ISRAEL JOURNAL OF MATHEMATICS 132 (2002), 359–372

CONSTRUCTING FINITARY ISOMORPHISMS WITH FINITE EXPECTED CODING TIMES

BY

ROBERT MOUAT AND SELIM TUNCEL

Department of Mathematics, University of Washington
Seattle, WA 98195, USA
e-mail: mouat@math.washington.edu, tuncel@math.washington.edu

ABSTRACT

This paper is motivated by the question of whether the invariants $\beta, \Delta, c\Delta$ completely characterize isomorphism of Markov chains by finitary isomorphisms that have finite expected coding times (fect). We construct a finitary isomorphism with fect under an additional condition. Whether coincidence of $\beta, \Delta, c\Delta$ implies the required condition remains open.

1. Introduction

In the wake of the results of Keane and Smorodinsky [KS79a, KS79b] on finitary isomorphisms of Markov chains, there was considerable activity on finitary isomorphisms with finite expected coding times (fect). It became clear that entropy was no longer a complete invariant for (aperiodic) Markov chains, and further invariants were constructed [Par79, Kri83, Tun81, PS84, Sch84]. As the smoke cleared in 1983, two questions emerged. The question of whether the classifications by almost block isomorphism and finitary isomorphism with fect were identical was answered negatively in [MT91]. The other question, whether the three invariants $\beta, \Delta, c\Delta$ (which will be defined below) form a complete set for finitary isomorphism with fect, remains open after nearly 20 years.

In this paper we will make some progress on this question by constructing a finitary isomorphism with fect under an additional positivity assumption. In particular, working over a suitable ring, it is known that there exists a matrix

Received April 30, 2001

intertwining the defining matrices of the two Markov chains; a corollary of our result is that the Markov chains are finitarily isomorphic with fect whenever there exists an intertwining matrix with a positive row and a positive column.

We start by defining Markov chains as in [PT82, MT91]. An **exponential function** is a function $\mathbb{R} \to \mathbb{R}^+$ given by $t \mapsto a^t$ for some $a > 0$. Let $\exp = \{a^t : a > 0\}$, and $\mathbb{Z}^+$ denote the non-negative integers. Consider a matrix $A = A(t)$ over $\mathbb{Z}^+[\exp]$, the set of polynomials in exp with non-negative integral coefficients. We can write

$$A(I,J) = \sum_{m \in \exp} a_{I,J,m} m,$$

with $a_{I,J,m} \in \mathbb{Z}^+$ and $a_{I,J,m}$ nonzero for only finitely many m. Let A_t denote the non-negative real matrix resulting from evaluating A at $t \in \mathbb{R}$. We will assume that the matrices we use are primitive (aperiodic). Everything we say has an extension to the periodic case, which we leave to the reader.

When $t = 0$, the matrix A_0 determines a directed graph $G(A)$ with edges $E(A)$ and states (vertices) $S(A)$, and a shift of finite type (X_A, σ_A). We can associate with each edge, e, in $G(A)$ an element of exp called the **weight** of the edge, $wt_A(e)$. The matrix A also defines a probability measure, μ_A, giving the Markov chain (X_A, σ_A, μ_A) as in [MT91]. A matrix P over $\mathbb{Z}^+[\exp]$ is called **stochastic** if the non-negative real matrix P_1 is stochastic. By Proposition 1.3 of [MT91], there exists a unique stochastic matrix P defining the same Markov chain as A.

A Markov chain is traditionally defined via a non-negative real-valued stochastic matrix M. Putting $P(I,J) = M(I,J)^t$ whenever $M(I,J) \neq 0$ and $P(I,J) = 0$ when $M(I,J) = 0$, we get a stochastic matrix over $\mathbb{Z}^+[\exp]$ which has $P_1 = M$ and defines the same Markov chain.

Define the beta function β_A of A by letting $\beta_A(t)$ equal the spectral radius of A_t, for $t \in \mathbb{R}$.

Let G be a primitive weighted graph with weights in a multiplicative Abelian group; for example, $G = G(A)$. For a path $\gamma = e_1 \cdots e_n$ in G, its **length** is the number of edges, $l(\gamma) = n$, while its **weight** is the product of the weights of the edges, $wt(\gamma) = \prod_{i=1}^n wt(e_i)$. We define the **delta group**, Δ, **gamma group**, Γ, and **distinguished coset**, $c\Delta$ for G as in [PS84].

$$\Delta = \{wt(\gamma)wt(\gamma')^{-1} : \gamma, \gamma' \text{ are cycles in } G \text{ with } l(\gamma) = l(\gamma')\},$$
$$c\Delta = \{wt(\gamma)wt(\gamma')^{-1} : \gamma, \gamma' \text{ are cycles in } G \text{ with } l(\gamma) = l(\gamma') + 1\},$$

and Γ is the multiplicative group generated by $\{wt(\gamma) : \gamma \text{ is a cycle in } G\}$. Parry and Schmidt [PS84] showed that the quotient group Γ/Δ is cyclic and generated by $c\Delta$.

Two Markov chains (X_A, σ_A, μ_A) and (X_B, σ_B, μ_B) are **isomorphic** if, almost surely, there exists a measure-preserving bijection ϕ: $X_A \to X_B$ with $\phi\sigma_A = \sigma_B\phi$. The isomorphism ϕ is **finitary** if both ϕ and ϕ^{-1} are continuous a.e., in which case we can find a_ϕ, m_ϕ: $X_A \to \mathbb{Z}^+$ defined a.e. so that if $x'_i = x_i$ for $-m_\phi(x) \leq i \leq a_\phi(x)$ then $\phi(x)_0 = \phi(x')_0$, and similarly $a_{\phi^{-1}}, m_{\phi^{-1}}$: $X_B \to \mathbb{Z}^+$. A finitary isomorphism is said to have **finite expected coding times (fect)** if $a_\phi, m_\phi, a_{\phi^{-1}}, m_{\phi^{-1}}$ may be chosen so that $\int(a_\phi + m_\phi)d\mu_A < \infty$ and $\int(a_{\phi^{-1}} + m_{\phi^{-1}})d\mu_B < \infty$.

