(Page proofs)

[3 2]

ISRAEL JOURNAL OF MATHEMATICS, 33 (1991), 1–23

THE ENTROPIES OF RENEWAL SYSTEMS*

BY

JACOB GOLDBERGER

Department of Mathematics, Tel Aviv University, Tel Aviv, Israel

AND

DOUGLAS LIND

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

AND

MEIR SMORODINSKY

Department of Statistics, Tel Aviv University, Tel Aviv, Israel

ABSTRACT

Renewal systems are symbolic dynamical systems originally introduced by Adler. If W is a finite set of words over a finite alphabet A, then the renewal system generated by W is the subshift $X_W \subset A^{\mathbb{Z}}$ formed by bi-infinite concatenations of words from W. Motivated by Adler's question of whether every irreducible shift of finite type is conjugate to a renewal system, we prove that for every shift of finite type there is a renewal system having the same entropy. We also show that every shift of finite type can be approximated from above by renewal systems, and that by placing finite-type constraints on possible concatenations, we obtain all sofic systems.

1. Introduction

Let A be a finite alphabet, and $W \subset A^*$ be a finite collection of words over A. Form the compact set X_W of $A^{\mathbb{Z}}$ consisting of all bi-infinite sequences of symbols that can be factored as a bi-infinite product of words from W. Then X_W is clearly

*The authors were supported in part by NFS grants DMS-8706284, DMS-8814159 and DMS-8820716.
Received February 21, 1990

invariant under the shift σ_W one symbol to the left. These symbolic dynamical systems were introduced by Adler, who called them **renewal systems** in analogy with renewal theory from probability. They can be described as the possible bi-infinite trips on a graph with one central node, and one loop for each word of W. Thus they are also called **loops systems**, or **flower automata** [BP]. This description shows that renewal systems are sofic.

Adler's motivation was an approach to the shift equivalence problem of R. Williams. If it were true that every irreducible shift of finite type were topologically conjugate to a renewal system, the additional renewal structure would provide a tool with which to attack shift equivalence. Since every renewal system is forwardly transitive (see Lemma 3.1), the questions can be stated as follows.

Adler's Problem: Is every irreducible shift of finite type topologically conjugate to a renewal system?

This problem is still open. However, we show here that for every shift of finite type there is a renewal system having the same entropy.

The paper is organized as follows. In §2 we discuss "sentences" that can be parsed into words from W, and show that the growth rate of the number of sentences of length n gives the topological entropy $h(\sigma_W)$. We also describe some motivating examples, including one due to S. Williams of a sofic system that cannot be conjugate to any renewal system. In §3 we observe in Proposition 3.1 that σ_W is topologically transitive in each direction, and in Proposition 3.2 that σ_W is topologically mixing if $\gcd\{|w| : w \in W\} = 1$. If σ is an irreducible shift of finite type, then in Theorem 3.3 we prove that for every $\varepsilon > 0$ there is a renewal system σ_W that is also a shift of finite type which factors onto σ with $h(\sigma_W) < h(\sigma) + \varepsilon$. The proof of our main result Theorem 4.1 on entropy is contained in §4, where the main ingredients are the techniques from [L1] and [L2] for constructing nonnegative matrices with prescribed spectral radius, a result of Handelman [H] on integral bases for eventually nonnegative matrices, and a method for introducing controlled ambiguity into sets of words. Finally, in §5 we show that if concatenations are constrained by a finite-type condition, then all sofic systems can be obtained from such systems.

The authors are grateful to Mike Keane for many fruitful discussions. In particular, his contribution to the proof of Theorem 4.1 was significant. This collaboration was a result of the Workshop on Ergodic Theory and Symbolic Dynamics held in Seattle during the summer of 1989. The Workshop was supported by grants from the National Science Foundation, IBM, the Milliman Endowment, and the University of Washington.

2. Words, Sentences, and Parsing

Let A be a finite alphabet, and $W \subset A^*$ be a finite set of words over A. We call W a **vocabulary**. A **sentence** over W is a string $s \in A^*$ that has at least one factorization, or parsing, into a concatenation of words from W. If each sentence has exactly one such parsing, then W is called **uniquely decipherable**. Vocabularies obeying this condition are also called **codes** [BP].

For any word $v \in A^*$, we denote its length (i.e. the number of letters it contains) by $|v|$. The empty word ε has length 0. If $s = w_1 \cdots w_r$ is a sentence over W, then its length is $|s| = |w_1| + \cdots + |w_r|$. Let $s_n(W)$ denote the number of sentences of length n, where $s_0(W) = 1$ since we consider ε as a concatenation of 0 words from W. It is a well-known fact from automata theory [BP] that the sequence $\{s_n(W)\}$ obeys a finite-order recurrence relation. Thus the generating function

$$S_W(u) = \sum_{n=0}^{\infty} s_n(W) u^n$$

is rational, say $p_W(u)/q_W(u)$. The growth rate of $s_n(W)$ is then $1/\lambda$, where λ is the smallest positive root of $q_W(u)$. The first result shows that this growth rate is also the topological entropy of σ_W.

