Annals of Mathematics **173** (2011), 1529–1586
doi: 10.4007/annals.2011.173.3.7

The conjugacy problem in ergodic theory

By Matthew Foreman, Daniel J. Rudolph, and Benjamin Weiss

Abstract

All common probability preserving transformations can be represented as elements of MPT, the group of measure preserving transformations of the unit interval with Lebesgue measure. This group has a natural Polish topology and the induced topology on the set of ergodic transformations is also Polish. Our main result is that the set of ergodic elements T in MPT that are isomorphic to their inverse is a complete analytic set. This has as a consequence the fact that the isomorphism relation is also a complete analytic set and in particular is not Borel. This is in stark contrast to the situation of unitary operators where the spectral theorem can be used to show that conjugacy relation in the unitary group is Borel.

This result explains, perhaps, why the problem of determining whether ergodic transformations are isomorphic or not has proven to be so intractable. The construction that we use is general enough to show that the set of ergodic T's with nontrivial centralizer is also complete analytic.

On the positive side we show that the isomorphism relation is Borel when restricted to the rank one transformations, which form a generic subset of MPT. It remains an open problem to find a good explicit method of checking when two rank one transformations are isomorphic.

In Memoriam: Prior to the final proofs of our paper our dear friend, Dan Rudolph, died just short of his 60th birthday, after a long and valiant struggle with ALS. He took an active part in this work to the very end despite his illness. This is not the place to describe in detail his lasting contributions to the modern theory of measurable dynamics; suffice it to say that his loss will be deeply felt not only by his close collaborators but by all who have an interest in this field. We want to dedicate our part of this paper to his memory.

1. Introduction

The isomorphism problem has been the focus of much of the work in ergodic theory ever since it was formulated in 1932 by von Neumann in his pioneering paper [21]. Following Koopman's definition of the unitary operator

M. Foreman was supported in part by NSF grant DMS-0701030. D.J. Rudolph was supported in part by NSF grant DMS-0618030. M. Foreman and B. Weiss were partially supported by the US-Israel BSF grant BSF-2006312.

associated to a measure preserving transformation, von Neumann first used the newly developed spectral theory to prove the ergodic theorem and then showed that for ergodic measure preserving transformations with pure point spectrum the spectrum is a complete isomorphism invariant. The property of pure point spectrum corresponds to mechanical systems with very regular behavior. It was quickly realized that von Neumann's hope that spectrum would be a complete invariant turned out to be too optimistic.

The next significant advance came with the introduction of the Kolmogorov-Sinai entropy as an invariant capable of distinguishing between different chaotic systems, such as geodesic flows on surfaces of negative curvature or independent processes. Later Ornstein showed that, in fact, this entropy was sufficient to classify a large class of these highly random systems.

In the framework of this program many properties of transformations such as *mixing* of various types, or *finite rank* have been studied and characterized. In spite of all of this work, no new invariants as powerful as entropy have been found.

Of particular interest is the class of *ergodic transformations*, as these are the irreducible objects in this category and the Ergodic Decomposition Theorem says that every measure preserving transformation can be written as an integral of ergodic transformations.

Despite the fact that complete invariants have been discovered for some classes of ergodic transformations, such as the measure distal transformations of a particular height, the general problem remained intractable. Recent developments have established some *anticlassification* results that demonstrate in a rigorous way that classification is not possible. Our main result is exactly in this direction.

The key notion in the anticlassification results is the Borel/non-Borel distinction. Behind this distinction is the idea that saying a set X (or function) is Borel is a very liberal way of saying that membership in X can be determined (or f computed) using a countable protocol whose basic input is membership in open sets. Saying that a set (or function) is not Borel is saying that no amount of countable resources can be applied to decide membership in that set (or be used to compute the function); i.e., the issue of membership is highly intractable.

Hjorth ([10]) introduced the notion of *turbulence* and used it in [11] to show that there is no Borel way of attaching algebraic invariants to ergodic transformations that completely determine isomorphism. Foreman and Weiss [6] improved this result to show that the action of the group of measure preserving transformations on the ergodic transformations is turbulent. A consequence of this is that no generic class can have a complete collection of algebraic invariants.

Passing from a single transformation to pairs (S, T) of measure preserving transformations and asking whether the equivalence relation defined by isomorphism is a Borel set, Hjorth gave a negative answer. His proof used nonergodic transformations in an essential way, the complexity being due to the ergodic decomposition.[1] The main result of this paper is that the isomorphism relation for *ergodic* transformations is not Borel.

This can be interpreted as saying that there is no method or protocol that involves a countable amount of information and countable number of steps that reliably distinguishes between nonisomorphic ergodic measure preserving transformations. We view this as a rigorous way of saying that the classification problem for ergodic measure preserving transformations is intractable.

We also note the contrast between the situation for the ergodic measure preserving transformations and that of the unitary operators, where the Spectral Theorem can be used to show that the isomorphism relation *is* Borel.

On the other hand we point out that there is a well-studied generic set of transformations — the rank one transformations — where the isomorphism relation is Borel.

For these results to make sense we must find a natural model or space in which to put the measure preserving transformations. This is discussed further in Section 2.2. We give only the basic model here. As far as is known, all models are equivalent in the sense of the Borel/non-Borel distinction and in which collections of transformations are generic.

A result of von Neumann that every nonatomic separable probability space is measure theoretically isomorphic to Lebesgue measure on the unit interval $[0, 1]$ shows that every measure preserving transformation has an isomorph that maps from the unit interval to itself.

After identifying two invertible measure preserving transformations on the unit interval, if they agree on a set of full measure, we call the resulting group **MPT**. The classification problem can now be rephrased as a conjugacy problem, as two measure preserving transformations on the unit interval are isomorphic just in case they are conjugate in this group.

The strong operator topology on **MPT** is compatible with a complete separable metric space. This topology can be described by saying that two transformations are close if and only if they send a very fine partition of the unit interval to very similar places.

The topology endows the space with a natural Borel structure and allows us to give qualitative information about classes of measure preserving transformations. Natural classes of transformations, such as the Bernoulli transformations or translations on a compact group, form Borel sets in **MPT**.

[1] The ergodic components of Hjorth's transformations are irrational rotations of the circle.

Others, such as the *measure distal* transformations form co-analytic, non-Borel sets ([3]).

The topology also allows discussion of *generic* classes, i.e. classes which contain a dense $\mathcal{G}_\delta$ sets in the topology. The collection E of ergodic transformations form a dense $\mathcal{G}_\delta$ set as do the *rank one* transformations. Thus the induced topology on E is Polish; i.e. compatible with a complete separable metric; hence the product space $E \times E$ is also a Polish space. The main result, Theorem 7, has as a corollary:

THEOREM. *The collection*

$$\{(S,T) : S \textit{ and } T \textit{ are ergodic and conjugate}\} \subseteq E \times E$$

is not Borel.

In fact we prove more. A continuous image of a Borel set is called *analytic* (or Σ^1_1). The collection of analytic sets is strictly larger than the collection of Borel sets [20]. There is a natural notion of complexity among analytic sets and our result shows that the set of pairs of conjugate, ergodic transformations is an analytic set of maximal complexity.

This is done by showing that the collection of transformations that are isomorphic to their inverse is a maximally complicated analytic set. For this to happen there must exist some transformations that are not isomorphic to their inverses. Such examples were first constructed by Anzai ([1]) refuting a conjecture of Halmos and von Neumann ([9]) that they could not exist.

A closely related result shows that the collection of ergodic transformations that have nontrivial centralizer is also a maximally complicated analytic set and hence not Borel. This perhaps explains why transformations with trivial centralizer were originally difficult to construct ([16]).

Abstract considerations of Polish group actions imply that there is a generic class on which the orbit equivalence relation is Borel. In this case we use a theorem of King [14] to identify the generic class, the space of rank one transformations:

THEOREM. *The collection of pairs (S,T) such that S and T are ergodic and rank one and are conjugate is Borel subset of $E \times E$.*

1.1. *Outline of the paper.* To show that a set $C \subset Y$ is not Borel it suffices to take an existing example of a non-Borel set $A \subset X$ and define a Borel function $f : X \to Y$ such that $C = f[A]$. The function f will be called a *reduction.* A canonical example of such a pair (A, X) can be described as follows. We consider the space $\mathcal{T}rees$ consisting of trees of finite sequences of elements of a countable set. Among the trees, some have infinite branches and we let A be be this collection.

The crux of the paper is to continuously associate to each tree $\mathcal{T}$ a transformation $T = F(\mathcal{T})$ such that $T \cong T^{-1}$ just in case $\mathcal{T}$ has an infinite branch.

We do so by constructing a shift invariant closed subset $\mathbb{K}(\mathcal{T})$ of $\{0,1\}^{\mathbb{Z}}$ that has a unique invariant measure (which is then automatically ergodic). The shift invariant subset is constructed by describing in succession collections of finite blocks $\mathcal{W}_1, \mathcal{W}_2, \ldots, \mathcal{W}_n \cdots$ in the basic alphabet $\{0,1\}$. The words in $\mathcal{W}_n$ all have the same length l_n, and the words in $\mathcal{W}_{n+1}$ are concatenations of words from $\mathcal{W}_n$. The tree $\mathcal{T}$ controls exactly how $\mathcal{W}_{n+1}$ is formed from $\mathcal{W}_n$.

Paralleling the construction of the words $\mathcal{W}_n$, we construct groups of involutions $G_s^n, n \geq s$, with the G_s^n having a direct limit G_s as n tends to infinity. There are also surjective homomorphisms $\rho_{s+1}^n : G_{s+1}^n \to G_s^n$ which converge to a surjective homomorphism $\rho_{s+1} : G_{s+1} \to G_s$. The system (G_s, ρ_s) has an inverse limit denoted by G, and it is nontrivial if and only if $\mathcal{T}$ has an infinite branch.

As s goes to infinity, the elements of the G_s's give more and more precise information about an isomorphism of T with T^{-1} or a nontrivial isomorphism of T with itself. It will be relatively easy to see that an element of the inverse limit of the G_s's are isomorphisms between T and $T^{\mp 1}$. The difficult part is showing that there are no other isomorphisms.

This is done by giving a complete analysis of the joinings between T and $T^{\mp 1}$. We find a canonical sequence of factors $\mathbb{K}_s$ of $\mathbb{K}(T)$ corresponding to some equivalence relations $\mathcal{Q}_s$. These factors have $\mathbb{K}$ as their inverse limit. In passing from $\mathcal{W}_s$ to $\mathcal{W}_{s+1}$ we will have randomly concatenated words from $\mathcal{W}_s$. This will allow us to show that any joining of T with $T^{\mp 1}$ is either an independent joining over an isomorphism of $\mathbb{K}_s$ with $\mathbb{K}_s^{\mp 1}$ arising from an element of G_s or comes from some element of the inverse limit of the G_s's.

Here is how the paper is organized. The second part of this section establishes our notational conventions. Section 2 develops the tools necessary to give a precise statement of our theorem and gives pointers to elements of ergodic theory and descriptive set theory that we use in the proof. Section 3 introduces the groups associated with a tree and describes the relationship between ill-founded trees and nontrivial inverse limits of these groups.

In Section 4, we begin specifying the properties of our reduction. The basic specifications describe how we get transformations by building collections of words and give some sufficient conditions for our transformations to be ergodic. The group specifications establish the relationships between the groups associated to a tree and the equivalence relations on words.

The first nine specifications suffice to allow us to define a tower of factors that exhaust the σ-algebras associated with our transformations. Section 5 is where this is done. In Section 6 we first show that branches through a tree give appropriate isomorphisms. Then we introduce the final two specifications and use them to establish the relationships between the canonical factors. Along the way we show that the odometer factor is the Kronecker factor.

Having shown that the existence of a continuous function from $\mathcal{T}rees$ into collections of words satisfying our specifications suffices to prove our theorem, it still remains to construct the words. In Section 7, we collect and simplify the specifications with the aim of showing how to accomplish this. The actual construction is done by probabilistic arguments and is given in Section 8. Section 9 makes some remarks about the consequences of Theorem 7.

In Section 10 we explain why isomorphism of rank one transformations is Borel.

1.2. *Notation.* We finish this section by giving some notational conventions for the paper. These are given here primarily for later reference. We will let $\mathbb{N} = \{0, 1, 2, \dots\}$ and identify each $n \in \mathbb{N}$ with $\{0, 1, \dots, n-1\}$. We will write X^n for the n-fold Cartesian product of X with itself. We identify X^n with sequences of elements of X of length n. The notation $X^{<\mathbb{N}}$ will denote all finite sequences of elements of X.[2] Clearly $X^{<\mathbb{N}} = \bigcup_{n\in\mathbb{N}} X^n$. We will write $\mathbb{Z}_n$ for the abelian group $\mathbb{Z}/n\mathbb{Z}$.

We will often discuss "words" in an alphabet Σ. These are simply strings of symbols from Σ. If w is a finite or infinite word, we will write $\mathrm{rev}(w)$ for the reverse word of w. If $\mathcal{W}$ is a collection of words, we write $\mathrm{rev}(\mathcal{W})$ for the collection of reverse words from $\mathcal{W}$. We will write $\Sigma^{\mathbb{Z}}$ for the doubly infinite sequences of elements of Σ. If $x \in \Sigma^{\mathbb{Z}}$, then we will write x^{-1} for the reverse of x, i.e. the sequence defined by setting $x^{-1}(n) = x(-n)$. If $\mathbb{K} \subseteq \Sigma^{\mathbb{Z}}$ we will write $\mathbb{K}^{-1}$ for the collection of reverses of elements of $\mathbb{K}$. Analogously, if $\mathcal{B}$ is a Σ-algebra on $\mathbb{K}$, then $\mathcal{B}^{-1}$ is the image of $\mathcal{B}$ under the homeomorphism of $\mathbb{K}$ with $\mathbb{K}^{-1}$ defined by $x \mapsto x^{-1}$.

In some contexts we will want to compare finite shifts of finite words u and v. In these cases we will view $u : n \to \Sigma$ and $v : m \to \Sigma$. The shift of v by k is viewed as a function $\mathrm{sh}^k(v) : [-k, m-k-1] \to \Sigma$. The comparison will be done on the overlap of u and $\mathrm{sh}^k(v)$, i.e. $u \restriction [0, \min(n, m-k-1)]$ and $\mathrm{sh}^k(v) \restriction [0, \min(n, m-k-1)]$. This allows us to talk about relative occurrences of words (u', v') as subwords of $(u, \mathrm{sh}^k(v))$–meaning the occurrences on the overlap.

A *Polish* space is a topological space whose topology is compatible with a complete separable metric.

When we discuss subsets of the domain of a measure we will be implicitly assuming they are measurable. The sets constructed in this paper are all measurable.

[2]The empty sequence is a member of $X^{<\mathbb{N}}$.

2. **Background information and precise statements**

In this section we present the barest outline of the background necessary for the proof of the main theorem. In explanation we can do no better than to quote the historian of science, George Sarton ([19]):

> "As far as scientific matters are concerned, I try to say enough to refresh the reader's memory but do not attempt to provide complete explanations, which would be equally unbearable to those who know and to those who do not."

2.1. *Trees and their topologies.* A *tree* on a set X is a subset $\mathcal{T} \subseteq X^{<\mathbb{N}}$ that is closed under initial segments.[3] In this paper we will primarily be considering trees $\mathcal{T} \subseteq \mathbb{N}^{<\mathbb{N}}$. *Level* s of $\mathcal{T}$ is defined to be the collection elements of $\mathcal{T}$ that have length s. We will write $\mathrm{lh}(\tau)$ for the length of τ. If σ is an initial segment of τ then σ is a *predecessor* of τ and τ is a *successor* of σ.

A *branch* through $\mathcal{T}$ is a function b into $\mathcal{T}$ having domain either some $n \in \mathbb{N}$ or $\mathbb{N}$ itself such that $\mathrm{lh}(b(s)) = s$, and if $s+1$ is in the domain of b, then $b(s+1)$ is an immediate successor of $b(s)$.

We will occasionally abuse our conventions by writing "successor" to mean "immediate successor" and "branch" to mean "infinite branch".

If $\mathcal{T}$ does not have an infinite branch, then we will say that $\mathcal{T}$ is *well-founded.* If it has an infinite branch, then $\mathcal{T}$ is *ill-founded.*

In many places in the paper we will refer to a fixed enumeration $\langle \sigma_n : n \in \mathbb{N}\rangle$ of $\mathbb{N}^{<\mathbb{N}}$ with the property that every proper predecessor of σ_n is some σ_m for $m < n$. We identify subsets A of $\mathbb{N}^{<\mathbb{N}}$ with their characteristic functions $\chi_A : \mathbb{N}^{<\mathbb{N}} \to \{0,1\}$. The collection of such χ_A can be viewed as the members of an infinite product space $\{0,1\}^{\mathbb{N}^{<\mathbb{N}}}$ homeomorphic to the Cantor space. Each function $a : \{\sigma_m : m < n\} \to \{0,1\}$ determines a basic open set

$$\langle a \rangle = \{\chi : \chi \restriction \{\sigma_m : m < n\} = a\} \subseteq \{0,1\}^{\mathbb{N}^{<\mathbb{N}}}.$$

The collection of all such $\langle a\rangle$ form a basis for the topology. In this topology the collection of trees is a closed (hence compact) subset of $\{0,1\}^{\mathbb{N}^{<\mathbb{N}}}$. The collection of trees containing arbitrarily long finite sequences is a dense $\mathcal{G}_\delta$ subset of $\{0,1\}^{\mathbb{N}^{<\mathbb{N}}}$. In particular it is a Polish space. We will refer to the space of trees with arbitrarily long finite sequences as $\mathcal{T}rees$.

We can understand this topology by viewing the basic open sets as giving us a finite amount of information about the trees in it, namely which among the first n finite sequences belong to the trees. This allows us to adopt the heuristic of building continuous (or Borel) functions f from the space $\mathcal{T}rees$ to

[3] Explicitly: if $\sigma = (x_0, \dots, x_n) \in \mathcal{T}$ and $s \leq n$, then $\tau = (x_0, \dots, x_s) \in \mathcal{T}$. All of our trees are rooted in the sense that they all contain the empty sequence.

an arbitrary topological space $\mathcal{X}$ using constructions that associate an element of $\mathcal{X}$ to each tree $\mathcal{T}$ by taking more and more finite information about which σ_n belong to $\mathcal{T}$ and building a decreasing sequence of open (or Borel) sets whose intersection is a point $x \in \mathcal{X}$.

Definition 1. We will say that σ_m and σ_n are *consecutive* elements of $\mathcal{T}$ if $\sigma_m, \sigma_n \in \mathcal{T}$ and there is no j between m and n with $\sigma_j \in \mathcal{T}$.

We need an artifact for later use in the construction.

Definition 2. Define a map

$$M : \mathcal{T}rees \to \mathbb{N}^{\mathbb{N}}$$

by setting $M(\mathcal{T})(s) = n$ if and only if n is the least number such that $\sigma_n \in \mathcal{T}$ and $\mathrm{lh}(\sigma_n) = s$. If $\mathcal{T}$ is clear from context we will frequently write $M(s)$, instead of $M(\mathcal{T})(s)$. Dually we define a function

$$s : \mathcal{T}rees \to \mathbb{N}^{\mathbb{N}}$$

by setting $s(\mathcal{T})(n)$ to be the length of the longest sequence σ_m such that $\sigma_m \in \mathcal{T}$ and $m \leq n$. When $\mathcal{T}$ is clear from context we write $s(n)$ instead of $s(\mathcal{T})(n)$.

Note that $s(n) \leq n$.

LEMMA 3. *Endowing* $\mathbb{N}$ *with the discrete topology and* $\mathbb{N}^{\mathbb{N}}$ *with the product topology*, M *and* s *are continuous functions.*

2.2. *Transformations and their topologies.* We now describe some topologies on the space of measure preserving transformations. There are several models for the space of measure preserving transformations. These are discussed in more detail in the forthcoming paper [5]. Perhaps the two most prominent are

1. The group **MPT** of measure preserving transformations of $[0, 1]$ with the Halmos metric.
2. The space of shift invariant measures on $\Sigma^{\mathbb{Z}}$ with the weak* topology, where Σ is any countable set.[4]

The difference between the two models is that in the first we fix the measure and look at transformations that preserve the measure; in the second we fix the transformations and look at the measure preserved by the transformation.

[4]This second model captures all ergodic transformations and all free actions. It is not a suitable model for periodic transformations on a nonatomic measure space. It is adequate for our purposes since we are concerned with ergodic transformations in this paper.

In the second model, we will refer to measures as *ergodic* if the shift transformation is ergodic relative to the measure.

In both of the models the ergodic transformations form a dense $\mathcal{G}_\delta$ set. In [5] a stronger property is shown: there is a Borel bijection ψ between the aperiodic transformations on $[0,1]$ and nonatomic measures on $\Sigma^{\mathbb{Z}}$ with the property that

$$([0,1], \mathcal{B}, \lambda, T) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, \psi(T), \mathrm{sh}).$$

and, moreover, ψ takes comeager sets to comeager sets. From the point of view of the Borel/analytic distinction the choice of model is not important.

In the proof of Theorem 7 we will be building transformations as subshifts of some $\Sigma^{\mathbb{Z}}$ it is natural to adopt the second model, although we could easily work with the other model by viewing our transformations as being built by "cut and stack" methods on the unit interval. This alternative would build continuous reductions into the space of measure preserving transformations and its square. In Theorem 51 it is easier to adopt the the first model.

Basic open sets in the space of invariant measures can be specified by taking finitely many functions $a : [0, n-1] \to \Sigma$ and specifying the measures of

$$\langle a \rangle = \{x \in \Sigma^{\mathbb{Z}} : x \restriction [0, \ldots, n-1] = a\}$$

within some $\epsilon > 0$. For ergodic measures, this is equivalent to specifying the frequency the translates of a occur in a typical x within ϵ.

Notation. We will denote the Polish space of ergodic measures on $\Sigma^{\mathbb{Z}}$ as $\mathcal{E}$.

2.3. *A precise statement of the theorem.* If X is a topological space, then the collection of Borel sets is the smallest σ-algebra containing the open sets. A set is *analytic* if and only if it is the continuous image of a Borel set.[5]

Central to our arguments is the idea of a reduction. If X and Y are Polish spaces, $A \subseteq X, B \subseteq Y$ and $f : X \to Y$ is a function, then f *reduces* A to B if $f^{-1}(B) = A$. Informally: the question of an arbitrary $x \in X$ belonging to A can be reduced to whether $f(x)$ belongs to B.[6]

We will choose our functions to be either continuous (a *continuous reduction*) or Borel (a *Borel reduction*). In either category we get a transitive relation on the class of subsets of Polish spaces.

For sufficiently complex sets we interpret A being reducible to B as saying that B contains all of the complexity that A has. For example, if f is a Borel function and B is a Borel set, then $f^{-1}(B)$ is Borel. In particular, if f is

[5]For Polish spaces, this is equivalent to being the Borel image of an Borel set, and also equivalent of being the projection to X of a Borel set $B \subset X \times Y$ for some Polish Y.

[6]There is a related, but distinct, notion of reduction for equivalence relations. See [12].

a Borel function and A is *not* Borel, then B is not Borel. Similarly, if f is continuous and B is a $\mathcal{G}_\delta$, then A must be a $\mathcal{G}_\delta$.

Among the analytic sets there are some that are maximally complex under the notion of reducibility. These are the sets to which any analytic set can be reduced by a Borel function and are called *complete* analytic. By the transitivity of reduction, if A is complete analytic and can be reduced to an analytic set B, then B is complete analytic.

We will use the following classical result (see e.g. [12]).