A finitary isomorphism is called **hyperbolic** if it preserves the stable and unstable manifolds a.e. Schmidt [Sch86, Sch87] showed that a finitary isomorphism with fect is hyperbolic, and that $\beta, \Delta, c\Delta$ are invariants of hyperbolic finitary isomorphism. An affirmative answer to the question of whether coincidence of $\beta, \Delta, c\Delta$ implies finitary isomorphism with fect would in particular show that the finitary isomorphisms with fect and hyperbolic finitary isomorphisms yield the same classifications, that is, if there exists a hyperbolic finitary isomorphism between two Markov chains then there also exists a finitary isomorphism with finite expected coding times.

Our main result is the following.

1.1. THEOREM: *Let P, Q be primitive stochastic matrices over $\mathbb{Z}^+[\exp]$ with $\beta_P = \beta_Q$, $\Delta_P = \Delta_Q$, and $c_P\Delta_P = c_Q\Delta_Q$. Suppose there are states $I_0 \in S(P)$ and $J_0 \in S(Q)$, a nontrivial column vector $\mathbf{v}_r$ and a row vector $\mathbf{v}_l$ over $\mathbb{Z}^+[\exp]$ with $(A^n\mathbf{v}_r)(I_0) = (\mathbf{v}_l B^n)(J_0)$ for all $n \in \mathbb{Z}^+$. Then there exists a finitary isomorphism ϕ: $X_P \to X_Q$ with finite expected coding times.*

It was shown in [PT81] that $\beta_P = \beta_Q$ implies the existence of a nontrivial matrix V over $\mathbb{Z}[\exp]$ such that $PV = VQ$. In the case there is a V with a non-negative row and a non-negative column we have the following.

COROLLARY 1.2: *Let P, Q be primitive stochastic matrices over $\mathbb{Z}^+[\exp]$ with $\beta_P = \beta_Q$, $\Delta_P = \Delta_Q$, and $c_P\Delta_P = c_Q\Delta_Q$. Suppose that we have a matrix V over $\mathbb{Z}[\exp]$ such that $PV = VQ$ and the entries of at least one nontrivial row and one column of V belong to $\mathbb{Z}^+[\exp]$. Then there exists a finitary isomorphism ϕ: $X_P \to X_Q$ with finite expected coding times.*

The Markov chains (X_P, σ_P, μ_P) and (X_Q, σ_Q, μ_Q) are **finitely equivalent** if there exist a Markov chain (X_R, σ_R, μ_R) and bounded-to-one continuous surjections ψ: $X_R \to X_P$, θ: $X_R \to X_Q$ that commute with the shift transformation and preserve measure. If, in addition, the maps ψ and θ are one-to-one a.e., then (X_P, σ_P, μ_P) and (X_Q, σ_Q, μ_Q) are said to be **almost block isomorphic**. An almost block isomorphism is of course a (very special) finitary isomorphism with fect. It was shown in [PT81] that if $PV = VQ$ for a matrix V over $\mathbb{Z}^+$ [exp] then (X_P, σ_P, μ_P) and (X_Q, σ_Q, μ_Q) are finitely equivalent. The expectation was that the existence of such a V would follow from the equality $\beta_P = \beta_Q$ (see [PT81, Par91]). Combined with Ashley's replacement theorem [Ash90], this would then yield an almost block isomorphism from the equalities $\beta_P = \beta_Q$, $\Delta_P = \Delta_Q$, $c_P\Delta_P = c_Q\Delta_Q$. These hopes were dashed by examples given in [MT91] (see Example 2.2 below). Corollary 1.2 reveals that a milder positivity requirement on V is sufficient for finitary isomorphism with fect.

If H is a (multiplicative) subgroup of the positive reals, we will denote by $\mathbb{Z}[H]$ and $\mathbb{Z}^+[H]$ the corresponding sets of integral and non-negative integral combinations of the exponentials a^t with $a \in H$.

We do not know whether the existence of $I_0, J_0, \mathbf{v}_r, \mathbf{v}_l$ as in 1.1 can be deduced from the coincidence of $\beta, \Delta, c\Delta$. We will see below that we may restrict to $\mathbb{Z}[\Delta]$ and, since $\mathbb{Z}[\Delta]$ is naturally isomorphic to the Laurent polynomial ring $\mathcal{R} = \mathbb{Z}\left[x_1^{\pm}, \ldots, x_d^{\pm}\right]$, work over $\mathcal{R}$. Then P, Q are replaced by matrices A, B over $\mathcal{R}^+ = \mathbb{Z}^+\left[x_1^{\pm}, \ldots, x_d^{\pm}\right]$. Letting m, p denote the sizes of A, B, for each $I_0 \in S(A)$ and $J_0 \in S(B)$ the set

$$M_{I_0,J_0} = \{(v, w) \in \mathcal{R}^{m+p}\colon (A^n v)(I_0) = (wB^n)(J_0) \text{ for all } n \in \mathbb{Z}^+\}$$

is an $\mathcal{R}$-submodule of $\mathcal{R}^{m+p}$. We have $M_{I_0,J_0} \neq \{0\}$ since there exists V over $\mathcal{R}$ with $AV = VB$. In addition, by the primitivity of A, if M_{I_0,J_0} contains a nontrivial element of $(\mathcal{R}^+)^{m+p}$ then it contains an element of $(\mathcal{R}^+ \setminus \{0\})^{m+p}$. This motivated us to give in [EMT] a characterization of the submodules of $\mathcal{R}^N$ that contain an element of $(\mathcal{R}^+ \setminus \{0\})^N$. It is possible that this characterization will help to decide if the existence of I_0, J_0 with $M_{I_0,J_0} \cap (\mathcal{R}^+ \setminus \{0\})^{m+p} \neq \emptyset$ is a consequence of the equality of $\beta, \Delta, c\Delta$.

