ISRAEL JOURNAL OF MATHEMATICS, Vol. 38, No. 4, 1981

THE CLASSIFICATION OF TOPOLOGICAL MARKOV CHAINS ADAPTED SHIFT EQUIVALENCE

BY

WILLIAM PARRY[†]

ABSTRACT

This paper is motivated by Williams's problem: to show that shift equivalence of non-negative irreducible matrices implies topological equivalence of their associated topological Markov chains. Instead we prove that *adapted* shift equivalence implies (and is implied by) topological equivalence.

Let S be an irreducible $k \times k$ zero-one matrix and let X_S be the shift invariant subset of $\Pi_{n=-\infty}^{\infty}\{1,2,\cdots,k\}$ consisting of those sequences $x = \{x_n\}$ such that $S(x_n, x_{n+1}) = 1$ for all $n \in \mathbf{Z}$. (The shift sends x to y where $y_n = x_{n+1}$.) The restriction of the shift to X_S is denoted by σ_S and (S_X, σ_S) is called a *topological Markov chain.*

The main problem which concerns us here is the topological classification of topological Markov chains. Under what conditions does there exist a topological conjugacy between two topological Markov chains (X_S, σ_S), (X_T, σ_T), i.e., a homeomorphism ϕ of X_S onto X_S such that $\phi\sigma_S = \sigma_T\phi$?

In [3], Williams defines S, T to be *strong shift equivalent* if there exist non-negative integral rectangular matrices such that $S = U_1V_1$, $V_1U_1 = U_2V_2$, $V_2U_2 = \cdots = T$. He proves that σ_S, σ_T are topologically conjugate if and only if S, T are strong shift equivalent.

It is easy to see, as Williams shows, that strong shift equivalence implies *shift equivalence* in the sense that there exists a positive integer l and non-negative integral rectangular matrices U, V such that

[†] Supported by the Deutsche Forschungsgemeinschaft.
Received April 25, 1980

$$US = TU, \quad SV = VT, \quad UV = T^l, \quad VU = S^l.$$

(l is called the *lag* of the equivalence.)

It is conjectured (cf. [3]) that shift equivalence implies topological conjugacy but a proof is still lacking. I was invited to discuss this problem by the Sonderforschungsbereich Stochastische Mathematische Modelle, Institut für Angewandte Mathematik, Universität Heidelberg (Summer 1979), and I would like to take this opportunity to thank Professor W. Krieger for the invitation and for the useful discussions which ensued.

§1. On shift equivalence

This seems an appropriate place to clarify an obscurity, pointed out by Krieger, which occurs in [12] concerning the definition of shift equivalence and a weaker version (cf. Remark 4 after Proposition 5.1). Here is a complete proof.

PROPOSITION 1. *In the definition of shift equivalence of irreducible non-negative integral matrices, U, V may be taken as rectangular matrices over* $\mathbf{Z}$ *rather than over* $\mathbf{Z}^+$.

PROOF. Suppose $US = TU$, $SV = VT$, $UV = T^l$, $VU = S^l$, where U, V are defined over $\mathbf{Z}$ and S, T are defined over $\mathbf{Z}^+$. We shall assume that S, T are aperiodic. A similar proof can be given in the periodic case. If β is the maximum eigenvalue of S and T then S^n/β^n converges to a matrix with columns $\lambda_1 r, \cdots, \lambda_k r$ where $Sr = \beta r$ and r is a strictly positive vector. Moreover $\beta(Ur) = T(Ur)$ and since Ur cannot be the zero vector by virtue of the equation $VU = S^l$, Ur must be "the" vector corresponding to the maximum eigenvalue β of T, i.e., Ur must be strictly positive or strictly negative. Hence US^n/β^n converges to $\lambda_1 Ur, \cdots, \lambda_k Ur$. Thus US^n/β^n converges to a matrix, each column of which is pure positive or pure negative. A similar argument can be given concerning the rows of the limit of US^n/β^n. Hence US^n/β^n converges to a strictly positive or strictly negative matrix. Thus for large n US^n is "pure". Similarly VT^n is "pure" for large n. (We can select the same integer n.)

Moreover $US^nVT^n = UVT^{2n} = T^{2n+l}$ which is strictly positive. Hence US^n, VT^n have the same sign, which we may take to be positive, by replacing U, V by $-U$, $-V$ if necessary. In conclusion we see that

$$(US^n)S = T(US^n), \qquad S(VT^n) = (VT^n)T,$$

$$(US^n)(VT^n) = T^{2n+l}, \qquad (VT^n)(US^n) = S^{2n+l},$$

i.e., S, T are shift equivalent with lag $2n + l$.

