International Journal of Bifurcation and Chaos, Vol. 14, No. 1 (2004) 309–319

ON THE SPATIAL ENTROPY OF TWO-DIMENSIONAL GOLDEN MEAN

JONQ JUANG
Department of Applied Mathematics, National Chiao Tung University, Hsinchu, Taiwan
jjuang@math.nctu.edu.tw

SHIH-FENG SHIEH
Department of Mathematics, National Tsing Hua University, Hsinchu, Taiwan
sfshieh@am.nthu.edu.tw

Received August 12, 2002; Revised October 21, 2002

The aim of this paper is to derive a sharper lower bound for the spatial entropy of two-dimensional golden mean.

Keywords: Spatial entropy; subshift of finite type; cellular neural networks; two-dimensional golden mean.

1. Introduction

The dynamical properties of one-dimensional subshifts of finite type (Markov shifts) are well understood. However, not much is known for a general theory of higher dimensional subshifts. For instance, the spatial entropy of subshifts of finite type is known to be the logarithm of the largest eigenvalue of its corresponding transition matrix. On the other hand, very little is known on the spatial entropy of higher dimensional subshifts. Even the "trivially" looking problem of the spatial entropy of two-dimensional golden mean $\mathbf{H} = \mathbf{V} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ remains open (see e.g. [Schmidt, 1990]). For the difficulties associated with higher dimensional Markov shifts, we refer to [Schmidt, 1990]. The two-dimensional golden mean problem corresponds to fill $\mathbb{Z}^2$ lattice with $\{1, 2\}$ with the following rules

$$\begin{array}{ccccc} * & & & 1 & \\ 1 & * & & 2 & 1\,, \end{array}$$

where $*$ indicates no restriction on what 1 can be adjacent to. Such a pattern can also be generated by cellular neural networks (CNNs) (see e.g. [Chua & Yang, 1988a, 1988b; Juang & Lin, 2000] and the work cited therein). More specifically, consider CNNs of the form

$$\frac{dx_{ij}}{dt} = -x_{ij} + z + \sum_{|k|\le 1,\, |l|\le 1} a_{k,l} f(x_{i+k,j+l}), \quad (i,j) \in \mathbb{Z}\,,$$

$$x_{i,j}(0) = x_{i,j}^0\,.$$

Here the nonlinearity f is a piecewise-linear function of the form

$$f(x) = \frac{1}{2}\left(|x+1| - |x-1|\right).$$

The numbers $a_{k,l}$, $|k| \le 1$, $|l| \le 1$, $k, l \in \mathbb{Z}$ are arranged in a 3×3 matrix form, which is called a space-invariant $\mathbf{A}$-template

$$\mathbf{A} = \begin{pmatrix} a_{-1,1} & a_{0,1} & a_{1,1} \\ a_{-1,0} & a_{0,0} & a_{1,0} \\ a_{-1,-1} & a_{0,-1} & a_{1,-1} \end{pmatrix}.$$

Now, set

$$\mathbf{A} = \begin{pmatrix} 0 & a\varepsilon & 0 \\ a\varepsilon & a\varepsilon & a\varepsilon \\ 0 & a\varepsilon & 0 \end{pmatrix}.$$

By choosing z, a and ε approximately, say $(z, a) \in [5, 1]_\varepsilon$, where $a\varepsilon < 0$ (see Theorem 3.5 of [Juang & Lin, 2000]), we see, via Lemma 4.1 of [Juang & Lin, 2000], that any positively saturated cell, denoted by 1, can be adjacent to either positively saturated cell or negatively saturated cell, denoted by 2. Moreover, any negatively saturated cell must be adjacent to at least four positively saturated cells. These mosaic patterns are exactly generated by two-dimensional golden mean. It is easy to see that the entropy h of such problem satisfies the inequality

$$\frac{1}{2}\log 2 < h < \log\frac{1+\sqrt{5}}{2},$$

but the precise value of h is still not known (see e.g. [Markely & Paul, 1981a, 1981b]). In this paper, we will give a nontrivial lower bound of h. We also note that most of the discussion of higher dimensional Markov shifts is restricted to examples of special nature (see e.g. [Baxter, 1982; Kasteleyn, 1961; Lieb, 1967; Schmidt, 1990; Temperley & Lieb, 1971]). We conclude this introductory section by summarizing the organization of this paper. In Sec. 2, we recall some needed notations, definitions and known results. In Sec. 3, we define a class of Markov measures associated with a transition matrix $\mathbf{A}$. Such class of the measures is then used to compute the measure theoretic entropy of the shift map $\sigma_{\mathbf{A}}$. In Sec. 4, we combine the results from Secs. 2 and 3 to get a nontrivial lower bound of the spatial entropy of two-dimensional gold mean.

2. Preliminaries

To make the paper self-contained, we recall some definitions and results. Let N be a positive integer with $N \geq 2$, let $\mathcal{S} = \{1, 2, \ldots, N\}$. Denote by $\mathbb{Z}^d$ the integer lattice on $\mathbb{R}^d$ where $d \geq 1$ is a positive integer representing the lattice dimension. The set of all functions $\mathbf{u} : \mathbb{Z}^d \to \mathcal{S}$ is denoted by $\mathcal{S}^{\mathbb{Z}^d}$. For $\alpha \in \mathbb{Z}^d$, we write $\mathbf{u}(\alpha)$ as $\mathbf{u}_\alpha$. The kth shift operator on $\mathcal{S}^{\mathbb{Z}^d}$ is defined by

$$(\sigma_k \mathbf{u})_\alpha = u_{\alpha+\mathbf{e}_k},$$

where $\alpha \in \mathbb{Z}^d$ and $\mathbf{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0)$ is the usual unit vector in the direction of the kth coordinate. For convenience we also write $\Sigma_N = \mathcal{S}^{\mathbb{Z}^d}$. We define a metric d on Σ_N as follows.

$$d(\mathbf{u}, \mathbf{v}) = \sum_{k\in\mathbb{Z}^d} \frac{\delta(u_k, v_k)}{3^{|k|}}, \tag{1}$$

where

$$\delta(i, j) = \begin{cases} 0, & \text{if } i = j, \\ 1, & \text{if } i \neq j, \end{cases}$$

and $|k| = \max\{k_1, k_2, \ldots, k_d\}$ for $k = (k_1, k_2, \ldots, k_d) \in \mathbb{Z}^d$. The space Σ_N with the shift operators, $(\Sigma_N; \sigma_1, \ldots, \sigma_d)$, is called the *symbol space on N symbols*, or the *full N-shift space*.

