

# Subsystems, Perron numbers, and continuous homomorphisms of Bernoulli shifts

SELIM TUNCEL

*Department of Mathematics, University of Washington, Seattle, WA 98195, USA*

(Received 26 May 1988)

**Abstract.** Let  $S, T$  be subshifts of finite type, with Markov measures  $p, q$  on them, and let  $\phi: (S, p) \rightarrow (T, q)$  be a block code. Let  $I_p, I_q$  denote the information cocycles of  $p, q$ . For a subshift of finite type  $\bar{T} \subset T$ , the pressure of  $-I_q|_{\bar{T}}$  equals that of  $-I_q|_{\phi^{-1}(\bar{T})}$ . Applying this to Bernoulli shifts and using finiteness conditions on Perron numbers, we have the following. If the probability vector  $p = (p_1, \dots, p_{k+1})$  is such that the distinct transcendental elements of  $\{p_1/p_{k+1}, \dots, p_k/p_{k+1}\}$  are algebraically independent then the Bernoulli shift  $B(p)$  has finitely many Bernoulli images by block codes.

## 1. Introduction

Let  $S, T$  be subshifts of finite type, with Markov measures  $p, q$  on them, and let  $\phi: S \rightarrow T$  be a continuous homomorphism (block code), so that  $p \circ \phi^{-1} = q$ . Let  $I_p, I_q$  denote the information cocycles of  $p, q$ . We show that, for a subshift of finite type  $\bar{T} \subset T$ , the pressure of  $-I_q|_{\bar{T}}$  equals that of  $-I_p|_{\phi^{-1}(\bar{T})}$ , taken in  $\bar{T}$  and  $\phi^{-1}(\bar{T})$ , respectively. Applying this, we first find that subsystems of Markov chains (in the sense of [6]) are respected by block codes. Then we specialize to Bernoulli shifts. We denote by  $B(p)$  the Bernoulli shift defined by the probability vector  $p = (p_1, p_2, \dots, p_{k+1})$ , put  $\tau_1 = p_1/p_{k+1}, \tau_2 = p_2/p_{k+1}, \dots, \tau_k = p_k/p_{k+1}$ , and consider the case where the distinct transcendental elements of  $\{\tau_1, \tau_2, \dots, \tau_k\}$  are algebraically independent. This includes the case where  $\{\tau_1, \dots, \tau_k\}$  consists of algebraic numbers. Relating to Perron numbers the entries of a vector  $q$  for which there exists a block code  $\phi: B(p) \rightarrow B(q)$ , we use finiteness conditions involving Perron numbers to show that there are finitely many Bernoulli images of  $B(p)$  by block codes. The clustering results of [4] and [5] may also be established by such arguments; the paper ends with some related questions.

Throughout the paper, the case of interest is that of infinite-to-one codes, though we never make this assumption; for the bounded-to-one case stronger results are already known.

Most of the notation and terminology we use is well-known. For specificity, we mostly follow [2] and make the definitions explicit when they differ from [2]. As in [2], we use the same symbol to denote a subshift of finite type, its space, and its transformation. Whenever we consider a Markov chain  $(T, q)$  we mean the support of  $q$  to be the whole of  $T$ .

**Acknowledgement.** This paper benefitted from discussions with Mike Boyle, Mike Keane and, especially, Brian Marcus. The author was partially supported by NSF Grant DMS-8803260.

2. Subsystems and pressure

Let  $(S, p)$  be a Markov chain, and let  $\bar{S} \subset S$  be a subshift. We write  $\mathcal{P}$  for pressure. In particular, restricting the information cocycle  $I_p$  to  $\bar{S}$ , we write  $\mathcal{P}(-I_p|_{\bar{S}})$  for the pressure of  $-I_p|_{\bar{S}}$  taken on  $\bar{S}$ . See [11] for a detailed account of pressure.

**THEOREM 1.** *Let  $\phi : (S, p) \rightarrow (T, q)$  be a clock code between Markov chains, and let  $\bar{T} \subset T$  be a subshift of finite type. Then  $\phi^{-1}(\bar{T})$  is also of finite type, and*

$$\mathcal{P}(-I_p|_{\phi^{-1}(\bar{T})}) = \mathcal{P}(-I_q|_{\bar{T}}).$$

*Proof.* Without loss of generality, assume that  $S, T$  and  $\bar{T}$  are 1-step subshifts of finite type and  $\phi$  is a 1-block map. It is well known that  $\phi^{-1}(\bar{T})$  is then a (possibly reducible) 1-step subshift of finite type – a simple proof of this appears in [6]. Observe that  $\mathcal{P}(-I_q|_{\bar{T}})$  is the growth rate of

$$\beta_n = \sum_{j_0 \cdots j_n \in \bar{T}} q[j_0 j_1 \cdots j_n],$$

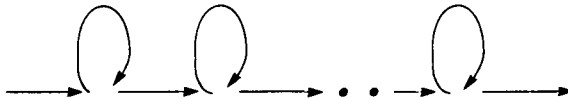
where the sum is over all  $\bar{T}$ -words of length  $n + 1$ . Since  $\phi$  is measure-preserving, we have

$$\beta_n = \sum_{j_0 \cdots j_n \in \bar{T}} q[j_0 j_1 \cdots j_n] = \sum_{\phi(i_0 \cdots i_n) \in \bar{T}} p[i_0 i_1 \cdots i_n],$$