LEMMA 2.1: *Let $W \subset A^*$ be a vocabulary, and (X_W, σ_W) be the corresponding renewal system. If $s_n(W)$ is the number of sentences over W of length n, then*

$$h(\sigma_W) = \limsup_{n\to\infty} \frac{1}{n} \log\, s_n(W).$$

Proof: Let $b_n(W)$ denote the number of blocks in A^n that can occur in the points of X_W. By definition of entropy,

$$h(\sigma_W) = \limsup_{n\to\infty} \frac{1}{n} \log\, b_n(W).$$

Since a sentence is an allowed block in X_W, clearly $s_n(W) \le b_n(W)$, so that

$$\limsup_{n\to\infty} \frac{1}{n} \log\, s_n(W) \le h(\sigma_W).$$

On the other hand, let $L = \max_{w\in W} |w|$. Since each block of length n must occur as a subblock of a sentence of length $n + l$ for some l with $0 \le l \le 2L - 2$, and there are $l + 1$ choices for its position, we have that

$$b_n(W) \le \sum_{l=0}^{2L-2} (l+1) s_{n+l}(W).$$

Let $a = \limsup_{n\to\infty}(1/n)\log s_n(W)$, and suppose $\varepsilon > 0$. Then $s_n(W) \le e^{n(a+\varepsilon)}$ for large enough n, so that

$$b_n(W) \le \sum_{l=0}^{2L-2}(l+1)e^{(n+l)(a+\varepsilon)} \le (2L-1)^2 e^{(2L-2)(a+\varepsilon)}e^{n(a+\varepsilon)}.$$

Thus

$$h(\sigma_W) = \limsup_{n\to\infty}\frac{1}{n}\log\, b_n(W) \le a+\varepsilon$$

for all $\varepsilon > 0$, so that $h(\sigma_W) \le a$. □

We now give some motivational examples of these ideas.

Example 2.2: Recall that a vocabulary is uniquely decipherable if every sentence can be parsed uniquely into words from W. In this case $S_W(u)$ has a simple expression. Since each sentence ends in a unique word, $s_n(W)$ obeys the recurrence relation

$$s_n(W) = \sum_{w\in W} s_{n-|w|}(W) = \sum_{k=1}^{L} c_k s_{n-k}(W), \tag{2.1}$$

where c_k denotes the number of words in W of length k, and $L = \max_{w\in W}|w|$. Using the initial conditions

$$s_n(W) = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } -L+1 \le n < 0, \end{cases} \tag{2.2}$$

multiplying (2.1) by u^n and summing over $n \ge 1$, we obtain that

$$S_W(u) - 1 = \left(\sum_{k=1}^{L} c_k u^k\right) S_W(u).$$

Thus

$$S_W(u) = \frac{1}{1-\sum_{k=1}^{L} c_k u^k} = \frac{1}{1-\sum_{w\in W} u^{|w|}}.$$

Since $q_W(u) = 1 - \sum_{k=1}^{L} c_k u^k \to -\infty$ as $u \to \infty$, is decreasing in u, and $q_W(0) = 1$, it follows that $q_W(u)$ has a unique positive root $1/\lambda$. Since the $c_k \ge 0$, a simple argument shows that $1/\lambda$ is the closest root to 0, so that the growth rate of $s_n(W)$, which is $h(\sigma_W)$ by Lemma 2.1, equals $\log\lambda$. This λ satisfies the polynomial $t^L - \sum_{k=1}^{L} c_k t^{L-k}$, so that λ has no other positive Galois conjugates. This immediately shows that certain entropies $\log\lambda$ cannot be the entropy of a uniquely decipherable renewal system, since certain λ have positive Galois conjugates. One example is $\lambda = (3+\sqrt{5})/2$. □

Our method of construction is to build vocabularies with certain ambiguities, so that some sentences can be parsed in more than one way. The ambiguity will be controlled, so that $S_W(u)$ can still be easily computed. The next example is a simple prototype of this method.

Example 2.3: Let $A = \{a, b\}$, and $W = \{a, ab, ba\}$. Notice that the sentence aba has the two parsings $a(ba)$ and $(ab)a$. Let $\mathcal{S}_n$ denote the collection of sentences of length n. Clearly

$$\mathcal{S}_n = \mathcal{S}_{n-1}a \cup \mathcal{S}_{n-2}ab \cup \mathcal{S}_{n-2}ba.$$

However, each sentence ending in aba occurs twice in this union, once in $\mathcal{S}_{n-1}a$ and once in $\mathcal{S}_{n-2}ba$. All other sentences occur once. It follows that $\{s_n(W)\}$ obeys the recurrence relation

$$s_n(W) = s_{n-1}(W) + 2s_{n-2}(W) - s_{n-3}(W),$$

where the important feature is the negative term resulting from the ambiguity. Using the initial conditions (2.2) with $L = 3$, we obtain as in Example 2.2 that the generating function is

$$S_W(u) = \frac{1}{1 - u - 2u^2 + u^3}.$$

Here $h(\sigma_W) = \log\ \lambda$, where $\lambda \cong 1.80194$ is a root of $t^3 - t^2 - 2t + 1$, and has a Galois conjugate $\lambda_2 \cong 0.44504 > 0$. Thus even this small amount of ambiguity is enough to obtain renewal systems that cannot be conjugate to uniquely decipherable renewal systems. It is the introduction of such ambiguities that provides the flexibility to obtain all entropies. □

Example 2.4: (S. Williams [W]): Let $X = \{0,1\}^{\mathbb{Z}}$ be the full 2-shift, and define a 2-block map ψ by $\psi(00) = \psi(11) = a$, $\psi(01) = b$, and $\psi(10) = c$. The image $\psi(X) \subset \{a, b, c\}^{\mathbb{Z}}$ is a sofic system, and S. Williams has shown by an argument using fixed points that this system cannot be topologically conjugate to any renewal system.