Definition 2.1. An $N \times N$ matrix $\mathbf{A} = (a_{ij})$ is said to be a transition matrix if

(i) $a_{ij} = 0$ or 1 for all $1 \leq i, j \leq N$,
(ii) $\sum_i a_{ij} \geq 1$ for all $1 \leq j \leq N$,
(iii) $\sum_j a_{ij} \geq 1$ for all $1 \leq i \leq N$.

Definition 2.2. Given d transition matrices, $\mathbf{A}_k = (a_{ij}^k)_{N\times N}$, $k = 1, \ldots, d$, let

$$\Sigma_{\mathbf{A}_1,\ldots,\mathbf{A}_d} = \{\mathbf{u} \in \Sigma_N | a^k_{u_\alpha, u_{\alpha+\mathbf{e}_k}} = 1, \quad \text{for all } \alpha \in \mathbb{Z}^d,\ 1 \leq k \leq d\}.$$

which determines all the admissible transitions between symbols $1, \ldots, N$. Each element in $\Sigma_{\mathbf{A}_1,\ldots,\mathbf{A}_d}$ is called a *pattern*. The shift operators $\sigma_1, \ldots, \sigma_d$ restricted on $\Sigma_{\mathbf{A}_1,\ldots,\mathbf{A}_d}$ are called the *subshifts of finite type* for matrices $\mathbf{A}_1, \ldots, \mathbf{A}_d$.

We shall write $\Sigma_{\mathbf{A}_1,\ldots,\mathbf{A}_d}$ as Σ_d provided no confusion arises. It is clear that Σ_d is closed with respect to the metric defined in (1) and translation invariant, that is,

$$\sigma_k(\Sigma_d) = \Sigma_d$$

for all $1 \leq k \leq d$. To measure the complexity of Σ_d, we compute the growth rate of the number of patterns on a parallelepiped of size $N_1 \times N_2 \times \cdots N_d$ on the lattice as $N_1, \ldots, N_d$ go to infinity.

Definition 2.3. The spatial entropy $h(\Sigma_d)$ is defined by

$$h(\Sigma_d) = \lim_{N_1,\ldots,N_d\to\infty} \frac{\log \Gamma_{N_1,\ldots,N_d}(\Sigma_d)}{N_1, N_2 \cdots N_d}. \tag{2}$$

Here $\Gamma_{N_1,\ldots,N_d}(\Sigma_d)$ is the number of distinct patterns that one observes among the elements of Σ_d by restricting one's observation to a parallelepiped of size $N_1 \times N_2 \times \cdots N_d$ on the lattice. The limit in (2) is well-defined and exists (see e.g. [Chow *et al.*, 1996]). Moreover, if Σ_d is replaced by $\mathcal{U}$ where $\mathcal{U} \subset \Sigma_N$ and satisfies

$$\sigma_1^{p_1}(\mathcal{U}) = \sigma_2^{p_2}(\mathcal{U}) = \cdots = \sigma_d^{p_d}(\mathcal{U}) = \mathcal{U}$$

for some $(p_1, p_2, \ldots, p_d) \in \mathbb{Z}^d$, the well-definedness and existence of the limit in (2) remain true (see e.g. [Juang *et al.*, 2002]).

Theorem 2.1 (see e.g. Theorem VIII-1.9 of [Robinson, 1993]). *For $d = 1$, let $\mathbf{A}$ be a transition matrix on N symbols, so $\mathbf{A}$ is $N \times N$. Then*

$$h(\Sigma_{\mathbf{A}}) = \log \lambda_1$$

where λ_1 is the dominant eigenvalue of $\mathbf{A}$.

Definition 2.4. Let $f : X \to X$ be a continuous map on the space X with metric d. For n a positive integer and $\varepsilon > 0$, a set $S \subset X$ is called (n, ε)-*separated* for f provided for every pair of distinct points $\mathbf{x}, \mathbf{y} \in S$, there is at least one k with $0 \le k < n$ such that $d(f^k(\mathbf{x}), f^k(\mathbf{y})) > \varepsilon$.

The number of different orbits of length n (as measured by ε) is defined by

$$\begin{aligned} r(n, \varepsilon, f) = \max\{\#(S) | S \subset X \text{ is a } (n, \varepsilon) \\ - \text{separated set for } f\}, \end{aligned}$$

where $\#(S)$ is the number (cardinality) of elements in S. To measure the growth rate of $r(n, \varepsilon, f)$ as n increases, we define

$$h(\varepsilon, f) = \limsup_{n\to\infty} \frac{\log r(n, \varepsilon, f)}{n}.$$

If $r(n, \varepsilon, f) = e^{n\tau}$, then $h(\varepsilon, f) = \tau$. Thus, $h(\varepsilon, f)$ means the "exponent" of the manner in which $r(n, \varepsilon, f)$ grows with respect to n. Finally, we consider the way that $h(\varepsilon, f)$ varies as ε goes to zero, and define the topological entropy of f as

$$h(f) = \lim_{\varepsilon\to 0} h(\varepsilon, f).$$

We note that for $0 < \varepsilon_1 < \varepsilon_2$, $r(n, \varepsilon_2, f) \ge r(n, \varepsilon_2, f)$, so $h(\varepsilon, f)$ increases as ε decreases and, hence, the limit defining $h(f)$ exists. If f is C^1 on a compact space, then it has been proven that $h(f) < \infty$ (see e.g. [Bowen, 1971, 1988]).

The following theorem shows that $h(\sigma_{\mathbf{A}}) = h(\Sigma_{\mathbf{A}})$.

