

MATRICES OF POLYNOMIALS, POSITIVITY, AND FINITE EQUIVALENCE OF MARKOV CHAINS

BRIAN MARCUS AND SELIM TUNCEL

Dedicated to the memory of Audrey Adler (1932–1990)

1. INTRODUCTION

We consider the ring $R = \mathbb{R}[x_1^\pm, \dots, x_k^\pm]$ of Laurent polynomials with real coefficients in the variables x_1, \dots, x_k and its positive cone $R^+ = \mathbb{R}^+[x_1^\pm, \dots, x_k^\pm]$. Letting \mathbb{R}^{++} denote the positive reals, we also have the set $R^{++} = \mathbb{R}^{++}[x_1^\pm, \dots, x_k^\pm]$ of polynomials with positive coefficients, so that $R^+ = R^{++} \cup \{0\}$. We say that $p \in R$ is *numerically positive* if p gives a positive value whenever we substitute positive numbers for x_1, \dots, x_k , that is, if $p(x_1, \dots, x_k) > 0$ for $x_1, \dots, x_k > 0$. Every element of R^{++} is numerically positive; however, there are numerically positive polynomials that are not elements of R^{++} . This distinction lies at the heart of the paper; we will discuss it in greater detail in §2.

Let B be a square matrix over R^+ . Whenever we substitute positive numbers for x_1, \dots, x_k , we get a nonnegative real-valued matrix. We assume that the matrix resulting from one, hence every, such evaluation is irreducible and, for $x_1, \dots, x_k > 0$, let $\beta(x_1, \dots, x_k) > 0$ be the maximum eigenvalue of $B(x_1, \dots, x_k)$ furnished by the Perron-Frobenius theorem. The function $\beta = \beta_B: (\mathbb{R}^{++})^k \rightarrow \mathbb{R}^{++}$ is called the β -function of B ; it satisfies the characteristic polynomial χ_B of B , a monic polynomial whose coefficients lie in R . Thus, β is algebraic over R . We assume β has degree one; that is, we consider the case $\beta \in R$. This is a significant case. For instance, for any $\beta \in R^{++}$, we obtain numerous examples of matrices with $\beta_B = \beta$ by simply requiring that B have its row (or column) sums equal to β . There are also many examples of B with $\beta_B \in R \setminus R^+$ (see [D]). Since $\beta \in R$, the entries of the adjoint $\text{Adj}(\beta I - B)$ belong to R . Moreover, by Perron-Frobenius theory [Se], the entries are numerically positive and any column r of $\text{Adj}(\beta I - B)$ satisfies $Br = \beta r$. So, as an immediate consequence of Perron-Frobenius theory, we find an eigenvector r whose entries are numerically positive polynomials. One of the main purposes of this paper is to show that B has an eigenvector whose entries lie in R^{++} .

Received by the editors August 29, 1991.

1991 *Mathematics Subject Classification*. Primary 28D20; Secondary 15A48, 15A54, 60J10, 06F25, 52B20.

The second author was partially supported by NSF Grant DMS-9004253.

In order to better describe our results, let us agree to call a polynomial $p \in R$ Handelman if there exists $q \in R^{++}$ such that the product $qp \in R^{++}$. This is one of the equivalent formulations of Handelman [H1]; in §2, we will take a more verifiable condition as the definition of a Handelman polynomial and recall from [H1] the equivalence of that definition to the criterion we have just given. Returning to B , what we actually establish is that the entries of $\text{Adj}(\beta I - B)$ are Handelman. Intimately tied with this are facts concerning submatrices of B : A square matrix A is a submatrix of B if its indexing set S is a subset of that of B and, for $I, J \in S$, the polynomials $A(I, J)$ and $B(I, J) - A(I, J)$ are Handelman or zero. Letting A be such a matrix, we show the following (Theorem 1). The nonzero entries of $\text{Adj}(\beta I - A)$ are Handelman. If A is irreducible, every entry of $\text{Adj}(\beta I - A)$ is Handelman. If $A \neq B$ then $\chi_A(\beta) = \det(\beta I - A)$ is Handelman.

In fact, we can relax the requirement that B be over R^+ and allow its nonzero entries to be Handelman. Note that we require $\beta = \beta_B$ to be a polynomial, but no restriction is imposed on β_A . Following the preliminaries of §2, Theorem 1 is proved in §§3 and 4. Among the results established in the course of this proof is the fact that β itself is Handelman.

Our interest in these positivity issues arises from problems in ergodic theory. As explained in [H2], these issues also relate to areas such as operator theory, K -theory, and probability. Theorem 1 may also be regarded as generalizing parts of Perron-Frobenius theory from $(\mathbb{R}, \mathbb{R}^+)$ to $(\mathbb{R}[x_1^\pm, \dots, x_k^\pm], \mathbb{R}^+[x_1^\pm, \dots, x_k^\pm])$. Although a construction used in the proof of Theorem 1 is closely related to [MT], the treatment given in §§2–4 is self-contained. We hope this will make §§2–4 accessible to those outside of ergodic theory.

In §5 we turn to Markov chains and their classifications. We adopt the viewpoint and notation spelled out in §1 of [MT]. Accordingly, Markov chains $(\Sigma_A, \sigma_A, \mu_A)$ are defined by matrices A over $\mathbb{Z}^+[\text{exp}]$; we abbreviate $(\Sigma_A, \sigma_A, \mu_A)$ to Σ_A . We have the basic invariants $\beta_A, \Delta_A, c_A \Delta_A$, due to [T, Kr, PS]. By choosing bases for suitable finitely generated (multiplicative) groups of exponentials, we can work over $\mathbb{R}[x_1^\pm, \dots, x_k^\pm]$ and apply Theorem 1.

Let Σ_B be a Markov chain with $\beta_B \in \mathbb{R}[\text{exp}]$. Combining Corollary 1 of §2 with the work in [T, PT], we show that a Markov chain Σ_A is finitely equivalent to Σ_B if and only if $\beta_A = \beta_B$. In particular, $\beta_A \in \mathbb{Z}^+[\text{exp}]$ if and only if Σ_A is finitely equivalent to the Bernoulli shift based on β_A . Making use of results of [PS, Ash], we also find that, in the case $\beta_B \in \mathbb{R}[\text{exp}]$, the Markov chains Σ_A, Σ_B are almost block isomorphic if and only if they have $\beta_A = \beta_B, \Delta_A = \Delta_B, c_A \Delta_A = c_B \Delta_B$, and the same period.

Thus, when the β -function is an element of $\mathbb{R}[\text{exp}]$, it completely characterizes finite equivalence. The completeness of the β -function as an invariant of finite equivalence was conjectured in general in [PT]; counterexamples to the general conjecture are given in [MT].

For the case $\beta_B \in \mathbb{R}[\text{exp}]$, Theorem 1 also gives a necessary condition for Σ_A to be a subsystem of Σ_B in the sense of [KT2]: After normalization, $\chi_A(\beta_B)$ and the entries of $\text{Adj}(\beta_B I - A)$ must be Handelman. We further show that,

up to normalization, $\chi_A(\beta_B)$ and the entries of $\text{Adj}(\beta_B I - A)$ are Handelman if and only if Σ_A is finitely equivalent to a subsystem of Σ_B .

2. HANDELMAN POLYNOMIALS

Let $p \in R = \mathbb{R}[x_1^\pm, \dots, x_k^\pm]$. Write $p = \sum_w a_w x_1^{w_1} x_2^{w_2} \dots x_k^{w_k}$, with $w = (w_1, \dots, w_k) \in \mathbb{Z}^k$, $a_w \in \mathbb{R}$, and the a_w are nonzero for only finitely many w . The *Newton polytope* of p , denoted $W(p)$, is the convex hull, taken in \mathbb{Q}^k , of the finite set $\{w \in \mathbb{Z}^k : a_w \neq 0\}$. For each face F of $W(p)$, let $p_F = \sum_{w \in F} a_w x_1^{w_1} \dots x_k^{w_k}$. We call p_F the *face* of p corresponding to F . Thus, there are finitely many faces p_F of p and, in particular, corresponding to $F = W(p)$ we have p itself. We call p a *Handelman polynomial* if all of its faces p_F are numerically positive. The significance of this condition comes from the following fundamental result of Handelman.

