving diffeomorphisms of the torus are unclassifiable. *TO APPEAR*, pages 1–102, 2018.

[8] Matthew Foreman and Benjamin Weiss. Odometer based systems. *TO APPEAR*, 2019.

[9] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory*. Princeton University Press, Princeton, N.J., 1981. M. B. Porter Lectures.

[10] Hillel Furstenberg and Benjamin Weiss. A mean ergodic theorem for $(1/N) \sum_{n=1}^{N} f(T^n x) g(T^{n^2} x)$. In *Convergence in ergodic theory and probability (Columbus, OH, 1993)*, volume 5 of *Ohio State Univ. Math. Res. Inst. Publ.*, pages 193–227. de Gruyter, Berlin, 1996.

[11] Eli Glasner. *Ergodic theory via joinings*, volume 101 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003.

[12] Paul R. Halmos. *Lectures on ergodic theory*. Chelsea Publishing Co., New York, 1960.

[13] Paul R. Halmos and John von Neumann. Operator methods in classical mechanics. II. *Ann. of Math. (2)*, 43:332–350, 1942.

[14] Anatole Katok. *Combinatorial constructions in ergodic theory and dynamics*, volume 30 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2003.

[15] Karl Petersen. *Ergodic theory*, volume 2 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1989. Corrected reprint of the 1983 original.

[16] Daniel J. Rudolph. *Fundamentals of measurable dynamics*. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1990. Ergodic theory on Lebesgue spaces.

[17] Paul C. Shields. *The ergodic theory of discrete sample paths*, volume 13 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996.

[18] William A. Veech. A criterion for a process to be prime. *Monatsh. Math.*, 94(4):335–341, 1982.

[19] J. von Neumann. Zur Operatorenmethode in der klassischen Mechanik. *Ann. of Math. (2)*, 33(3):587–642, 1932.

[20] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[21] Benjamin Weiss. *Single orbit dynamics*, volume 95 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, Providence, RI, 2000.

[22] Robert J. Zimmer. Ergodic actions with generalized discrete spectrum. *Illinois J. Math.*, 20(4):555–588, 1976.