Theorem 2.2 (see e.g. Theorem VIII.1.9 of [Robinson, 1993]). *Let $\sigma : \Sigma_N \to \Sigma_N$ be the full shift of N symbols (either one side of two). Assume $X \subset \Sigma_N$ is a closed invariant subset. Let Γ_n be the number of words of length n in X, i.e.*

$$\begin{aligned} \Gamma_n = \#\{(s_0, \ldots, s_{n-1}) | s_j = x_j, \\ \mathit{for}\ 0 \le j \le n\ \mathit{for\ some}\ \mathbf{x} \in X\}. \end{aligned}$$

Then

$$h(\sigma|_X) = \limsup_{n\to\infty} \frac{\log \Gamma_n}{n}.$$

We also need to recall two recursive formulas, which was derived in [Juang *et al.*, 2000] for computing the spatial entropy of two-dimensional golden mean. In the following, we first introduce some notations and concepts.

Given a transition matrix $\mathbf{A} = (a_{i,j})_{n\times n}$. A word $\omega = (\omega_0, \omega_1, \ldots, \omega_{k-1})$ of length k is called admissible (allowable) if $a_{\omega_{j-1},\omega_j} = 1$ for $j = 1, 2, \ldots, k-1$. Let $\mathbf{A}$ be a transition matrix. The set of admissible words of length m whose first symbol is ω_0 is to be denoted by $\omega(\omega_0, m; \mathbf{A})$. Set

$$\begin{aligned} \omega(m; \mathbf{A}) &= \text{set of all admissible words of length } m \\ &= \bigcup_{1\le\omega_0\le n} \omega(\omega_0, m; \mathbf{A}). \end{aligned}$$

To save notation, the transition matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ introduced in Definition 2.2 will be denoted by $\mathbf{H} = (h_{i,j})_{N\times N}$ and $\mathbf{V} = (v_{i,j})_{N\times N}$, respectively, called horizontal and vertical transition matrices. Then $\mathrm{Card}(\omega(m; \mathbf{H})) = \sum_{i,j=1}^{n} (\mathbf{H}^{m-1})_{i,j} =: N_m$. Here $\mathbf{H}^0$ = identity matrix. Using these N_m symbols, we may define a transition matrix $T^{(m)}_{\mathbf{H},\mathbf{V}} = (t^{(m)}_{i,j})$ of size $N_m \times N_m$ as follows. We begin with giving a lexicographic order for elements in $\omega(m; \mathbf{H})$. Specifically, let $\mathbf{s} = (s_1 s_2 \cdots s_m)$ and $\mathbf{p} = (p_1 p_2 \cdots p_m) \in \omega(m; \mathbf{H})$, and suppose that j is the smallest index for which $s_j \ne p_j$, then we define

$$\mathbf{s} < \mathbf{p} \quad \text{if } s_j < p_j. \tag{3}$$

With such ordering, the sets $\omega(m; \mathbf{H})$ and $\{1, 2, 3, \ldots, N_m\}$ can have an association that is one to one, onto and order preserving.

Definition 2.5. If $\mathbf{s}$ and $\mathbf{p}$ in $\omega(m; \mathbf{H})$ are associated with positive integers k and l, where $1 \le k$, $l \le N_m$ respectively, then we define the (k, l)-entry or $(\mathbf{s}, \mathbf{p})$-entry of $T^{(m)}_{\mathbf{H},\mathbf{V}}$ as

$$t^{(m)}_{\mathbf{s},\mathbf{p}} = t^{(m)}_{k,l} = v_{s_1,p_1} \cdot v_{s_2,p_2} \cdots v_{s_m,p_m} := \prod_{i=1}^{m} v_{s_i,p_i}, \tag{4}$$

i.e. $t^{(m)}_{k,l} = 1$ provided that for all $1 \le i \le m$, the words $\begin{pmatrix} s_i \\ p_i \end{pmatrix}$ are admissible with respect to $\mathbf{V}$. Otherwise, $t^{(m)}_{k,l} = 0$. For convenience, we shall use $t^{(m)}_{\mathbf{s},\mathbf{p}}$ to denote $t^{(m)}_{k,l}$. We shall call $T^{(m)}_{\mathbf{H},\mathbf{V}}$ the m-transition matrices with respect to the horizontal and vertical transition matrices $\mathbf{H}$ and $\mathbf{V}$, or for short, the m-transition matrix. If we start out with a lexicographic order for elements in $\omega(m;\mathbf{V})$, we shall obtain the so-called m-transition matrix $T^{(m)}_{\mathbf{V},\mathbf{H}}$ with respect to $\mathbf{V}$ and $\mathbf{H}$.

The relationship between m-transition matrix $T^{(m)}_{\mathbf{H},\mathbf{V}}$ and $h(\Sigma_{\mathbf{H},\mathbf{V}})$ is given in the following.

Proposition 2.1 (Proposition 2.1 of [Juang *et al.*, 2000]). *Let $T^{(m)}_{\mathbf{H},\mathbf{V}}$ be the m-transition matrix with respect to* $\mathbf{H}$ *and* $\mathbf{V}$. *Let $\rho(T^{(m)}_{\mathbf{H},\mathbf{V}})$ be the maximal eigenvalue of $T^{(m)}_{\mathbf{H},\mathbf{V}} = (t^{(m)}_{\mathbf{s},\mathbf{p}})$, where $t^{(m)}_{\mathbf{s},\mathbf{p}}$ are given in (4). Then*

$$h(\Sigma_2) = \lim_{m\to\infty} \frac{\log\, \rho(T^{(m)}_{\mathbf{H},\mathbf{V}})}{m}\,, \tag{5}$$

where $\Sigma_2 = \Sigma_{\mathbf{H},\mathbf{V}}$.

The following recursive formula for constructing $T^{(m)}_{\mathbf{H},\mathbf{V}}$ can also be found in [Juang *et al.*, 2000]. Note first that $T^{(m)}_{\mathbf{H},\mathbf{V}}$ can be written as the following block structure

$$T^{(m)}_{\mathbf{H},\mathbf{V}} = (T^{(m)}_{i,j})\,, \quad 1 \le i, j \le n\,, \tag{6}$$

where $T^{(m)}_{i,j}$ is a matrix of size $\mathrm{Card}(\omega(i,m;\mathbf{H}))\times \mathrm{Card}(\omega(j,m;\mathbf{H}))$. Let $1 \le k \le \mathrm{Card}(\omega(i,m;\mathbf{H}))$ and $1 \le l \le \mathrm{Card}(\omega(j,m;\mathbf{H}))$. Via the lexicographic order defined in (3), there exist $\mathbf{s} \in \omega(i,m;H)$ and $\mathbf{p} \in \omega(j,m;H)$ whose associated numbers are k and l, respectively. Then the (k,l)-entry, or simply $(\mathbf{s},\mathbf{p})$-entry, of the matrix $T^{(m)}_{i,j}$ is 1 provided that for all $1 \le r \le m$, $\begin{pmatrix} s_r \\ p_r \end{pmatrix}$ is an admissible word of size two with respect to vertical transition matrix $\mathbf{V}$. Otherwise, the entry is zero. We are now ready to state the following result.

