Almost topological classification of finite-to-one factor maps between shifts of finite type

ROY ADLER[1], BRUCE KITCHENS[2] AND BRIAN MARCUS[3]

[1,2] *IBM Research Laboratory, Yorktown Heights, New York* 10598, *USA:* [3] *IBM Research Laboratory, San Jose, California* 95193, *USA* and *University of North Carolina, Chapel Hill, North Carolina* 27514, *USA*

Abstract. We classify finite-to-one factor maps between shifts of finite type up to almost topological conjugacy.

0. *Introduction*

In [2], shifts of finite type (SFT) were classified up to almost topological conjugacy. The purpose of this paper is to classify finite-to-one continuous factor maps between shifts of finite type. We consider two equivalence relations. The first is very strong: two maps are equivalent if they have the same range shift and they differ by a special kind of almost continuous change of variables in the domain; for a fixed range shift and a fixed generic cardinality of the fibre of the factor maps (a natural invariant), we get infinitely many equivalence classes. The second (more natural) equivalence relation allows the same kind of change of variables in the range (as well as in the domain) and gives only finitely many equivalence classes (for a fixed entropy class and generic cardinality of the fibre). For the latter classification, we reduce the problem to a group action problem which was solved by us in [1]. Our results are completely analogous to and were inspired by those of D. Rudolph [10] in the measure-theoretic category (i.e. classification of finite-to-one factor maps between Bernoulli shifts). Our work is basically a more concrete version of Rudolph's work. In particular, let T be a Bernoulli shift whose measure-theoretic entropy $(\log(\lambda))$ is the same as the topological entropy of an aperiodic SFT; then for any finite-to-one factor map of T, we construct (theorem 4.3) an equivalent (in the measure-theoretic sense) factor map $\pi: \Sigma_A \to \Sigma_B$, where Σ_B is an *arbitrary* aperiodic SFT of entropy $\log(\lambda)$ and Σ_A is some aperiodic SFT of entropy $\log(\lambda)$. (In general, Σ_A cannot be chosen arbitrarily.)

It would perhaps have been most natural for us to classify our factor maps with respect to continuous changes of variables (i.e. topological conjugacy) since, after all, our factor maps are continuous themselves. But the simplest case of this, when the factor maps involved are identity maps, reduces to the classical topological conjugacy problem for shifts of finite type—which is not yet solved satisfactorily.

For background, we basically refer to [2]. We use the notation Σ_A for an SFT defined by 0-1 transition matrix A, σ_A for the shift, and L_A for the alphabet (or

states or symbols) used. The notation $i \rightarrow j$ indicates an allowable transition from symbol i to symbol j. We assume that the reader is familiar with the definitions of irreducible (ergodic) and aperiodic (mixing) and also the graph-theoretic description of SFT's.

We shall be interested in *continuous factor maps* (i.e. onto, continuous, shift-commuting maps) $\pi: \Sigma_A \rightarrow \Sigma_B$ between aperiodic SFT's. It is well known that every such map can be represented as a sliding block map which depends on finitely many coordinates. We shall frequently use the fact that every continuous factor map can be recoded to a one-block map (i.e. a sliding block map depending on only one coordinate:

$$\pi(\cdots x_{-1}x_0x_1\cdots) = \cdots \pi^*(x_{-1})\pi^*(x_0)\pi^*(x_1)\cdots,$$

where π^* is some map $\pi^*: L_A \rightarrow L_B$).

The following is a basic fact:

PROPOSITION 0.1 ([3, theorems 3.3 and 5.7]). *Let* $\pi: \Sigma_A \rightarrow \Sigma_B$ *be a continuous shift-commuting map between irreducible SFT's. Then, any two of the following conditions imply the third condition*:
  (1) π *is onto*;
  (2) $h(\Sigma_A) = h(\Sigma_B)$ *(where h denotes topological entropy)*;
  (3) π *is finite-to-one (i.e. every point has finitely many inverse images)*.

From the point of view of ergodic theory, every finite-to-one continuous factor map is k-1 a.e. for some k. (Here 'a.e.' means a.e. with respect to any fully supported ergodic measure, e.g. the unique measure of maximal entropy.) But more is true, as the following indicates:

PROPOSITION 0.2. *Let* $\pi: \Sigma_A \rightarrow \Sigma_B$ *be a finite-to-one continuous factor map between irreducible SFT's. Then there is a positive integer k such that*:
(a) *Every point in* Σ_B *has at least k* mutually separated *inverse images (if π is a one-block map, then a pair of points* $x, y \in \Sigma_A$ *is* mutually separated *if for all i, $x_i \neq y_i$)*.
(b) *Every doubly transitive point in* Σ_B *(i.e. a point whose forward and backward orbits are dense) has exactly k inverse images, whence, by* (a), *all its inverse images are mutually separated*.

For the proof of the above, we refer the reader to [3, theorem 6.5]; see also [5]. We also have:

PROPOSITION 0.3. *A continuous factor map between irreducible SFT's is finite-to-one if and only if it is k-1 a.e. for some k.*

Proof. (only if) Proposition 0.2(b) is a strong version of this.
  (if) follows from application of k-1 a.e. to the measure of maximal entropy to obtain condition (2) of proposition 0.1. □

The following characterizes maps which are constant-to-one.

PROPOSITION 0.4 ([8, Theorem 6.3], see also [5]). *A continuous factor map is k-to-1* everywhere *for some k if and only if it is right and left closing. (Right (left) closing means that the map does not identify two negatively (positively) asymptotic points.)*

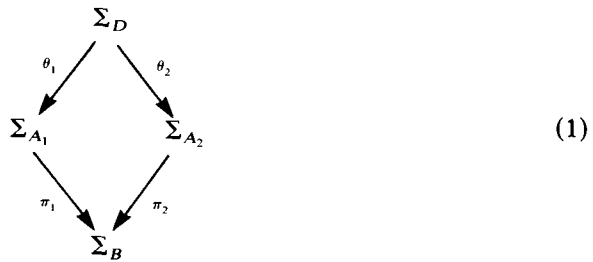For some illuminating examples, we refer the reader to those constructed by P. Shields [**10**, p. 258].

