Ergod. Th. & Dynam. Sys. (2004), **00**, 1–43
Printed in the United Kingdom

Anosov families, renormalization and nonstationary subshifts

Pierre Arnoux† and Albert M. Fisher‡

† *Pierre Arnoux, Institut de Mathématiques de Luminy (UPR 9016), 163 Avenue de Luminy, case 930 13288 Marseille Cedex 9, France(e-mail: arnoux@iml.univ-mrs.fr)*

‡ *Albert M. Fisher, Dept Mat IME-USP, Caixa Postal 66281, CEP 05315-970 São Paulo, Brazil (e-mail: afisher@ime.usp.br)*

(*Received ¡dates¿*)

Abstract. We introduce the notion of an Anosov family, a generalization of an Anosov map of a manifold. This is a sequence of diffeomorphisms along compact Riemannian manifolds such that the tangent bundles split into expanding and contracting subspaces.

We develop the general theory, studying sequences of maps up to a notion of isomorphism and with respect to an equivalence relation generated by two natural operations, gathering and dispersal.

Then we concentrate on linear Anosov families on the two-torus. We study in detail a basic class of examples, the multiplicative families, and a canonical dispersal of these, the additive families. These form a natural completion to the collection of all linear Anosov maps.

A renormalization procedure constructs a sequence of Markov partitions consisting of two rectangles for a given additive family. This codes the family by the nonstationary subshift of finite type determined by exactly the same sequence of matrices.

Any linear positive Anosov family on the torus has a dispersal which is an additive family. The additive coding then yields a combinatorial model for the linear family, by telescoping the additive Bratteli diagram. The resulting combinatorial space is then determined by the same sequence of nonnegative matrices, as a nonstationary edge shift. This generalizes and provides a new proof for theorems of Adler and Manning.

Contents

1. *Introduction*

When studying dynamics, one usually considers the iterates of a single transformation on a fixed space. Here we are interested in the dynamical behavior of a sequence of maps, along a sequence of spaces. Although at first glance the dynamics of a sequence of maps is trivial (since it is wandering), for the examples we will study here, one can find all of the richness of a recurrent dynamical system.

1.1. *Motivation: a completion for Anosov maps.* Sequences of maps arise naturally in a variety of ways; see Examples 1-6 in §1. The objects studied in this paper can be looked at from a corresponding variety of viewpoints.

We choose one of these as our primary motivating idea. We consider the set of all orientation-preserving linear Anosov diffeomorphisms on the two-torus, and ask this question: what would be a natural notion of completion for this collection of dynamical systems? For a first attempt at an answer, let us associate to a map A its pair of expanding and contracting foliations. These are numerically special, as the slopes of the foliations belong to the (dense) set of quadratic irrationals. Looking only at the foliations, it is natural to take as a completion, then, the collection of *all* pairs of transverse linear foliations.

For a notion of completion to be reasonable, the important dynamical structures associated with the individual maps should pass over to the new points, by continuity.

In fact, certain structures do extend naturally to this provisional completion: the unstable flows and their return (holonomy) maps, which in this case is the completion from circle rotations of quadratic type to all rotations.

Moreover, the usual Berg-Adler-Weiss construction of Markov partitions [**AW70**], [**Adl98**], (or at least its first step, the construction of a pair of parallelograms from eigenline segments) depends in no way on the special character of the Anosov slopes and so extends to this completion.

But what is iteration, where is the hyperbolic dynamics, when one begins just with a pair of foliations? Indeed, we say "first step" of the construction here because the second step is more problematic: making a generator (a partition which separates points under iteration) from this pair of boxes. Here we recall the process Adler and Weiss used in building a generator: one pulls back the two-box partition from the future via the map, and then takes connected components of the resulting intersections with the present partition. See Fig. 12.

Apparently this uses in an essential way the dynamics. But in fact, if we consider closely the geometry of the resulting partition, we can find a way to build this partition without any reference to the hyperbolic dynamics.

The key observation is that the new partition also can be specified directly from the pair of circle rotations, in terms of their joint renormalization; one is renormalized and the other un-renormalized. So at least some sense can be made of generation in this setting. And in this way, even without dynamics, one can define the "itinerary" of a point, associating to it a symbol sequence.

However the mystery remains: where is the dynamics? And how can that be seen as an extension of the dynamics of the individual maps to this completion?

What we do is change the initial set-up, and instead of representing each Anosov map A by its pair of foliations, we represent it by a biinfinite sequence of maps, repeating periodically, given by a factorization of A. To do this in a canonical way we use a (well-known) representation of orientation-preserving nonnegative (2×2) integer matrices as a product of two positive Dehn twists, see Lemma 3.11. Then, our completion will be all sequences of maps, given by those generators; and the dynamics takes us along a sequence of distinct copies of the torus.

It then turns out that the two-box Markov partition sequence is itself a generator for this sequence of maps, and moreover that the previous construction (pulling back and taking connected components) can be understood directly from that fact. See Lemma 5.4.

Having defined our candidate for the completion, a first question is whether the basic defining property of the Anosov maps (the hyperbolic splitting with contraction bounds) extends to this completion. The answer is yes, if we first remove a countable dense set. These correspond to the rational angles, which are also exactly those angles for which the renormalization of the rotations breaks down after a finite number of steps.

1.2. *Structure of the paper.* There are three main circles of results in this paper. The first, given in §2, involves general results on mapping and Anosov families. This provides an appropriate context and language for the rest of the paper.

In the second part, in §3, we introduce the symbolic dynamics for families, given by sequences of partitions. For the case of Anosov families, this specializes to Markov partition sequences and the corresponding symbolic spaces, given by a nonstationary sequence of $0-1$ matrices and geometrically represented by two-sided Bratteli diagrams.

In the third circle of results, presented in §§4 and 5 , we focus on a class of specific examples on the two-torus, which illustrate and give form to these abstract ideas.

These examples, the multiplicative and additive families, are related to the idea of a completion of the Anosov toral maps discussed above.

In the rest of the Introduction we summarize the main results.

1.3. *Anosov families.* To understand the new elements of the above completion, and what is the nature of dynamics for them, we introduce this notion:

Definition. An *Anosov family* is a (biinfinite) sequence of diffeomorphisms along a sequence of compact Riemannian manifolds, with an invariant sequence of splittings of the tangent bundle into expanding and contracting subspaces, and with a uniform upper bound for the contraction and lower bound for the expansion.

Similarly, an *eventually Anosov family* has the invariant splitting but now the contraction and expansion only are required to happen after an (unbounded) number of iterates.

Proposition. *An (eventually) Anosov splitting is unique if it exists.*

(See Proposition 2.2.)

It is important to know this, since with wandering dynamics, any splitting can be made invariant simply by transporting it along the sequence of spaces.

Here are our main examples.

Definition. Given a sequence $(n_i) = (\dots, n_{-1}, n_0, n_1 \dots)$ for $n_i \in \mathbb{N}^* = \{1, 2, \dots\}$, we define matrices $A_i \in SL(2,\mathbb{Z})$ by: $A_i = \begin{bmatrix} 1 & 0 \\ n_i & 1 \end{bmatrix}$ for i even, $A_i = \begin{bmatrix} 1 & n_i \\ 0 & 1 \end{bmatrix}$ for i odd. We let the A_i act on the torus; each choice of (n_i) defines a sequence of maps along a sequence of tori, which we shall call a *multiplicative family* because of its relation to the multiplicative continued fraction.

Our second basic example, the *additive family*, is given by factoring a multiplicative family; thus we take all biinfinite sequences of matrices of the form $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, the *additive generators*, which is nontrivial in the sense that the type of the matrix changes infinitely often at $\pm\infty$.

We show:

Theorem. *Each multiplicative family is an Anosov family; indeed the slopes of the contracting and expanding directions on the component M_i are related in a simple way to the continued fractions*

$$[n_i n_{i+1} \dots] = \cfrac{1}{n_i + \cfrac{1}{n_{i+1} + \cdots}}$$

and $[n_{i-1} n_{i-2} \dots]$ respectively (for the precise statement see Proposition 4.1.)

We have as a consequence:

Theorem. *Any nontrivial additive family is an eventually Anosov family; its eigendirections are given by the corresponding additive continued fractions.*

Our completion, the collection of multiplicative families, is parametrized by the product $\Pi_{-\infty}^{\infty} \mathbb{N}^*$, with the Anosov maps themselves corresponding exactly to periodic sequences (n_i).

1.4. *Anosov families as mapping families: morphisms, gatherings and dispersals.* The sequences of maps discussed above are part of a larger abstract context, which provides a clearer perspective on the Anosov and eventually Anosov families.

To describe this general setting, we consider sequences of maps as a category with respect to certain types of homomorphism, and also introduce a notion of equivalence generated by two natural operations: gathering and dispersal.

Definition. A *mapping family* is a sequence of continuous maps along a sequence of compact metric spaces, called *components*. A *uniform conjugacy* from one mapping family to another is given by an equicontinuous sequence of conjugating maps.

Remark. This notion of morphism makes the collection of all mapping families into a category. The reason for not considering the topological category- and using topological instead of metric spaces, and topological instead of uniform conjugacy-is that for mapping families *there is no topological dynamics*: every family is *topologically* conjugate to the trival family whose maps are all the identity, see Proposition 2.1.

For the uniform category, on the other hand, despite the fact that a mapping family has no recurrence (each component is a wandering set) the basic dynamical structures of stable and unstable sets make sense, as we have by Proposition 2.2 and Corollary 2.3:

Theorem. *Stable and unstable sets are preserved by uniform conjugacies hence are well-defined notions in the category of mapping families.*

Definition. For the category of smooth mapping families ($\mathcal{C}^1$ mapping families along a sequence of compact manifolds with Riemannian metrics) we define the morphisms to be *bounded conjugacies* (to have uniform bounds on the derivatives).

Thus bounded conjugacies preserve expansion bounds up to a constant. Hence we have (Proposition 2.14):

Theorem. *The Anosov and eventually Anosov families are categories with respect to bounded conjugacy.*

Definition. A *gathering* of a mapping family is a family given by taking of partial compositions along a subsequence; a *dispersal* is its converse, a factorization.

A basic example of dispersal is given by the insertion of extra copies of the identity map. For a second example, note that an additive family gathers to the corresponding multiplicative family and that conversely the additive family is a dispersal of the multiplicative family.

Remark. These are natural operations to consider for general mapping families but one must be careful: by contrast with uniform conjugacy, the operations of gathering and dispersal do not preserve the dynamical structures of stable and unstable sets; see Proposition 2.5.

However, as we show in Corollary 2.13, see also Remark 2.13:

Theorem. *Within the category of eventually Anosov families, stable and unstable sets are preserved by the operations of gathering and dispersal.*

1.5. *Coding the additive and multiplicative families: nonstationary subshifts.* A key tool for the further study of additive and multiplicative families is the extension of the notion of Markov partition to sequences of maps. The associated symbolic representation will be a combinatorially defined mapping family which we call a nonstationary subshift of finite type (*nsft*).

We begin in the setting of mapping families. One still has (as for single maps) a notion of coding the dynamics by the itineraries of a point, but now instead of this being described by where the orbit is located with respect to a single partition, this "name" of the point will now be given by a sequence of partitions along the sequence of spaces.

We generalize Markov partitions in the natural way; the new phenomenon is that in this setting, the symbolic space is now defined by a sequence of rectangular transition matrices (i.e. with entries 0 and 1), replacing the single square matrix which defines a subshift of finite type (*sft*).

The resulting space is a two-sided version of what Vershik and Livshits [**VL92**] call a *Markov compactum*.

As such, this space has no shift defined on it as the transitions keep changing. We introduce shift dynamics by taking the disjoint union of all compacta defined by shifts of the transition matrix sequence. This space is now a mapping family, whose components are simply all the shifts of the original Markov compactum.

Definition. This total space (the compactum together with all its shifts), is a nonstationary subshift of finite type (*nsft*).

This gives a second, combinatorial mapping family, which as is the case for a single map, factors onto the original family, provided the partition sequence generates (Proposition 3.6).

For our main examples, the additive and multiplicative families, we construct explicit Markov partitions. To do this, first we define a sequence of pairs of parallelograms (*boxes*) given by a simple algorithmic procedure related to continued fractions. For the additive family things are simplest. We show (Lemma 5.2):

Theorem. *Given a nontrivial additive family, the box pairs give a generating Markov partition sequence.*

For the multiplicative families, the two-box partition is Markov but does not generate. The same idea used by Adler and Weiss for a single map works here:

Theorem. *Given a multiplicative family, the partition sequence given by taking connected components of the join of each two-box partition with the pullback of the succeeding one gives a Markov partition sequence which generates.*

See Proposition 4.6. Now we apply the abstract machinery developed before (gatherings of mapping families) to see this construction in a new way, see Lemma 5.4:

Theorem. *The (generating) connected component partitions of a multiplicative family are alternatively given by gathering the two-box partition sequence of the corresponding additive family.*

Next we describe the nonstationary subshifts of finite type defined by these Markov partition sequences. First we have, for a nontrivial additive family (Theorem 5.3):

Theorem. *The mapping family is symbolically represented by the* nsft *given by exactly the same sequence of* $0-1$ *matrices, now interpreted as transition matrices.*

See Lemma 5.4.

The symbolic version of gathering corresponds to what is known in the theory of Bratteli diagrams as the *telescoping* of the diagram. The general theory, developed in §§3.1- 3.5, allows us to extend the previous result to multiplicative families, and further to the mapping family defined by a general sequence of (2×2) matrices in $SL(2,\mathbb{N})$, the semigroup of matrices with determinant 1 and with nonnegative integer entries.

The conclusion is (Theorems 5.1 and 5.6):

Theorem. *Given a mapping family defined by a nontrivial sequence of matrices in* $SL(2,\mathbb{N})$ *acting on the two-torus, then:*

(i) this is an Anosov family, and

(ii) it has a generating Markov partition sequence which codes it as the nsft *defined as an edge shift by exactly the same sequence of matrices.*

Restricting to the case of a single map gives a new proof of a theorem of Adler on codings of Anosov maps: that an orientation-preserving (2×2) Anosov matrix with nonnegative entries has a Markov partition for which it itself gives the edge shift space. See Theorem 5.3. Anthony Manning [**Man02**] proved a result similar to Adler's, but got the transpose matrix instead.

Remark. To get a statement like Adler's one needs to consistently choose the same convention for matrix actions, using either the action on rows or on columns both for the toral action and for defining the combinatorial spaces. Throughout this paper we use the column-vector convention. Mixing the conventions gives the transpose, as in Manning's version of the theorem. The reason for this becomes especially transparent when seen from the mapping family point of view; see Remark 5.3.

1.6. *Further generalizations.* There are several natural directions for extending this work and that of [**AF01**] and [**AF02a**]: to the orientation-reversing case, to higher genus surfaces, to higher dimensional tori, and to nonlinear Anosov maps. We plan to develop this material in a series of papers. In particular, we have linear, topological and smooth classification theorems of Anosov families on the two-torus, and a topological classification on the n-torus. As a corollary we can solve a question asked by Kifer (Conjecture A1 of [**Kif00**]). We mentioned above that there is no shift dynamics on a single component of an *nsft*; however one can define there a transversal dynamics, given by Vershik's adic transformations; see [**Ver94**]. Equivalently the tranversal dynamics can be defined by nonstationary substitution systems. We do not discuss these fascinating topics in the present paper; see [**AF02b**], [**AF02a**], [**AF01**]. In work of A. F. with M. Urbanski, we prove the existence of a Markov partitions for an Anosov family, generalizing the construction of [**Bow75**], [**Bow77**]. The main case studied in the present paper is simpler so we can we construct the partitions explicitly, see §4.

Another natural generalization is from discrete to continuous time: from mapping families to what could be called *flow families.* One example is the suspension flow of a mapping family. The suspension of a multiplicative family is an interesting object for an entirely different reason: it models the *scenery flow* of the transverse irrational circle rotation. See [**AF02a**], [**AF01**]. A further potentially interesting class of examples to consider are the flow families given by nonautonomous differential equations, where the orbits are integral curves of time-varying vector fields.

1.7. *Acknowledgements.* This work draws from many different parts of dynamical systems theory, and so builds on research of many people, too numerous to mention here. The project originally came out of an attempt to put together the ideas of [**Fis92**] on the scenery flow for a fractal set and [**Arn94**] on the Teichmüller flow, and in particular owes much to the work of Veech. The connections of these flows with the present work is made in [**AF02a**]; see also [**Fis04**], [**Fis03**].

The name "families" we borrowed from David Rand's *Markov families* of expanding maps on the interval [**Ran88**], see Remark 2.17 below. We mention that V.I. Bakhtin also has considered non-random hyperbolic sequences of mappings [**Bak95a**], [**Bak95b**], studying nonlinear theory; we thank Y. Kifer for this citation.

The second-named author would like to thank the following institutions and agencies for their support and hospitality during the course of the writing: IHES; the Graduate Center at CUNY; the CNRS and Univ. Paris XIII; the IMS and SUNY Stony Brook; the CNPQ and FAPERGS at UFRGS Porto Alegre, Brazil; IMPA; the CNPQ at the University of São Paulo; the CNRS and the University of Marseilles, Luminy.

We would like to thank all the many colleagues who have inspired and aided us along the way. Special thanks for related conversations and for their interest, inspiration and encouragement goes to: Elise Cawley, Matthias Gundlach, Pascal Hubert, Jeremy Kahn, Rick Kenyon, Curt McMullen, Yair Minsky, Alberto Pinto, Tom Schmidt, and Dennis Sullivan.

2. *Mapping families*

2.1. *Sequences of maps; conjugacy and equivalence.* We begin with a discussion of sequences of maps in several categories, considering topological, uniform and bounded differentiable conjugacies as providing the morphisms. Then we introduce two operations, gathering and dispersal. Lastly we study the interplay between change of metric and dynamics, for which purpose we consider isometric conjugacy.

A key conclusion will be that for families, the spaces should have a uniform structure, with the morphisms preserving that uniformity. For simplicity, we work in the category of metric spaces with uniformly continuous maps. We show that furthermore, gathering and dispersal are not well behaved operations in complete generality, but do make sense in the presence of hyperbolicity. Finally, we see that any dynamics can be pushed entirely to the metrics, and that a converse is also true, provided the minimum necessary requirement of having isometric spaces is satisfied. This abstract framework will be illustrated in the next section by concrete examples.