the last sum being over all  $S$ -words  $i_0 i_1 \cdots i_n$  such that  $\phi(i_0 i_1 \cdots i_n)$  is a  $\bar{T}$ -word. On the other hand,  $\mathcal{P}(-I_p|_{\phi^{-1}(\bar{T})})$  is the growth rate of

$$\alpha_n = \sum_{i_0 i_1 \cdots i_n} p[i_0 i_1 \cdots i_n],$$

where the sum is over all  $S$ -words  $i_0 i_1 \cdots i_n$  such that  $(\phi(i_0 i_1 \cdots i_n))$  is a  $\bar{T}$ -word and) there exists a point  $x = (x_m) \in S$  with  $x_0 x_1 \cdots x_n = i_0 i_1 \cdots i_n$  and  $\phi(x) \in \bar{T}$ . Clearly,  $\alpha_n \leq \beta_n$ . Moreover, there exists  $l$  such that, for all large  $n$  and all  $i_0 i_1 \cdots i_n$  for which  $\phi(i_0 i_1 \cdots i_n)$  is a  $\bar{T}$ -word, the sub-word  $i_l \cdots i_{n-l}$  appears in an element of  $\phi^{-1}(\bar{T})$ . This is because every long  $i_0 i_1 \cdots i_n$  consists of periodic stretches separated by connecting blocks. That is,  $i_0 i_1 \cdots i_n$  looks like



and the length of the word before the first loop and that of the word following the last loop are bounded above by an integer,  $l$ , independent of  $n$ . We can delete the word before the first loop and that following the last loop and extend the remaining sub-word of  $i_0 i_1 \cdots i_n$  to an element  $x \in S$  by cycling around the first and last loops; if  $\phi(i_0 i_1 \cdots i_n)$  is a  $\bar{T}$ -word, we will have  $\phi(x) \in \bar{T}$ . It follows that  $\beta_{n-2l} \leq \alpha_n \leq \beta_n$  and that

$$\mathcal{P}(-I_p|_{\phi^{-1}(\bar{T})}) = \limsup \frac{1}{n} \log \alpha_n = \limsup \frac{1}{n} \log \beta_n = \mathcal{P}(-I_q|_{\bar{T}}). \quad \square$$

We first use Theorem 1 in the context of the induced measures defined in [6]. Let  $\bar{S} \subset S$  be an irreducible subshift of finite type. The measure induced on  $\bar{S}$  by  $p$  is defined to be the unique equilibrium state of  $-I_p|_{\bar{S}}$ . This is a Markov measure, denoted by  $p_{\bar{S}}$ . (See [6].)

**COROLLARY.** *Let  $\phi: (S, p) \rightarrow (T, q)$  be a block code between Markov chains, and let  $\bar{T} \subset T$  be an irreducible subshift of finite type. If  $\bar{S}$  is an irreducible component of  $\phi^{-1}(\bar{T})$  with  $\mathcal{P}(-I_p|_{\bar{S}}) = \mathcal{P}(-I_p|_{\phi^{-1}(\bar{T})})$ , then the restriction  $\phi: \bar{S} \rightarrow \bar{T}$  is surjective and satisfies  $p_{\bar{S}} \circ \phi^{-1} = q_{\bar{T}}$ .*

*Proof.* Since  $\mathcal{P}(-I_p|_{\phi^{-1}(\bar{T})})$  is the maximum of  $\mathcal{P}(-I_p|_{\bar{S}})$  over the finitely many irreducible components  $\bar{S}$  of  $\phi^{-1}(\bar{T})$ , there exists  $\bar{S}$  with  $\mathcal{P}(-I_p|_{\bar{S}}) = \mathcal{P}(-I_p|_{\phi^{-1}(\bar{T})})$ . Suppose, for such an  $\bar{S}$ , we have  $\phi(\bar{S}) \neq \bar{T}$ . It is not hard to see that we could then find a subshift of finite type  $\tilde{T}$  with  $\phi(\bar{S}) \subset \tilde{T} \subset \bar{T}$  and  $\tilde{T} \neq \bar{T}$ . As  $\tilde{T} \subset \bar{T}$  and  $\tilde{T} \neq \bar{T}$ , we have  $\mathcal{P}(-I_q|_{\tilde{T}}) < \mathcal{P}(-I_q|_{\bar{T}})$ . On the other hand, using Theorem 1,

$$\mathcal{P}(-I_q|_{\bar{T}}) = \mathcal{P}(-I_p|_{\phi^{-1}(\bar{T})}) = \mathcal{P}(-I_p|_{\bar{S}}) \leq \mathcal{P}(-I_p|_{\phi^{-1}(\tilde{T})}) = \mathcal{P}(-I_q|_{\tilde{T}}).$$

This contradiction shows that  $\phi(\bar{S}) = \bar{T}$ . Write

$$c = \mathcal{P}(-I_p|_{\bar{S}}) = \mathcal{P}(-I_p|_{\phi^{-1}(\bar{T})}) = \mathcal{P}(-I_q|_{\bar{T}}).$$