3. Renewal Systems

Let $W \subset A^*$ be a vocabulary, and denote the associated renewal system by (X_W, σ_W). In this section we discuss the mixing properties of σ_W. We also show that every irreducible shift of finite type is a factor of a renewal system with slightly higher entropy.

PROPOSITION 3.1: *Every renewal system is topologically transitive in each direction.*

Proof: This follows easily from the observation that every sentence can both precede and follow every other sentence. ☐

PROPOSITION 3.2: *Let W be a vocabulary. If $\gcd\{|w| : w \in W\} = 1$, then σ_W is topologically mixing of all orders.*

Proof: This follows from the fact that for n sufficiently large, there is a sentence over W of length n. ☐

Remark: The vocabulary $W = \{00, 01, 10, 11\}$ shows that the converse to Proposition 3.2 is false.

The following shows that shifts of finite type can at least be approximated by renewal systems.

THEOREM 3.3: *Let (X, σ) be an irreducible shift of finite type. For every $\varepsilon \geq 0$ there is a uniquely decipherable renewal system (X_W, σ_W) that is a shift of finite type and a continuous map $\psi : X_W \to X$ such that $\sigma\psi = \psi\sigma_W$ and* Use \colon
$h(\sigma_W) < h(\sigma) + \varepsilon$.

Proof: First assume that σ is mixing. There is an alphabet B, and a zero-one matrix T indexed by B such that $(X, \sigma) \cong (X_T, \sigma_T)$, where

$$X_T = \{x = (\xi_i) \in B^{\mathbf{Z}} : T_{\xi_i \xi_{i+1}} = 1 \text{ for } i \in \mathbf{Z}\},$$ $X_T =$

and σ_T is the shift on X_T. Since σ_T is mixing, for all sufficiently large n we have that $(T^n)_{\xi\eta} > 0$ for all $\xi, \eta \in B$. Let $M_n = \max\{(T^n)_{\xi\eta} : \xi, \eta \in B\}$. For each pair $\xi, \eta \in B$ there is a map $\varphi_{\xi\eta}$ from $\{1, \ldots, M_n\}$ onto the set of blocks in X_T of length $n+1$ beginning with ξ and ending with η.

Form the alphabet A consisting of all triples (ξ, k, r), where $\xi \in B$, $1 \leq k \leq n$, and $1 \leq r \leq M_n$. Construct the vocabulary W to contain all words of the form

$$w_{\xi,r} = (\xi, 1, r)(\xi, 2, r) \cdots (\xi, n, r).$$

Thus W contains $|B| M_n$ words of length n, and no others. Clearly W is uniquely decipherable.

Define an n-block map $\psi : X_W \to X_T$ as follows. If the n-block is a word $w_{\xi,r} \in W$, put $\psi(w_{\xi,r}) = \xi$. If the n-block is a terminal segment of one word followed by an initial segment of another, then it has the form

$$b = (\xi, k, r)(\xi, k+1, r) \cdots (\xi, n, r)(\eta, 1, s)(\eta, 2, s) \cdots (\eta, k-1, s).$$

Let $\varphi_{\xi,\eta}(r) = \xi\xi_2\xi_3 \dots \xi_n\eta$, and then put $\psi(b) = \xi_k$. The ψ defines a surjective shift-invariant map from X_W to X_T.

Since W is uniquely decipherable and all words in W have length n, we see that

$$s_{kn}(W) = s_n(W)^k = (|B|M_n)^k.$$

By Lemma 2.1,

$$h(\sigma_W) = \frac{1}{kn} \log\ s_{kn}(W) = \frac{1}{n} \log\ |B| + \frac{1}{n} \log\ M_n.$$

Noting that $(1/n) \log\ M_n \to h(\sigma_T)$, the result follows.

If (X_T, σ_T) is irreducible with period $p > 1$ then there is a subset $C \subset B$ such that for any $\xi, \eta \in C$ we have that $(T^{np})_{\xi\eta} > 0$ for all large enough n, while $(T^{np})_{\xi\eta} = 0$ if $\xi \in C$ and $\eta \notin C$. The construction then works as before, with C replacing B. $\square$

4. Entropies

In this section we prove our main result, that for every shift of finite type there is a renewal system having the same entropy.

Call an algebraic integer λ a **Perron number** if $\lambda \geq 1$ and λ is strictly greater than the absolute value of its other Galois conjugates. Denote the set of Perron numbers by $\mathbf{P}$. In [L1] it is shown that the entropies of mixing shifts of finite type are exactly the numbers $\log\ \lambda$ for $\lambda \in \mathbf{P}$. Call λ **weak Perron** if $\lambda \geq |\lambda_i|$ for every Galois conjugate λ_i of λ, or, equivalently, if $\lambda^k \in \mathbf{P}$ for some $k \geq 1$. Then the entropies of general shifts of finite type are just the numbers $\log\ \lambda$ with λ weak Perron. We will concentrate on the Perron case first, indicating briefly in the proof of Theorem 4.5 the modifications needed for the weak Perron case.