Given a sequence of sets M_i for $i \in \mathbb{Z}$, the **disjoint union** written $M = \coprod M_i$ is the *coproduct* in the category of sets, see e.g. [**Hun74**]. We recall that this is simply the indexed union; that is, a point in M is a point p in some M_i together with the index i. For convenience we will suppress this index and write p for this point; thus, any $p \in M$ belongs to exactly one M_i. We refer to the M_i as **components** of M. If the components are topological spaces, then we topologize M by putting these spaces together discretely. By this we mean the topology is generated by the union of all those topologies, so in particular each M_i itself is open and closed in M. (Warning: a component M_i may itself have more than one *topological* component; an example is given by the nonstationary shift space of §3.1).

Definition 2.1. Let M_i for $i \in \mathbb{Z}$ be a sequence of metric spaces with metrics ρ_i. Assume for simplicity that the diameter of each space is less than or equal to 2. We give the total space the metric $\rho(x, w) = \rho_i(x, w)$ if $x, w \in M_i$, $= 1$ if they are in different components (the bound of 2 ensures that one has the triangle equality). We assume $f_i : M_i \to M_{i+1}$ are continuous functions. We define the **total map** $f : M \to M$ on the disjoint union to be equal to f_i on each component M_i. The n^{th} composition f^n maps M_i to M_{i+n} and is equal to $f_{i+n} \circ \cdots \circ f_i$ on M_i for each i. We will call the resulting pair (M, f) a **mapping family**. We say (M, f) is **invertible** if all the maps are homeomorphisms.

For the simplest case, all the maps and spaces are identical:

Definition 2.2. Given a homeomorphism f_a of a metric space M_a, we define the **constant family** associated to f_a (M, f) to be the following family of maps: $M = \coprod M_i$ where each M_i is a copy of M_a (with the same metric); $f_i : M_i \to M_{i+1}$ is equal to f_a modulo this identification.

We will say that the mapping family (M, f) is a **lift** of the dynamical system (M_a, f_a). This is the simplest example of a mapping family; a particularly trivial case is the **identity family** on M, that is, the constant mapping family which is a lift of the identity map on M.

We next examine what should be the morphisms for this collection of objects, in order to make it into a category. We begin with the most obvious, but in fact quite wrong notion:

Definition 2.3. Given two mapping families (M, f) and (N, g), with metrics d and ρ, a **topological conjugacy** is a homeomorphism h from M to N which conjugates f and g, i.e. such that $h \circ f = g \circ h$; if h is a continuous map but not a homeomorphism, we say this is a topological semiconjugacy.

The problem with this definition is that the category then becomes trivial, up to isomorphism:

Proposition 2.1. *Any invertible mapping family (M, f) is topologically conjugate to the identity family on M_0.*

Proof Let (N, g) be the identity family on M_0. We define $h : M \to N$ by $h_0 =$ the identity; for $n > 0$, $h_n = (f_{n-1} \circ \cdots \circ f_0)^{-1}$; and for $n < 0$, $h_n = f_{-1} \circ \cdots \circ f_n$; we check that this conjugates (M, f) to (N, g). □

Thus even the lift of a single map is trivial up to topological conjugacy. Our choice of morphism will be the following, which works better, fortunately:

Definition 2.4. A **uniform (semi)conjugacy** is a (semi)conjugacy which in addition is a uniformly continuous map (on the total space); equivalently, the sequence of conjugating maps $\{h_i\}$ is uniformly equicontinuous on $\{M_i\}$.

Definition 2.5. Let M be a metric space and $f : M \to M$ a homeomorphism. The **stable set** $W^s(x)$ for $x \in M$ is $\{y : \text{ dist}(f^n x, \ f^n y) \to 0 \text{ as } n \to \infty\}$. The **unstable set** $W^u(x)$ is the stable set of f^{-1}.

Given a mapping family (M, f), we apply this definition to the total map and we have:

Proposition 2.2. *Stable and unstable sets are preserved by uniform semiconjugacy, more precisely we have* $h(W^s(x)) \subseteq W^s(h(s))$*; for a uniform conjugacy there is equality.*

Proof Immediate from the equicontinuity. □

Hence:

Corollary 2.3. *Stable and unstable sets are well-defined notions in the category of mapping families.* □

Remark 2.1. From the definition, the unstable set in the k^{th} component only depends on the past of the sequence of maps, i.e. on f_i for $i < k$, while the stable set only depends upon the future $i \geq k$; see Example 8(i) and Proposition 2.16.

We next define:

Definition 2.6. Given a mapping family (M, f), a second mapping family $(\widetilde{M}, \tilde{f})$ is a **gathering** of (M, f) if there exists a strictly increasing biinfinite subsequence (n_i) of the integers $\mathbb{Z}$ such that $\widetilde{M}_i = M_{n_i}$ and $\tilde{f}_i = f_{n_{i+1}-1} \circ \cdots \circ f_{n_{i+1}} \circ f_{n_i}$.

If a family $(\widetilde{M}, \tilde{f})$ is a gathering of (M, f), we say that (M, f) is a **dispersal** of $(\widetilde{M}, \tilde{f})$. That is, dispersal is the converse procedure of gathering.

Remark 2.2. One can think of a dispersal as given by inserting extra spaces and maps along the way so that the new family gathers to the original one. The simplest way to do this is to insert copies of the identity map on a component. For a single map, one can think of these as "fillers" coming from adding (full) levels in a tower, in ergodic theory parlance; if the map originates as a flow cross-section, this is like inserting extra cross-section levels.

With this simplest type of dispersal we have introduced a sequence of time delays, without changing the "dynamics" in any essential way.

As we shall soon see, however, things are not quite so simple for general dispersals. But first we consider how a mapping family (M, f) is related to the shifted family (N, g) defined by $N_i = M_{i+1}$, $g_i = f_{i+1}$. Surely the dynamics here has not changed in a major way, so our operations should reflect this:

Proposition 2.4. *The shifted family (N, g) is a gathering of (M, f), by taking $n_k = k + 1$; it is uniformly conjugate to (M, f) via the conjugacy $h_i = f_i$ if and only if the collection $\{f_i\}_{i\in\mathbb{Z}}$ is equicontinuous.*

Proof For both statements the proof is immediate from the definitions. The second can be understood by the following commutative diagram:

$$\begin{array}{ccccccc} M_0 & \xrightarrow{f_0} & M_1 & \xrightarrow{f_1} & M_2 & \xrightarrow{f_2} & M_3 \\ \cdots\cdots \big\downarrow f_0 & & \big\downarrow f_1 & & \big\downarrow f_2 & & \big\downarrow f_3 \cdots\cdots \\ M_1 & \xrightarrow{f_1} & M_2 & \xrightarrow{f_2} & M_3 & \xrightarrow{f_3} & M_4 \end{array}$$

□

Remark 2.3. Now we discuss some trouble one can get into using this idea. A first problem is that the notion of stable sets, which seems to be basic to the study of mapping families, is not preserved by gathering; the stable set of a point in the gathered family is contained in that of the original family, but may have strictly increased: simply insert maps which move the points a definite distance apart again.

A related problem is shown in a strong form in the next proposition. Here we see that in the category of mapping families, the equivalence relation generated by these two operations is trivial in that there is only one equivalence class. The situation could thus appear hopeless; but as we shall see below, this difficulty will be resolved once we introduce hyperbolicity. Indeed, *within* the class of Anosov families gathering is well behaved in that it does preserve stable sets, see Corollary 2.13.

Proposition 2.5. *Any invertible mapping family (M, f) has a dispersal, which has a gathering, which is equal to the identity family on M_0.*

Proof This is a corollary of Proposition 2.1; indeed, let (h_i) be the sequence of conjugating maps given there; then for $g_i =$ the identity, $g_i = h_{i+1} \circ f_i \circ h_i^{-1}$. The dispersal is the sequence of maps $\ldots, h_0^{-1}, f_0, h_1, h_1^{-1}, f_1, h_2, h_2^{-1}, \ldots$; associating these as $(h_1 \circ f_0 \circ h_0^{-1})$ gives the gathering equal to g_i. □

We have seen above the simplest example, the constant family, where the metrics do not change. But in general the dynamics of a mapping family is determined by the interplay of the sequence of maps, and of metrics, both of which may be "changing". A family may be purely of one or the other type; thus one has the case where the spaces are all isometric and all the dynamics is carried by the map, and the opposite situation, where each map is the identity, and hence all the change is carried by the metrics.

In fact these two extremes are quite different, as one is general while the other is limited, as the next proposition shows.

Proposition 2.6. *Given an invertible mapping family (M, f), there exists an isometrically isomorphic family (N, g), with all maps g_i being the identity map on a single topological space $N_i = N_0$, and with changing metrics ρ_i.*

Given a mapping family (N, g) with all maps g_i the identity map on $N_i = N_0$, and with metrics ρ_i on N_i, then there exists an isometrically isomorphic family (M, f) such that each M_i is the same metric space, if and only if the components N_i with metric ρ_i are all isometric.

Proof Thus, let (M, f) be a mapping family with f_i homeomorphisms, along a sequence of metric spaces (M_i, ρ_i). Now we simply pull back the metrics to the space M_0. More precisely, define a second mapping family (N, g) as follows: $N_0 = M_0$ with metric $\hat{\rho}_0 = \rho_0$; g_i is the identity map for all $i \in \mathbb{Z}$; $h_i : M_i \to N_i$ is $(f_n \circ \cdots \circ f_1 \circ f_0)^{-1}$; $\hat{\rho}_i$ is the image of the metric ρ_i under h_i. The two families are isometrically isomorphic via h, and in the new family the dynamics has been moved entirely to the metrics.

For the second part, let us assume that we are given a family (N, g) such that all maps g_i are the identity on a topological space $N_0 = N_i$, with N_i given a metric $\hat{\rho}_i$.

Now if there is a second family (M, f) such that the spaces M_i are identical copies of a single metric space (M_0, ρ_0), and if this is isometrically isomorphic to the family (N, g), then certainly the components N_i are all isometric, so that is a necessary condition. We now show this condition is sufficient. So we assume we are given isometries g_i from N_i to N_{i+1}. Then we define (M, f) by $M_i = N_0$ with metric $\rho_0 = \hat{\rho}_0$, for all i. We define f_i from M_i to M_{i+1} by $f_i = g_i^{-1}$. The conjugating maps are as indicated in the commutative diagram.

$$\begin{array}{ccccccccc}
\cdots & N_{-1} & \xrightarrow{I} & N_0 & \xrightarrow{I} & N_1 & \xrightarrow{I} & N_2 & \cdots \\
 & \downarrow{\scriptstyle g_{-1}^{-1}} & & \downarrow{\scriptstyle I} & & \downarrow{\scriptstyle g_0^{-1}} & & \downarrow{\scriptstyle (g_1)^{-1} \circ g_0^{-1}} & \\
\cdots & M_{-1} & \xrightarrow{f_{-1}} & M_0 & \xrightarrow{f_0} & M_1 & \xrightarrow{f_1} & M_2 & \cdots
\end{array}$$

□

Remark 2.4. In Example 4 we will see an Anosov family defined by changing the metrics; a specific case of the above relationships is studied in Proposition 4.4.

2.2. *Hyperbolicity.* Now we move to the smooth category, where our primary interest is a natural generalization of Anosov diffeomorphisms. Some good references regarding the case of a single map [**Bow75**], [**Shu87**].

Definition 2.7. An **Anosov family** is a mapping family (M, f) such that:

- (i) the components M_i for $i \in \mathbb{Z}$ are a sequence of Riemannian manifolds (i.e. compact $\mathcal{C}^\infty$ manifolds with fixed Riemannian metrics) and the maps $f_i : M_i \to M_{i+1}$ are $\mathcal{C}^1$ diffeomorphisms,
- (ii) the tangent bundle TM has a continuous splitting $E^s \oplus E^u$ which is f-invariant, and
- (iii) there exist constants $\lambda > 1$ and $c > 0$ such that for each $n \geq 1$, for each i, for every point $p \in M_i$ one has:

  $$||D(f_i^{-n})(v)|| \leq c^{-1}\lambda^{-n}||v||$$

  for every vector $v \in E^u_p$, and

  $$||D(f_i^{n})(v)|| \leq c^{-1}\lambda^{-n}||v||$$

  for every $v \in E^s_p$.

  (Here $E_p = E^s_p \oplus E^u_p$ is the tangent space based at p.)

Note that without loss of generality, $c \leq 1$ (otherwise it can be replaced by 1); if we can take $c = 1$, then we say the family is **strictly Anosov**.

Remark 2.5. An interesting difference between mapping families and single maps is that for families there are always *many* invariant continuous splittings, while e.g. for an Anosov map there is essentially one, because of the density of periodic points (though there is some flexibility as eigenspaces can be combined in different ways). However there is a unique *hyperbolic* splitting: see Proposition 2.12 and Remark 2.12.

We next see that condition (iii) in the above definition actually gives us twice as much for free:

Lemma 2.7. *The statement*

$$||D(f_i^{-n})(v)|| \leq c^{-1}\lambda^{-n}||v||$$

for each $n \geq 1$, for each i, for every vector $v \in E^u_p$, is equivalent to:

$$||D(f_i^{n})(v)|| \geq c\lambda^{n}||v||$$

for each $n \geq 1$, for each i, for each $v \in E^u_p$. The same is true for the condition on the stable subspace.

Proof Write $w = D(f_i^{-n})(v)$. Then $||w|| = ||D(f^{-n_i})(v)|| \leq c^{-1}\lambda^{-n}||v|| = c^{-1}\lambda^{-n}||D(f^{-n_i}))(w)||$ so

$$||D(f^{-n_i}))(w)|| \geq c\lambda^{n}||w||.$$

The derivative maps give isomorphisms of the unstable subspaces, so any element $w \in E^u_q$ occurs here, where $q = f^{n_i}(p)$. □

Remark 2.6. For this reason the unstable and stable subspaces can also be called the *expanding* and *contracting* subspaces; $\lambda > 1$ gives a lower bound for the expanding constant.

From the Lemma we have:

Corollary 2.8. *Condition (iii) above can be replaced by:*

- *(iii′) for each $n \geq 1$, for each i,*
$$||(D(f_i^{-n}))^u|| \leq c^{-1}\lambda^{-n},\ ||D(f_i^n)^s|| \leq c^{-1}\lambda^{-n}$$
where $(D(f_i^{-1}))^u, D(f_i)^s$ are the restrictions of those linear maps to the unstable and stable subspaces;

  by

- *(iii″) for each $n \geq 1$, for each i,*
$$||D(f_i^n)(v)|| \geq c\lambda^n ||v|| \quad \textit{for every } v \in E_p^u,$$
and
$$||D(f_i^n)(v)|| \leq c^{-1}\lambda^{-n}||v|| \textit{ for } v \in E_p^s;$$
or by:

- *(iii‴)*
$$||D(f_i^{-n})(v)|| \leq c^{-1}\lambda^{-n}||v|| \quad \textit{and } ||D(f_i^n)(v)|| \geq c\lambda^n||v||$$
for every vector $v \in E_p^u$,
$$||D(f_i^n)(v)|| \leq c^{-1}\lambda^{-n}||v|| \textit{ and } ||D(f_i^{-n})(v)|| \geq c\lambda^n||v||$$
for every $v \in E_p^s$.

□

We shall need:

Lemma 2.9. *Let A be an invertible linear map between inner product spaces. Let c_1, c_2 be the minimum and maximum radii of the ellipsoid which is the image of the unit ball, so that $c_1 = \inf\{||Av||/||v||\}$ and $c_2 = \sup\{||Av||/||v||\}$. Then $c_2 = ||A||$ and $1/c_1 = ||A^{-1}||$.* □

Remark 2.7. Conditions (iii) and (iii') only speak of contraction of the unstable and stable subspaces, for past and future times respectively; as this is an upper bound, it can be conveniently phrased in terms of the norm of the matrix, as in (iii'). Condition (iii'') only refers to future time, saying that the unstable space expands and the stable space contracts with uniform lower and upper bounds respectively, as time goes to $+\infty$. (By the Lemma, this lower bound for expansion can be expressed using the reciprocal of the norm of the inverse matrix, which would return us to condition (iii').) Condition (iii''') gives the complete picture, saying that the unstable space expands with a uniform lower bound as time goes to $+\infty$ and contracts with a uniform upper bound as time goes to $-\infty$, with the reverse for the stable space.

We note that conditions (iii), (iii'), and (iii''') are symetric with respect to time reversal. We now formalize this idea for families:

Definition 2.8. The **inverse family** of an invertible mapping family (M, f) is the mapping family (N, g) with $N_i = M_{-i}$ and $g_i = (f_{-i-1})^{-1}$.

Remark 2.8. This is not a composition inverse as there is no notion of composition of families; rather, the inverse family is the family of made up of inverse maps. There is however a duality, as the inverse family of the inverse family is the original family.

Proposition 2.10.

(a) The family (M, f) is Anosov iff the inverse family is.

(b) For a strictly Anosov family it is enough to state condition (iii) for $n = 1$, i.e. to know that

$$||(D(f_i^{-1}))(v)|| \leq \lambda^{-1}||v|| \text{ for } v \in E^u, \text{ and } ||D(f_i)(v)|| \leq \lambda^{-1}||v|| \text{ for } v \in E^s.$$

Equivalently, for the strict Anosov case we can replace (iii) by:

(iii''')

$$||(D(f_i^{-1}))^u|| \leq \lambda^{-1}, \ ||D(f_i)^s|| \leq \lambda^{-1}.$$

(c) Given an Anosov family, there exists a C^1-uniformly bounded change of metrics on the components which makes the family strictly Anosov.

Proof

Part (a) follows since condition (iii) of Corollary 2.8 is symmetric with respect to time reversal, switching the unstable and stable spaces for the inverse family. Part (b) is immediate. Part (c) is a version for families of a well-known lemma of Mather; the proof for the case of a single map of [**Shu87**] goes through for families. □

See also Proposition 2.14.

A more general context is this:

Definition 2.9. Given a family (M_i, f_i) of diffeomorphisms of (not necessarily compact) Riemannian manifolds, suppose there is an invariant set $\Lambda = (\Lambda_i)$ for the total map f on $M = \coprod M_i$, such that TM_Λ has an invariant splitting, and such that constants $\lambda > 1$, $c > 0$ exist as above for all $p \in M$. We then call (M, f, Λ) a **hyperbolic family**.