THEOREM 4. *As a subset of the Polish space of infinite trees, the collection of ill-founded trees is a complete analytic set. In particular it is not Borel.*

It follows easily that:

COROLLARY 5. *The collection of trees with at least two infinite branches is a complete analytic set.*

We remind the reader of a standard definition:

Definition 6. Suppose that $(X, \mathcal{B}, \mu)$ is a measure space and $T : X \to X$ is an invertible measure preserving transformation. Then the centralizer of T is defined as the collection of $S : X \to X$ such that S is an invertible measure preserving transformation and $ST = TS$. We will denote it by $C(T)$.

The powers T^k belong to $C(T)$ for $k \in \mathbb{Z}$. If $C(T) \neq \{T^k : k \in \mathbb{Z}\}$, then we say that T has nontrivial centralizer.

We are now in a position to state our main theorem, which has as a corollary the fact that, as a subset of **MPT**×**MPT**, the collection of pairs (S, T) such that S is conjugate to T is a complete analytic set. Before we give a precise statement we note that there is a canonical map $\nu \mapsto \nu^{-1}$ with the property that $(\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh}^{-1}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}^{-1}, \nu^{-1}, \mathrm{sh})$. We will equate these two systems.

The main theorem is that the collection of ill-founded trees is Borel reducible to $\{\nu \in \mathcal{E} : (\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh}^{-1})\}$ by a continuous map F. If we define $D : \mathcal{E} \to \mathcal{E} \times \mathcal{E}$ by setting $D(\nu) = (\nu, \nu^{-1})$ (where ν^{-1} is the measure on $\Sigma^{\mathbb{Z}}$ taking a set X to $\nu(X^{-1})$) then

$$D \circ F : \mathcal{T}rees \to \mathcal{E} \times \mathcal{E}$$

is a continuous reduction of the ill-founded trees to

$$\{(\nu, \mu) : (\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, \mu, \mathrm{sh})\} \subseteq \mathcal{E} \times \mathcal{E}.$$

THEOREM 7 (Main result). *There is a continuous one-to-one map*

$$F : \mathcal{T}rees \to \mathcal{E}$$

that

A. *reduces the collection of ill-founded trees to the collection of ergodic measures ν such that $(\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh}^{-1})$, and*

B. *reduces the collection of trees with at least two infinite branches to the collection of ergodic measures* ν *such that* $(\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh})$ *has nontrivial centralizer.*

In our theorem the reduction F will be strongly one-to-one in that it takes distinct trees to measures yielding nonisomorphic transformations.

COROLLARY 8. *As a subset of* $\mathcal{E} \times \mathcal{E}$, *the collection of pairs* (ν, μ) *such that* $(\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh})$ *is conjugate to* $(\Sigma^{\mathbb{Z}}, \mathcal{D}, \mu, \mathrm{sh})$ *is a complete analytic set.*

COROLLARY 9. *As a subset of* $\mathcal{E}$, $\{\nu : (\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh})$ *has nontrivial centralizer*$\}$ *is a complete analytic set.*

Proof. Granting the theorem, all that is required to prove Corollaries 8 and 9 is to show that both the set of isomorphic pairs (ν, μ) and the set of ν with nontrivial centralizer are analytic.

This follows immediately by noting that the set of triples (ν, μ, ϕ) such that ν is ergodic and $\phi : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ is an invertible measure preserving map from $(\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu)$ to $(\Sigma^{\mathbb{Z}}, \mathcal{D}, \mu)$ is a Polish space $\mathbb{P}$ and the collection of triples where ϕ commutes with the shift map is a closed subset of $\mathbb{P}$. The set of isomorphic pairs (ν, μ) is the projection to the first two coordinates of this closed subset.

The second corollary follows similarly: we consider the collection of $(\nu, \nu, \phi) \in \mathbb{P}$ such that ϕ differs from sh^k for all k and ϕ commutes with the shift. □

By virtue of the equivalence of the two universal models for measure preserving systems we can immediately restate Theorem 7 as saying:

THEOREM. *There is a continuous one-to-one map*

$$F : \mathcal{T}rees \to \{T : [0,1] \to [0,1] | T \text{ is invertible and ergodic}\}$$

that

A. *reduces the collection of ill-founded trees to the collection of ergodic* T *such that* $T \cong T^{-1}$, *and*
B. *reduces the collection of trees with at least two infinite branches to the collection of invertible ergodic* T *that have nontrivial centralizer.*

This theorem has corollaries analogous to Corollaries 8 and 9. For future reference we now make explicit what we have to show to prove Theorem 7.

We will build a continuous function $F : \mathcal{T}rees \to \mathcal{E}$ such that

Claim 10. 1. If $\mathcal{T}$ has an infinite branch, then

$$(\Sigma^{\mathbb{Z}}, \mathcal{C}, F(\mathcal{T}), \mathrm{sh}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, F(\mathcal{T}), \mathrm{sh}^{-1}).$$

2. If $\mathcal{T}$ has at least two infinite branches, then there is an element of the centralizer of $(\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh})$ that is not of the form sh^j for any $j \in \mathbb{Z}$.

The other claim that must be proved is:

Claim 11. 1. If $(\Sigma^{\mathbb{Z}}, \mathcal{C}, F(\mathcal{T}), \mathrm{sh}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, F(\mathcal{T}), \mathrm{sh}^{-1})$, then $\mathcal{T}$ has an infinite branch.

2. If there is an element of the centralizer of $(\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh})$ that is not of the form sh^j for any $j \in \mathbb{Z}$, then $\mathcal{T}$ has at least two infinite branches.

2.4. *Odometer maps.* A particular class of transformations will be important for our arguments. Let $\langle q_i : i \in \mathbb{N} \rangle$ be an infinite sequence of integers with $q_i \geq 2$. Then the sequence q_i determines an *odometer* transformation with domain the compact space

$$O =_{\mathrm{def}} \prod_i \mathbb{Z}_{q_i}.$$

The space O is naturally a monothetic compact abelian group, which, in order to avoid confusion, we denote by $\mathfrak{O}$ with the operation of addition and "carrying right". We will denote the group element $(1, 0, 0, 0, \dots)$ by $\bar{1}$, and the result of adding $\bar{1}$ to itself j times by $\bar{j}$. The collection of $\bar{j}$ for $j \in \mathbb{N}$ coincides with those elements of $\prod_i \mathbb{Z}_{q_i}$ that are eventually 0. For negative j, the $\bar{j}$'s correspond to those group elements that eventually agree with the sequence $\langle q_i - 1 : i \in \mathbb{N} \rangle$. In particular, $\{\bar{j} : j \in \mathbb{Z}\}$ is dense in $\prod_i \mathbb{Z}_{q_i}$.

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

The odometer transformation $\mathcal{O} : O \to O$ is defined by taking an $x \in \prod_i \mathbb{Z}_{q_i}$ and adding the group element $\bar{1}$, More explicitly, $\mathcal{O}(x)(0) = x(0) + 1 \mod (q_0)$ and $\mathcal{O}(x)(1) = x(1)$ unless $x(0) = q_0 - 1$, in which case we "carry one" and set $\mathcal{O}(x)(1) = x(1) + 1 \mod (q_1)$, etc.

The map $\mathcal{O} : O \to O$ is a topologically minimal, uniquely ergodic invertible homeomorphism that preserves the measure μ.

As with any abelian group the map $x \mapsto -x$ is an automorphism of $\mathfrak{O}$ of order 2. In particular it is an isomorphism between the measure preserving transformations $\mathcal{O}$ and $\mathcal{O}^{-1}$.

The characters $\chi \in \hat{\mathfrak{O}}$ are eigenfunctions for the odometer since

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

Since the characters form a basis for $L^2(\mathfrak{O})$, the odometer map has discrete spectrum. It is instructive for our purposes to describe these eigenfunctions explicitly.

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

Define

$$\mathcal{R}_n = \sum_{k=0}^{l_n-1} (e^{2\pi i/l_n})^k \chi_{A_k}.$$

Define $U_{\mathcal{O}} : L^2(\mathfrak{O}) \to L^2(\mathfrak{O})$ by setting $U_{\mathcal{O}}(f) = f \circ \mathcal{O}$. Then

1. $U_{\mathcal{O}}$ is the canonical unitary operator associated with $\mathcal{O}$;
2. $\mathcal{R}_n$ is an eigenvector of $U_{\mathcal{O}}$ with eigenvalue $e^{2\pi i/l_n}$;
3. $(\mathcal{R}_n)^{q_n} = \mathcal{R}_{n-1}$;
4. $\{(\mathcal{R}_n)^k : k \leq l_n, n \in \mathbb{N}\}$ form a basis for $L^2(\prod_i \mathbb{Z}_{q_i})$.

2.5. *Factors.* For measure preserving systems $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$, a factor map is a measure preserving map $\pi : X \to Y$ such that $\pi \circ T = S \circ \pi$ with $\pi^{-1}(\mathcal{C}) \subseteq \mathcal{B}$. There are several equivalent ways of viewing a factor Y. We will alternate between the following without further comment:

1. As an invariant complete sub-σ-algebra of $\mathcal{B}$.
2. As a closed subspace of $L^2(X)$ containing the constant functions that is closed under multiplication of bounded functions, truncation and the map $f \mapsto \bar{f}$, its complex conjugate.

2.6. *Joinings.* In this section we summarize some basic facts about joinings. Proofs of these facts can be found in [18].

We start with measure preserving systems $\mathbb{X} = (X, \mathcal{B}, \mu, T)$ and $\mathbb{Y} = (Y, \mathcal{C}, \nu, S)$. A *joining* of $\mathbb{X}$ and $\mathbb{Y}$ is a $T \times S$ invariant measure η on $X \times Y$ for which

1. all sets of the form $B \times C$ are measurable when $B \in \mathcal{B}$ and $C \in \mathcal{C}$ and is such that
2. for all $B \in \mathcal{B}, \eta(B \times Y) = \mu(B)$ and for all $C \in \mathcal{C}, \eta(X \times C) = \nu(C)$.[7]

We will let $\mathcal{J}(\mathbb{X}, \mathbb{Y})$ be the space of joinings under the weak*-topology.[8] The space is nonempty since it always contains the product measure $\mu \otimes \nu$. Then $\mathcal{J}(\mathbb{X}, \mathbb{Y})$ is a compact, convex Polish space. If $\mathbb{X}$ and $\mathbb{Y}$ are ergodic, then the extremal points of $\mathcal{J}(\mathbb{X}, \mathbb{Y})$ are the ergodic joinings.

A factor map $\phi : X \to Y$ gives rise to a joining defined by setting $\eta(B \times C) = \mu(B \cap \phi^{-1}(C))$. This joining is supported on the graph of ϕ. Conversely, any joining that is supported on the graph of a measurable function corresponds to a factor map. If this map is invertible, then it is a conjugacy. Thus, understanding the joinings gives a complete understanding of factor maps and conjugacies between $\mathbb{X}$ and $\mathbb{Y}$. Joinings supported on graphs of functions from X to Y we will call *graph joinings*.

[7] i.e. η has marginals μ and ν.

[8] The weak* topology can be defined independently of any particular topologies on $\mathbb{X}$ and $\mathbb{Y}$. See [7], [8] or [18] for more information.

Let $\phi : X \to Y$ be a factor map. For $A \subset X$ and $y \in Y$, let $A^y = \{x : \phi(x) = y\}$. We can define the *disintegration* of $\mathbb{X}$ over $\mathbb{Y}$. This is a family of measures $\langle \nu_y : y \in Y \rangle$, such that

1. ν_y is a standard probability measure on X^y;
2. For $A \in \mathcal{B}$, and $r \in [0,1]$ and $\epsilon > 0$, $\{y : A^y$ is ν_y measurable and $|\nu_y(A^y) - r| < \epsilon\}$ is ν-measurable; and
3. $\mu(A) = \int \nu_y(A^y) d\nu(y)$.

If η is a graph joining corresponding to a function ϕ, then η concentrates on $\{(x,y) : \phi(x) = y\}$ and we can identify it with the *disintegration* of μ over ν.

Suppose that $\mathbb{X}$ and $\mathbb{Y}$ are ergodic and we have factor maps $\phi_1 : \mathbb{X} \to \mathbb{Z}$ and $\phi_2 : \mathbb{Y} \to \mathbb{Z}$, where $\mathbb{Z} = (Z, \mathcal{D}, \rho, W)$. Then we get disintegrations $\langle \mu_z : z \in Z \rangle$ of μ and $\langle \nu_z : z \in Z \rangle$ of ν over ρ.

Definition 12. The *relatively independent joining* of $\mathbb{X}$ and $\mathbb{Y}$ over $\mathbb{Z}$ is the unique measure η on $X \times Y$ such that for all $A \in \mathcal{B}$ and $B \in \mathcal{C}$,

$$\eta(A \times B) = \int \mu_z(A^z)\nu_z(B^z)d\rho(z).$$

The measure η concentrates on $\{(x,y) : \phi_1(x) = \phi_2(y)\}$. If $\mathbb{Z}$ is trivial, then ρ is the product measure and we will refer to ρ simply as the *independent joining*.

Note that a relatively independent joining over some common factors will come from a conjugacy if and only if the two factors are the full algebras to begin with. Thus if we verify a joining is a relatively independent joining over some nontrivial factors then it does not come from a conjugacy.

In later sections will be working in a situation where there are $\psi_1 : \mathbb{X} \to \mathbb{Z}_1$ and $\phi_2 : \mathbb{Y} \to \mathbb{Z}_2$ and an isomorphism $\phi : \mathbb{Z}_1 \to \mathbb{Z}_2$. Using Definition 12 with $\phi_1 = \phi \circ \psi_1$ we are able to extend the invertible joining of $\mathbb{Z}_1$ and $\mathbb{Z}_2$ via ϕ to a joining of $\mathbb{X}$ and $\mathbb{Y}$.

We will use the following standard fact about self-joinings of monothetic groups that can be proved combining results of [8] and Corollary 2 of [9]:

THEOREM 13. *Let Γ be a compact group and $g \in \Gamma$ be such that $\{g^n : n \in \mathbb{Z}\}$ is dense in Γ.*[9] *Let T_g be the Haar measure preserving transformation given by translating by g. If η is an ergodic self-joining of T_g, then η is supported on the graph of the translation by some $h \in \Gamma$; i.e., is of the form*

$$\eta(A \times B) = \mu(A \cap h^{-1}B),$$

where μ is the Haar measure. Moreover, each $h \in \Gamma$ determines an invertible self-joining.

We will sometimes identify h with the self-joining determined by h.

[9] i.e., g is a witness that Γ is monothetic.

2.7. *Kronecker factors.* All ergodic measure preserving transformations have a maximal factor $\mathcal{K}$ isomorphic to a translation on a compact abelian group, in the sense that every other compact group factor is a factor of $\mathcal{K}$.

We now give a brief description of this factor, called the *Kronecker factor*. (See [7] or [23] for more information.) The Kronecker factor is trivial just in case the transformation is weakly mixing.

Given a standard probability space $(X, \mathcal{B}, \mu)$ and an invertible measure preserving transformation T, one can associate a unitary operator $U_T : L^2(X) \to L^2(X)$. The collection of eigenvalues for U_T form a countable subgroup G of the unit circle. It is possible to choose eigenfunctions $\{g_\lambda : \lambda \in G\}$ so that $g_{\lambda_1\lambda_2} = g_{\lambda_1} g_{\lambda_2}$.

The dual of G, $\widehat{G}$ is a compact monothetic subgroup. The powers of the inclusion character $i : G \to \mathbb{T}$ are dense in $\widehat{G}$. Letting ν be the Haar measure on $\widehat{G}$, the map T_i defined by $g \mapsto [i]g$ is an ergodic measure preserving transformation of $(\widehat{G}, \mathcal{C}, \nu)$. Each $\lambda \in G$ gives an eigenfunction f_λ for $\widehat{G}$ defined by setting $f_\lambda(\chi) = \chi(\lambda)$. The $\{f_\lambda : \lambda \in G\}$ span $L^2(\widehat{G})$ and the embedding e defined by sending $f_\lambda \mapsto g_\lambda$ gives a linear, multiplicative injection of $L^2(\widehat{G})$ into $L^2(X)$ of norm 1. In particular there is a factor map $\pi : X \to \widehat{G}$ corresponding to this embedding.

The *Kronecker factor* of $\mathbb{X}$ is this $\mathcal{K} = (\widehat{G}, \mathcal{C}, \nu, T_i)$. Clearly if $\phi : X \to Y$ is an isomorphism between transformations T and S, then ϕ induces a canonical isomorphism between the Kronecker factors of Y and X.

More generally, if $(K, \mathcal{D}, \eta, T_k)$ is an ergodic translation on a compact group K and $\pi : X \to K$ is a factor map, then the map $f \mapsto f \circ \pi : L^2(K) \to L^2(X)$ is an isometry into $L^2(X)$ that sends eigenfunctions to eigenfunctions. Since K is a compact group we can choose a spanning set $\{k_\lambda : \lambda \in G'\}$ for $L^2(K)$ consisting of eigenvectors such that $k_{\lambda_1\lambda_2} = k_{\lambda_1} k_{\lambda_2}$. Then the map $k_\lambda \mapsto f_\lambda$ determines a multiplicative isometry of $L^2(K)$ into $L^2(\widehat{G})$, and hence a factor map of $\widehat{G}$ to K.

2.8. *Generic sequences.* We will use some facts about symbolic shifts that are corollaries of the ergodic theorem. (See [17] for more details.) We recall that for a finite sequence $s \in \Sigma^{<\mathbb{N}}$ the frequency of occurrences of s in $x \in \Sigma^{\mathbb{Z}}$ is defined as

$$\lim_{n\to\infty} \frac{1}{n} |\{0 \leq i < n : \text{s occurs in } x \text{ starting at } i\}|.$$

An equivalent formulation of this is:

$$\lim_{n\to\infty} \frac{1}{n} \Sigma_0^{n-1} \chi_{\langle s\rangle} \mathrm{sh}^i(x),$$

where $\chi_{\langle s\rangle}$ is the characteristic function of the basic open set $\langle s\rangle$. Not all sequences s have a frequency in an arbitrary x.

LEMMA 14. *Suppose that* $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, \mathrm{sh})$ *is a measure preserving system. Then for* ν*-almost all* x, *and all finite sequences* $s \in \Sigma^{<\mathbb{N}}$ *the frequency of the occurrences of* s *in* x *exists. Moreover,* ν *is ergodic if and only if for* ν*-almost all* x, *for all finite sequences* $s \in \Sigma^{<\mathbb{N}}$, *the frequency of* s *in* x *is* $\nu(\langle s \rangle)$.

We will use the following generalization:

LEMMA 15. *Let* $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, \mathrm{sh})$ *be a measure preserving system. Then there is a set* G *of* ν *measure one such that for all finite sequences* $s \in \Sigma^{<\mathbb{N}}$ *and all sequences* $a_n \leq 0 \leq b_n$ *with* $b_n - a_n \to \infty$ *as* $n \to \infty$, *and all* $x \in G$ *the frequency of the occurrences of* s *in* x *exists and is equal to*

$$\lim_{n \to \infty} \left(\frac{1}{b_n - a_n + 1} \right) |\{ i : a_n \leq i \leq b_n : s \textit{ occurs in } x \textit{ starting at } i \}|.$$

Moreover, ν *is ergodic iff for* ν*-almost all* x, *and all* s *this limit is equal to* $\nu(\langle s \rangle)$.

We will call an x with the properties of the lemma *generic* for ν.

3. Trees, groups and equivalence relations

In this section we introduce the groups we will use to approximate conjugacies and state some lemmas relating them to trees and equivalence relations.

3.1. *Groups of involutions and trees.* The groups we build in our construction will be direct sums of $\mathbb{Z}_2$. We will call such a group a *group of involutions.*

If $G = \sum_{i \in I} (\mathbb{Z}_2)_i$ and we have a distinguished basis $B = \{ r_i : i \in I \}$ for G, then there is a well-defined notion of *parity*: a $g \in G$ is *even* if it can be written as a sum of an even number of elements of B and odd otherwise.

Parity is preserved under homomorphisms in the sense that if G has distinguished basis B and H has distinguished basis C and $\phi : H \to G$ is a homomorphism sending C to B, then ϕ sends even elements to even elements and odd elements to odd elements.

As a consequence, if we are given an inverse limit system of groups of involutions $\{G_s : s \in S\}$ over a linearly ordered set $(S, <_S)$ with maps $\{\rho_{t,s} : s <_S t\}$, where each G_s is a group of involutions with a distinguished set of generators and $G = \varprojlim G_s$ is the inverse limit, then the elements of G have a well-defined parity.

If $\mathcal{T} \subset \mathbb{N}^{<\mathbb{N}}$ is a tree, either finite or infinite, then to each level s of $\mathcal{T}$ we can associate a group of involutions $G_s(\mathcal{T})$ by taking the sum of copies of $\mathbb{Z}_2$ indexed by the nodes on $\mathcal{T}$ at level s. We will view the nodes of $\mathcal{T}$ at level s as the distinguished generators of $G_s(\mathcal{T}) = \sum_{\tau \in \mathcal{T},\ \mathrm{lh}(\tau) = s} (\mathbb{Z}_2)_\tau$. If $s < t$ are levels of $\mathcal{T}$, then we get a canonical homomorphism $\rho_{t,s} : G_t(\mathcal{T}) \to G_s(\mathcal{T})$ that sends a generator τ of $G_t(\mathcal{T})$ at level t to the unique generator σ of $G_s(\mathcal{T})$ that is an initial segment of τ.

Definition 16. $G_\infty(\mathcal{T})$ is the inverse limit of $\langle G_s(\mathcal{T}), \rho_{t,s} : s < t < \infty\rangle$. Let $\rho_s : G_\infty(\mathcal{T}) \to G_s(\mathcal{T})$ be the projection map to $G_s(\mathcal{T})$.

The following lemma is immediate as there is a one-to-one correspondence between the infinite branches of $\mathcal{T}$ and infinite sequences of canonical generators $g_s \in G_s(\mathcal{T})$ such that $\rho_{t,s}(g_t) = g_s$ for $t > s$.

LEMMA 17. [10] *Let $\mathcal{T} \subset \mathbb{N}^{<\mathbb{N}}$ be a tree. Then*

1. *$G_\infty(\mathcal{T})$ has a nonidentity element of odd parity if and only if $\mathcal{T}$ is ill-founded.*
2. *$G_\infty(\mathcal{T})$ has a nonidentity element of even parity if and only if $\mathcal{T}$ has at least two infinite branches.*

3.2. *Group actions.* Let X be a set and Q an equivalence relation on X. We will consider group actions on quotient spaces X/Q, or equivalently group actions on the partition associated with Q. During the construction we will need to control systems of such group actions on finer and finer equivalence relations. We now develop some definitions for this purpose.[11]

Definition 18. Suppose:

1. Q and R are equivalence relations on a set X with R refining Q.
2. G and H are groups with G acting on X/Q and H acting on X/R.
3. $\rho : H \to G$ is a homomorphism.

Then we will say that the H action is *subordinate* to the G action if for all $x \in X$, whenever $[x]_R \subset [x]_Q$ we have $h \cdot [x]_R \subseteq \rho(h) \cdot [x]_Q$.

If X is a set we will write ${}^nX$ for the n-fold concatenations $x_0x_1 \cdots x_{n-1}$ of elements from X. While ${}^nX$ can be identified with the Cartesian product X^n, it is convenient in some contexts to distinguish the two objects. For example, if X is a set of words W in a language Σ, then ${}^nX$ is a natural way of creating another collection of words, namely the n-fold concatenations of elements of X. If G acts on X, then there is a canonical *diagonal* action of G on ${}^nX$ defined by $g(x_0x_1 \cdots x_{n-1}) = gx_0\ gx_1 \cdots gx_{n-1}$.

If G is a group of involutions with a distinguished collection of free generators, then we define the *skew diagonal* action on ${}^nX$ by setting

$$g(x_0x_1x_2 \cdots x_{n-1}) = gx_{n-1}\ gx_{n-2} \cdots gx_0$$

[10]This lemma can obviously be rephrased as saying certain sets are complete analytic. As this is not essential to the development of the theorem we do not do it here.

