EVENTUAL EXTENSIONS OF FINITE CODES
MIKE BOYLE
(Communicated by Kenneth R. Meyer)

ABSTRACT. Suppose $S$ and $T$ are shift equivalent mixing shifts of finite type,
and $f$ is a conjugacy from a subsystem of $S$ to a subsystem of $T$. Then for any
sufficiently large $n$, $f$ extends to a conjugacy of $S^n$ and $T^n$. A consequence of
the proof is a fortified version of Wagoner’s Stable FOG Theorem.

Recall [Wi] that a nonnegative integral matrix $A$ is the adjacency matrix of a
directed graph whose arc set is the coordinate state space of a shift of finite type
(SFT), which we denote $(X_A, S_A)$. In this paper we prove the following theorem.

**EVENTUAL EXTENSION THEOREM.** Let $(X, S)$ be a subshift contained in a
mixing SFT $(X_A, S_A)$. If $f$ be a continuous injective map $f: X \to X_B$ such that
$f S = S_B f$. Then the following are equivalent.

1. $A$ and $B$ are shift equivalent.
2. For any sufficiently large $n$, there exists a homeomorphism $f: X_A \to X_B$
such that $f S = S_B f$.

**COROLLARY.** Let $(X, S)$ be a mixing SFT and suppose $U$ is an automorphism
of a subshift of $(X, S)$. Then for all large $n$, $U$ extends to an automorphism of
$(X, S^n)$.

As an application of the ideas involved in the proof of the theorem (specifically
of Lemma 1), we provide in the appendix a more direct proof of Wagoner’s Stable
FOG Theorem [Wa2]. This proof also yields technical improvements in the result
which lead to a concrete presentation theorem for automorphisms of the shift which
lie in the kernel of the dimension representation.

We find the theorem and its corollary of interest for three reasons. The first is its
relevance to a fundamental unsolved problem of symbolic dynamics, due essentially
to R. F. Williams: when does an automorphism of a subsystem of a mixing SFT
extend to an automorphism of the SFT? This question is basic to understanding
the dynamics: one wants to know whether isomorphic subsystems, especially finite
subsystems such as fixed points, can sit within the SFT in essentially different ways.
Also, Williams has pointed out that this extension problem provides a test for his
conjecture [Wi] that shift equivalence implies conjugacy: for example, it is possible
that a transposition of fixed points may extend to an automorphism in one SFT
but not in a shift equivalent SFT. The corollary provides some insight into this

Received by the editors December 1, 1987.
1980 Mathematics Subject Classification (1985 Revision). Primary 54H20; Secondary 58F15,
28D20.

This research was partially supported by the National Science Foundation under grant DMS-
8601619.

©1988 American Mathematical Society
0002-9939/88 $1.00 + .25 per page

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
problem, by placing limits on possible obstructions to extension. For more on the extension problem see [BK, BLR, N2, B2, Wa2, F].

The second reason is that the theorem gives a complete answer to the extension question in the “eventual” category, i.e., with respect to all sufficiently large powers of the transformation. This has emerged as a natural and significant category for understanding SFTs. An “eventual” result gives a lot of information and for some purposes is as good as the corresponding “noneventual” result; still the eventual category allows enough freedom to answer some questions completely and elegantly. The main example is the classification of SFTs up to eventual conjugacy by shift equivalence [Wi, KR]. Another is the characterization for existence of a closing factor map between mixing SFTs of equal entropy [BMT]. Also, eventual results can have noneventual consequences. For example, the Eventual Factors Theorem in [BMT] led to the classification of Markov shifts of maximal type by regular isomorphism, and Wagoner’s “stable FOG” theorem implied new constraints for the action of an automorphism of an SFT on periodic points [Wa2].

The last reason for our interest is that the present result fits a pattern which suggests some conceptual relationship among several open problems in symbolic dynamics. When are two mixing SFTs conjugate [Wi]? When is there a closing factor map between them [BMT]? What is the image of the dimension representation of the automorphism group [BLR]? In each case one has an answer with respect to all sufficiently large powers of the shift, but the original problem is open. Their solutions may involve a common idea.