Remark 2.9. So by definition, an Anosov family is a hyperbolic family for which $\Lambda = M$ and each M_i is compact.

Example 1. The simplest example is the constant family (M, f) defined as the lift of an Anosov map f_a of a Riemannian manifold M_a where the manifolds M_i and maps f_i are identical, but distinct, copies of M_a and f_a.

Example 2. A more interesting example is given by random perturbations of this. For $\alpha > 0$, let Ω_α be an α-neighborhood of f_a in the $\mathcal{C}^{1+1}$-norm (sometimes called "1 plus Lipschitz", thus derivatives are Lipschitz close; this Banach space strictly contains $\mathcal{C}^2$). Define $g_i : M_i \to M_{i+1}$ to be an arbitrary sequence chosen from Ω_α. Then Proposition 2.2 of L.S. Young [**You86**] says exactly the following, in our language: for α sufficiently small, for all such choices of g_i, (M, g) is an Anosov family.

Remark 2.10. Note that (M, g) is also a small perturbation of the constant family (M, f). However as families these are not actually so different in their dynamics: as we show in a forthcoming paper, one has structural stability for Anosov families, which in this case says exactly that the constant and perturbed families are boundedly conjugate, see Definition 2.12 below. (This strengthening of Young's result gives good additional evidence for the naturalness of that choice for the morphisms.)

Example 3. Examples of Anosov families which are also nontrivial topologically can be built as follows; this is the main class of examples we will be interested in for this paper. Let (n_i) for $i \in \mathbb{Z}$ be a sequence of integers ≥ 1, and let, for each i, M_i be distinct copies of the torus $\mathbb{R}^2/\mathbb{Z}^2$ with the Riemannian metric inherited from the plane. Define $f_i : M_i \to M_{i+1}$ to be the map given by the matrix A_i multiplying column vectors on the left, where

$$A_i = \begin{bmatrix} 1 & 0 \\ n_i & 1 \end{bmatrix} \quad \text{for } i \text{ even,}$$

$$A_i = \begin{bmatrix} 1 & n_i \\ 0 & 1 \end{bmatrix} \quad \text{for } i \text{ odd.}$$

Then (M, f) is an Anosov family (see Proposition 4.1). We call this particular mapping family on the torus the **multiplicative family** determined by the sequence (n_i). See §4. There, we will find explicitly the eigenspaces E_i^s and E_i^u, and the sequence of eigenvalues; see Proposition 4.1 below.

Example 4. Let M_0 be a smooth manifold, with a continuous splitting of TM_0 denoted $E_0^s \oplus E_0^u$. Define M_i to be identical, but distinct, copies of M_0, and let $f_i : M_i \to M_{i+1}$ be the identity map (modulo this identification) for each i. Let ρ_0 be a Riemannian metric on M_0; for simplicity one might choose this so E_0^s and E_0^u are everywhere orthogonal. For $n \neq 0$ in $\mathbb{Z}$ define ρ_n to expand and contract ρ_0 exponentially along those subspaces and to extend to the (unique) inner product on the tangent space; thus for some chosen $\lambda > 1$ set $\rho_n = \lambda^{-n}\rho_0$ on E^s and $\rho_n = \lambda^n \rho_0$ on E^u. Then (M, f) is an Anosov family. See Proposition 4.4 for a concrete case of this.

This example generalizes naturally to pseudo-Anosov families, (extending Thurston's notion of pseudo-Anosov map), by now permitting a finite number of fixed singular points; see also [**AF01**].

Example 5. Let $\gamma(t)$ for $t \in \mathbb{R}$ be a geodesic in the hyperbolic disk, parameterized by hyperbolic length. As is well known, the disk (or equivalently the upper half plane $\mathbb{H}$) is naturally identified with the Teichmüller space of the torus $\mathbb{T}^2$. Let $\dots t_0, t_1, \dots t_i \dots$ be a biinfinite sequence of reals tending toward $\pm\infty$ as $i \to \pm\infty$, such that $t_{i+1} - t_i$ is bounded below. Let M_i be the torus with the flat metric determined by the conformal structure of $\gamma(t_i)$, and let $f_i : M_i \to M_{i+1}$ be the corresponding Teichmüller maps. Then (M, f) is an Anosov family, providing in addition that the endpoints of the geodesic are irrational numbers in the upper halfspace model. This is related to the previous example; see [**AF02a**], [**AF01**].

We will also want the following more general notion.

Definition 2.10. A mapping family is **eventually Anosov** if there exists an f-invariant splitting, as before, but now with sequences $\lambda_i^u, \lambda_i^s > 0$ for $i \in \mathbb{Z}$ defined by

$$\lambda_i^u \equiv \inf\{||D(f)(v)||/||v|| \text{such that } v \in E^u\},\ \lambda_i^s \equiv \sup\{||D(f)(v)||/||v|| \text{such that } v \in E^s\},$$

or equivalently (by Lemma 2.9)

$$1/\lambda_i^u = ||(D(f)^u)^{-1}||,\ \lambda_i^s = ||D(f)^s||,$$

where as before $D(f)^u, D(f)^s$ denote the linear maps restricted to these subspaces, with these sequences satisfying, for some (and hence for all) $k \in \mathbb{Z}$,

$$\Pi_k^{k+n} \lambda_i^u \to +\infty \text{ and } \Pi_{k-n}^k (\lambda_i^u)^{-1} \to 0 \text{ as } n \to +\infty, \tag{1}$$

and

$$\Pi_k^{k+n} \lambda_i^s \to 0 \text{ and } \Pi_{k-n}^k (\lambda_i^s)^{-1} \to +\infty \text{ as } n \to +\infty.$$

Remark 2.11. The requirement says that each vector in E^u be eventually expanded at $+\infty$ and eventually contracted at $-\infty$, with the reverse for the stable eigenspaces; this is similar to version (iii''') of Definition 2.7. However we emphasize that *there is no analogue of version* (iii) *or* (iii') *for the eventually Anosov case;* see Example 9 below for a counterexample.

We note that, just as in Proposition 2.10, the inverse family of an eventually Anosov family is eventually Anosov. We mention that the above conditions can be more succinctly (but perhaps less clearly) be stated as:

$$\Pi_k^{k+n} \lambda_i^u \to +\infty$$

and

$$\Pi_{k-n}^k (\lambda_i^s)^{-1} \to +\infty$$

as $n \to \pm\infty$.

We observe that:

Proposition 2.11. *An Anosov family is eventually Anosov. An eventually Anosov family has a gathering which is strictly Anosov. An Anosov family has a* constant length *gathering which is strictly Anosov.* □

Here is an example:

Example 6.

Definition 2.11. Given a sequence (n_i) for $i \in \mathbb{Z}$ of positive integers, we let, for each $k \in \mathbb{Z}$, M_k to be the torus $\mathbb{R}^2/\mathbb{Z}^2$ with the metric inherited from the plane, and define $f_k : M_k \to M_{k+1}$ to be the map given by the elementary matrix A_k acting on column vectors, with

$$A_k = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

for k satisfying $n_i \leq k \leq n_{i+1} - 1$ if i is even,

$$A_k = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

for k as above when i is odd. We will call this the **additive family** determined by the sequence (n_i) because of its connection with what is called the *additive continued fraction*: expressing n_k as a sum of $1's$ as we have done here.

Then (M, f) is an eventually Anosov family, since a subsequence of its partial products is exactly that from Example 3. Thus, the multiplicative family of Example 3 is a gathering of Example 6.

The following proposition shows that the concept of eventually Anosov families has meaning.

Proposition 2.12. *Given an eventually Anosov family (M, f), the splitting $E^s \oplus E^u$ of the tangent bundle TM is unique.*

Proof Choose one component, say M_0; the splitting is invariant hence is determined by the splitting there. Suppose there is a second hyperbolic splitting $\widetilde{E}^s \oplus \widetilde{E}^u$; let $\tilde{v}_0^s$ be a vector in $\widetilde{E}^s$, based at some point p, which is not in E^s at p. Now this can be expressed as a sum in the first splitting, $\tilde{v}_0^s = av_0^u + bv_0^s$ with $a \neq 0$. Applying the maps f_i, this vector expands at $+\infty$, since it has a non-zero unstable component. However from part (iii) of the definition, the stable space is required to contract as time goes to $+\infty$, giving a contradiction. Hence we must have $\widetilde{E}^s \subset E^s$. By the symmetric argument, $E^s \subset \widetilde{E}^s$. The same proof applied to the inverse family shows that $\widetilde{E}^u = E^u$, and hence the splitting is unique. □

Remark 2.12. For an Anosov family (as contrasted to the case of a single map) the are many *invariant* splittings: simply choose a splitting at time 0 and transport it forward and backward to the other components. The Proposition shows that any other such invariant splitting cannot be hyperbolic.

Corollary 2.13. *A gathering of an eventually Anosov family is again eventually Anosov. Within the class of eventually Anosov families, gathering preserves the splitting of the tangent bundle, and also preserves stable and unstable sets.*

Proof Let us suppose we are given an eventually Anosov family (M, f) and family (N_n, g) which is the gathering along the components $N_n = M_{i_n}$, for some strictly increasing subsequence of the integers $i_n : n \in \mathbb{Z}$. The statement means that on each of these components the splittings are the same. Now it is clear a fortiori that the inherited splitting fits the definition of an eventually Anosov splitting. Hence (N_n, g) is an eventually Anosov family. By Proposition 2.2 this splitting is unique; thus the splittings are preserved by the gathering. By the stable manifold theorem, the proof of which we will give elsewhere, the stable set is determined by (and determines) the splitting, as its tangent space, hence this passes to the stable (and unstable) sets. □

Remark 2.13. It follows that given an eventually Anosov family (M, f), if a second family (N, g) is a dispersal of this *and is eventually Anosov*, then the stable and unstable sets are preserved. What went wrong in Proposition 2.5 is that there we had a dispersal which took us out of the eventually Anosov category. Equivalently: the converse to Proposition 2.11 is false; that is, there exist mapping families for which there is an Anosov gathering but which themselves are not eventually Anosov.

A further class of examples come from random dynamical systems; for related work see also [**Bog93**], [**BG95**]:

Example 7. Consider a skew product transformation on $X \times \mathbb{T}^d$, with base map $T : X \to X$ invertible and preserving a probability measure, and with fiber the d-dimensional torus. Suppose the skewing function A takes values in $SL(d, \mathbb{Z})$, with the matrix $A(x)$ acting on the fibers by multiplying a column vector y on the left, so the skew product is the map $(x, y) \mapsto (T(x), (A(x)) \cdot y)$. Assume that $\log ||A||, \log ||A^{-1}||$ are integrable functions, and that the Lyapunov exponents are nonzero for a.e. choice of $x \in X$. Then the matrix product given by following the orbit of x defines, for almost every x, an eventually Anosov family on the d-torus. This is immediate from Osceledec' multiplicative ergodic theorem. Example 3, the multiplicative family, can be looked at from this point of view, taking for the base the natural extension of the continued fraction transformation with, say, the extension of Gauss measure. See §3.3. However our basic perspective here is different from this, as we are interested in *all* the multiplicative families, and not just a measure-one subset; also from our point of view it is important to study them individually as well as when taken all together in the skew product. See also §5.4.

Definition 2.12. Two Riemannian metrics on a manifold are **boundedly equivalent** if the ratios of the norms they induce on the tangent space at each point are bounded away from 0 and ∞ by constants. Two smooth mapping families are **boundedly conjugate** if there exists a differentiable conjugating map which induces a bounded equivalence of the metrics. From Lemma 2.9, a conjugacy h is bounded iff there exists $c > 0$ such that $||Dh_i|| < c$ and $||D(h_i)^{-1}|| < c$ for all i.

We note that having a bounded conjugacy is equivalent to having a smooth and uniformly Lipschitz conjugacy.

Proposition 2.14. *Let (M, f) be an eventually Anosov family. Let (N, g) be another family of diffeomorphisms of Riemannian manifolds (also with a definite Riemannian metric), and let $h : M \to N$ be a diffeomorphism which conjugates f and g. Assume that h is a bounded conjugacy. Then (N, g) is an eventually Anosov family. If (M, f) is an Anosov family, then so is (N, g), with the same expansive constant λ.*

Proof We are given that there exists a constant $c > 0$ such that $||Dh_i^{-1}||, ||Dh_i|| < c$ for each $i \in \mathbb{Z}$. We push forward the splitting $E^u \oplus E^s$ of TM to get a splitting $\widetilde{E}^u \oplus \widetilde{E}^s$ of TN; this is also invariant. Let sequences $\lambda_i^u, \lambda_i^s > 0$ and $\tilde{\lambda}_i^u, \tilde{\lambda}_i^s > 0$ be defined for the families (M, f) and (N, g) as in Definition 2.10. We need to verify condition (1) for the $\tilde{\lambda}_i$. We have from the Chain Rule that $\tilde{\lambda}_i^s = ||Dg^s|| \leq ||Dh||_{f\circ h^{-1}(p)} \cdot ||Df^s||_{h^{-1}(p)} \cdot ||Dh^{-1}||_p \leq c^2\lambda_i^s$. Hence for each i, $\tilde{\lambda}_i^s \leq c^2\lambda_i^s$, and similarly, $c^{-2}\lambda_i^u \leq \tilde{\lambda}_i^u$. But this constant c^2 remains the same if instead we consider partial compositions of the maps. Therefore condition (1) holds, so (N, g) is also eventually Anosov. The case of an Anosov family is similar. □

Remark 2.14. Proposition 2.10 can now be restated to say: given an Anosov family (M, f), there exists (N, g) boundedly conjugate to the such that (N, g) is strictly Anosov.

We now examine the difference between mapping and Anosov families. Consider the following mapping family:

Example 8.

(i) Let M_k be the torus and let $f_k : M_k \to M_{k+1}$ be defined by these matrices acting on column vectors:

$$A_k = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \text{ for } k \geq 0, \; A_k = \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix} \text{ for } k \leq -1.$$

Proposition 2.15. *Example 8(i) is eventually Anosov. Indeed, its (unique, by Proposition 2.2 above) hyperbolic splitting is given on component M_0 by the unstable direction of $A = \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}$ and the stable direction of $B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$.*

Proof Since these are positive matrices, they map the positive cone of $\mathbb{R}^2$ into itself and their inverses map the cone $\{(x, y) : y \geq 0, x \leq 0\}$ into itself. A vector in the unstable space of A contracts in the mapping family as time goes toward $-\infty$; at time 0 (i.e. for component M_0) it lies in the positive quadrant, and by the Perron-Frobenius Theorem, any vector in the positive cone is attracted to the unique expanding eigendirection of the matrix B and hence expands as time goes to $+\infty$. The argument for the stable space is similar. □

Remark 2.15. Note that the unstable spaces are constant for times ≤ 0, and that for times $i > 0$ they keep changing direction, converging at $+\infty$ to the unstable space for the matrix B.

This illustrates an important more general phenomenon:

Proposition 2.16. *Let (M, f) be an eventually Anosov family. Let $E_0^s \oplus E_0^u$ be the hyperbolic splitting of TM_0. Then E_0^u depends only on the past, and E_0^s only on the future. That is, if $(M, \tilde{f})$ is another eventually Anosov family such that $f_i = \tilde{f}_i$ for $i \leq -1$, then $E_0^u = \widetilde{E}_0^u$, and similarly for the stable space for $i \geq 0$.*

Proof This is just like the proof of the uniqueness of the splitting: in getting the contradiction there we in fact only used the past of the sequence. □

Remark 2.16. To understand this statement it is useful to think of two complementary ways to define the unstable space at a point $p \in M_0$. For the first, E_0^u is the set of all vectors v such that $||Df_i(v)|| \to 0$ as $n \to -\infty$. This is similar to the definition of stable set, but for the inverse map. The second way of defining the unstable space is constructive, but only is easy to state for certain specific examples, for instance for nonnegative matrices; there, the unstable space is the intersection of all images of the positive cone mapped forward from times $n < 0$ to time 0 by the derivative map.

The reason this definition is harder to state in general is that one does not always have an analogue of the positive cone, which for the case of the nonnegative matrices is definitely disjoint from all the stable subspaces. A similar phenomenon occurs in Furstenberg's definition of the boundary of a matrix group [**Fur63**], [**Fur71**].

Now both these definitions agree in saying that the unstable space only depends upon the past.

Thus it might seem that the definition of eventually Anosov could be simplified to require only that vectors in the unstable space contract toward $-\infty$, without requiring that they also expand toward $+\infty$; and indeed, that works in the Anosov case. But here is an example which shows that for eventually Anosov families this is not such a good idea; see also Remark 2.11:

Example 9.

(ii) We take

$$A_k = A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \text{ for } k \geq 0, \;\; A_k = B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \text{ for } k \leq -1.$$

This mapping family is not eventually Anosov. If we try to find the unstable eigenspace by either of the above two methods, cones or stable direction for the inverse map, we will get the space E_0^u for the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, since this gives the past sequence. Similarly, using the future sequence $B = A^{-1}$ to determine the stable space, it again converges, but to the same space E_0^u. Hence for this family,

E^u_0 expands both toward $+\infty$ and $-\infty$, while E^s_0 for the matrix A contracts for both. Thus there is no hyperbolic splitting.

A third interesting example of a mapping family which is not Anosov is this:

Example 10.

(iii) We define

$$A_k = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ for } k \geq 0, \; A_k = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} \text{ for } k \leq -1.$$

We will examine all three examples more closely in a later paper.

Remark 2.17. One can generalize Anosov families in a number of obvious ways; we mention just two here. Example 5 also makes sense for higher genus surfaces. This leads to the related notion of a **pseudo-Anosov family**, which is hyperbolic except for an invariant collection of singularities, finite in number in each component. Essentially one is looking at Thurston's theory of measured foliations from a different perspective, adding a dynamics (of a mapping family) in the case when there does not exist a single map for which the foliation is invariant.

For the second generalization, let (M, f) be a noninvertible mapping family along a sequence of compact manifolds. We say this is an **expanding family** if and only if: the maps f_i are $\mathcal{C}^1$ differentiable but not necessarily 1-1, with branch points taken to branch points, and there exists $\lambda > 1$ as above such that

$$||D(f)|| \geq \lambda.$$

We say (M, f) is *two-sided* if $i \in \mathbb{Z}$ and *one-sided* if $i \in \mathbb{N} = \{0, 1, 2, \dots\}$.