[11]If Q and R are equivalence relations with R refining Q, then, considered as a set of ordered pairs, $R \subseteq Q$. Hence we write $R \subseteq Q$ to mean that R refines Q.

for g a canonical generator. Note that the skew diagonal actions of elements of G with odd parity reverse the orders of the x_i's, while elements of even parity preserve the order.

If we have an equivalence relation Q on X, then we can define the *product* equivalence relation ${}^nQ$ on ${}^nX$ by setting

$$x_0x_1\cdots x_{n-1} \sim x'_0x'_1\cdots x'_{n-1}$$

if and only if for all i, $x_i \sim x'_i$.

Using the obvious identification of ${}^nX/{}^nQ$ with ${}^n(X/Q)$, we can extend an action of G on X/Q to the diagonal and skew diagonal actions on ${}^nX/{}^nQ$.

4. General facts about our transformations

To prove Theorem 7, we will build a continuous, one-to-one function

$$F : \mathcal{T}rees \to \mathcal{E}$$

that reduces the collection of ill-founded trees to the set of ergodic ν isomorphic to ν^{-1} and the collection of trees with at least two infinite branches to the set of ergodic transformations with nontrivial centralizer.

One heuristic view of our construction is that we are being given information about a particular tree $\mathcal{T}$ by specifying which σ_n belong to $\mathcal{T}$. As we learn more about which σ_n belong to $\mathcal{T}$ we give a decreasing sequence of open sets in $\mathcal{E}$ by giving information about the measures of the basic open sets determined by longer and longer finite words in $\Sigma^{<\mathbb{N}}$. These numbers give a complete description of the measure $F(\mathcal{T})$.

An equivalent alternative point of view is that we are simultaneously constructing $F(\mathcal{T})$ for all trees $\mathcal{T}$ by specifying, for all possible tree-like subsets T of $\{\sigma_m : m \leq n\}$, a collection of words in $\Sigma^{<\mathbb{N}}$ in a coherent way.

In the first part of our exposition we gradually develop a collection of specifications on the sets of words that we build, verifying their sufficiency for the necessary claims as we present them. In the second part of the exposition, we collect these specifications and prove that we can construct sets of words that satisfy them.

All of our words will be in a basic language $\Sigma = \{0, 1\}$. We fix in advance a fast growing sequence of odd prime numbers $\langle p_n : n \in \mathbb{N}\rangle$.[12] For each n with $\sigma_n \in \mathcal{T}$ we will construct a set of words $\mathcal{W}_n = \mathcal{W}_n(\mathcal{T})$. To construct $\mathcal{W}_n(\mathcal{T})$ we need only have knowledge of $\mathcal{T} \cap \{\sigma_m : m \leq n\}$. By this we mean that if:

$$\mathcal{T} \cap \{\sigma_m : m \leq n\} = \mathcal{T}' \cap \{\sigma_m : m \leq n\}, \tag{1}$$

[12] The only point of the prime numbers is to make sure that our function takes distinct trees to transformations with nonisomorphic Kronecker factors. This is not strictly necessary for the proof of the theorem.

then for all $m \leq n$

$$\mathcal{W}_m(\mathcal{T}) = \mathcal{W}_m(\mathcal{T}'). \tag{2}$$

This allows us to view this as a construction of $F(\mathcal{T})$ for all $\mathcal{T} \in \mathcal{T}rees$ simultaneously.

We will let $\mathcal{W}(\mathcal{T}) = \bigcup_{\sigma_n \in \mathcal{T}} \mathcal{W}_n(\mathcal{T})$. Having described $\mathcal{W}(\mathcal{T})$ we will get a dynamical system as a subshift $\mathbb{K} \subseteq \Sigma^{\mathbb{Z}}$ by taking as its domain those doubly infinite sequences of words in Σ all of whose finite subwords occur somewhere in some $w \in \mathcal{W}(\mathcal{T})$. Clearly $\mathbb{K}$ is shift invariant, and we show that it is a uniquely ergodic and topologically minimal system. The measure $F(\mathcal{T})$ we associate with $\mathcal{T}$ is the unique shift invariant measure on $\mathbb{K}$.

4.1. *Basic specifications.* To start off we let $\mathcal{W}_0 = \Sigma$ and $l_0 = 1$. We construct $\mathcal{W}_n(\mathcal{T})$ just in case $\sigma_n \in \mathcal{T}$ and the construction will depend only on $\mathcal{T} \cap \{\sigma_m : m \leq n\}$. For simplicity we write $\mathcal{W}_n$ and $\mathcal{W}$ for $\mathcal{W}_n(\mathcal{T})$ and $\mathcal{W}(\mathcal{T})$.

- E1. Any pair w_1, w_2 of words in $\mathcal{W}_n$ have the same length l_n. The number l_n will be a product of powers of $\{p_i : \sigma_i \in \mathcal{T}, i \leq n\} \cup \{2\}$. The cardinality W_n of $\mathcal{W}_n$ will be a power of 2 and $W_{n+1} > W_n$.
- E2. If σ_m and σ_n are consecutive elements of $\mathcal{T}$, then every word in $\mathcal{W}_n$ is built by concatenating words in $\mathcal{W}_m$. There is an integer f_m such that every word in $\mathcal{W}_m$ occurs in each word of $\mathcal{W}_n$ exactly f_m times. The number f_m will be a power of p_n

If $f_m = p_n^2$ we see that $l_n = l_m p_n^2 W_m$. The next specification allows us to parse our sequences precisely.

- E3. (Unique readability) If σ_m and σ_n are consecutive elements of $\mathcal{T}$ and $w \in \mathcal{W}_n$ and
$$w = bw_1 \cdots w_k e,$$
where b or e have length less than l_m and each $w_i \in \mathcal{W}_m$, then both b and e are the empty word. If $w, w' \in \mathcal{W}_n$ and $w = w_1 w_2 \cdots w_{l_n/l_m}$ and $w' = w'_1 w'_2 \cdots w'_{l_n/l_m}$ with $w_i, w'_i \in \mathcal{W}_m$, and $k = [l_n/2l_m] + 1$, then $w_k w_{k+1} \cdots w_{l_n/l_m} \neq w'_1 w'_2 \cdots w_{l_n/l_m - [k] - 1}$; i.e., the first half of w' is not equal to the second half of w.

We now sketch the following elementary lemma:

PROPOSITION 19. *Let Σ be a finite language, and $\mathcal{W} = \bigcup_{\sigma_n \in \mathcal{T}} \mathcal{W}_n$, where the sequence $\langle \mathcal{W}_n : \sigma_n \in \mathcal{T} \rangle$ satisfies the specification* E1–E3. *Suppose that $\mathbb{K} \subseteq \Sigma^{\mathbb{Z}}$ is the set of doubly infinite sequences all of whose finite subwords occur somewhere in some $w \in \mathcal{W}$. Then $\mathbb{K}$ is a shift invariant closed subset of $\Sigma^{\mathbb{Z}}$ that is topologically minimal and uniquely ergodic.*

Proof. That $\mathbb{K}$ is closed and shift invariant is immediate. We first show that $\mathbb{K}$ is topologically minimal. Since $\{\mathrm{sh}^k\langle w \rangle : k \in \mathbb{Z}$ and $w \in \mathcal{W}_n$, for some

n with $\sigma_n \in \mathcal{T}\}$ form a subbase for the topology of $\mathbb{K}$, it suffices to show that if $w \in \mathcal{W}_n$ for some n and $x \in \mathbb{K}$ there is an $k \in \mathbb{Z}$ such that $x \restriction [k, k+l_n) = w$.

Suppose that $m > n$ with $\sigma_m \in \mathcal{T}$. The sequence $x \restriction [0, 2l_m)$ must contain some word $w' \in \mathcal{W}_m$. By E2, w must occur at least once in w'. Hence there is such a k and, moreover, $0 < k < k + l_n < 2l_m$.

The existence of at least one invariant measure follows from the compactness of $\mathbb{K}$. Suppose that μ is an arbitrary invariant measure. Then by the Ergodic Theorem and specification E2, for each $w \in \mathcal{W}_m$, $\mu(\langle w \rangle) = f_m/l_n$, where $n > m$ is least such that $\sigma_n \in \mathcal{T}$. Since the elements of $\bigcup_n \mathcal{W}_n$ determine a subbase for the open sets this implies that μ is the unique invariant measure. □

We now show that our specifications imply that the function F is continuous.

Proposition 20. *The function $F : \mathcal{T}rees \to \mathcal{E}$ sending a tree $\mathcal{T}$ to the unique shift invariant measure on $\mathbb{K}$ is continuous.*

Proof. Each function $r : \Sigma^l \to [0, 1]$ and $\epsilon > 0$ determines a basic open set

$$N_{r,\epsilon} = \{\nu : \text{for all } w \in \Sigma^l, |\nu(\langle w \rangle) - r(w)| < \epsilon\}$$

in the weak* topology relative to the shift invariant measures on $\Sigma^{\mathbb{Z}}$. Let $C \subseteq \mathbb{N}$ be an arbitrary fixed infinite set. Then the collection of $N_{r,\epsilon}$, where r has domain Σ^l and $l \in C$ forms a subbase for the weak* topology on the invariant measures.

Fix such an r and ϵ. Let $\mathcal{T}$ be such that $F(\mathcal{T}) \in N_{r,\epsilon}$. We can assume that the domain of r consists of words of length l_n for some n with $\sigma_n \in \mathcal{T}$. Let $N > n$ be such that $\sigma_N \in \mathcal{T}$. The set of sequences that belong to some $\mathcal{W}_m(\mathcal{T})$ with $m \leq N$ as well as the set of numbers f_m and l_m with $m < N$ are determined by $\mathcal{T} \cap \{\sigma_m : m \leq N\}$. Since these determine the values of $F(\mathcal{T})$ on the basic open sets in $\{\langle w \rangle \subseteq \Sigma^{\mathbb{Z}} : w \in \Sigma^{l_n}\}$, we see that if $\mathcal{T}' \cap \{\sigma_m : m \leq N\} = \mathcal{T} \cap \{\sigma_m : m \leq N\}$, then $F(\mathcal{T}') \in N_{r,\epsilon}$. If follows that $F^{-1}(N_{r,\epsilon})$ is an open set in $\mathcal{T}rees$. □

4.2. *Specifications for the group actions.* In order to make the elements of $G_\infty(\mathcal{T})$ correspond to conjugacies we build symmetries into our construction using finite approximations to $G_\infty(\mathcal{T})$. Towards this end we let G_0^n be the trivial group and for $s > 0$,[13]

$$G_s^n = \sum \{(\mathbb{Z}_2)_\tau : \tau \in \mathcal{T} \cap \{\sigma_m : m \leq n\},\ \mathrm{lh}(\tau) = s\}.$$

[13]Since the G_s^n's depend on $\mathcal{T}$, perhaps we should write $G_s^n(\mathcal{T})$ instead of G_s^n.

Clearly, the G^n_s's vary continuously with $\mathcal{T}$ and G^{n+1}_s factors naturally as $G^{n+1}_s = G^n_s \oplus H$ where H is either trivial or $\mathbb{Z}_2$. Moreover, the direct limit of the G^n_s's via these embeddings is $G_s(\mathcal{T})$. Since the enumeration of $\mathbb{N}^{<\mathbb{N}}$ has the property that every initial segment of a given $\sigma \in \mathbb{N}^{<\mathbb{N}}$ is enumerated before σ, the restrictions of the canonical homomorphisms $\rho_{s,s-1} : G_s \to G_{s-1}$ map G^n_s to G^n_{s-1}. We will use the observation that $|\bigcup_s G^m_s| \leq 2^m$.

Recall from Definition 2 that the continuous function $s : \mathcal{T}rees \to \mathbb{N}^{\mathbb{N}}$ is defined by setting $s(\mathcal{T})(n) = s$ if and only if s is the length of the longest sequence in $\mathcal{T} \cap \{\sigma_m : m \leq n\}$ and write $s(n)$ for $s(\mathcal{T})(n)$ if $\mathcal{T}$ is fixed in context. For those $s \leq s(n)$, there will be equivalence relations $\mathcal{Q}^n_s$ on $\mathcal{W}_n$. The $\mathcal{Q}^n_s$ will naturally induce an equivalence relation on $\mathrm{rev}(\mathcal{W}_n)$ which we will also call $\mathcal{Q}^n_s$. We will write Q^n_s for the number of equivalence classes in $\mathcal{Q}^n_s$. Each class of $\mathcal{Q}^n_s$ will have the same number of elements which we will write C^n_s.

Some numerics. We will have a decreasing sequence of numbers ϵ_n and δ_s that go to zero rapidly. We will have the following numerical relations.

If $\mathcal{T} = \langle \sigma_{n_i} : i \in \mathbb{N}\rangle$, then

$$2\epsilon_{n_i} W^2_{n_i} < \epsilon_{n_{i-1}} \tag{3}$$

and

$$\epsilon_{n_i}(l_{n_i}/l_{n_{i-1}})W^{-2}_{n_{i-1}} \to \infty \text{ as } n \to \infty. \tag{4}$$

We want our ϵ_n's to go to zero fast enough:

$$\prod_{n\in\mathbb{N}} (1 - \epsilon_n) > 0. \tag{5}$$

We note that all of our numerical requirements can be satisfied by deciding on W_n, taking ϵ_n small enough and l_n large enough.

Some specifications. The groups and equivalence relations will satisfy the following specifications. The first two describe the diameter and separation of the $\mathcal{Q}_s$ classes. To start we let $\mathcal{Q}^0_0$ be the equivalence relation on $\mathcal{W}_0 = \Sigma$ which has one equivalence class; i.e., any two elements of Σ are equivalent.

Q4. Suppose that $n = M(s)$. Then any two words in the same $\mathcal{Q}^n_s$ equivalence class agree on an initial segment of length at least $(1 - \epsilon_n)l_n$.

Q5. For $n \geq M(s) + 1$, $\mathcal{Q}^n_s$ is the product equivalence relation of $\mathcal{Q}^{M(s)}_s$.[14] Hence we can view $\mathcal{W}_n/\mathcal{Q}^n_s$ as sequences of elements of $\mathcal{W}_{M(s)}/\mathcal{Q}^{M(s)}_s$ and similarly for $\mathrm{rev}(\mathcal{W}_n)/\mathcal{Q}^n_s$.

It follows that $\mathcal{Q}^n_0$ is the equivalence relation on $\mathcal{W}_n$ which has one equivalence class.

[14] This makes sense since the members of $\mathcal{W}_n$ are made by concatenating members of $\mathcal{W}_s$.

Specification Q5 links $\mathcal{Q}_s^n$ directly to $\mathcal{Q}_s^{M(s)}$. When the exponent is not relevant we will refer to the $\mathcal{Q}_s^n$ as $\mathcal{Q}_s$. For $u \in \mathcal{W}_n$ we will write $[u]_s$ for its $\mathcal{Q}_s^n$ class.

Q6. $\mathcal{Q}_{s+1}^n$ refines $\mathcal{Q}_s^n$ and each $\mathcal{Q}_s^n$ class contains $2^{k(n)}$ many $\mathcal{Q}_{s+1}^n$ classes for some number $k(n) \to \infty$ as $n \to \infty$.

We introduce another numerical requirement: If σ_m and σ_n are consecutive elements of $\mathcal{T}$, then

$$2^{-k(n)+m} < \epsilon_m. \tag{6}$$

This can be satisfied by making the $k(n)$ grow fast enough.

A7. G_s^n acts freely on $\mathcal{W}_n/\mathcal{Q}_s^n \cup \mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$ and the G_s^n action is subordinate to the G_{s-1}^n action via the natural homomorphism $\rho_{s,s-1}$ from G_s^n to G_{s-1}^n.

A8. The canonical generators of $G_s^{M(s)}$ send elements of $\mathcal{W}_{M(s)}/\mathcal{Q}_s^{M(s)}$ to elements of $\mathrm{rev}(\mathcal{W}_{M(s)})/\mathcal{Q}_s^{M(s)})$ and vice versa by reversing the words.

A9. If $M(s) < n$, σ_m and σ_n are consecutive elements of $\mathcal{T}$ and we view $G_s^n = G_s^m \oplus H$, then the action of G_s^m on $\mathcal{W}_m/\mathcal{Q}_s^m \cup \mathrm{rev}(\mathcal{W}_m/\mathcal{Q}_s^m)$ is extended to an action on $\mathcal{W}_n/\mathcal{Q}_s^n \cup \mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$ by the skew diagonal action. If H is nontrivial, then its canonical generator maps $\mathcal{W}_n/\mathcal{Q}_s^n$ to $\mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$.

We will refer to the words in some $\mathcal{W}_n$ as having *even parity* and the words in some $\mathrm{rev}(\mathcal{W}_n)$ as having *odd parity*.

These specifications suffice to verify directly that if $\mathcal{T}$ has an infinite branch, then $F(\mathcal{T})$ is conjugate to $F(\mathcal{T})^{-1}$, and if there are at least two branches then $F(\mathcal{T})$ has nontrivial centralizer. For a more unified presentation we wait to show these properties until we discuss the canonical factors.

We will prove a lemma that we will use for later calculations.

Proposition 21. *Let $g \in G_s^i$ for $i < n$ and suppose that σ_m and σ_n are consecutive elements of $\mathcal{T}$. Let $\mathcal{W}_n' = \mathcal{W}_n$ or $\mathrm{rev}(\mathcal{W}_n)$, according to whether g has even or odd parity. Among the pairs $(u,v) \in \mathcal{W}_n \times \mathcal{W}_n'$ for which $g[u]_s = [v]_s$, the proportion for which there is an $h \in G_{s+1}^m$ with $h[u]_{s+1} = [v]_{s+1}$ is at most*

$$\frac{|G_{s+1}^m|}{2^{k(n)}}. \tag{7}$$

Proof. Fix a $\mathcal{Q}_s^m$ class C. Then $C \times gC = \bigcup\{C_1 \times C_2 : C_1 \subset C, C_2 \subset gC$ and C_1, C_2 are $\mathcal{Q}_{s+1}^m$-classes$\}$. The number of pairs (C_1, C_2) with $C_1 \times C_2 \subseteq C \times gC$ for which there is some $h \in G_{s+1}^m$ with $hC_1 = C_2$ is $|G_{s+1}^m| 2^{k(n)}$. This has proportion $|G_{s+1}^m| 2^{-k(n)}$ in $C \times gC$. Since C was arbitrary, the lemma is proved. $\square$

5. **The canonical factors**

In this section we describe a canonical tower of factors we will use to understand the joinings between the transformations we build. Specifications E1–E3 determine an odometer factor of our transformations. It will turn out that this odometer factor is the Kronecker factor of each the transformations. The odometer factor will determine a "blocking" of the elements of our space and the equivalence classes that appear in the blocking give the canonical factors.

For the moment we will assume that we are working with a fixed arbitrary measure ν of the form $F(\mathcal{T})$, where $\mathcal{T} \in \mathcal{T}rees$. We let $\mathbb{K} \subseteq \Sigma^{\mathbb{Z}}$ be the space corresponding to ν.

5.1. *The odometer factor.* Let $x \in \mathbb{K}$. Suppose that σ_m and σ_n are consecutive members of $\mathcal{T}$. Then by specification E3 there are unique $w, w' \in \mathcal{W}_n$ and $k \in [0, l_n - 1]$ such that $x \restriction [-k, -k + 2l_n - 1] = ww'$. We will call the interval of integers $[-k, -k + l_n - 1]$ the *n-block of x containing* 0.

If the m-block of x containing 0 is $[-k', -k' + l_m - 1]$, then there is a unique $\pi(x, n)$ such that $k' = -k + \pi(x, n)l_m$. Less formally, if u is the word in $\mathcal{W}_m$ sitting on the m-block in x containing 0, then u is the $\pi(x, n)^{\text{th}}$ word occurring in the word in $\mathcal{W}_n$ sitting on the n-block of x containing 0.

We can view $\pi(x, n) \in \mathbb{Z}_{l_n/l_m} = \mathbb{Z}_{p_n^2 W_m}$ and view $\pi(x, n)$ as a function of x:

$$\pi : \mathbb{K} \to \prod\{\mathbb{Z}_{p_n^2 W_m} : \sigma_m \text{ and } \sigma_n \text{ are succesive elements of } \mathcal{T}\}.$$

When we look at $\pi(sh(x), \cdot)$ we see that the element $\pi(x, \cdot)$ has been changed by adding $\bar{1}$ in the odometer.

Let $\mathcal{O}_{\mathcal{T}}$ be the odometer transformation on $O_{\mathcal{T}} =_{\text{def}} \prod\{\mathbb{Z}_{p_n^2 W_m} : \sigma_m$ and σ_n are successive elements of $\mathcal{T}\}$. We will use the following facts. The second assertion follows easily from the results of Section 2.4:

LEMMA 22. *The map $\pi : \mathbb{K} \to \prod_{\sigma_n \in \mathcal{T}} \mathbb{Z}_{p_n^2 W_m}$ is a factor map from $F(\mathcal{T})$ to $\mathcal{O}_{\mathcal{T}}$. If $p > 2$ is a prime number, then $e^{2\pi i/p}$ is an eigenvalue of the unitary operator associated with $\mathcal{O}_{\mathcal{T}}$ if and only if $p = p_n$ for some n with $\sigma_n \in \mathcal{T}$.*

We can associate the same odometer transformation $\mathcal{O}_{\mathcal{T}}$ to $\mathbb{K}^{-1}$ with the analogous procedure, Let $\pi^* : \mathbb{K}^{-1} \to \prod_{\sigma_n \in \mathcal{T}} \mathbb{Z}_{p_n^2 W_m}$ be the analogous map. If i is the involution of $\mathcal{O}_{\mathcal{T}}$ given by $x \mapsto -x$, then $\pi^*(x^{-1}) = i(\pi(x))$.

Since π is the initial element in a sequence of projections about to be defined we will refer to it from now on as π_0.

5.2. *The canonical factors.* In this section we define a canonical sequence of invariant sub-σ-algebras of the algebra $\mathcal{B}(\mathbb{K})$ of measurable subsets of $\mathbb{K}$.

Let $\mathbb{K}_0$ be $O_{\mathcal{T}}$. For each s, $\mathcal{Q}_s^{M(s)}$ is an equivalence relation on $\mathcal{W}_{M(s)}$. Enumerate the classes $\{c_j : j < Q_s^{M(s)}\}$. A typical $x \in \mathbb{K}$ gives a well-defined bi-infinite sequence of such classes. This gives a shift invariant map from $\mathbb{K}$ to $\{0, 1, \ldots, Q_s^{M(s)} - 1\}^{\mathbb{Z}}$.

More formally: define a map

$$\pi_s^- : \mathbb{K} \to \{0, 1, \ldots, Q_s^{M(s)} - 1\}$$

as follows. For $x \in \mathbb{K}$ consider the word $w \in \mathcal{W}_{M(s)}$ on the $M(s)$-block of x containing 0. If $[w]_s = c_j$, let $\pi_s^-(x) = j$.

Let $\mathbb{K}_s$ be the space $\{0, 1, \ldots, Q_s^{M(s)} - 1\}^{\mathbb{Z}} \times O_{\mathcal{T}}$. The map π_s^- determines a continuous shift invariant map

$$\pi_s : \mathbb{K} \to \mathbb{K}_s$$

defined by letting the first coordinate of $\pi_s(x)$ be the $\mathbb{Z}$-sequence $\langle \pi_s^-(x)(k) : k \in \mathbb{Z}\rangle$ and the second coordinate be $\pi_0(x)$.[15]

We can describe a nice base for the topology on $\mathbb{K}_s$. For $n \geq M(s), w \in \mathcal{W}_n$ and $k < l_n$, we let $\langle [w]_s, k\rangle$ be the collection of $(x, y) \in \mathbb{K}_s$ such that 0 is at the k^{th} place in the n-block B of x containing it, and if $v \in \{0, 1, \ldots, Q_s^{M(s)} - 1\}^{l_n}$ is the word in x at B, then v is the sequence of $\mathcal{Q}_s^{M(s)}$ classes given by $[w]_s$. The collection of $\langle [w]_s, k\rangle$ for $\sigma_n \in \mathcal{T}$, $w \in \mathcal{W}_n$ and $k \leq l_n$ form a basis for the topology of $\mathbb{K}_s$ consisting of clopen sets.