In showing that  $\phi: \bar{S} \rightarrow \bar{T}$  sends  $p_{\bar{S}}$  to  $q_{\bar{T}}$ , we will consider positive functions on cylinders  $[j_0 j_1 \cdots j_n]$  of  $\bar{T}$  and, for two such functions  $f, g$ , write  $f \sim g$  if there exist positive constants  $k, K$  such that

$$kf[j_0 j_1 \cdots j_n] \leq g[j_0 j_1 \cdots j_n] \leq Kf[j_0 j_1 \cdots j_n]$$

for all  $[j_0 j_1 \cdots j_n]$ . For instance, the definition of  $q_{\bar{T}}$  implies that

$$q_{\bar{T}}[j_0 j_1 \cdots j_n] \sim \frac{1}{c^n} q[j_0 j_1 \cdots j_n].$$

Assume, without loss of generality, that  $S, T$  and  $\bar{T}$  are 1-step subshifts of finite type and  $\phi$  is a 1-block map, and let  $l$  be as in the proof of Theorem 1. Considering all possible extensions  $j_{-l} \cdots j_{-1} j_0 \cdots j_n j_{n+1} \cdots j_{n+l}$  of  $j_0 \cdots j_n$  in  $\bar{T}$ , we have

$$q_{\bar{T}}[j_0 \cdots j_n] \sim \frac{1}{c^n} \sum_{j_{-l} \cdots j_{n+l}} q[j_{-l} \cdots j_{n+l}] = \frac{1}{c^n} \sum_{j_{-l} \cdots j_{n+l}} p[i_{-l} \cdots i_{n+l}],$$

where the last sum is over all  $S$ -words  $i_{-l} \cdots i_{n+l}$  such that  $\phi(i_{-l} \cdots i_{n+l})$  is a  $\bar{T}$ -word and  $\phi(i_{-l} \cdots i_{n+l}) = j_{-l} \cdots j_{-1} j_0 \cdots j_n j_{n+1} \cdots j_{n+l}$  for some  $j_{-l} \cdots j_{-1}$  and  $j_{n+1} \cdots j_{n+l}$ . Let  $W$  denote the set of all  $S$ -words  $i_0 \cdots i_n$  with the property that  $\phi(i_0 \cdots i_n) = j_0 \cdots j_n$  and there exists an extension  $i_{-l} \cdots i_{n+l}$  of  $i_0 \cdots i_n$  such that  $\phi(i_{-l} \cdots i_{n+l})$  is a  $\bar{T}$ -word. Using the last sum, it is easy to see that

$$q_{\bar{T}}[j_0 \cdots j_n] \sim \frac{1}{c^n} \sum_{i_0 \cdots i_n \in W} p[i_0 \cdots i_n].$$

Moreover, our choice of  $l$  implies (see the proof of Theorem 1) that

$$W = \{i_0 \cdots i_n : i_0 \cdots i_n \text{ is a } \phi^{-1}(\bar{T})\text{-word and } \phi(i_0 \cdots i_n) = j_0 \cdots j_n\}.$$

Hence,

$$q_{\bar{T}}[j_0 \cdots j_n] \sim \frac{1}{c^n} \sum_{i_0 \cdots i_n \in W} p[i_0 \cdots i_n] \geq \frac{1}{c^n} \sum_{\substack{\bar{S}\text{-words} \\ i_0 \cdots i_n \in W}} p[i_0 \cdots i_n] \sim p_{\bar{S}}(\phi^{-1}[j_0 \cdots j_n]),$$

where the sum on the right hand side of the inequality is over all  $\bar{S}$ -words  $i_0 \cdots i_n$  with  $\phi(i_0 \cdots i_n) = j_0 \cdots j_n$ , and  $\phi^{-1}[j_0 \cdots j_n]$  stands for the inverse image of  $[j_0 \cdots j_n]$  under the restriction  $\phi: \bar{S} \rightarrow \bar{T}$ . This shows  $p_{\bar{S}} \circ \phi^{-1} \ll q_{\bar{T}}$  and, using ergodicity, we conclude that  $p_{\bar{S}} \circ \phi^{-1} = q_{\bar{T}}$ . □

### 3. Bernoulli shifts and Perron numbers

We now restrict to Bernoulli shifts. We use the same symbol, say  $p$ , to denote a probability vector and the Bernoulli measure defined by this vector;  $B(p)$  denotes the resulting Bernoulli shift. Let  $p_1, \dots, p_k, p_{k+1}$  be the distinct entries of  $p$ , and let  $a_i$  be the multiplicity with which  $p_i$  occurs, so that  $\sum_{i=1}^{k+1} a_i p_i = 1$ . For  $i = 1, \dots, k$ , put  $\tau_i = p_i / p_{k+1}$ . (This notation differs slightly from the one used in the introduction.)

Having fixed  $p$ , let  $q = (q_1, q_2, \dots, q_l)$  be another probability vector. The vector  $q$  is called a clustering of  $p$  if there exist non-negative integers  $b_{ij}$ ,  $1 \leq i \leq k+1$ ,  $1 \leq j \leq l$  such that  $\sum_{j=1}^l b_{ij} = a_i$  for all  $i$  and  $\sum_{i=1}^{k+1} b_{ij} p_i = q_j$  for all  $j$ . In this case, there is a 1-block map  $\phi: B(p) \rightarrow B(q)$ . It was conjectured in [4] that the existence of a block code  $\phi: B(p) \rightarrow B(q)$  forces  $q$  to be a clustering of  $p$ . [4] and [5] contain positive results in this direction. Counter-examples are given in [2] and [9]. According to [4] and [10], for  $q \neq p$ , a block code  $\phi: B(p) \rightarrow B(q)$  must be infinite-to-one.

A Perron number is an algebraic integer  $\lambda$  such that  $\lambda \geq 1$  and  $\lambda$  is strictly greater than the absolute values of its conjugates (other than itself). According to [7], a number is Perron if and only if it is the spectral radius of an aperiodic non-negative integral matrix. Furthermore, a positive real number is the spectral radius of a non-negative integral matrix if and only if a (positive integral) power of it is a Perron number; we will call these PF numbers (for Perron–Frobenius). For details, and other properties of Perron numbers, see [7].