Handelman’s Theorem [H1]. *Let $p \in R = \mathbb{R}[x_1^\pm, \dots, x_k^\pm]$ be a numerically positive polynomial. The following are equivalent.*

- (a) p is Handelman.
- (b) There exists $q \in R$ such that the product $qp \in R^{++} = \mathbb{R}^{++}[x_1^\pm, \dots, x_k^\pm]$.
- (c) There exists $\bar{q} \in R^{++}$ such that $\bar{q}p \in R^{++}$.

In one variable, a polynomial $p \in \mathbb{R}[x^\pm]$ satisfies the conditions of this theorem if and only if it is numerically positive. (This is a result of Poincaré, a proof of which may be found in [HLP].) For every $k > 1$, however, the class of numerically positive polynomials is larger than that of Handelman polynomials. For example, in two variables, $p(x, y) = 1 - 2x + x^2 + y = (1 - x)^2 + y$ is numerically positive but not Handelman. Not only does p contain negative coefficients but so does every product qp . By contrast, $1 - x + x^2 + y$ is Handelman; in fact, $(1 + x)(1 - x + x^2 + y) = 1 + x^3 + y + xy$ has positive coefficients.

Note that sums and products of Handelman polynomials are Handelman.

Let us write H for the subset of R consisting of Handelman polynomials and the zero polynomial. Let B be an $n \times n$ matrix over H . So every entry of B is either Handelman or zero. We refer to the elements of the indexing set $\{1, \dots, n\}$ of B as the *states* of B . As in the introduction, we assume that B is irreducible and, for $x_1, \dots, x_k > 0$, define $\beta(x_1, \dots, x_k) > 0$ to be the maximum eigenvalue of $B(x_1, \dots, x_k)$. We consider the case where $\beta = \beta_B \in R$.

We say that a square matrix A is a *submatrix* of B if A is indexed by a subset of $\{1, \dots, n\}$ and, for elements I, J of this subset, both $A(I, J)$ and $B(I, J) - A(I, J)$ belong to H . We will prove the following.

Theorem 1. *Let B be an irreducible matrix over H such that its β -function $\beta = \beta_B$ is a (Laurent) polynomial. Let A be a submatrix of B .*

- (a) *If $A \neq B$, then $\chi_A(\beta) = \det(\beta I - A)$ is Handelman.*
- (b) *The nonzero entries of the adjoint $\text{Adj}(\beta I - A)$ are Handelman. If A is irreducible, every entry of $\text{Adj}(\beta I - A)$ is Handelman.*

Corollary 1. *The entries of $\text{Adj}(\beta I - B)$ are Handelman. Corresponding to β , the matrix B has left and right eigenvectors whose entries belong to $\mathbb{R}^{++}[x_1, \dots, x_k]$.*

By Perron-Frobenius theory [Se], $\text{Adj}(\beta I - B)$ is numerically positive and its rows and columns are left and right eigenvectors of B corresponding to β . Hence, Corollary 1 follows from Theorem 1 by taking $A = B$ in (b) and using Handelman's theorem.

Remark 1. The polynomial β is Handelman. This is a fact that we will establish in the course of the proof of Theorem 1. (It can also be seen by arguing as in the proof of (5.1) of [MT].) On the other hand, there are many examples in which β belongs to $R \setminus R^+$, even when B is over R^+ (see [H3, D]). DeAngelis [D] also has examples of B over R^+ such that $\beta \in R^+$, but $\text{Adj}(\beta I - B)$ is not over R^+ .

Henceforth, we focus on right eigenvectors; left eigenvectors can be treated in an entirely similar way (or just consider B^T). An eigenvector of special interest is the one whose entries are coprime, obtained by dividing by any nontrivial common factors the entries might have, and, if necessary, multiplying by -1 to ensure numerical positivity. This eigenvector is unique up to multiplication by a monomial.

Corollary 2. *Let r be the numerically positive right eigenvector of B whose entries are coprime. The entries of r are Handelman.*

Corollary 2 follows from Corollary 1 by the following lemma of Handelman [H1]. This lemma can be checked directly from the definition of a Handelman polynomial. (It also follows from Handelman's theorem.) We first realized its validity several years ago on Interstate 280 in California.

Lemma 280. *If a product pq is Handelman, then either both p and q are Handelman, or both $-p$ and $-q$ are. Hence, a polynomial is Handelman if and only if (up to changes of sign) all of its irreducible factors are.*

Before proceeding to the proof of Theorem 1, we remark that analogous results are valid if, instead of $\mathbb{R}[x_1^\pm, \dots, x_k^\pm]$, we work over $\mathbb{Z}[x_1^\pm, \dots, x_k^\pm]$, or over either of the polynomial rings $\mathbb{R}[x_1, \dots, x_k]$ and $\mathbb{Z}[x_1, \dots, x_k]$. In fact, in §§3 and 4 we will mainly work over $\mathbb{R}[x_1, \dots, x_k]$ and, for the proof of Theorem 1, we will observe that, upon multiplication by a monomial, we may assume that B is over $\mathbb{R}[x_1, \dots, x_k]$

3. FACES

We now start to work toward the proof of Theorem 1. Let $p \in \mathbb{R}[x_1^\pm, \dots, x_k^\pm]$. We put $y = x_k$ and denote by p_0 the coefficient, in $\mathbb{R}[x_1^\pm, \dots, x_{k-1}^\pm]$, of the lowest power of y appearing in p . Thus, $p_0 \in \mathbb{R}[x_1^\pm, \dots, x_{k-1}^\pm]$, $p_0 \neq 0$, and, for suitable $l \in \mathbb{Z}$ and $q \in \mathbb{R}[x_1^\pm, \dots, x_{k-1}^\pm, y]$, we have $p = p_0 y^l + q y^{l+1}$. For $\alpha = (\alpha_1, \dots, \alpha_k) \in \mathbb{R}^k$, we write $p_0(\alpha) = p_0(\alpha_1, \dots, \alpha_{k-1})$.

Lemma 1. *If $p \in \mathbb{R}[x_1^\pm, \dots, x_k^\pm]$ is numerically positive and $\alpha = (\alpha_1, \dots, \alpha_k) \in (\mathbb{R}^{++})^k$ then $p_0(\alpha) = p_0(\alpha_1, \dots, \alpha_{k-1}) \geq 0$.*

Proof. If $p_0(\alpha_1, \dots, \alpha_{k-1}) < 0$ then $p(\alpha_1, \dots, \alpha_{k-1}, y) < 0$ for small positive y . \square

Express $p \in \mathbb{R}[x_1^\pm, \dots, x_k^\pm]$ as a sum

$$p = \sum_{w \in E} a_w x_1^{w_1} \cdots x_k^{w_k}$$

with $a_w \neq 0$ for $w \in E \subset \mathbb{Z}^k$. Let $v \in \mathbb{Z}^k$. Define $\delta_v(p) = \min\{v \cdot w : w \in E\}$ and

$$p_v = \sum_{v \cdot w = \delta_v(p)} a_w x_1^{w_1} \cdots x_k^{w_k}.$$

When $v = (0, \dots, 0, 1)$, we shall write $\delta(p) = \delta_{(0, \dots, 0, 1)}(p)$. It is easy to see that $\delta(p)$ gives the lowest y -degree of the terms of p . Writing $p = p_0 y^l + q y^{l+1}$ as before, we have $\delta(p) = l$ and $p_{(0, \dots, 0, 1)} = p_0 y^l$.

Note that $F = \{w \in W(p) : v \cdot w = \delta_v(p)\}$ is a face of $W(p)$ and, by definition, $p_F = p_v$. Moreover, for any face F of $W(p)$ there exists $v \in \mathbb{Z}^k$ such that $F = \{w \in W(p) : v \cdot w = \delta_v(p)\}$. These observations lead immediately to the following proposition.

Proposition 1. *A polynomial $p \in \mathbb{R}[x_1^\pm, \dots, x_k^\pm]$ is Handelman if and only if $p_v(\alpha) > 0$ for every $v \in \mathbb{Z}^k$ and $\alpha \in (\mathbb{R}^{++})^k$.*

Now suppose $p \in \mathbb{R}[x_1, \dots, x_k]$. For $\alpha = (\alpha_1, \dots, \alpha_k) \in \mathbb{R}^k$, we shall denote by $m_\alpha(p)$ the *multiplicity* (or *order*) of p at α . When p is expressed as a polynomial \bar{p} in $x_1 - \alpha_1, \dots, x_k - \alpha_k$, the order $m_\alpha(p)$ is the lowest (total) degree of the terms of \bar{p} . In particular, if $p(\alpha) \neq 0$ then $m_\alpha(p) = 0$.