Define a measure $\nu_s =_{\text{def}} \pi_s^*(\nu)$, where ν is the measure on $\mathbb{K}$ given by $F(\mathcal{T})$. The measure ν_s can be described explicitly using specifications Q6 and E2: if A and B are sets from our basis arising from two words in $\mathcal{W}_n$, then $\nu_s(A) = \nu_s(B) = 1/Q_s^n$.

Let $\mathcal{H}_0$ be the shift invariant sub-σ-algebra of $\mathcal{B}(\mathbb{K})$ generated by the collection of $\pi_0^{-1}(B)$, where B is a basic open set in $O_{\mathcal{T}}$. Let $\mathcal{H}_s$ be the shift invariant sub-σ-algebra of $\mathcal{B}(\mathbb{K})$ generated by $\{\pi_s^{-1}(B) : B$ is a basic open set in $\{0, 1, \ldots, Q_s^{M(s)} - 1\}^{\mathbb{Z}} \times O_{\mathcal{T}}\}$. Then $\mathcal{H}_s$ is the sub-σ-algebra determined by the factor map π_s.

If $l = l_{M(s+1)}/l_{M(s)}$, then the equivalence relation $\mathcal{Q}_{s+1}^{M(s+1)}$ refines ${}^l\mathcal{Q}_s^{M(s)}$. Consequently the algebra $\mathcal{H}_{s+1} \supseteq \mathcal{H}_s$. The factor map $\pi_{s+1,s} : \mathbb{K}_{s+1} \to \mathbb{K}_s$ can be computed explicitly by noting that each $\mathcal{Q}_{s+1}^{M(s+1)}$ class is contained in a $\mathcal{Q}_s^{M(s+1)}$ class, which in turn is an $l_{M(s+1)}/l_{M(s)}$-tuple of $\mathcal{Q}_s^{M(s)}$-classes. Hence a $\mathbb{Z}$-sequence of $\mathcal{Q}_{s+1}^{M(s+1)}$ classes determines a $\mathbb{Z}$ sequence of $\mathcal{Q}_s^{M(s)}$-classes. The factor map is continuous.

[15]It will follow from later specifications that $\mathbb{K}_0$ is a factor of the projection of $\mathbb{K}$ to $\{0, 1, \ldots, Q_s^{M(s)} - 1\}^{\mathbb{Z}}$, making the second term of $\mathbb{K}_s$ redundant. However we do not need this for the argument.

The disintegration $\langle \nu_x : x \in \mathbb{K}_s \rangle$ of the measure ν_{s+1} over ν_s are similarly approximated: $v \in \mathcal{W}_n$ and $k \in \mathbb{Z}$ determine a basic open set $\langle [v]_s, k\rangle$ in $\mathbb{K}_s$. The inverse image of $\langle [v]_s, k\rangle$ in $\mathbb{K}_{s+1}$ is a finite union of basic open sets $\langle [w]_{s+1}, k\rangle$. Each of these has the same ν_{s+1} measure and the same measure by ν_x for each $x \in \langle [v]_s, k\rangle$.

Proposition 23. *$\mathcal{B}(\mathbb{K})$ is the smallest invariant σ-algebra that includes $\bigcup_{s\in\mathbb{N}} \mathcal{H}_s$.*

Proof. It suffices to show that if $u \in \mathcal{W}_m$, then the basic open set $\langle u\rangle \subseteq \mathbb{K}$ is arbitrarily well-approximated in measure by elements of $\bigcup_{s\in\mathbb{N}} \mathcal{H}_s$. We will use specification Q4.

Fix $\epsilon > 0$. Choose an $n > m$ such that $n = M(s)$ for some s and large enough that $\epsilon_n + \frac{l_m}{l_n} < \epsilon$.

Let $G \subset \mathbb{K}$ be the collection of x such that $x(0)$ is not among the last $(\epsilon_n + \frac{l_m}{l_n})l_n$ letters of the word in $\mathcal{W}_n$ sitting at the n-block of x containing 0. By the ergodic theorem, the measure of G is at least $1 - \epsilon$.

Let $x \in G \cap \langle u\rangle$ and $w \in \mathcal{W}_n$ sit at the n-block $B = [-k, l_n - k)$ of x containing 0. If c_j is the $\mathcal{Q}_s^{M(s)}$ class of w, then by specification Q4, $\mathrm{sh}^{-k}(x) \in \pi_s^{-1}(\langle c_j, 0\rangle) \subseteq \mathrm{sh}^{-k}\langle u\rangle$. Hence $G \cap \langle u\rangle$ is a union of shifts of sets of the form $\pi_s^{-1}(\langle c_j, 0\rangle)$. □

By the Ergodic Theorem if $n = M(s)$, then the set of x such that $x(0)$ is in the last $\epsilon_n l_n$ segment of the n-block containing $x(0)$ has measure ϵ_n. Let L be the collection of $x \in \mathbb{K}$ such that for all large s, if $n = M(s)$ then $x(0)$ is not in the last $\epsilon_n l_n$ segment of the n-block containing $x(0)$. By equation (5) and the Borel-Cantelli Lemma, L has measure one.

An argument similar to the one for Proposition 23 shows:

Proposition 24. *For all $x \neq y$ belonging to L, there is an open set $S \in \bigcup_{s\in\mathbb{N}}\mathcal{H}_s$ such that $x \in S$ and $y \notin S$.*

We note that there is a set $L_0 \subset \mathbb{K}_0$ such that $L = \pi_0^{-1}(L_0)$.

Proposition 25. *For all $s \geq 1, \mathcal{H}_s$ is a strict subalgebra of $\mathcal{H}_{s+1}$. Moreover, if $\langle \nu_x : x \in \mathbb{K}_s\rangle$ is the disintegration of ν_{s+1} over ν_s, then for ν_s-*a.e. *x, ν_x is nonatomic.*

Proof. View ν_{s+1} as a measure on $\{(x, y) : x \in \mathbb{K}_s, y \in \mathbb{K}_{s+1}$ and $\pi_{s+1,s}(y) = x\}$. Let $A = \{(x, y) : y$ is in the nonatomic part of $\nu_x\}$. Then A is invariant and hence of ν_{s+1}-measure zero or one. We claim it is measure one. If not, then we have that $A_{\alpha,\beta} = \{(x, y) : y$ is an atom of ν_x with measure between α and $\beta\}$. Then $A_{\alpha,\beta}$ is invariant and, intersecting countably many, we see that there is some $\gamma \in (0, 1)$ for almost all (x, y), y is an atom of ν_x with measure γ.

It follows that for some k, for ν_s-almost all x, there are exactly k atoms and each atom has measure $1/k$.

Thus we can assume that A is a Borel set of ν_{s+1}-measure one such that for ν_s almost all $x \in \mathbb{K}_s$, $|\{y : (x, y) \in A\}| = k$ and $\nu_x(\{y : (x, y) \in A\}) = 1$.

Choose an n such that $2^{-k(n)}k < 1/2$. We can find a basic open set O in $\mathbb{K}_s$ and $\{u_1, \ldots, u_k\} \subseteq \mathcal{W}_n$ and such that for all $x \in O$, the n-block B of x containing 0 starts at t and for ν_s measure 9/10's of the $x \in O$,

$$\{y : (x, y) \in A\} \subseteq \{x\} \times \bigcup_{i \leq k} \langle [u_i]_{s+1}, t\rangle. \tag{8}$$

For each $x \in O$, there are $2^{k(n)}$ many $\mathcal{Q}^n_{s+1}$ classes inside $[x \restriction B]_s$. All basic open sets $\langle [u]_{s+1}, t\rangle$ with $[u]_{s+1} \subseteq [x \restriction B]_s$ have the same ν_x-measure. This implies that the union on the right-hand side of equation (8) has measure less than 1/2, a contradiction. □

6. **An analysis of joinings**

In this section we give a complete analysis of the self-joinings of $F(\mathcal{T})$ and the joinings of between $F(\mathcal{T})$ and $F(\mathcal{T})^{-1}$. Theorem 7 will be an immediate consequence of this analysis.

6.1. *Branches give graph joinings.* Let $\mathbb{K}'$ be $\mathbb{K}$ or $\mathbb{K}^{-1}$, and $\mathbb{K}'_s$ the corresponding factor. Let $\langle \eta_s : s \in \mathbb{N}\rangle$ be a sequence with each η_s a joining of $\mathbb{K}_s$ with $\mathbb{K}'_s$. We will say that the sequence is *coherent* if η_s is the projection of η_{s+1} to a joining of $\mathbb{K}_s$ with $\mathbb{K}'_s$ via $\pi_{s+1,s} \times \pi_{s+1,s}$.

LEMMA 26. *Suppose that $\langle \eta_s : s \in \mathbb{N}\rangle$ is a coherent sequence of invertible graph joinings. Then there is a unique invertible graph joining η of $\mathbb{K}$ with $\mathbb{K}'$ such that for all s, η_s is the projection of η to a joining on $\mathbb{K}_s$ with $\mathbb{K}'_s$.*

Proof. For each s, let $g_s : \mathbb{K}_s \to \mathbb{K}'_s$ be the invertible measure preserving transformation from η_s. Since the g_s cohere, their inverse limit defines a measure preserving isomorphism between the sub-algebra of $\mathcal{B}(\mathbb{K})$ generated by $\bigcup_s \mathcal{H}_s$ and the subalgebra of $\mathcal{B}(\mathbb{K}')$ generated by $\bigcup_s \mathcal{H}'_s$. By Proposition 23, this extends uniquely to a measure preserving isomorphism ι between $\mathcal{B}(\mathbb{K})$ and $\mathcal{B}(\mathbb{K}')$. Proposition 24 implies that we can find a set $D \subset \mathbb{K}$ of ν measure zero such that that ι determines a shift invariant spatial isomorphism I of $\mathbb{K}\backslash D$ with $\mathbb{K}'\backslash D'$. The graph joining η determined by I projects to each η_s. □

The proof of Lemma 26 gives more information. Let L be the set of measure one defined in Proposition 24, on which $\mathcal{B}(\mathbb{K})$ separates points. Then without loss of generality we can assume that $L \supseteq \mathbb{K}\backslash D$ and $L' \supseteq \mathbb{K}'\backslash D'$. On L, the $\mathcal{Q}_s$ classes of x on the blocks containing 0 determine x and similarly for L'.

Each of the maps g_s is continuous; hence their inverse limit is a continuous shift invariant conjugacy I defined pointwise on a set of measure one.

LEMMA 27. *Let $s \in \mathbb{N}$ and $g \in G_s^n$. Let $\mathbb{K}'$ be $\mathbb{K}$ if g has even parity and $\mathbb{K}^{-1}$ otherwise. Then there is a unique invertible graph joining η_g of $\mathbb{K}_s$ with $\mathbb{K}_s'$ such that for all $m \geq n$*

$$\{(x,y) : \text{ for all } w \in \mathcal{W}_m \text{ if } x \in \langle [w]_s, 0\rangle \text{ then } y \in \langle g[w]_s, 0\rangle\} \tag{9}$$

has η_g measure one. The projection of this joining to $\mathbb{K}_0 \times \mathbb{K}_0'$ concentrates on the diagonal. Moreover, if $s' \geq s$, $g \in G_s^n$ and $g' \in G_{s'}^n$ with $\rho_{s',s}(g') = g$, then the projection of $\eta_{g'}$ to $\mathcal{H}_s \times \mathcal{H}_s'$ is η_g.

We note that η_g concentrates on a homeomorphism between $\mathbb{K}_s$ and $\mathbb{K}_s$.

Proof. Specifications A7, A8 and A9 show that the map sending the basic open interval $\langle [w]_s, 0\rangle \subseteq \mathbb{K}_s$ to $\langle g[w]_s, 0\rangle \subseteq \mathbb{K}_s'$ determines a measure preserving homeomorphism G from $\mathbb{K}_s$ to $\mathbb{K}_s'$. We set η_g to be this graph joining.

If we are given $g \in G_s^m$ and $g' \in G_{s'}^m$ with $\rho_{s,s'}(g') = g$, then specification A7 tells us that the action of g' is subordinate to the action of g. This translates to the fact that if G and G' are the graphs of the joinings η_g and $\eta_{g'}$,

$$\pi_{s',s}(G'(x)) = G(\pi_{s',s}(x)). \tag{10}$$

It follows that the projection of $\eta_{g'}$ is η_g. □

We note that Lemma 27 tells us that a $g \in G_\infty(\mathcal{T})$ yields a coherent sequence of invertible graph joinings of $\mathcal{H}_s \times \mathcal{H}_s^{\pm 1}$. By Lemma 26, such a g gives a graph joining of $\mathbb{K}$ with $\mathbb{K}'$, where $\mathbb{K}'$ is $\mathbb{K}$ if g has even parity and $\mathbb{K}'$ is $\mathbb{K}^{-1}$ otherwise. We will call this the *graph joining determined by g* and denote it as η_g.

COROLLARY 28. *Suppose that $\mathcal{T}$ has an infinite branch. Then*

$$(\Sigma^{\mathbb{Z}}, \mathcal{C}, F(\mathcal{T}), \mathrm{sh}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, F(\mathcal{T}), \mathrm{sh}^{-1}).$$

If $\mathcal{T}$ has at least two infinite branches, then there is an automorphism G between $(\Sigma^{\mathbb{Z}}, \mathcal{C}, F(\mathcal{T}), \mathrm{sh})$ that is not a power of the shift map.

Proof. If $\mathcal{T}$ has an infinite branch then, by Lemma 17, $G_\infty(\mathcal{T})$ has an element g of odd parity. The sequence $\eta_{\rho_s(g)}$ gives a coherent sequence of invertible graph joinings between $\mathbb{K}_s$ and $\mathbb{K}_s^{-1}$ and hence an invertible graph joining of $\mathbb{K}$ with $\mathbb{K}^{-1}$.

If $\mathcal{T}$ has at least two infinite branches, then there is a nonidentity $g \in G_\infty(\mathcal{T})$ with even parity. This gives a nontrivial element S of the centralizer of $F(\mathcal{T})$ that projects to the identity on $\mathbb{K}_0$. Hence S must be different from the shift map. □

This proves Claim 10, and half of what we need to see that F is a reduction.

6.2. *The joining specifications.* To finish the proof of the theorem we need to prove Claim 11, namely that joinings give branches through trees. We will need two more specifications about our words to do this. The specifications Q4 to A9 build certain symmetries into our words in a manner that allows nodes in the tree $\mathcal{T}$ to give increasingly precise information about invertible graph joinings.

The intent of the joinings specifications J10 and J11 is to give a mechanism for showing that any joining not arising from branches through our trees are independent joinings over a canonical factor.

Our approach to characterizing joinings η will be to calculate the measures of basic open sets $\langle a, b\rangle \subseteq \mathbb{K} \times \mathbb{K}'$ by determining the density of appearances of the finite sequences (a, b) in η-generic pairs $(x, y) \in \mathbb{K} \times \mathbb{K}'$. For m, where l_m is much larger than the lengths of a and b, this density can be well-approximated by counting the appearances of pairs (u', v') from $\mathcal{W}_m \times \mathcal{W}'_m$ in (x, y). In the cases we are concerned with, joining specifications will say that each pair (u', v') occurs with essentially the same frequency in (x, y), and hence (a, b) occurs with essentially the same frequency in each generic (x, y).

A complication to this outline is that the elements of the odometer associated to x and y may not be the same. If this happens, then the block structures of x and y are not synchronized. In this case a given m-block of x may overlap two m-blocks of y. The joining specifications imply that these overlaps are either insignificant or each pair (u', v') occurs overlapped essentially the same number of times with the same overlap. Again for large m this suffices to show that the occurrences of (a, b) are independent of the choice of x and y.

We will achieve the desired frequencies by randomizing the words as much as possible, subject to the restrictions imposed by Q4–A9.

Our first joining specification, J10, is aimed at the latter situation where the odometer factors are not synchronized. We consider m and n with σ_m and σ_n consecutive elements of $\mathcal{T}$. We take arbitrary words u, v in $W_n \cup \operatorname{rev}(W_n)$ and shift one of them by a multiple k of the lengths of m words. Then on the overlap both words are concatenations of strings of $(l_n/l_m) - k$ many m words. For a given pair (u', v') of appropriate m words we can count their simultaneous occurrence in the overlap. The assertion is that as long as k is nontrivial and the overlap is nontrivial, then the occurrence is close to random.

The first joining specification. Fix a $\mathcal{T} \in \mathcal{T}rees$ and let $\mathcal{W}_i$ denote $\mathcal{W}_i(\mathcal{T})$.

J10. Suppose that σ_m and σ_n are consecutive elements of $\mathcal{T}$. Let u and v be elements of $\mathcal{W}_n \cup \operatorname{rev}(\mathcal{W}_n)$. Let $1 \leq k < (1 - \epsilon_n)(l_n/l_m)$. Then for each pair $u', v' \in \mathcal{W}_m \cup \operatorname{rev}(\mathcal{W}_m)$ such that u' has the same parity as u and v' has the same parity as v, let $r(u', v')$ be the number of occurrences of (u', v') in $(\operatorname{sh}^{kl_m}(u), v)$ on their overlap. Then

$$\left|\frac{r(u',v')}{(l_n/l_m)-k} - \frac{1}{W_m^2}\right| < \epsilon_n. \tag{11}$$

In view of equation (3), we can assume that ϵ_n is much less than W_m^{-2}.

If we set $d = l_n/l_m - k$, then d is the number of $\mathcal{W}_m$ words in the overlap. This is at least $\epsilon_m(l_n/l_m)$. By equation (4), for large m and n, this is much larger than W_m^2. We can rewrite equation (11) as

$$\left|r(u',v') - \frac{d}{W_m^2}\right| < d\epsilon_n.$$

PROPOSITION 29. *Let u and v be as in specification $J10$. Then we can remove a portion of the overlap of $\mathrm{sh}^{kl_m}(u)$ and v of proportion at most $2\epsilon_n W_m^2$ consisting of pairs of words from $\mathcal{W}_m \cup \mathrm{rev}(\mathcal{W}_m)$, so that each pair (u',v') of elements from $\mathcal{W}_m \cup \mathrm{rev}(\mathcal{W}_m)$ with the correct parity occurs the same number of times in what remains. In particular, each occurs in what remains with density exactly W_m^2.*

The second joining specification. Our final specification, J11, is directed at joinings that preserve the odometer factor. It will imply that the only graph joinings preserving the odometer factor are those we built into the construction by Q4–A9.

Let σ_m and σ_n be consecutive members of $\mathcal{T}$. Suppose that $u \in \mathcal{W}_n$ and $v \in \mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$. We let $s = s(u,v)$ be the maximal i such that there is a $g \in G_i^m$ such that $g[u]_i = [v]_i$. Then there is a unique such $g \in G_s^m$ with this property, which we will call $g(u,v)$.

The action of $g(u,v)$ introduces a common pattern between u and v. Specification J11 will say that relative to this pattern all occurrences of pairs (u',v') occur randomly.

The words in $\mathcal{W}_n$ are concatenations of length l_n/l_m of words in $\mathcal{W}_m$. We can write $u = u_0u_1\cdots u_{(l_n/l_m-1)}$ with $u_i \in \mathcal{W}_m$ and $v = v_0v_1\cdots v_{(l_n/l_m-1)}$ with $v_i \in \mathcal{W}'_m$, where $\mathcal{W}'_m$ is either $\mathcal{W}_m$ or $\mathrm{rev}(\mathcal{W}_m)$ depending on the parity of g. Since $g[u]_s = v_s$, we know that $g[u_i]_s = [v_i]_s$ if g has even parity and $g[u_i]_s = [v_{l_n/l_m-i-1}]_s$ if g has odd parity.

Given a particular $\mathcal{Q}_s^m$ class $\mathcal{C}$, a portion of approximately $1/Q_s^m$ of the u_i's in u come from $\mathcal{C}$. At the same locations in v we have occurrences of v_i's from $g\mathcal{C}$. Specification J11 says that in these locations all pairs from $\mathcal{C} \times g\mathcal{C}$ occur with about the same frequency, $|\mathcal{C}|^{-2}$.

J11. Let σ_m and σ_n be successive members of $\mathcal{T}$. Suppose that $u \in \mathcal{W}_n$ and $v \in \mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$. Let $(u',v') \in \mathcal{W}_m \times \mathcal{W}'_m$ be such that $g[u']_s = [v']_s$. Let $r(u',v')$ be the cardinality of $\{i : (u_i,v_i) = (u',v')\}$. Then

$$\left|r(u',v') - \frac{l_n}{l_m}\frac{1}{Q_s^m}\left(\frac{1}{C_s^m}\right)^2\right| < \epsilon_n(l_n/l_m).$$

Since $C_s^m Q_s^m = W_m$ it follows from equation (3), that for large m, ϵ_n is much less than $\frac{1}{Q_s^m}(\frac{1}{C_s^m})^2$. Setting $l = l_n/l_m$ to be the number of m words occurring in an n word, we see that alternate form of specification J11 is

$$\left|\frac{r(u',v')}{l} - \frac{1}{Q_s^m}\left(\frac{1}{C_s^m}\right)^2\right| < \epsilon_n.$$

This has the corollary analogous to Proposition 29:

Proposition 30. *Let (u,v) be as in the hypothesis of* J11. *Then we can delete a portion of (u,v) made up of some (u_i,v_i)'s of proportion at most $2\epsilon_n Q_s^m (C_s^m)^2$ so that in what remains each $(u',v') \in \mathcal{W}_m \times \mathcal{W}'_m$ with $g[u']_s = [v']_s$ occurs the same number of times. In particular, each such (u',v') occurs with proportion exactly $1/Q_s^m(C_s^m)^2$.*

The next consequence of specification J11 is a strengthening of Proposition 30. It says that for each $m < n$ after we delete a relatively small number of pairs in $\mathcal{W}_m \times \mathcal{W}'_m$ that occur in a pair $(u,v) \in \mathcal{W}_n \times \mathcal{W}'_n$, every permissible pair from $\mathcal{W}_m \times \mathcal{W}'_m$ occurs the same number of times in what is left.

Proposition 31. *Suppose that $\mathcal{T}$ is $\langle \sigma_{n_k} : k \in \mathbb{N}\rangle$. Let $n = n_l$. Suppose that $u \in \mathcal{W}_n$ and $v \in \mathcal{W}'_n$, where $\mathcal{W}'_n = \mathcal{W}_n$ or $\mathcal{W}'_n = \mathrm{rev}(\mathcal{W}_n)$. Let $s = s(u,v)$ and $g = g(u,v)$. Suppose that $g \in G_s^{n_j}$, where $n_j < n$. Let $k \in [j,l)$ and $m = n_k$. Then we can remove pairs $(u',v') \in \mathcal{W}_m \times \mathcal{W}'_m$ occurring in (u,v) leaving a subsequence of such pairs of proportion at least $(1-\epsilon_{n_{l-1}})\prod_{k\le i<l-1}(1-\epsilon_{n_i})^2$ in which every pair $(u',v') \in \mathcal{W}_m \times \mathcal{W}'_m$ with $g[u']_s = [v']_s$ occurs the same number of times.*

Proof. We go by induction on l for a fixed k. Suppose that $l=k+1$. In this case σ_m and σ_n are consecutive elements of $\mathcal{T}$. It follows from Proposition 30 that we can remove $2\epsilon_n Q_s^m (C_s^m)^2$ portion from (u,v) and have each appropriate (u',v') occur the same number of times. Since $Q_s^m(C_s^m)^2 < W_m^2 < W_n^2$, equation (3) implies that we have removed a portion of size less than $\epsilon_m = \epsilon_{n_{l-1}}$.