To prepare for the proof of the theorem, we recall how an SSE (strong shift equivalence) of matrices defining two SFTs yields a conjugacy. This correspondence was introduced by Williams [Wi]. Our viewpoint is also influenced by Parry and Tuncel [PT, Chapter V, Theorem 20] and Nasu [N1].

An $n \times n$ matrix $A$ over $\mathbb{Z}+$ is the adjacency matrix of a (directed) graph with $n$ vertices, with $A_{ij}$ arcs from vertex $i$ to vertex $j$. The arc set $\mathcal{A}(A)$ is the coordinate symbol space for $X_A$, the bisequence space of walks through this graph; $S_A$ is the shift map defined by setting $(S_A x)_i = x_{i+1}$ for $x$ in $X_A$ and $i$ in $\mathbb{Z}$; $(X_A, S_A)$ is the SFT defined by $A$. If $U$ and $V$ are matrices over $\mathbb{Z}+$ with $A = UV$, $B = VU$, then the elementary SSE $(U, V)$ from $A$ to $B$ induces a conjugacy from $(X_A, S_A)$ to $(X_B, S_B)$ as follows. Let $C$ be the matrix $\begin{pmatrix} 0 & U \\ V & 0 \end{pmatrix}$, so

$$C^2 = \begin{pmatrix} UV & 0 \\ 0 & VU \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$ 

Let $(X_C^{(1)}, (S_C)^2)$ and $(X_C^{(2)}, (S_C)^2)$ be the two components of $(X_C, (S_C)^2)$ corresponding to $UV$ and $VU$. Define bijections of arcs

$$\rho : \mathcal{A}(A) \to \mathcal{A}(UV), \quad \bar{\rho} : \mathcal{A}(B) \to \mathcal{A}(VU).$$

These induce one-block conjugates $h_A : (X_A, S_A) \to (X_C^{(1)}, S_C^2)$ and $h_B : (X_B, S_B) \to (X_C^{(2)}, S_C^2)$. Then $(h_B)^{-1} S_C h_A$ is a conjugacy from $(X_A, S_A)$ to $(X_B, S_B)$. This conjugacy, which we denote $c(U, V)$ as in [Wa1], is uniquely determined by the bijections $\rho, \bar{\rho}$ (which we usually suppress from the notation: for example, $f = c(U, V)$ means that for some choice of bijections $f = c(U, V)$). Very concretely, then, we think of $\mathcal{A}(UV)$ as a subset of $\mathcal{A}(U) \mathcal{A}(V)$. Likewise $\mathcal{A}(VU)$ is a subset
of $\mathcal{A}(V),\mathcal{A}(U)$, and after the identifications of arcs we think of a sequence in $X_A$ as

$$\cdots x_0x_1x_2\cdots = \cdots (u_0v_0)(u_1v_1)(u_2v_2)\cdots$$

being mapped by $c(U,V)$ to

$$\cdots y_0y_1y_2\cdots = \cdots (v_0u_1)(v_1u_2)(v_2u_3)\cdots.$$ 

We say that a conjugacy $f$ from $(X_A, S_A)$ to $(X_B, S_B)$ is associated to an SSE $(U_1, V_1), \ldots, (U_k, V_k)$ if (for suitable choices of bijections of arcs)

$$f \circ (S_A)^p = c(U_k, V_k) \circ \cdots \circ c(U_1, V_1)$$

for some $p$ in $\mathbb{Z}$. For example, the identity map is associated to $(I, A)$ (with $p = 0$) and the shift $S_A$ is associated to $(A, I)$ (with $p = 0$). Implicit in [Wi] is the result that any conjugacy between SFTs $(X_A, S_A), (X_B, S_B)$ is associated to some SSE, and in fact the number $p$ above can be chosen to satisfy $0 \leq p \leq k$. (Alternatively, see [PT, Proposition V.19] or [Wal, Proposition 3.6].)