Given the existence of a block code  $\phi: B(p) \rightarrow B(q)$ , we will relate the entries of  $q$  to Perron numbers and use algebraic arguments to recover results of [4, 5] and establish further finiteness results. To illustrate this method, we start with a proof of the following.

**THEOREM [5].** *Suppose  $p_1, \dots, p_{k+1}$  are rational and  $N$  is their common denominator. If there exists a block code  $\phi: B(p) \rightarrow B(q)$  then  $q_j = b_j / N$  for some  $b_j \in \mathbb{N}$  and  $j = 1, \dots, l$ .*

**LEMMA 1.** *If  $\lambda, \mu$  and  $\beta$  are PF numbers and  $\lambda + \mu = \beta$ , then  $\lambda, \mu \in \mathbb{Q}(\beta)$ .*

The proof of Lemma 1 is similar to that of Proposition 5 of [7]. Lemma 1 is also a special case of Lemma 2, whose proof is given in the next section.

*Proof of the theorem.* Applying Theorem 1 to each fixed point of  $B(q)$ , we find that each  $q_j$  is the spectral radius of a matrix whose entries lie in  $\{0, p_1, \dots, p_{k+1}\}$ .

Extracting the common denominator  $N$ , we have

$$q_j = \frac{1}{N} \operatorname{spr}(A_j) = \frac{1}{N} \lambda_j,$$

where  $A_j$  is a non-negative integral matrix and  $\lambda_j = \operatorname{spr}(A_j)$  is its spectral radius. Since  $\sum_{j=1}^l q_j = 1$ , we get  $\sum_{j=1}^l \lambda_j = N$ , and Lemma 1 shows that  $\lambda_j \in \mathbb{Q}$ . Since  $\lambda_j$  is an algebraic integer and  $\lambda_j \geq 1$ , this implies  $\lambda_j \in \mathbb{N}$ .  $\square$

A consequence of the foregoing theorem is the fact [4] that the only continuous Bernoulli images of  $B(N^{-1}, \dots, N^{-1})$  are its clusterings. A generalization of the proof will give us:

**THEOREM 2.** *Suppose  $\tau_1, \dots, \tau_k$  are algebraic numbers. Then there exists a Perron number  $\alpha$  with the following property: if  $B(q)$  is a continuous image of  $B(p)$  then each  $q_j = \lambda_j p_{k+1} / \alpha$  for some PF number  $\lambda_j \in \mathbb{Q}(\alpha)$ . In particular, there are finitely many Bernoulli images of  $B(p)$  by block codes.*

*Proof.* Put  $K = \mathbb{Q}(\tau_1, \dots, \tau_k)$ . By 5.3.8. of [12] and 3.1 of [3], there exists an algebraic integer  $\alpha \in K$  with  $\alpha > 1 > |\theta(\alpha)|$  for all monomorphisms  $\theta: K \rightarrow \mathbb{C}$  other than the identity. Replacing  $\alpha$  by a large power  $\alpha^n$  if necessary, we assume that  $\sigma_1 = \alpha\tau_1, \dots, \sigma_k = \alpha\tau_k$  are also Perron and  $K = \mathbb{Q}(\sigma_1) = \dots = \mathbb{Q}(\sigma_k) = \mathbb{Q}(\alpha)$ . Write  $\sigma_{k+1} = \alpha$ .

Suppose  $\phi: B(p) \rightarrow B(q)$  is a block code. For each fixed point  $j^\infty$  of  $B(q)$ , the singleton  $\{j^\infty\}$  is a subshift of finite type. As before, an application of Theorem 1 to  $\{j^\infty\}$  shows that  $q_j$  is the spectral radius of a matrix whose entries lie in  $\{0, p_1, \dots, p_{k+1}\}$ . Extracting  $p_{k+1} / \alpha$  from this matrix, we have

$$q_j = \frac{p_{k+1}}{\alpha} \operatorname{spr}(A_j) = \frac{p_{k+1}}{\alpha} \lambda_j$$

for a matrix  $A_j$  with entries in  $\{0, \sigma_1, \dots, \sigma_{k+1}\}$  and  $\lambda_j = \operatorname{spr}(A_j)$ . By Perron-Frobenius theory [8],  $\lambda_j$  is an eigenvalue of  $A_j$ . Since the non-zero entries of  $A_j$  are Perron numbers, it follows that  $\lambda_j$  is an algebraic integer and  $\lambda_j \geq 1$ . We will show that each  $\lambda_j$  is a PF number. Assume this for the moment. The equation  $\sum_{i=1}^{k+1} a_i p_i = 1$  implies

$$\frac{1}{p_{k+1}} = \sum_{i=1}^k a_i \tau_i + a_{k+1} \quad \text{and} \quad \frac{\alpha}{p_{k+1}} = \sum_{i=1}^{k+1} a_i \sigma_i$$

which, in turn, show that  $\alpha / p_{k+1}$  is a Perron number and  $1 / p_{k+1} \in K = \mathbb{Q}(\alpha)$ . Since  $\sum_{j=1}^l q_j = 1$ , we have  $\sum_{j=1}^l \lambda_j = \alpha / p_{k+1}$ . Now Lemma 1 implies  $\lambda_j \in \mathbb{Q}(\alpha / p_{k+1}) = \mathbb{Q}(\alpha) = K$ . Hence, each  $\lambda_j$  is a PF number,  $\lambda_j \in \mathbb{Q}(\alpha)$  and  $\lambda_j = q_j \alpha / p_{k+1} \leq \alpha / p_{k+1}$ . As in Proposition 3 of [7], it follows that there are finitely many possibilities for the  $\lambda_j$ , so that there are only a finite number of possibilities for  $q$  also.