Suppose now that it is true for l. We show it for $n = n_{l+1}$. We perform two deletion operations and apply induction. Let $(u,v) \in \mathcal{W}_n \times \mathcal{W}'_n$. By Proposition 30 and equation (3) we can remove a portion of (u,v) consisting of pairs of words from $\mathcal{W}_{n_l} \times \mathcal{W}'_{n_l}$ so that:

1. the portion of (u,v) remaining has proportion $(1-\epsilon_{n_l})$ and
2. every pair $(u',v') \in \mathcal{W}_{n_l} \times \mathcal{W}'_{n_l}$ with $g[u']_s = [v']_s$ occurs the same number of times.

By Proposition 21, we can remove a portion of the remainder that consists of pairs of $(u^\dagger,v^\dagger) \in \mathcal{W}_{n_l} \times \mathcal{W}'_{n_l}$ of proportion at most $|G_{s+1}^{n_{l-1}}|2^{-k(n_l)}$ so that for the remaining pairs $(u^\dagger,v^\dagger)$ there is no $h \in G_{s+1}^{n_{l-1}}$ with $h[u^\dagger]_{s+1} = [v^\dagger]_{s+1}$.

Since $|G^{n_{l-1}}_{s+1}| \leq 2^{n_{l-1}}$ we can use equation (6) to see that what is removed in the second deletion operation has proportion less than $\epsilon_{n_{l-1}}$.

After these two deletions we are left with at least $(1-\epsilon_{n_{l-1}})(1-\epsilon_{n_l})$ proportion of (u,v) and for each remaining pair $(u^\dagger, v^\dagger) \in \mathcal{W}_{n_l} \times \mathcal{W}'_{n_l}$ we have $s = s(u^\dagger, v^\dagger)$ and $g = g(u^\dagger, v^\dagger)$. Hence we can apply induction to delete portions of each such $(u^\dagger, v^\dagger)$ keeping at least proportion $(1-\epsilon_{n_{l-1}})\prod_{k\leq i<l-1}(1-\epsilon_{n_i})^2$ so that every pair $(u',v') \in \mathcal{W}_m \times \mathcal{W}'_m$ occurs the same number of times in the remaining portion of $(u^\dagger, v^\dagger)$.

What is left after these deletions is a portion of the original pair (u,v) having proportion at least $(1-\epsilon_{n_l})\prod_{k\leq i<l}(1-\epsilon_{n_i})^2$ and in which each appropriate pair $(u',v') \in \mathcal{W}_m \times \mathcal{W}'_m$ occurs the same number of times. □

Again we remark that if each appropriate pair (u',v') appears the same number of times in the remainder after the deletions, each appropriate pair occurs with proportion $\frac{1}{Q^m_s}\left(\frac{1}{C^m_s}\right)^2$.

6.3. *The Kronecker factor.* For the next few sections we will be considering one tree $\mathcal{T}$ at a time. For notational simplicity, if $\{n_j : j \in \mathbb{N}\}$ is an enumeration of $\{n : \sigma_n \in \mathcal{T}\}$ in increasing order, then we will write "σ_j" instead of "σ_{n_j}", "$\mathcal{W}_j$" instead of "$\mathcal{W}_{n_j}$", "l_j" instead of "l_{n_j}" and "$M(i) = j$" instead of "$M(\mathcal{T})(i) = \sigma_{n_j}$".

The new notation has the effect that σ_j and σ_{j+1} denote consecutive elements of the tree $\mathcal{T}$. Moreover the odometer factor of $F(\mathcal{T})$ is now written $\prod_j \mathbb{Z}_{l_{j+1}/l_j}$.

Suppose that η is an ergodic joining of $\mathbb{K}$ with $\mathbb{K}'$ for $\mathbb{K}' = \mathbb{K}^{\pm 1}$. Since $\mathcal{O}_{\mathcal{T}} \times \mathcal{O}_{\mathcal{T}}$ is a factor of $\mathbb{K} \times \mathbb{K}'$, η projects to a joining of the odometer factor with itself. By Theorem 13, there is a $g \in \mathcal{O}_{\mathcal{T}}$ such that the projection of η to $\mathcal{O}_{\mathcal{T}} \times \mathcal{O}_{\mathcal{T}}$ is given by the graph joining determined by g. We can describe g explicitly.

For each $\vec{a} \in \prod_{j<J} \mathbb{Z}_{l_{j+1}/l_j}$, the collection of pairs $(x,y) \in \mathbb{K} \times \mathbb{K}'$ such that the element of $\prod_j \mathbb{Z}_{l_{j+1}/l_j}$ given by $\pi_0(x) - \pi_0(y)$ starts with $\vec{a}$ either has η measure zero or measure one and for each J there is a unique $\vec{a}_J$ which gives a set of measure one. If $g \in \prod_j \mathbb{Z}_{l_{j+1}/l_j}$ is such that for all J, $g \restriction J = \vec{a}_J$, then the projection of η to $\mathcal{O}_{\mathcal{T}} \times \mathcal{O}_{\mathcal{T}}$ is the graph joining given by g.

Proposition 32. *Suppose that $\mathcal{T} \in \mathcal{T}rees$ and η is an ergodic joining between $\mathbb{X} = (\Sigma^{\mathbb{Z}}, \mathcal{C}, F(\mathcal{T}), \mathrm{sh})$ and $\mathbb{Y} = (\Sigma^{\mathbb{Z}}, \mathcal{C}^{\pm 1}, F(\mathcal{T})^{\pm 1}, \mathrm{sh})$. Suppose that $\eta \restriction \mathcal{H}_0 \times \mathcal{H}_0$ is supported on the graph of some $g \neq \bar{j}$ for any $j \in \mathbb{Z}$. Then η must be the relatively independent joining of $\mathbb{X}$ and $\mathbb{Y}$ over the graph joining given by g.*

Remark As the set of $(x,y) \in \mathbb{K} \times \mathbb{K}'$ which projects to a pair of elements of $\mathcal{O}_{\mathcal{T}}$ whose difference is g is a closed set, the proposition proves the unique ergodicity of sh $\times$ sh restricted to this closed set. In particular, this implies that there is a unique ergodic joining η that lifts the joining on the graph of g and η must be the relatively independent joining over g.

Proof of Proposition 32. Let $\mathbb{K}'$ be either $\mathbb{K}$ or $\mathbb{K}^{-1}$ depending on whether $\mathbb{Y}$ comes from $F(\mathcal{T})$ or from $F(\mathcal{T})^{-1}$ and similarly we let $\mathcal{W}'_j$ be either $\mathcal{W}_j$ or $\mathrm{rev}(\mathcal{W}_j)$.

Suppose $g \in O_{\mathcal{T}}$ with $g \neq \bar{j}$ for any $j \in \mathbb{Z}$. Let s and t be arbitrary elements of $O_{\mathcal{T}}$ such that $s - t = g$. Let x and y be arbitrary elements of $\mathbb{K}$ and $\mathbb{K}'$ such that $\pi_0(x) = s$ and $\pi_0(y) = t$.

Claim. If $\langle a \rangle \times \langle b \rangle$ is a cylinder set in $\mathbb{K} \times \mathbb{K}'$, then the density of the occurrences of (a, b) in (x, y) does not depend on the choice of x, y, s or t.

To see that the claim suffices for Proposition 32, we note that if η is an ergodic joining then $\eta(\langle a \rangle \times \langle b \rangle)$ is given by the frequency of occurrences of (a, b) in a generic (x, y). Since this frequency is independent of the choice of (x, y) there is a unique such η, which must be the relatively independent joining.

Proof of claim. For each j, s partitions $\mathbb{Z}$ into "j-blocks" of length l_j corresponding to where the words from $\mathcal{W}_j$ lie in x and similarly t gives a partition of $\mathbb{Z}$ corresponding to the occurrence of words in $\mathcal{W}_j$ or $\mathrm{rev}(\mathcal{W}_j)$ in y.

Each j-block of x overlaps two j-blocks of y that have lengths g_j and $l_j - g_j$ respectively for some numbers g_j. We will call these the "first overlaps" and the "second overlaps" respectively. Note that for each j the lengths of the overlaps of all j-blocks are the same and this length is determined by g. It is thus independent of x, y, s and t.

If $g \in \prod_{j \in \mathbb{N}} \mathbb{Z}_{l_{j+1}/l_j} = O_{\mathcal{T}}$ is given by some sequence $\langle k_j : j \in \mathbb{N} \rangle$, then we can compute the values g_j recursively by setting $g_0 = k_0$ and $g_j = g_{j-1} + k_j l_{j-1}$.

Since $g \neq \bar{q}$ for any $q \in \mathbb{Z}$, there is an infinite set J of values j where $l_j/l_{j-1} - 1 > k_j > 0$. Moreover, the lengths of the overlaps grow to infinity as j goes to infinity.

Our strategy is to focus on an arbitrary j-block of x and count instances of the occurrence of (a, b) in the corresponding section of (x, y). As j gets large this will be increasingly independent of the choices of x, y, s, t and the choice of j-block.

This suffices since occurrences of (a, b) in (x, y) are either inside a j-block or overlap two j-blocks of x. Let k be a natural number bigger than the lengths of a and b. As j goes to infinity, the portion of the j-block that is covered by an overlapping occurrence of (a, b) is bounded by $2k/l_j$, which goes to zero as j gets large. Hence the density of occurrences (a, b) in (x, y) is given by averaging the densities of occurrences of (a, b) along j-blocks and letting j go to infinity.

We start by choosing a j with $k_j \neq 0$ that is so large that l_{j-1} is very large relative to the lengths of a and b, i.e. k/l_{j-1} is very small and, moreover, l_{j-1}/l_j is very small. We also assume that $\epsilon_j(l_j/l_{j-1})$ is much larger than W_{j-1}^2 and that $\epsilon_j W_{j-1}^2$ is very small.

We show that for each pair of $j-1$-words $(u', v') \in \mathcal{W}_{j-1} \times \mathcal{W}'_{j-1}$, the number of times where u' and v' occur overlapped as first or second overlaps of $j-1$-words in this arbitrary j-block is approximately $\frac{1}{|\mathcal{W}_{j-1}|^2}$, and hence nearly independent of the choices of x, y, s and t. Since the number and extent of the overlaps of pairs of $j-1$-words closely determines the density of occurrences of (a, b), this will suffice.

Let u be the $\mathcal{W}_j$ word occurring in x at an arbitrary j-block B, and v_l and v_r be the $\mathcal{W}'_j$ words occurring in y as the first and second overlaps. Each occurrence in B of a $u' \in \mathcal{W}_{j-1}$ overlaps either one or two words in $\mathcal{W}'_{j-1}$.

An occurrence of (a, b) along B can occur in the following ways:

1. completely contained in both a $j-1$ word in u and a $j-1$ word in v_l or v_r,
2. properly overlapping two $j-1$ words occurring in u,
3. properly overlapping two $j-1$ words occurring in v_l or v_r,
4. in the first or last l_{j-1} letters of u or
5. in the last l_{j-1} letters of v_l or the first l_{j-1} letters of v_r.

The section of B in which (a, b) can occur in ways 2–5 has size bounded by

$$2(k/l_{j-1})(l_j/l_{j-1}) + 2(k/l_{j-1})(l_j/l_{j-1}) + 2l_{j-1} + 2l_{j-1} \tag{12}$$

which has proportion of B at most:

$$4k/l_{j-1}^2 + 4(l_{j-1}/l_j). \tag{13}$$

An occurrence of (a, b) of type one can occur in two ways:

L) in a portion of B corresponding the first overlap,
R) in a portion of B corresponding to the second overlap.

If either the first overlap or the second overlap has size less than $\epsilon_j l_j + l_{j-1}$, then we neglect it, as it contributes a proportion of size

$$\epsilon_j + l_{j-1}/l_j \tag{14}$$

of the block B. Otherwise we can treat the two overlaps separately.

Consider the overlap of u with v_r. This section B_r of x has various $j-1$-blocks that may or may not be synchronous with $j-1$-blocks of y. If they are not synchronous; then the $j-1$-blocks of u also have first and second overlaps. We deal with the case that they are not synchronous; the synchronous case is similar and easier.

By shifting x right by g_{j-1} or left by $l_{j-1} - g_{j-1}$, the $j-1$-blocks of x line up with the $j-1$-blocks of y. Depending on whether we shift right or left, the first or last $j-1$-blocks on B_r are lost.

Either way we shift to make the $j-1$ words line up, we can apply Proposition 29 to find a portion of the shifted overlap of proportion at least $1 - 2\epsilon_j W_{j-1}^2 > 1 - \epsilon_{j-1}$ so that every pair $(u', v') \in \mathcal{W}_{j-1} \times \mathcal{W}'_{j-1}$ occurs with density W_{j-1}^{-2}. Taking into account the blocks that might be lost, we can find a portion of B_r of proportion at least $1 - \epsilon_{j-1} - 2l_{j-1}/|B_r|$, where each pair $(u', v') \in \mathcal{W}_j \times \mathcal{W}'_j$ occurs in both first overlaps and second overlaps with proportion exactly W_{j-1}^{-2}.

We now do the same thing on the section B_l where B overlaps with v_l. The result of deleting these portions of B_r and B_l is a section of B of proportion at least

$$1 - \epsilon_{j-1} - 4(l_{j-1}/l_j) \tag{15}$$

on which each pair $(u', v') \in \mathcal{W}_j \times \mathcal{W}'_j$ occurs in both first overlaps and second overlaps with proportion exactly W_{j-1}^{-2}.

Thus neglecting occurrences of (a, b) in portions of B corresponding to equations (13), (14) and (15) we are neglecting a portion of B of density at most

$$\varepsilon = 4k/l_{j-1}^2 + 9(l_{j-1}/l_j) + \epsilon_j + \epsilon_{j-1}.$$

The remaining "typical" occurrences reside in the interior of overlaps of pairs of words $(u', v') \in \mathcal{W}_{j-1} \times \mathcal{W}'_{j-1}$ as either right or left overlaps in the portion of B not discarded in equation (15). The density of these typical occurrences is determined by the amount of shift and the density of the occurrences of the pairs (u', v'). The amount of shift is given by g as a right shift of g_{j-1} and a left shift of $l_{j-1} - g_{j-1}$. The density of occurrences of each (u', v') is W_{j-1}^{-2}. Since these numbers are independent of x, y, s and t we see that up to an error of ε, the density of occurrences of (a, b) in j-blocks is also independent of x, y, s and t. Since ε goes to 0 as j goes to ∞, the density of occurrences of (a, b) is independent of the choice of x, y, s and t. □

Corollary 33. *$\mathcal{O}_\mathcal{T}$ is the Kronecker factor of $\mathbb{K}$.*

Proof. If not, then $\mathcal{O}_\mathcal{T}$ is a proper factor of the Kronecker factor $\mathcal{K}$. Let $t \in \mathcal{K}$ be such that the factor transformation on the group $\mathcal{K}$ is given by translation by t. If $\pi'_0 : \mathcal{K} \to \mathcal{O}_\mathcal{T}$ is the factor map, then $\pi'_0(t) = \bar{1}$.

Choose an element $t' \in \mathcal{K}$ such that $\pi'_0(t') = g \neq \bar{1}$. Then t' determines a graph joining of $\mathcal{K}$ with $\mathcal{K}$ that can be lifted to the relatively independent joining η' of $\mathbb{K}$ with $\mathbb{K}$. This joining concentrates on $\{(x, y) \in \mathbb{K} \times \mathbb{K}' : \pi_0(x) - \pi_0(y) = g\}$. Since $\mathrm{sh} \times \mathrm{sh}$ is uniquely ergodic on this set, η' must be the relatively independent joining over the graph joining of $\mathcal{O}_{\mathcal{T}}$ with $\mathcal{O}_{\mathcal{T}}$ by g. But this implies that $\mathcal{K}$ is a trivial extension of $\mathcal{O}_{\mathcal{T}}$. $\square$

If $\mathcal{T} \neq \mathcal{T}'$, then $\{p_n : \text{a } p_n^{\text{th}} \text{ root of unity is an eigenvalue of the unitary operator associated with } F(\mathcal{T})\}$ is different from $\{p_n : \text{a } p_n^{\text{th}} \text{ root of unity is an eigenvalue of the unitary operator associated with } F(\mathcal{T}')\}$. In particular:

Corollary 34. *If $\mathcal{T} \neq \mathcal{T}'$, then $F(\mathcal{T})$ is not conjugate to $F(\mathcal{T}')^{\pm 1}$.*

Remark. With slightly more effort, we could reorganize our argument so that the odometer factor of each $F(\mathcal{T})$ was the same and arrange that for $\mathcal{T} \neq \mathcal{T}'$ any joining between $F(\mathcal{T})$ and $F(\mathcal{T}')^{\pm 1}$ is the relatively independent joining over this common Kronecker factor.

6.4. *Arbitrary joinings.* We begin by separating the various joinings.
For $g \in G_s$, let

$$\begin{aligned} I_g &= \{(x,y) : g\pi_s(x) = \pi'_s(y)\} \subseteq \mathbb{K} \times \mathbb{K}', \\ T_g &= I_g \cap \{(x,y) : \text{for no } g^* \in G_{s+1} \text{ is } g^*\pi_{s+1}(x) = \pi'_{s+1}(y)\}. \end{aligned}$$

The next two lemmas are routine to verify.

Lemma 35. *Let $s, t \in \mathbb{N}$ and $g \in G_s, h \in G_t$ be different. Then:*

1. *Each I_g is a closed set, as is $\bigcup_{g \in G_s} I_g$,*
2. *T_g is a $\mathcal{G}_\delta$ set,*
3. *$T_g \cap T_h = \emptyset$,*
4. *If η is the relatively independent joining of $\mathbb{K}$ with $\mathbb{K}'$ over the graph joining η_g of $\mathbb{K}_s$ with $\mathbb{K}'_s$, then $\eta(T_g) = 1$,*
5. *If $g_\infty \in G_\infty(\mathcal{T})$, η is the graph joining of $\mathbb{K}$ with $\mathbb{K}'$ determined by g_∞ and $g_s = \rho_s(g_\infty)$, then $\bigcap_{s \in \mathbb{N}} I_{g_s}$ is measure one for η.*

Lemma 36. *Let $\mathcal{I}$ be the collection of joinings of $\mathbb{K}$ with $\mathbb{K}'$ of the following form:*

1. *the relatively independent joining of $\mathbb{K}$ with $\mathbb{K}'$ over the graph joining η_g of $\mathbb{K}_s$ with $\mathbb{K}'_s$ for some $g \in G_s$,*
2. *the graph joining of $\mathbb{K}$ with $\mathbb{K}'$ determined by some $g \in G_\infty(\mathcal{T})$.*

Then any two distinct members of $\mathcal{I}$ are mutually singular measures.

Given a measure preserving transformation $\phi : \mathbb{K}' \to \mathbb{K}'$ that commutes with the shift map and a joining η of $\mathbb{K} \times \mathbb{K}'$, we get another joining we will denote $\eta \circ (1, \phi)$ defined by setting

$$\eta \circ (1,\phi)(X) = \eta(\{(k,k') : (k, \phi(k')) \in X\}).$$

If η is ergodic, then so is $\eta \circ (1, \phi)$.

We can now finish verifying that F is a reduction by proving Claim 11. The following proposition suffices.

Proposition 37. *Let η be an ergodic joining of $\mathbb{K}$ with $\mathbb{K}'$. Then exactly one of the following holds*:

1. *For some s, some $j \in \mathbb{Z}$ and some $g \in G_s$, η is the relatively independent joining of $\mathbb{K}$ with $\mathbb{K}'$ over the graph joining $\eta_g \circ (1, \mathrm{sh}^{-j})$ of $\mathbb{K}_s \times \mathbb{K}'_s$.*
2. *There is a $g \in G_\infty(\mathcal{T})$ and a $j \in \mathbb{Z}$ such that $\eta \circ (1, \mathrm{sh}^{-j})$ is the graph joining of $\mathbb{K}$ with $\mathbb{K}'$ determined by g.*

Proof. Fix such an ergodic η. By Theorem 13 it induces a graph joining on $\mathcal{O}_\mathcal{T}$. By Proposition 32, if η is not a relatively independent joining of $\mathbb{K} \times \mathbb{K}'$ over a graph joining of $\mathcal{O}_\mathcal{T} \times \mathcal{O}_\mathcal{T}$, then there is a $j \in \mathbb{Z}$ such that the projection of η to a joining of $\mathcal{O}_\mathcal{T} \times \mathcal{O}_\mathcal{T}$ is the graph joining given by some $\bar{j}$. In particular if we set $\eta' = \eta \circ (1, \mathrm{sh}^{-j})$ we get another ergodic joining of $\mathbb{K} \times \mathbb{K}'$, where the induced joining on $\mathcal{O}_\mathcal{T} \times \mathcal{O}_\mathcal{T}$ is the graph joining coming from the identity map. Moreover, if η is not a relatively independent joining over $\mathcal{H}_s \times \mathcal{H}'_s$, then η' is not.

Hence without loss of generality we can assume that the projection of η to a joining of $\mathcal{O}_\mathcal{T} \times \mathcal{O}_\mathcal{T}$ is the graph joining induced by the identity.

Let (x, y) be generic for η. Assume that we are not in case 2. Then there is a maximal s such that for some $g \in G_s$ such that $g\pi_s(x) = \pi_s(y)$. This g is unique and η concentrates on the set T_g. We show that η is the relatively independent joining over η_g.

For each j let $B_j = [a_j, b_j]$ be the interval of integers on which the j-block of x containing 0 lies.

Claim. There is an infinite set J of j such that for all $g' \in G^j_{s+1}$

$$g'(\pi_{s+1}(x) \restriction B_{j+1}) \neq \pi_{s+1}(y) \restriction B_{j+1}. \tag{16}$$

To see the claim suppose that $g \in G^{i_0}_s$. For each $g' \in G^i_{s+1} \backslash G^{i-1}_{s+1}$ with $i \geq i_0$, if there is a j such that $g'(\pi_{s+1}(x) \restriction B_j) = \pi_{s+1}(y) \restriction B_j$, then we have $\rho_{s+1,s}(g') = g$ and the collection of such j is a finite interval of the form $[i, j_{max}]$. Moreover, if $g'_1 \neq g'_2$ then the corresponding intervals are disjoint. We must show that there are infinitely many j not appearing in any such interval. Fix such a g' and let $j = j_{max}$. Then for all $g^* \in G^j_{s+1}$ different than g',

$$g^*(\pi_{s+1}(x) \restriction B_j) \neq g'(\pi_{s+1}(x) \restriction B_j) = \pi_{s+1}(y) \restriction B_j.$$

Hence, $g^*(\pi_{s+1}(x) \restriction B_{j+1}) \neq \pi_{s+1}(y) \restriction B_{j+1}$. Since $g'(\pi_{s+1}(x) \restriction B_{j+1}) \neq \pi_{s+1}(y) \restriction B_{j+1}$, we see that $j \in J$. This proves equation (16).

Continuing the proof of Proposition 37, we follow the same strategy as in Proposition 32: we show that all generic pairs (x, y) for any ergodic η

concentrating on T_g have the same frequencies of occurrence of finite sequences (a, b). Hence the shift map on T_g is uniquely ergodic and thus any joining is the relatively independent joining.[16]

Fix (x, y) and (x', y') that are generic for η and η' concentrating on T_g, and let J and J' be the sets corresponding to equation (16) for each pair and B_j and $B'_{j'}$ the corresponding blocks in x and x'.