For $k \geq 1$, the systems $(X^k_A, (S_A)^k)$ and $(X^{k*}_A, S^{k*}_A)$ are topologically conjugate [Wi]. Given $k$, for a conjugacy $\gamma_A : (X_A, (S_A)^k) \to (X_B^{k*}, S^{k*}_B)$ we choose a bijection $g$ from the words $x_0 \cdots x_{k-1}$ of length $k$ in $X_A$ to words of length $1$ in $X_B$ such that each path $x_0 \cdots x_{k-1}$ shares with the arc $g(x_0 \cdots x_{k-1})$ the same initial vertex and the same terminal vertex. Then $\gamma_A$ is given by the block code

$$(\gamma_A x)_i = g(x_{ik} \cdots x_{ik+k-1}), \quad x \in X_A, \; i \in \mathbb{Z}.$$ 

Below, we will tacitly replace systems $(X_A, (S_A)^k)$ with systems $(X^{k*}_A, S^{k*}_A)$, using for each $k$ some fixed conjugacy $\gamma_A$ ($\gamma_A$ depends on $k$, but we suppress this from the notation). For example, when $f : (X_A, S_A) \to (X_B, S_B)$ and we refer to $f$ "as a map from $(X^{k*}_A, S^{k*}_A)$ to $(X^{k*}_B, S^{k*}_B)$", we are referring to the map $\gamma_B f(\gamma_A)^{-1}$.

**Lemma 1.** Suppose $f : X_A \to X_B$, $p$ is an integer, and

$$f \circ (S_A)^p = c(U_k, V_k) \circ \cdots \circ c(U_1, V_1)$$

for an SSE $(U_1, V_1), \ldots, (U_k, V_k)$ from $A$ to $B$. Let $U = U_1U_2\cdots U_k, \; V = V_k\cdots V_2V_1$. Suppose $m$ and $n$ are nonnegative integers. Then as a conjugacy from $(X_A^{*+m+n}, S_A^{*+m+n})$ to $(X_B^{*+m+n}, S_B^{*+m+n})$,

$$f \circ (S_A)^{p+m} = c(A^mU, VA^n).$$

**Proof.** There are bijections of arcs giving $c(I, A) = \text{Id}$ and $c(A, I) = S_A$; consequently

$$f \circ (S_A)^{p+m} = c(U_k, V_k) \circ \cdots \circ c(U_1, V_1) \circ (c(A, I))^m \circ (c(I, A))^n$$

and therefore it suffices to prove the case $m = n = 0$. The proof of this case amounts to understanding a certain picture, which we give for the case $k = 3$:

This is a picture of how the elementary conjugacies

$$A_{i-1} = U_iV_i, \quad A_i = V_iU_i$$

with the given bijections of arcs determine $c(U_k, V_k) \circ \cdots \circ c(U_1, V_1)$. In the picture, for each $i$ the recurring symbols $u_i$ may be different—we suppress this from the notation, and use subscripts just to indicate the associated matrix. The bijections
of arcs for $V_{i-1}U_{i-1} = A_{i-1} = U_i V_i$ determine which pair $u_i v_i$ occurs directly beneath a pair $v_{i-1} u_{i-1}$. The map $f$ sends the sequence of $a$'s on the top line to the sequence of $b$'s on the bottom line. If $x$ is in $X_A$ and $y$ in $X_B$ is its image under $c(U_k, V_k) \circ \cdots \circ c(U_1, V_1)$ and $x_0 \cdots x_{2k-1}$ is the word $a \cdots a$ on the top line, then $y_0 \cdots y_{k-1}$ is the word $b \cdots b$ on the bottom line. The idea of the lemma is that the circled words of the picture correspond to arcs in a diagram

![Diagram](image)

for the bijections $\rho, \bar{\rho}$ which define the desired map $c(U, V)$.