§2. Conjugacy and partitions

The last section was something of a diversion. Let (X_S, σ_S), (X_T, σ_T) be two topological Markov chains which are topologically conjugate by ϕ. Let α be the *standard partition* (or *state partition*) of X_S, i.e., $\alpha = (A_1, \cdots, A_k)$ where $A_i = \{x : x_0 = i\}$. If α' is the state partition of X_T then it can be pulled back to X_S using ϕ^{-1} and a suitable power of σ_S to yield a partition β satisfying

$$\text{(2.1)} \qquad \begin{cases} \text{(a)}\ \alpha^n = \alpha \vee \sigma_S^{-1}\alpha \vee \cdots \vee \sigma_S^{-n}\alpha \geqq \beta, \\ \text{(b)}\ \beta^n \geqq \sigma_S^{-n}\alpha. \end{cases}$$

For the moment we shall drop the suffix S and derive a sequence of "simpler" relationships between α, β. The following is a consequence of (2.1)(a):

$$\text{(2.2)} \qquad \begin{array}{l} (\alpha^{2n})^1 \geqq \alpha^{2n} \vee \beta^{n+1} \geqq \alpha^{2n} \\ (\alpha^{2n} \vee \beta^{n+1})^1 \geqq \alpha^{2n} \vee \beta^{n+2} \geqq \alpha^{2n} \vee \beta^{n+1} \\ \text{-------------------------} \\ (\alpha^{2n} \vee \beta^{2n-1})^1 \geqq \alpha^{2n} \vee \beta^{2n} \geqq \alpha^{2n} \vee \beta^{2n-1}. \end{array}$$

Given two partitions ξ, η we define zero-one matrices indexed by $\xi \times \eta$:

$$\begin{aligned} (\xi, \eta)(A, B) &= 1 \quad \text{if } A \cap B \neq \emptyset, \\ &= 0 \quad \text{otherwise;} \\ (\xi, \eta)_\sigma(A, B) &= 1 \quad \text{if } A \cap \sigma^{-1}B \neq \emptyset, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

When $\xi \vee \sigma^{-1}\xi \geqq \eta \geqq \xi$ it can be proved (cf. [2]) that

$$(\xi, \eta)(\eta, \xi)_\sigma = (\xi, \xi)_\sigma,$$

$$(\eta, \xi)_\sigma(\xi, \eta) = (\eta, \eta)_\sigma.$$

As a consequence of (2.2) we then see that $(\alpha^{2n}, \alpha^{2n})_\sigma$ and $(\alpha^{2n} \vee \beta^{2n}, \alpha^{2n} \vee \beta^{2n})_\sigma$ are strong shift equivalent and the number of steps in this equivalence is n.

However, using (2.1)(b) rather than (2.1)(a) (and noticing that (a) and (b) are symmetrically related with respect to negative and positive iterations of σ) we see that $(\beta^{2n}, \beta^{2n})_\sigma$ and $(\alpha^{2n} \vee \beta^{2n}, \alpha^{2n} \vee \beta^{2n})_\sigma$ are strong shift equivalent "in n steps".

To summarise we have

PROPOSITION 2. *$(\alpha^{2n}, \alpha^{2n})_\sigma$ and $(\beta^{2n}, \beta^{2n})_\sigma$ are strong shift equivalent in $2n$ steps.*

COROLLARY (Williams [3]). *If σ_S and σ_T are topologically conjugate then S and T are strong shift equivalent.*

PROOF. In fact $(\alpha, \alpha)_\sigma$ is strong shift equivalent to $(\alpha^{2n}, \alpha^{2n})_\sigma$ (in $2n$ steps) and $(\beta, \beta)_\sigma$ is strong shift equivalent to $(\beta^{2n}, \beta^{2n})_\sigma$ (in $2n$ steps). To complete the proof we note that with a suitable indexing $(\alpha, \alpha)_\sigma = S$ and $(\beta, \beta)_\sigma = T$.

If we interpret a $k \times k$ zero-one matrix S as a directed graph, that is as k vertices with transitions between i and j allowed if and only if $S(i, j) = 1$, then from the matrix S we can create new matrices S_n with θ_n vertices, where θ_n is the number of $(i_0, i_1, \cdots, i_n)$ such that $S(i_m, i_{m+1}) = 1$. The vertices are precisely such "allowable words" $(i_0, i_1, \cdots, i_n)$ and $S_n(i_0, i_1, \cdots, i_n; j_0, \cdots, j_n) = 1$ if and only if $i_1 = j_0, \cdots, i_n = j_{n-1}$.