Theorem 2.3. *Let $T^{(m+1)}_{\mathbf{H},\mathbf{V}}$ and $T^{(m)}_{\mathbf{H},\mathbf{V}}$ be, respectively, $(m+1)$- and m-transition matrix with respect to horizontal and vertical transition matrices $\mathbf{H} = (h_{i,j})$ and $\mathbf{V} = (v_{i,j})$. Let $\alpha(i) = \{q \in \mathbb{N} : 1 \le q \le n, h_{i,q} = 1\}$ and $Card(\alpha(i)) = \alpha_i$. Moreover, we set $\alpha(i) = \{i_1, i_2, \ldots, i_{\alpha_i}\}$ with the following order $i_1 \le i_2 \le \cdots \le i_{\alpha_i}$. Then $T^{(m)}_{\mathbf{H},\mathbf{V}}$ can be defined recursively as follows:*

$$\begin{aligned} &T^{(1)}_{\mathbf{H},\mathbf{V}} = \mathbf{V}\,,\\ \textit{and}\quad &T^{(m+1)}_{\mathbf{H},\mathbf{V}} = (T^{(m+1)}_{k,l})_{n\times n}, \quad 1 \le k, l \le n\,. \end{aligned} \tag{7a}$$

Here the block matrices $T^{(m+1)}_{k,l}$ are of following form

$$T^{(m+1)}_{k,l} = v_{k,l}\begin{pmatrix} T^{(m)}_{k_1,l_1} & T^{(m)}_{k_1,l_2} & \cdots & T^{(m)}_{k_1,l_{\alpha_l}} \\ T^{(m)}_{k_2,l_1} & T^{(m)}_{k_2,l_2} & \cdots & T^{(m)}_{k_2,l_{\alpha_l}} \\ \vdots & \vdots & \ddots & \vdots \\ T^{(m)}_{k_{\alpha_k},l_1} & T^{(m)}_{k_{\alpha_k},l_2} & \cdots & T^{(m)}_{k_{\alpha_k},l_{\alpha_l}} \end{pmatrix}, \tag{7b}$$

where $k_i \in \alpha(k)$, $l_i \in \alpha(l)$, $T^{(m+1)}_{k,l}$ and $T^{(m)}_{k_v,l_q}$, $1 \le v \le \alpha_k$ and $1 \le q \le \alpha_i$, are defined as in (6).

We next recall some basic definitions and well-known results from ergodic theory. Let $(X, \mathcal{B}, m)$ be a measure space. Here $\mathcal{B}$ denotes the σ-algebra of all measurable sets in X and m denotes the measure on X. Let $f : X \to X$ be a measurable function. f is said to be *measure preserving* with respect to the measure m if $m(S) = m(f^{-1}(S))$ for all $S \in \mathcal{B}$. Here m is called an *invariant measure* for f.

Definition 2.6. Let f be measure preserving on $(X, \mathcal{B}, m)$. A set $S \in \mathcal{B}$ is called *f-invariant* if $f^{-1}(S) = S$. f is said to be *ergodic* if every f-invariant set has measure 0 or full measure.

We are now ready to state a well-known theorem in ergodic theory.

Theorem 2.4 (Birkhoff Ergodic Theorem (see e.g. [Mane, 1983]). *Let f be measure preserving on $(X, \mathcal{B}, m)$ and g be in $L^1(X)$.*

(1) *There exists an integrable function g^* such that*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(f^k(x)) = g^*(x) \tag{8}$$

for almost every point $x \in X$.

(2) *For all $k \in \mathbb{N}$,*

$$g^*(f^k(x)) = g^*(x) \quad a.e.$$

(3) *If $m(X) = 1$, then*

$$\int_X g dm = \int_X g^* dm\,. \tag{9}$$

Here the left side of (8) is called the *ergodic average*. In the following corollary we see that under the condition of ergodicity, the ergodic average is equal to the "Riemann sum" of $\int_X gdm$.

Theorem 2.5 (see e.g. [Mane, 1983]). *If f is ergodic and $m(X) = 1$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(f^k(x)) = \int_X g dm \quad a.e. \tag{10}$$

The next part of our preliminaries is about the measure theoretic entropy.

Definition 2.7. Let $(X, \mathcal{B}, m)$ be a measure space and $\mathcal{P}$ be a partition of X, the entropy of partition $\mathcal{P}$ is defined to be

$$H(\mathcal{P}) = -\sum_{P\in\mathcal{P}} m(P) \log m(P).$$

Let $f : X \to X$ be measure preserving. The entropy of f with respect to $\mathcal{P}$ is defined by

$$h(f, \mathcal{P}) = \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P})\right). \tag{11}$$

Here the notation $\bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P})$ denotes the partition whose elements are of the form $A_0 \cap \cdots \cap A_{n-1}$ for $A_i \in f^{-j}(\mathcal{P})$, $i = 0, \ldots, n-1$, satisfying $m(A_0 \cap \cdots \cap A_{n-1}) \neq 0$. The measure theoretic entropy of f is then given by

$$h_m(f) = \sup_{\mathcal{P}:\text{ partition}} h(f, \mathcal{P}).$$

Proposition 2.2 (Proposition IV.3.2 of [Mane, 1983]). *The limit in (11) is well defined and exists.*

Let $\mathbf{A}$ be an $n\times n$ *transition matrix.* $\mathbf{P} = (p_{ij}) \in M_{n\times n}(\mathbb{R})$ *is said to be a stochastic matrix associated with* $\mathbf{A}$ *if*

1. $p_{ij} = 0$ *if and only if* $a_{ij} = 0$ *for* $1 \le i, j \le n$.
2. $0 \le p_{ij} \le 1$ *for all* $1 \le i, j \le n$.
3. $\sum_j p_{ij} = 1$.