An example of this second idea is due to David Rand. Following [**Ran88**], a *Markov family* is a one- or two- sided $\mathcal{C}^{1+\alpha}$ expanding (noninvertible) mapping family on the interval $M_i = [0, 1]$ with the Euclidean metric for all i, and such that the family is supplied with a sequence of partitions consisting of intervals which satisfy the natural version of the Markov condition. Rand's motivation was to use this idea in renormalization theory; see [**Ran88**], [**Pin91**], [**PR95a**], [**PR95b**], [**Sta88**].

3. *Symbolic dynamics of mapping families.*

3.1. *Partition sequences.* In the next sections we develop the machinery of symbolic dynamics in the setting of mapping families. For simplicity, we will from now on assume the maps are invertible, although what we do can be generalized to noninvertible families. Indeed, that was Rand's setting for Markov families, see Remark 2.17.

For the case of Anosov families, the result will be a Markov partition sequence which codes the family as a nonstationary version of a subshift of finite type; see Proposition 3.6.

We begin with the more general setting where the partitions are not necessarily Markov.

Definition 3.1. A **partition** $\mathcal{Q}$ of a compact metric space X is a finite collection $\mathcal{Q} = \{Q_i : i \in I\}$ of closed subsets of X such that:

- each Q_i is the closure of its interior, and its boundary is nowhere dense in X;
- the Q_i have disjoint interiors;
- $\cup Q_i = X$.

An **ordered partition** is a partition with totally ordered index set I; up to renaming, we can always take in this case $\mathcal{Q} = \{Q_0 \dots Q_l\}$.

We define a partition $\mathcal{P}$ of a mapping family (M, f) to be a sequence $\mathcal{P}_i$ of partitions in the above sense of the components M_i.

We say the partition **generates** for the mapping family if it separates points outside of the meagre set consisting of the pullback to the component M_0 of all future and past partition boundaries, i.e. if for each x, y in the G_δ subset which is the complement of the meagre set, there exists $n \in \mathbb{Z}$ such that $f^n(x)$ and $f^n(y)$ are in different elements of the partition $\mathcal{P}_n$ of M_n.

Note that since by our simplifying assumption the f_i are invertible maps of compact metric spaces, and hence are homeomorphisms, in the above definition of generation we could equivalently pull back to any other component M_i.

Recall that:

Definition 3.2. The **join** of two partitions $\mathcal{R}$ and $\mathcal{Q}$ of a space X, written $\mathcal{R} \vee \mathcal{Q}$, is the partition whose elements consist of the intersections of elements in each which are nontrivial in that they have nonempty interior. If $\mathcal{R}$ and $\mathcal{Q}$ have index sets I and J, then we index $\mathcal{R} \vee \mathcal{Q}$ by the subset of $I \times J$ corresponding to the nontrivial intersections. We extend this definition in the natural way to a finite number of partitions.

The notions of gathering and dispersal have natural counterparts for a partition on a mapping family:

Definition 3.3. Let (M, f) be a mapping family with partition $\mathcal{P}$, and let $(\widetilde{M}, \tilde{f})$ be a second family which is a gathering of (M, f) along the subsequence (n_i). For the gathered family we define a partition $\widetilde{\mathcal{P}}$ by $\widetilde{\mathcal{P}}_i = \mathcal{P}_{n_i} \vee f_{n_i}^{-1}(\mathcal{P}_{n_i+1}) \vee \cdots \vee (f_{n_{i+1}-2} \circ \cdots \circ f_{n_i})^{-1}(\mathcal{P}_{n_{i+1}-1})$. We call $\widetilde{\mathcal{P}}$ the **gathered partition**. Note that for this we have taken the join from time n_i to time $n_{i+1} - 1$.

We define a second partition $\widehat{\mathcal{P}}$ by including one more unit of time, taking the join from n_i to time n_{i+1}. We call this the **augmented gathered partition.**

Remark 3.1. Thus for example taking the trivial gathering of the family with partition $\mathcal{P}$, i.e. gathering along the subsequence $n_i = i$, the gathered partition does not change, so $\mathcal{P} = \widetilde{\mathcal{P}}$, while for the augmented gathered partition $\widehat{\mathcal{P}}_i = \mathcal{P}_i \vee f_i^{-1}(\mathcal{P}_{i+1})$.

We mention that symbolically $\widehat{\mathcal{P}}$ is a *two-block code* of $\mathcal{P}$, and corresponds to the edge rather than vertex labels on a Bratteli diagram; this shall be explained below.

We have:

Proposition 3.1. *If the partition $\mathcal{P}$ generates for (M, f), then the gathered partition and augmented gathered partitions $\widetilde{\mathcal{P}}$, $\widehat{\mathcal{P}}$ generate for the gathered family $(\widetilde{M}, \tilde{f})$. Conversely, if the gathered or augmented gathered partition generates for the gathered family, then the original partition $\mathcal{P}$ generates for the first family (M, f).* □

(The proof is immediate.)

Remark 3.2. Note that the augmented gathered partition is slightly less efficient, as there is redundancy: the partitions at times n_i are each included twice.

Given a mapping family with generating partition sequence, we extend this sequence to a dispersal of the family as follows: we simply take the trivial partition $\widetilde{\mathcal{P}}_i = \{\widetilde{M}_i\}$ on the new components. Clearly, this sequence generates for the dispersed family.

3.2. *Markov partitions.* Given an (invertible) mapping family, following the case of a single transformation as treated in [**Bow75**], [**Bow77**], we write $W^s_\varepsilon(p)$, $W^u_\varepsilon(p)$ for ε-disks in the stable and unstable leaves of a point p; we say the family has **canonical coordinates** if given $\varepsilon > 0$, there exists $\delta > 0$ such that if x and y are within δ of each other then $W^s_\varepsilon(x) \cap W^u_\varepsilon(y)$ consists of a single point, in which case we write $[x, y]$ for this point.

The existence of canonical coordinates for Axiom A maps is proved in [**Bow75**]; the idea is that there exists ε sufficiently small that $W^s_\varepsilon(x) \cap W^u_\varepsilon(x)$ is a single point and that this property is preserved under small perturbations. The proof goes through for families; we shall present that elsewhere.

The condition that $W^s_\varepsilon(x) \cap W^u_\varepsilon(x)$ be a singleton may fail for large ε for two reasons: first, one of the leaves may curve back to meet the other a second time; for this it needs enough room to turn around. Second, even if it is not curved, it may return by wrapping around the manifold. Both of these possibilities are eliminated with small enough ε.

The existence of locally defined canonical coordinates is used by Bowen to produce small rectangles. For the main examples studied in the present paper we want to allow for large rectangles; we give a different definition which works for our specific case, where the geometry is very simple, without striving for complete generality; here rectangles will be constructed explicitly.

So, supposing now that the components of our Anosov family are the flat two-torus and that the stable and unstable manifolds W^s, W^u are linear foliations, a **rectangle** will be a (filled-in) parallelogram with sides in W^s and W^u. See Fig. 1; in this case the eigendirections happen to be orthogonal, since the matrix $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ is symmetric.

We define $W^s(p, R)$ to be the connected component of $W^s(p) \cap \overset{\circ}{R}$ which contains p, and similarly for $W^u(p, R)$. If $W^s(x, R) \cap W^u(y, R)$ consists of a single point, we define $[x, y]$ to be this point. Canonical coordinates in this sense clearly exist for x, y in the interior of the rectangle.

The reason the boundary points have been excluded can be seen in Fig. 1; the larger of the two paralellograms wraps around the torus and so taking y in its unstable boundary and x in its interior, the unstable segment containing y meets the stable containing x in two points.

Now we return to Bowen's situation; we include this here to indicate what the two cases have in common, and to show how general hyperbolic sets for mapping families can be treated. Thus, given a mapping family with canonical coordinates in Bowen's sense, we define $R \subseteq M_i$ to be a **(small) rectangle** if:

(i) for any $x,\ y \in R$, $[x, y]$ is defined; and

(ii)for $x,\ y \in R$, $[x, y] \in R$. For $p \in R$ we then define $W^s(p, R)$ to be $W^s_\varepsilon(p) \cap R$ where ε is small and the diameter of R is less than ε, similarly for $W^u(p, R)$. Note that for such an R, give two points $x, y \in R$, then $W^s(x, R) \cap W^u(y, R)$ consists of a single point, $[x, y]$. Note that for small rectangles there is no need to exclude the boundary points.

From now on we follow the rest of Bowen's presentation, which works for both of the above types of rectangles.

We say a rectangle is **proper** if R is the closure of its interior $\overset{\circ}{R}$.

Definition 3.4. For a mapping family (M_i, f_i), a **Markov partition** is a sequence of finite partitions $\mathcal{R}_i$ of M_i, i.e. coverings of M_i by closed sets with disjoint interiors, such that each partition element is a proper rectangle, and such that the **Markov condition** is satisfied: for $R^i_j \in \mathcal{R}_i$ and $R^{i+1}_k \in \mathcal{R}_{i+1}$, such that $x \in R^i_j$ and $f_i(x) \in R^{i+1}_k$, then

$$f_i(W^u(x, R^i_j)) \supseteq W^u(x, R^{i+1}_k)$$

and

$$f_i(W^s(x, R^i_j)) \subseteq W^s(x, R^{i+1}_k).$$

Note that from the definition of proper rectangle, the partition boundaries are closed nowhere dense sets, so for a generating partition, the complement of the union of all pullbacks of partition boundaries to a single component is a dense G_δ. It is for these points that the symbolic dynamics will be defined.

We say a rectangle R **passes completely through** a second rectangle S in the stable (respectively unstable) direction if for a point $x \in R$, $W^s(x, R) \supseteq W^s(x, S)$, resp. $W^u(x, R) \supseteq W^u(x, S)$.

Then the Markov condition implies this geometric fact about partition intersections:

Lemma 3.2. *A Markov partition sequence $\mathcal{R}_i$ for an invertible mapping family (M, f) satisfies the* **geometric Markov property***: the preimage in the component M_i of each element R^j_{i+1} of a partition $\mathcal{R}_{i+1}$ under the map f_i either misses an element of $\mathcal{R}_i$ or passes completely through it in the stable direction. Similarly,*

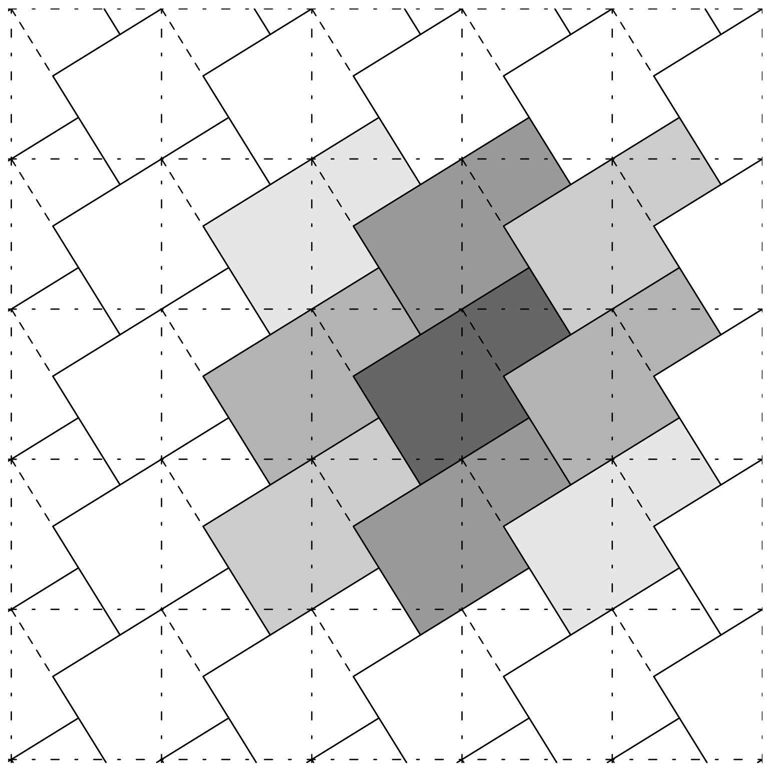

FIGURE 1. Generating Markov partition for the additive golden family, $\langle n\rangle = (\dots 111 \dots)$, for parity (+). For parity (−), the picture is reflected in the line $y = -x$. See §4.1.

elements of the push-forward of $\mathcal{R}_{i-1}$ to M_i pass completely through in the unstable direction.

Proof For the case of the square torus with parallelograms, this follows immediately from the Markov condition; see Fig. 3.2. For the general case, one can follow the proof of Lemma 3.17 of [**Bow75**]. □

The combinatorial consequence of this (see Proposition 3.6) is that a Markov partition gives a nice symbolic dynamics for the Anosov family: a mapping family along a sequence of combinatorially defined compact metric spaces, defined in the following section.

3.3. *Nonstationary subshifts.*

Definition 3.5. Let $\mathcal{A}_i$ for $i \in \mathbb{Z}$ be a sequence of finite nonempty sets, called **alphabets**, the elements of which shall be termed **symbols**. For definiteness, with $\#\mathcal{A}_i = a_i$, we take $\mathcal{A}_i = \{0, 1, \dots, a_{i-1}\}$. A **transition matrix** is a rectangular matrix with entries 0 or 1. Given a a sequence $(T_i)_{i\in\mathbb{Z}}$ of $(a_{i+1}) \times (a_i)$ transition matrices, an **allowed string** is a finite or infinite sequence (x_i) such that the $(x_{i+1}x_i)-$entry of T_i is equal to 1. A finite allowed string will be called a **word**.

We let T denote the entire sequence of transition matrices $(T_i)_{-\infty}^{\infty}$. We write $\Sigma^0 = \Pi_{-\infty}^{\infty}\mathcal{A}_i$ and define the subset Σ_T^0 to be the collection of allowed two-sided infinite strings $x = (\dots x_{-1}x_0x_1\dots) \in \Sigma_T^0$. We say the matrix sequence is **nondegenerate** iff each column and row has at least one nonzero entry.

Remark 3.3. In defining this space we have chosen the **column-vector convention.** We will see later in the paper why this choice has been made, rather than the more standard *row-vector convention*, for which the $(x_ix_{i+1})-$entry indicates the transition, with the matrices then $(a_i) \times (a_{i+1})$.

Next we introduce shift dynamics.

Definition 3.6. Let σT denote the left-shifted sequence of matrices, i.e. $(\sigma T)_i = T_{i+1}$. We define $\Sigma_T^k = \Sigma_{\sigma^k T}^0$ for $k \in \mathbb{Z}$. We set $\Sigma_T = \coprod \Sigma_T^k$, the disjoint union (i.e. the indexed union, see §2.1). We call Σ_T^k the k^{th} **component** of Σ_T, which we call the **total space.** We define a map σ on Σ_T, the **shift**, to be the map given by shifting a string to the left. We call Σ_T together with σ the **nonstationary subshift of finite type** (*nsft*) defined by $(\mathcal{A}_i)$ and (T_i).

The **present** coordinate of a point x in Σ_T^k is the symbol x_0; its **future** coordinates are x_i for $i \geq 1$, its **past** for $i \leq -1$.

Remark 3.4. For $n \in \mathbb{Z}$, the power σ^k maps the i^{th} to the $(i+k)^{\text{th}}$ component; thus for $x \in \Sigma_T^0$, with $x = (\dots x_{-1}x_0x_1\dots) \in \Pi_{-\infty}^{\infty}\mathcal{A}_i$, the point $\sigma^k x$ is the biinfinite string of symbols defined by $(\sigma^k x)_i = x_{i+k}$ is in Σ_T^k, which is a subset of $\Pi_{i=-\infty}^{\infty}\mathcal{A}_{i+k}$.

We emphasize that for x in Σ_T^k, x_0 denotes the 0^{th} coordinate in *that* component, not in Σ_T^0. A point $x = (\dots x_i \dots)$ in Σ_T is in some definite component, so it carries with it that definition of present (sometimes indicated by placing a "decimal point" to the left of x_0), and not that of the 0^{th} component.

Next we define a topology and metric on Σ_T:

Definition 3.7. We give each alphabet $\mathcal{A}_i$ the discrete topology and Σ_T^k the product topology. We define a topology on the disjoint union Σ_T as we did in §2.1 for general disjoint unions, combining these topologies discretely.

A **cylinder set** is a set of the form $[x_m \dots x_n] \equiv \{w \in \Sigma_T^k \ : w_i = x_i, m \leq i \leq n\}$ for some allowed finite string $x_n \dots x_m$, for some n, m in $\mathbb{Z}$.

Let $w_l(-j,k)$ denote the number of allowed words in Σ_T^l from $-j$ to k; this is also the number of cylinder sets in Σ_T^l of the form $[x_{-j} \dots x_0 \dots x_k]$. Note that by this definition $w_l(0,0) = a_l$ (the number of symbols in the alphabet at position 0 in Σ_T^l. Given x, y in the same component Σ_T^l, we define $d_l(x,y) = 1$ if $x_0 \neq y_0$; otherwise, hence assuming $x_0 = y_0$, we let j, m be the largest nonnegative integers such that $x_i = y_i$ for $-j \leq \ i \leq m$, and set

$$d_l(x,y) = \max\{(w_l(-j,0))^{-1}, (w_l(0,m))^{-1}\}.$$

We call this metric, extended discretely to the total space as above, the **word metric** on Σ_T.

We define a second metric on our *nsft*, the **θ- metric**, for $\theta \in (0,1)$, by: $d_\theta(x,y) = \theta^N$ where N is the largest integer ≥ 0 such that $x_i = y_i$ for all $|i| \leq N$.

Remark 3.5. As is easy to see, the word- and θ- metrics are equivalent (for any θ), i.e. they give the same topology. The θ- metric is the standard one for a subshift of finite type, see [**PP90**]. For the special case of a constant or periodic *nsft*, as we show elsewhere. the two metrics are comparable in a strong sense (this implies e.g. that the classes of Hölder functions are the same); but in general they are quite different. We shall see the naturalness of the word metric in the proofs of Proposition 3.9 below and in [**AF02b**].