THEOREM 4.1: *Let λ be a Perron number. Then there is a finite alphabet A and a vocabulary $W \subset A^*$, such that the renewal system (X_W, σ_W) is topologically mixing and has topological entropy $h(\sigma_W) = \log\ \lambda$.*

We shall first prove a series of auxiliary results that will be used in the proof. Let λ have degree d. We may assume that $d \geq 2$ since the theorem is trivial for $\lambda \in \mathbb{N}$. We will first describe our method for constructing polynomials with largest root λ. In Lemma 4.2 we apply this method to produce a class of such polynomials $f_{B,\mathbf{n}}$ indexed by $\mathbf{n} \in \mathbb{N}^d$ with coefficients whose size and location obey certain inequalities. Lemma 4.3 is our device for controlling ambiguity, and Lemma 4.4 allows us to combine vocabularies over disjoint alphabets. By placing conditions on $\mathbf{n}$ and using these lemmas, we produce vocabularies W so that the recurrence relation obeyed by $s_n(W)$ is "dominated" by $_{B,\mathbf{n}}(t)$. The proof concludes by adding enough new words to W to form a new vocabulary $\bar{W}$ with

$$S_{\bar{W}}(u) = \frac{1}{u^m f_{B,\mathbf{n}}(1/u)},$$

where $m = \deg f_{B,\mathbf{n}}$, so that $h(\sigma_{\bar{W}}) = \log \lambda$.

Let C be the companion matrix of the minimal polynomial of λ. Since $\deg \lambda \geq 2$, an eigenvector for λ cannot be rational. Hence by a result of Handelman [H], there is an integral basis for $\mathbb{Z}^d$ with respect to which C is eventually positive. Let B be the matrix of C in this basis, so that $B^n > 0$ for large enough n.

The method in [L1] of constructing aperiodic nonnegative integral matrices with spectral radius λ can be described as follows. Suppose we find integral vectors $\mathbf{z}_1, \ldots, \mathbf{z}_n$ in $\mathbb{Z}^d$, all with positive coordinate in the dominant eigendirection, and such that

$$B\mathbf{z}_j = \sum_{i=1}^{n} m_{ij}\mathbf{z}_i, \tag{4.1}$$

where the m_{ij} are nonnegative integers. Then $M = [m_{ij}]$ has spectral radius λ. In fact, every nonnegative aperiodic integral matrix with spectral radius λ arises in this fashion [L1, Thm. 2].

Since B is eventually positive, there is a simple way to generate the required integral vectors. Let $\mathbf{e}_j$ denote the jth elementary column unit vector. Fix a d-tuple $\mathbf{n} = [n_1, \ldots, n_d] \in \mathbb{N}^d$. We will require that $\mathbf{n}$ be large enough so that

$$B^{n_j} > 0 \quad \text{for } 1 \leq j \leq d. \tag{4.2}$$

We will also require that

$$\gcd\{n_1, \ldots, n_d\} = 1 \tag{4.3}$$

for aperiodicity. Consider the integral vectors $\mathbf{z}_{i,j} = B^i\mathbf{e}_j$ for $1 \leq j \leq d$ and $0 \leq i \leq n_j - 1$. We write the images under B of these vectors by use of the relations

$$B\mathbf{z}_{i,j} = B^{i+1}\mathbf{e}_j = \mathbf{z}_{i+1,j} \quad (0 \leq i \leq n_j - 2),$$

$$B\mathbf{z}_{n_j-1,j} = B^{n_j}\mathbf{e}_j = \sum_{i=1}^{d}(B^{n_j})_{i,j}\mathbf{e}_i = \sum_{i=1}^{d}(B^{n_j})_{i,j}\mathbf{z}_{0,i},$$

where all the coefficients are nonnegative by (4.2). This produces a nonnegative integral matrix $M_{B,\mathbf{n}}$ of size $|\mathbf{n}| = n_1 + \cdots + n_d$. Strict positivity of the $(B^{n_j})_{i,j}$ shows that $M_{B,\mathbf{n}}$ is irreducible, and condition (4.3) shows that $M_{B,\mathbf{n}}$ is aperiodic.

Using row eliminations, the determinant defining the characteristic polynomial of $M_{B,\mathbf{n}}$ can be reduced to the d-dimensional determinant

$${}_{B,\mathbf{n}}(t) = \det\left[(t^{n_1}I - B^{n_1})\mathbf{e}_1, \ldots, (t^{n_d}I - B^{n_d})\mathbf{e}_d\right],$$

where I is the $d \times d$ identity matrix.

To obtain a more compact description of ${}_{B,\mathbf{n}}(t)$, put $(tI)^{\mathbf{n}} = [t^{n_1}I, \ldots, t^{n_d}I]$, $B^{\mathbf{n}} = [B^{n_1}, \ldots, B^{n_d}]$, both having size $d \times d^2$, and let

$$H = H_d = \left[\mathbf{e}_1\mathbf{e}_1^T, \ldots, \mathbf{e}_d\mathbf{e}_d^T\right]^T,$$

where $\mathbf{e}_j^T$ is the jth elementary row vector. Thus H is $d^2 \times d$. Then $(tI)^{\mathbf{n}}H = \operatorname{diag}[t^{n_1}, \ldots, t^{n_d}]$, and $B^{\mathbf{n}}H = \left[B^{n_1}\mathbf{e}_1, \ldots, B^{n_d}\mathbf{e}_d\right]$. Thus

$${}_{B,\mathbf{n}}(t) = \det\left[(tI)^{\mathbf{n}}H - B^{\mathbf{n}}H\right].$$