1. *The strong relation*

In this section, Σ_{A_1}, Σ_{A_2} and Σ_B are aperiodic SFT's all with the same topological entropy. Let $\pi_1 \colon \Sigma_{A_1} \to \Sigma_B$ and $\pi_2 \colon \Sigma_{A_2} \to \Sigma_B$ be continuous factor maps.

Definition. We say that π_1 and π_2 are *almost topologically conjugate* over Σ_B (A.T.C. over Σ_B) if there exists an aperiodic SFT Σ_D and 1–1 a.e. continuous factor maps

$$\theta_i \colon \Sigma_D \to \Sigma_{A_i} \qquad i = 1, 2$$

such that the diagram:

$$\begin{array}{ccc} & \Sigma_D & \\ {\scriptstyle \theta_1} \swarrow & & \searrow {\scriptstyle \theta_2} \\ \Sigma_{A_1} & & \Sigma_{A_2} \\ {\scriptstyle \pi_1} \searrow & & \swarrow {\scriptstyle \pi_2} \\ & \Sigma_B & \end{array} \tag{1}$$

commutes.

Remark. It follows that Σ_D has the same entropy as Σ_{A_1} and Σ_{A_2} and so the maps θ_1 and θ_2 must be finite-to-one (by proposition 0.1).
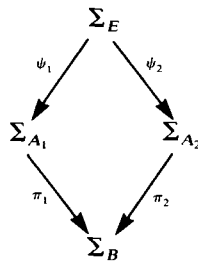
Definition. The *fibred product* over π_1 and π_2 is the SFT

$$\Sigma_E = \{(x, y) \in \Sigma_{A_1} \times \Sigma_{A_2} \colon \pi_1(x) = \pi_2(y)\}$$

together with the factor maps

$$\psi_1 \colon \Sigma_E \to \Sigma_{A_1}, \qquad \psi_1(x, y) = x;$$
$$\psi_2 \colon \Sigma_E \to \Sigma_{A_2}, \qquad \psi_2(x, y) = y.$$

Note that the diagram

$$\begin{array}{ccc} & \Sigma_E & \\ {\scriptstyle \psi_1} \swarrow & & \searrow {\scriptstyle \psi_2} \\ \Sigma_{A_1} & & \Sigma_{A_2} \\ {\scriptstyle \pi_1} \searrow & & \swarrow {\scriptstyle \pi_2} \\ & \Sigma_B & \end{array}$$

commutes. Also, each ψ_i must be finite-to-one.

THEOREM 1.1. *The factor maps π_1 and π_2 are A.T.C. over Σ_B if and only if there is an irreducible component $\Sigma_{E'}$ of the fibred product, Σ_E, with maximal entropy (in Σ_E) such that the restrictions $\psi_1/\Sigma_{E'}$ and $\psi_2/\Sigma_{E'}$ are 1–1 a.e.*

Proof. (if) Set $\theta_i = \psi_i$, $i = 1, 2$ and observe that the θ_i must be onto since $\Sigma_{E'}$ has full entropy and ψ_i are finite-to-one (proposition 0.1). Also, $\Sigma_{E'}$ is aperiodic since ψ_1 is 1–1 a.e. and Σ_{A_1} is aperiodic.

(only if) Let θ_1 and θ_2 be the maps that provide the A.T.C. over Σ_B. Now, define the map:

$$\varphi : \Sigma_D \to \Sigma_E, \qquad \varphi(z) = (\theta_1(z), \theta_2(z)).$$

This is well-defined since the diagram (1) commutes. Since Σ_D is aperiodic, so is its image $\varphi(\Sigma_D)$.

Since the θ_i are assumed to be 1-1 a.e., they are finite-to-one by proposition 0.3. Thus, φ is finite-to-one. So, $\varphi(\Sigma_D)$ has the same entropy as Σ_D, which is the same entropy as Σ_{A_i} and Σ_E (since the ψ_i and θ_i are finite-to-one and onto). So, $\Sigma_{E'} \equiv \varphi(\Sigma_D)$ has maximal entropy in Σ_E; since $\Sigma_{E'}$ is also aperiodic, it follows that $\Sigma_{E'}$ is an irreducible component of maximal entropy in Σ_E.

Moreover the diagram:
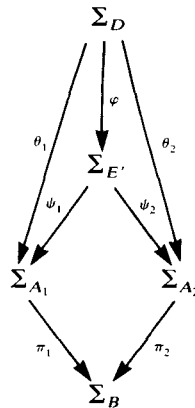
commutes. Now, since θ_i are 1-1 a.e. $i = 1, 2$, it follows that $\psi_i|\Sigma_{E'}$ are 1-1 a.e. $i = 1, 2$ as desired.   □

Remarks. (i) The classification given above is completely effective since: (1) the construction of the fibred product is constructive (assuming by recoding that π_1 and π_2 are 1-block maps, then Σ_E is the SFT with alphabet

$$L_E = \{(a_1, a_2) \in L_{A_1} \times L_{A_2} : \pi_1(a_1) = \pi_2(a_2)\},$$

and transitions $(a_1, a_2) \to (a_1', a_2')$ if and only if $a_1 \to a_1'$ and $a_2 \to a_2'$); (2) the determination of irreducible components is constructive; and (3) there is a finite procedure for deciding whether or not a factor map is 1-1 a.e. (see [4, 3.4]).

(ii) The generic cardinality of the fibre for a finite-to-one continuous factor map (which always exists (proposition 0.2)), is clearly an invariant of A.T.C. over Σ_B (i.e. if π_1 and π_2 are A.T.C. over Σ_B and π_1 is k_1-to-1 a.e. and π_2 is k_2-to-1 a.e., then $k_1 = k_2$).

(iii) From a measure-theoretic point of view, the factor map $\pi_1 : \Sigma_{A_1} \to \Sigma_B$ decomposes Σ_{A_1} as a skew-product over Σ_B with k-point fibres (where k is the generic cardinality of the fibre): namely, Σ_{A_1} can be written as $\Sigma_B \times \{1, \ldots, k\}$ and the shift σ_{A_1} can be represented as

$$\sigma_{A_1}(x, i) = (\sigma_B(x), f_1(x)(i)), \tag{2}$$

where $f_1: \Sigma_B \to S_k$ is a measurable (in fact, almost continuous) map into the symmetric group on k letters. Here, the measure structure is given by the measure of maximal entropy, and almost continuous means continuous on some subset of full measure. Similarly, we get a function $f_2: \Sigma_B \to S_k$ for the map π_2.