We have:

Proposition 3.3. *The nsft* (Σ_T, σ) *is a mapping family.*

Proof It fits the definition given in §2.1: the metric and topology are obviously compatible; each component Σ_T^k is a compact metric space; indeed cylinder sets are clopen sets, and if infinitely many of the alphabets have at least two symbols, then it is topologically a Cantor set, and the total map σ is a sequence of homeomorphisms from one component to the next. □

Proposition 3.4. *If two matrix sequences* T, T' *defining mapping families* Σ_T *and* $\Sigma_{T'}$ *are nondegenerate, then these mapping families are the same iff the sequences* T, T' *are equal.*

Proof Knowing the matrix sequence is equivalent to knowing the allowed words. Nondegeneracy implies (indeed is equivalent to) that any finite allowed string can be continued infinitely in both directions. By compactness there exists a point in Σ_T^0 which has the name of such a string. Thus, knowing the space, i.e. knowing the infinite allowed strings, is equivalent to knowing the matrix sequences. □

Remark 3.6. Easy examples show that nondegeneracy is necessary; e.g. the constant sequences $T_i = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, $T'_i = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $T''_i = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ all define the same *nsft*, with only one point in $\Sigma_T^0 = \Sigma_{T'}^0 = \Sigma_{T''}^0$, the string $(\ldots 000 \ldots)$; one can easily see that a degenerate matrix sequence can be simplified by eliminating the symbols in each alphabet that belong to no allowed biinfinite string, thus producing a canonical nondegenerate matrix sequence with the same *nsft*.

Remark 3.7. A major difference between an *nsft* and a general mapping family is that for an *nsft*, each component carries all the dynamical information; simply by shifting it, all other components are reconstructed. Nevertheless, on a single component *there is no shift dynamics*, because even if the shifted symbol string is allowed - for example in the *constant* case, when the cardinalities $\#\mathcal{A}_i$ and the matrices T_i are all the same- the index has changed; we have moved to a different component of the disjoint union Σ_T.

As always for mapping families, we supress the index for points. Thus, two identical biinnfinite sequences x and y represent the same point in Σ_T if and only if they not only have the same string of symbols, $x_i = y_i$ for all i, but also belong to the same component. In particular, even in the constant case, *for an* nsft *there are no periodic points.*

In the constant case by forgetting the index, $\Sigma_T^k = \Sigma_T^0$ for all k, and Σ_T^0 is equal to the subshift of finite type (*sft*) Σ_A defined by the matrix. In this way the total space Σ_T naturally projects to Σ_A, with the map σ projecting to the usual shift map (also denoted σ) on Σ_A.

A reason for considering the *nsft* rather than the *sft* even in this constant case is that the *nsft* provides more flexibility. For example, the Gibbs theory of the two spaces is totally different; with the formalism of *nsfts* we can allow for a *sequence* of Hölder functions, one on each component, i.e. on each copy Σ_T^k of Σ_A. See [**AF02b**]. A related construction has been studied by Ferrero and Schmidt in [**FB88**], motivated by random dynamics. An additional, completely different reason for considering sequences of potentials is seen in work of Ruelle and Ledrappier [**Rue72**], [**Led77**]. Their idea is to use a nonstationary potential (i.e. a nonconstant sequence) to help study a stationary one; the nonstationary potential provides a direction in the function space in which to perturb the potential of interest. The nonstationary potential (or *interaction*) supplies a "small external field"; if the derivative in all "averageable" such noninvariant directions of the pressure function exists at this point (at the invariant potential), then there is a unique equilibrium state (of completely positve entropy, for [**Led77**]), and conversely. Ruelle's setting is that of lattice models of statistical physics; Ledrappier extends this to the dynamical setting, of a *weakly expansive* map on a compact metric space.

Remark 3.8. To represent the *nsft* by specific matrices we have made use of the order on the alphabets. Changing the order corresponds to conjugating the matrices with permutation matrices; since these may be a sequence as well, the appearance of the matrix sequence might change drastically. Thus fixing an order and hence a matrix representation is more important for an *nsft* than for the usual case of an *sft*.

3.4. *Symbolic dynamics for Anosov families.*

Here we shall see that an *nsft* gives exactly the symbolic representation for an Anosov family which is provided by a Markov partition sequence.

Lemma 3.5. *Given an invertible mapping family (M, f), assume it has a Markov partition sequence $\mathcal{R}_i$. If a finite sequence of partition elements $R_j, R_{j+1}, \dots R_{j+m}$ with $R_i \in \mathcal{R}_i$ has successively pairwise nonempty intersection when pulled back, i.e. if $R_{j+i} \cap f_{j+i}^{-1} R_{j+i+1} \neq \emptyset$, then the simultaneous intersection of the pullbacks to a single component is nonempty:*

$$R_{j+i} \cap f_{j+i}^{-1} R_{j+i+1} \cap \dots \cap f_{j+m-11}^{-1} R_{j+m} \neq \emptyset.$$

Proof This is immediate from the geometrical Markov property, Lemma 3.2. □

$$\begin{array}{ccccccccc}
 & \Sigma_T^{-1} & \xrightarrow{\sigma} & \Sigma_T^{0} & \xrightarrow{\sigma} & \Sigma_T^{1} & \xrightarrow{\sigma} & \Sigma_T^{2} & \\
\cdots\cdots & \downarrow \pi & & \downarrow \pi & & \downarrow \pi & & \downarrow \pi & \cdots\cdots \\
 & M_{-1} & \xrightarrow{f_{-1}} & M_0 & \xrightarrow{f_0} & M_1 & \xrightarrow{f_1} & M_2 &
\end{array}$$

FIGURE 2. Symbolic dynamics for a mapping family

Given an invertible mapping family and generating Markov partition, let a_k denote the number of elements of $\mathcal{R}_k$. We order each partition, and define the ij^{th} entry of an $(a_{k+1}) \times (a_k)$ matrix T_k to be 1 exactly when $f_k^{-1}(R_i^{k+1})$ meets R_j^k, where R_m^l denotes the m^{th} element of $\mathcal{R}_l$.

We have:

Proposition 3.6. *The map* $\pi : \Sigma_T^0 \to M_0$ *defined by* $x \mapsto \cap f^{-i} R_{x_i}^i$ *for* $i \in \mathbb{Z}$ *is* $1-1$ *off the set of boundary pullbacks. This is a topological semiconjugacy from the mapping family* (Σ_T, σ) *to* (M, f)*.*

Proof The key observation is that if a finite string is allowed in our *nsft*, the corresponding successive rectangles have pairwise disjoint intersection when pulled back, so by the previous Lemma, there is a point in the space which has that finite name. By compactness of the rectangles and the components M_i this extends to infinite allowed strings.

Then each Σ_T^k corresponds naturally to M_k via the projection map π_k. The disjoint union Σ_T projects to M via the map π, defined to be equal to π_k on each component Σ_T^k. The left shift σ maps Σ_T^k to Σ_T^{k+1}, so it projects to $f_k : M_k \to M_{k+1}$, and the total map σ on Σ_T projects to the total map f on M, and the diagram in Figure 2 commutes. □

Remark 3.9. As we have seen in §2.1, it is desirable for mapping families to have a semiconjugacy which is not just topological but is uniform; for the coding map π both uniform and nonuniform examples occur, see Remark 5.2.

We remark that sometimes it is useful to consider the total map as a single map rather than as a sequence; from that point of view, for instance, the Markov partition for the Anosov family is a Markov partition in the usual sense for the total map, but with countably many elements; also, the total map f on M itself is hyperbolic, hence almost fits the definition of an Anosov diffeomorphism (see e.g. [**Bow75**], [**Bow77**], [**Shu87**]); what is missing is that the total space is not compact. And, of course, there is no recurremce and there are no periodic points. Nevertheless, this point of view can be of use e.g. when studying the local theory.

3.5. *Nonstationary vertex and edge shift spaces and and Bratteli diagrams.* So far we have defined *nsfts* when given a sequence of $0-1$ matrices. We now generalize this to matrix sequences with nonnegative integer entries, using the same edge-shift rather than vertex-shift idea as for the familiar case of *sfts*, see [**Fra82**] p. 20 or

[**LM95**] p. 43 and §4 below. As for *sfts*, it shall be of great help to represent the transitions pictorially, by means of a graph, the difference being that here we shall need an infinite graph.

Definition 3.8. A **Bratteli diagram** is a directed graph defined by a sequence of finite vertex sets $\mathcal{V}_i$ and edge sets $\mathcal{E}_i$ indexed by $i \in \mathbb{Z}$. Each edge $e \in \mathcal{E}_i$ has a **source** and **range** $s(e) \in \mathcal{V}_i, r(e) \in \mathcal{V}_{i+1}$, and is drawn as an arrow with tail at the source and head at the range.

Definition 3.9. We say the diagram is **nondegenerate** if each vertex has at least one arrow coming into it, and at least one arrow leading out of it.

Remark 3.10. We borrow this tool from the study of C^*- algebras, see e.g. [**BH94**], [**ES80**], [**DHS99**], [**Dur98**]. Here we use two-sided diagrams, with index set $\mathbb{Z}$, rather than the one-sided diagrams which are more usual. Also, we draw our diagrams horizontally with arrows from source to range and pointing from left to right, as the index i increases.

We now describe the edge and vertex labels and associated matrix sequences.

Definition 3.10. First we identify the vertex set $\mathcal{V}_i$ with an ordered alphabet $\mathcal{A}_i = \{0, \dots, a_i - 1\}$; the resulting Bratteli diagram is the **vertex-labelled** diagram. We associate to the diagram a sequence of $(a_{i+1}) \times (a_i)$ matrices $(F_i)_{i\in\mathbb{Z}}$ with nonnegative integer entries by setting the ml^{th} entry of F_i equal to k iff there are k edges connecting vertex l in $\mathcal{V}_i$ to vertex m in $\mathcal{V}_{i+1}$ (and to be 0 if there are no such edges).

We say the diagram is **single-edged** iff there is at most one edge from a given source to a given range, or equivalently, iff the matrices F_i have entries 0 or 1.

For the second, we identify the edge set $\mathcal{E}_i$ with an ordered alphabet $\mathcal{A}_i$. This gives an **edge-labelled** diagram. We define a sequence (T_i) of matrices by taking the ml^{th} entry of T_i to be 1 iff the edge m in $\mathcal{E}_{i+1}$ follows the edge l in $\mathcal{E}_i$, and otherwise to be 0.

Proposition 3.7. *In each case, the diagram determines and is determined by the matrix sequence (F_i), which have nonnegative integer entries for the vertex labels and which always have $0-1$ entries for the edge labels.* □

Remark 3.11. The specific matrix sequence is defined in each case by the order on the alphabet, corresponding to ordering the vertices or edges respectively.

As for *nsfts*, here we have adopted the column vector convention when defining the matrices.

An *nsft* corresponds to a sequence of $0-1$ matrices. If the diagram happens to have single edges, therefore, there are two natural ways to define a sequence of $0-1$ matrices and hence an associated *nsft*, by labelling the vertices or the edges.

Definition 3.11. Given a Bratteli diagram with edge labels, we call the *nsft* defined by the $0-1$ sequence (F_i) the **edge-shift space** of the diagram. If it is a vertex-labelled and single-edged diagram, we we call the resulting *nsft* the **vertex-shift space** of the diagram.

Proposition 3.8. *Two nondegenerate Bratteli diagrams determine the same edge shifts iff they are equal; for single-edged diagrams the same holds for the vertex shifts.*

Proof This is just like the proof of Proposition 3.4. □

Remark 3.12. Given a vertex-labelled diagram with single edges, there is a natural choice for naming the edges: if an edge e connects symbols $s(e) = i$ and $r(e) = j$, then this edge is named ij; these can then be ordered lexicographically, producing an edge-labelled diagram. This is illustrated in Figures 3, 4, where we have started with the vertex-labelled Bratteli diagram corresponding to the sequence

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \text{ for } i \text{ even}; \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ for } i \text{ odd}$$

and where the vertices are listed in increasing order from top to bottom.

The edge-labelled diagram defines an *nsft*, the **additive golden shift**, which in turn gives a new vertex-labelled diagram, see Figure 5. Symbolically, this construction (the new *nsft* produced by the change of symbols from vertices to edges) has this description:

Definition 3.12. A **two-block code** of an *nsft* is the new *nsft* produced by taking as the alphabet allowed words of length two, with the inherited transition rules.

As in the case of *sfts*, see Proposition 3.2 of [**Fra82**], we have:

Proposition 3.9. *An* nsft *is topologically conjugate to its edge shift* nsft *via the two-block coding. If $a_i = \#\mathcal{A}_i \le l$ for all $i \in \mathbb{Z}$, then the conjugacy is uniform; in fact, for x, y in the first* nsft*, then writing x' for the image of x,*

$$d(x, y) \le d(x', y') \le l \cdot d(x, y).$$

Proof Write Σ_T for the first *nsft*, $\Sigma_{T'}$ for the second. There are several cases to consider. First, x and y are in different components of Σ_T iff x', y' are, in which case by definition, $d(x, y) = 1 = d(x', y')$. So assume x, y are in some component, say in Σ_T^0. Now if $x_0 \ne y_0$ then also $x'_0 \ne y'_0$ hence again $d(x, y) = d(x', y') = 1$. So assume $x_0 = y_0$. The key observation is that for any integers $j < k$, the allowed blocks $(x_j \dots x_k)$ in Σ_T^0 are in $1-1$ correspondence with blocks $(x'_j \dots x'_{k-1})$ in $\Sigma_{T'}^0$, by taking $x'_j = x_j x_{j+1}$ and so on; this implies that

$$w_0(j, k) = w'_0(j, k-1).$$

Now let j, m be the largest integers ≥ 0 such that $x_i = y_i$ for $-j \le i \le m$. Consider first the case where $m \ge 1$. Then $x'_i = y'_i$ for $-j \le i \le m-1$ where these are the largest such nonnegative integers. By the above observation, $w_0(-j, 0) = w'_0(-j, 0)$ and $w_0(0, m) = w'_0(0, m-1)$. So also in this case, $d(x, y) = d(x', y')$.

The last case to consider is when $m = 0$, i.e. when $x_0 = y_0$ but $x_1 \ne y_1$. Here the distances are in general different; indeed $x'_0 \ne y'_0$ so $d(x', y') = 1$, while $d(x, y) = \max\{(w_0(-j, 0))^{-1}, (w_0(0, m))^{-1}\} = \max\{(w_0(-j, 0))^{-1}, (a_0)^{-1}\} \ge (a_0)^{-1} \ge 1/l$. Thus in this case $d(x', y') = 1 \ge d(x, y) \ge 1/l = d(x', y')/l$, which gives the statement. □

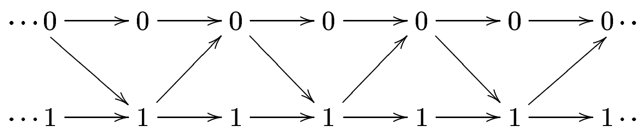

FIGURE 3. The vertex-labelled Bratteli diagram for the additive golden shift.

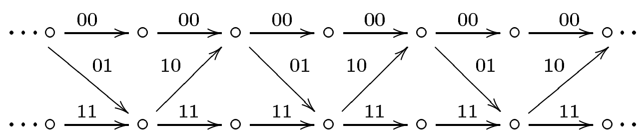

FIGURE 4. The corresponding edge-labelled diagram.

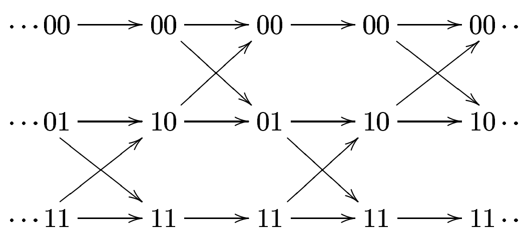

FIGURE 5. The vertex-labelled diagram which gives the same *nsft* as Figure 4, and which is a two-block code of Figure 3.

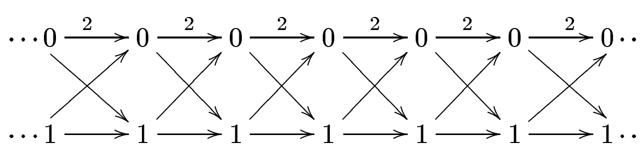

FIGURE 6. Telescoped diagram corresponding to the gathering along even times of the additive golden shift, with number of edges indicated when > 1.

There are two natural operations on Bratteli diagrams, see e.g. [**DHS99**], [**BH94**], [**GJ98**]:

Definition 3.13. Given an increasing subsequence $\dots n_{-1}n_0n_1\dots$ of $\mathbb{Z}$, the **telescoping** of a Bratteli diagram is the diagram defined by the sequence of matrices given by partial compositions of the original sequence; conversely, a **microscoping** is given by a factoring of the sequence.

Remark 3.13. Since we are using the column-vector convention, the order is reversed in taking this product.

We now see that this gives a symbolic version of our operations on mapping families:

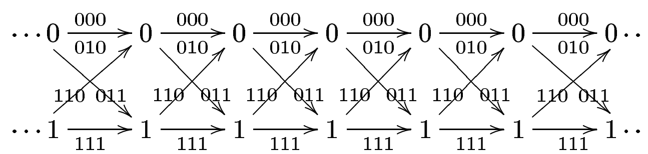

FIGURE 7. The same diagram with its additive edge labels; the upper two labels are for two edges.

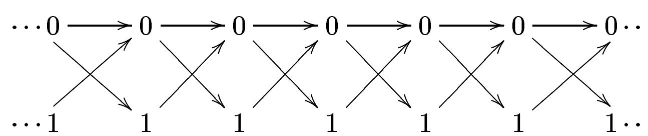

FIGURE 8. A second diagram which telescopes to that of Figure 6; now the edge labels are different from those of Figure 7.

Proposition 3.10. *Given a generating Markov partition $\mathcal{P}$ of a mapping family (M, f), the operations of gathering and of dispersal via the insertion of identity maps for the family correspond to telescoping and microscoping of the Bratteli edge diagram. That is, the nsft for the augmented trivial gathering of $\mathcal{P}$, which is the partition $\widetilde{\mathcal{P}}$ defined by $\widetilde{\mathcal{P}}_i = \mathcal{P}_i \vee \mathcal{P}_{i+1}$, is equal to the nsft for the edge labeling of the Bratteli diagram associated to $\mathcal{P}$; and telescoping this diagram along an increasing subsequence n_i gives an edge-labelled diagram whose nsft is identical to that for the augmented gathered partition taken along that subsequence.* □

The proof is immediate from the definitions; See Remark 3.1.

3.6. *The (2×2) case: additive sequences and canonical labelings.* Now we restrict attention to Bratteli diagrams with two vertices, i.e. those given by a sequence of (2×2) nonnegative matrices. This is the case we shall need for the rest of the paper.