Lemma 2. *Let r be an n -vector whose entries are numerically positive polynomials in x_1, \dots, x_k . Let $v \in \mathbb{Z}^k$ and $\alpha \in (\mathbb{R}^{++})^k$. Put $m = m_\alpha(r_v) = \max\{m_\alpha(r(I)_v) : 1 \leq I \leq n\}$ and $S = S_\alpha(r_v) = \{1 \leq I \leq n : m_\alpha(r(I)_v) = m\}$. Suppose that for $I \in S$, a nonempty subset $T \subset \{1, \dots, n\}$, polynomials $U(J)$, $J \in T$, with $U(J)_v(\alpha) > 0$, and a polynomial p we have*

$$pr(I) = \sum_{J \in T} U(J)r(J).$$

Then $S \cap T \neq \emptyset$ and $p_v(\alpha) > 0$.

Corollary 3. *If a (Laurent) polynomial β is the β -function of a matrix over H , it is Handelman.*

Proof of Lemma 2. We first observe that it suffices to prove the lemma for the case $v = (0, \dots, 0, 1)$. Suppose the lemma holds in this case, and let $v \in \mathbb{Z}^k$. If v is the zero vector, then $p_v = p$ and we have $p(\alpha) > 0$ as a result of the numerical positivity of r . Otherwise, we may assume that the entries of v are coprime. Extend v to a basis of \mathbb{Z}^k to find $M \in GL(k, \mathbb{Z})$ which has v

as its k th row. Consider $P = M^{-1}$. Observe that $vP = (0, \dots, 0, 1)$. Let $\tilde{x}_i = x_1^{P(1,i)} x_2^{P(2,i)} \dots x_k^{P(k,i)}$ and $\tilde{p}(\tilde{x}_1, \dots, \tilde{x}_k) = p(x_1, \dots, x_k)$. Note that under this change of variables a monomial $\prod_{i=1}^k x_i^{w_i}$ is sent to the monomial $\prod_{i=1}^k \tilde{x}_i^{(Mw)_i}$ and, considering the respective exponents,

$$v \cdot w = (vP) \cdot (Mw) = (0, \dots, 0, 1) \cdot (Mw).$$

It follows that $p_v(x_1, \dots, x_k) = \tilde{p}_{(0, \dots, 0, 1)}(\tilde{x}_1, \dots, \tilde{x}_k)$. Similarly,

$$r(I)_v(x_1, \dots, x_k) = \widetilde{r(I)}_{(0, \dots, 0, 1)}(\tilde{x}_1, \dots, \tilde{x}_k).$$

If necessary, we multiply \tilde{p} , each entry of \tilde{r} , and each $\tilde{U}(J)(\tilde{x}_1, \dots, \tilde{x}_k) = U(J)(x_1, \dots, x_k)$ by a monomial $\tilde{q} = \tilde{x}_1^{u_1} \dots \tilde{x}_k^{u_k}$ to put them in $\mathbb{R}[\tilde{x}_1, \dots, \tilde{x}_k]$. Taking $\tilde{\alpha}_i = \prod_{j=1}^k \alpha_j^{P(j,i)}$, we find that

$$m_{\tilde{\alpha}}((\tilde{q}\tilde{r})_{(0, \dots, 0, 1)}) = m_{\alpha}(r_v) + u_k$$

and

$$S_{\tilde{\alpha}}((\tilde{q}\tilde{r})_{(0, \dots, 0, 1)}) = S = S_{\alpha}(r_v).$$

It then follows from the equation

$$(\tilde{q}\tilde{p})(\tilde{q}\tilde{r}(I)) = \sum_{J \in T} (\tilde{q}\tilde{U}(J))(\tilde{q}\tilde{r}(J))$$

that $S \cap T \neq \emptyset$ and $p_v(\alpha) = \tilde{p}_{(0, \dots, 0, 1)}(\tilde{\alpha}) > 0$.

It remains to establish the lemma for $v = (0, \dots, 0, 1)$. In this case p_v consists of the lowest y -degree terms of p ; that is, $p_v = p_0 y^l$, with $l = \delta(p)$, and our task is to prove $p_0(\alpha) = p_0(\alpha_1, \dots, \alpha_{k-1}) > 0$. Let $\delta = \min\{\delta(U(J)r(J)) : J \in T\}$ and $T' = \{J \in T : \delta(U(J)r(J)) = \delta\}$. Clearly, $T' \neq \emptyset$. The lowest y -degree terms appearing in the sum $\sum_{J \in T} U(J)r(J)$ are given by

$$s_0 = \sum_{J \in T'} U(J)_0 r(J)_0.$$

The danger is that the terms of the last sum could cancel out and render s_0 the zero polynomial. We rule out this possibility. Let $J' \in T'$. By assumption, $U(J)_0(\alpha) > 0$ for every $J \in T$. Hence there exists a neighbourhood of α in $(\mathbb{R}^{++})^k$ on which each $U(J)_0$ is strictly positive. Using this and the fact that $r(J')_0$ is not the zero polynomial, we can find $\alpha' \in (\mathbb{R}^{++})^k$ such that $U(J)_0(\alpha') > 0$ for every $J \in T$ and $r(J')_0(\alpha') \neq 0$. Since each $r(J)_0(\alpha') \geq 0$ by Lemma 1, we then have

$$s_0(\alpha') \geq U(J)_0(\alpha') r(J')_0(\alpha') > 0.$$

It follows that s_0 is nontrivial and

$$\left(\sum_{J \in T} U(J)r(J) \right)_0 = s_0 = \sum_{J \in T'} U(J)_0 r(J)_0.$$

This fact and the equation $pr(I) = \sum_{J \in T} U(J)r(J)$ imply

$$(*) \quad p_0 r(I)_0 = \sum_{J \in T'} U(J)_0 r(J)_0.$$

Consider the order of $(*)$ at α . By assumption, $U(J)_0(\alpha) > 0$, so that $m_\alpha(U(J)_0) = 0$. With

$$m = m_\alpha(r_{(0, \dots, 0, 1)}) = \max\{m_\alpha(r(J)_0) : 1 \leq J \leq n\},$$

we also have $m_\alpha(r(J)_0) \leq m$ for every J . Therefore, the order at α of the RHS of $(*)$, and hence of $p_0 r(I)_0$, is at most m . In addition, this order equals m if and only if $m_\alpha(r(J)_0) = m$ for every $J \in T'$. Since $m_\alpha(r(I)_0) = m$, we conclude that $m_\alpha(r(J)_0) = m$ for every $J \in T'$ and that $m_\alpha(p_0) = 0$. Hence, $T' \subset S \cap T$. Furthermore, $p_0(\alpha) \neq 0$ and, using $(*)$ and the numerical positivity of r , we see $p_0(\alpha) > 0$. \square

4. PROOF OF THEOREM 1

Lemma 3. *Let $\beta \in \mathbb{R}[x_1, \dots, x_k]$ and let A be a $d \times d$ matrix whose nonzero entries are Handelman polynomials in x_1, \dots, x_k . Let $v \in \mathbb{Z}^k$ and $\alpha \in (\mathbb{R}^{++})^k$. Suppose $\chi_{A'}(\beta)_v(\alpha) > 0$ for every submatrix A' of A whose size is strictly less than d . Then every nonzero entry $\text{Adj}(\beta I - A)(I, J)$ of $\text{Adj}(\beta I - A)$ has*

$$\text{Adj}(\beta I - A)(I, J)_v(\alpha) > 0.$$

If A is irreducible, $\text{Adj}(\beta I - A)$ has no zero entries and the above inequality is valid for every entry.