Clearly, there exists a left eigenvector $\mathbf{q} = (q_1, \ldots, q_n)^T$ satisfying the following:

$$\mathbf{q}^T\mathbf{P} = \mathbf{q}^T, \tag{12a}$$

and

$$\sum_{i=1}^n q_i = 1. \tag{12b}$$

We define a Markov measure $\mu = \mu_{\mathbf{P},\mathbf{q}}$ associated with $(\mathbf{P}, \mathbf{q})$ by

$$\mu(C(i_0, i_1, \ldots, i_k)) = q_{i_0} p_{i_0,i_1} \cdots p_{i_{k-1},i_k}, \tag{13}$$

where $C(i_0, i_1, \ldots, i_k) = \{(j_0, j_1, \ldots) \in \Sigma_{\mathbf{A}} | j_0 = i_0, \ldots, j_k = i_k\}$ is called a *cylinder*.

Proposition 2.3 (see e.g. Theorem I-10.1 of [Mane, 1983]). $\mu = \mu_{\mathbf{P},\mathbf{q}}$ *is an invariant measure of the Markov shift* $\sigma_{\mathbf{A}}$.

Theorem 2.6. *Let* $\mathbf{A}$ *be an* $n \times n$ *transition matrix and* $\mu_{\mathbf{P},\mathbf{q}} = \mu$ *be the invariant Markov measure defined by* $(\mathbf{P}, \mathbf{q})$ *associated with* $\mathbf{A}$. *Then*

(i) (*see e.g. p. 221 of* [*Mane, 1983*]) $h_\mu(\sigma_{\mathbf{A}}) = -\sum_{ij} q_i p_{ij} \log p_{ij}$.
(ii) [*Parry, 1964*] *If* $\sigma_{\mathbf{A}}$ *is topological mixing, then for any invariant measure* μ',

$$h_{\mu'}(\sigma_{\mathbf{A}}) \le \log \lambda_1$$

where λ_1 *is the dominant eigenvalue of* $\mathbf{A}$. *Moreover, there is a unique measure such that the equality attains.*

It has been shown in Theorem 2.1 that $h_{\text{top}}(\sigma_{\mathbf{A}}) = \log \lambda_1$. Theorem 2.6 states that for topological mixing Markov shifts, the topological entropy is the maximal of measure theoretic entropy. This is also true for a general class of maps [Misiurewicz, 1976].

3. Shift Map and Entropy

Let $\mathbf{A}$ be an $n \times n$ transition matrix, and let $\mathcal{P}$ be a set of vectors satisfying the following

$$\mathcal{P} = \left\{\mathbf{x} = (x_1, x_2, \ldots, x_n)^T : \right.$$
$$\left. x_i > 0 \text{ for all } i \text{ and } \sum_{i=1}^n x_i = 1\right\}. \tag{14}$$

Given $\mathbf{x} \in \mathcal{P}$, we set

$$s_i := \frac{(\mathbf{A}\mathbf{x})_i}{x_i},$$

where $(\mathbf{A}\mathbf{x})_i$ is the ith-component of vector $\mathbf{A}\mathbf{x}$. Since $\text{diag}(s_1^{-1}, \ldots, s_n^{-1})\mathbf{A}\mathbf{x} = \mathbf{x}$, there exists a left eigenvector $\mathbf{y}$ satisfying the following.

$$\mathbf{y}^T \text{diag}(s_1^{-1}, \ldots, s_n^{-1})\mathbf{A} = \mathbf{y}^T \tag{15a}$$

and

$$\mathbf{y}^T\mathbf{x} = 1. \tag{15b}$$

We note that if, in addition, $\mathbf{A}$ is symmetric, then

$$\mathbf{y}^T = \frac{\mathbf{x}^T\mathbf{A}}{\mathbf{x}^T\mathbf{A}\mathbf{x}}.$$

Now, if we set

$$\mathbf{P_x} = (\operatorname{diag}\mathbf{Ax})^{-1}\mathbf{A}\operatorname{diag}\mathbf{x}\,, \tag{16a}$$

where $\operatorname{diag}\mathbf{Ax} = \operatorname{diag}((\mathbf{Ax})_1, \ldots, (\mathbf{Ax})_n)$ and $\operatorname{diag}\mathbf{x}$ is also defined similarly, and

$$\mathbf{q}_{\mathbf{x}}^T = \mathbf{y}^T \operatorname{diag}\,\mathbf{x}\left(= \frac{\mathbf{x}^T\mathbf{A}(\operatorname{diag}\,\mathbf{x})}{\mathbf{x}^T\mathbf{Ax}} \text{ if } \mathbf{A} \text{ is symmetric}\right). \tag{16b}$$

Clearly, $\mathbf{P_x}$ is a stochastic matrix associated with $\mathbf{A}$ and $\mathbf{q_x}$ is the left eigenvector of $\mathbf{P_x}$ satisfying (12). We are now ready to state the main result of this section.

Theorem 3.1. *Let $\mathbf{A}$ be an $n \times n$ transition matrix which is irreducible. Let $\mathbf{x} \in \mathcal{P}$, and $\mathbf{P_x}$ and ${\mathbf{q_x}}^T$ are defined as in (16a) and (16b), respectively. Let $\mu_{\mathbf{P_x},\mathbf{q_x}} = \mu_{\mathbf{x}}$ be the Markov measure given as in (13). Then*

(1)

$$h_{\mu_{\mathbf{x}}} = \mathbf{y}^T \log\operatorname{diag}(s_1, s_2, \ldots, s_n)\mathbf{x}\,. \tag{17a}$$

If, in addition, $\mathbf{A}$ is symmetric, then

$$h_{\mu_{\mathbf{x}}} = \frac{\mathbf{x}^T\mathbf{A}\log\operatorname{diag}(s_1, s_2, \ldots, s_n)\mathbf{x}}{\mathbf{x}^T\mathbf{Ax}}\,. \tag{17b}$$

(2)

$$h_{\mu_{\mathbf{x}}} \le \log\lambda \ \textit{for any}\ \mathbf{x} \in \mathcal{P}\,. \tag{18}$$

Here λ is the maximal eigenvalue of $\mathbf{A}$. The equality can be achieved by choosing $\mathbf{x}$ to be the left eigenvector of $\mathbf{A}$ associated to eigenvalue λ with $\sum_{i=1}^n x_i = 1$.