We noted above that for a single-edged Bratteli diagram, a vertex labelling gives natural labels for the edges. In general for multiple edges, there is no canonical way to pass from vertex to edge labels; however as we shall see, in the (2×2) case this is in fact possible. This labelling will be needed later, in Theorem 5.6.

We recall that $SL(2, \mathbb{Z})$ is the group of (2×2) matrices with integer entries and with determinant 1. We shall write $SL(2, \mathbb{N})$ for the subsemigroup whose entries are all ≥ 0.

The next lemma is well-known and we do not know the proper attribution; we learned this simple proof a long time ago, perhaps from Rauzy.

Lemma 3.11. *$SL(2, \mathbb{N})$ is the free semigroup generated by the* **additive generators**

$$M = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and } N = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Proof We note that the identity $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is included here as $I = M^0 = N^0$. Let $A \in SL(2, \mathbb{N})$, with $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. We claim that if $A \neq I$, then either the first column is $\geq$ the second, in the sense that $a \geq b$ and $c \geq d$, or the reverse.

If both of these conditions fail then either $a > b$ and $c < d$ or the reverse. However the reverse ($b < a$ and $d > c$) cannot happen as this would imply that $bc > ad$ so $ad - bc < 0$, but by assumption the determinant is one.

Since therefore $a > b$ and $c < d$, we have: $a \geq b + 1$ and $d \geq c + 1$ so the determinant is:

$$ad - bc \geq (b + 1)(c + 1) - bc = bc + b + c + 1 - bc = b + c + 1.$$

Since $\det A = 1$, we have b and $c = 0$ in which case $A = I$, as claimed.

Now we show that $A \in SL(2, \mathbb{N})$ can be factored as a product of nonnegative powers of M and N. Writing $A = A_0$, if $A_0 \neq I$ then remove the smaller column from the larger to form A_1. This amounts to writing

$$A_1 = A_0 M^{-1} \quad \text{or } A_1 = A_0 N^{-1};$$

note that the new matrix A_1 is still in $SL(2, \mathbb{N})$. If A_1 again has one column larger than the other then we continue, producing a sequence $A_0, A_1, \ldots, A_n$. This process terminates with a matrix A_n with determinant one and which has neither column larger than the other. So as shown above, $A_n = I$. Thus, reversing the process, we have factored A as a product of nonnegative powers of M and N.

We have proved a little more: the preceding argument shows that an element of $SL(2, \mathbb{N})$ which is not the identity can be factored either as $A = A_1 N$ or as $A = A_1 M$, but not both. Therefore the decomposition of A in terms of M and N is unique, and this implies that there can be no nontrivial relations in the semigroup $SL(2, \mathbb{N})$; hence it is free. □

Remark 3.14. What the proof has shown is that there is a unique way to get to a given positive matrix with determinant one by starting with I, adding one column to the other and repeating the operation. From a different point of view, this is exactly the construction of the Farey tree.

Of course one would like such a theorem for the (3×3) case, but there is no hope of such a simple structure, as it is known that $SL^+(3, \mathbb{Z})$ is neither finitely generated nor free. (Joël Rivat, private communication).

A sequence $(A_i)_{i \in \mathbb{Z}}$ such that each matrix either equals one of the additive generators M, N we shall call an **additive sequence**. We call the Bratteli diagram defined by such a sequence an **additive diagram**.

Given a sequence (F_i), from $SL(2, \mathbb{N})$, consider the Bratteli diagram determined by (F_i); thus there are two symbols 0 and 1 and the matrix entries specify the number of edges. In general for a multiple-edged diagram there is no natural way to label the edges. However in this case we have such a way, given by the Lemma:

Corollary 3.12.

(i) A sequence (F_i) with $F_i \in SL(2,\mathbb{N})$ uniquely determines an additive sequence (A_j) such that $F_0 = A_n \cdots A_1 A_0$, $F_1 = A_{n+k} \cdots A_{n+1}$ and so on for F_i with $i \in \mathbb{Z}$.

(ii) Fixing the location of time 0, there is a unique additive diagram, the **additive microscoping**, *which telescopes to the Bratteli diagram for (F_i). The gathered labels on the edges of this diagram specify the paths of the additive microscoping.* □

We call the resulting canonical edge labels for the diagram its **additive labels**. See Fig. 3.5.

Thus, given a sequence (F_i) of (2×2) matrices with nonnegative integer entries, in view of Corollary 3.12, we can associate an *nsft* to (F_i) of that special type in a canonical way:

Definition 3.14. Given sequence (F_i) with $F_i \in SL(2,\mathbb{N})$, we define Σ_F to be the *nsft* determined up to by gathering the *nsft* determined by the additive factorization, with its time-zero partition, and with the gathered partition elements labelled accordingly, and ordered in some chosen way.

4. *The multiplicative family.*

For the rest of the paper we focus on certain Anosov families on the two-torus, given by a sequence of (2×2) matrices. As we will see, the corresponding codings will also come from (2×2) matrices, as studied in §3.6. Our main example is this:

Definition 4.1. Given the choice of a sequence $\langle n \rangle \equiv (\dots n_{-1} n_0 n_1 \dots) \in \Pi_{-\infty}^{+\infty} \mathbb{N}^*$ for $\mathbb{N}^* = \{1, 2, \dots\}$ and a **parity** $p \in \{+, -\}$, the **square torus version** of the **multiplicative family** determined by $\langle n \rangle$ and p is the mapping family $(M, f) = (M_i, f_i)$ along a sequence of two-tori, defined as follows.

We take first the case of parity $(+)$. We set

$$A_i = \begin{bmatrix} 1 & 0 \\ n_i & 1 \end{bmatrix} \text{ for } i \text{ even; } \quad \begin{bmatrix} 1 & n_i \\ 0 & 1 \end{bmatrix} \text{ for } i \text{ odd.}$$

For each $i \in \mathbb{Z}$, let $M_i = \mathbb{R}^2/\mathbb{Z}^2$, with the standard Euclidean metric and thought of as column vectors.

We define $f_i : M_i \to M_{i+1}$ to be the function given by left multiplication by A_i; that is, for (x, y) in $M_i = \mathbb{R}^2/\mathbb{Z}^2$,

$$f_i : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto A_i \begin{bmatrix} x \\ y \end{bmatrix}.$$

For parity $(-)$ the transposes are used instead.

We will show (M, f) is an Anosov family, by describing the expanding and contracting eigendirections and eigenvalues explicitly. Changing to eigencoordinates will give a second version of the multiplicative family, which is boundedly conjugate to the square torus version; see Proposition 4.3.

We shall use this terminology: an individual matrix A_i as above has parity $(+)$, $(-)$ if it is lower, respectively upper, triangular; thus the *sequence* $(A_i)_{i\in\mathbb{Z}}$ has parity $(+)$ exactly when each of the matrices at even times has parity $(+)$.

From the sequence (A_i) we will define a sequence of matrices (B_i), the columns of which will give our eigenvectors. These matrices will be in a set $\mathcal{B}_0 = \mathcal{B}_{0,+} \cup \mathcal{B}_{0,-} \subseteq SL(2,\mathbb{R})$ with $B = \begin{bmatrix} a & c \\ -b & d \end{bmatrix}$ satisfying the conditions:

1. (*i*) $a, b, c, d \geq 0$

   (*ii*) $\det B = 1$

   (*iii*) for $B \in \mathcal{B}_{0,+}$, $0 < a < 1 \leq b = 1$ and $d < c$

   for $B \in \mathcal{B}_{0,-}$, $0 < b < 1 \leq a = 1$ and $c < d$.

We say $B \in \mathcal{B}_0$ has parity $(+)$ or $(-)$ when it is in $\mathcal{B}_+$ or $\mathcal{B}_-$ respectively.

Given a sequence of integers $m_0, m_1, \ldots$ with $m_i \geq 1$, we shall use this notation for the (*multiplicative*) continued fraction:

$$[m] = [m_0 m_1 \ldots] = \cfrac{1}{m_0 + \cfrac{1}{m_1 + \cdots}}$$

The correspondence $(m_i)_{i\in\mathbb{N}} \mapsto [m]$ defines a bijection from $\Pi_0^{+\infty}\mathbb{N}^*$ and the set of irrationals in $(0,1)$.

Proposition 4.1. *Given a choice of* $\langle n \rangle \in \Pi_{-\infty}^{+\infty}\mathbb{N}^*$ *and parity* $(+)$ *or* $(-)$, *then the mapping family* (M, f) *as defined above is an Anosov family. The eigenspaces* E_i^s *and* E_i^u *of* $M_i = \mathbb{R}^2/\mathbb{Z}^2$ *are spanned by vectors* $v_i^s = \begin{bmatrix} a_i \\ -b_i \end{bmatrix}$, $v_i^u = \begin{bmatrix} c_i \\ d_i \end{bmatrix}$, *with eigenvalue sequences* $(\lambda_i^{-1}) < 1$ *and* $(\lambda_i) > 1$ *respectively, where these are defined by the condition* $a_i d_i + c_i b_i = 1$ *together with:*

$$\text{for } A_i \text{ with parity } (+): \quad a_i = [n_i n_{i+1} \ldots], \quad b_i = 1, \quad \frac{d_i}{c_i} = [n_{i-1} n_{i-2} \ldots], \text{ and } \quad \lambda_i = \frac{1}{a_i}$$

$$\text{for } A_i \text{ with parity } (-): \quad b_i = [n_i n_{i+1} \ldots], \quad a_i = 1, \quad \frac{c_i}{d_i} = [n_{i-1} n_{i-2} \ldots], \text{ and } \quad \lambda_i = \frac{1}{b_i}$$

Proof We began with the sequence $\langle n \rangle$, parity $p = +$ or $(-)$, and the associated sequence of matrices (A_i). From the positive real numbers a_i, b_i, c_i, d_i determined by this as in the statement of the Proposition, we define

$$B_i = \begin{bmatrix} a_i & c_i \\ -b_i & d_i \end{bmatrix}$$

Since $a_i d_i + c_i b_i = 1$, we have $\det B_i = 1$. Thus, from the sequence (A_i), or equivalently from $\langle n \rangle \in \Pi_{-\infty}^{+\infty}\mathbb{N}^*$ together with a choice of parity p we have defined a sequence of matrices $B_i \in \mathcal{B}_0$, with parity of B_i the same as that for A_i.

Next we define the diagonal matrices

$$D_i = \begin{bmatrix} \lambda_i & 0 \\ 0 & \lambda_i^{-1} \end{bmatrix}$$

We claim that the sequences A_i, B_i, D_i are related as follows:

$$A_iB_iD_i = B_{i+1}.$$

This will finish the proof, since the vectors v_i^s, v_i^u are the columns of B_i, and we have from the above equation that

$$A_iv_i^s = A_iB_i\begin{bmatrix}1\\0\end{bmatrix} = B_{i+1}D_i^{-1}\begin{bmatrix}1\\0\end{bmatrix} = \lambda_0^{-1}v_{i+1}^s$$

as claimed.

To verify the claim is immediate; we check this for $i = 0$ and parity $(+)$. We have

$$A_0B_0 = \begin{bmatrix}a_0 & c_0\\ -(b_0 - n_0a_0) & d_0 + n_0c_0\end{bmatrix} \tag{2}$$

and so:

$$b_1 = \frac{b_1}{a_1} \equiv [n_1n_2\ldots] = \frac{1}{a_0} - n_0 = \frac{b_0 - n_0a_0}{a_0} = \lambda_0(b_0 - n_0a_0), \tag{3}$$

and similarly for the other matrix entries. Hence indeed $A_0B_0D_0 = B_1$, as claimed. □

We use the above diagonalization to define a second mapping family, which expresses the multiplicative family in eigencoordinates.

We write $N_i \equiv \mathbb{R}^2/\Lambda_i$ where Λ_i is the lattice spanned by the columns of B_i^{-1}, with the Euclidean metric inherited from $\mathbb{R}^2$, and define maps $g_i : N_i \to N_{i+1}$ by

$$g_i : \begin{bmatrix}x\\y\end{bmatrix} \mapsto D_i^{-1}\begin{bmatrix}x\\y\end{bmatrix}.$$

We call (N, g) the **eigencoordinate version** of the multiplicative family (M, f).

Lemma 4.2. *For a matrix in $\mathcal{B}_0$, the row vectors have length in the interval $(1/2, \sqrt{2})$ and the angle θ between them satisfies $\sin(\theta) \geq 1/2$.*

Proof Consider the case of parity $(+)$. From the conditions $1(i), (ii), (iii)$ which define $\mathcal{B}_0$, we have $b = 1$ and $a, c, d \leq 1$. Hence the vectors $(-b, d)$ and (a, c) have length no more than $\sqrt{2}$.

Writing $v = (-b, d)$ and $w = (a, c)$, we know the parallelogram spanned by v, w has area 1, so $1 = ||v|| \cdot ||w|| \sin(\theta)$ where θ is the angle between them. So from the length estimate, $\sin(\theta) = 1/(||v|| \cdot ||w||) \geq 1/2$ as claimed. □

Note: these estimates are not optimal, but we need only rough bounds.

Proposition 4.3. *The Anosov families (M, f) and (N, g) are boundedly conjugate via the maps $\tilde{h}_i : M_i \to N_i$ defined by*

$$\begin{bmatrix}x\\y\end{bmatrix} \mapsto B_i^{-1}\begin{bmatrix}x\\y\end{bmatrix}.$$

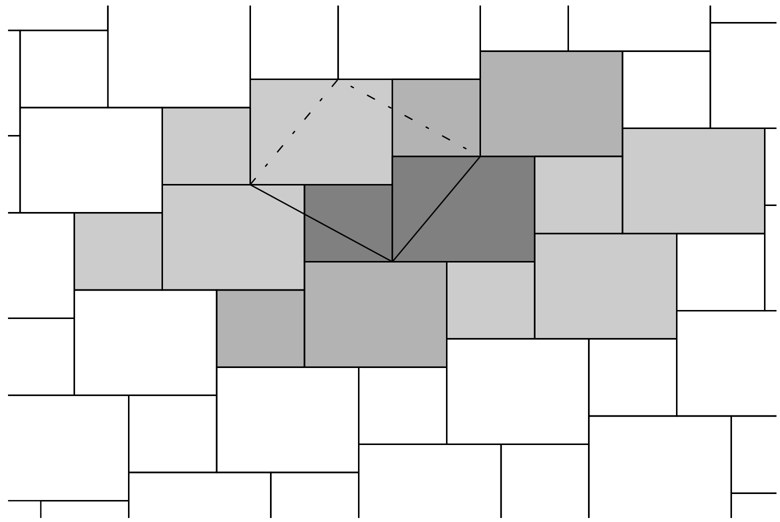

FIGURE 9. Two fundamental domains for the lattice, in eigencoordinates: the parallelogram and the two-box partition $\mathcal{R}_0$. The picture has been rotated counterclockwise by 90° for convenience. Thus in the above picture, the horizontal axis gives the expanding direction.

For (N, g), the eigenvector sequences

$$w_i^s = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } w_i^u = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

have the same eigenvalues, (λ_i^{-1}) and (λ_i) respectively.

Proof From the formula

$$B_0 \mapsto A_0 B_0 D_0 = B_1$$

we have the commutative diagram

$$\begin{array}{ccccccccc} & M_0 & \xrightarrow{A_0} & M_1 & \xrightarrow{A_1} & M_2 & \xrightarrow{A_2} & M_3 & \\ \cdots\cdots & \uparrow B_0 & & \uparrow B_1 & & \uparrow B_2 & & \uparrow B_3 & \cdots\cdots \\ & N_0 & \xrightarrow{D_0^{-1}} & N_1 & \xrightarrow{D_1^{-1}} & N_2 & \xrightarrow{D_2^{-1}} & N_3 & \end{array}$$

for the action on column vectors, and D_i^{-1} is the diagonalization of A_i with respect to the eigenbases. The conjugacy of the families is then simply left multiplication of column vectors by B_i^{-1}. This shows the conjugacy on $\mathbb{R}^2$. It remains to check that N_i maps to N_{i+1}. It is enough to show that $D_i^{-1}(\Lambda_i) = \Lambda_{i+1}$. Taking the inverse of the formula $A_i B_i D_i = B_{i+1}$ we have

$$B_{i+1}^{-1} = D_i^{-1} B_i^{-1} A_i^{-1};$$

now since $A_i^{-1} \in SL(2, \mathbb{Z})$, right multiplication by this gives a matrix whose columns span the same lattice in $\mathbb{R}^2$, as we wanted.

The statement about eigenvectors follows immediately from the diagonalization.

Now we show the conjugacy is bounded. From Lemma 4.2, the ellipses which are the images of the unit ball by the matrices B_i acting on row vectors have a uniformly bounded inner and outer radius. Hence the same is true for the matrices B_i^{-1} acting on columns, as the columns of B_i^{-1} are a rotation by $\pi/2$ of the rows of B_i. Thus by Lemma 2.9, this is a bounded conjugacy. □

Remark 4.1. The above conjugacy, while bounded, is not an isometry. Indeed, for the family (M, f) the metric on each torus M_i is the same, while the tori N_i in general are not even isometric, as for instance the length of the minimum closed geodesic will change. For M_i the metric has been chosen so that the standard basis vectors are orthogonal and have length 1, while the eigenvectors are in general neither orthogonal nor of length 1; for (N, g), the eigenvectors are instead chosen to be orthonormal.

The next construction provides a different perspective on this example. Let us write ρ_0 for the Riemannian metric on N_0 inherited from the plane. We define a third mapping family $(\widetilde{N}, \tilde{g})$ with $\widetilde{N}_i = N_0$ and with g_i the identity map for all $i \in \mathbb{Z}$; we define Riemannian metrics $\tilde{\rho}_i$ on $\widetilde{N}_i$ by setting for $i > 0$ $\tilde{\rho}_i = \lambda_{i-1} \cdots \lambda_0 \rho_0$ on E^u, $\tilde{\rho}_i = \lambda_{i-1}^{-1} \cdots \lambda_0^{-1} \rho_0$ on E^s; for $i < 0$ we take $\tilde{\rho}_i = \lambda_i^{-1} \cdots \lambda_{-1}^{-1} \rho_0$ on E^u, $\tilde{\rho}_i = \lambda_i \cdots \lambda_{-1} \rho_0$ on E^s; we extend this by linearity to an inner product on each tangent space, i.e. to a unique Riemannian metric.