Evidently $(\alpha^n, \alpha^n)_\sigma = S_n$ and the following definition is indicated: The zero-one matrices S, T are said to be *adapted* (adapted shift equivalent) if there exists $n \in \mathbf{Z}^+$ such that S_n and T_n are shift equivalent with lag n. In view of the above remarks we have proved (with respect to the lag $2n$ of Proposition 2):

THEOREM 1. *If σ_S and σ_T are topologically conjugate then S and T are adapted.*

So far we have been considering two-sided topological Markov chains, that is, shifts defined on doubly infinite sequences. One-sided topological Markov chains (X'_S, σ'_S) are defined as shifts on the space of one-sided sequences

$$X'_s = \left\{ x \in \prod_{n=0}^{\infty} \{1, 2, \cdots, k\} : S(x_n, x_{n+1}) = 1 \right\},$$

$$\sigma'_S(x) = y \qquad \text{where } y_n = x_{n+1}.$$

If we define $V_n(S)$ to be the group of integer valued functions f on X'_S which are dependent on only the first $n+1$ variables ($f(x) = f(x_0, x_1, \cdots, x_n)$) then it is easy to see that the transpose S^* of the matrix S has an interpretation as a homomorphism of $V_0(S)$ to itself:

$$(S^* f)(x) = \sum_{\sigma_S y = x} f(y).$$

The transpose S^*_n of the matrix S_n then has the interpretation

$$(S^*_n f)(x) = \sum_{\sigma_S y = x} f(y)$$

which is a homomorphism of $V_n(S)$ into $V_{n-1}(S) \subset V_n(S)$. Evidently adapted shift equivalence has the following meaning: There exists $n \in \mathbf{Z}^+$ and order-preserving homomorphisms $A: V_n(S) \to V_n(T)$, $B: V_n(T) \to V_n(S)$ such that

$$AS_n^* = T_n^*A, \qquad S_n^*B = BT_n^*,$$

$$BA = S_n^{*n}, \qquad AB = T_n^{*n}.$$

§3. Adapted shift equivalence

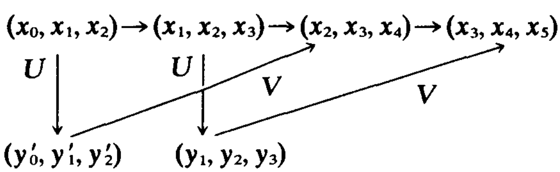
In the last section we showed that topological conjugacy implies adapted shift equivalence. We shall now prove the converse.

THEOREM 2. *If S, T are adapted shift equivalent then the topological Markov chains (X_S, σ_S), (X_T, σ_T) are topologically conjugate.*

PROOF. For convenience of presentation we shall assume that the lag is 2 so that

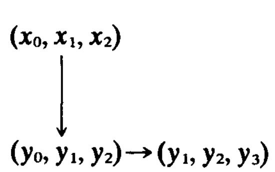
$$US_2 = T_2U, \qquad S_2V = VT_2,$$

$UV = T_2^2$, $VU = S_2^2$, for some non-negative integral matrices U, V. Arrows will indicate allowable transitions. Vertices for S_2 are triples, so that we have, for example,

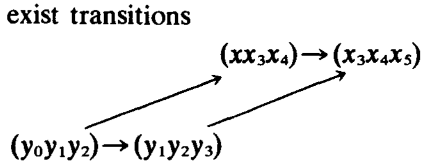


The question arises as to whether there is a transition from (y_0', y_1', y_2') to (y_1, y_2, y_3).

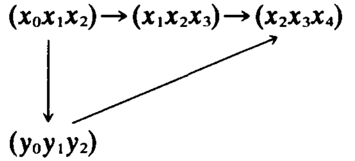
We see from commutativity that there must exist (y_0, y_1, y_2) such that



and therefore there exist transitions



However "x" above must necessarily be x_2 to comply with the commutativity

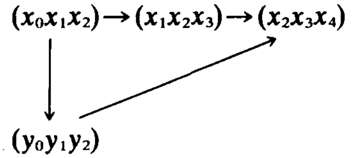


Thus there are the two paths from (x_0, x_1, x_2) to (x_2, x_3, x_4) passing through the "y" triples (y_0', y_1', y_2') and (y_0, y_1, y_2) which correspond to a unique "x" path $(x_0, x_1, x_2) \to (x_2, x_3, x_4)$. We conclude that $(y_0, y_1, y_2) \equiv (y_0', y_1', y_2')$ and therefore the transition $(y_0', y_1', y_2') \equiv (y_0, y_1, y_2) \to (y_1, y_2, y_3)$ is allowed.