Proof. For a $d \times d$ matrix C and a set $T \subset \{1, \dots, d\}$ of states of C , we let $C\langle T \rangle$ denote the matrix obtained from C by deleting the (rows and columns of) the states in T . We write $\text{Adj} = \text{Adj}(\beta I - A)$. By definition, a diagonal entry

$$\text{Adj}(J, J) = \det(\beta I - A\langle J \rangle) = \chi_{A\langle J \rangle}(\beta)$$

and, since $A\langle J \rangle$ has size $d - 1$, we have $\chi_{A\langle J \rangle}(\beta)_v(\alpha) > 0$. To handle off-diagonal entries, let C be a $d \times d$ matrix and fix *distinct* states $I, J \in \{1, \dots, d\}$. Let X be the set of all permutations π of $\{1, \dots, d\}$ with $\pi(I) = J$. Considering the adjoint $\text{Adj}(C)$ of C , recall that

$$(\dagger) \quad C(I, J) \text{Adj}(C)(J, I) = \sum_{\pi \in X} \text{sgn}(\pi) \left(\prod_{K=1}^d C(K, \pi(K)) \right).$$

(This is exactly why we have an I th row expansion

$$\det(C) = \sum_{J=1}^d C(I, J) \text{Adj}(C)(J, I)$$

for the determinant of C .) Let \mathcal{C} be the set of all *cyclic* permutations γ of $\{1, \dots, d\}$ with $\gamma(I) = J$. Denote the orbit of $(I$ under) $\gamma \in \mathcal{C}$ by $\text{orb}(\gamma)$. Every $\pi \in X$ can be written uniquely as the product of an element

$\gamma \in \mathcal{E}$ with a permutation of ρ of $\{1, \dots, d\} \setminus \text{orb}(\gamma)$. In fact, this gives a bijection between X and pairs (γ, ρ) where $\gamma \in \mathcal{E}$ and ρ is a permutation of $\{1, \dots, d\} \setminus \text{orb}(\gamma)$. We use this bijection to partition the sum in (\dagger) as

$$C(I, J) \text{Adj}(C)(J, I) = \sum_{\gamma \in \mathcal{E}} \text{sgn}(\gamma) \left(\prod_{K \in \text{orb}(\gamma)} C(K, \gamma(K)) \right) \det(C \langle \text{orb}(\gamma) \rangle).$$

Hence,

$$\text{Adj}(C)(J, I) = \sum_{\gamma \in \mathcal{E}} (-1)^{|\text{orb}(\gamma)|+1} \left(\prod_{\substack{K \in \text{orb}(\gamma) \\ K \neq I}} C(K, \gamma(K)) \right) \det(C \langle \text{orb}(\gamma) \rangle).$$

Taking $C = \beta I - A$,

$$\begin{aligned} \text{Adj}(J, I) &= \sum_{\gamma \in \mathcal{E}} (-1)^{|\text{orb}(\gamma)|+1} \left(\prod_{\substack{K \in \text{orb}(\gamma) \\ K \neq I}} (-A(K, \gamma(K))) \right) \det(\beta I - A \langle \text{orb}(\gamma) \rangle) \\ &= \sum_{\gamma \in \mathcal{E}} \left(\prod_{\substack{K \in \text{orb}(\gamma) \\ K \neq I}} A(K, \gamma(K)) \right) \det(\beta I - A \langle \text{orb}(\gamma) \rangle). \end{aligned}$$

Since $A \langle \text{orb}(\gamma) \rangle$ has size less than d , we know $\det(\beta I - A \langle \text{orb}(\gamma) \rangle)_v(\alpha) > 0$. In addition, if an entry $A(K, \gamma(K))$ of A is nonzero, it is Handelman, so that $A(K, \gamma(K))_v(\alpha) > 0$. It follows that either $\text{Adj}(J, I) = 0$ or $\text{Adj}(J, I)_v(\alpha) > 0$.

Moreover, $\text{Adj}(J, I) = 0$ if and only if we have

$$\prod_{\substack{K \in \text{orb}(\gamma) \\ K \neq I}} A(K, \gamma(K)) = 0$$

for every $\gamma \in \mathcal{E}$. If A is irreducible, we can find states J_1, \dots, J_h such that I, J, J_1, \dots, J_h are distinct and the entries $A(J, J_1), A(J_1, J_2), \dots, A(J_{h-1}, J_h), A(J_h, I)$ are nonzero, giving us a cyclic permutation $\gamma = (I, J, J_1, \dots, J_h)$ with

$$\prod_{\substack{K \in \text{orb}(\gamma) \\ K \neq I}} A(K, \gamma(K)) \neq 0.$$

Consequently, in the case of irreducible A , no entry of $\text{Adj} = \text{Adj}(\beta I - A)$ is zero. \square

Proof of Theorem 1. It is easy to see that, upon multiplication by a suitable monomial, we may assume that B is an $n \times n$ matrix whose nonzero entries are Handelman polynomials in $\mathbb{R}[x_1, \dots, x_k]$. Let $v \in \mathbb{Z}^k$ and $\alpha \in (\mathbb{R}^{++})^k$. In view of Proposition 1 and Lemma 3, to prove the theorem, it suffices to show that $\chi_A(\beta)_v(\alpha) > 0$ whenever A is a proper submatrix of B . Let r denote a numerically positive right eigenvector of B with polynomial entries. As in Lemma 2, we have

$$m = \max\{m_\alpha(r(I)_v) : 1 \leq I \leq n\}$$

and

$$S = \{1 \leq I \leq n : m_\alpha(r(I)_v) = m\}.$$

Let C be the submatrix of B corresponding to S ; that is, C is the submatrix of B obtained by deleting (the rows and columns of) the states outside of S . We will show $S = \{1, \dots, n\}$ and conclude $C = B$. First we establish the desired result for submatrices of C .

Lemma 4. *Let A be a submatrix of C . Unless $A = B$, we have $\chi_A(\beta)_v(\alpha) > 0$.*

Proof. We will use induction on the size d of A . First we consider the case $d = 1$. In this case A is a polynomial $q = A = \beta_A$ and our task is to prove $(\beta - q)_v(\alpha) > 0$. Order the states of B so that A corresponds to the $(1, 1)$ entry of B . In particular, $1 \in S$. Put $U(1) = B(1, 1) - q$ and $U(J) = B(1, J)$ for $2 \leq J \leq n$. Then, for every $1 \leq J \leq n$, the polynomial $U(J)$ is Handelman, provided $U(J) \neq 0$. In particular, setting $T = \{1 \leq J \leq n : U(J) \neq 0\}$, we have $U(J)_v(\alpha) > 0$ for every $J \in T$. Note that $T \neq \emptyset$ as long as $A \neq B$. The equation $Br = \beta r$ implies

$$(\beta - q)r(1) = \sum_{J \in T} U(J)r(J).$$

Since $1 \in S$, we have $(\beta - q)_v(\alpha) > 0$ by Lemma 2. This proves Lemma 4 in the case $d = 1$.

Now suppose Lemma 4 has been proved when the size of the submatrix is less than d , and let A be $d \times d$. Assume $A \neq B$. Order the states of B so that A corresponds to $\{1, \dots, d\}$. In particular, $\{1, \dots, d\} \subset S$. For $1 \leq I \leq d$ put

$$U(I, J) = \begin{cases} B(I, J) - A(I, J) & \text{if } 1 \leq J \leq d, \\ B(I, J) & \text{if } d + 1 \leq J \leq n. \end{cases}$$

Then the nonzero entries of the $d \times n$ matrix U are Handelman. Moreover, $U \neq 0$, because $A \neq B$. The first d rows of equation $Br = \beta r$ give

$$(\beta I - A)r' = Ur,$$

where r' is the vector consisting of the first d entries of r . Hence

$$\det(\beta I - A)r' = \text{Adj}(\beta I - A)Ur.$$

Put $V = \text{Adj}(\beta I - A)U$. Observe that, by our inductive assumption and Lemma 3, we have $\text{Adj}(\beta I - A)(I, J)_v(\alpha) > 0$ whenever the entry is nonzero. Similarly, whenever an entry $U(I, J)$ is nonzero, it is Handelman and so $U(I, J)_v(\alpha) > 0$. It follows that every nonzero entry $V(I, J)$ of V has $V(I, J)_v(\alpha) > 0$. Moreover, no row of V is trivial, because $\chi_A(\beta)r' = Vr$ and $A \neq B$. Therefore, by Lemma 2, we have $\chi_A(\beta)_v(\alpha) > 0$. \square

We now complete the proof of Theorem 1. Suppose, for a contradiction, that S is a proper subset of $\{1, \dots, n\}$. Order the states of B so that $S = \{1, 2, \dots, |S|\}$ and accordingly write B and r in block form

$$B = \begin{bmatrix} C & U \\ W & Z \end{bmatrix}, \quad r = \begin{bmatrix} r' \\ r'' \end{bmatrix}.$$