Proposition 4.4. *The family $(\widetilde{N}, \tilde{g})$ is isometrically isomorphic to (N, g).*

Proof The conjugacy is: h_0 =identity; $h_1 = g_0^{-1}$, $h_2 = g_0^{-1} \circ g_1^{-1}$ and so on. □

Remark 4.2. In Proposition 2.6, we discussed the two extreme cases for mapping families, one where the metric is fixed and the other where the maps are the identity and only the metrics are changing, the general case being a mixture of these two. From the proof, we see that $(\widetilde{N}, \tilde{g})$ is exactly the family constructed in the first part of that proposition. That is, all of the dynamics of the family (N, g) has been pushed to the metrics.

The square torus version represents the other extreme, where all the components are isometric and the dynamics is carried entirely by the maps. The eigencoordinate version (N, g) is an intermediate case; the metrics are also changing as the component spaces are in general not isometric to each other, as noted above. It is possible however that the manifold M_i happens to return to being isometric to M_0 (by another map). This is the case for the multiplicative family if and only if the family is periodic. This statement is also true for $(\widetilde{N}, \tilde{g})$, in which case one can then recover a constant family which is isometric to a gathering of the family (N, g) by way of the construction in the proof of the second part of Proposition 2.6.

We note also that the family $(\widetilde{N}, \tilde{g})$ is like that constructed in Example 4.

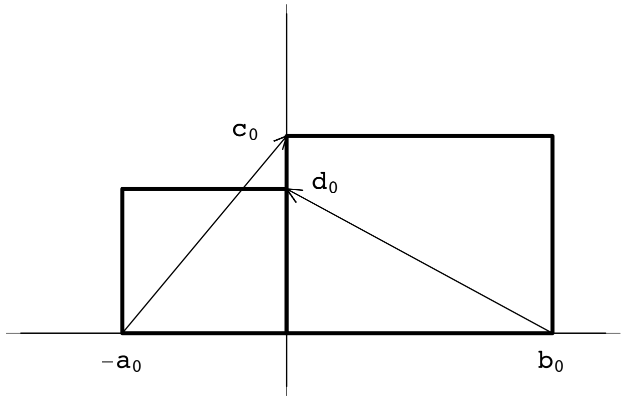

FIGURE 10. The box pair in eigencoordinates, for i with parity $(+)$: the row vectors of B_i are the coordinates of the vectors which generate the parallelogram lattice. The picture has been rotated 90° counterclockwise. For parity $(-)$, this picture is reflected in the vertical axis. In both cases, the left box is R_i^1 and the right box is R_i^0.

4.1. *Box renormalization and the construction of Markov Partitions.* Although the proof of Proposition 4.1 above appears to be purely formal and algebraic, there is a simple geometrical interpretation behind the formulas, which we will explain in this section.

Thus in particular the appearance of the continued fractions is no accident; it comes from the renormalization of circle rotations. From this perspective, the matrices A_i and B_i which occur in the key formula which gave the diagonalization, $A_i B_i D_i = B_{i+1}$, have two quite different interpretations. The matrices A_i provide that renormalization, so in one set of coordinates they are purely combinatoric (a change of basis in a lattice) while in other coordinates they are hyperbolic maps.

The matrices will have two corresponding interpretations: as a change of basis matrix which diagonalizes the Anosov family, and as the coordinates of a Markov partition for the family. And, as we shall see in [**AF02a**], there is a third interpretation for B_i, as a unit tangent vector to the Teichmüller space of the torus.

We first construct the Markov partitions in the square torus version. As we saw in the proof of Proposition 4.1, the columns of the matrix B_i give the eigenvector sequence v_i^s, v_i^u. The maps A_i are in $SL(2,\mathbb{Z})$, hence are a mapping family along the sequence of square tori $M_i = \mathbb{R}^2/\mathbb{Z}^2$. We now use the matrix B_i to build two parallelograms which will be the Markov partition for the i^{th} component M_i. These are: the parallelogram with sides spanned by the vectors $d_i v_i^s$, $a_i v_i^u$, and that spanned by $c_i v_i^s$, $-b_i v_i^u$. Note that the union of these two parallelograms gives a second fundamental domain for the lattice $\mathbb{Z}^2$ acting on $\mathbb{R}^2$; thus they partition the torus. Indeed the condition $\det B_i = 1$ implies that the areas of the parallelograms $a_i d_i \cdot \det(B_i)$, $|-b_i c_i| \cdot \det(B_i)$ add up to one.

See Fig. 1 for the simplest example (with $n_i = 1$ for all i); in this case the eigendirections are orthogonal, as noted in §3.2, and moreover the boxes happen to

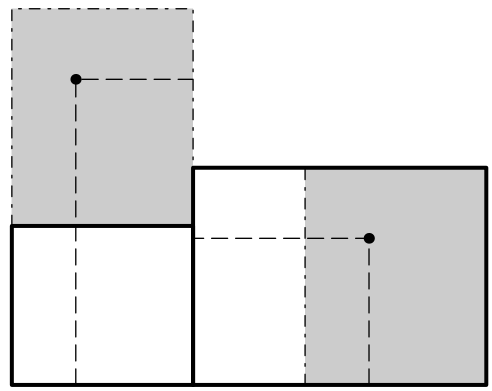

FIGURE 11. Box renormalization, in the eigencoordinates, showing two points equivalent under the lattice action, after rotation of the coordinates by 90° counterclockwise.

be square, as is easily calculated from the above formulas.

In the eigencoordinates, the first parallelogram is spanned by the vectors $c_i w_i^s$, $-b_i w_i^u$, the second by $d_i w_i^s$, $a_i w_i^u$.

The parallelograms have become rectangles as the eigenvectors are now orthogonal. We call these rectangles **boxes** and write $\mathcal{R}_i = \{R_i^0, R_i^1\}$ for the resulting partition of $N_i \equiv \mathbb{R}^2/\Lambda_{B_i}$.

Note that in the eigencoordinates the lattice which defines the torus has become a parallelogram lattice. The vectors (d, b) and $(-c, a)$ generate this lattice; these are the columns of B_i^{-1}.

It will be convenient to rotate the eigencoordinates counterclockwise by 90° for the illustrations (see Figures 9, 10, 11, 12). Then the vector w_i^u is $(-1, 0)$ after rotation, while w_i^s is $(0, 1)$, so the rectangle R_i^0 is the right-hand box; changing the orientation of the horizontal axis to agree with the standard coordinates, this box has as base the interval $[0, b_i]$ on the horizontal axis and height the interval $[0, c_i]$ on the vertical axis; the box R_i^1 has base $[-a_i, 0]$ and height $[0, d_i]$.

Therefore the boundary of the partition $\mathcal{R}_i$ is, in the rotated coordinates, the segment $[-a_i, b_i]$ on the x-axis union the segment $[0, \max\{c_i, d_i\}]$ on the y-axis. See Figure 10.

Note that in this picture, our choice of $b_i = 1$, for parity $(+)$ (hence for i even) has this geometrical meaning: the larger of the two boxes (with respect to both height and width) is on the right, while for parity $(-)$ this switches.

In the proof of Proposition 4.1, we encountered the formula

$$A_0 B_0 D_0 = B_1.$$

This formula has the following geometrical interpretation. Consider the related operation

$$B_0 \mapsto A_0 B_0 = B_1 D_0^{-1}.$$

We use the new matrix $A_0 B_0$ to define a new pair of boxes on the same torus, $N_0 \equiv \mathbb{R}^2/\Lambda_{B_0}$, in the same way as just described for B_0. We call this procedure of

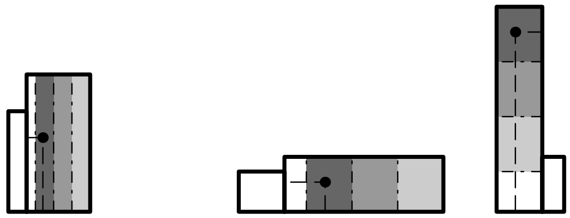

FIGURE 12. From left to right: two-box partition $\mathcal{R}_0$ with $b_0 = 1$, cut to give generating partition $\mathcal{P}_0$ with $n_0 + 1$ elements labelled, reading from left to right, 1111, 0000, 0001, 0011, and 0111 (see the end of §4) ; rescaling by the diagonal matrix so $a_1 = 1$; stacking to get b_1, c_1, d_1 and new two-box partition $\mathcal{R}_1$. The last two pictures overlap, and connected components of $\mathcal{R}_1$ join the rescaled $\mathcal{R}_0$ give the rescaled $\mathcal{P}_0$.

passing from one pair of boxes to the next **box renormalization.** It has a simple algorithmic description, as we shall show:

Proposition 4.5. *Consider the multiplicative family given by a choice of $\langle n\rangle$ and parity. In the eigencoordinate representation (N, g), the pullback of the partition $\mathcal{R}_i$ on N_i to the torus $N_0 \equiv \mathbb{R}^2/\Lambda_{B_0}$ gives the sequence of partitions defined as above from the matrices $B_0, A_0B_0 = B_1D_0^{-1}, A_1A_0B_0 = B_2D_0^{-1}D_1^{-1}, \dots$. Equivalently, the sequence of box pairs is given by box renormalization: first, remove as many copies as possible of the smaller base interval, e.g. for parity $(+)$ that with length a_0 from the right side of the larger, with length $b_0 = 1$. Next, cut the larger box into corresponding pieces and restack these parts above the smaller one, producing the second pair of boxes. For the third pair, begin with the right box, as the parity has changed and that is now shorter. The partition sequence $\mathcal{R}_i$ is a Markov partition sequence for the mapping family (N, g).*

Proof Note that the number of copies of the interval of length a_0 removed is the greatest integer $\left[\frac{b_0}{a_0}\right] = \left[\frac{1}{a_0}\right] = n_0$. And so from equation (2) the box renormalization produces a new pair of boxes, whose sides have new relative lengths given by the matrix A_0B_0. □

4.2. *Producing a generator: the Adler-Weiss method.* This describes a partition sequence with the Markov property. However the $\mathcal{R}_k$ do not generate, i.e. separate points. To produce a generator from $\mathcal{R}$ we employ a method of Adler and Weiss [**AW70**]. We define $\mathcal{P}_k$ to be the partition consisting of the closures of the connected components of $\mathring{\mathcal{R}}_k \vee g_k^{-1}(\mathring{\mathcal{R}}_{k+1})$. As in §3.1, $\mathcal{P} \vee \mathcal{Q}$ denotes the join of two partitions. (By $\mathring{\mathcal{R}}$ we mean the union of the interiors of the elements of the partition $\mathcal{R}$.) We need to take interiors here as we want the new boundary to be the same as the old, and not to have two pieces accidentally glued together.) We shall call $\mathcal{P}_k$ the **connected-component partition** for the multiplicative family.

Proposition 4.6. *The connected-component partition $\mathcal{P} = (\mathcal{P}_k)_{k\in\mathbb{Z}}$ is a generating Markov partition sequence for the multiplicative family (N, g). This passes by* *conjugacy over to a generating Markov partition sequence for the square-torus multiplicative family (M, f). For a chosen parity $(+)$ or $(-)$, the partition of any given component varies continuously with respect to the Hausdorff metric, with respect to $\langle n \rangle \in \Pi_{-\infty}^{+\infty} \mathbb{N}^*$.*

Proof The boundary of $\mathcal{P}_0$ now consists of the two segments $[-a_0, b_0]$ and $[0, d_0 + n_0 c_0] = [0, \lambda_0 b_1]$. Therefore it satisfies the Markov property, as above. Now the elements of the partition $(g_{-n})^n \mathcal{P}_n \vee \dots \mathcal{P}_0 \vee \dots \vee (g_n)^{-n} \mathcal{P}_n$ are also connected sets. Hence their diameters $\to 0$ as $n \to \infty$, so there is at most one point in a given infinite intersection of the interiors. Continuity of the partition boundaries is immediate from the formula for the coordinates of the two boxes, and this passes over to the connected components. □

In the next section we will describe another way to construct this generator, which automatically leads us to natural labels for the partition elements.

5. *An extension of theorems of Adler and Manning to mapping families*

Recall that $SL(2, \mathbb{R}), SL(2, \mathbb{Z})$ are the the groups of (2×2) matrices with real (respectively integer) subgroups with determinant 1. An element $A \in SL(2, \mathbb{Z})$ induces an orientation-preserving automorphism of the two-torus (thought of as the factor group $\mathbb{R}^2 / \mathbb{Z}^2$) by left mutiplication of column vectors. Such an automorphism is *hyperbolic* when it has two eigenvalues λ, λ^{-1} with $\lambda > 1$.

In [**Adl98**] Adler proves a theorem which includes the following: for any orientation-preserving (2×2) hyperbolic toral automorphism A with nonnegative entries there exists a Markov partition such that the map codes as an edge shift using exactly the same matrix. Thus, when

$$A = \begin{bmatrix} p & q \\ r & s \end{bmatrix},$$

there is a generating partition of the torus with $p + q + r + s$ elements, which is coded as an edge shift on two vertices, labelled 0 and 1, such that there are (in the column vector convention) p edges from 0 to itself, q from 1 to 0 and so on; that is, the toral automorphism is coded as an *sft* with edges for symbols, where one symbol can follow another exactly when that edge can follow the other edge in the graph.

Anthony Manning proved a similar result independently. See [**Man02**] and Remark 5.3.

The theorem is remarkable in that it gives a direct link between two apparently completely different uses of the same matrix: to define an edge shift space, and to define a toral map. The existence of *some* coding as an *sft* has been known for a long time (since [**AW70**]) but it seems that this special coding had not been observed except in certain simple cases.

If one considers for example the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

then the *sft* defined by the edge shift is given by a (5×5) $0-1$ matrix. To write this down, we need to choose an order for the edges. Taking the additive labels of Figure 3.5, we order them as $(110, 111, 000, 010, 011)$; this corresponds to the geometric order from left to right of the partition elements, when the unstable leaf is horizontal. The transition matrix (with column convention) is then

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

The utility of the edge shift presentation, as given by the (2×2) matrix A, is evident; this form is much more concise. Adler and Manning's discovery was that when the Markov partition is chosen carefully, this form moreover reveals everything about the dynamics, as the map has exactly the same matrix.

In Adler's proof the generating partition is produced is the following way: he begins with a non-generating partition $\mathcal{R}$ into two parallelograms; these define a single-edge graph with two vertices and corresponding *sft* on two symbols, labelled 0 and 1; the generating partition consists of the connected components of the join of $\mathcal{R}$ with $f^{-1}(\mathcal{R})$; the edge graph for this is given by replacing each single edge by the number of connected components in the corresponding intersection. Thus, for the matrix

$$A = \begin{bmatrix} p & q \\ r & s \end{bmatrix},$$

there are q connected components in $f^{-1}(R_0) \cap R_1$ and so on.

In this section we shall show how the Adler-Manning coding extends to sequences of matrices. In the process we discover a new way of understanding his result.

We will conclude:

Theorem 5.1. *For any nontrivial sequence $(A_i)_{i\in\mathbb{Z}}$ with $A_i \in SL(2,\mathbb{N})$, the mapping family determined by this sequence is an Anosov family. There exists a generating Markov partition sequence which codes the family as the Bratteli edge diagram with transitions given by the same sequence (A_i), and such that, moreover, the generating partitions consist of connected components of sucessive joins from a two-box partition sequence, with the number of multiple edges between two successive vertices equal to the corresponding number of components.*

Here *nontrivial* means not eventually either lower or upper triangular at $(+)$ or $(-)$ infinity.

The idea is as follows: we define the coding first for additive sequences, where the geometric explanation comes directly from box renormalization and is there completely transparent; then we show that these codings behave well with respect to the operation of gathering, and that connected components correspond exactly to different names along the additive factorization. The final step is to remember from Lemma 3.11 that any sequence of nonnegative matrices has a canonical factorization in terms of the additive generators, allowing us to apply the previous result.

5.1. *Additive and linear families.* We call a sequence $F_i \in SL(2,\mathbb{N})$, for $i \in \mathbb{Z}$, a (nonnegative) **linear sequence**. A linear sequence defines a mapping familiy by its action on column vectors, with spaces the square torus $\mathbb{R}^2/\mathbb{Z}^2$, i.e. with the standard Euclidean metric inherited from the plane. We call this the **linear family** determined by $(F_i)_{i\in\mathbb{Z}}$.

Now we treat Example 6 from §2 in detail. When the linear sequence is such that each matrix F_i either equals one of the **additive generators**

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

we call this an **additive family**; we say these generators have parity $(+)$ and $(-)$ respectively. We say an additive familiy is **nontrivial** if the matrices change parity infinitely often at both $+\infty$ and $-\infty$.

Every nontrivial additive family gathers to a multiplicative family, collecting neighboring upper and lower triangular matrices in the obvious way, possibly after a shift to have the change of parity occur between times -1 and 0. The converse holds also, since of course the mutiplicative family factors as powers of the generators, giving a (nontrivial) additive family.

More generally, by Lemma 3.11, any nonnegative linear sequence factors uniquely as an additive family, as in part (i) of Corollary 3.12.

(Removing the trivial additive families i.e. the eventually constant sequences corresponds exactly to removing the rational directions for the stable and unstable foliations, and is a countable set.)

5.2. *A Markov Partition and symbolic dynamics for the additive family.* We describe next a Markov partition sequence for a nontrivial additive family. Here we have a generating sequence from the outset, without needing to take connected components of joins. This construction will then give a new way of getting generators for the multiplicative families.

We begin with a sequence $\langle n\rangle = (\dots n_0 n_1 \dots) \in \Pi_{-\infty}^{+\infty}\mathbb{N}^*$ and assume for simplicity as above that the additive sequence (A_i) changes parity between coordinates -1 and 0. The additive family then gathers to the multiplicative family which we shall now write as $(\widetilde{A}_i)$.

Corresponding to the map

$$\widetilde{B}_i \mapsto \widetilde{A}_i\widetilde{B}_i\widetilde{D}_i = \widetilde{B}_{i+1},$$

which describes algebraically the operation of multiplicative box renormalization, we now have the equation

$$B_k \mapsto A_k B_k D_k = B_{k+1},$$

which corresponds to the renormalization being done additively, step by step; that is, at each stage i we remove and stack one box, repeating this n_i times. The diagonal matrix D_k is defined so as to renormalize the boxes, keeping the largest width always equal to 1.