Let $x \in X_S$, then we define

$$\phi(x_0x_1x_2x_3x_4) = (y_0y_1y_2)$$

where the transitions

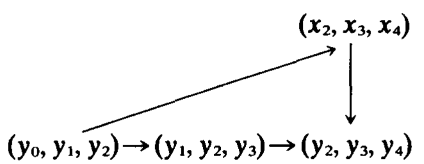


are allowed. Evidently

$$\phi(x) = \{\phi(x_n, x_{n+1}, \cdots, x_{n+4})\}$$

is then a well defined continuous map of X_S into X_T.

In a similar way we can define, for $y \in X_T$, $\psi(y_0, y_1, y_2y_3y_4) = (x_2, x_3, x_4)$ where the transitions



are allowed and

$$\psi(y) = \{\psi(y_ny_{n+1}, \cdots, y_{n+4})\}.$$

Again ψ is a well defined continuous map of X_T into X_S. It is clear that

$$\phi\sigma_S = \sigma_T\phi, \qquad \psi\sigma_T = \sigma_S\psi$$

and

$$\psi\phi = \sigma_S^2, \qquad \phi\psi = \sigma_T^2.$$

Hence ϕ, ψ are surjective and since σ_S, σ_T are homeomorphisms it follows that ϕ, ψ are also homeomorphisms. Thus σ_S, σ_T are topologically conjugate.

§4. Williams's problem

So far we are unable to solve Williams's problem: does shift equivalence imply topological conjugacy? We have presented Theorems 1 and 2 in the hope that they may lead to a solution, since there is some hope that shift equivalence may imply adapted shift equivalence.

We conclude with some observations in this direction using Williams's technique of splitting matrices into products of division and amalgamation matrices.

In the following we shall only consider rectangular non-negative integral matrices with non-trivial rows and non-trivial columns.

Any such matrix M can be written as the product $M = DA$ of a *division* matrix (all of whose columns are unit vectors, no trivial rows) and an *amalgamation* matrix (a transpose of a division matrix). Moreover if $M = D_1A_1$ where D_1, A_1 are division and amalgamation matrices respectively then $D_1 = DP$ and $A_1 = P^{-1}A$, where P is a permutation matrix. The product of two divisions is a division and the product of two amalgamations is an amalgamation.

Suppose we are given a shift equivalence between two zero-one irreducible matrices: can we use this information to construct an adapted shift equivalence? Specifically suppose

$$US = TU, \quad SV = VT, \quad UV = T^2, \quad VU = S^2; \tag{4.1}$$

does there exist U_2, V_2 such that

$$U_2S_2 = T_2U_2, \quad S_2V_2 = V_2T_2, \quad U_2V_2 = T_2^2, \quad V_2U_2 = S_2^2? \tag{4.2}$$

Without going into the lengthy details we can answer this question negatively with the following example:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

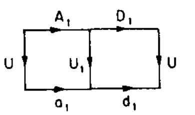
$$\begin{pmatrix}1&0\\0&1\\1&0\\0&1\end{pmatrix}\begin{pmatrix}1&1&1&1\\1&1&1&1\end{pmatrix}=\begin{pmatrix}1&1&0&0\\0&0&1&1\\0&0&1&1\\1&1&0&0\end{pmatrix}^2, \quad \begin{pmatrix}1&1&1&1\\1&1&1&1\end{pmatrix}\begin{pmatrix}1&0\\0&1\\1&0\\0&1\end{pmatrix}=\begin{pmatrix}1&1\\1&1\end{pmatrix}^2.$$

(However, if

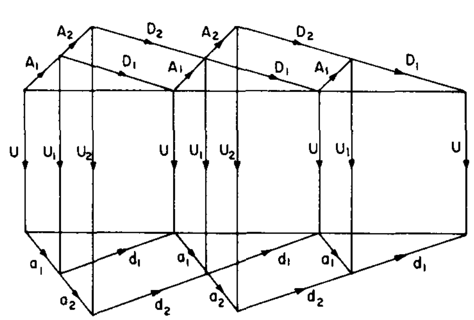
$$U=\begin{pmatrix}1&0\\0&1\\1&0\\0&1\end{pmatrix} \text{ is replaced by } \begin{pmatrix}1&0\\0&1\\0&1\\1&0\end{pmatrix}$$

then matrices U_2, V_2 satisfying (4.2) can be found. Brian Marcus brought my attention to the curious nature of this shift equivalence, in a rather different context.)