Proof. We first prove (17). Let $\mathbf{P} = (p_{ij})$, and so $(p_{ij}) = (\frac{x_j}{(\mathbf{Ax})_i}a_{ij})$. Set $\tilde{\mathbf{P}} = (p_{ij}\log p_{ij})$, and $\mathbf{e} = (1, \ldots, 1)^T$, it follows from (16a), (16b) and Theorem 2.6(i) that

$$\begin{aligned} h_{\mu_{\mathbf{P},\mathbf{q}}}(\sigma_{\mathbf{A}}) &= -\mathbf{q}^T\tilde{\mathbf{P}}\mathbf{e} \\ &= -\mathbf{y}^T(\operatorname{diag}\,\mathbf{x})\tilde{\mathbf{P}}\mathbf{e}\,. \end{aligned} \tag{19}$$

Now,

$$\begin{aligned} \tilde{\mathbf{P}}\mathbf{e} &= \left(\frac{x_j}{(\mathbf{Ax})_i}a_{ij}\log\left(\frac{x_j}{(\mathbf{Ax})_i}a_{ij}\right)\right)_{n\times n}\mathbf{e} \\ &= (\operatorname{diag}\mathbf{Ax})^{-1}\left(a_{ij}\log\left(\frac{x_j}{(\mathbf{Ax})_i}a_{ij}\right)\right)_{n\times n}\operatorname{diag}(x_1, \ldots, x_n)\mathbf{e} \\ &= (\operatorname{diag}\mathbf{Ax})^{-1}\left(a_{ij}\log\left(\frac{x_j}{(\mathbf{Ax})_i}a_{ij}\right)\right)_{n\times n}\mathbf{x}\,. \end{aligned} \tag{20}$$

Moreover, we have that

$$\begin{aligned} &-\left(a_{ij}\left(\log\frac{x_j}{(\mathbf{Ax})_i}a_{ij}\right)\right)_{n\times n} \\ &\quad= -(a_{ij}\log a_{ij})_{n\times n} + (a_{ij}\log(\mathbf{Ax})_i)_{n\times n} \\ &\qquad -(a_{ij}\log x_j)_{n\times n}\,. \end{aligned} \tag{21}$$

Since either $a_{ij} = 0$ or $a_{ij} = 1$, we see that $a_{ij}\log a_{ij} = 0$. We also note that

$$(a_{ij}\log(\mathbf{Ax})_i)_{n\times n} = \log(\operatorname{diag}\mathbf{Ax})\mathbf{A}$$

and

$$(a_{ij}\log x_j)_{n\times n} = \mathbf{A}\log\operatorname{diag}\mathbf{x}\,.$$

Substituting (21) into (20), we get that

$$\begin{aligned} -\tilde{\mathbf{P}}\mathbf{e} &= (\operatorname{diag}\mathbf{Ax})^{-1}\log(\operatorname{diag}\mathbf{Ax})\mathbf{Ax} \\ &\quad - (\operatorname{diag}\mathbf{Ax})^{-1}\mathbf{A}(\log\operatorname{diag}\mathbf{x})\mathbf{x}\,. \end{aligned} \tag{22}$$

Here $\log\mathbf{A} = (\log a_{ij})$. To further simplify (19), we note that

$$\mathbf{y}^T\operatorname{diag}\mathbf{x}(\operatorname{diag}\mathbf{Ax})^{-1}\mathbf{A} = \mathbf{y}^T \tag{23}$$

and

$$\begin{aligned} &\mathbf{y}^T\operatorname{diag}\mathbf{x}(\operatorname{diag}\mathbf{Ax})^{-1}\log(\operatorname{diag}\mathbf{Ax})\mathbf{Ax} \\ &\quad= \mathbf{y}^T\log(\operatorname{diag}\mathbf{Ax})(\operatorname{diag}\mathbf{x})(\operatorname{diag}\mathbf{Ax})^{-1}\mathbf{Ax} \\ &\quad= \mathbf{y}^T\log(\operatorname{diag}\mathbf{Ax})(\operatorname{diag}\mathbf{x})\mathbf{e} \\ &\quad= \mathbf{y}^T\log(\operatorname{diag}\mathbf{Ax})\mathbf{x}\,. \end{aligned} \tag{24}$$

It then follows from (22)–(24) and (19) becomes

$$\begin{aligned} h_{\mu_{\mathbf{P},\mathbf{q}}}(\sigma_{\mathbf{A}}) &= \mathbf{y}^T\log(\operatorname{diag}\mathbf{Ax})\mathbf{x} - \mathbf{y}^T(\log\operatorname{diag}\mathbf{x})\mathbf{x} \\ &= \mathbf{y}^T\log(\operatorname{diag}(s_1, \ldots, s_n))\mathbf{x}\,. \end{aligned}$$

The inequality in (18) is a direct consequence of Theorem 2.6(ii). A direct calculation would yield the last assertion of the theorem. We thus complete the proof of the theorem. ■

4. Two-Dimensional Golden Mean

In this section, we study the two-dimensional golden mean, that is, the two-dimensional subshifts of finite type with $\mathbf{H} = \mathbf{V} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Recall in Sec. 2 that $\mathbf{T}^{(m)}_{\mathbf{H},\mathbf{V}}$, (resp. $\mathbf{T}^{(m)}_{\mathbf{H},\mathbf{V}}$) represents the m-transition matrices with respect to $\mathbf{H}$ and $\mathbf{V}$ (resp. $\mathbf{V}$ and $\mathbf{H}$). Since $\mathbf{H} = \mathbf{V}$, we see that $\mathbf{T}^{(m)}_{\mathbf{V},\mathbf{H}} = \mathbf{T}^{(m)}_{\mathbf{H},\mathbf{V}} (:= \mathbf{T}^{(m)})$. Applying Theorem 2.3, $\mathbf{T}^{(m)}$ can be written recursively as follows:

$$\mathbf{T}^{(1)} = \left(\begin{array}{c|c} 1 & 1 \\ \hline 1 & 0 \end{array}\right) = \left(\begin{array}{c|c} \mathbf{T}^{(0)}_{1,1} & \mathbf{T}^{(0)}_{1,2} \\ \hline \mathbf{T}^{(0)}_{2,1} & \mathbf{T}^{(0)}_{2,2} \end{array}\right)$$

and

$$\mathbf{T}^{(m+1)} = \left(\begin{array}{cc|c} \mathbf{T}^{(m)}_{1,1} & \mathbf{T}^{(m)}_{1,2} & \mathbf{T}^{(m)}_{1,1} \\ \mathbf{T}^{(m)}_{2,1} & 0 & \mathbf{T}^{(m)}_{2,1} \\ \hline \mathbf{T}^{(m)}_{1,1} & \mathbf{T}^{(m)}_{1,2} & 0 \end{array}\right). \tag{25}$$