Let $\mathcal{W}_k(A)$ be the set of $X_A$-words of length $k$. A word $W$ in $\mathcal{W}_k(A)$ determines a word $\rho(W) = u_1 \cdots u_k v_k \cdots v_1$, whose initial and terminal vertices agree with those of $W$, via the given bijections (as in the picture: $W$ is the circled word $a \cdots a$; now just fill in under $W$ to get $\rho(w)$ from the circled words). The resulting word $u_1 \cdots u_k v_k \cdots v_1$ must be a path (i.e., the terminal vertex of any symbol must equal the initial vertex of the next symbol). Conversely, any path $u_1 \cdots u_k v_k \cdots v_1$ equals $\rho(W)$ for a unique word $W$ from $\mathcal{W}_k(A)$ (let $u_1 \cdots u_k$ and $v_k \cdots v_1$ be the circled words in the picture and fill in back up to $a \cdots a = W$). Now for each pair $i, j$ the entry $U_{ij}$ is the number of paths $u_1 u_2 \cdots u_k$ from $i$ to $j$, and we identify the $U$-arcs from $i$ to $j$ with these paths. Likewise, the paths $v_k \cdots v_1$ are the $V$-arcs. With these identifications and the identifications $\mathcal{W}_k(A) \leftrightarrow \mathcal{A}(A^k)$ which define $\gamma_A$, we have a bijection

$$\rho: \mathcal{A}(A^k) \leftrightarrow \mathcal{A}(UV).$$

Similarly we get a bijection

$$\bar{\rho}: \mathcal{A}(B^k) \rightarrow \mathcal{A}(UV)$$

using the circled segments of the picture $b \cdots b \leftrightarrow v_k \cdots v_1 \bar{u}_1 \cdots \bar{u}_k$. The bijections $\rho, \bar{\rho}$ define a conjugacy $c(U, V): (X_A^*, S_A^*) \rightarrow (X_B^*, S_B^*)$ which agrees with

$$\gamma_B \circ c(U_k, V_k) \circ \cdots \circ c(U_1, V_1) \circ (\gamma_A)^{-1}. \quad \square$$

**Lemma 2.** Let $(X_A, S_A)$ and $(X_B, S_B)$ be mixing SFTs with the same zeta function. Suppose there are subshifts

$$(X, S) \subset (X', S') \subsetneq (X_A, S_A)$$
and an embedding
\[(X, S) \hookrightarrow (X_B, S_B).\]

Then this embedding extends to an embedding
\[(X', S') \hookrightarrow (X_B, S_B).\]

**Proof.** This is a marker exercise: mix the embedding theorem proof of [K] and the extension lemma viewpoint of [B1]—see the remark following Lemma (2.4) in [B1]. □

**Proof of the Theorem.** (2)⇒(1) Given (2), for all large \(n\) the shifts \((X_A, (S_A)^n), (X_B, (S_B)^n)\) are conjugate. It follows from Williams [Wi] that for all large \(n\), the matrices \(A^n, B^n\) are shift equivalent. Kim and Roush [KR, Theorem 3.3] proved that this forces \(A\) and \(B\) to be shift equivalent.

(1)⇒(2) We may assume that \(A\) and \(B\) are primitive. Also, we may apply Lemma 2 to extend \(f\) to some SFT properly contained in \((X_A, S_A)\); so, we may assume that \((X, S)\) is an SFT to begin with. After passage to a higher block presentation, we may assume that \((X, S)\) and its image under \(f\) are defined by matrices \(A\) and \(B\), and that \(A\) and \(B\) are primitive with the forms

\[A = \begin{pmatrix} \bar{A} & \ast \\ \ast & \ast \end{pmatrix}, \quad B = \begin{pmatrix} \bar{B} & \ast \\ \ast & \ast \end{pmatrix}.\]