For the moment we shall ignore this counterexample and work positively toward a solution of (4.2). We assume that (4.1) holds. In particular we have $US = TU$. This equation alone implies the existence of U_1 and amalgamations A_1, a_1 and divisions D_1, d_1 such that $S = D_1A_1$, $T = d_1a_1$, where

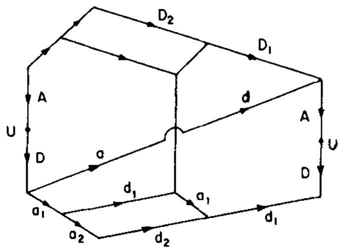


is commutative (cf. [1]). We can now repeat the same argument with respect to the equation $U_1(A_1D_1) = (a_1d_1)U_1$ (i.e. $U_1S_1 = T_1U_1$) to obtain the commutative diagram:

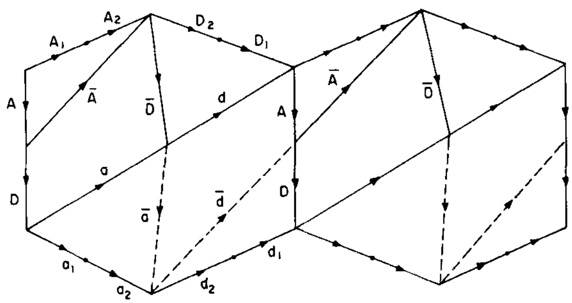


Thus we have $U_2S_2 = T_2U_2$ since $S_2 = A_2D_2$, $T_2 = a_2d_2$. In a similar way we can obtain an equation $S_2V_2 = V_2T_2$. In other words the first two equations of (4.2) can be obtained.

Now let us look at the third and fourth equations of (4.2). From (4.1) we obtain the commutative diagram:

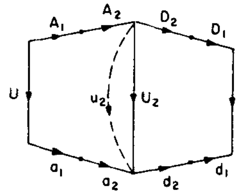


where $U = DA$, $V = da$ are division-amalgamation splittings. aD can also be split, say, $aD = \bar{D}\bar{A}$. Thus $d\bar{D}\bar{A}A = D_1D_2A_2A_1$ are two splittings. Hence by a suitable permutation modification, if necessary, we have the commutative diagram:



where the amalgamation and division matrices $\bar{a}$, $\bar{d}$ are provided in an analogous way to a, d (i.e. $Ad = \bar{d}\bar{a}$, $d_1d_2a_2a_1 = D\bar{d}\bar{a}a$).

If we define $u_2 = \bar{a}\bar{D}$ and $v_2 = \bar{A}\bar{d}$ then $v_2u_2 = S_2^2$, $u_2v_2 = T_2^2$ so that the third and fourth equations of (4.2) are satisfied. However, from our counterexample we cannot claim that $u_2 = U_2$, $v_2 = V_2$, i.e., we cannot claim that all four equations of (4.2) are satisfied simultaneously. Although we cannot guarantee that $u_2S_2 = T_2u_2$, $S_2v_2 = v_2T_2$, it is obvious that $u_2S_2^2 = T_2^2u_2$, $S_2^2v_2 = v_2T_2^2$. The problem is that we have two maps u_2, U_2 satisfying the commutative diagram:



If we write $u_2 = \delta\alpha$ and $U_2 = \delta'\alpha'$ as amalgamation-division splittings then it is clear that $\delta = \delta' P$ and $\alpha = Q\alpha'$ for some permutations P, Q. In other words, u_2, U_2 differ by the existence of a permutation "inside" the amalgamation-division splitting.

This is as far as we can go. Since we have a counter-example we know that there can exist a permutation obstruction to deriving (4.2) from (4.1) by way of our procedure. Perhaps we may have to replace (4.1) by other shift equivalences before adopting our procedure. In fact, in our counter-example if we replace U by

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

then (4.2) can be satisfied. Perhaps instead of U one should consider US^n for various $n \in \mathbf{Z}^+$ and instead of V one should consider S^nV.

References

1. W. Parry, *A finitary classification of topological Markov chains and Sofic systems*, Bull. London Math. Soc. **9** (1977), 86–92.

2. W. Parry and R. F. Williams, *Block coding and a zeta function for finite Markov chains*, Proc. London Math. Soc. (3) **35** (1977), 483–495.

3. R. F. Williams, *Classification of subshifts of finite type*, Ann. of Math. **98** (1973), 120–153; *Errata*, Ann. of Math. **99** (1974), 380–381.

Mathematics Institute
University of Warwick
Coventry CV4 7AL, England