Suppose that $M = [m_{ij}]$ is a $d \times d$ matrix. If $J \subset \{1, \ldots, d\}$, let $M\langle J\rangle$ denote the determinant of the principal submatrix $[m_{ij}]_{i,j\in J}$. If $\mathbf{n} \in \mathbb{N}^d$, and $J = \{j_1, \ldots, j_r\}$, let $\mathbf{n}_J = [n_{j_1}, \ldots, n_{j_r}]$. When $J = \emptyset$ we put $M\langle\emptyset\rangle = 1$ and $|\mathbf{n}_\emptyset| = 0$. Expansion of the determinant shows that

$$\det\left[(tI)^{\mathbf{n}}H - M\right] = \sum_{J\subset\{1,\ldots,d\}} (-1)^{|J|}M\langle J\rangle t^{|\mathbf{n}|-|\mathbf{n}_J|},$$

where the sum is taken over all subsets J of $\{1, \ldots, d\}$, including $J = \emptyset$.

The following lemma shows that for every principal submatrix of $B^{\mathbf{n}}H$ of size ≥ 2, the product of the diagonal terms dominates the determinant.

LEMMA 4.2: *Let $\delta > 0$. Then there is an $N(B,\delta)$ such that if $\min(\mathbf{n}) \geq N(B,\delta)$, then for every $J \subset \{1, \ldots, d\}$ with $|J| \geq 2$ we have*

$$\delta \prod_{j\in J}(B^{\mathbf{n}}H)_{jj} > |(B^{\mathbf{n}}H)\langle J\rangle|. \qquad (4.4)$$

$\langle J \rangle$

Proof: There is a simple motivating idea for this result. Let $\mathbf{v}$ be a positive eigenvector for B with eigenvalue λ. Then there is a $\rho < \lambda$ and $\gamma_j > 0$ for $1 \le j \le d$ such that $B^{n_j}\mathbf{e}_j = \gamma_j \lambda^{n_j}\mathbf{v} + O(\rho^{n_j})$. Thus the entries in $B^{n_j}H$ are all of order λ^{n_j}, but the columns are nearly collinear. Under expansion of the determinant $(B^{\mathbf{n}}H)\langle J\rangle$, by multilinearity, each term involving two or more dominant factors $\gamma_j\lambda^{n_j}\mathbf{v}$ vanishes, and what is left has size $o(\lambda^{|\mathbf{n}_J|})$.

Since the characteristic polynomial of B is the minimal polynomial of λ, it is irreducible over $\mathbb{Q}$, so that the eigenvalues $\lambda_1 = \lambda, \lambda_2, \dots, \lambda_d$ are all simple. For each λ_j, there is a corresponding eigenvector $\mathbf{v}_j \in \mathbb{C}^d$. Since B is eventually positive, by the Perron-Frobenius theorem [S] we may assume that the entries of $\mathbf{v} = \mathbf{v}_1$ are strictly positive. Let $\Lambda = \operatorname{diag}[\lambda, \dots \lambda_d]$, and $V = [\mathbf{v}_1, \dots, \mathbf{v}_d]$, so that $BV = V\Lambda$. Thus $V^{-1}B = \Lambda V^{-1}$, showing that the first row of V^{-1} is a left eigenvector for B with eigenvalue λ, so this row is positive since $VV^{-1} = I \ge 0$. The equation $VV^{-1} = I$ shows that

$$\mathbf{e}_j = \sum_{i=1}^{d} (V^{-1})_{ij}\mathbf{v}_i.$$

Let $\gamma_j = (V^{-1})_{1j}$, so $\gamma_j > 0$ by the above. This shows that the $\mathbf{v}$-coordinates γ_j of the $\mathbf{e}_j$ form a left eigenvector for B with eigenvalue λ, hence are all positive. Let $\rho = \max_{2 \le j \le d} |\lambda_j|$. Then

$$B^n\mathbf{e}_j = \sum_{i=1}^{d} (V^{-1})_{ij}\lambda_i^n\mathbf{v}_i = \gamma_j\lambda^n\mathbf{v} + O(\rho^n) \quad \text{as } n \to \infty. \tag{4.5}$$

From this it follows that if $|J| \ge 2$ and $\mathbf{v} = [v_1, \dots, v_d]^T$, then

$$\prod_{j\in J} (B^{\mathbf{n}}H)_{jj} = \Big(\prod_{j\in J}\gamma_j v_j\Big)\lambda^{|\mathbf{n}_J|} + o(\lambda^{|\mathbf{n}_J|}). \tag{4.6}$$

On the other hand, expansion of $(B^{\mathbf{n}}H)\langle\langle J\rangle\rangle$ be using (4.5) and multilinearity, and observing that any determinant with two columns that are multiples of $\mathbf{v}$ must vanish, we obtain that

$$|(B^{\mathbf{n}}H)\langle\langle J\rangle\rangle| = o(\lambda^{|\mathbf{n}_J|}). \tag{4.7}$$

Comparison of (4.6) and (4.7) shows that (4.4) holds provided that $\min(\mathbf{n})$ is large enough. □

The next lemma describes our mechanism for building ambiguity into a vocabulary. It is an elaboration of the idea in Example 2.3.