We then have $Cr' + Ur'' = \beta r'$ so that

$$\det(\beta I - C)r' = \text{Adj}(\beta I - C)Ur''.$$

It was shown in the proof of Lemma 4 that no row of $V = \text{Adj}(\beta I - C)U$ is trivial and every nonzero entry $V(I, J)$ has $V(I, J)_v(\alpha) > 0$. (This is part of the inductive step for $d = |S|$, $A = C$.) For $I \in S$,

$$\chi_C(\beta)r(I) = \sum_{J \in T} V(I, J)r(J),$$

where $T = \{|S| + 1 \leq J \leq n : V(I, J) \neq 0\}$. Note that $T \neq \emptyset$ and $T \cap S = \emptyset$. On the other hand, according to Lemma 2, we have $S \cap T \neq \emptyset$! This contradiction shows that $S = \{1, \dots, n\}$. So $C = B$ and, by Lemma 4, $\chi_A(\beta)_v(\alpha) > 0$ for every proper submatrix A of B . \square

The proof of Theorem 1 contains the fact that $S = \{1, \dots, n\}$, which means that, for any $v \in \mathbb{Z}^k$ and $\alpha \in (\mathbb{R}^{++})^k$, the multiplicity $m_\alpha(r(I)_v)$ is independent of $I \in \{1, \dots, n\}$. This gives the following strengthening of Lemma 2.

Theorem 2. *Let B be an irreducible $n \times n$ matrix over H such that $\beta = \beta_B$ is a (Laurent) polynomial. Let r be a numerically positive right eigenvector with polynomial entries. Suppose that for a state I , a nonempty subset $T \subset \{1, \dots, n\}$, polynomials $U(J)$, $J \in T$, and a polynomial p we have*

$$pr(I) = \sum_{J \in T} U(J)r(J).$$

- (a) *If $v \in \mathbb{Z}^k$, $\alpha \in (\mathbb{R}^{++})^k$, and $U(J)_v(\alpha) > 0$ for every $J \in T$ then $p_v(\alpha) > 0$.*
- (b) *If the polynomials $U(J)$, $J \in T$, are Handelman, so is p .*

5. FINITE EQUIVALENCE OF MARKOV CHAINS

We now turn to Markov chains. We adopt the viewpoint, terminology, and notation spelled out in §1 of [MT]. Accordingly, an irreducible matrix A over $\mathbb{Z}^+[\text{exp}]$ defines a Markov chain $(\Sigma_A, \sigma_A, \mu_A)$, which we abbreviate to Σ_A . Without loss of generality, we assume the defining matrix A to be in stochastic form. As in §1 of [MT], the associated basic invariants are β_A , Δ_A , $c_A \Delta_A$; these were first introduced in [T, Kr, PS], respectively.

At any point our discussion will involve only a finite collection $\{a^t\}$ of exponential functions. By considering the multiplicative subgroup of \mathbb{R}^{++} generated by the corresponding set $\{a\}$ of bases (or any larger finitely generated subgroup of \mathbb{R}^{++}) and choosing a basis b_1, \dots, b_k for this group, we can identify the ordered ring $(\mathbb{R}[b_1^t, \dots, b_k^t], \mathbb{R}^+[b_1^t, \dots, b_k^t])$ with the ordered Laurent polynomial ring $(\mathbb{R}[x_1^\pm, \dots, x_k^\pm], \mathbb{R}[x_1^\pm, \dots, x_k^\pm])$. Under this process, $(\mathbb{Z}[b_1^t, \dots, b_k^t], \mathbb{Z}^+[b_1^t, \dots, b_k^t])$ is identified with $(\mathbb{Z}[x_1^\pm, \dots, x_k^\pm], \mathbb{Z}^+[x_1^\pm, \dots, x_k^\pm])$. (See [PT, Lemma 23].) Making implicit use of such identifications, we will apply Theorem 1 and its corollaries to Markov chains. As an explicit instance of the use of these identifications, we explain what it means

for $\alpha \in \mathbb{R}[\text{exp}]$ to be Handelman. Write $\alpha = \sum_{i=1}^l n_i a_i^t$, with $n_i \in \mathbb{Z}$ and $a_i \in \mathbb{R}^{++}$, and consider the multiplicative subgroup of G of \mathbb{R}^{++} generated by $\{a_i : 1 \leq i \leq l\}$. Let b_1, \dots, b_k be a basis for G . Writing $x_1 = b_1^t, \dots, x_k = b_k^t$ and expressing each a_i as a product of the generators b_1, \dots, b_k , we obtain from α an element $\bar{\alpha}$ of $\mathbb{R}[x_1^\pm, \dots, x_k^\pm]$. Define α to be *Handelman* if and only if the corresponding polynomial $\bar{\alpha} \in \mathbb{R}[x_1^\pm, \dots, x_k^\pm]$ is Handelman. This definition is independent of the choice of generators for G . Moreover, it is unaffected by the choice, instead of G , of a finitely generated subgroup of \mathbb{R}^{++} which contains G . This is equivalent to the statement that if x_1, \dots, x_k are multiplicatively independent monomials in $\mathbb{R}[y_1^\pm, \dots, y_m^\pm]$ then a polynomial $p \in \mathbb{R}[x_1^\pm, \dots, x_k^\pm] \subset \mathbb{R}[y_1^\pm, \dots, y_m^\pm]$ is Handelman as an element of $\mathbb{R}[x_1^\pm, \dots, x_k^\pm]$ if and only if it is Handelman as an element of $\mathbb{R}[y_1^\pm, \dots, y_m^\pm]$.

Recall that two Markov chains Σ_A, Σ_B are *finitely equivalent* if there exists a Markov chain Σ_C and bounded-to-one block homomorphisms $\phi: \Sigma_C \rightarrow \Sigma_A, \psi: \Sigma_C \rightarrow \Sigma_B$. An *almost block isomorphism* between Σ_A and Σ_B is a finite equivalence in which the maps ϕ and ψ are both one-to-one a.e. Throughout our discussion Σ_B will be a Markov chain such that $\beta_B \in \mathbb{R}[\text{exp}]$.

Remark 2. We do not assume $\beta_B \in \mathbb{Z}[\text{exp}]$, for there exist irreducible matrices B over $\mathbb{Z}^+[\text{exp}]$ with $\beta_B \in \mathbb{R}[\text{exp}] \setminus \mathbb{Z}[\text{exp}]$. For an instance of this, let $a \in \mathbb{R}^{++}, a \neq 1$, and write $x = a^t$. Let $\rho = (1 + \sqrt{5})/2$ be the golden mean. It is easily checked that the matrix

$$B = \begin{bmatrix} 2 + x & x^2 \\ 1 & 2 \end{bmatrix}$$

has $\beta_B = 2 + \rho x$. Thus, there are polynomial beta functions β_B whose coefficients are real but not integral. As we have already remarked, the coefficients of β_B may contain negative numbers.

Theorem 3. *Let Σ_B be a Markov chain with $\beta_B \in \mathbb{R}[\text{exp}]$. A Markov chain Σ_A is finitely equivalent to Σ_B if and only if $\beta_A = \beta_B$.*

Proof. The β -function was established as an invariant of finite equivalent in [T]. For the converse, let $\beta = \beta_A = \beta_B$ and use Corollary 1 to find a row vector l and a column vector r , both over $\mathbb{R}^{++}[\text{exp}]$, such that $Ar = \beta r$ and $lB = \beta l$. Then the matrix $F' = rl$ is over $\mathbb{R}^{++}[\text{exp}]$ and, for every positive integer h , we have $A(hF') = (hF')B$. Moreover, for every $\varepsilon > 0$, we can find a positive integer h such that every nonzero coefficient of every entry of hF' is within ε of a positive integer. For small enough $\varepsilon > 0$, we obtain a matrix F over $\mathbb{Z}^{++}[\text{exp}]$ such that $AF = FB$. This, by (22) of [PT], implies the finite equivalence of Σ_A and Σ_B . \square

Bernoulli shifts correspond to elements of $\mathbb{Z}^{++}[\text{exp}]$. Their finite equivalence classes are of special interest.