It remains to show that  $\lambda_j$  is a PF number. If  $\alpha \in \mathbb{Q}$  then  $\sigma_1, \dots, \sigma_{k+1} \in \mathbb{Q}$  and, being algebraic integers,  $\sigma_1, \dots, \sigma_{k+1} \in \mathbb{N}$ , so that  $\lambda_j$  is a clearly a PF number. So, let us assume  $\alpha \notin \mathbb{Q}$ , and write  $\lambda = \lambda_j$ ,  $A = A_j$ . Let  $C$  be the companion matrix of  $\alpha = \sigma_{k+1}$ . The spectrum  $\sigma(C)$  of  $C$  consists of the conjugates of the Perron number  $\alpha$ , and  $\alpha$  is a simple eigenvalue. Letting  $d$  denote the degree of  $\alpha$ , an eigenvector corresponding to  $\alpha$  is  $(1, \alpha, \dots, \alpha^{d-1})$ . Since  $\alpha \notin \mathbb{Q}$ , one of the results in [3] yields

a matrix  $M \in SL(d, \mathbb{Z})$  such that  $B = MCM^{-1}$  is eventually positive, that is,  $B^n$  is strictly positive for all large  $n$ . Write  $B_{k+1} = B$ . Let  $\theta_1, \dots, \theta_d$  be the distinct embeddings (monomorphisms) of  $K$  into  $\mathbb{C}$  and, for, for  $i = 1, \dots, k$ , let  $\pi_i \in \mathbb{Q}[x]$  be such that  $\pi_i(\alpha) = \sigma_i$ . Put  $B_i = \pi_i(B)$ . Recalling  $\mathbb{Q}(\sigma_i) = K$ , we see that the spectrum  $\sigma(B_i) = \pi_i(\sigma(B)) = \pi_i(\sigma(C))$  then consists of the  $d$  distinct conjugates

$$\theta_1(\sigma_i) = \pi_i(\theta_1(\alpha)), \dots, \theta_d(\sigma_i) = \pi_i(\theta_d(\alpha))$$

of the Perron number  $\sigma_i$ . Moreover, for a vector  $v$  and  $1 \leq h \leq d$ , we have  $Bv = \theta_h(\alpha)v$  if and only if  $B_iv = \pi_i(\theta_h(\alpha))v$ . It follows that each  $B_i$  is eventually positive. In fact, there exists  $L$  such that:

any product  $B_{c_1}B_{c_2} \cdots B_{c_n}$ , with  $c_1, \dots, c_n \in \{1, \dots, k+1\}$  is strictly positive as long as  $n \geq L$ . (\*)

Let  $v > 0$  be such that  $Bv = \alpha v$ . By restricting  $A$  to an irreducible component with spectral radius  $\lambda$ , we assume that  $A$  is irreducible and let  $r > 0$  be such that  $Ar = \lambda r$ . Putting  $B_0 = 0, \sigma_0 = 0$ , we have  $B_iv = \sigma_iv$  for  $i = 0, 1, \dots, k+1$ . Define a new matrix  $D$  as the block matrix obtained by replacing, in  $A$ , each entry that equals  $\sigma_i$  by the corresponding matrix  $B_i$ . Let  $w$  be the vector obtained from  $r$  on replacing each entry  $r(s)$  by  $r(s)v$ . Explicitly, if we index  $A$  by  $\{s\}$  and  $B_i$  by  $\{h\}$  and identify  $B_{\sigma_i} = B_i$ , then  $D$  is indexed by pairs  $\{(s, h)\}$  and

$$D((s, h), (s', h')) = B_{A(s,s')}(h, h').$$

Also,  $w(s, h) = r(s)v(h)$ . Now

$$\begin{aligned} (Dw)(s, h) &= \sum_{(s', h')} D((s, h), (s', h'))r(s')v(h') \\ &= \sum_{s'} r(s') \sum_{h'} B_{A(s,s')}(h, h')v(h') \\ &= \sum_{s'} r(s')A(s, s')v(h) \\ &= v(h) \sum_{s'} A(s, s')r(s') = v(h)\lambda r(s) = \lambda w(s, h). \end{aligned}$$

That is,  $w$  is an eigenvector of  $D$  corresponding to  $\lambda$ . Moreover, (\*) implies that  $D$  is eventually non-negative and, because we also have  $w > 0$ , the Perron-Frobenius theorem shows that  $\lambda = \text{spr}(D)$ . Since  $D$  is a rational matrix, we can write  $D = N^{-1}E$  for some integral matrix  $E$  and  $N \in \mathbb{N}$ . Eventual non-negativity then gives  $n \in \mathbb{N}$  and a PF number  $\mu$  such that  $\lambda^n = \mu/N^n$ . We already know that  $\lambda$  is an algebraic integer and  $\lambda \geq 1$ . It follows that  $\lambda^n = \mu/N^n$  is a PF number. □

I would like to thank Brian Marcus for pointing out, while I was trying to carry out the above proof with non-negative matrices, that Handelman’s theorem can be used to make the argument work in the setting of eventually positive matrices. Motivated by this, Doug Lind has proved, in general, that if the non-zero entries of a matrix are PF numbers, then its spectral radius is also a PF number.