The generating Markov partitions now are easy to describe: they are for each component M_k just the pair of boxes. Let us write this partition sequence as $\mathcal{P}_k$; we shall compare this to the multiplicative partition already defined, which we shall now write as $\widetilde{\mathcal{P}}_k$. We have:

Lemma 5.2. *The sequence $\mathcal{P}_k$ generates for the additive family.*

Proof We give two proofs. The first is direct: note that the intersections of the interiors of the partition elements after being pulled back by the appropriate maps are always connected, while clearly the widths (and the heights for negative time) go to 0. Hence they separate points i.e. generate.

The second proof is more abstract: it begins with the generating partition for the multiplicative family given in Proposition 4.6. Next, by Lemma 5.4 which follows, this partition, the connected component partition, is equal to augmented gathered partition; finally, by the converse part of Proposition 3.1, the partition sequence $\mathcal{P}$ therefore generates for the additive family. □

It turns out that an especially nice symbolic dynamics will result from labelling the two boxes of $\mathcal{P}_k$ in the following way:

Theorem 5.3. *For the additive family (M, f) given by the additive sequence $A = (A_i)_{i\in\mathbb{Z}}$ acting on column vectors, then if we assign to the left box the symbol 1 and to the right-hand box the symbol 0, the corresponding nonstationary sft defined by the above coding has for its defining transition matrices the same sequence of matrices. That is, the nsft is Σ_A.*

Proof Let us suppose that the parity of the additive family given by A_i is $(+)$; thus we have $A_i = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. As in Figure 11, the wider box is on the right, and this has label 0; the pullback of the partition at time $i+1$ meets this exactly in the renormalized boxes defined by the matrix B_{i+1}, and so the points of box 0 at time i can either be in box 0 or 1 at time $i+1$, while those in box 1 will necessarily next be in box 1. Hence the column transition matrix is $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, exactly the same matrix. □

5.3. *Gathering, connected components, and symbolic dynamics.* The connection with the generating partition sequence $\widetilde{\mathcal{P}}_k$ of $\widetilde{M}_k$, for the multiplicative family, is given by:

Lemma 5.4. *For the multiplicative family $(\widetilde{M}, \tilde{f})$, the connected-component partition $\widetilde{\mathcal{P}}_i$ is the augmented gathered partition of the partition $\mathcal{P}$, taken along the subsequence of times (n_i). Thus, $\widetilde{\mathcal{P}}_i$ is equal to the join of the pullbacks to the component $\widetilde{M}_i = M_{n_i}$ of the (n_i+1) partitions $\mathcal{P}_k$ into two boxes, for $n_i \le k \le n_{i+1}$.*

Proof This is clear from the definitions; see Figure 12. Thus, e.g. for $n_0 = 1$,

$$\widetilde{\mathcal{P}}_0 = \mathcal{P}_0 \vee (g_0)^{-1}(\mathcal{P}_1)$$

which has three elements. □

In fact something similar is true in much more generality:

Lemma 5.5. *Let (A_k) be an additive family, and let $\mathcal{P}_k$ be its two-box (generating) partition. Let (m_i) be* any *increasing subsequence with $m_i \in \mathbb{Z}$ and define $\widehat{A}$ to be the family gathered along that subsequence of times. Define $(\widehat{\mathcal{P}}_i)$ to be the connected-component partition sequence for (m_i); that is, the elements of $\widehat{\mathcal{P}}_i$ are the connected components of the join of $\mathcal{P}_{m_i}$ with the pullback of $\mathcal{P}_{m_{i+1}}$ by the composed map. Then $(\widehat{\mathcal{P}}_i)$ is equal to the augmented gathered partition taken along the subsequence (m_i) of the additive partition $\mathcal{P}$.*

Proof We observe that in the eigencoordinate version of the multiplicative family, the boundaries of the two boxes at time m_i are formed by a union of two segments in the x and y axes, each containing the point $(0, 0)$ and forming a T shape. Pulling back the corresponding two intervals from the future time m_{i+1} gives a subinterval of the unstable interval and a segment which contains the stable interval. We are to show that the connected components of the complement of these four closed intervals in the torus is the same as the partition given by the join of all box-pairs for all the times up to and including time m_{i+1}, which gives the augmented gathered partition. But the two pullback segments respectively are contained in and contain all of these in-between pulled back segments, as they are in the expanding and contracting leaves. Hence no new boundaries are added in the augmented gathering (it gives exactly the same vertical striping of the two original boxes), verifying the claim. □

Remark 5.1. Note that for the additive family, the obvious two-box partition sequence already generates, with no need to take connected components.

We remark that the additive partition sequence is slightly more efficient than the multiplicative one. This one sees from the expression in Lemma 5.4; note that in defining $\widetilde{\mathcal{P}}_i$ there is a bit of redundancy, as each partition $\mathcal{P}_k$ for $k = \dots, n_0, n_0 + n_1, \dots$ gets included twice: this is exactly the difference between the gathered and augmented gathered partitions of $\mathcal{P}$ along the sequence (n_i).

The lemma shows the relationship between the connected-component method and augmented gathering along the additive family, and hence the connection with telescoping of Bratteli diagrams, by Proposition 3.10, leading to the following:

Theorem 5.6. *Given (F_i) in $SL(2, \mathbb{N})$, the linear family has a Markov partition which codes it as a Bratteli edge diagram with additive labels determined by the same sequence (F_i).*

That is, the unique additive sequence (A_j) determined by the sequence (F_i) from Lemma 3.11 serves two purposes: it gives the unique additive dispersal of the family (F_i), and for this dispersal, the two-box partition sequence generates and is coded by the sequence (A_j); microscoping this single-edged diagram gives a multiple-edged diagram on two symbols with additive labels, and with the number of edges given by the sequence (F_i). Taking the connected components of pullbacks of the two-box partitions for (F_i) gives a generating Markov partition, which is the augmented gathering of the additive partition taken along that subsequence. □

We recall from Definition 3.14 that a (2×2) sequence of nonnegative integer matrices (F_i) determines an *nsft* called Σ_F. We write (M, f) for the mapping family on the torus defined by (F_i) acting on columns.

Corollary 5.7. *The coding by the* nsft Σ_F *of the linear family* (F_i) *defines a topological conjugacy* π *from the mapping family* (Σ_F, σ) *to the family* (M, f).

Remark 5.2. It is natural to ask when this topological conjugacy is in fact uniform (from the word metric on the combinatorial space to the Eucldean metric on the torus). One can show that for multiplicative family it is always uniform, in fact is uniformly Lipschitz; for additive families this is true if and only if the sequence n_i is bounded above. Furthermore, uniformity is preserved by gatherings. We shall present these results in a later paper.

Example. Because a positive linear family (F_i) has a unique additive sequence, the elements of the generating Markov partition described above inherit natural labels. We describe the labels for the case of the multiplicative family, see Figure 12. For k even, taking for example $k = 0$ and $n_0 = 3$, as in the Figure, we have that the elements of the partition $\mathcal{P}_0$, reading from left to right, are labelled as follows: the smaller (leftmost) of the two boxes is labelled 1111, and the larger box is cut into pieces labelled $0000, 0001, 0011$, and 0111 respectively; taking e.g. $n_1 = 3$, for the partition $\mathcal{P}_1$, with parity reversed, we have that its smaller box, now on the right, is labelled 0000, and $0, 1, 2, 3$ are labelled (from right to left) $1111, 1000, 1100$, and 1000. We shall see the usefulness of these labels in a later paper.

Remark 5.3. In defining shift spaces, one can let the transition matrices act "on columns" as we do here, or "on rows" (the more usual choice, in Markov chain theory; for Bratteli diagrams, the column convention is often used). Similarly, matrices can be chosen to act on the torus by multiplying row vectors or columns (the more common convention in dynamical systems theory). Thus, it is usual in dynamics to have a mixture of the conventions. Ordinarily this makes little difference, but when taking a sequence of matrices, as we do here, making a consistent choice is notationally important, as otherwise the order of matrix multiplication is reversed. And even for a single Anosov map, consistency is crucial if one wants to get the nice Adler-Manning-type coding. Thus in Adler [**Adl98**] the matrices act on the right on row vectors, as shift space are defined in the usual (row) way. Once we know this result, then an immediate corollary is that if instead the column vector convention is used both for the maps and transitions, one again gets the same matrix for both; whereas if a mixed convention is chosen, the coding matrix will be the *transpose* of that which gives the map.

Anthony Manning (personal communication) in the early 1980's, independently of Adler, discovered such a coding of Anosov toral automorphisms, using a mixed convention and thus ending up with the transpose matrix.

In fact we rediscovered this by accident; unaware of Adler or Manning's work, and studying the case of additive families, we had initially used a mixed convention but with switched labels 0 and 1 on the boxes. In this case the two operations cancel, yielding exactly the same sequence of matrices. Furthermore this extends to multiplicative families by gathering, providing a nice theorem. However when one gathers further-thus mixing parities- then the coding gets hopelessly confused. This is due to the noncommutativity of matrix multiplication- if the conventions are mixed, in one case the order of multiplication will be reversed, yielding a coding matrix which looks nothing like the original.

Upon encountering Adler's paper, we saw that by switching the labels, we could almost recover his result- and extend it to arbitrary linear families as above- except that we kept getting the transpose sequence. Finally the importance of a consistent choice became clear, leading to this version.

We mention that [**Adl98**] and [**Man02**] treat cases not considered here, nonpositive (2×2) matrices and certain positive $(n \times n)$ matrices respectively.

5.4. *Application: An Adler-Manning type coding for a random dynamical system* As an illustration of the above, consider the following skew product transformation, which can be thought of as a *random dynamical system.*

Write $\Omega = \Pi_{-\infty}^{\infty}\{0,1\}$ with left shift map σ for the Bernoulli shift space with invariant measure μ the product measure $(1/2, 1/2)$; that is, each symbol has equal probablity and is chosen independently of the others.

Consider the product space $\widehat{\Omega} = \Omega \times \mathbb{T}^2$ where $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ is the torus. Choose two matrices A_0, A_1 with nonnegative integer entries and determinant 1, e.g. $A_0 = \begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix}$ and $A_1 = \begin{bmatrix} 4 & 1 \\ 11 & 3 \end{bmatrix}$. Define now a skew product transformation $\hat{\sigma}$ acting on $\widehat{\Omega}$ with measure $\hat{\mu} = \mu \times \lambda$ where λ is Lebesgue measure on the torus, by: $\hat{\sigma} : (\omega, v) \mapsto (\sigma(\omega), A_{\omega_0} v)$. We have:

Theorem 5.8. *There exists a continuous choice of Markov partitions into two parallelograms on each torus fiber above each point $x \in \Sigma$, such that the transformation is coded as a random subshift of finite type, by a Bratteli edge diagram given by exactly the same random product of matrices.*

The proof is immediate from Theorem 5.6, and indeed has nothing to do with the choice of measure on Ω, since for *any* (not almost any) choice of ω, the sequence of matrices $\bar{A}_\omega = \dots, A_{\omega_{-1}}, A_{\omega_0}, A_{\omega_1} \dots$ defines a nontrivial linear family on the torus. The corresponding partition sequence $\mathcal{P}_\omega = \dots, \mathcal{P}_{\omega_{-1}}, \mathcal{P}_{\omega_0}, \mathcal{P}_{\omega_1} \dots$ is a Markov partition sequence, and codes the linear family as the *nsft* $\Sigma_{\bar{A}_\omega}$. The partition $\mathcal{P}_{\omega_0}$ varies continuously with respect to ω, in the Hausdorff metric on partition boundaries, by Proposition 4.6.

Putting these together for all $\omega \in \Omega$ gives a **random Markov partition** and corresponding **random subshift of finite type** in the sense of Bogenschütz [**Bog93**].

We make this remark on the measure theory: by well-known results, the skew product is ergodic, and as we shall show elsewhere (see also [**AF01**]), the measure on the fiber over a given point has a nice combinatorial description as a Shannon-Parry measure for the nonstationary subshift of finite type over that point.

References

[Adl98] Roy L. Adler. Symbolic dynamics and Markov partitions. *Bulletin of the American Mathematical Society*, 35(1):1–56, 1998.

[AF01] Pierre Arnoux and Albert M. Fisher. The scenery flow for geometric structures on the torus: the linear setting. *Chinese Ann. of Math.*, 4:427–470, 2001.

[AF02a] Pierre Arnoux and Albert M. Fisher. Anosov families, random dynamics and the scenery flow. Technical report, 2002. preprint.

[AF02b] Pierre Arnoux and Albert M. Fisher. The Ruelle-Perron-Frobenius Theorem for nonstationary and random subshifts, 2002.

[Arn94] Pierre Arnoux. Le codage du flot géodésique sur la surface modulaire. *l'enseignement Mathématique*, 40:29–48, 1994.

[AW70] Roy L. Adler and Benjamin Weiss. Similarity of automorphisms of the torus. *Memoirs of the American Mathematical Society*, 98:1–43, 1970.

[Bak95a] V.I. Bakhtin. Random processes generated by a hyperbolic sequence of mappings I. *Russian Acad. Sci. Izv. Math.*, 44(2):247–279, 1995.

[Bak95b] V.I. Bakhtin. Random processes generated by a hyperbolic sequence of mappings II. *Russian Acad. Sci. Izv. Math.*, 44(2):617–627, 1995.

[BG95] T. Bogenschütz and M. Gundlach. Ruelle's transfer operator for random subshifts of finite type. *Ergodic Th. and Dynam. Sys.*, 15:413–447, 1995.

[BH94] Mike Boyle and David Handelman. Entropy versus orbit equivalence for minimal homeomorphisms. *Pacific Jour. Math.*, 164(1):1–13, 1994.

[Bog93] T. Bogenschütz. Entropy, pressure and a variational principle for random dynamical systems. *Random and Computational Dynamics*, 1:99–116, 1992-93.

[Bow75] Rufus Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms, SLN 470.* Springer Verlag, 1975.

[Bow77] Rufus Bowen. On Axiom A diffeomorphisms. *Conference Board Math. Sciences*, 35, 1977.

[DHS99] Fabien Durand, Bernard Host, and Christian Skau. Substitution dynamical systems, Bratteli diagrams and dimension groups. *Ergod. Th. and Dynam. Sys.*, 19:953–993, 1999.

[Dur98] Fabien Durand. A characterization of substitutive sequences using return words. *Discrete Mathematics*, 179:89–101, 1998.

[ES80] Edward G. Effros and Chao-Liang Shen. Approximately finite C^*-algebras and continued fractions. *Indiana University Math. Journal*, 29(2):191–204, 1980.

[FB88] P. Ferrero and B.Schmitt. Produits aléatoires d'operateurs matrices de transfert. *Probability Theory*, 79:227–248, 1988.

[Fis92] Albert M. Fisher. Integer Cantor sets and an order-two ergodic theorem. *Ergod. Th. and Dynam. Sys.*, 13:45–64, 1992.

[Fis03] Albert M. Fisher. Small-scale Structure and Randomness: the Scenery Flow in Dynamics, Geometry and Probability. http://ime.usp.br/∼afisher, March 2003. thesis for livre docencia, Univ. of São Paulo.

[Fis04] Albert M. Fisher. Small-scale structure via flows. In *Progress in Probability*, pages 59–78. Birkhäuser, 2004. Conference Proceedings, Fractal Geometry and Stochastics III, Friedrichroda, Germany, March 2003.

[Fra82] John Franks. *Homology and dynamical systems*, volume 49 of *CBMS*. AMS, 1982.

[Fur63] Hillel Furstenberg. A Poisson formula for semi-simple Lie groups. *Ann. of Math.*, 77:335–386, 1963.

[Fur71] Hillel Furstenberg. *Random walks and discrete subgroups of Lie groups*, volume 1, pages 3–63. Dekker, New York, 1971.

[GJ98] Richard Gjerde and Orjan Johansen. Bratteli-Vershik models for Cantor minimal systems. Technical report, 1998.

[Hun74] Thomas W. Hungerford. *Algebra*. Holt, Rinehardt and Winston, 1974.

[Kif00] Yuri Kifer. Random f-expansions. *Proceedings Symp. Pure Math.*, 00, 2000. to appear.

[Led77] F. Ledrappier. Mesures d'equilibre d'entropie compétement positive. *Astérisque*, 50:251–272, 1977.

[LM95] Douglas Lind and Brian Marcus. *Symbolic dynamics and coding*. Cambridge University Press, 1995.

[Man02] Anthony Manning. A Markov partition that reflects the geometry of a hyperbolic toral automorphism. *Trans. AMS*, 354:2865–2895, 2002.

[Pin91] Alberto Pinto. *Convergence of Renormalisation and Rigidity of Dynamical Systems*. PhD thesis, Warwick University, 1991. advisor: David Rand.

[PP90] William Parry and Mark Pollicott. Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics. *Astérisque*, 187-188:1–268, 1990.

[PR95a] Alberto Pinto and David A. Rand. Classifying $\mathcal{C}^{1+}$ structures on dynamical fractals: 1. The moduli space of solenoid functions for Markov maps on train tracks. *Ergod. Th. and Dynam. Sys.*, 15:697–734, 1995.

[PR95b] Alberto Pinto and David A. Rand. Classifying $\mathcal{C}^{1+}$ structures on dynamical fractals: 2. Embedded trees. *Ergod. Th. and Dynam. Sys.*, 15:969–992, 1995.

[Ran88] David A. Rand. Global phase space universality, smooth conjugacies and renormalization: I. the $\mathcal{C}^{1+\alpha}$ case. *Nonlinearity*, 1:181–202, 1988.

[Rue72] David Ruelle. On the Use of "Small External Fields" in The Problem of Symmetry Breakdown in Statistical Mechanics. *Annals of Physics*, 69:364–374, 1972.

[Shu87] Michael Shub. *Global Stability of Dynamical Systems*. Springer Verlag, 1987.

[Sta88] Jaroslav Stark. Smooth conjugacy and renormalization for diffeomorphisms of the circle. *Nonlinearity*, 1:541–575, 1988.

[Ver94] Anatoly M. Vershik. Locally transversal symbolic dynamics. *Algebra i Analiz*, 6:94–106, 1994. in Russian.

[VL92] A. M. Vershik and A. N. Livshits. Adic models of ergodic transformations, spectral theory, substitutions, and related topics. *Advances in Soviet Mathematics*, 9:185–204, 1992.

[You86] Lai-Sang Young. Stochastic stability of hyperbolic attractors. *Ergod. Th. and Dynam. Sys.*, 6:311–319, 1986.