2. Laurent polynomials

We will restate the conditions of 1.1 in a more convenient form.

Suppose P, Q are primitive stochastic matrices over $\mathbb{Z}^+$ [exp] with $\beta_P = \beta_Q$, $\Delta_P = \Delta_Q$, and $c_P\Delta_P = c_Q\Delta_Q$. Write $\Delta = \Delta_P = \Delta_Q$. According to Proposition 1.17 of [MT91], we can find diagonal matrices D_1, D_2 over exp and $m \in$ exp such that the matrices

$$A = \frac{D_1^{-1} P D_1}{m} \quad \text{and} \quad B = \frac{D_2^{-1} Q D_2}{m}$$

are over $\mathbb{Z}^+[\Delta]$. It is easily seen that $\beta_A = \beta_B$, $\Delta_A = \Delta_B = \Gamma_A = \Gamma_B = \Delta$, and A, B define the same Markov chains as P, Q.

If $I_0, J_0, \mathbf{v}_r$ and $\mathbf{v}_l$ are as in 1.1, let $\hat{\mathbf{v}}_r = mD_1(I_0, I_0)D_1^{-1}\mathbf{v}_r$ and $\hat{\mathbf{v}}_l = \mathbf{v}_l D_2 D_2^{-1}(J_0, J_0)m$. Find a finitely generated subgroup H of the positive reals such that $\hat{\mathbf{v}}_r, \hat{\mathbf{v}}_l$ are over $\mathbb{Z}[H]$ and H contains Δ. Then write $\hat{\mathbf{v}}_r = \sum_i a_i^t \mathbf{v}_{r,i}$, $\hat{\mathbf{v}}_l = \sum_i a_i^t \mathbf{v}_{l,i}$, where $\mathbf{v}_{r,i}, \mathbf{v}_{l,i} \in \mathbb{Z}[\Delta]$, the index i runs through the (finite) set of cosets H/Δ and a_i is a representative of the i-th coset in H/Δ. For each i the equation $(A^n \mathbf{v}_{r,i})(I_0) = (\mathbf{v}_{l,i}B^n)(J_0)$ holds for all $n \in \mathbb{Z}^+$. Choosing i such that $\mathbf{v}_{r,i} \neq 0$ and replacing $\mathbf{v}_r, \mathbf{v}_l$ by $\mathbf{v}_{r,i}, \mathbf{v}_{l,i}$, we may assume that $\mathbf{v}_r$ and $\mathbf{v}_l$ are over $\mathbb{Z}^+[\Delta]$.

Now pick a basis $b_1, \ldots, b_d$ of the free Abelian group Δ. Every $\delta \in \Delta$ can be expressed uniquely as a product of integral powers of $b_1, \ldots, b_d$. Putting $x_i = b_i^t$ we can view δ^t as a monomial of $\mathcal{R} = \mathbb{Z}\left[x_1^\pm, \ldots, x_d^\pm\right]$, thus identifying $\mathbb{Z}[\Delta]$ with $\mathcal{R} = \mathbb{Z}\left[x_1^\pm, \ldots, x_d^\pm\right]$ and $\mathbb{Z}^+[\Delta]$ with $\mathcal{R}^+ = \mathbb{Z}^+\left[x_1^\pm, \ldots, x_d^\pm\right]$. The matrices A, B are then over $\mathcal{R}^+$, and weights of edges in the corresponding graphs are monomials of $\mathcal{R}$. For $x_1, \ldots, x_d > 0$ we have the spectral radius $\beta_A(x_1, \ldots, x_d)$ of the nonnegative matrix $A(x_1, \ldots, x_d)$ and, as in the proof of Theorem 5.1 in [MT91], we find that $\beta_A(x_1, \ldots, x_d) = \beta_B(x_1, \ldots, x_d)$. This gives us the conditions that will be used for the construction of a finitary isomorphism with fect.

2.1. CONDITION:

1. *$A, B, \mathbf{v}_r, \mathbf{v}_l$ are over $\mathbb{Z}^+[\Delta] \subset \mathcal{R}$ and $\Delta_A = \Delta_B = \Delta$.*
2. *$\beta_A(x_1, \ldots, x_d) = \beta_B(x_1, \ldots, x_d)$ for all $x_1, \ldots, x_d > 0$.*
3. *$(A^n \mathbf{v}_r)(I_0) = (\mathbf{v}_l B^n)(J_0)$ for all $n \in \mathbb{Z}^+$.*

2.2. Example ([MT91] 5.7 and 5.8): Let $p(x) = 2x$ and $q(x) = 1 + x^2$. Consider

$$A = \begin{pmatrix} p & 1 \\ y & q \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ pq+y & 0 & p & q \\ y & 0 & p & 0 \\ y & 0 & 0 & q \end{pmatrix} \text{ over } \mathbb{Z}^+\left[x^\pm, y^\pm\right].$$

[MT91] showed that X_A and X_B are not finitely equivalent (hence not almost block isomorphic), and also explicitly constructed a finitely isomorphism with fect between them. Let

$$V = \begin{pmatrix} y & 0 & p & q \\ qy & y & 0 & q^2 - pq \end{pmatrix}.$$

Then $AV = VB$. Since the first row and second column of V are nontrivial and over $\mathbb{Z}^+[x, y]$, we let $I_0 = 1$, $J_0 = 2$, $\mathbf{v}_r = Ve_{J_0}$, and $\mathbf{v}_l = e_{I_0}V$. Then

Condition 2.1 is satisfied, and the existence of a finitary isomorphism with fect follows from 1.1.