Now, one can easily show that if π_1 and π_2 are A.T.C. over Σ_B, then there exists a measurable (in fact, almost continuous) map $h: \Sigma_B \to S_k$ such that the cohomology equation:

$$f_1 = (h \circ \sigma_B) \cdot f_2 \cdot h^{-1} \tag{3}$$

holds a.e., (where for $x \in \Sigma_B$, $h^{-1}(x)$ denotes the inverse permutation of $h(x)$) (see [10, p. 256]).

(iv) If π_1 and π_2 are k-to-1 *everywhere*, then f_1 and f_2 can (in (2)) be chosen continuous and in fact depend on only finitely many coordinates (see [8]). It follows then that the solution h to (3) (if it exists) must depend on only finitely many coordinates as well, and in fact (3) holds everywhere. The proof of this is similar to the results [6, theorem 1 and remark 2] [9, 2.42] and is deferred to the appendix. The continuity of h then yields:
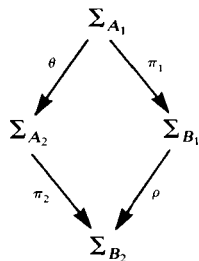
PROPOSITION 1.2. *If two k-to-1 everywhere extensions π_1 and π_2 of Σ_B are A.T.C. over Σ_B, then in fact they are topologically conjugate over Σ_B (i.e. there is a topological conjugacy $\bar{h}: \Sigma_{A_1} \to \Sigma_{A_2}$ such that $\pi_1 = \pi_2 \circ \bar{h}$).*

From this, one can use the form (2) to construct, for each SFT Σ_B and integer $k \geq 2$, infinitely many inequivalent (in the sense of A.T.C. over Σ_B) k-to-1 everywhere extensions of Σ_B. (Just play with the zeta functions of the domains.)

2. *The weak relation*: *reduction to constant-to-one maps*
In this and the next section, Σ_{A_1}, Σ_{A_2}, Σ_{B_1} and Σ_{B_2} are aperiodic SFT's of the same entropy. Let $\pi_1: \Sigma_{A_1} \to \Sigma_{B_1}$ and $\pi_2: \Sigma_{A_2} \to \Sigma_{B_2}$ be continuous onto factor maps (hence finite-to-one).

Definition. We say that π_2 is an *almost topological factor* of π_1 (and π_1 is an almost *topological extension* of π_2) if there exist 1–1 a.e. continuous factor maps $\theta: \Sigma_{A_1} \to \Sigma_{A_2}$ and $\rho: \Sigma_{B_1} \to \Sigma_{B_2}$ such that the diagram:

$$
\begin{array}{ccc}
 & \Sigma_{A_1} & \\
{}^{\theta}\swarrow & & \searrow{}^{\pi_1} \\
\Sigma_{A_2} & & \Sigma_{B_1} \\
{}^{\pi_2}\searrow & & \swarrow{}^{\rho} \\
 & \Sigma_{B_2} &
\end{array}
$$

commutes.

Definition. We say that π_1 and π_2 are *almost topologically conjugate* (A.T.C.) if they have a common almost conjugate extension (cf. [2]).

We show, in this section, that we may as well assume that our factor maps are *constant*-to-one (i.e. k-to-1 *everywhere* for some k).

THEOREM 2.1. *Let Σ_A and Σ_B be aperiodic* SFT's *and let $\pi: \Sigma_A \to \Sigma_B$ be a k-to-1 a.e. continuous factor map. Then there is an almost topological extension of π which is k-to-1 everywhere.*

Proof. We may assume that π is a 1-block map by the usual recoding argument.

Define $\Sigma_{B'}$ as follows: the alphabet of $\Sigma_{B'}$ is the set of all distinct *un*ordered k-tuples of symbols from L_A with the same π-image:

$$L_{B'} = \{\{a_1, \ldots, a_k\}: a_1, \ldots, a_k \in L_A \text{ are distinct and } \pi(a_1) = \pi(a_2) = \cdots = \pi(a_k)\}$$

and transitions

$$\{a_1, \ldots, a_k\} \to \{a'_1, \ldots, a'_k\}$$

if and only if there is a permutation $q \in S_k$ such that for each $i = 1, \ldots, k$

$$a_i \to a'_{q(i)} \quad \text{in } \Sigma_A.$$

Let $\rho: \Sigma_{B'} \to \Sigma_B$ be defined in the obvious manner; namely:

$$\rho(\{a_1, \ldots, a_k\}) = \pi(a_1).$$

Then Coven and Paul ([4, 3.4]) proved that there is an irreducible component of $\Sigma_{B'}$ such that the restriction of ρ to this component is finite-to-one, onto and 1–1 a.e. (We will abuse notation and call this component $\Sigma_{B'}$ as well.)

Now let $\Sigma_{A'}$ be the SFT defined by

$$L_{A'} = \{(a, \{a_1, \ldots, a_k\}): a \in L_A, \{a_1, \ldots, a_k\} \in L_{B'} \text{ and } a \in \{a_1, \ldots, a_k\}\}$$
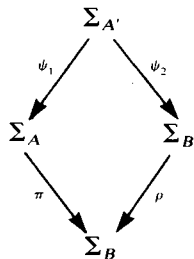
with transitions:

$$(a, \{a_1, \ldots, a_k\}) \to (a', \{a'_1, \ldots, a'_k\})$$

if and only if

$$a \to a' \quad \text{and} \quad \{a_1, \ldots, a_k\} \to \{a'_1, \ldots, a'_k\}.$$

Define $\psi_1: \Sigma_{A'} \to \Sigma_A$ by $\psi_1(x, y) = x$ and $\psi_2: \Sigma_{A'} \to \Sigma_{B'}$ by $\psi_2(x, y) = y$. (Note that we have here a subshift of the fibred product over π and ρ.)

Of course the diagram:

commutes.