Now, taking some SSE of \(\bar{A}, \bar{B}\) and applying Lemma 1, we obtain matrices \(\bar{U}, \bar{V}\) and \(k\) in \(\mathbb{N}\) such that for all nonnegative integers \(m, n\) there exists an integer \(t = t(m, n)\) such that \(\bar{f} \circ (S_A)^t\) as a conjugacy of \((X_A, (S_A)^{k+m+n}), (X_B, (S_B)^{k+m+n})\) is associated to the elementary SSE \((\bar{A}^m \bar{U}, \bar{V} \bar{A}^n)\). Also we obtain matrices \(U, V\) and \(j\) in \(\mathbb{N}\) such that for all large \(m, n\), \((A^m U, VA^n)\) is an elementary SSE of \(A^{j+m+n}\) and \(B^{j+m+n}\). We may assume \(j = k\). Then we can extend some \(\bar{f}(S_A)^j\) to a conjugacy \(g\) compatible with the elementary SSE \((A^m U, VA^n)\) if the upper left corners of \(A^m U\) and \(VA^n\) dominate the matrices \(\bar{A}^m \bar{U}\) and \(\bar{V} \bar{A}^n\). Because \(A\) and \(B\) are primitive the Perron theorem implies that all the entries of \(A^m U\) and \(VA^n\) grow exponentially with \(m\) and \(n\) like the spectral radius of \(A\), which strictly exceeds the spectral radius of \(\bar{A}\). Therefore for all large \(m\) and \(n\) we have the required domination. Now we regard \(g\) as a conjugacy from \((X_A, (S_A)^{m+n+k})\) to \((X_B, (S_B)^{m+n+k})\) and let \(f = g \circ (S_A)^{-t(m, n)}\) be the desired extension. □

I thank Jack Wagoner, for detailed comments on an earlier version of this paper which strongly influenced its revision; and I thank Ulf Fiebig for a helpful discussion on the appendix.

**Appendix.** We will use Lemma 1 to give an alternate proof of the following theorem of Wagoner.

**Stable FOG Theorem** [Wa2]. Let \(\alpha\) be an automorphism of an SFT \((X_A, S_A)\) in the kernel of the dimension representation. There is an integer \(k_0 \geq 1\) such that if \(k \geq k_0\), then \(\alpha\) is a product of homeomorphisms of \(X_A\) of finite order which commute with \((S_A)^k\). Furthermore, each of these finite order elements is a simple automorphism of \((X_A, (S_A)^k)\).

Here, FOG stands for finite order generation. The FOG conjecture is that the kernel of the dimension representation of the automorphism group of a shift of finite
type is generated by elements of finite order. A proof of this conjecture would have tremendous implications for the extension problem (e.g., see [BK, BLR, Wa2]).

An automorphism \( f \) of an SFT \((X_A, S_A)\) is simple if \( f = \beta^{-1} g \beta \), where \( \beta \) is a conjugacy to some SFT \((X_B, S_B)\), and \( g \) is given there by a graph automorphism fixing every vertex. (In detail: \( g \) is a 1-block map, state symbols are arcs of a graph with adjacency matrix \( B \), and the permutation of state symbols/arcs by \( g \) determines a graph automorphism fixing every vertex.) Simple automorphisms were introduced and analyzed by Nasu [Na2]. The action of compositions of simple automorphisms on finite subsystems is essentially completely understood [Na2, B2]. The Simple FOG conjecture holds that the kernel of the dimension representation is generated by simple automorphisms. (We echo Wagoner’s question in [Wa2]: does FOG imply Simple FOG?)

Let \( \text{Aut}(S_A) \) denote the automorphism group of an SFT \((X_A, S_A)\). The action of \( \text{Aut}(S_A) \) on certain subsets of \( X_A \) induces via Krieger’s theory a homomorphism (the dimension representation \( \rho \)) from \( \text{Aut}(S_A) \) into the group of automorphisms of the direct limit group \( G_A = \lim \rightarrow X = \lim \rightarrow A \mathbb{Z}^N \) (where \( A = N \) by \( N \)). This group is the quotient of \( \{(v, i) : v \in \mathbb{Z}^N, i \in \mathbb{Z}\} \) by the equivalence relation \( (v, i) \sim (u, j) \) if \( \exists k \in \mathbb{N} \) such that \( vA^{i+k} = uA^{j+k} \). (For background, see [BLR, §6; BK, BMT, Wa2] and their references.) For the proof below, we need just one fact: if \((U_1, V_1), \ldots, (U_k, V_k)\) is an SSE from \( A \) to \( A \), and \( f = c(U_k, V_k) \circ \cdots \circ c(U_1, V_1) \), then \( \rho(f) : G_A \rightarrow G_A \) is given by

\[
\rho(f) : [(v, i)] \mapsto [(vU, i)], \quad v \in \mathbb{Z}^N, i \in \mathbb{Z}.
\]

This fact follows from adapting the argument of [BLR, Lemma 6.2], or by combining that lemma with Lemma 1 of this paper.