3. Weights and boxes

We will need a couple of lemmas regarding the weights of cycles in our graphs.

3.1. LEMMA (cf. [MT91] 5.13): *For A, B satisfying Condition 2.1 and for every δ in Δ there exist cycles α, $\bar{\alpha}$ in A starting at I_0 and cycles β, $\bar{\beta}$ in B starting at J_0 such that*

$$\delta = \frac{wt(\alpha)}{wt(\bar{\alpha})} = \frac{wt(\beta)}{wt(\bar{\beta})},$$
$$l(\alpha) = l(\bar{\alpha}) = l(\beta) = l(\bar{\beta}),$$
$$wt(\alpha) = wt(\beta), \text{ and}$$
$$wt(\bar{\alpha}) = wt(\bar{\beta}).$$

Proof: Since $\delta \in \Delta$ we can find a', $\bar{a}'$ starting at I_0 and passing through all states of A, and b', $\bar{b}'$ starting at J_0 and passing through all states of B with

$$\delta = \frac{wt(a')}{wt(\bar{a}')} = \frac{wt(b')}{wt(\bar{b}')},$$
$$l(a') = l(\bar{a}'), \text{ and}$$
$$l(b') = l(\bar{b}').$$

Let $a = {a'}^{l(b')}$, $\bar{a} = \bar{a}'{a'}^{l(b')-1}$, $b = {b'}^{l(a')}$ and $\bar{b} = \bar{b}'{b'}^{l(a')-1}$. The ratio of weights is still δ, but now the lengths are all equal.

Now recall from [MT91] the weight-per-symbol polytopes $\overline{WPS_A}$, $\overline{WPS_B}$ and the fact (see also [Tun92]) that $\beta_A = \beta_B$ implies $\overline{WPS_A} = \overline{WPS_B}$. Hence, $wt(b)^{\frac{1}{l(b)}} \in \overline{WPS_A}$ and there exists a finite number of cycles, α_i, in A and positive integer n such that

$$wt(\{\alpha_i\}) := \prod_i wt(\alpha_i) = wt(b^n), \text{ and}$$
$$l(\{\alpha_i\}) := \sum_i l(\alpha_i) = l(b^n).$$

Similarly, since $wt(a))^{1/l(a)} \in \overline{WPS_B}$ there exists a finite number of cycles, β_i, in B and positive integer m such that

$$wt(\{\beta_i\}) := \prod_i wt(\beta_i) = wt(a^m), \text{ and}$$
$$l(\{\beta_i\}) := \sum_i l(\beta_i) = l(a^m).$$

Since a passes through every state of A we can splice the cycles α_i into the cycle a; write this as $a\{\alpha_i\}$. Similarly splice the β_i into b and denote the resulting cycle by $b\{\beta_i\}$. Finally, let

$$\begin{aligned}\alpha &:= a^m\{\alpha_i\},\\ \bar{\alpha} &:= \bar{a}a^{m-1}\{\alpha_i\},\\ \beta &:= b^n\{\beta_i\}, \text{ and}\\ \bar{\beta} &:= \bar{b}b^{n-1}\{\beta_i\},\end{aligned}$$

to get the desired result. ∎

3.2. Lemma: *There exist cycles γ_A, γ'_A in A starting at I_0 and passing through every state of A, and a cycle γ_B in B starting at J_0 and passing through every state of B, such that $wt(\gamma_A) = wt(\gamma'_A) = wt(\gamma_B)$, and $l(\gamma_A) = l(\gamma'_A) + 1 = l(\gamma_B)$. Moreover, if $L \in \mathbb{N}$, we can make sure that γ_A, γ_B are not periodic of period p for any $p \leq L$.*

Proof: Pick cycles c_A, c'_A in A starting at I_0 and passing through every state of A, and a cycle c_B in B starting at J_0 and passing through every state of B, with $l(c_A) = l(c'_A) + 1 = l(c_B)$. Now, using the cycles found in Lemma 3.1, let

$$\begin{aligned}\gamma_A &= c_A\bar{\alpha}_{wt(c_A)}\alpha_{wt(c'_A)}\alpha_{wt(c_B)},\\ \gamma'_A &= c'_A\alpha_{wt(c_A)}\bar{\alpha}_{wt(c'_A)}\alpha_{wt(c_B)}, \text{ and}\\ \gamma_B &= c_B\beta_{wt(c_A)}\beta_{wt(c'_A)}\bar{\beta}_{wt(c_B)}.\end{aligned}$$

Clearly, c_A, c_B may be chosen to also ensure that γ_A, γ_B are not periodic of period p for any $p \leq L$. ∎

For each $I \in S(A)$ we view $\mathbf{v}_r(I)$ as a sum of monomials each of which gives a weighted edge from I to J_0. Similarly, for each $J \in S(B)$, we view $\mathbf{v}_l(J)$ as specifying weighted edges from I_0 to J. Taking $n = 0$ in 3 of Condition 2.1 we get $\mathbf{v}_r(I_0) = \mathbf{v}_l(J_0)$, which means that $\mathbf{v}_r(I_0)$ and $\mathbf{v}_l(J_0)$ define the same set of edges from I_0 to J_0. When e is an edge in A, B, $\mathbf{v}_l$, or $\mathbf{v}_r$ we will use α_e, $\bar{\alpha}_e$, β_e, and $\bar{\beta}_e$ to refer to the cycles provided by Lemma 3.1 for δ equal to the weight of e.