LEMMA 4.3: *Let $n_1, \dots, n_r$ be distinct positive integers with $n_r > n_1 + \cdots + n_{r-1}$, and let $c_1, \dots, c_r$ be positive integers such that $2c_1 \cdots c_{r-1}$ divides c_r. Then there is a vocabulary W such that*

$$S_W(u) = \frac{1}{1 - \sum_{k=1}^{r} c_k u^{n_k} + \frac{1}{2}(c_1 c_2 \cdots c_r) u^{n_1 + \cdots + n_r}}.$$

Proof: For $1 \le k \le r-1$ and $1 \le i \le c_k$ introduce distinct symbols $a_{k,i}$. Augment these with additional distinct symbols b_m with $1 \le m \le c_r/(2c_1 \cdots c_{r-1})$ to form an alphabet A. Let $W \subset A^*$ be the vocabulary consisting of the words

$$a_{k,i}^{n_k} \qquad (1 \le k \le r-1, 1 \le i \le c_k)$$

and

$$a_{1,i_1}^{n_1} a_{2,i_2}^{n_2} \cdots a_{r-1,i_{r-1}}^{n_{r-1}} b_m^{n_r - (n_1 + \cdots + n_{r-1})},$$

$$b_m^{n_r - (n_1 + \cdots + n_{r-1})} a_{1,i_1}^{n_1} a_{2,i_2}^{n_2} \cdots a_{r-1,i_{r-1}}^{n_{r-1}},$$

where $1 \le i_k \le c_k$ for $1 \le k \le r-1$, and $1 \le m \le c_r/(2c_1 \cdots c_{r-1})$. Thus W contains c_k words of length n_k for $1 \le k \le r$, and no others.

Let $s_n(W)$ be the number of sentences of length n, where we put $s_0(W) = 1$, and $s_n(W) = 0$ for $n < 0$. If $\mathcal{S}_n$ denotes the collection of sentences of length n over W, then

$$\mathcal{S}_n = \bigcup_{w \in W} \mathcal{S}_{n-|w|} w.$$

As in Example 2.3, sentences ending in

$$a_{1,i_1}^{n_1} \cdots a_{r-1,i_{r-1}}^{n_{r-1}} b_m^{n_r - (n_1 + \cdots + n_{r-1})} a_{1,j_1}^{n_1} \cdots a_{r-1,j_{r-1}}^{n_{r-1}} \tag{4.8}$$

occur exactly twice in this union, while all other sentences occur once. There are

$$(c_1 \cdots c_{r-1})(c_r/2c_1 \cdots c_{r-1})(c_1 \cdots c_{r-1}) = \frac{1}{2} c_1 \cdots c_r$$

words of the form (4.8). Thus $s_n(W)$ obeys the recurrence relation

$$s_n(W) = \sum_{k=1}^{r} c_k s_{n-n_k}(W) - \frac{1}{2}(c_1 \cdots c_r) s_{n-(n_1 + \cdots + n_r)}(W) \quad (n \ge 1).$$

Multiplying this by u^n, summing over $n \geq 0$, and using the initial conditions shows that

$$S_W(u)\left(1 - \sum_{k=1}^{r} c_k u^{n_k} + \frac{1}{2}(c_1 \cdots c_r)u^{n_1+\cdots+n_r}\right) = 1.$$

□

The following result will allow us to apply Lemma 4.3 to to $\mathbf{n}_J$ for each subset $J \subset \{1, \ldots, d\}$ individually, and then compute the generating function for the combined vocabularies.

LEMMA 4.4: *Suppose that $A_1, \ldots, A_m$ are pairwise disjoint alphabets, that $W_i \subset A_i^*$ are vocabularies, and that the generating functions $S_{W_i}(u)$ all have the form*

$$S_{W_i}(u) = \frac{1}{1 - p_i(u)}.$$

If $W = W_1 \cup \cdots \cup W_m$, then

$$S_W(u) = \frac{1}{1 - p_1(u) - \cdots - p_m(u)}.$$

Proof: By induction it suffices to consider the case $m = 2$. Let $W = W_1 \cup W_2$. Since the alphabets are disjoint, each sentence in W^* ends in a unique maximal subsentence from W_1 or from W_2. Let $r_n(W_i)$ denote the number of such sentences of length n ending in a sentence from W_i. By convention, we put $r_0(W_i) = 1$. Thus

$$s_n(W) = \begin{cases} r_n(W_1) + r_n(W_2), & \text{if } n \geq 1, \\ r_n(W_1) + r_n(W_2) - 1 & \text{if } n = 0. \end{cases} \tag{4.9}$$

Removal of the maximal subsentence shows that

$$\begin{aligned} r_n(W_1) &= \sum_{k=1}^{n} r_{n-k}(W_2) s_k(W_1), \\ r_n(W_2) &= \sum_{k=1}^{n} r_{n-k}(W_1) s_k(W_2). \end{aligned} \tag{4.10}$$

Let $S_{W_i}^*(u) = S_{W_i}(u) - 1 = p_i(u)/(1 - p_i(u))$. Put $R_i(u) = \sum_{n=0}^{\infty} r_n(W_i)u^n$. Multiplication of (4.10) by u^n and summing over $n \geq 0$ gives

$$\begin{aligned} R_1(u) &= R_2(u) S_{W_1}^*(u) + 1, \\ R_2(u) &= R_1(u) S_{W_2}^*(u) + 1. \end{aligned}$$