Fix a pair (a, b) of finite sequences from Σ. Without loss of generality, we can assume that (a, b) occurs somewhere in $\mathbb{K} \times \mathbb{K}'$ (in particular they have the right relative parity) and that they have length bounded by some k. We show that the density of occurrences of (a, b) is the same in (x, y) as it is in (x', y'). Since each pair is generic for its respective measure it suffices to show:

Claim. For each $\delta > 0$ for arbitrarily large $j \in J$ and $j' \in J'$ the density of occurrences of (a, b) in $(x, y) \upharpoonright B_j$ is within δ of the density of occurrences of (a, b) in $(x', y') \upharpoonright B'_{j'}$.

Fix such a δ and let $j < j'$ be elements of J and J' so large that

$$\prod_{j-1 \le q \le j'} (1 - \epsilon_q)^2 - 2k/l_{j-1} > 1 - \delta. \tag{17}$$

By Proposition 31 we can delete $j-1$-blocks in $(x, y) \upharpoonright B_j$ and $(x', y') \upharpoonright B'_{j'}$ to get D and D' so that a proportion of at least $\prod_{j-1 \le q \le j'} (1 - \epsilon_q)^2$ remains in each and every pair $(u', v') \in \mathcal{W}_{j-1} \times \mathcal{W}'_{j-1}$ with $g[u']_s = [v']_s$ occurs with proportion $\frac{1}{Q_s^m} \left(\frac{1}{C_s^m}\right)^2$.

In the remaining sections (a, b) can occur

a) overlapping two $j-1$-blocks,
b) in the interior of $j-1$-blocks.

The proportion of a j or j'-block that is within k of the end or beginning of a $j-1$-block is less than $2k/l_{j-1}$. We delete these sections from D and D' and are left with remainders with proportion at least $1 - \delta$ of the original blocks in which all occurrences of (a, b) occur in the interiors of appropriate (u', v'). Since each appropriate (u', v') occurs with the same proportion in D and D', we see that density of occurrences of (a, b) is the same in the two remainders. Hence the density in $(x, y) \upharpoonright B_j$ and $(x', y') \upharpoonright B'_{j'}$ are within δ. $\square$

COROLLARY 38. *Suppose that $\mathbb{K}'$ is $\mathbb{K}^{\pm 1}$. Suppose that ϕ is an isomorphism between $\mathbb{K}$ and $\mathbb{K}'$. Then there is a $j \in \mathbb{Z}$ such that the joining associated with ϕ is a graph joining of the form $\eta_g \circ (1, \mathrm{sh}^j)$ for some $g \in G_\infty(\mathcal{T})$. Moreover, if g has odd parity, then $\mathbb{K}' = \mathbb{K}^{-1}$ and if g has even parity, then $\mathbb{K}' = \mathbb{K}$.*

[16] We remind the reader that for a generic pair (x, y) for η the block structure of x exactly synchronized with the block structure of y.

While the following is not strictly necessary for our theorem it seems worth noting.

PROPOSITION 39. *Each relatively independent joining of $\mathbb{K}$ with $\mathbb{K}'$ over a graph joining of $\mathbb{K}_s$ with $\mathbb{K}'_s$ determined by $g \in G_s$ is ergodic. In particular, the collection of joinings of $\mathbb{K}$ with $\mathbb{K}'$ is the closed convex hull of*

1. *the relatively independent joinings of $\mathbb{K}$ with $\mathbb{K}'$ over some graph joining η_g of $\mathbb{K}_s$ with $\mathbb{K}'_s$ for some $g \in G_s$,*
2. *the graph joinings of $\mathbb{K}$ with $\mathbb{K}'$ determined by some $g \in G_\infty(\mathcal{T})$.*

Lemma 17, together with Corollary 38 finish the proof of Claim 11 and thus Theorem 7. For if there is a conjugacy between $F(\mathcal{T})$ and $F(\mathcal{T})^{-1}$ it must correspond to some $g \in G_\infty(\mathcal{T})$, in particular $G_\infty(\mathcal{T})$ is not the trivial group. By Lemma 17, we see $\mathcal{T}$ has an infinite branch.

If there is a nontrivial element of the centralizer of $F(\mathcal{T})$, then there must be a graph joining of $F(\mathcal{T})$ with $F(\mathcal{T})$ that is not of the form $(1, \mathrm{sh}^j)$ for any $j \in \mathbb{Z}$. By Corollary 38, there must be a nontrivial element $G_\infty(\mathcal{T})$ with even parity. Again by Lemma 17, this implies that $\mathcal{T}$ has at least two infinite branches.

Finally, from Proposition 39 we get more information:

COROLLARY 40. *Suppose that ϕ is an invertible measure preserving transformation and $\phi F(\mathcal{T})\phi^{-1} = F(\mathcal{T})^{\pm 1}$. Then there is a $j \in \mathbb{Z}$ and a $g \in G_\infty(\mathcal{T})$ such that $\eta_g \circ (1, \mathrm{sh}^j)$ is supported on the graph of ϕ. If $\phi F(\mathcal{T})\phi^{-1} = F(\mathcal{T})^{-1}$, then g has odd parity; otherwise g has even parity.*

7. A recapitulation of the specifications

For the reader's convenience we now collect the various specifications about the words and numerical assumptions we made in order to prove Theorem 7. Collecting these here will make it easier to verify that the construction given in the next section works. We also take the chance to simplify the specifications a bit by making them stronger.

As earlier in the text we will write $\mathcal{W}_n$ for $\mathcal{W}_n(\mathcal{T})$ with the understanding that $\mathcal{W}_n$ is determined by $\mathcal{T} \cap \{\sigma_m : m \leq n\}$.

7.1. *The numerical parameters.* We have various numerical parameters:

1. W_m the number of elements of $\mathcal{W}_m$. Q^m_s and C^m_s, the number of classes and sizes of each class of $\mathcal{Q}^m_s$ respectively.
2. $2^{k(n)}$ the number of $\mathcal{Q}^m_{s+1}$ classes inside each $\mathcal{Q}^m_s$ class. The numbers $k(n)$ will be chosen to grow fast enough that

$$2^m 2^{-k(n)} < \epsilon_m. \tag{18}$$

If s is the maximal length of an element of $\mathcal{T} \cap \{\sigma_m : m \leq n\}$, then we set $C^n_s = 2^{k(n)}$ as well. This makes W_m, Q^m_s and C^m_s be powers of 2.

3. Numbers ϵ_n and l_n. The first goes to zero rapidly and the last is an integer going to infinity rapidly. They will satisfy the following numerical relations:
   (a) If $\mathcal{T} = \langle \sigma_{n_i} : i \in \mathbb{N} \rangle$, then

$$2\epsilon_{n_i} W^2_{n_i} < \epsilon_{n_{i-1}} \tag{19}$$

   and

$$\epsilon_{n_i}(l_{n_i}/l_{n_{i-1}}) W^{-2}_{n_{i-1}} \to \infty \text{ as } n \to \infty. \tag{20}$$

   (b)

$$\prod_{n \in \mathbb{N}} (1 - \epsilon_n) > 0. \tag{21}$$

   (c) There will be prime numbers p_{n_i} such that $l_{n_i} = p^2_{n_i} W_{n_{i-1}} l_{n_{i-1}}$. The p_n's will satisfy some fast growth conditions specified during the word construction itself.

7.2. *The specifications*. The words in $\mathcal{W}_n$ are sequences of elements of $\Sigma = \{0, 1\}$. We construct $\mathcal{W}_n$ just in case $\sigma_n \in \mathcal{T}$. Which words are in $\mathcal{W}_n$ depends only on $\mathcal{T} \cap \{\sigma_m : m \leq n\}$. $\mathcal{W}_0 = \{0, 1\}$ and $\mathcal{Q}^0_0$ is the trivial equivalence relation with one class.

E1. Any pair w_1, w_2 of words in $\mathcal{W}_n$ have the same length l_n.

E2. If σ_m and σ_n are consecutive elements of $\mathcal{T}$, then every word in $\mathcal{W}_n$ is built by concatenating words in $\mathcal{W}_m$. Every word in $\mathcal{W}_m$ occurs in each word of $\mathcal{W}_n$ exactly p^2_n times.

E3. (Unique readability). If σ_m and σ_n are consecutive elements of $\mathcal{T}$ and $w \in \mathcal{W}_n$ and

$$w = bw_1 \cdots w_k e$$

where b or e have length less than l_m and each $w_i \in \mathcal{W}_m$, then both b and e are the empty word. If $w, w' \in \mathcal{W}_n$ and $w = w_1 w_2 \cdots w_{l_n/l_m}$ and $w' = w'_1 w'_2 \cdots w'_{l_n/l_m}$ with $w_i, w'_i \in \mathcal{W}_m$, and $k = [l_n/2l_m] + 1$, then $w_k w_{k+1} \cdots w_{l_n/l_m} \neq w'_1 w'_2 \cdots w_{l_n/l_m - [k] - 1}$; i.e., the first half of w' is not equal to the second half of w.

The equivalence relations $\mathcal{Q}^n_s$ are defined for all $s \leq s(n)$.

Q4. Suppose that $n = M(s)$. Then any two words in the same $\mathcal{Q}^n_s$ equivalence class agree on an initial segment of length at least $(1 - \epsilon_n) l_n$.

Q5. For $n \geq M(s) + 1$, $\mathcal{Q}^n_s$ is the product equivalence relation of $\mathcal{Q}^{M(s)}_s$. Hence we can view $\mathcal{W}_n / \mathcal{Q}^n_s$ as sequences of elements of $\mathcal{W}_{M(s)} / \mathcal{Q}^{M(s)}_s$ and similarly for $\text{rev}(\mathcal{W}_n) / \mathcal{Q}^n_s$.

Q6. $\mathcal{Q}^n_{s+1}$ refines $\mathcal{Q}^n_s$ and each $\mathcal{Q}^n_s$ class contains $2^{k(n)}$ many $\mathcal{Q}^n_{s+1}$ classes.

A7. G_s^n acts freely on $\mathcal{W}_n/\mathcal{Q}_s^n \cup \mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$ and the G_s^n action is subordinate to the G_{s-1}^n action via the natural homomorphism $\rho_{s,s-1}$ from G_s^n to G_{s-1}^n.

A8. The canonical generators of $G_s^{M(s)}$ send elements of $\mathcal{W}_{M(s)}/\mathcal{Q}_s^{M(s)}$ to elements of $\mathrm{rev}(\mathcal{W}_{M(s)})/\mathcal{Q}_s^{M(s)}$ and vice versa.

A9. If $M(s) < n$, σ_m and σ_n are consecutive elements of $\mathcal{T}$ and we view $G_s^n = G_s^m \oplus H$, then the action of G_s^m on $\mathcal{W}_m/\mathcal{Q}_s^m \cup \mathrm{rev}(\mathcal{W}_m/\mathcal{Q}_s^m)$ is extended to an action on $\mathcal{W}_n/\mathcal{Q}_s^n \cup \mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$ by the skew diagonal action. If H is nontrivial, then its canonical generator maps $\mathcal{W}_n/\mathcal{Q}_s^n$ to $\mathrm{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$.

For specifications J10 and J11, σ_m and σ_n are consecutive elements of $\mathcal{T}$.

J10. Let u and v be elements of $\mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$. Let $1 \leq k < (1-\epsilon_n)(l_n/l_m)$. Then for each pair $u', v' \in \mathcal{W}_m \cup \mathrm{rev}(\mathcal{W}_m)$ such that u' has the same parity as u and v' has the same parity as v, let $r(u', v')$ be the number of occurrences of (u', v') in $(\mathrm{sh}^{kl_m}(u), v)$ on their overlap. Then

$$\left| \frac{r(u',v')}{(l_n/l_m) - k} - \frac{1}{W_m^2} \right| < \epsilon_n. \tag{22}$$

J11. Suppose that $u \in \mathcal{W}_n$ and $v \in \mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$. We let $s = s(u,v)$ be the maximal i such that there is a $g \in G_i^m$ such that $g[u]_i = [v]_i$. Let $g = g(u,v)$ be the unique g with this property and $(u',v') \in \mathcal{W}_m \times \mathcal{W}'_m$ be such that $g[u']_s = [v']_s$. Let $r(u',v')$ be the number of occurrences of (u',v') in (u,v). Then

$$\left| r(u',v') - \frac{l_n}{l_m} \frac{1}{Q_s^m} \left(\frac{1}{C_s^m} \right)^2 \right| < \epsilon_n (l_n/l_m).$$

8. **The word construction**

To finish the proof of Theorem 7, for each tree $\mathcal{T} \in \mathcal{T}rees$, we must build a sequence of collections of words $\{\mathcal{W}_n(\mathcal{T}) : \sigma_n \in \mathcal{T}\}$ satisfying the specifications E1–J11. This must be done so that the members of $\mathcal{W}_n(\mathcal{T})$ are entirely determined by $\mathcal{T} \cap \{\sigma_m : m \leq n\}$. We organize our construction so that for each n and for each subtree $\mathcal{S} \subseteq \{\sigma_m : m \leq n\}$ and each $\sigma_m \in \mathcal{S}$ we build $\mathcal{W}_m(\mathcal{S})$. To pass from stage $n-1$ to stage n we inductively assume that we have constructed $\{\mathcal{W}_m(\mathcal{S}) : \sigma_m \in \mathcal{S}\}$ for each subtree $\mathcal{S} \subseteq \{\sigma_m : m \leq n-1\}$. Our task is to construct the sequences of words for subtrees of $\{\sigma_m : m \leq n\}$. For those trees $\mathcal{S}$ with $\sigma_n \notin \mathcal{S}$, there is nothing to do.

This leaves finitely many trees $\mathcal{S}$ with $\sigma_n \in \mathcal{S}$. List these trees in any order as $\{\mathcal{S}_0, \ldots, \mathcal{S}_E\}$. Let $\mathcal{T}$ be the e^{th} tree on this list and suppose that we have constructed the words $\mathcal{W}_n$ for $\{\mathcal{S}_0, \ldots, \mathcal{S}_{e-1}\}$. Let P be the collection of prime numbers occurring in the prime factorization of any of the lengths of any of the words that we have constructed so far.

Let m be the largest number less than n such that $\sigma_m \in \mathcal{T}$. Our induction assumption tells us that we have the collection of words, equivalence relations and groups $\mathcal{W}_m(\mathcal{T})$, $\mathcal{Q}_s^m(\mathcal{T})$, $G_s^m(\mathcal{T})$ that satisfy the specifications for $m' \leq m$ and we need to construct $\mathcal{W}_n(\mathcal{T})$, $\mathcal{Q}_s^n(\mathcal{T})$ and $G_s^n(\mathcal{T})$. For notational simplicity, we will write $\mathcal{W}_n$ for $\mathcal{W}_n(T)$, G_s^n for $G_S^n(\mathcal{T})$ etc.

Some of the specifications are redundant and some require no explicit efforts to satisfy. For example specification E3 is follows from specification J10. To see this note that the first part of the statement follows from the second part: if we write $w \in \mathcal{W}_n$ as

$$w = bw_1 \cdots w_k e$$

where $w_i \in \mathcal{W}_m$ and it is also written as

$$w = u_1 u_2 \cdots u_{l_n/l_m},$$

then one of w_1 or w_k must overlap u_1 or u_{l_n/l_m} by at least half of the length of a word. Specification J10 implies that no significant end segment of u_1 can agree with an initial segment of w_1 and no significant initial segment of u_{l_n/l_m} can agree with a tail segment of w_k. Hence we do not need to separately verify E3.

Here is what we need to do:

1. Specifications E1 and E2 give some ground rules for constructing our words. Specification E1 will hold because we build the words by concatenating words in previous stages.
2. We build our words by iteratively randomly substituting $\mathcal{Q}_{s+1}^n$ classes into strings of $\mathcal{Q}_s^n$ classes. Specification Q6 will hold automatically from the form of our "Substitution Lemma". At each stage in the iteration we will use a "Finishing Lemma" to modify words that satisfy an approximation to E2 to make E2 hold precisely at that stage.
3. If $n = M(s)$, then we need to define the equivalence relation $\mathcal{Q}_s^n$. After this is done the equivalence relations $\mathcal{Q}_s^m$ are determined by specification Q5.
4. Specifications A7, A8 and A9 describe how the groups act on the words. If $m \geq M(s)$, and G_s^m acts freely on $\mathcal{W}_m/\mathcal{Q}_s^m \cup \text{rev}(\mathcal{W}_m/\mathcal{Q}_s^m)$, and the G_s^m action is subordinate to the G_{s-1}^m action, and we extend the G_s^m action to $\mathcal{W}_n/\mathcal{Q}_s^n \cup \text{rev}(\mathcal{W}_n/\mathcal{Q}_s^n)$ by the skew symmetric action, then we automatically get a free subordinate action.

   One of two possible requirements must be addressed at stage n:

   (a) If σ_n has length s and $m \geq M(s)$, then $G_s^n = G_s^m \oplus \mathbb{Z}_2$ where the additional generator corresponds to σ_n. This requires extending the skew symmetric action of G_s^m on the n-words to an action of all of G_s^n.

(b) In the other case σ_n has a length $s(n)$ longer than any of the sequences occurring in $\mathcal{T} \cap \{\sigma_m : m < n\}$. In this case $G^n_{s(n)} = \mathbb{Z}_2$ and the action of $G^n_{s(n)}$ must be defined. Note that for $s < s(n)$ we have $G_s^n = G_s^m$ so the G_s^n action is already defined.

5. Our main task is to satisfy the joining specifications J10 and J11. This is done by creating sequence of $\mathcal{W}_m/\mathcal{Q}_1^m$ words (and their reverses), substituting $\mathcal{W}_m/\mathcal{Q}_2^m$ words into the $\mathcal{Q}_1^m$ classes, $\mathcal{W}_m/\mathcal{Q}_3^m$ words into the $\mathcal{Q}_2^m$ classes, and so on until we get a sequence of $\mathcal{W}_m$ words. The iterative process uses the Substitution and Finishing Lemmas at each stage.

If $w \in {}^lX$ is a word in X and $x \in X$, then we will write $r(x, w)$ for the number of times that x occurs in w and $freq(x, w)$ for $r(x, w)/l$. Similarly, if $(w, w') \in {}^lX \times {}^lX$ and $(x, y) \in X \times X$, we will write $r(x, y, w, w')$ for the number of $i < l$ such that x is the i^{th} member of w and y is the i^{th} member of w'; i.e. the number of occurrences of (x, y) in (w, w'). We let $freq(x, y, w, w') = r(x, y)/l$.

8.1. *The basic lemmas.* We begin with three lemmas. The first lemma says that if we can build a collection of words that roughly satisfies our specifications, then we can "finish" them to exactly satisfy the specifications. The second is an application of the law of large numbers that says that we can iteratively substitute into equivalence relations to satisfy approximations to specifications J10 and J11. The third lemma gives our mechanism for extending group actions.

Here is our finishing lemma.

LEMMA 41 (Finishing). *Suppose that $w \in {}^lX$ where $l = k|X|$, and for each $x \in X$*

$$|r(x, w) - k| < \delta l,$$

then there is a w' that differs from w in at most $2\delta|X|l$ places such that for all $x \in X, r(x, w') = k$.

Proof. We argue as in Proposition 29. We can remove at most $2\delta|X|l$ many places in w and be left with a word w^* in which each element of X occurs exactly the same number of times. Then each element of X can be filled back into the empty slots the same number of times. The result is the word w'. □

We note that the proportion of w' that differs from w is $2\delta|X|$.

Our next lemma uses the following version of the Law of Large Numbers:

Law of Large Numbers. *Suppose that* $\{X_i : i \in \mathbb{N}\}$ *is a sequence of independent identically distributed* 2*-value random variables defined on a measure space* $(\mathcal{X}, \mathcal{B}, \mu)$ *taking value* 1 *with probability* p *and* 0 *with probability* $1-p$. *Let* $\delta > 0$. *Then*

$$P\left(\left|\frac{1}{n}\sum_{i=0}^{n-1} X_i - p\right| \geq \delta\right) < e^{-n\delta^2/4}. \tag{23}$$

We use this to show that randomly substituting element of an equivalence relation preserves the estimates of specifications J10 and J11.

The substitution will have the following initial data:

- An alphabet X and an equivalence relation $\mathcal{Q}$ on X with Q classes each class having cardinality C.
- Groups of involutions G and H with distinguished generators.
- A homomorphism $\rho : H \to G$ that preserves the distinguished generators. The kernel of ρ is H_0 and the range of ρ is G'.
- a free G action on $X/\mathcal{Q}$ and a free H action on X such that the H action is subordinate to the G action via ρ. In particular, for each l the skew diagonal actions of G on ${}^l X/\mathcal{Q}$ and H on ${}^l X$ are defined.
- $\epsilon_b < \epsilon_a$.

The Substitution Lemma says that for a collection of words of large enough length in the alphabet $X/\mathcal{Q}$, by relaxing ϵ_b to ϵ_a it is possible to substitute elements of X into the equivalence classes of $\mathcal{Q}$ in such a manner that preserves the statistics comparing words except as required by the H action.

Definition 42. Let $w \in {}^l(X/\mathcal{Q})$ and $w' \in {}^l X$. We will say that w' is a substitution instance of w if and only if

$$w' = x_0 x_1 \cdots x_{l-1} \text{ and } w = [x_0]_{\mathcal{Q}}[x_1]_{\mathcal{Q}} \cdots [x_{l-1}]_{\mathcal{Q}}.$$

If $\mathcal{W} \subseteq {}^l X/\mathcal{Q}$ is a collection of words, then we will say that $\mathcal{W}$ is symmetric if $\text{rev}(\mathcal{W}) = \mathcal{W}$.

Proposition 43 (Substitution Lemma). *There is a number* l_0 *depending on* $(\epsilon_b, \epsilon_a, Q, C, W, K \cdot |H_0|)$ *such that for all numbers* $l \geq l_0$ *and all symmetric* $\mathcal{W} \subseteq {}^l(X/\mathcal{Q})$ *with cardinality* W *that are closed under the skew diagonal action of* G, **if** *for all* k *with* $1 \leq k \leq (1-\epsilon_b)l$, $u, v \in X/\mathcal{Q}$ *and* $w, w' \in \mathcal{W}$:

$$\left|\frac{r(u, v, \text{sh}^k(w), w')}{(l-k)} - \frac{1}{Q^2}\right| < \epsilon_b \tag{24}$$

and each $u \in X/\mathcal{Q}$ *occurs with frequency* $1/Q$ *in each* $w \in \mathcal{W}$,

then *there is a collection of words* $S \subseteq {}^l X$ *consisting of substitution instances of* $\mathcal{W}$ *such that if* $\mathcal{W}' = HS \cup \text{rev}(HS)$ *we have*:

1. *every element of* $\mathcal{W}'$ *is a substitution instance of an element of* $\mathcal{W}$ *and each element of* $\mathcal{W}$ *has exactly* $K \cdot |H_0|$ *many substitution instances in* $\mathcal{W}'$;
2. *for each* $x \in X$ *and each* $w \in \mathcal{W}'$

$$\left|\frac{r(x,w)}{l} - \frac{1}{|X|}\right| < \epsilon_a, \tag{25}$$

i.e. the frequency of x *in* w *is within* ϵ_a *of* $1/|X|$;

3. *If* $w_1, w_2 \in S \cup \mathrm{rev}(S)$ *are different,* $x, y \in X, h \in H_0$ *and* $[w_1]_{\mathcal{Q}} = [w_2]_{\mathcal{Q}}$ *and then*

$$\left|\frac{r(x,y,w_1,hw_2)}{l} - \frac{1}{C^2}\right| < \epsilon_a; \tag{26}$$

4. *for all* k *with* $1 \le k \le (1-\epsilon_a)l,\ x, y \in X,\ w_1, w_2 \in \mathcal{W}'$

$$\left|\frac{r(x,y,\mathrm{sh}^k(w_1),w_2)}{l-k} - \frac{1}{|X|^2}\right| < \epsilon_a; \tag{27}$$

and

5. *for all* $x, y \in X$ *and all* $w_1, w_2 \in \mathcal{W}'$ *with different* H *orbits,*

$$\left|\frac{r([x]_{\mathcal{Q}},[y]_{\mathcal{Q}},[w_1]_{\mathcal{Q}},[w_2]_{\mathcal{Q}})}{l} - c\right| < \epsilon_b \tag{28}$$

implies that

$$\left|\frac{r(x,y,w_1,w_2)}{l} - \frac{c}{C^2}\right| < \epsilon_a. \tag{29}$$

Proof. Consider "rev" as a function on words, we start by choosing a set $R \subset \mathcal{W}$ that intersects each $G' \cup \{\mathrm{rev}\}$ orbit exactly once. If $\{S_r : r \in R\}$ is such that $S_r \subseteq X^l$ and S_r is a collection of K substitution instances of r, then

$$\bigcup_{r\in R} HS_r \cup \mathrm{rev}\Big(\bigcup_{r\in R} HS_r\Big)$$

is a candidate for $\mathcal{W}'$.