Since π and ρ are finite-to-one and onto, so are ψ_1 and ψ_2. So, if we replace $\Sigma_{A'}$ by any one of its irreducible components of maximal entropy (which we will still call $\Sigma_{A'}$) we can assume (by proposition 0.1) that $\Sigma_{A'}$ is irreducible *and* that ψ_1 and ψ_2 are finite-to-one and onto. Moreover, since ρ is 1–1 a.e. so is ψ_1. So, the factor map ψ_2 is an almost topological extension of π. It remains to show that ψ_2 is k-to-1

everywhere. The proof of this follows from the following:

LEMMA 2.2. *If* $\{a_1, \ldots, a_k\} \to \{a'_1, \ldots, a'_k\}$ *in* $\Sigma_{B'}$ *then for all* $i = 1, \ldots, k$ *there exists a unique* $j \in \{1, \ldots, k\}$ *such that* $a_i \to a'_j$.

Proof. By definition of the transition

$$\{a_1, \ldots, a_k\} \to \{a'_1, \ldots, a'_k\}$$

for each i, there exists j such that

$$a_i \to a'_j.$$

Suppose there were two such j (call them j_1 and j_2). Let w be a doubly transitive point in $\Sigma_{B'}$. Assume also that

$$w_0 = \{a_1, \ldots, a_k\}, \qquad w_1 = \{a'_1, \ldots, a'_k\}.$$

Then $\rho(w)$ is a doubly transitive point in Σ_B and $\pi^{-1}(\rho(w))$ contains two mutually separated points x and y where

$$x_0 = a_i, \quad x_1 = a'_{j_1}, \quad y_1 = a'_{j_2}.$$

But then the point u defined by

$$u_l = \begin{cases} x_l & \text{if } l \leq 0, \\ y_l & \text{if } l > 0, \end{cases}$$

is also in $\pi^{-1}(\rho(w))$; also u is *not* mutually separated from x (nor y); this contradicts the fact (proposition 0.2) that the inverse images of a doubly transitive point (in this case $\rho(w)$) are mutually separated. This completes the proof of lemma 2.2.

$\square$

One might ask for more than the preceding result yields: namely, if $\pi: \Sigma_A \to \Sigma_B$ is a k-to-1 a.e. factor map, then can π be written as
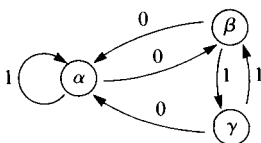
$$\pi = \pi_1 \circ \pi_2,$$

where the (range of π_2) = (domain of π_1) is an SFT and
   (i) π_1 is k-to-1 everywhere? or
   (ii) π_2 is k-to-1 everywhere?
The following examples show neither of these to be true.

Example 1. Let Σ_A be the SFT defined by the graph:

The labels on the edges define a factor map $\pi: \Sigma_A \to \Sigma_B$ where Σ_B is the full 2-shift. The map is 2–1 a.e. since the collapsing of vertices β and γ decomposes π into a 1–1 a.e. map followed by a 2-to-1 map.

However, if π could be factored $\pi = \pi_2 \circ \pi_1$ where π_1 was 2-to-1 everywhere, then every fibre of π would have an even number of points. But $\pi^{-1}(\ldots 1111 \ldots)$ has 3 points.

Example 2. Let Σ_A and Σ_B be the SFT's defined by:
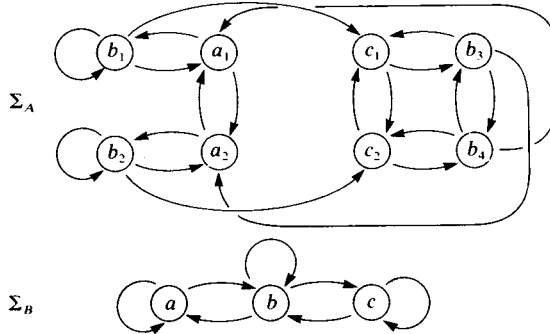
The map $\pi:\Sigma_A \to \Sigma_B$ is defined by 'dropping subscripts'. One easily sees that this map is 2-to-1 a.e. Suppose that π factored as $\pi = \pi_2 \circ \pi_1$ where π_2 was 2-to-1 *everywhere*. Then, by proposition 0.4, π_2 would be left closing. Then, π_1 must identify the following 2 pairs of points:

(a) $\ldots b_1 b_1 b_1 b_1 c_1 c_2 c_1 c_2 \ldots$ and $\ldots b_4 b_3 b_4 b_3 c_1 c_2 c_1 c_2 \ldots$;

(b) $\ldots b_2 b_2 b_2 b_2 c_2 c_1 c_2 c_1 \ldots$ and $\ldots b_3 b_4 b_3 b_4 c_2 c_1 c_2 c_1 \ldots$.

But then by the continuity of π_1, it must also identify the pairs of points:

(c) $\ldots b_1 b_1 b_1 b_1 \ldots$ and $\ldots b_4 b_3 b_4 b_3 \ldots$;

(d) $\ldots b_2 b_2 b_2 b_2 \ldots$ and $\ldots b_3 b_4 b_3 b_4 \ldots$.

Thus, π_1 must identify all the points in the π-fibre of $\bar{b} = \ldots bbb \ldots$. So, $\pi_2^{-1}(\bar{b})$ is a single point. But this contradicts the assumption that π_2 was 2-to-1 everywhere.

Finally we remark that the direct product map of the factor maps π constructed in examples 1 and 2 provides an example of a 4-to-1 a.e. map which cannot be factored: $\pi = \pi_2 \circ \pi_1$ so that either π_1 or π_2 is 4-to-1 everywhere.

3. *The weak relation*: *reduction to group actions*

In this section, we reduce the classification of factor maps to the classification of certain group actions (theorem 3.1 below); this is very much analogous to [10]. However, our group actions are continuous, a fact that follows basically from theorem 2.1.