Solving for the R_i and adding gives

$$R_1(u) + R_2(u) = \frac{2 + S^*_{W_1}(u) + S^*_{W_2}(u)}{1 - S^*_{W_1}(u)S^*_{W_2}(u)}.$$

By (4.9), the required generating function is

$$S_W(u) = R_1(u) + R_2(u) - 1 = \frac{1 + S^*_{W_1}(u) + S^*_{W_2}(u) + S^*_{W_1}(u)S^*_{W_2}(u)}{1 - S^*_{W_1}(u)S^*_{W_2}(u)}.$$

Using the relations $S^*_{W_i}(u) = p_i(u)/(1 - p_i(u))$ and some manipulation, this reduces to

$$S_W(u) = \frac{1}{1 - p_1(u) - p_2(u)}.$$

□

Proof of Theorem 4.1: We use the notation and assumptions of this section. Use B and $\mathbf{n}$ to form the polynomial replace u with t (my mistake)

$$_{B,\mathbf{n}}(t) = \det\big[(tI)^{\mathbf{n}}H - B^{\mathbf{n}}H\big] = \sum_{J\subset\{1,\ldots,d\}} (-1)^{|J|} c_J u^{|\mathbf{n}|-|\mathbf{n}_J|},$$

where $c_J = (B^{\mathbf{n}}H)\langle\langle J\rangle\rangle$. Abbreviate $c_{\{k\}}$ to c_k, and note from equation (4.5) that

$$c_k = (\gamma_k v_k)\lambda^{n_k} + o(\lambda^{n_k}), \tag{4.11}$$

where $\gamma_k, v_k > 0$. Assume that $\mathbf{n}$ is large enough so that (4.2) and (4.3) hold, and further that

$$n_k > n_1 + \cdots n_{k-1} \quad \text{for } 2 \le k \le d \tag{4.12}$$

This condition guarantees that all subsums $|\mathbf{n}_J|$ are distinct, and that we can apply Lemma 4.3 to $\mathbf{n}_J$. The exponential estimate (4.11) and Lemma 4.2 show that by making $n_k >> n_1 + \cdots + n_{k-1}$ we can find positive integers $c_j^{(J)}$ indexed by $J \subset \{1,\ldots,d\}$ with $|J| \ge 2$ and by $1 \le j \le d$ meeting the following conditions.

(i) $c_j^{(J)} = 0$ if $j \notin J$.
(ii) $c_j^{(J)} < 2^{-d}c_{\cancel{J}}$. Replace J with j (my mistake) l.c. j
(iii) If $J = \{j_1, \ldots, j_r\}$, then $2\prod_{k=1}^{r-1} c_{j_k}^{(J)}$ divides $c_{j_r}^{(J)}$.
(iv) $(1/2)\prod_{j\in J} c_j^{(J)} > |c_J|$.

For each $J \subset \{1, \dots, d\}$ with $|J| \geq 2$, using conditions (4.12) and (iii), we may apply Lemma 4.3 to find an alphabet A_J and a vocabulary $W_J \subset A_J^*$ such that

$$S_{W_J}(u) = \frac{1}{1 - \sum_{j \in J} c_j^{(J)} u^{n_j} + \frac{1}{2}\left(\prod_{j \in J} c_j^{(J)}\right) u^{|\mathbf{n}_J|}}.$$

The alphabets A_J can be taken pairwise disjoint, so application of Lemma 4.4 shows that for $W = \bigcup_{|J| \geq 2} W_J$ we have

$$S_W(u) = \frac{1}{1 - \sum_{|J| \geq 2} \sum_{j \in J} c_j^{(J)} u^{n_j} + \sum_{|J| \geq 2} \frac{1}{2}\left(\prod_{j \in J} c_j^{(J)}\right) u^{|\mathbf{n}_J|}}.$$

Noting from (4.12) that the exponents $|\mathbf{n}_J|$ are pairwise distinct, and using the estimates (ii) and (iv), we see that

$$S_W(u)^{-1} - \sum_{J \subset \{1, \dots, d\}} (-1)^{|J|} c_J u^{|\mathbf{n}_J|} = \sum_{J \subset \{1, \dots, d\}} r_J u^{|\mathbf{n}_J|}$$

has nonnegative coefficients $r_J \geq 0$. For each J with $|J| \geq 1$ let A'_J be an alphabet of r_J letters distinct from all others and each other, and let $W'_J = \{a^{|\mathbf{n}_J|} : a \in A'_J\}$. Since W'_J is uniquely decipherable,

$$S_{W'_J}(u) = \frac{1}{1 - r_J u^{|\mathbf{n}_J|}}.$$

Finally, let $\bar{A} = \bigcup_J (A_J \cup A'_J)$, and $\bar{W} = \bigcup_J (W_J \cup W'_J)$. Since the alphabets are disjoint, another application of Lemma 4.4 shows that \overbar W

$$S_{\bar{W}}(u) = \frac{1}{\sum_{J \subset \{1, \dots, d\}} (-1)^J c_J u^{|\mathbf{n}_J|}}.$$

\overbar W

Since λ is the largest root of $f_{B,\mathbf{n}}(t)$, $1/\lambda$ is the smallest root of $S_{\bar{W}}(u)$, proving that $h(\sigma_W) = \log \lambda$. $\square$ f ⊂ / \overbar W

The arguments here can easily be adapted to the weak Perron case.