Our isomorphism will be constructed by gluing together **boxes**. The top corners of the box will be states $S(A)$, for example I_1, I_2, the top edge, a, will represent a path in A from I_1 to I_2. The bottom corners will be states in $S(B)$, for example J_1, J_2, and the bottom edge, b, will represent a path in B of the same length as a. The left and right sides will be weighted edges $I_1 \to J_1$ and $I_2 \to J_2$, respectively. Finally, whenever we form a box, the product of the weights of the top and right sides will equal the product of the weights of the left and bottom sides. (See Figure 1.)

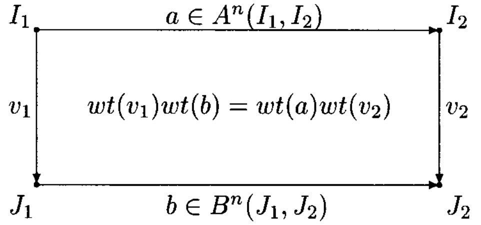

Figure 1. The product of the weights of the top and right sides of our boxes will equal the product of the weights of the left and bottom sides.

In Condition 2.1 we can assume that $\mathbf{v}_r$ and $\mathbf{v}_l$ are over $\mathbb{Z}^+[\Delta]\setminus\{0\}$ by using the fact that A and B are primitive to replace $\mathbf{v}_r$ by $A^N\mathbf{v}_r$ and $\mathbf{v}_l$ by $\mathbf{v}_lB^N$ for suitable $N \in \mathbb{N}$. We pick one of the edges in $\mathbf{v}_r(I_0) = \mathbf{v}_l(J_0)$ and call it v_0.

3.3. Lemma: *There exists a B-cycle w starting at J_0 and, for each $I \in S(A)$ and each edge $v \in \mathbf{v}_r(I)$, there exists an A-path $u(v)$ from I to I_0, such that $l(u(v)) = l(w)$ and $wt(u(v))wt(v_0) = wt(v)wt(w)$. Moreover, we can assume that all the $u(v)$ are different. (See Figure 2.)*

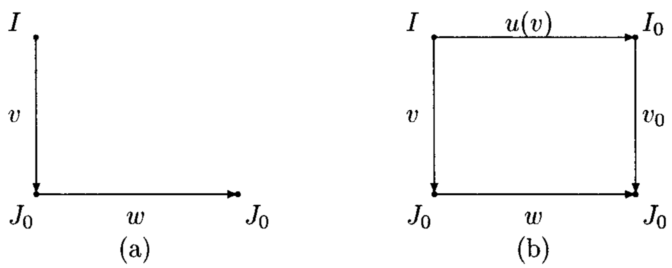

Figure 2. By Lemma 3.3, there exists a word w such that for any edge $v \in \mathbf{v}_r(I)$, as in (a), there exists $u(v)$ to complete the box shown in (b).

Proof: Write $\mathcal{S} = S(A)$ and $\mathbf{v} = \mathbf{v}_r$. We will use a tree in the manner of [AM79]. The tree will have $\{(I,v) \in \mathcal{S} \times \mathbf{v}\colon v \in \mathbf{v}(I)\}$ for its set of vertices, and will be rooted at (I_0, v_0). To grow the tree, we start with a single vertex (I_0, v_0), and iterate the following step for $k \leq |\mathcal{S}| - 1$.

Iteration: Pick $I_k \in \mathcal{S} \setminus \{I_0, \dots, I_{k-1}\}$ such that there exists an edge, a, in A from I_k to I_j for some $j < k$. Now pick $v_k \in \mathbf{v}(I_k)$ and add the vertex (I_k, v_k) to the tree with an edge e_k from (I_k, v_k) to (I_j, v_j). The edges e_k $(k = 1, \dots, |\mathcal{S}| - 1)$ will be called type I edges.

Now for each $i \in \{0, \dots, |\mathcal{S}| - 1\}$ and each $v \in \mathbf{v}(I_i) \setminus v_i$ add a vertex (I_i, v) and a type II edge $(I_i, v) \to (I_i, v_i)$. Label these type II edges $e_{|\mathcal{S}|}, \dots, e_n$, where n is the number of edges in the tree.

For every $I \in \mathcal{S}$ and every $v \in \mathbf{v}(I)$ there is a path from (I, v) to (I_0, v_0) which consists of zero or one type II edges followed by zero or more type I edges. Note that the edges are always transversed with decreasing values of k.

The tree is complete, and we are ready to start the construction of w.

If edge e_k is type I $(k < |\mathcal{S}|)$, then $e_k\colon (I_k, v_k) \to (I_j, v_j)$ for some $j < k$ and there exists an edge a in A with $a\colon I_k \to I_j$. Now let

$$w_k = \gamma_B \bar{\beta}_{v_k} \beta_a \beta_{v_j},$$
$$u_k = a\left(\gamma'_A \alpha_{v_k} \bar{\alpha}_a \bar{\alpha}_{v_j}\right)_{(I_j)},$$
$$u'_k = \gamma_A \bar{\alpha}_{v_k} \alpha_a \alpha_{v_j},$$

where we used the notation $c_{(I)}$ to denote a rotation of the cycle c so that it starts at I. (This is possible since γ'_A passes through every state of A.)