If S satisfies conclusion 3, then distinct elements of S have distinct H orbits. Since the H action is subordinate to the G action and H acts freely on HS_r, the number of substitution instances of r in HS_r is $|H_0|K$. We will use the law of large numbers to show that for large l and for most choices of S_r, $\mathcal{W}' = \bigcup_{r\in R} HS_r \cup \mathrm{rev}(\bigcup_{r\in R} HS_r)$ satisfies equations (25), (26), (27) and (29).

For counting purposes we let $\mathcal{X}_l$ be the collection of all ordered sets S of the form $\bigcup_{r\in R} S_r$. We put the counting measure on $\mathcal{X}_l$. We can identify $\mathcal{X}_l$ with a large product space:

$$\mathcal{X}_l = \prod_{r\in R}\prod_{q=0}^{K-1} S(r,q) \tag{30}$$

where, if $r = [x_0]_{\mathcal{Q}}[x_1]_{\mathcal{Q}} \cdots [x_{l-1}]_{\mathcal{Q}}$ then $S(r,q) = \prod_{j=0}^{l-1}[x_j]_{\mathcal{Q}}$. With this definition elements of $S(r,q)$ consist of all substitution instances of r.

Thus we can view $S(r,q)$ as the product probability space of the finite space with C elements, giving equal weight to each element. Substituting particular x's into equivalence classes in different coordinates of the product give independent events of probability $1/C$.

Equation (25) is the most straightforward. Each instance is determined by fixing $r \in R$, $q<K$, $h \in H$ and $x \in X$. For $\vec{w} \in \mathcal{X}_l$, let w be its component in $S(r,q)$ For $i < l$, X_i be the random variable defined on $\mathcal{X}_l$ that takes value 1 if x occurs in the i^{th} place in hw and 0 otherwise. Then the X_i are independent and identically distributed. By the law of large numbers, there is an $l(\delta, C)$ and a $\gamma < 1$ so that for $l \geq l(\delta, C)$, all but $(1-\gamma^l)$ portion of $\mathcal{X}_l$ has $|\frac{1}{l}\sum_{i=0}^{l-1} X_i - \frac{1}{|X|}| < \delta$. Since the number of requirements is $|R|K|H||X|$ (i.e. fixed) we can find an l' depending on ϵ_a, Q, C, W, K such that for $l \geq l'$ the collection N_0 of elements of $\mathcal{X}_l$ that satisfy equation (25) is the vast majority of members of $\mathcal{X}_l$.

Equation (26) follows similarly. Depending on whether w_1 and w_2 are in S or $\mathrm{rev}(S)$ we have three different requirements. They are handled identically so we assume we are dealing with the case that w_1 and w_2 both belong to S. Here each requirement is determined by $r \in R, q_1, q_2 < K$, $h \in H_0$ and $(x,y) \in X \times X$. For each class $u \in X/\mathcal{Q}$ we let O_u be those $j < l$ such that u is in the j^{th} place of r. By assumption $|O_u| = l/Q$. For $\vec{w} \in \mathcal{X}_l$, let w_1 be its component in $S(r,q_1)$ and w_2 its component in $S(r,q_2)$. For $j \in O_u$, let X_j be the random variable which takes value 1 if (x,y) occur in the j^{th} place in (w_1, hw_2) and zero otherwise. Then the X_j's for $j \in O_u$ are independent and identically distributed with mean $1/C^2$. Since O_u contains a fixed proportion of l, the law of large numbers can again be applied to calculate a particular $l(\epsilon_a, C)$ for all $l \geq l(\epsilon_a, C)$,

$$\left|\frac{r(x,y,w_1,hw_2)}{l} - \frac{1}{C^2}\right| < \epsilon_a$$

for all but a portion of $\mathcal{X}_l$ shrinking exponentially in l. Since the number of requirements is fixed as $|R|K^2|H_0|$ we can calculate an l' determined by the numbers $\epsilon_a, Q, C, K|H_0|$ such that for all $l \geq l'$ the collection N_1 of those $\vec{w} \in \mathcal{X}_l$ that satisfy conclusion 3 is the vast majority of $\mathcal{X}_l$.

Each of the requirements of equation (27) is determined by fixing a $r_1, r_2 \in R$, $q_1, q_2<K$, $h_1, h_2 \in H$, $x, y \in X$ and a k between 1 and $(1-\epsilon_a)l$ and considering $w_1 \in S(r_1, q_1), w_2 \in S(r_2, q_2)$ or their reverses. For notational simplicity we neglect the reverses — they add three more cases that are handled in exactly the same manner.

Let $g_1 = \rho(h_1)$ and $g_2 = \rho(h_2)$. Let O be the collection of i such that $([x]_{\mathcal{Q}}, [y]_{\mathcal{Q}})$ occur at the i^{th} place in the overlap of $\text{sh}^k(g_1 r_1)$ and $g_2 r_2$. By equation (24), $O/(l-k)$ is within ϵ_b of $1/Q^2$. For $i \in O$, let X_i be the random variable that takes value 1 if the pair (x, y) occurs in the i^{th} place in the overlap of $sk^k(h_1 w_1)$ and $h_2 w_2$, and 0 otherwise.

Since the X_i are independent and identically distributed and $k < (1-\epsilon_b)l$, for any particular δ we can apply the law of large numbers to find an $l(\delta, k, Q, C, \epsilon_b)$ and a $\gamma < 1$ such that if $l \geq l(\delta, k, Q, C, \epsilon_b)$ all but $1 - \gamma^{l-k}$ portion of the elements of $\mathcal{X}_l$ have $|\frac{1}{|O|}\sum X_i - 1/C^2| < \delta$. Noting that $|X| = CQ$, we see that we can take δ very small and use equation (24) to see that for a $1 - \gamma^{l-k}$ portion of the elements of $\mathcal{X}_l$ satisfy

$$\left| \frac{r(x, y, \text{sh}^k(h_1 w_1), h_2 w_2)}{l-k} - \frac{1}{|X|^2} \right| < \epsilon_a. \tag{31}$$

Keeping $x, y, r_1, r_2, q_1, q_2, h_1, h_2$ fixed, letting $\delta' > 0$ and summing over k we see that for each δ' we can find an $l(\delta', Q, C, \epsilon_b)$ such that for all $l \geq l(\delta, Q, C, \epsilon_b)$, a portion of $\mathcal{X}_l$ of proportion at least $1 - \delta'$, we have that for all k with $1 < k < (1-\epsilon_b)l$ equation (31) holds. By taking δ' small enough we can arrange that for all but ϵ_a portion of $\mathcal{X}_l$ for all $x, y, r_1, r_2, q_1, q_2, h_1, h_2$ equation (31) holds. Let N_2 be this portion of $\mathcal{X}_l$.

To see that there is a choice of $w \in \mathcal{X}_l$ such that $\bigcup_r HS_r \cup \text{rev}(\bigcup_r HS_r)$ satisfies conclusion 5 as well, we use a similar law of large numbers argument. In this case we are given $u, v \in X$, $h_1, h_2 \in H$ and $w_1 \in S(r_1, k_1), w_2 \in S(r_2, k_2)$ (or their reverses — which we again neglect) in different H orbits for which $r(u, v, [w_1]_{\mathcal{Q}}, [w_2])$ is very close to cl, and we count how many times the pair (x, y) is substituted into occurrences of (u, v) in $(\rho(h_1)[w_1], \rho(h_2)[w_2])$ to get $(h_1 w_1, h_2 w_2)$.

For particular $r_1, r_2, k_1, k_2, h_1, h_2, x, y$ these substitutions again give us a collection of independent random variables, and the law of large numbers gives us an $l(\delta, \epsilon_b, C)$ so that for all $l \geq l(\delta, \epsilon_b, C)$

$$\left| \frac{r(x, y, w_1, w_2)}{l} - \frac{c}{C^2} \right| < \delta \tag{32}$$

for the vast majority of $w \in \mathcal{X}_l$.

By choosing δ small enough we get an $l(\epsilon_b, \epsilon_a, Q, C)$ such that for $l \geq l(\epsilon_b, \epsilon_a, Q, C)$ there is a set N_3 containing the vast majority of $\mathcal{X}_l$ for which conclusion 5 holds.

Since N_0, N_1, N_2 and N_3 have the vast majority of $\mathcal{X}_l$ for large l the intersection is nonempty. Any element w of the intersection yields a collection $\mathcal{W}'$ satisfying the conclusions of the lemma. □

In the next lemma we take l_0 to be as in Proposition 43 for the constants $(\epsilon_b, \epsilon_a^2/10|X|, Q, C, W, K \cdot |H_0|)$.

Proposition 44. *Suppose $\epsilon_b < \epsilon_a^2/5|X|$ and $l \geq l_0$ is a multiple of $|X|$. Then there is a collection S satisfying the conclusions of Proposition 43 such that every $x \in X$ occurs with frequency $1/|X|$ in each $w \in HS \cup \mathrm{rev}(HS)$.*

Proof. Since the action of H on X is free it suffices to find an S such that every $x \in X$ occurs with frequency $1/|X|$ in S and the conclusions hold for $HS \cup \mathrm{rev}(HS)$. Apply Proposition 43 for some $\delta = \epsilon_a^2/10|X|$ to get a collection of words S_0 satisfying the conclusions of the lemma for $\epsilon'_a = \delta$.

For each $w \in S_0$ and each $u \in X/\mathcal{Q}$ we can apply the finishing lemma to change w on the places where u occurs in $[w]_{\mathcal{Q}}$ so that each $x \in u$ occurs exactly C times. The result is a word w^* in which each x occurs exactly $1/|X|$ times, w^* is still a substitution instance of $[w]_{\mathcal{Q}}$ and for which w and w^* agree on a set of places of proportion at least $1 - 2\delta|X|$.

We check that the conclusions hold for $S = \{w^* : w \in S_0\}$ and ϵ_a. Since H acts freely and subordinately to G, conclusion 1 follows as before.

For $h \in H$ and $w \in S_0$ the difference between hw and hw^* is a fraction of at most $2\delta|X|$. It follows easily that conclusions 2, 3 and 5 hold with ϵ_a replacing δ.

We must verify equation (27). Suppose that $w_1^*, w_2^* \in HS \cup \mathrm{rev}(HS)$, $1 \leq k \leq (1-\epsilon_a)l$, and $x, y \in X$. Then there are words $s_1, s_2 \in S_0 \cup \mathrm{rev}(S_0)$ and $h_1, h_2 \in H$ such that $w_1^* = h_1 s_1^*, w_2^* = h_2 s_2^*$. Let $w_i = h_i s_i$. The $w_i \in HS_0 \cup \mathrm{rev}(HS_0)$ and w_i and w_i^* differ on at most $2\delta|X|l$ places. Hence

$$|r(x, y, \mathrm{sh}^k(w_1), w_2) - r(x, y, \mathrm{sh}^k(w_1^*), w_2^*)| < 4\delta|X|l.$$

As a consequence

$$\left| \frac{r(x, y, \mathrm{sh}^k(w_1), w_2)}{k - l} - \frac{r(x, y, \mathrm{sh}^k(w_1^*), w_2^*)}{k - l} \right| < \frac{4\delta|X|}{\epsilon_a} < (4/5)\epsilon_a.$$

Equation (27) follows by the triangle inequality and the fact that $(4/5)\epsilon_a + \delta < \epsilon_a$. $\square$

An argument similar, but simpler than that of Proposition 43 gives the minimal orbit structure coming from the system of groups $\langle G_s^m : s \leq s(m) \rangle$. For $s < s(m)$, let H_{s+1} be the kernel of $\rho_{s+1,s} \restriction G_{s+1}^m$. Define $\kappa_1 = |G_1^m|$ and $\kappa_{i+1} = \kappa_i |H_{i+1}|$.

Lemma 45. *Let M be such that $2^M \geq K_s$. Then for all large multiples of $\mathcal{W}_m$ l, there is a collection of words $\mathcal{W} \subseteq {}^l\mathcal{W}_m$ such that:*

1. $|\mathcal{W}| = 2^M$,
2. $|\mathcal{W}/\mathcal{Q}_s^m| = \kappa_s$ *for all* $s \leq s(m)$,

3. G^m_s *acts freely on each* $\mathcal{W}/\mathcal{Q}^m_s$,
4. *each* $u \in \mathcal{W}_m$ *occurs in each* $w \in \mathcal{W}$ *the same number of times.*

We will use the following extension lemma to build our group actions:

LEMMA 46. *Let* X *be a set and* $R \subseteq Q$ *equivalence relations on* X. *Suppose that*

1. $\rho : H \to G \times \mathbb{Z}_2$ *is a homomorphism.*
2. $G \times \mathbb{Z}_2$ *acts on* X/Q.
3. H *acts on* X/R *by an action subordinate to the* $G \times \mathbb{Z}_2$ *action.*
4. *If* $H_0 = \{h \in H : \rho(h) = (0, i), i \in \{0, 1\}\}$, *then every orbit of the* $\mathbb{Z}_2$ *factor of* $G \times \mathbb{Z}_2$ *contains an even number of* H_0 *orbits.*

Then there is a free action of $H \times \mathbb{Z}_2$ *on* X/R *subordinate to the* $G \times \mathbb{Z}_2$ *action via the map* $\rho'(h, i) = \rho(h) + (0, i)$.[17]

8.2. *The initial data.* We recall from Definition 2 the notation $s(m)$ for the maximum length of a sequence in $\{\sigma'_m : m' \leq m\} \cap \mathcal{T}$. In our inductive construction we are assuming that we are given $\mathcal{W}_m, G^m_s, \mathcal{Q}^m_s$ for $s \leq s(m)$, together with the appropriate group actions. We are also given some numerical parameters $W_m, Q^m_s, C^m_s, k(m), \epsilon_m$ and l_m.

The numbers W_n, Q^n_s, C^n_s are determined by $k(n)$. We start by choosing $k(n)$ large enough that $2^m 2^{-k(n)} < \epsilon_m$. Next we choose $\epsilon_n < 2^{-n-1}$ small enough that $2\epsilon_n W_n^2 < \epsilon_m$. Finally we will choose a prime number $p_n > 2^{k(n)}$ larger than the maximum of P and so large that if we set $l_n = p_n^2 W_m l_m$ then $\epsilon_n (l_n/l_m) W_m^{-2} > n$. Exactly how large we must choose p_n to be is determined in the next section by the substitution lemma. Any choices made this way satisfy our "numerical requirements".

8.3. *Rolling the dice.* We break our word construction into two cases, according to whether $s(n) = s(m)$ or $s(n) = s(m) + 1$. In both cases we apply the substitution and finishing lemmas $s(n)+1$ -times inductively to build collections of words, with the i^{th} collection a subset of $(\mathcal{W}_m/\mathcal{Q}^m_i)$.

Choose $\delta_0, \delta_1 \cdots \delta_{s(n)}$ so that $\delta_i < \delta^2_{i+1}/10W_m$ and $\delta_{s(n)} < \epsilon^2_n/100W_m$ and let $\delta_{s(n)+1} = \epsilon_n/100$. For $i \leq s(n)$, let $C = 2^{k(n)}$ and $Q_i = (2^{k(n)})^i$, $X_i = \mathcal{W}_m/\mathcal{Q}^m_i$, $X_{s(m)+1} = \mathcal{W}_m$, $G_i = G^m_i$ and $G_{s(m)+1}$ be the trivial group acting trivially on $\mathcal{W}_m$. Choose K_i so that $K_i|ker(\rho_{i,i-1})| = 2^{k(n)}$ and $K_{s(n)+1} = 2^{k(n)}$. The substitution lemma gives us numbers

$$l_i = l(\delta_i, \delta^2_{i+1}/10|X_i|, Q_i, C, W_m, 2^{k(n)})$$

for each i.

[17]This lemma is easy to prove because the groups in question are Abelian. Analogous extension results for arbitrary groups are shown in the forthcoming paper [4].

Choose a prime number p_n large enough to satisfy the numerical requirements and larger than any element of P so that

$$100/p_n < \epsilon_n^2 \text{ and } p_n W_m > l(\delta_i, \delta_{i+1}, Q_i, C, W_m, 2^{k(n)})$$

for all i. Let $l_n = p_n^2 W_m l_m$.

Case 1: $s(n) = s(m)$. We apply Proposition 44 and substitute $s(m)+1$-times. The result is a collection of words $\mathcal{W}_n$ that satisfy E1–E3, specifications Q4–A9 hold relative to the groups G_s^m and the equivalence relations $\mathcal{Q}_s^n$ for $s \leq s(m)$. The specifications J10 and J11 hold with ϵ_n replaced by $\epsilon_n/100$. To finish we need only define the group actions G_s^n for $s \leq s(n)$. The groups $G_s^n = G_s^m$ for all s except $s_0 = \mathrm{lh}(\sigma_n)$. Here $G_{s_0}^n = G_{s_0}^m \oplus \mathbb{Z}_2$. By lemma 46, we can extend the $G_{s_0}^m$ action to a $G_{s_0}^n$ action subordinate to the $G_{s_0-1}^n$-action. This finishes the construction in this case.

Case 2: $s(n) = s(m)+1$. In this case σ_n is the only sequence in $\mathcal{T} \cap \{\sigma_{m'} : m' \leq n\}$ of length $s(n)$. The group $G_{s(n)}^n = \mathbb{Z}_2$. In addition to building $\mathcal{W}_n$ we need to define $\mathcal{Q}_{s(n)}$ and the action of $\mathbb{Z}_2$ on $\mathcal{W}_n/\mathcal{Q}_{s(n)}$.

We will build two collections of collections of words $\mathcal{W}^\dagger$ and $\mathcal{W}^{\dagger\dagger}$. We will define $\mathcal{W}_n = \{w_0 \frown w_1 : w_0 \in \mathcal{W}^\dagger, w_1 \in \mathcal{W}^{\dagger\dagger}\}$ and take $\mathcal{Q}_{s(n)}^n$ to be the equivalence relation that partitions $\mathcal{W}_n$ into pieces of the form $P_{w_0} = \{w_0 \frown w_1 : w_1 \in \mathcal{W}^{\dagger\dagger}\}$.

Let $\mathcal{W}^{\dagger\dagger} \subseteq {}^{p_n W_m}\mathcal{W}_m$ be a collection of words of length $p_n W_m l_m$ of cardinality $2^{k(n)}$ built by Lemma 45, with each word in $\mathcal{W}_m$ occurring p_n times.

Let $\mathcal{W}^\dagger$ be the result of applying Proposition 44 $s(n)$-times with K_i taken so that $K_i \kappa_i |H_0| = 2^{k(n)}$ and $l = (p_n^2 - p_n) W_m l_m$. The result is that $\mathcal{W}^\dagger$ is a collection of words of length $(p_n^2 - p_n) W_m l_m$ satisfying the specifications E1–J11 with respect to the groups G_s^n for $s < s(n)$.

We need to check that the specifications continue to hold for $\mathcal{W}_n$. The specifications E1–A9 are easy to verify. We need to check that the tiny modifications made at the end of words in $\mathcal{W}^\dagger$ do not affect the specifications J10 and J11.

We verify J10. Let u and v be elements of $\mathcal{W}_n \cup \mathrm{rev}(\mathcal{W}_n)$ and $1 \leq k < (1-\epsilon_n)(l_n/l_m)$. Let $u', v' \in \mathcal{W}_m \cup \mathrm{rev}(\mathcal{W}_m)$ have the correct parities. Write $u = u_0 \frown u_1$ and $v = v_0 \frown v_1$. Let $r^\dagger(u', v')$ be the number of occurrences of (u', v') in the overlap of $\mathrm{sh}^{kl_m}(u_0)$ and v_1. Then

$$\left| r^\dagger(u', v') - r(u, v) \right| \leq p_n W_m.$$

So,

$$\left| \frac{r(u', v')}{k(l_n/l_m)} - \frac{r^\dagger(u', v')}{k(l_n/l_m)} \right| \leq \frac{1}{\epsilon_n p_n} < \epsilon_n/100. \tag{33}$$

Moreover,

$$\left|\frac{r^\dagger(u',v')}{\epsilon_n p_n^2 W_m} - \frac{r^\dagger(u',v')}{\epsilon_n (p_n^2 - p_n) W_m}\right| < \epsilon_n/50 \tag{34}$$

and finally by the substitution construction

$$\left|\frac{r^\dagger(u',v')}{\epsilon_n (p_n^2 - p_n) W_m} - \frac{1}{W_m^2}\right| < \epsilon_n/100. \tag{35}$$

Putting equations (33), (34) and (35) together the triangle inequality gives us:

$$\left|\frac{r(u',v')}{k(l_n/l_m)} - \frac{1}{W_m^2}\right| < \epsilon_n$$

as needed.

The verification of J11 is similar and easier.

What remains is to define the $G^n_{s(n)}$ action. But it follows from Lemma 46 that there is an action of $\mathbb{Z}_2$ on the $\{C_w : w \in \mathcal{V}_n\}$ subordinate to the action of $G^m_{s(m)}$. We take the $G^m_{s(n)}$ action to be this action.

This completes the construction of the words and hence the proof of Theorem 7.

9. **Marginalia**

In this section we tie up some loose ends.

9.1. *Positive entropy and continuous conjugacies.* We begin with two remarks dealing with obvious questions: what can you say about positive entropy transformations, and what happens if you ask that the conjugacies be continuous.

9.1.1. *Positive entropy.* The transformations we build are easily seen to be zero entropy, as the number of words of length n grows slowly. However, if $\mathcal{B}$ is a Bernoulli shift, $\mathcal{T} \in \mathcal{T}rees$ and we take the product transformation $\mathcal{C} = \mathcal{F}(\mathcal{T}) \times \mathcal{B}$, then $F(\mathcal{T})$ is the Pinsker algebra of $\mathcal{C}$. It follows that $\mathcal{C} \cong \mathcal{C}^{-1}$ if and only if $F(\mathcal{T}) \cong F(\mathcal{T})^{-1}$. Thus for any fixed entropy h, the collection of pairs of ergodic transformations (S, T) that have entropy h and are isomorphic is a complete analytic set.

9.1.2. *Continuous conjugacies.* As remarked after Lemma 26, we have proved something stronger. Either:

1. $F(\mathcal{T})$ is not congruent to $F(\mathcal{T})^{-1}$

   or
2. There is a $\mathcal{G}_\delta$ set of measure one $L = L_\mathcal{T}$ on which $F(\mathcal{T})$ is continuous and a continuous isomorphism between $F(\mathcal{T}) \restriction L$ and $F(\mathcal{T})^{-1} \restriction L^{-1}$.

The set $L_\mathcal{T}$ had an explicit description, making it easily computable from $\mathcal{T}$.