Corollary 4. *A Markov chain Σ_A has $\beta_A \in \mathbb{Z}^{++}[\text{exp}]$ if and only if it is finitely equivalent to the Bernoulli shift based on β_A .*

Corollary 5. *Let Σ_B be a Markov chain with $\beta_B \in \mathbb{R}[\text{exp}]$. A Markov chain Σ_A is almost block isomorphic to Σ_B if and only if they have $\beta_A = \beta_B$, $\Delta_A = \Delta_B$, $c_A \Delta_A = c_B \Delta_B$, and the same period.*

Proof. The pair $(\Delta, c\Delta)$ was established as an invariant of almost block isomorphism in [PS]. As shown in (5.9) of [MT], the converse follows from Corollary 1 by (6.1) of [Ash]. \square

In the zero or one variable case (that is, when the defining matrices are over \mathbb{Z} or $\mathbb{Z}[x^\pm]$), Theorem 3 and Corollary 5 hold without the restriction that $\beta_A = \beta_B$ be polynomial (see [P1, AM, P2, MT]). Contrary to the conjecture of [PT], however, in the higher variable cases there are examples Σ_A, Σ_B which are not finitely equivalent even though $\beta_A = \beta_B$ (see [MT]).

As in §1 of [MT], for a cycle γ in the graph $G(A)$ of A we denote by $l(\gamma)$, $\text{wt}_A(\gamma)$ the length and weight, respectively, of γ . The *weight-per-symbol* of γ is defined to be $\text{wps}_A(\gamma) = \frac{1}{l(\gamma)} \log(\text{wt}_A(\gamma))$. We will encounter in the sequel maps $\pi: \Sigma_A \rightarrow \Sigma_B$ which are continuous, shift-commuting, bounded-to-one, and weight-preserving in the following sense. There exists $c \in \mathbb{R}^{++}$ such that, for every cycle γ in the graph $G(A)$ of A , we have $c \text{wps}_B(\pi(\gamma)) = \text{wps}_A(\gamma)$. When π is injective, π is called an *embedding*, and the Markov chain Σ_A is a subsystem of Σ_B ; we denote this by $\pi: \Sigma_A \hookrightarrow \Sigma_B$. (See Remark 3.)

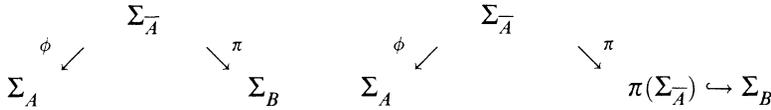
Suppose now that the map $\pi: \Sigma_A \rightarrow \Sigma_B$ is 1-block, so that it is given by a map $\pi: E(A) \rightarrow E(B)$ from the edges $E(A)$ of $G(A)$ to $E(B)$. Such a map is called *right resolving* if, for $e_1 \in E(A)$ and $f_1, f_2 \in E(B)$ with $\tau(f_1) = i(f_2)$ and $\pi(e_1) = f_1$, there exists at most one $e_2 \in E(A)$ with $\tau(e_2) = i(e_2)$ and $\pi(e_2) = f_2$. Left resolving maps are similarly defined. Left and right resolving maps are necessarily bounded-to-one [AM].

We remark that, as a result of the construction in (22) of [PT], in Theorem 3 and its corollaries the finite equivalence may be chosen so that one of the maps is left resolving and the other is right resolving. We will next observe that the construction in (22) of [PT] works in a more general context.

We say that Σ_A is *finitely equivalent to a subsystem* of Σ_B if there exists a Markov chain $\Sigma_{\bar{A}}$, a bounded-to-one block homomorphism $\Sigma_{\bar{A}} \rightarrow \Sigma_A$, and a continuous, shift-commuting, bounded-to-one map $\Sigma_{\bar{A}} \rightarrow \Sigma_B$ which is weight-preserving in the above sense. We pause to elaborate on this definition.

Remark 3. [KT1] introduced a class of measures for sofic systems. Elements of this class were called semigroup measures in [KT1, KT2]; here, we refer to them as *finitary* measures. In the case of a shift of finite type, the class of finitary measures supported by the shift of finite type contains all Markov measures supported by it. Suppose S is a sofic system contained in the shift of finite type Σ_B underlying the Markov chain (Σ_B, μ_B) . According to [KT2], μ_B induces a finitary measure μ on S and the pair (S, μ) is called a *subsystem* of (Σ_B, μ_B) . Moreover, by [KT1], there exists a Markov chain $(\Sigma_{\bar{A}}, \mu_{\bar{A}})$ and a continuous, shift-commuting, bounded-to-one surjection $\bar{\pi}: \Sigma_{\bar{A}} \rightarrow S$ with $\mu_{\bar{A}} \circ \bar{\pi}^{-1} = \mu$. Composing $\bar{\pi}$ with the inclusion of S in Σ_B , we obtain a map $\pi: \Sigma_{\bar{A}} \rightarrow \Sigma_B$ and [KT2] shows that there exists $c \in \mathbb{R}^{++}$ such that $c \text{wps}_B(\pi(\gamma)) = \text{wps}_{\bar{A}}(\gamma)$ for every cycle γ of $G(\bar{A})$. Conversely, suppose $\pi: \Sigma_{\bar{A}} \rightarrow \Sigma_B$ is a continuous, shift-

commuting, bounded-to-one, weight-preserving map. Then $\pi(\Sigma_A)$ is a sofic system and, considering π as a map onto $\pi(\Sigma_A)$ and letting μ be the measure induced on $\pi(\Sigma_A)$ by μ_B , we have a measure-preserving map of $(\Sigma_{\bar{A}}, \mu_{\bar{A}})$ onto the subsystem $(\pi(\Sigma_{\bar{A}}), \mu)$. (In particular, if π is injective then $(\Sigma_{\bar{A}}, \mu_{\bar{A}})$ is block isomorphic to $(\pi(\Sigma_{\bar{A}}), \mu)$, so $(\Sigma_{\bar{A}}, \mu_{\bar{A}})$ may be regarded as a subsystem of (Σ_B, μ_B) .) It follows from these observations that the statement that Σ_A is finitely equivalent to a subsystem of Σ_B may be visualized in two ways:



In the first of these pictures the bounded-to-one, continuous, shift-commuting, weight-preserving map π need not be onto. In the second picture π is a measure-preserving map onto the subsystem $(\pi(\Sigma_{\bar{A}}), \mu)$ of (Σ_B, μ_B) . In each picture, ϕ is a bounded-to-one block homomorphism.

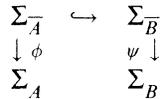
The proof of the following proposition is just like the construction in (22) of [PT] (see the description of this construction given at the beginning of the proof of (5.9) of [MT]).

Proposition 2. *Let A, B be irreducible matrices over $\mathbb{Z}^+[\text{exp}]$, and let F be a nontrivial matrix over $\mathbb{Z}^+[\text{exp}]$ such that $AF \leq FB$ (i.e., such that the matrix $FB - AF$ is over $\mathbb{Z}^+[\text{exp}]$). Then Σ_A is finitely equivalent to a subsystem of Σ_B by a finite equivalence in which the map onto Σ_A is left resolving and the other map is right resolving.*

The following gives dynamical characterizations of the condition that $\chi_A(\beta)$ and the entries of $\text{Adj}(\beta I - A)$ be Handelman. Note, in particular, the equivalence of (i), (iv), and (vii).

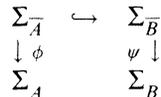
Theorem 4. *Let Σ_A, Σ_B be Markov chains, with $\beta = \beta_B \in \mathbb{R}[\text{exp}]$. The following are equivalent.*

- (i) *There exists $c \in \mathbb{R}^{++}$ and a nontrivial matrix F over $\mathbb{Z}^+[\text{exp}]$ such that $AF \leq c^t FB$ (i.e., such that the matrix $c^t FB - AF$ is over $\mathbb{Z}^+[\text{exp}]$).*
- (ii) *Σ_A is finitely equivalent to a subsystem of Σ_B by a finite equivalence in which the map onto Σ_A is left resolving and the other is right resolving. (See Remark 3.)*
- (iii) *We have a diagram*



in which $\Sigma_{\bar{A}}, \Sigma_{\bar{B}}$ are Markov chains, $\Sigma_{\bar{A}}$ is a subsystem of $\Sigma_{\bar{B}}$, ϕ and ψ are block homomorphisms, ϕ is left resolving, and ψ is right resolving.