4. *When the transcendental elements of  $\{\tau_1, \dots, \tau_k\}$  are algebraically independent*

Retaining the notation of the preceding section, we now consider the case where

$B(p)$  is such that the transcendental elements of  $\{\tau_1, \dots, \tau_k\}$  are algebraically independent. Re-ordering  $\tau_1, \dots, \tau_k$  if necessary, we assume that  $\tau_1, \dots, \tau_r$  are algebraically independent and  $\tau_{r+1}, \dots, \tau_k$  are algebraic numbers. We put  $K = \mathbb{Q}(\tau_{r+1}, \dots, \tau_k)$ ,  $R = \mathbb{Z}[\tau_1, \dots, \tau_r]$ , let  $\mathcal{D}$  denote the ring of integers of  $K$ , and let  $F = R/R$  be the field of fractions of  $R$ . If a number  $\lambda$  satisfies a monic polynomial whose coefficients are elements of  $R$ , then  $\lambda$  is called an algebraic integer over  $R$ . Since  $R$  is a unique factorization domain, an algebraic integer  $\lambda$  over  $R$  has a minimal polynomial which is monic, over  $R$ , and divides every polynomial, over  $F$ , that  $\lambda$  satisfies. Thus, we can talk of the conjugates of  $\lambda$  over  $R$ . We say that a number  $\lambda$  is Perron over  $R$  if  $\lambda$  is an algebraic integer over  $R$  and  $\lambda$  is strictly greater than the absolute values of its other conjugates over  $R$ . Similarly,  $\lambda$  is PF over  $R$  if  $\lambda^n$  is Perron over  $R$  for some  $n \in \mathbb{N}$ . We will need the following lemmas.

LEMMA 2. *If  $\lambda, \mu, \beta$  are PF numbers over  $R$  and  $\lambda + \mu = \beta$ , then  $\lambda, \mu \in F(\beta)$ .*

LEMMA 3. *If  $\tau_1, \dots, \tau_r$  are algebraically independent and  $\alpha$  is an algebraic number then  $\sigma_1 = \alpha\tau_1, \dots, \sigma_r = \alpha\tau_r$  are algebraically independent.*

LEMMA 4. *If  $\sigma_1, \dots, \sigma_r$  are algebraically independent then they are also independent over  $\mathcal{D}$ , so that  $\mathcal{D}[\sigma_1, \dots, \sigma_r]$  is isomorphic to the polynomial ring  $\mathcal{D}[x_1, \dots, x_r]$  in  $r$  variables.*

Lemmas 3 and 4 can be established by the usual arguments involving symmetric polynomials. Note that Lemma 1 is a degenerate case of Lemma 2, with  $R = \mathbb{Z}$ . The following proof is similar to that of Proposition 5 of [7].

*Proof of Lemma 2.* Suppose  $\lambda \notin F(\beta)$ . Let  $f(x) = \prod_{i=1}^n (x - \lambda_i)$ , with  $\lambda_1 = \lambda$  and  $n \geq 2$ , be the minimal polynomial of  $\lambda$  over  $F(\beta)$ . Then the polynomial  $g(x) = \prod_{i=1}^n (x - \beta + \lambda_i)$  is also irreducible over  $F(\beta)$ . Since  $g(\mu) = 0$ ,  $g$  must be the minimal polynomial, over  $F(\beta)$ , of  $\mu$ . Hence,  $\beta - \lambda_2$  is a conjugate of  $\mu$ . Using the fact that  $\mu, \lambda$  are PF, we find

$$\mu \geq |\beta - \lambda_2| > |\beta - \lambda| = \mu,$$

which is impossible. Thus, we must have  $\lambda, \mu \in F(\beta)$ . □

Suppose  $\phi : B(p) \rightarrow B(q)$  is a block code. Initially, we argue as in the proof of Theorem 2: Find  $\alpha \in K$  such that  $\sigma_{r+1} = \alpha\tau_{r+1}, \dots, \sigma_k = \alpha\tau_k, \sigma_{k+1} = \alpha$  are Perron numbers and  $K = \mathbb{Q}(\sigma_i)$  for  $i = r+1, \dots, k+1$ . For  $i = r+1, \dots, k$  let  $\pi_i \in \mathbb{Q}[x]$  be such that  $\pi_i(\alpha) = \sigma_i$ . Put  $\sigma_1 = \alpha\tau_1, \dots, \sigma_r = \alpha\tau_r$  also. Use Theorem 1 to write

$$q_j = \frac{p_{k+1}}{\alpha} \text{spr}(A_j) = \frac{p_{k+1}}{\alpha} \lambda_j,$$

where  $A_j$  is an irreducible matrix whose non-zero entries are elements of  $\{\sigma_1, \dots, \sigma_{k+1}\}$  and  $\lambda_j = \text{spr}(A_j)$ . Hence,  $\lambda_j$  satisfies a monic polynomial whose coefficients lie in  $\mathbb{Z}[\sigma_1, \dots, \sigma_{k+1}]$ . By the usual arguments involving symmetric polynomials, it follows that  $\lambda_j$  is an algebraic integer over  $R$ . Use Handelmann's theorem [3] to conjugate the companion matrix of  $\alpha$  to an eventually positive integral matrix  $B = B_{k+1}$  with  $\text{spr}(B) = \alpha = \sigma_{k+1}$ . Put  $B_i = \pi_i(B)$  for  $i = r+1, \dots, k$ ,

and let  $\sigma_0 = 0$ . Replace, in  $A_j$ , each entry that equals  $\sigma_i$  by  $B_i$  for  $i = r + 1, \dots, k + 1$ , and by  $\sigma_i I$  for  $i = 0, 1, \dots, r$ . The result is an eventually non-negative matrix  $D_j$  with  $\text{spr}(D_j) = \lambda_j$ . Furthermore, the entries of  $D_j$  lie in  $\mathbb{Q} \cup \{\sigma_1, \dots, \sigma_r\}$ . Eventual non-negativity gives  $n, N \in \mathbb{N}$  and  $\mu_j$  such that  $\mu_j$  is PF over  $R$  and  $\lambda_j^n = \mu_j / N^n$ . As we already know  $\lambda_j$  to be an algebraic integer over  $R$ , it follows that  $\lambda_j$  is PF over  $R$ .