If edge e_k is type II $(k \geq |\mathcal{S}|)$, then $e_k\colon (I_i, v) \to (I_i, v_i)$ for some i. Now let

$$w_k = \gamma_B \gamma_B \bar{\beta}_v \beta_{v_i},$$
$$u_k = (\gamma_A \alpha_v \bar{\alpha}_{v_i})\, \gamma_A, \text{ and}$$
$$u'_k = \gamma_A\, (\gamma_A \bar{\alpha}_v \alpha_{v_i}).$$

Here we used two copies of γ_A and γ_B so that we can ensure $u_k \neq u'_k$ by taking $L = 2\max\{l(\alpha_v)\colon v \in \mathbf{v}_r \text{ or } v \in \mathbf{v}_l\}$ in 3.2. Notice that in each case

$$l(w_k) = l(u_k) = l(u'_k), \text{ and}$$
$$wt(w_k) = wt(u'_k).$$

Also, for edges of type I $(e_k\colon (I_k, v_k) \to (I_j, v_j))$ we have

$$wt(u_k)wt(v_j) = wt(v_k)wt(w_k),$$

while for edges of type II (e_k: $(I_i, v) \to (I_i, v_i)$) we have

$$wt(u_k)wt(v_i) = wt(v)wt(w_k).$$

This means that for every edge e_k: $(I, v) \to (I', v')$ we can make boxes as in Figure 3.

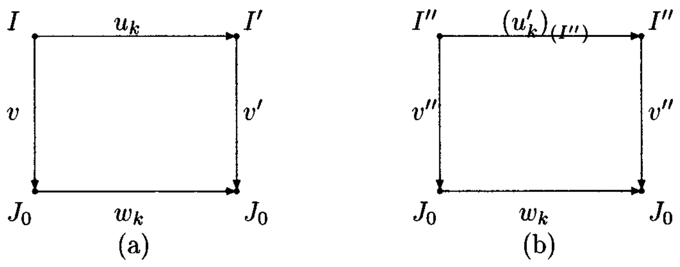

Figure 3. Boxes for (a) moving down the tree and (b) staying at a vertex of the tree.

We define the word w to be the concatenation $w = w_n w_{n-1} \cdots w_1$. Each word $u(v)$ is put together in a similar way, using u_k and u'_k: Look at the path $(I, v) \to (I_0, v_0)$. If e_k is on this path, put $u(v)_k = u_k$; otherwise, put $u(v)_k = u'_k$. Then rotate each $u(v)_k$ appropriately to define $u(v)$ as the concatenation $u(v) = u(v)_n u(v)_{n-1} \cdots u(v)_1$. Since u'_k allows us to stay at a vertex of the tree while traversing w_k in B and u_k allows us to move down the tree while traversing w_k, the proof of the lemma is completed by picking (see Figure 3) the corresponding boxes and gluing them together. ∎

4. Proof of the theorem

We will continue to regard $\mathbf{v}_r(I)$ as specifying edges from I to J_0, and $\mathbf{v}_l(J)$ as specifying edges from I_0 to J. If a is one of these edges, or if a is a path (word) in one of the graphs $G(A)$ or $G(B)$, we will denote its starting state by $s(a)$ and its terminal state by $t(a)$.

The proof of Theorem 1.1 consists of three steps.

STEP 1: *Construction of markers.* Let the B-word w_B and the A-words $u_A(v)$ be as in 3.3. Similarly find an A-word w_A from I_0 to I_0 and, for each $v \in \mathbf{v}_l(J)$, a corresponding B-word $u_B(v)$ from I_0 to J such that $l(w_A) = l(u_B(v))$ and

$wt(w_A)wt(v) = wt(v_0)wt(u_B(v))$. If we put $w_A(v) = u_A(v)w_A$ for $v \in \mathbf{v}_r$ and $w_B(v) = w_Bu_B(v)$ for $v \in \mathbf{v}_l$ and define $\mathcal{W}_A = \{w_A(v): v \in \mathbf{v}_r\}$ and $\mathcal{W}_B = \{w_B(v): v \in \mathbf{v}_l\}$, then for each $v \in \mathbf{v}_r(I)$ and each $v' \in \mathbf{v}_l(J)$ we obtain from Lemma 3.3 a box as in Figure 4.

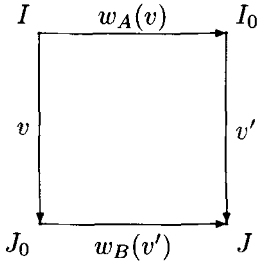

Figure 4. For each $v \in v_r(I)$ and $v' \in v_l(J)$ we have a box with $wt(v)wt(w_B(v')) = wt(w_A(v))wt(v')$.

We will modify the words $w_A(v), w_B(v')$ to ensure that, in addition, words in $\mathcal{W}_A$ and $\mathcal{W}_B$ do not overlap.

Take a simple cycle $a = i_0 \cdots i_m$ starting and ending at I_0, and label the states it passes through $I_0, \ldots, I_m$ so that $i_k\colon I_k \to I_{k+1}$ (with the convention that $I_{m+1} = I_0$). Find I_k so that there exists an edge $j_0 \neq i_k$ with $s(j_0) = I_k$. Then find a path $j_0 \cdots j_p$ of minimal length back to one of the states I_l. If $l = 0$ let $b = i_0 \cdots i_{k-1} j_0 \cdots j_p$; otherwise let $b = i_0 \cdots i_{k-1} j_0 \cdots j_p i_{l+1} \cdots i_m$. The cycles a and b only pass through the state I_0 at their start and end, hence they have no nontrivial overlap.

There exists $k > 3$ so that $a^k b^k$ is not contained in $w\gamma_A^n$ for any $w \in \mathcal{W}_A$ and $n \in \mathbb{Z}^+$. Let $c_1 = aaaa^kb^k$ and $c_2 = aaba^kb^k$. Considering A-words g_A and $n \in \mathbb{Z}^+$ with $l(c_1), l(c_2), l(g_A) < l(w\gamma_A^n)$, note that no two words of the form $w\gamma_A^n c_1 g_A c_2$ overlap. Similarly we can find words d_1, d_2 such that, for $n \in \mathbb{Z}^+$ and $w \in \mathcal{W}_B$ with $l(d_1), l(d_2), l(g_B) < l(\gamma_B^n w)$, no two words of the form $d_1 g_B d_2 \gamma_B^n w$ overlap.

