PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 92, Number 2, October 1984

A NOTE ON STATE SPLITTING

JOEL FRIEDMAN

ABSTRACT. Using state splitting methods, we construct nice matrix representatives for arbitrary subshifts of finite type.

In this paper, we will prove the following result.

THEOREM 1. *Let Σ_A be a subshift of finite type (SSFT) with entropy $h(\Sigma_A)$ satisfying*

$$\log n \leq h(\Sigma_A) < \log n + 1.$$

Then there is an SSFT, Σ_B, topologically conjugate to Σ_A with
(i) *each row sum of B either n or $n+1$,*
(ii) *the set of column sums of B is the same as that of A.*

It is known that using the method of symbol splitting (see [**ACH** and **M2**]), one can construct SSFT's conjugate to Σ_A either with all row sums $\geq n$ or with all row sums $\leq n+1$. In our theorem, the inequalities $\geq n$ and $\leq n+1$ are achieved simultaneously. It is known that the above theorem would follow from the lemma below, which was conjectured by Ethan Coven and Brian Marcus at the 1982 AMS Summer Conference in Ergodic Theory (see the methods of [**ACH** and **M1**]).

In addition, given Σ_A as above, we can first apply the theorem to obtain Σ_B with row sums n or $n+1$, and then apply the theorem to Σ_{B^t} to obtain:

THEOREM 2. *Let Σ_A be as in Theorem 1. Then there is an SSFT, conjugate to Σ_A with each row sum and each column sum either n or $n+1$.*

LEMMA. *Let $X_1, \ldots, X_k, n, M$ be positive integers satisfying*

$$\text{each } X_i \leq M, \tag{1}$$

$$\text{some } X_i < M, \tag{2}$$

$$nM \leq \sum_{i=1}^{k} X_i \leq (n+1)M. \tag{3}$$

Then there exist positive integers P_1, P_2, and a nontrivial partition of $\{1, \ldots, k\}$ into sets E_1, E_2, such that

$$M = P_1 + P_2 \tag{4}$$

and

$$nP_j \leq \sum_{i \in E_j} X_i \leq (n+1)P_j, \qquad j = 1, 2. \tag{5}$$

Received by the editors October 11, 1983.
1980 *Mathematics Subject Classification*. Primary 28D20; Secondary 54H20, 58F11.

PROOF. *Case* 1. $\sum_{i=1}^{k} X_i = (n+1)M$. By virtue of (1) and (2), we have $k > n+1$. Consider the sums

$$X_1, X_1 + X_2, \ldots, X_1 + X_2 + \cdots + X_k.$$

Since there are k of these sums, two of them must be congruent modulo $n+1$. Their difference is of the form $X_r + X_{r+1} + \cdots + X_s$, where $1 < r \leq s$; this sum is positive, divisible by $n+1$, and does not exceed

$$\sum_{i=2}^{k} X_i = (n+1)M - X_1 < (n+1)M.$$

Setting $E_1 = \{r, r+1, \ldots, s\}$, $P_i = (X_r + \cdots + X_s)/(n+1)$, $P_2 = M - P_1$ and $E_2 = \{1, 2, \ldots, k\} - E_1$, we have $0 < P_1 < M$, $P_1 + P_2 = M$, and

$$\sum_{i \in E_j} X_i = (n+1)P_j, \qquad j = 1, 2.$$

Hence, this choice of P_1, P_2, E_1, E_2 satisfies (4) and (5).

Case 2. $\sum_{i=1}^{k} X_i < (n+1)M$. Let $\alpha = (\sum X_i)/M$. We have $n \leq \alpha < n+1$. By (1) and (2), we have $k > n$. Consider the sums

$$X_1, X_1 + X_2, \ldots, X_1 + \cdots + X_k$$

modulo α. Since $k \geq n+1 > \alpha$, two of these sums must be less than unit distance apart modulo α; i.e.,

$$X_r + \cdots + X_s = t\alpha + \theta \tag{6}$$

for integers r, s, t with $1 < r \leq s$, $t < M$, and θ real with $|\theta| < 1$. Let $P_1 = t$, $E_1 = \{r, r+1, \ldots, s\}$, $P_2 = M - t$, $E_2 = \{1, \ldots, k\} - E_1$. From (6), it follows that

$$t\alpha - 1 < X_r + X_{r+1} + \cdots + X_s < t\alpha + 1,$$

or

$$P_1\alpha - 1 < \sum_{i \in E_1} X_i < P_1\alpha + 1, \tag{7}$$

and thus

$$P\ n - 1 < \sum_{i \in E_1} X\ < P\ (n+1) + 1. \tag{8}$$

Since the expressions in (8) are all integers, it follows that the strict inequalities in (8) imply

$$P_1 n \leq \sum_{i \in E_1} X_i \leq P_1(n+1). \tag{9}$$

From (7), it follows that $-P_1\alpha - 1 < -\sum_{i \in E_1} X_i < -P_1\alpha + 1$, so

$$-M\alpha + P_2\alpha - 1 < -\sum_{i \in E_1} X_i < -M\alpha + P_2\alpha + 1$$

or $P_2\alpha - 1 < M\alpha - \sum_{i \in E_1} X_i < P_2\alpha + 1$ or $P_2\alpha - 1 < \sum_{i \in E_2} X_i < P_2\alpha + 1$. Thus, as before,

$$P_2 n \leq \sum_{i \in E_2} X_i \leq P_2(n+1). \tag{10}$$

By virtue of (9) and (10) we have

$$\sum_{i\in E_j} X_i \le P_j(n+1), \qquad j = 1, 2.$$

Since E_1 and E_2 are nonempty, it follows that P_1 and P_2 are positive. Hence, this choice of P_1, P_2, E_1, E_2 satisfies the desired conditions.

References

[**ACH**] Roy L. Adler, Don Coopersmith and Martin Hassner, *Algorithms for sliding block codes*, IEEE Trans. Inform. Theory **IT-29** (1983).

[**M1**] Brian Marcus, *Factors and extensions of full shifts*, Monatsh. Math. **88** (1979), 239–247.

[**M2**] ———, *Sofic systems and encoding data* (to appear).

IBM Research Laboratory, San Jose, California 95193

Current address: Department of Mathematics, Harvard University, Cambridge, Massachusetts 02138