- (iv) *Σ_A is finitely equivalent to a subsystem of Σ_B . (See Remark 3.)*
- (v) *We have a diagram*



in which $\Sigma_{\bar{A}}$, $\Sigma_{\bar{B}}$ are Markov chains, $\Sigma_{\bar{A}}$ is a subsystem of $\Sigma_{\bar{B}}$, ϕ and ψ are bounded-to-one block homomorphisms.

(vi) We have a diagram

$$\begin{array}{ccc} \Sigma_{\bar{A}} & \hookrightarrow & \Sigma_{\bar{B}} \sim \Sigma_B \\ \phi \downarrow & & \\ \Sigma_A & & \end{array}$$

in which $\Sigma_{\bar{A}}$, $\Sigma_{\bar{B}}$ are Markov chains, $\Sigma_{\bar{A}}$ is a subsystem of $\Sigma_{\bar{B}}$, ϕ is a bounded-to-one block homomorphism, and $\Sigma_{\bar{B}}$ is finitely equivalent to Σ_B .

(vii) There exists $c \in \mathbb{R}^{++}$ such that $\chi_A(c^t \beta)$ and the entries of $\text{Adj}(c^t \beta I - A)$ are Handelman (except that $\chi_A(c^t \beta) = 0$ in the case $\beta_A = c^t \beta$).

Remark 4. (a) In the exceptional case $\beta_A = c^t \beta$, we necessarily have $c = 1$ and $\beta_A = \beta$ as a result of our standing assumption that A and B are stochastic forms so that, considering Theorem 3, all of the above (equivalent) conditions holds.

(b) The symmetry of (vii) of Theorem 4 implies that there exists a nontrivial $\mathbb{Z}^+[\text{exp}]$ matrix F with $AF \leq c^t FB$ if and only if there is such a matrix \bar{F} with $\bar{F}A \leq c^t B\bar{F}$, so that the roles of left and right resolving maps may be interchanged in (ii) and (iii).

Proof of Theorem 4. By Proposition 2, (i) implies (ii). It is clear that (iii) implies (iv) and (v) implies (vi).

(ii) \Rightarrow (iii). We start with the first picture in Remark 3, ϕ being left resolving and π right resolving. According to Theorem 5 of [AMPT], there exists an irreducible graph $G \supset G(\bar{A})$ and a map (labeling) $\psi: E(G) \rightarrow E(B)$ from the edges of G to those of $G(B)$ which extends $\pi: E(\bar{A}) \rightarrow E(B)$ and gives a right resolving map $\psi: \Sigma \rightarrow \Sigma_B$ of the shift of finite type Σ defined by G onto Σ_B . Associate to each $e \in E(G)$ the exponential corresponding to $\psi(e) \in E(B)$ to obtain a matrix \bar{B} over $\mathbb{Z}^+[\text{exp}]$. Then the map $\psi: \Sigma_{\bar{B}} \rightarrow \Sigma_B$ satisfies $\mu_{\bar{B}} \circ \psi^{-1} = \mu_B$ (see [T, K, MT]), so it is a right resolving block homomorphism. The natural injection $\Sigma_{\bar{A}} \hookrightarrow \Sigma_{\bar{B}}$ arising from the inclusion $G(\bar{A}) \subset G = G(\bar{B})$ is an embedding of $\Sigma_{\bar{A}}$ into $\Sigma_{\bar{B}}$.

(iv) \Rightarrow (v). This is established by arguing just like in the proof of (ii) \Rightarrow (iii), using Theorem 4 of [AMPT] instead of Theorem 5 of that paper.

(vi) \Rightarrow (vii). Considering the embedding $\pi: \Sigma_{\bar{A}} \hookrightarrow \Sigma_{\bar{B}}$, let $c \in \mathbb{R}^{++}$ be such that $\text{wps}_{\bar{A}}(\gamma) = c \text{wps}_{\bar{B}}(\pi(\gamma))$ for all cycles γ of A . With the aim of applying Lemma 3, we let A' be an irreducible submatrix of A and check that $\chi_{A'}(c^t \beta)$ is Handelman. Find an irreducible matrix \bar{A}' over $\mathbb{Z}^+[\text{exp}]$ such that $\Sigma_{\bar{A}'} \subset \Sigma_{\bar{A}}$ and $\phi(\Sigma_{\bar{A}'}) = \Sigma_{A'}$, and we have $\text{wps}_{\bar{A}'}(\gamma) = \text{wps}_{A'}(\phi(\gamma))$ for all cycles γ of $G(\bar{A}')$. Note that, because $\Sigma_{\bar{B}}$ and Σ_B are finitely equivalent, $\beta_{\bar{B}} = \beta = \beta_B$. Passing to a higher block system of $\Sigma_{\bar{B}}$ if necessary, we can find a submatrix \bar{B}' of \bar{B} such that $\pi(\Sigma_{\bar{A}'}) = \Sigma_{\bar{B}'}$. It follows from Theorem 1 that $\chi_{\bar{B}'}(\beta)$ is Handelman, unless $A' = A$ and $\beta_A = c^t \beta$. Using Corollary 3, we find that $\chi_{c^t \bar{A}'}(\beta)$ and $\chi_{\bar{A}'}(c^t \beta)$ are Handelman unless $A' = A$ and $\beta_A = c^t \beta$.

Also, by Corollary 2 of [K], $\chi_{A'}(c^t\beta)$ divides $\chi_{A'}(c^t\beta)$, modulo powers of β . Using Lemma 280 and Corollary 3, we conclude that $\chi_{A'}(c^t\beta)$ is Handelman for every irreducible submatrix A' of A , unless $A' = A$ and $\beta_A = c^t\beta$. Taking $A' = A$, we see that $\chi_A(c^t\beta)$ is Handelman, except that $\chi_A(c^t\beta) = 0$ when $\beta_A = c^t\beta$. Moreover, if A'' is any submatrix of A then $\chi_{A''}(c^t\beta)$ is the product of a power of $c^t\beta$ with the $\chi_{A'}(c^t\beta)$ for the irreducible submatrices A' of A corresponding to the irreducible components of A'' . Using this fact and Lemma 3, we see that $\text{Adj}(c^t\beta - A)$ has Handelman entries.

(vii) \Rightarrow (i). Let l be a vector over $\mathbb{R}^{++}[\text{exp}]$ such that $lB = \beta l$, and let r be a column of $\text{Adj}(c^t\beta I - A)$. Remark 4 dealt with the case $\beta_A = c^t\beta$. Suppose $\beta_A \neq c^t\beta$. Since

$$(c^t\beta I - A) \text{Adj}(c^t\beta I - A) = \chi_A(c^t\beta)I$$

and $\chi_A(c^t\beta)$ is Handelman, we have $Ar = c^t\beta r - u$ for a nontrivial vector u whose nonzero entries are Handelman. The matrix $F' = rl$ is nontrivial and $c^tF'B - AF' = ul$. Since the nonzero entries of F' and ul are Handelman, we can apply Handelman's theorem to find $\alpha \in \mathbb{R}^+[\text{exp}]$ such that the nonzero entries of $F'' = \alpha F'$ and $D'' = \alpha ul$ are in $\mathbb{R}^{++}[\text{exp}]$. Then, for every positive integer h , the difference $c^t(hF'')B - A(hF'') = hD''$ is over $\mathbb{R}^+[\text{exp}]$. Furthermore, for every $\varepsilon > 0$, we can find a positive integer h such that every nonzero coefficient of every entry of hF'' and hD'' is within ε of a positive integer. For small enough $\varepsilon > 0$, we obtain nontrivial matrices F, D over $\mathbb{Z}^+[\text{exp}]$ with $c^tFB - AF = D$. \square

Corollary 6. *If $\beta_B \in \mathbb{R}[\text{exp}]$ and Σ_A is a proper subsystem of Σ_B , there exists $c \in \mathbb{R}^{++}$ such that $\chi_A(c^t\beta_B)$ and the entries of $\text{Adj}(c^t\beta_B I - A)$ are Handelman.*

Corollary 7. *If both $\beta_A, \beta_B \in \mathbb{R}[\text{exp}]$ and $\beta_A \neq \beta_B$, then (i)-(vii) of Theorem 4 are also equivalent to*

(viii) *There exists $c \in \mathbb{R}^{++}$ such that $c^t\beta_B - \beta_A$ is Handelman.*