Let a_n be the size of $\mathbf{T}^{(n)}$, then a_n satisfies the following recursive formula.

$$a_{n+1} = a_n + a_{n-1} \tag{26a}$$

and

$$a_1 = 2, \quad a_2 = 3. \tag{26b}$$

Proposition 4.1. *For each $n \geq 1$, $\mathbf{T}^{(n)}$ is symmetric and irreducible.*

Proof. It is easy to see that $\mathbf{T}^{(n)}$ is symmetric for all $n \geq 1$. We next prove that each $\mathbf{T}^{(n)}$ is eventually positive. Since $\mathbf{T}^{(1)}$ and $\mathbf{T}^{(2)}$ are eventually positive, we assume that $\mathbf{T}^{(n-1)}$ and $\mathbf{T}^{(n)}$ are eventually positive for some n. Then there exists $m > 0$ such that

$$(\mathbf{T}^{(n)})^m > \mathbf{E}_n \quad \text{and} \quad (\mathbf{T}^{(n-1)})^m > \mathbf{E}_{n-1}.$$

Here $\mathbf{E}_n = (1)_{a_n \times a_n}$. We observe in (25) that $\mathbf{T}^{(n+1)}_{11} = \mathbf{T}^{(n)}$, thus the matrix multiplication gives

$$(\mathbf{T}^{(n+1)})^{m+1} > \left(\begin{array}{c|c} \mathbf{E}_n & \mathbf{E}_n \begin{pmatrix} \mathbf{T}^{(n)}_{11} \\ \mathbf{T}^{(n)}_{21} \end{pmatrix} \\ \hline \begin{pmatrix} \mathbf{T}^{(n)}_{11} & \mathbf{T}^{(n)}_{12} \end{pmatrix} \mathbf{E}_n & \mathbf{E}_{n-1} \end{array}\right)$$
$$> 0.$$

An inductive argument then leads to the assertion of the proposition. ■

Letting $\mathbf{e}_n = (1, \ldots, 1)^T \in \mathbb{R}^{a_n}$, then $h_{\mu_{\mathbf{e}_n}}$ as defined in (17b) becomes

$$h_{\mu_{\mathbf{e}_n}} = \lambda_n = \frac{\mathbf{e}_n^T \mathbf{T}^{(n)} \log(\operatorname{diag} \mathbf{e}_n^T \mathbf{T}^{(n)}) \mathbf{e}_n}{\mathbf{e}_n^T \mathbf{T}^{(n)} \mathbf{e}_n}. \tag{27a}$$

Moreover, if we let $\mathbf{e}_n^T \mathbf{T}^{(n)} =: \mathbf{v}^{(n)} = (v_i^{(n)}) \in \mathbb{R}^{a_n}$ and $s_n = \sum_{i,j=1}^{a_n} (\mathbf{T}^{(n)})_{ij}$ be the sum taken over all entries of $\mathbf{T}^{(n)}$. Then

$$\lambda_n = \frac{\sum_{i=1}^{a_n} v_i^{(n)} \log v_i^{(n)}}{s_n}. \tag{27b}$$

We remark that s_n satisfy the following recursive formulas:

$$s_{n+1} = 2s_n + s_{n-1} \tag{28a}$$

and

$$s_1 = 3, \quad s_2 = 7. \tag{28b}$$

Applying Theorem 3.1 and (27), and Proposition 2.1 we obtain the following lower bound for $h(\Sigma_{\mathbf{H},\mathbf{V}})$ of two-dimensional golden mean.

Theorem 4.1

$$h(\Sigma_{\mathbf{H},\mathbf{V}}) \geq \limsup_{n \to \infty} \frac{\sum_{i=1}^{a_n} v_i^{(n)} \log v_i^{(n)}}{n s_n}. \tag{29}$$

The remainder of the section is to compute the sum of the infinite series as given in (29).

We first observe that

$$\begin{array}{ll|l|ll|l|ll|l|llll}
\mathbf{v}^{(1)} = (2, & 1) & & & & & & & & & & & \\
\mathbf{v}^{(2)} = (3, & 2, & 2) & & & & & & & & & & \\
\mathbf{v}^{(3)} = (5, & 3, & 4, & 3, & 2) & & & & & & & & \\
\mathbf{v}^{(4)} = (8, & 5, & 6, & 6, & 4, & 5, & 3, & 4) & & & & & \\
\mathbf{v}^{(5)} = (13, & 8, & 10, & 9, & 6, & 10, & 6, & 8, & 8, & 5, & 6, & 6, & 4).
\end{array}$$

To derive a recursive formula for $\mathbf{v}^{(n)}$, we first write $\mathbf{v}^{(n)}$ as

$$\mathbf{v}^{(n)} = (\mathbf{u}^{(n+1)}, \mathbf{u}^{(n)}) .$$

Here $\mathbf{u}^{(n+1)}$ and $\mathbf{u}^{(n)}$ are row vectors whose dimensions are $1 \times a_{n-1}$ and $1 \times a_{n-2}$, respectively. For instance, $\mathbf{v}^{(4)} = (\mathbf{u}^{(5)}, \mathbf{u}^{(4)})$, where $\mathbf{u}^{(5)} = (8, 5, 6, 6, 4)$ and $\mathbf{u}^{(4)} = (5, 3, 4)$. Clearly, $\mathbf{u}^{(n+1)}$ can be recursively defined as

$$\begin{aligned}\mathbf{u}^{(n+1)} &= (\mathbf{u}^{(n)} + \mathbf{v}^{(n-2)}, 2\mathbf{u}^{(n-1)}) \\ &= (\mathbf{u}^{(n)} + (\mathbf{u}^{(n-1)}, \mathbf{u}^{(n-2)}), 2\mathbf{u}^{(n-1)}) \qquad (30)\end{aligned}$$

with

$$\mathbf{u}^{(1)} = 1 , \quad \mathbf{u}^{(2)} = 2 , \quad \mathbf{u}^{(3)} = (3, 2) .$$