Let Σ_A and Σ_B be irreducible SFT's of the same entropy and let $\pi:\Sigma_A \to \Sigma_B$ be a k-to-1 everywhere factor map. As mentioned before (§ 1, remark (iv)), up to a topological conjugacy of π, Σ_A can be written as a continuous skew-product over Σ_B and, in fact, the skewing function depends on finitely many coordinates; by recoding, we can actually assume that the skewing function depends on only two coordinates: i.e.

$$\Sigma_A = \Sigma_B \times \{1, \ldots, k\}$$

and in this form

$$\pi(x, i) = x,$$
$$\sigma_A(x, i) = (\sigma_B(x), f(x_0 x_1)(i)), \tag{4}$$

where f is some function on 2-blocks taking values in S_k (the symmetric group).

So $L_A = L_B \times \{1, \ldots, k\}$ and the transitions are given by

$$(b, i) \to (b', j)$$

if and only if

$$b \to b' \text{ in } \Sigma_B \quad \text{and} \quad j = f(bb')(i).$$

As in [**10**], we introduce the *full extension* of π as the SFT

$$\Sigma_\pi = \Sigma_B \times S_k,$$
$$\sigma_\pi(x, g) = (\sigma_B(x), f(x_0 x_1)g).$$

Definition. Let G be a finite group and Σ an SFT. A $\mathbb{Z} \times G$ *action on* Σ is the action of the shift together with a continuous action of the group G on Σ in which each element of G commutes with the shift. *We assume that the G part of the action is free*, i.e. that no non-trivial element of G fixes any element of Σ.

The $\mathbb{Z} \times G$ action we are most interested in is the following: let S_k, the symmetric group on k letters, act on Σ_π by its natural action on the alphabet of Σ_π namely, $\bar{g}(x, g) = (x, g\bar{g})$. We call this the *full $\mathbb{Z} \times S_k$ action* on Σ_π. We let G act on the right since the shift σ_π acts on the left. (This makes the G-action commute with σ_π.)

Definition. Let G be a finite group and let Σ_1 and Σ_2 be two SFT's, each with a $\mathbb{Z} \times G$ action. We say that these two $\mathbb{Z} \times G$ actions are *almost topologically conjugate* (A.T.C.) if there is another SFT Σ_3 with a $\mathbb{Z} \times G$ action and 1-1 a.e. continuous factor maps $\Sigma_3 \to \Sigma_1$ and $\Sigma_3 \to \Sigma_2$ which also commute with the G-actions.

The reason for our interest in group actions is the following:

THEOREM 3.1. *Let $\pi_1 : \Sigma_{A_1} \to \Sigma_{B_1}$ and $\pi_2 : \Sigma_{A_2} \to \Sigma_{B_2}$ be two k-to-1 everywhere continuous factor maps. Then π_1 and π_2 are A.T.C. if and only if the full $\mathbb{Z} \times S_k$ actions on Σ_{π_1} and Σ_{π_2} are A.T.C.*

Proof. (only if) Let π_1 and π_2 be A.T.C. and let $\pi_3 : \Sigma_{A_3} \to \Sigma_{B_3}$ be the common almost topological extension of π_1 and π_2. Clearly, π_3 is k-to-1 a.e. We may assume by theorem 2.1 that π_3 is in fact k-to-1 everywhere and that all maps π_1, π_2, π_3 are in the form (4) with Σ_{A_i} built with skewing function f_i over Σ_{B_i}.

We have the commutative diagram:

$$(5)$$
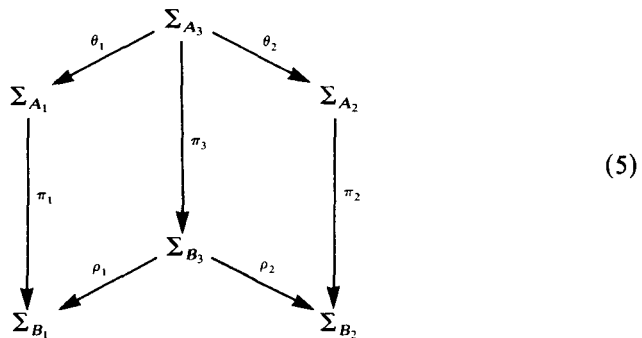
Since π_1, π_2 and π_3 are simply projections onto the first coordinate of (4), we have

$$\theta_1(x, i) = (\rho_1 x, l_1(x)(i))$$

and

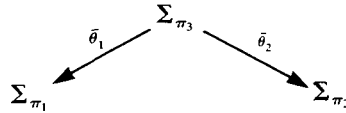$$\theta_2(x, i) = (\rho_2 x, l_2(x)(i))$$

for some $l_1, l_2 : \Sigma_{B_3} \to S_k$. Clearly, l_1 and l_2 are continuous (since ρ_i, θ_i are). Now define

$$\bar{\theta}_1 : \Sigma_{\pi_3} \to \Sigma_{\pi_1}, \qquad \bar{\theta}_1(x, g) = (\rho_1 x, l_1(x)g),$$

and

$$\bar{\theta}_2 : \Sigma_{\pi_3} \to \Sigma_{\pi_1}, \qquad \bar{\theta}_2(x, g) = (\rho_2 x, l_2(x)g).$$

We leave it to the reader to verify that

is an A.T.C. of group actions.

For the converse, let Σ_3 be a common almost topological extension of Σ_{π_1} and Σ_{π_2} (in the $\mathbb{Z} \times S_k$ category). We would like to produce a continuous factor map π_3 for which the full action is Σ_3. This is done as follows: we construct two equivalence relations on Σ_3:

$$x \sim y \quad \text{if there exists } g \in S_k \text{ such that } gx = y,$$

$$x \approx y \quad \text{if there exists } g \in S_k \text{ such that } gx = y \text{ and } g(1) = 1.$$

(When we say $g(1) = 1$, we are thinking of g acting on $\{1, \ldots, n\}$.) Let

$$\Sigma_{A_3} = \Sigma_3 / \approx$$

$$\Sigma_{B_3} = \Sigma_3 / \sim$$

and let $\pi_3 : \Sigma_{A_3} \to \Sigma_{B_3}$ be the obvious map. One can verify that Σ_{A_3} and Σ_{B_3} are indeed SFT's. Moreover, since the continuous factor maps $\Sigma_3 \to \Sigma_{\pi_1}$ and $\Sigma_3 \to \Sigma_{\pi_2}$ commute with the S_k actions, they naturally induce 1–1 a.e. maps $\theta_1, \rho_1, \theta_2, \rho_2$ such that the diagram (5) above commutes, i.e. an A.T.S. between π_1 and π_2. □

Finally, we mention that the work in this section can be viewed in a slightly more concrete manner:

Namely, Σ_π can be defined as the SFT whose alphabet is the set of *ordered* k-tuples $(b_1, \ldots, b_k)$ where $b_1, \ldots, b_k$ are distinct elements of L_A all with the same π-image, b in L_B (the transition are the ones naturally inherited from Σ_A). In this setting, S_k acts by permuting the ordering.