From  $\sum_{j=1}^l q_j = 1 = \sum_{i=1}^{k+1} a_i p_i$  we obtain

$$\sum_{j=1}^l \lambda_j = \frac{\alpha}{p_{k+1}} = \sum_{i=1}^{k+1} a_i \sigma_i = \sum_1^r a_i \sigma_i + \sum_{r+1}^{k+1} a_i \sigma_i.$$

Putting  $s = \sum_{j=1}^l \lambda_j - \sum_{i=1}^r a_i \sigma_i = \sum_{i=r+1}^{k+1} a_i \sigma_i$ , we see that  $s \in \mathcal{D}$ ,  $s$  is Perron, and

$$\sum_{j=1}^l \lambda_j = \sum_{i=1}^r a_i \sigma_i + s.$$

By Lemma 2, we have  $\lambda_j \in F(\sum_{i=1}^r a_i \sigma_i + s) = F(s)$ . Since

$$R \subset \mathcal{D}[\sigma_1, \dots, \sigma_r] \quad \text{and} \quad \mathbb{Q}(\sigma_1, \dots, \sigma_r, s) = \mathcal{D}[\sigma_1, \dots, \sigma_r] / \mathcal{D}[\sigma_1, \dots, \sigma_r],$$

this implies  $\lambda_j \in \mathbb{Q}(\sigma_1, \dots, \sigma_r, s)$ .

Now note that  $K = \mathbb{Q}(s)$  and that  $\mathcal{D} \subset \mathbb{Q}(s) = \mathcal{D} / \mathcal{D}$  is integrally closed. By Corollary 2 on p. 312 of [1].

$$\mathcal{D}[\sigma_1, \dots, \sigma_r] \subset \mathbb{Q}(\sigma_1, \dots, \sigma_r, s) = \mathcal{D}[\sigma_1, \dots, \sigma_r] / \mathcal{D}[\sigma_1, \dots, \sigma_r]$$

is also integrally closed. That is, if an element of  $\mathbb{Q}(\sigma_1, \dots, \sigma_r, s)$  is integral over  $\mathcal{D}[\sigma_1, \dots, \sigma_r]$  then it is, in fact, in  $\mathcal{D}[\sigma_1, \dots, \sigma_r]$ . Hence,  $\lambda_j \in \mathcal{D}[\sigma_1, \dots, \sigma_r]$ .

Let  $d$  be the block-length of  $\phi$ . Then the inverse image  $\phi^{-1}[j]$  of the cylinder  $[j]$  is a disjoint union of cylinders of length  $d$ , which allows us to write

$$\frac{p_{k+1}}{\alpha} \lambda_j = q_j = \left( \frac{p_{k+1}}{\alpha} \right)^d \pi,$$

where  $\pi$  is a sum of monomials of degree  $d$  in  $\sigma_1, \dots, \sigma_{k+1}$ . Therefore, we have an equation over  $\mathcal{D}[\sigma_1, \dots, \sigma_r]$ :

$$\left( \sum_{i=1}^r a_i \sigma_i + s \right)^{d-1} \lambda_j = \pi. \tag{**}$$

By Lemmas 3 and 4,  $\mathcal{D}[\sigma_1, \dots, \sigma_r]$  is a polynomial ring in  $r$  variables. As an element of  $\mathcal{D}[\sigma_1, \dots, \sigma_r]$ , the degree of  $\pi$  is at most  $d$ , and comparison of the degrees of the two sides of (\*\*) reveals that the degree of  $\lambda_j$  is, at most, 1. Write  $\lambda_j = \alpha_0 + \alpha_1 \sigma_1 + \dots + \alpha_r \sigma_r$  with  $\alpha_0, \alpha_1, \dots, \alpha_r \in \mathcal{D}$ . Then the coefficient of  $\sigma_i^d$  is  $a_i^d \alpha_i$  on the left hand side of (\*\*), while it is a non-negative integer on the right hand side. Hence, we can find  $v_j \in \mathcal{D}$  and non-negative  $b_{1j}, b_{2j}, \dots, b_{rj} \in \mathbb{Z}$  such that

$$\lambda_j = b_{1j} \sigma_1 + \dots + b_{rj} \sigma_r + v_j.$$

Now

$$\sum_{i=1}^r a_i \sigma_i + s = \frac{\alpha}{p_{k+1}} = \sum_{j=1}^l \lambda_j = \sum_{i=1}^r \left( \sum_{j=1}^l b_{ij} \right) \sigma_i + \sum_{j=1}^l v_j,$$

treated as an equation over  $\mathcal{D}[\sigma_1, \dots, \sigma_r]$ , implies that  $a_i = \sum_{j=1}^l b_{ij}$  for  $i = 1, \dots, r$  and  $\sum_{j=1}^l v_j = s$ . Go back to (\*\*) and consider the constant terms to see that  $v_j \geq 0$ .