Proof. The minimum polynomial of $\beta_A \in \mathbb{R}[\text{exp}]$ is $s(z) = z - \beta_A$. As s divides the characteristic polynomial χ_A , Lemma 280 shows that (vii) implies $s(c^t\beta_B) = c^t\beta_B - \beta_A$ is Handelman. Conversely, suppose $\beta_A, \beta_B \in \mathbb{R}[\text{exp}]$ and $c^t\beta_B - \beta_A$ is Handelman. Find $q \in \mathbb{R}^{++}[\text{exp}]$ such that $q(c^t\beta_B - \beta_A) \in \mathbb{R}^{++}[\text{exp}]$. Using Corollary 1, pick vectors l, r over $\mathbb{R}^{++}[\text{exp}]$ such that $Ar = \beta_A r$ and $lB = \beta_B l$. Let $F' = qrl$. For every positive integer h , the difference $c^t(hF')B - A(hF') = hq(c^t\beta_B - \beta_A)rl$ is over $\mathbb{R}^{++}[\text{exp}]$. Use an approximation argument, as at the end of the proof of Theorem 4, to find a nontrivial matrix F over $\mathbb{Z}^+[\text{exp}]$ such that $c^tFB - AF$ is over $\mathbb{Z}^{++}[\text{exp}]$. \square

Example. For an application of the above, let a, b be multiplicatively independent elements of \mathbb{R}^{++} and, taking $x = a^t, y = b^t$, consider $p(x, y) = 1 + x + x^4 + 2x^2y$ and $q(x, y) = 1 + x + x^4 + xy + x^3y + x^2y^2$. Let B be a matrix over $\mathbb{Z}^+[\text{exp}]$ such that $\beta_B = d^tq$, with $d \in \mathbb{R}^{++}$. In particular, we could take $B = d^tq$ and have as Σ_B the Bernoulli shift based on q . We claim that

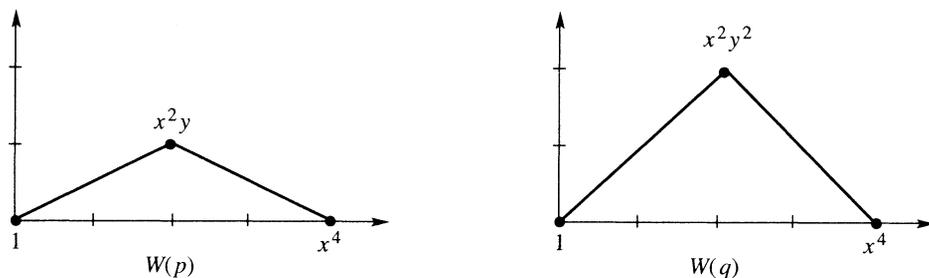


FIGURE 1

no Markov chain that is finitely equivalent to a subsystem of Σ_B can have β -function $\bar{d}^t p$, where $\bar{d} \in \mathbb{R}^{++}$. To verify this, let A be a matrix over $\mathbb{Z}^+[\text{exp}]$ with $\beta_A = \bar{d}^t p$ and suppose Σ_A is finitely equivalent to a subsystem of Σ_B . Then, by Corollary 7, there exists $c \in \mathbb{R}^{++}$ such that $c^t d^t q - \bar{d}^t p$ is Handelman. Note that $c^t d^t q, \bar{d}^t p \in \mathbb{Z}^{++}[\text{exp}]$. It is easily seen that if $\bar{q}, \bar{p} \in \mathbb{Z}^{++}[\text{exp}]$ and $\bar{q} - \bar{p}$ is Handelman then $W(\bar{p}) \subset W(\bar{q})$. Thus, $W(\bar{d}^t p) \subset W(c^t d^t q)$; that is, a translate of $W(p)$ is contained in $W(q)$. The polytopes $W(p), W(q)$ are depicted in Figure 1.

Clearly, the only translate of $W(p)$ that is contained in $W(q)$ is $W(p)$ itself, so $\bar{d} = cd$ and $q - p$ must be Handelman; but

$$q - p = (1 - x)^2 xy + x^2 y^2$$

is not Handelman since it has $(1 - x)^2 xy$ as a face. This contradiction shows that no Markov chain that is finitely equivalent to a subsystem of Σ_B can have $\bar{d}^t p$ as its β -function. In particular, the Bernoulli shift based on p is not finitely equivalent to a subsystem of the Bernoulli shift based on q .

ACKNOWLEDGMENTS

Over the years, we had discussions related to the contents of this paper with many mathematicians. We thank them all but cite only those whose conversations we benefited from within the past year or so: Jon Ashley, Valerio De Angelis and David Handelman. We are particularly grateful to Valerio De Angelis for pointing out a gap in the proof of Theorem 1 in an earlier draft of the paper.

REFERENCES

- [AM] R. Adler and B. Marcus, *Topological entropy and equivalence of dynamical systems*, Mem. Amer. Math. Soc., vol. 219, Amer. Math. Soc., Providence, RI, 1979.
- [Ash] J. Ashley, *Bounded-to-one factors of an aperiodic shift of finite type are 1-to-1 almost everywhere factors also*, Ergodic Theory Dynamical Systems **10** (1990), 615–625.
- [AMPT] J. Ashley, B. Marcus, D. Perrin, and S. Tuncel, *Surjective extensions of sliding block codes*, SIAM J. Discrete Math. (to appear).
- [D] V. De Angelis, Ph. D. thesis, Univ. of Washington, Seattle, 1992.

- [H1] D. Handelman, *Positive polynomials and product type actions of compact groups*, Mem. Amer. Math. Soc., vol. 320, Amer. Math. Soc., Providence, RI, 1985.
- [H2] ———, *Positive polynomials, convex integral polytopes, and a random walk problem*, Lecture Notes in Math., vol. 1282, Springer-Verlag, Berlin, Heidelberg, and New York, 1987.
- [H3] ———, *Eventual positivity and finite equivalence for matrices of polynomials*, preprint.
- [HLP] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge Univ. Press, Cambridge, 1959.
- [K] B. Kitchens, *Linear algebra and subshifts of finite type*, Contemp. Math., vol. 26, Amer. Math. Soc., Providence, RI, 1981, pp. 231–248.
- [KT1] B. Kitchens and S. Tuncel, *Finitary measures for subshifts of finite type and sofic systems*, Mem. Amer. Math. Soc., vol. 338, Amer. Math. Soc., Providence, RI, 1985.
- [KT2] ———, *On measures induced on subsystems*, Lecture Notes in Math., vol. 1342, Springer-Verlag, Berlin, Heidelberg, and New York, 1988, pp. 455–464.
- [Kr] W. Krieger, *On the finitary isomorphisms of Markov shifts that have finite expected coding time*, Z. Wahr. **65** (1983), 323–328.
- [MT] B. Marcus and S. Tuncel, *The weight-per-symbol polytope and scaffolds of invariants associated with Markov chains*, Ergodic Theory Dynamical Systems **11** (1991), 129–180
- [P1] W. Parry, *A finitary classification of topological Markov chains and sofic systems*, Bull. London Math. Soc. **9** (1977), 86–92.
- [P2] ———, *Problems and perspectives in the theory of Markov shifts*, Lecture Notes in Math., vol. 1342, Springer-Verlag, Berlin, Heidelberg, and New York, 1988, pp. 626–637.
- [PS] W. Parry and K. Schmidt, *Natural coefficients and invariants for Markov shifts*, Invent. Math. **76** (1984), 15–32.
- [PT] W. Parry and S. Tuncel, *On the classification of Markov chains by finite equivalence*, Ergodic Theory Dynamical Systems **1** (1981), 303–335.
- [Poin] H. Poincaré, *Sur les équations algébriques*, C. R. Acad. Sci. Paris **8** (1883), 1418.
- [Se] E. Seneta, *Non-negative matrices and Markov chains*, Springer-Verlag, Berlin, Heidelberg, and New York, 1981.
- [T] S. Tuncel, *Conditional pressure and coding*, Israel J. Math. **39** (1981), 101–112.

IBM ALMADEN RESEARCH CENTER, K65-802, 650 HARRY ROAD, SAN JOSE, CALIFORNIA 95120
E-mail address: marcus@almaden.ibm.com

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WASHINGTON 98195
E-mail address: tuncel@math.washington.edu