4. *Construction of factor maps*

Theorems 2.1 and 3.1, together with theorem 4.2 below, will give an effective procedure for classifying finite-to-one factor maps between shifts of finite type up to A.T.C. In particular, for a fixed entropy class and a fixed generic cardinality of the fibre, we will show that there are only finitely many equivalence classes all defined by the algebraic invariants of [**10**].

We first recall the results of [**1**]. Consider an irreducible SFT Σ_C, with period $p = p_C$. This means that Σ_C can be written as a disjoint union of p subsets Σ_C^i,

$i = 0, \ldots, p - 1$ such that

   (i) $\sigma(\Sigma_C^i) = \Sigma_C^{(i+1)(\mathrm{mod}\ p)}$; and

   (ii) σ^p / Σ_C^0 is aperiodic.

For a $\mathbb{Z} \times G$ action on Σ_C one defines

$$H_C^i = \{ g \in G : g(\Sigma_C^0) = \Sigma_C^i, i = 0, \ldots, p - 1 \}.$$

Note that $g(\Sigma_C^j) = \Sigma_C^{(i+j)\mathrm{mod}\ p}$ for all $g \in H_C^i$ and $j \in \{0, \ldots, p - 1\}$. The main result of [1] is:

THEOREM 4.1. ([1, theorem 6]) *Let Σ_{C_1} and Σ_{C_2} be irreducible SFT's. Two $\mathbb{Z} \times G$ actions, one on Σ_{C_1} and the other on Σ_{C_2}, are A.T.C. if and only if*

   (a) $h(\Sigma_{C_1}) = h(\Sigma_{C_2})$;

   (b) $p_{C_1} = p_{C_2}$; *and*

   (c) *for all i, $H_{C_1}^i = H_{C_2}^i$.*

We now restate this result in a slightly different form (similar to [10]). First observe that, since the G-part of a $\mathbb{Z} \times G$ action on Σ_C commutes with the shift, H_C^0 is a normal subgroup of G. Next observe that if $g_C \in H_C^1$, then for each $i = 0, \ldots, p - 1$ (where $p = p_{C_1} = p_{C_2}$)

$$H_C^i = (g_C)^i H_C^0.$$

So, each H_C^i is actually a coset of H_C^0; since $g_C^p \in H_C^0$ it follows that G / H_C^0 is a cyclic group with generator $g_C H_C^0$.

   Now, suppose for two $\mathbb{Z} \times G$ actions:

   (c') $H_{C_1}^0 = H_{C_2}^0$ and $g_{C_1} = g_{C_2} \bmod (H_{C_1}^0)$.

This is equivalent to saying

$$H_{C_1}^0 = H_{C_2}^0 \quad \text{and} \quad H_{C_1}^1 = H_{C_2}^1.$$

Then we also get

$$H_{C_1}^i = (g_{C_1})^i H_{C_1}^0 = (g_{C_2})^i H_{C_2}^0 = H_{C_2}^i$$

which is condition (c). So, condition (c') can replace condition (c) in theorem 4.1.

   Unfortunately, the SFT's we must consider may be reducible. However, they do satisfy a weaker property:

Definition. An SFT Σ_C is *non-wandering* if it is a disjoint union of irreducible components.

Note. In general, an SFT may have points that 'wander' from one component to another (see [2, p. 21]).

For any $\mathbb{Z} \times G$ action on a non-wandering SFT, the G-part of the action acts on the set of irreducible components of the SFT. If *this* G action is transitive, we say that the $\mathbb{Z} \times G$ action is *G-transitive*.

   For a non-wandering SFT, Σ_C, and a G-transitive $\mathbb{Z} \times G$ action on Σ_C, pick an arbitrary irreducible component $\Sigma_{\bar{C}}$ of Σ_C and let $G_{\bar{C}} = \{ g \in G : g(\Sigma_{\bar{C}}) = \Sigma_{\bar{C}} \}$. Define

H^0_C and g_C as $H^0_{\bar{C}}$ and $g_{\bar{C}}$(resp.) where the latter are defined as above for the restricted $\mathbb{Z} \times G_{\bar{C}}$ action on $\Sigma_{\bar{C}}$. Now, $H^0_C = H^0_{\bar{C}}$ is normal in $G_{\bar{C}}$, but not necessarily normal in G. However, $g_C \in N(H^0_C)$, the G-normalizer of H^0_C.

Now, suppose one picks two different components $\Sigma_{\bar{C}}$ and $\Sigma_{\hat{C}}$ of Σ_C. Let $g \in G$ map $\Sigma_{\bar{C}}$ onto $\Sigma_{\hat{C}}$; then

$$gH_{\bar{C}}g^{-1} = H_{\hat{C}}$$

and

$$g(g_{\bar{C}}H_{\bar{C}})g^{-1} = g_{\hat{C}}H_{\hat{C}}.$$

The latter is equivalent to:

$$gg_{\bar{C}}g^{-1} = g_{\hat{C}} \bmod H_{\hat{C}}.$$

So, our choices of H^0_C and g_C are unique only up to conjugation in G (and g_C is unique only up to that conjugation plus multiplication by an element of H^0_C). Using these observations, the following is straightforward:

THEOREM 4.2. *Let* Σ_{C_1} *and* Σ_{C_2} *be two non-wandering SFT's each with a transitive* $\mathbb{Z} \times G$ *action. Then these actions are A.T.C. if and only if*:
- (A) $h(\Sigma_{C_1}) = h(\Sigma_{C_2})$;
- (B) $H^0_{C_1}$ *is conjugate to* $H^0_{C_2}$ *(by a conjugacy g)*;
- (C) $gg_{C_1}g^{-1} = g_{C_2} \bmod H^0_{C_2}$.