This strengthens our result to the setting of measure preserving homeomorphism of $\mathcal{G}_\delta$ subsets of $\mathbb{K}$, with the equivalence relation arising by restricting to measure preserving conjugacies that are homeomorphisms.

9.2. *Bijective Borel reduction.* It follows from theorems of Kechris ([13]) and Louveau-St. Raymond ([15, Th. 4]) that:

Fact. If $A \subseteq X$ is a complete analytic set and $B \subseteq Y$ is an arbitrary analytic set there is a one-to-one, continuous function $f : Y \to X$ reducing B to A.

Thus if we have two complete analytic sets $A \subseteq X$ and $B \subseteq Y$ we can find one-to-one, continuous $f : X \to Y$ and $g : Y \to X$ reducing A to B and B to A respectively. Doing the usual Cantor-Bernstein argument (see, for example, [22]) we define

$$Y_0 = Y \setminus f(X) \text{ and } X' = X \setminus \bigcup_{n=0}^{\infty} (gf)^n g Y_0.$$

If we define $h : X \to Y$ by

$$h(x) = \begin{cases} f(x) & \text{if } x \in X' \\ g^{-1}(x) & \text{otherwise,} \end{cases}$$

then h is a (low level) Borel bijection between X and Y sending A to B.

The main theorem of this paper is that the collection of pairs (μ, ν) of ergodic measure preserving transformations such that

$$(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, \mathrm{sh}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh})$$

is a complete analytic subset of $\mathcal{E} \times \mathcal{E}$. As a consequence we get:

Proposition 47. *There is a Borel bijection F from $\mathcal{T}rees$ to $\mathcal{E} \times \mathcal{E}$ such that $\mathcal{T}$ is ill-founded if and only if $F(\mathcal{T}) = (\mu, \nu)$ and $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, \mathrm{sh}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh})$.*

In the concrete context of this paper, it is possible to explicitly exhibit the continuous injections between $\mathcal{T}rees$ and $\mathcal{E} \times \mathcal{E}$, and this gives some insight into the reason the two sets are equivalent. We now describe these injections.

For the rest of this section we will abuse notation to some extent by identifying a shift invariant measure μ on $\Sigma^{\mathbb{Z}}$ with the corresponding measure preserving system $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, \mathrm{sh})$. We will also assume that Σ has an ordering in type $\mathbb{N}$ so that for each natural number N, we can define $\Sigma \upharpoonright N$ to be the first N elements of Σ.

In the proof of Theorem 7 we built a continuous $F_0 : \mathcal{T}rees \to \mathcal{E}$ such that:

1. $\mathcal{T}$ is ill-founded if and only if $F_0(\mathcal{T}) \cong F_0(\mathcal{T})^{-1}$
   and
2. if $\mathcal{T} \neq \mathcal{T}'$ then $F_0(\mathcal{T}) \ncong F_0(\mathcal{T}')$.

The second item is clear from the construction since the eigenvalues of the system corresponding to $\mathcal{T}$ are different from the eigenvalues of the system corresponding to $\mathcal{T}'$. (This is noted explicitly in Corollary 34.)

The map $I : \mathcal{E} \to \mathcal{E} \times \mathcal{E}$ defined by $\mu \mapsto (\mu, \mu^{-1})$ is continuous; hence $F =_{\text{def}} I \circ F_0 : \mathcal{T}rees \to \mathcal{E} \times \mathcal{E}$ is a continuous, one-to-one reduction of the ill-founded trees to $\{(\mu, \nu) : \mu \cong \nu\}$.

To use the Cantor-Bernstein theorem to find a bijection, we must now define a continuous injection G from the set of pairs (μ, ν) of ergodic measures into $\mathcal{T}rees$ so that $\mu \cong \nu$ if and only if $G(\mu, \nu)$ is ill-founded.

The idea behind building G is that any conjugacy between μ and ν can be well approximated by a finite "stationary code". The collection of finite stationary codes naturally form a tree, and a branch through this tree gives a coherent collection of finite approximations that converge to a conjugacy between μ and ν.

To uniformize notation and terminology, we will give a brief overview of the theory of codes. We begin by defining a *code* to be a function $\phi : \Sigma^{[-N,N]} \to \Sigma$ for some $N \in \mathbb{N}$. Such a code induces a uniformly continuous function $\phi^* : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ by setting

$$\phi^*(x)(k) = \phi(x \restriction [k-N, k+N]).$$

Suppose μ is a shift invariant measure on $\Sigma^{\mathbb{Z}}$ and that $\psi : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ is a measurable function that commutes with the shift map. By approximating with continuous function we see that for each positive ϵ there is an N and a code $\phi : \Sigma^{[-N,N]} \to \Sigma$ such that $\{x : \phi(x \restriction [-N, N]) \neq \psi(x)(0)\}$ has measure less than ϵ.

To reduce our collection of approximations to a countable set we will want to use *partial codes of length N*. We begin by choosing N_μ large enough that if R is the collection of basic open intervals determined by functions $s : [-N, N] \to \Sigma \restriction N_\mu$, then

$$\mu\Big(\bigcup\{\langle s\rangle : s \in R\}\Big) > 1 - 2^{-N}.$$

A *partial code* is a function $\phi : (\Sigma \restriction N_\mu)^{[-N.N]} \to \Sigma$.

We will say a partial code $\phi_0 : (\Sigma{\restriction}N_\mu)^{[-N,N]} \to \Sigma$ has a property if and only if every code $\phi : \Sigma^{[-N,N]} \to \Sigma$ extending ϕ_0 has the property. For example, given an partial code $\phi_0 : (\Sigma \restriction N_\mu)^{[-N,N]} \to \Sigma$, let us s say that ϕ_0 (ϵ, n)-*approximates* the pair (μ, ν) if and only if for all basic open intervals $\langle s\rangle$ of length n, and all codes $\phi : \Sigma^{[-N,N]} \to \Sigma$ extending ϕ_0 we have $\phi^*(\langle s\rangle)$ has ν-measure within ϵ of $\mu(\langle s\rangle)$.

We observe that to determine whether ϕ_0 (ϵ, n)-approximates (μ, ν) we need only know the μ and ν measures of a finite number of basic open intervals in $\Sigma^{\mathbb{Z}}$. Moreover, if ψ is a measure preserving transformation and

$\phi : \Sigma^{[-N,N]} \to \Sigma$ is a code such that $\{x : \phi(x \restriction [-N,N]) \neq \psi(x)(0)\}$ has measure less than ϵ and N is large enough relative to ϵ, then any code ϕ' extending the restriction ϕ_0 of ϕ to $(\Sigma \restriction N_\mu)^{[-N,N]}$ has the same property.

There is a natural partition metric on codes. Given codes ϕ, ψ we can take

$$d(\phi,\psi) = \sum_{v\in\Sigma} \mu((\phi)^{-1}(v)\Delta(\psi)^{-1}(v)).$$

Following our convention we will say that a sequence of partial codes $\phi_0^i : (\Sigma \restriction (N_i)_\mu)^{[-N_i,N_i]} \to \Sigma$ is Cauchy if and only if for each sequence of codes ϕ^i with ϕ^i extending ϕ_0^i is Cauchy in this metric.

A sequence of partial codes $\langle \phi_0^i \rangle$ with $N_i \to \infty$ that are Cauchy in this metric such that ϕ_0^i $(1/i, i)$-approximates (μ,ν) determines a unique shift invariant, measure preserving map $\phi : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ such that for all $s : n \to \Sigma$

$$\mu(\phi^{-1}(\langle s\rangle)\Delta(\phi_i)^{-1}(\langle s\rangle)) \to_i 0.$$

This map ϕ must be a factor map from $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, \mathrm{sh})$ to $(\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh})$. To approximate a conjugacy we need also approximate ϕ^{-1}.

Suppose that ϕ_0 and ψ_0 are partial codes. We will say that ϕ_0 and ψ_0 are (ϵ, n)-*approximate inverses* if and only if for all ϕ extending ϕ_0 and ψ extending ψ_0 and all basic open intervals $\langle s\rangle$ of length less than or equal to n,

$$\mu((\psi^* \circ \phi^*)(\langle s\rangle)\Delta\langle s\rangle) < \epsilon.$$

We again note that one is able to determine whether ϕ_0 and ψ_0 are (ϵ, n)-approximate inverses from knowledge of a finite part of μ.

We now summarize our observations:

LEMMA 48. *Let μ, ν be ergodic measures on $\Sigma^{\mathbb{Z}}$. Then $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \mu, \mathrm{sh}) \cong (\Sigma^{\mathbb{Z}}, \mathcal{C}, \nu, \mathrm{sh})$ if and only if there is a sequence $\langle(\phi_i, \psi_i) : i \in \mathbb{N}\rangle$ of partial codes of length $N_i \to \infty$ such that*

1. *$d(\phi_i, \phi_{i+1}) < 1/i$ and $d(\psi_i, \psi_{i+1}) < 1/i$,*
2. *ϕ_i $(1/i, i)$-approximates (μ, ν) and ψ_i $(1/i, i)$-approximates (ν, μ), and*
3. *(ϕ_i, ψ_i) are $(1/i, i)$ approximate inverses.*

We can now describe a tree $\mathcal{S}(\mu, \nu)$ we associate with (μ, ν). A node in our tree will consist of a finite sequence $(t_0, \ldots, t_n)$ of 4-tuples $t_i = (\mu^*, \nu^*, \phi_i, \psi_i)$ where:

1. μ^* (respectively ν^*) is a list of rational numbers q_s indexed by functions $s : m \to \Sigma \restriction i$ for $m \leq i$ such that q_s is within $1/i$ of the μ-measures (respectively ν-measures) of basic open interval $\langle s\rangle$,
2. ϕ_i $(1/i, i)$-approximates (μ, ν) and ψ_i $(1/i, i)$-approximates (ν, μ),
3. (ϕ_i, ψ_i) are $(1/i, i)$ approximate inverses, and
4. both $d(\phi_i, \phi_{i+1})$ and $d(\psi_i, \psi_{i+1})$ are less than $1/i$.

By including the μ^* and ν^* we explicitly arrange that the map sending (μ, ν) to $S(\mu, \nu)$ is one-to-one. Moreover to determine whether $(t_0, \ldots, t_n)$ belongs to $\mathcal{S}(\mu, \nu)$ one only needs to know the μ and ν values of a finite number of basic open sets with sufficient accuracy. In particular, the collection of pairs (μ, ν) of ergodic measures such that $(t_0, \ldots, t_n) \in \mathcal{S}(\mu, \nu)$ is a weak*-open set.

Now take an arbitrary bijection b between the natural numbers and the set of 4-tuples $(\mu^*, \nu^*, \phi, \psi)$ that satisfy items 1–3. Define $\mathcal{T}_0(\mu, \nu)$ to be the collection of $(x_0, \ldots, x_n) \in \mathbb{N}^{<\mathbb{N}}$ with the property that $(b(x_0), \ldots, b(x_n)) \in \mathcal{S}(\mu, \nu)$. By the remarks in the previous paragraph, the map $(\mu, \nu) \mapsto \mathcal{T}_0(\mu, \nu)$ is a continuous, one-to-one map from the space of ergodic measures on $\Sigma^{\mathbb{Z}}$ with the weak* topology to the collection of trees with the induced topology from $\{0, 1\}^{\mathbb{N}^{<\mathbb{N}}}$ with the property that μ is isomorphic to ν if and only if $\mathcal{T}_0(\mu, \nu)$ has an infinite branch.

It could happen that $\mathcal{T}_0(\mu, \nu)$ is not in $\mathcal{T}rees$, because it does not have nodes of arbitrarily long length. To remedy this we fix a tree

$$\mathcal{T}^* \subseteq \{\text{odd numbers}\}^{<\mathbb{N}}$$

that belongs to $\mathcal{T}rees$ and has no infinite branch. Given an arbitrary tree $\mathcal{T}_0 \subseteq \mathbb{N}^{<\mathbb{N}}$, define

$$\mathcal{T}_1 = \{(2x_0, 2x_1, \ldots, 2x_n) : (x_0, x_1, \ldots, x_n) \in \mathcal{T}_0\}$$

and

$$\mathcal{T}_2 = \mathcal{T}_1 \cup \mathcal{T}^*.$$

The map $\mathcal{T}_0 \mapsto \mathcal{T}_2$ is a continuous injection from the collection of trees to $\mathcal{T}rees$. It follows that the map $G(\mu, \nu) = \mathcal{T}_2(\mu, \nu)$ is the desired injection.

Summarizing: Theorem 7 gives an injection from $\mathcal{T}rees$ to pairs of ergodic measures (μ, ν) that reduces the ill-founded trees to the collection of conjugate pairs. We have just described a method for building a continuous injection from the pairs of ergodic measures into $\mathcal{T}rees$ that reduces conjugate pairs to ill-founded trees. An application of the Cantor-Bernstein theorem finishes the alternate proof of Proposition 47.

9.3. *Flows.* Given a standard measure space $(X, \mathcal{B}, \mu)$ there is a natural topology on the space of measure preserving $\mathbb{R}$-flows. One can ask whether the isomorphism relation on ergodic flows is Borel.

Remark 49. There is a continuous map $\mathcal{S}$ from the ergodic transformations on $[0, 1]$ to the ergodic measure preserving $\mathbb{R}$-flows on $[0, 1] \times [0, 1]$ such that for all ergodic transformations S, T we have

$$S \cong T \text{ if and only if } \mathcal{S}(S) \cong_{\mathbb{R}} \mathcal{S}(T),$$

where $\cong_{\mathbb{R}}$ is isomorphism of $\mathbb{R}$-flows.

This remark immediately implies that the isomorphism relation of measure preserving $\mathbb{R}$-flows is complete analytic, and hence not Borel.

The map $\mathcal{S}$ is the standard "suspension" map. We briefly describe it. Given an invertible measure preserving transformation $T : [0,1] \to [0,1]$ we consider the equivalence relation on $[0,1] \times \mathbb{R}$ generated by setting $(x,s) \sim (T^n(x), s-n)$ for each n. We can take $[0,1] \times [0,1)$ as a fundamental domain for $([0,1] \times \mathbb{R})/\sim$. Then the transformation $\mathcal{S}(T)(t, [(x,s)]_\sim) = [(x, s+t)]_\sim$ defines a measure preserving flow on $[0,1] \times [0,1)$ with the product Lebesgue measure.

It is a standard fact that the map $\mathcal{S}$ is continuous and that it takes ergodic transformations to ergodic flows. Suppose now that $\mathcal{S}(S) \cong_{\mathbb{R}} \mathcal{S}(T)$. Define the "time one" measure preserving transformations of $[0,1] \times [0,1)$ by setting $S^*([(x,s)]_\sim) = [(x,s+1)]_\sim$. Equivalently, $S^*([(x,s)]_\sim) = [(Sx), s)]_\sim$. We define T^* similarly. Then S^* and T^* are not ergodic, but each ergodic component of S^* is isomorphic to S and each ergodic component of T^* is isomorphic to T. Consequently $S \cong T$.

10. **Rank one transformations**

In this section we use a theorem of King to show that the collection of pairs (S,T) of rank one transformations with S conjugate to T is a Borel set. This gives a dense $\mathcal{G}_\delta$ class of transformations on which the conjugacy relation is Borel. The proof is quite easy but unfortunately does not give a useful structure theorem for rank one transformations.

To prove this theorem, it is convenient to use the group of measure preserving transformations of the unit interval, **MPT**, as our model.

The main tool is the following theorem of King [14]:

Theorem: Suppose that T is an ergodic rank one transformation. Then the centralizer of T is the closure of $\{T^n : n \in \mathbb{Z}\}$ in **MPT**.

The Effros-Borel space of a Polish space is the collection $\mathcal{F}(X)$ of closed subsets of X with the σ-algebra generated by basic sets of the form:

$$\{F \in \mathcal{F}(X) : F \cap U \neq \emptyset\},$$

where $U \subset X$ is open. It is a standard Borel space in the sense that there is a Polish topology (the Fell topology) on $\mathcal{F}(X)$ for which it is the collection of Borel sets.

We will use the following facts whose proofs can be found in [2] and [12] respectively:

1. (Dixmier) If G is a Polish group, then there is a Borel $\mathfrak{T} \subset G \times \mathcal{F}(G)$ such that
   (a) $\{x : \mathfrak{T}(x, H)\} \neq \emptyset$ if and only if H is a closed subgroup of G, and

(b) if $\mathfrak{T}_H = \{x : \mathfrak{T}(x, H)\}$ and H is a closed subgroup of G, then $\mathfrak{T}_H$ is a transversal for G/H, the space of left cosets of H in G.

2. Suppose that X and Y are Polish spaces and $B \subset X \times Y$ is a Borel set such that for all $x \in X$ there is at most one $y \in Y$ with $(x, y) \in B$. Then $\{x : (\exists y)(x, y) \in B\}$ is Borel.

We now follow [2, Lemma 7.1.2]. Let R be the collection of ergodic rank one transformations.

LEMMA 50. *Define* $C : R \to \mathcal{F}(\mathbf{MPT})$ *by* $T \mapsto C(T)$. *Then* C *is a Borel map.*

Proof. It suffices to see that the C-inverse of a basic set is Borel. Suppose that $O = \{F \in \mathcal{F}(\mathbf{MPT}) : F \cap U \neq \emptyset\}$ for some open set U in $\mathbf{MPT}$. Then $C^{-1}(O) = \{T : C(T) \cap U \neq \emptyset\}$. By King's theorem,

$$C(T) \cap U \neq \emptyset \text{ if and only if } (\exists n) T^n \in U.$$

The collection of T such that for some $n, T^n \in U$ is an open set in $\mathbf{MPT}$; hence $C^{-1}(O)$ is open. $\square$

THEOREM 51. *The collection of pairs* (S, T) *of rank one ergodic transformations that are conjugate is a Borel subset of* $\mathbf{MPT} \times \mathbf{MPT}$.

Proof. Let $\mathfrak{T}_{C(T)}$ be the Borel transversal $\mathbf{MPT}/C(T)$. Define

$$B = \{(S, T, g) : gTg^{-1} = S \text{ and } g \in \mathfrak{T}_{C(T)}\} \subseteq (R \times R) \times \mathbf{MPT}.$$

Then for each $(S, T) \in R \times R$ there is at most one g with $(S, T, g) \in B$. Hence $\{(S, T) : \text{for some } g \in \mathbf{MPT}, (S, T, g) \in B\}$ is Borel. $\square$

References

[1] H. ANZAI, On an example of a measure preserving transformation which is not conjugate to its inverse, *Proc. Japan Acad.* **27** (1951), 517–522. MR 0047742. Zbl 0044.12502. Available at http://projecteuclid.org/euclid.pja/1195571227.

[2] H. BECKER and A. S. KECHRIS, *The Descriptive Set Theory of Polish Group Actions*, *London Math. Soc. Lecture Note Ser.* **232**, Cambridge Univ. Press, Cambridge, 1996. MR 1425877. Zbl 0949.54052. doi: 10.1017/CBO9780511735264.

[3] M. FOREMAN, A descriptive view of ergodic theory, in *Descriptive Set Theory and Dynamical Systems* (Marseille-Luminy, 1996), *London Math. Soc. Lecture Note Ser.* **277**, Cambridge Univ. Press, Cambridge, 2000, pp. 87–171. MR 1774425. Zbl 0962.03043.

[4] ———, Examples of ergodic transformations with complicated centralizers, in preparation.

[5] M. FOREMAN, D. J. RUDOLPH, and B. WEISS, Models for measure preserving transformations, in preparation.

[6] M. Foreman and B. Weiss, An anti-classification theorem for ergodic measure preserving transformations, *J. Eur. Math. Soc.* (*JEMS*) **6** (2004), 277–292. MR 2060477. Zbl 1063.37004. Available at http://link.springer.de/cgi/linkref?issn=1435-9855&year=04&volume=6&page=277.

[7] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton Univ. Press, Princeton, N.J., 1981. MR 0603625. Zbl 0459.28023.

[8] E. Glasner, *Ergodic Theory via Joinings*, *Math. Surveys Monogr.* **101**, Amer. Math. Soc., Providence, RI, 2003. MR 1958753. Zbl 1038.37002.

[9] P. R. Halmos and J. von Neumann, Operator methods in classical mechanics. II, *Ann. of Math.* **43** (1942), 332–350. MR 0006617. Zbl 0063.01888. doi: 10.2307/1968872.

[10] G. Hjorth, *Classification and Orbit Equivalence Relations*, *Math. Surveys Monogr.* **75**, Amer. Math. Soc., Providence, RI, 2000. MR 1725642. Zbl 0942.03056.

[11] ———, On invariants for measure preserving transformations, *Fund. Math.* **169** (2001), 51–84. MR 1852353. Zbl 0990.03036. doi: 10.4064/fm169-1-2.

[12] A. S. Kechris, *Classical Descriptive Set Theory*, *Grad. Texts in Math.* **156**, Springer-Verlag, New York, 1995. MR 1321597. Zbl 0819.04002.

[13] ———, On the concept of Π_1^1-completeness, *Proc. Amer. Math. Soc.* **125** (1997), 1811–1814. MR 1372034. Zbl 0864.03034. doi: 10.1090/S0002-9939-97-03770-2.

[14] J. King, The commutant is the weak closure of the powers, for rank-1 transformations, *Ergodic Theory Dynam. Systems* **6** (1986), 363–384. MR 0863200. doi: 10.1017/S0143385700003552.

[15] A. Louveau and J. Saint-Raymond, Borel classes and closed games: Wadge-type and Hurewicz-type results, *Trans. Amer. Math. Soc.* **304** (1987), 431–467. MR 0911079. Zbl 0655.04001. doi: 10.2307/2000725.

[16] D. S. Ornstein, On the root problem in ergodic theory, in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, *Vol.* II: *Probability Theory* (Univ. California, Berkeley, Calif., 1970/1971), Univ. California Press, Berkeley, CA, 1972, pp. 347–356. MR 0399415. Zbl 0262.28009.

[17] ———, *Ergodic Theory, Randomness, and Dynamical Systems*, *Yale Math. Monogr.* **5**, Yale Univ. Press, New Haven, Conn., 1974, James K. Whittemore Lectures in Mathematics given at Yale University. MR 0447525. Zbl 0296.28016.

[18] D. J. Rudolph, *Fundamentals of Measurable Dynamics*, *Oxford Sci. Publ.*, The Clarendon Press Oxford University Press, New York, 1990, Ergodic theory on Lebesgue spaces. MR 1086631. Zbl 0718.28008.

[19] G. Sarton, *A History of Science. Hellenistic Science and Culture in the Last Three Centuries B.C.*, Harvard Univ. Press, Cambridge, MA, 1959. Zbl 0092.24201.

[20] M. Suslin, Sur une définition des ensembles B sans nombres transfinis, *Compte Rendu Acad Science* **164** (1917), 88–91.

[21] J. von Neumann, Zur Operatorenmethode in der klassischen Mechanik, *Ann. of Math.* **33** (1932), 587–642. MR 1503078. Zbl 0005.12203. doi: 10.2307/1968537.

[22] S. Wagon, *The Banach-Tarski Paradox*, *Encyclopedia Math. Appl.* **24**, Cambridge Univ. Press, Cambridge, 1985, with a foreword by Jan Mycielski. MR 0803509. Zbl 0569.43001.

[23] P. Walters, *An Introduction to Ergodic Theory*, *Grad. Texts in Math.* **79**, Springer-Verlag, New York, 1982. MR 0648108. Zbl 0475.28009.

(Received: September 30, 2008)
(Revised: November 24, 2009)

University of California, Irvine, Irvine, CA 92697
E-mail: mforeman@math.uci.edu

Colorado State University, Fort Collins, CO 80523

The Hebrew University of Jerusalem, 91904 Jerusalem, Israel
E-mail: weiss@math.huji.ac.il