Let c' be an A-cycle at I_0 and d' a B-cycle at J_0 such that $l(c') = l(d_1d_2) + L$ and $l(d') = l(c_1c_2) + L$ for some $L \in \mathbb{Z}^+$. Put $g_A = \alpha_{d_1d_2}\bar{\alpha}_{c_1c_2}c'\alpha_{d'}\bar{\alpha}_{c'}$ and $g_B = \bar{\beta}_{d_1d_2}\beta_{c_1c_2}d'\bar{\beta}_{d'}\beta_{c'}$. Now $c_1g_Ac_2$ and $d_1g_Bd_2$ have the same length and weight. We replace each $u_A(v)w_A \in \mathcal{W}_A$ by $w_A(v) = u_A(v)w_A\gamma_A^n c_1 g_A c_2$ and each $w_Bu_B(v) \in \mathcal{W}_B$ by $w_B(v) = d_1 g_B d_2 \gamma_B^n w_B u_B(v)$. Then we have a box as in Figure 4 for each $v \in \mathbf{v}_r(I)$ and each $v' \in \mathbf{v}_l(J)$. In addition, words in $\mathcal{W}_A$ have no nontrivial overlaps, and neither do words in $\mathcal{W}_B$. The sets $\mathcal{W}_A$ and $\mathcal{W}_B$ will be used as markers for our isomorphism.

STEP 2: *Bijections between words not containing markers.* Let us denote the set of A-words of length n by $W_n(A)$ and put

$$\begin{aligned}\mathcal{A}'_n =& \{(a,v) \in W_n(A) \times \mathbf{v}_r \colon s(a) = I_0, t(a) = s(v)\},\\ \mathcal{A}_n =& \{(a,v) \in \mathcal{A}'_n \colon a \text{ does not contain a word in } \mathcal{W}_A\},\\ \mathcal{B}'_n =& \{(v,b) \in \mathbf{v}_l \times W_n(B) \colon t(v) = s(b), t(b) = J_0\},\\ \mathcal{B}_n =& \{(v,b) \in \mathcal{B}'_n \colon b \text{ does not contain a word in } \mathcal{W}_B\}.\end{aligned}$$

For $n = 1, 2, \ldots$, we will find bijections $\phi_n \colon \mathcal{B}_n \to \mathcal{A}_n$ such that $\phi_n(v,b) = (a,v')$ implies $wt(v)wt(b) = wt(a)wt(v')$. By Condition 2.1, we already have such weight-preserving bijections between $\mathcal{B}'_n$ and $\mathcal{A}'_n$. Note that all words in $\mathcal{W}_A$, $\mathcal{W}_B$ have the same length, L. If $n < L$, we have $\mathcal{A}_n = \mathcal{A}'_n$, $\mathcal{B}_n = \mathcal{B}'_n$ and we can take ϕ_n to be any bijection from $\mathcal{B}_n$ to $\mathcal{A}_n$. Assume, for an induction, that we have weight-preserving bijections $\phi_m \colon \mathcal{B}_m \to \mathcal{A}_m$ for $m < n$. We can then construct a bijection $\phi'_n \colon \mathcal{B}'_n \setminus \mathcal{B}_n \to \mathcal{A}'_n \setminus \mathcal{A}_n$ as follows. Consider $(v,b) \in \mathcal{B}'_n \setminus \mathcal{B}_n$ and, for suitable $p \in \mathbb{N}$, write

$$b = b_1 w_B(v_1) b_2 w_B(v_2) \cdots b_p w_B(v_p) b_{p+1},$$

where $w_B(v_i) \in \mathcal{W}_B$, $b_i \in \mathcal{B}_{l(b_i)}$, and b_1 and b_{p+1} are allowed to equal the empty word. Put $v_0 = v$ and, for $i = 1, \ldots, p+1$, let $(a_i, v'_i) = \phi_{l(b_i)}(v_{i-1}, b_i)$. Then define $v' = v'_{p+1}$,

$$a = a_1 w_A(v'_1) a_2 w_A(v'_2) \cdots a_p w_a(v'_p) a_{p+1},$$

and $\phi_n(v,b) = (a,v')$. This construction is depicted in Figure 5.

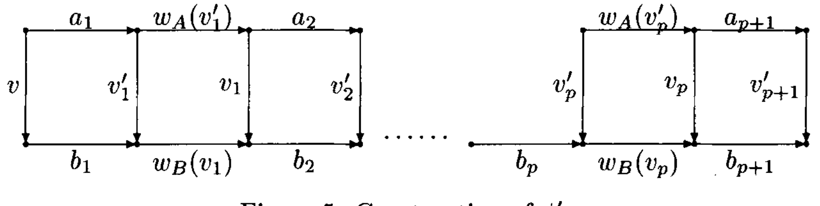

Figure 5. Construction of ϕ'_n.

Note that the markers in a and b occur in exactly the same places. Starting with (a,v') and using ϕ_m^{-1} instead of ϕ_m, we get ϕ_n^{-1} and see that $\phi_m^{-1}(a,v') = (b,v)$. Clearly, $wt(v)wt(b) = wt(a)wt(v')$. Thus, $\phi'_n \colon \mathcal{B}'_n \setminus \mathcal{B}_n \to \mathcal{A}'_n \setminus \mathcal{A}_n$ is a weight-preserving bijection. Since there also exists a weight-preserving bijection between $\mathcal{B}'_n$ and $\mathcal{A}'_n$, it is now easy to find such a bijection between $\mathcal{B}_n$ and $\mathcal{A}_n$.