We leave it to the reader (see [10]), to see that the full action for a factor map between aperiodic SFT's of the same entropy satisfies the above hypothesis: i.e. Σ_π is non-wandering and the full action is G-transitive. Now we prove:

THEOREM 4.3. *For any aperiodic SFT Σ_B and any subgroup H of S_k ($k \in \mathbb{N}$) and any element $g \in N(H)$, there is a factor map $\pi: \Sigma_A \to \Sigma_B$ where Σ_A is an SFT and π is k–1 everywhere and the full $\mathbb{Z} \times S_k$ action on Σ_π realizes the invariants H and g (i.e. there is an irreducible component Σ_C of Σ_π for which $H^0_C = H$ and $g_C = g$).*

One can show, as Rudolph did ([10, theorem 2]), that if H acts transitively on $\{1, \ldots, k\}$, then Σ_A can be chosen aperiodic. The reader should compare theorem 4.3 with the construction given at the end of [10].

Question. Given k, H and g, for which SFT's Σ_A does there exist an SFT Σ_B and k-to-1 a.e. factor map $\pi: \Sigma_A \to \Sigma_B$ realizing the invariant H and g? Arguments in Hedlund ([5]) can be used to show that for $k = 3$ and $\Sigma_A =$ (the full 2-shift) no such map exists (for any choice of H and g).

The proof of theorem 4.3 follows the lines of [10]. First we need:

LEMMA 4.4. *Let Σ_B be an aperiodic SFT. Let n be a positive integer. Then there is an SFT, topologically conjugate to Σ_B, which has a state with at least n successors (i.e. outgoing transitions).*

Proof. We do this by induction on n. For $n = 2$, this follows immediately from the fact that Σ_B is aperiodic. Now, assume that Σ_B has a state v with $n \geq 2$ successors;

we will produce a conjugate SFT which has a state with at least $n+1$ successors. Let $\eta = \eta_1 \eta_2 \cdots \eta_t$ be a simple path (in the defining graph for Σ_B) where η_1 has at least two predecessors (i.e. incoming transitions) and $\eta_t = v$. The existence of η follows from aperiodicity. By a *complete splitting* of the state v, we mean a new graph defined by splitting the vertex v into new vertices, one for each successor of v (see [7] for the notion of state splitting); here, each predecessor (in particular, η_{t-1}) of v is a predecessor of each of the new vertices, and each new vertex has exactly one successor, namely the successor that it represents.

before splitting   after splitting
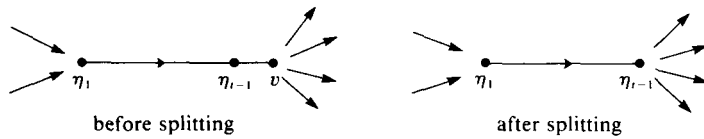
The SFT defined by this new graph is naturally conjugate to Σ_B and the state η_{t-1} now has at least n successors. So, by a conjugacy, we have effectively reduced the length of the path η. By continuing this inductively downward (shortening η), we obtain an SFT conjugate to Σ_B which has a state v with at least n successors and 2 predecessors. A complete splitting of v now gives us two distinct vertices v_1, v_2 each with at least n successors. Now, let η' be a simple path in the defining graph of the new SFT which originates at a successor s of v_1 and terminates at v_2.
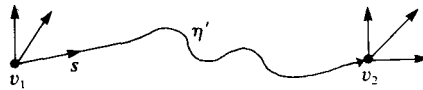
We may assume that η' does *not pass* through v_1 or any successors of v_1 (other than s). Again, we can do a complete splitting of v_2 and thereby effectively reduce the length of η'; continuing inductively downward (shortening η'), we can assume that v_2 is a successor of v_1. Now, a complete splitting of v_2 will provide v_1 with at least $2n-1 \geq n+1$ successors as desired.   □

We apply the lemma as follows. Let n be the order of H and write

$$H = \{h_1, \ldots, h_n\}.$$

Let m be the order of g in S_k and let p be the order of gH in $N(H)/H$.

By virtue of the lemma, we may as well assume that Σ_B itself has a state v with at least $n+2$ successors, $s_1, \ldots, s_{n+2}$. Moreover, using the aperiodicity assumption and replacing Σ_B by some higher-block presentation (see [2, p. 13]) of itself, we may assume that there exist *cycles* $\gamma_1, \ldots, \gamma_n, \alpha_1, \alpha_2$ such that:

(1) the initial edges of $\gamma_1, \ldots, \gamma_n, \alpha_1, \alpha_2$ are the edges $vs_1, \ldots, vs_{n+2}$;

(2) the set of noninitial edges of $\gamma_1, \ldots, \gamma_n, \alpha_1, \alpha_2$ is disjoint from the set of initial edges;

(3) the lengths of $\gamma_1, \ldots, \gamma_n$ are multiples of m; and

(4) the lengths of α_1, α_2 are multiples of p by a pair of relatively prime numbers.
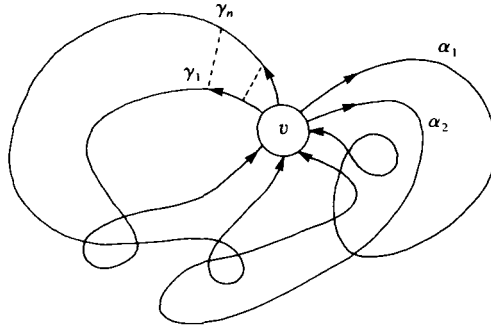
(We leave it to the reader to verify this.) Now, define f on the 2-blocks of Σ_B by:

$$f(bb') = \begin{cases} gh_i & \text{if } bb' = vs_i, \quad 1 \le i \le n \\ (g)^{1-|\alpha_i|} & \text{if } bb' = vs_{n+i}, \quad i = 1, 2 \\ g & \text{otherwise.} \end{cases}$$

(Here, $|\alpha_i|$ denotes the length of the cycle α_i.) Note that since $|\alpha_i|$ is a multiple of p, the order of gH in the group $N(H)/H$, it follows that $(g)^{-|\alpha_i|}$ belongs to H. So, the range of f lies in the coset gH. Using this f, let π be the map defined as in (4) of § 3.