**Theorem.** Suppose \( A \) is a nondegenerate (no zero row or column) \( N \times N \) nonnegative integral matrix, \( f \in \text{Aut}(S_A) \) and \( f \in \text{Ker}(\rho) \). Suppose \( 0 \leq p \leq k \) and \((U_1, V_1), \ldots, (U_k, V_k)\) is an SSE from \( A \) to \( A \) such that

\[
f \circ (S_A)^p = c(U_k, V_k) \circ \cdots \circ c(U_1, V_1).
\]

Let \( J \) be the smallest nonnegative integer such that \( \text{rank } A^J = \text{rank } A^{J+1} \) (so, \( J \leq N - 1 \)). Then for every \( t \geq K + 2J \), there exist simple automorphisms \( c_1, c_2 \) of \((X_A, (S_A)^t)\) such that \( f = c_2 \circ c_1 \).

**Proof.** Let \( U = U_1 \circ \cdots \circ U_k, V = V_k \circ \cdots \circ V_1 \). Consider nonnegative integers \( m \geq J, n \geq J, t = k + m + n \). By Lemma 1 (for suitable defining bijections of arcs),

\[
f \circ (S_A)^{p+m} = c(A^m U, V A^n)
\]
as an automorphism of \((X_{A^t}, S_{A^t})\). Because \( f \in \text{Ker}(\rho) \), the matrices \( A^p \) and \( U = U_1 \circ \cdots \circ U_k \) induce the same automorphism of \( \lim \rightarrow A \mathbb{Z}^N \). Therefore \( A^{p+J} = U A^J \). Similarly, \( A^{k-p+J} = V A^J \). Therefore,

\[
f \circ (S_A)^{p+m} = c(A^{p+m}, A^{k-p+n}).
\]

But for some choice of defining bijections of arcs

\[
(S_A)^{p+m} = c(A^{p+m}, A^{k-p+n}).
\]
The choice of bijections of arcs for \( c(A^{p+m}, A^{k-p+n}) \) is unique up to a simple (vertex-fixing) graph automorphism in the domain and another in the range. Therefore there are simple graph automorphisms \( d_1, d_2 \) such that
\[
 f \circ (S_A)^{p+m} = d_2(S_A)^{p+m} d_1 \quad \text{and} \quad f = d_2((S_A)^{p+m} d_1(S_A)^{-p-m}).
\]
Let \( c_2 = d_2 \), and \( c_1 = (S_A)^{p+m} d_1(S_A)^{-p-m} \). \( \square \)

The simple automorphisms obtained above are rather special. We devote a theorem to their description.

**Concrete Presentation Theorem.** In the notation of the theorem above, suppose \( r \) and \( t \) are integers such that \( t \geq k + 2J \) and \( (k-p+J) \leq r \leq t-(p+J) \). Then there exist permutations \( \pi_1, \pi_2 \) of the \( X_A \)-words of length \( t \), always fixing the initial vertex and the terminal vertex, and automorphisms \( c_1, c_2 \) of \( (X_A, (S_A)^t) \) defined by
\[
(c_1x)_{r+it} \cdots (c_1x)_{r+it+t-1} = \pi_1(x_{r+it} \cdots x_{r+it+t-1}), \quad i \in \mathbb{Z},
\]
\[
(c_2x)_{it} \cdots (c_2x)_{it+t-1} = \pi_2(x_{it} \cdots x_{it+t-1}), \quad i \in \mathbb{Z},
\]
such that
\[
f(x) = c_2 c_1(x), \quad x \in X_A.
\]

**Proof.** The proof of the previous theorem gives this result for \( r = m - p \), and therefore for \( r = t - m - p \). Given \( t = k + m + n \), the integer interval of allowed values for \( m \) is \([J, t-k-J]\), so the integer interval of allowed values for \( r = t - m - p \) is
\[
[t-(t-k-J)-p, t-J-p] = [k+J-p, t-J-p]. \quad \square
\]

**References**


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MARYLAND 20742