Considering terms of degree  $d - 1$ , we have

$$\nu_j \left( \sum_1^r a_i \sigma_i \right)^{d-1} + \left( \sum_{i=1}^r b_{ij} \sigma_i \right) \left( \sum_1^r a_i \sigma_i \right)^{d-1} s$$

on the left hand side and Perron coefficients on the right. In particular,  $\nu_j a_i^{d-1} + b_{ij} a_i^{d-2} s$ , the coefficient of  $\sigma_i^{d-1}$ , is Perron. As  $a_i > 0$  is an integer,  $\nu_j a_i + b_{ij} s > 0$  exceeds the absolute values of its other conjugates. We summarize these facts as:

**THEOREM 3.** *Suppose  $\tau_1, \dots, \tau_r$  are algebraically independent and  $\tau_{r+1}, \dots, \tau_k$  are algebraic numbers. Let  $\alpha \in \mathbb{Q}(\tau_{r+1}, \dots, \tau_k)$  be the Perron number chosen above, and put  $\sigma_1 = \alpha \tau_1, \dots, \tau_k = \alpha \tau_k, \sigma_{k+1} = \alpha$ , and  $s = \sum_{i=r+1}^{k+1} a_j \sigma_i$ . If  $B(q)$  is a continuous image of  $B(p)$  then*

$$q_j = b_{1j} \sigma_1 + \dots + b_{rj} \sigma_r + \nu_j$$

for non-negative integers  $b_{ij}$  with  $\sum_j b_{ij} = a_i, i = 1, \dots, r$ , and non-negative algebraic integers  $\nu_j \in \mathbb{Q}(\tau_{r+1}, \dots, \tau_k)$  such that  $\sum_j \nu_j = s$  and each  $\nu_j a_i + b_{ij} s$  is positive and exceeds the absolute values of its other conjugates. In particular, there are finitely many Bernoulli images of  $B(p)$  by block codes.

**COROLLARY [5].** *If  $\tau_1, \dots, \tau_k$  are algebraically independent then the only continuous Bernoulli images of  $B(p)$  are its clusterings.*

The remaining case is that where the transcendental elements of  $\{\tau_1, \dots, \tau_k\}$  are algebraically dependent: When  $\tau_1, \dots, \tau_r$  are algebraically independent and  $\tau_{r+1}, \dots, \tau_k$  are algebraic over  $\mathbb{Z}[\tau_1, \dots, \tau_r]$  (and not all algebraic over  $\mathbb{Q}$ ), our method leads to conditions involving Perron numbers over  $\mathbb{Z}[\tau_1, \dots, \tau_r]$  and, since  $\mathbb{Z}[\tau_1, \dots, \tau_r]$  is not discrete, we are unable to reach the desired finiteness conclusion from this.

*Remarks and questions.* Let us write  $p^n$  for the probability vector obtained by admitting, for each  $(i_1, \dots, i_n) \in \{1, \dots, k+1\}^n, a_{i_1} a_{i_2} \dots a_{i_n}$  copies of  $p_{i_1} p_{i_2} \dots p_{i_n}$  among its entries. The finiteness conclusion reached in Theorems 2 and 3 may be put thus: there exists an integer  $n(p) > 0$ , depending only on  $p$ , such that for every continuous image  $B(q)$  of  $B(p)$  the vector  $q$  is a clustering of  $p^{n(p)}$ . Does this hold when the transcendental elements of  $\{\tau_1, \dots, \tau_k\}$  are algebraically dependent? Is  $n(p)$  bounded by a simple function of  $p$ , such as the number of entries of  $p$ , or the number of distinct entries? Can every block code  $\phi : B(p) \rightarrow B(q)$  be replaced by one of block length (at most)  $n(p)$ ?

REFERENCES

[1] N. Bourbaki. *Commutative Algebra, Elements of Mathematics*. Addison-Wesley: Reading, Mass., 1972.  
 [2] M. Boyle & S. Tuncel. Infinite-to-one codes and Markov measures. *Trans. Amer. Math. Soc.* **285** (1984), 657-684.  
 [3] D. Handelmann. Positive matrices and dimension groups affiliated to  $C^*$ -algebras and topological Markov chains. *J. Operator Theory* **6** (1981), 55-74.  
 [4] A. del Junco, M. Keane, B. Kitchens, B. Marcus & L. Swanson. Continuous homomorphisms of Bernoulli schemes. *Progress in Math.* **10**, pp. 91-111. Birkhäuser: Boston, 1981.

- [5] G. Keller. To appear in *Proc. Amer. Math. Soc.*
- [6] B. Kitchens & S. Tuncel. On measures induced on subsystems. *Dynamical Systems*, SLN 1342, pp. 455–464. Springer: New York, 1989.
- [7] D. Lind. The entropies of topological Markov shifts and a related class of algebraic integers. *Ergod. Th. & Dynam. Sys.* **4** (1984), 283–300.
- [8] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer: New York, 1981.
- [9] M. Smorodinsky. Block codes for Bernoulli shifts. *Israel J. Math.* **49** (1984), 325–330.
- [10] S. Tuncel. Conditional pressure and coding. *Israel J. Math.* **39** (1981), 101–112.
- [11] P. Walters. *An Introduction to Ergodic Theory*. Springer: New York, 1982.
- [12] E. Weiss. *Algebraic Number Theory*. McGraw-Hill: New York, 1963.