THEOREM 4.5: *Let λ be a weak Perron number. Then there is a renewal system having topological entropy* $\log \lambda$.

Proof: Since λ is weak Perron, there is an integer $p \geq 1$ so that $\lambda^p \in \mathbf{P}$. By Theorem 4.1, there is an alphabet A and a vocabulary $W \subset A^*$ so that $h(\sigma_W) = p \log \lambda$. Define a monoid homomorphism $\varphi : A^* \to A^*$ by $\varphi(a) = a^p$ for $a \in A$. Then

$$h(\sigma_{\varphi(W)}) = \frac{1}{p} h(\sigma_W) = \log \lambda,$$

so the renewal system $(X_{\varphi(W)}, \sigma_{\varphi(W)})$ has the required properties. $\square$

5. Markov Renewal Systems

As Example 4.2 shows, not every sofic system is conjugate to a renewal system. We will show, however, that by introducing a finite-type constraint on the allowed concatenations of words, we can obtain all sofic systems.

Let $W \subset A^*$ be a vocabulary, and let T be a zero-one matrix indexed by W. Let $X_{T,W}$ denote the set of $x \in A^{\mathbb{Z}}$ that can be written as $\cdots w_{-1}w_0w_1\cdots$ with $w_i \in W$ and $T_{w_i w_{i+1}} = 1$ for all $i \in \mathbb{Z}$. Then $X_{T,W}$ is invariant under the shift $\sigma_{T,W}$. Call $(X_{T,W}, \sigma_{T,W})$ a **Markov renewal system.** When $T_{vw} = 1$ for all $v, w \in W$, this reduces to the renewal systems we have considered thus far. Every shift of finite type is trivially a Markov renewal system.

PROPOSITION 5.1: *A subshift is sofic if and only if it is topologically conjugate to a Markov renewal system.*

Proof: Let W and T define a Markov renewal system $(X_{T,W}, \sigma_{T,W})$. Replace each letter in each word of W by a new letter, so that all letters are distinct, creating a new vocabulary $\widehat{W}$. It is clear that $(X_{\widehat{W},T}, \sigma_{\widehat{W},T})$ is a shift of finite type, and that it factors onto $(X_{T,W}, \sigma_{T,W})$. Thus $(X_{T,W}, \sigma_{T,W})$ is sofic.

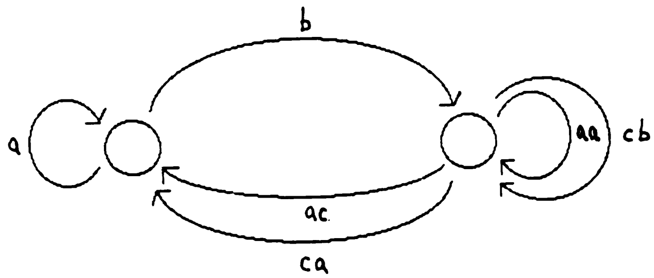
Conversely, suppose that (X,σ) is sofic. ~~By [BKM],~~ (X,σ) is the image under It is well-known that
a 2-block right resolving map of a shift of finite type. Thus there is a graph (see, for instance, [BKM])
$G = (V, E)$ and a labelling $l : E \to L$ of its edges, so that (X,σ) is conjugate to the sofic system $Y \subset L^{\mathbb{Z}}$ given by bi-infinite trips on the labelled graph.

Number the vertices of G as $V = \{i_1, \ldots, i_r\}$. For $1 \le k \le r$ let P_k denote the set of words w of length k so that there is a path labelled w starting at vertex i_k. The right resolving nature of the factor map means that if $w \in P_k$, then the path labelled w starting at i_k is unique. Let

$$W = \bigcup_{k=1}^{r} P_k \subset L^*.$$

Define T to be the zero-one matrix indexed by W so that $T_{v,w} = 1$ if and only if the terminal vertex of v equals the initial vertex of w. Then clearly any concatenation of words from W subject to the constraint from T gives a labelled path on G. Conversely, every bi-infinite labelled path on G can be decomposed into such a concatenation. Thus $Y = X_{T,W}$, concluding the proof. $\square$

Example 5.2: If this proof is carried out on Example 2.4 of a sofic system that is not conjugate to a renewal system, we obtain the following Markov renewal representation.



Replace with enclosed version.

References

[BP] J. Berstel and D. Perrin, *Theory of Codes*, Academic Press, New York, 1985.

[BKM] M. Boyle, B. Kitchens and B. Marcus, *A note on minimal covers for sofic systems*, Proc. Amer. Math. Soc. **95** (1985), 403–411.

[H] D. Handelman, *Positive matrices and dimension groups affiliated to C^*-algebras and topological Markov chains*, J. Operator Theory **6** (1981), 55–74.

[L1] D. Lind, *The entropies of topological Markov shifts and a related class of algebraic integers*, Ergodic Theory and Dynamical Systems **4** (1984), 283–300.

[L2] D. Lind, *The spectra of topological Markov shifts*, Ergodic Theory and Dynamical Systems **6** (1986), 571–582.

[S] E. Seneta, *Non-negative Matrices and Markov Chains*, Springer, New York, 1981.

[W] S. Williams, *Notes on renewal systems*, Proc. Amer. Math. Soc., ~~to appear.~~
110 (1990), 851-853.