We will show that $\Sigma_B \times H$ is a cyclic subset for an irreducible component of Σ_π. Since the range of f is contained in gH, this will show that the full $\mathbb{Z} \times G$ action for Σ_π realizes the invariants H and g. Let

$$Y = \bigcup_{i=0}^{p-1} \sigma_\pi^i(\Sigma_B \times H)$$

$$= \Sigma_B \times \left(\bigcup_{i=0}^{p-1} (g)^i H \right)$$

$$= \Sigma_B \times G(g, H)$$

where $G(g, H)$ is the group generated by g and H. We will show that Y is an irreducible component for Σ_π. Since σ_π operates in the 2nd coordinate only by elements of $G(g, H)$, it immediately follows that Y is an invariant set, and since Σ_B is a direct factor of Y, Y is a union of irreducible components. Thus it suffices to show that $\sigma_\pi | Y$ is transitive. In view of the irreducibility of Σ_B, it then suffices to show that for all pairs $h, h' \in H$ there is a path in the graph of Σ_π which starts at (v, h) and ends at (v, h'). One does this by 'lifting' the cycle γ_i where i is defined by

$$h_i = h'h^{-1},$$

namely: if $\gamma_i = x_1 \cdots x_l$ with $x_1 = v$ and $x_2 = s_i$ then the path

$$(x_1, h)(x_2, f(x_1x_2)h) \cdots (x_1, f(x_lx_1)f(x_{l-1}x_l) \cdots f(x_1x_2)h)$$

starts at (v, h) and ends at (v, h'), as desired. (Recall that l is a multiple of the order of g in S_k.) So, Y is an irreducible component. To see that $\Sigma_B \times H$ is a cyclic subset of Y with period p, it simply suffices to find two cycles in the graph of Σ_π which

start in $\Sigma_B \times H$ and whose lengths are multiples of p by a pair of relatively prime numbers. This is done by lifting the cycles α_1, α_2 to $\Sigma_B \times H$.

Appendix

Let Σ_B be an irreducible SFT and $f_1, f_2 \colon \Sigma_B \to S_k$ be locally constant functions. If h is an almost continuous solution to the cohomology equation:

$$f_1 = (h \circ \sigma_B) \cdot f_2 \cdot h^{-1} \quad \text{a.e.}$$

then h is locally constant and in fact depends on at most as many coordinates as f_1 and f_2 do.

Proof. Assume for simplicity that f_1 and f_2 depend on only 1 coordinate. Since h is almost continuous, there exists an element $\tau \in S_k$ and a cylinder set C on which $h = \tau$ modulo a set of measure zero. By recoding, we may assume that C is a 1-block cylinder $C = \{x \colon x_0 = c\}$.

For $d \in L_B$ let $w = w_0 \cdots w_l$ be a block such that $w_0 = c$ and $w_l = d$. Let

$$\alpha_w = \prod_{l-1}^{0} f_2(w_i).$$

(By this product, we mean $f_2(w_{l-1}) \cdots f_2(w_0)$.) Also let

$$\bar{\alpha}_w = \prod_{l-1}^{0} f_1(w_i).$$

Define $\bar{h} \colon L_B \to S_k$ by

$$\bar{h}(d) = \bar{\alpha}_w \tau \alpha_w^{-1}.$$

We claim that $\bar{h}$ is well-defined (i.e. does not depend on the choice of block w). To see this, let $w = w_0 \cdots w_l$ and $w^1 = w_0^1 \cdots w_{l'}^1$ with

$$w_0 = w_0^1 = c \quad \text{and} \quad w_l = w_{l'}^1 = d.$$

Let $z = z_1 z_2 \cdots z_k$ be a path which begins with d and ends with a predecessor of c. Then, wz and $w^1 z$ are both paths which begin with c and end with a predecessor of c. Then, using the cohomology equation and the fact that h is constant on C (mod 0), one sees that

$$\bar{\beta} \bar{\alpha}_w \tau (\beta \alpha_w)^{-1} = \tau$$

and

$$\bar{\beta} \bar{\alpha}_{w^1} \tau (\beta \alpha_{w^1})^{-1} = \tau$$

where $\beta = \prod_{i=k}^{1} f_2(z_i)$ and $\bar{\beta} = \prod_{i=k}^{1} f_1(z_i)$. From this, one sees that

$$\bar{\alpha}_w \tau \alpha_w^{-1} = \bar{\alpha}_{w^1} \tau \alpha_{w^1}^{-1}.$$

So, $\bar{h}$ is well-defined. Then $\bar{h}$ defines a 1-block solution to the cohomology equation. (The reader can now verify that in fact $h = \bar{h}$ mod 0).

REFERENCES

[1] R. Adler, B. Kitchens & B. Marcus. Finite group actions on shifts of finite type. *Ergod. Th. & Dynam. Sys.* **5** (1985), 1–25.

[2] R. Adler & B. Marcus. Topological entropy and equivalence of dynamical systems. *Mem. Amer. Math. Soc.* **20** (1979), #219.

[3] E. Coven & M. Paul. Endomorphisms of irreducible subshifts of finite type. *Math. Systems Theory* **8** (1974), 167–175.

[4] E. Coven & M. Paul. Finite procedures for sofic systems. *Monats. Math.* **83** (1977), 265–278.

[5] G. A. Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Math. Systems Theory* **3** (1969), 320–375.

[6] A. N. Livshits. Homology properties of Y-systems. *Mathematical Notes of Academy of Sciences USSR* **10** (1971), 751–763.

[7] B. Marcus. Factors and extensions of full shifts. *Monats. Math.* **88** (1979), 239–247.

[8] M. Nasu. Constant-to-one and onto global maps of homomorphisms between strongly connected graphs. *Ergod. Th. & Dynam. Sys.* **3** (1983), 387–414.

[9] W. Parry & S. Tuncel. *Classification Problems in Ergodic Theory.* London Math. Soc. Lecture Series #67, Cambridge University Press, 1982.

[10] D. Rudolph. Counting the relatively finite factors of a Bernoulli shift. *Israel J. Math.* **30** (1978), 255–263.