STEP 3: *Construction of the finitary isomorphism with fect.* We construct a bijection ϕ from the set of $x \in X_B$ which enter the closed-open set

$$\{x' \in X_B \colon x'_1 \cdots x'_L \in \mathcal{W}_B\}$$

infinitely often under both σ_B and σ_B^{-1}; the bijection will be onto the set of points in X_A which enter (the closed-open set corresponding to) $\mathcal{W}_A$ infinitely often under both σ_A and σ_A^{-1}. We describe $y = \phi(x)$ between two consecutive occurences of words in $\mathcal{W}_B$: Suppose $w_B(v_0), w_B(v_1) \in \mathcal{W}_B$ and $b_1 \in \mathcal{B}_n$ are such that for some $0 \le i < L + N$ we have

$$x_{-i}x_{-i+1} \cdots x_{-i+2L+n-1} = w_B(v_0)b_1w_B(v_1).$$

The word $w_B(v_0)b_1w_B(v_1)$ determines in y the word $a_1w_A(v'_1)$, where $\phi_n(v_0, b_1) = (a_1, v'_1)$ and the position of $a_1w_A(v'_1)$ in y is the same as that of $b_1w_B(v_1)$ in x. For three consecutive occurences of words in $\mathcal{W}_B$, the word determined in $y = \phi(x)$ by $w_B(v_{-1})b_0w_B(v_0)b_1w_B(v_1)$ may be pictured in terms of boxes as in Figure 6.

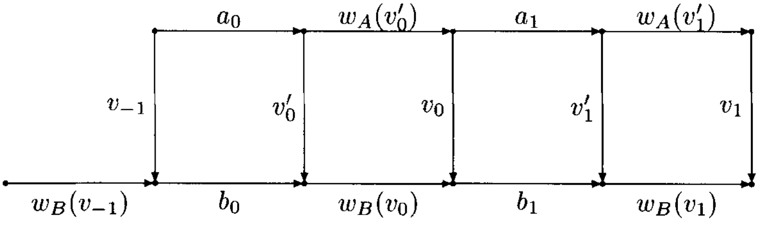

Figure 6. Construction of the isomorphism.

It is easy to see that the inverse ϕ^{-1} is constructed in the same way: the word $w_A(v'_0)a_1w_A(v'_1)$ in $y \in X_A$ determines in $x = \phi^{-1}(y)$ the word $w_B(v_0)b_1$, where $\phi_n^{-1}(a_1, v'_1) = (v_0, b_1)$ and $w_B(v_0)ab_1$ occurs in the same position in x as $w_A(v'_0)a_1$ does in y.

The maps ϕ, ϕ^{-1} are clearly defined a.e. and finitary. They are measure-preserving as a result of the weight-preserving condition on our boxes. Finally, they have fect since, by Kac's theorem [Pet83], the expected first-return time is finite for each of $\mathcal{W}_B, \mathcal{W}_A$.

References

[AM79] R. Adler and B. Marcus, *Topological entropy and equivalence of dynamical systems*, Memoirs of the American Mathematical Society **219** (1979), 1–84.

[Ash90] J. Ashley, *Bounded-to-1 factors of an aperiodic shift of finite type are 1-to-1 almost everywhere factors also,* Ergodic Theory and Dynamical Systems **10** (1990), 615–625.

[EMT] M. Einsiedler and R. Mouat and S. Tuncel, *When does a submodule of* $\mathbb{R}[x_1,\ldots,x_k]^n$ *contain a positive element?,* preprint.

[Kri83] W. Krieger, *On the finitary isomorphisms of Markov shifts that have finite expected coding time,* Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **65** (1983), 323–328.

[KS79a] M. Keane and M. Smorodinsky, *Bernoulli schemes of the same entropy are finitarily isomorphic,* Annals of Mathematics **109** (1979), 397–406.

[KS79b] M. Keane and M. Smorodinsky, *Finitary isomorphisms of irreducible Markov shifts,* Israel Journal of Mathematics **34** (1979), 281–286.

[MT91] B. Marcus and S. Tuncel, *The weight-per-symbol polytope and scaffolds of invariants associated with Markov chains,* Ergodic Theory and Dynamical Systems **11** (1991), 129–180.

[Par79] W. Parry, *Finitary isomorphisms with finite expected code lengths,* The Bulletin of the London Mathematical Society **11** (1979), 170–176.

[Par91] W. Parry, *Notes on coding problems for finite state processes,* The Bulletin of the London Mathematical Society **23** (1991), 1–33.

[Pet83] K. Petersen, *Ergodic Theory,* Cambridge University Press, 1983.

[PS84] W. Parry and K. Schmidt, *Natural coefficients and invariants for Markov shifts,* Inventiones Mathematicae **76** (1984), 15–32.

[PT81] W. Parry and S. Tuncel, *On the classification of Markov chains by finite equivalence,* Ergodic Theory and Dynamical Systems **1** (1981), 303–335.

[PT82] W. Parry and S. Tuncel, *On the topological and stochastic structure of Markov chains,* The Bulletin of the London Mathematical Society **14** (1982), 16–27.

[Sch84] K. Schmidt, *Invariants for finitary isomorphisms with finite expected code lengths,* Inventiones Mathematicae **76** (1984), 33–40.

[Sch86] K. Schmidt, *Hyperbolic structure preserving isomorphisms of Markov shifts,* Israel Journal of Mathematics **55** (1986), 213–228.

[Sch87] K. Schmidt, *Hyperbolic structure preserving isomorphisms of Markov shifts II,* Israel Journal of Mathematics **58** (1987), 225–242.

[Tun81] S. Tuncel, *Conditional pressure and coding,* Israel Journal of Mathematics **39** (1981), 101–112.

[Tun92] S. Tuncel, *Faces of Markov chains and matrices of polynomials,* Contemporary Mathematics **135** (1992), 391–422.