For example,

$$\begin{aligned}\mathbf{u}^{(6)} &= (\mathbf{u}^{(5)} + (\mathbf{u}^{(4)}, \mathbf{u}^{(3)}), 2\mathbf{u}^{(4)}) \\ &= ((8, 5, 6, 6, 4) + (5, 3, 4, 3, 2), 2(5, 3, 4)) \\ &= (13, 8, 10, 9, 6, 10, 6, 8) .\end{aligned}$$

We are ready to state the following useful proposition.

Proposition 4.2.

$$\mathbf{u}^{(n)} = (a_{n-1}, a_{n-2}\mathbf{u}^{(1)}, \ldots, a_{n-i-1}\mathbf{u}^{(i)}, \ldots, a_1\mathbf{u}^{(n-2)}) . \qquad (31)$$

Proof. Let $n = 3$, we see that 31 is clearly satisfied. Suppose 31 is true for $k = 3, \ldots, n$. Then

$$\begin{aligned}\mathbf{u}^{(n)} + (\mathbf{u}^{(n-1)}, \mathbf{u}^{(n-2)}) &= (a_{n-1}, a_{n-2}\mathbf{u}^{(1)}, \ldots, a_{n-i-1}\mathbf{u}^{(i)}, \ldots, a_2\mathbf{u}^{(n-3)}, a_1\mathbf{u}^{(n-2)}) \\ &\quad + (a_{n-2}, a_{n-3}\mathbf{u}^{(1)}, \ldots, a_{n-i-2}\mathbf{u}^{(i)}, \ldots, a_1\mathbf{u}^{(n-3)}, \mathbf{u}^{(n-2)}) \\ &= (a_n, a_{n-1}\mathbf{u}^{(1)}, \ldots, a_{n-i}\mathbf{u}^{(i)}, \ldots, a_3\mathbf{u}^{(n-3)}, (1 + a_1)\mathbf{u}^{(n-2)}) \\ &= (a_n, a_{n-1}\mathbf{u}^{(1)}, \ldots, a_{n-i}\mathbf{u}^{(i)}, \ldots, a_3\mathbf{u}^{(n-3)}, a_2\mathbf{u}^{(n-2)}) .\end{aligned}$$

Thus,

$$\begin{aligned}\mathbf{u}^{(n+1)} &= (\mathbf{u}^{(n)} + (\mathbf{u}^{(n-1)}, \mathbf{u}^{(n-2)}), 2\mathbf{u}^{(n-1)}) \\ &= (a_{n-1}, a_{n-2}\mathbf{u}^{(1)}, \ldots, a_{n-i-1}\mathbf{u}^{(i)}, \ldots, a_2\mathbf{u}^{(n-2)}, a_1\mathbf{u}^{(n-1)}) . \qquad \blacksquare\end{aligned}$$

To investigate $\sum_i v_i^{(n)} \log v_i^{(n)}$, we define $L : \mathbb{R}^N \to \mathbb{R}$, as

$$L(\mathbf{x}) = \sum_{i=1}^{N} x_i \log x_i ,$$

where $\mathbf{x} = (x_1, \ldots, x_N)^T$. Clearly, for any $c \in \mathbb{R}$, we have that

$$L(c\mathbf{x}) = \left(\sum_{i=1}^{N} x_i\right) c \log c + cL(\mathbf{x}) . \qquad (32)$$

Let $\mathbf{u}^{(n)} = (u_1^{(n)}, u_2^{(n)}, \ldots, u_{a_{n-1}}^{(n)})$. We set

$$\alpha_n = \sum_{i=1}^{a_{n-1}} u_i^{(n)} , \qquad (33a)$$

$$\beta_n = a_n \log a_n , \qquad (33b)$$

$$L_n = L(\mathbf{u}^{(n)}) , \qquad (33c)$$

$$p_n = \sum_{i=1}^{n-2} a_{n-i-1} L_i , \qquad (33d)$$

$$q_n = \beta_{n-1} + \sum_{i=1}^{n-2} \beta_{n-i-1}\alpha_i . \qquad (33e)$$

Applying (32) and (33a), we have that

$$\begin{aligned}L(\mathbf{u}^{(n)}) &= a_{n-1} \log a_{n-1} + \sum_{i=1}^{n-2} L(a_{n-i-1}\mathbf{u}^{(i)}) \\ &= p_n + q_n . \qquad (34)\end{aligned}$$

Proposition 4.3. *α_n, p_n and q_n satisfy the following recursive formulas*

(i) $\alpha_{n+1} = 2\alpha_n + \alpha_{n-1} , \alpha_1 = 1 , \alpha_2 = 2 ,$ (35a)

(ii) $p_{n+1} = p_n + 3p_{n-1} + p_{n-2} + 2q_{n-1} + q_{n-2} ,$ (35b)

(iii) $q_{n+1} = 2q_n + q_{n-1} + (\beta_n - \beta_{n-1} - \beta_{n-2}) .$ (35c)

Proof. Clearly, $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_3 = 5$. Hence $\alpha_3 = 2\alpha_2 + \alpha_1$. Assume that $\alpha_{k+1} = 2\alpha_k + \alpha_{k-1}$ holds for $2 \le k \le n-1$. Then

$$\alpha_{n+1} = \sum_{i=1}^{a_n} u_i^{(n)} = \alpha_n + 3\alpha_{n-1} + \alpha_{n-2}\,. \quad (36)$$

We have used (30) to justify the last equality in (36). Using the inductive hypothesis, we get that $\alpha_{n+1} = 2\alpha_n + \alpha_{n-1}$. The proof of the first assertion of the proposition is thus complete. To prove (ii), we see that

$$\begin{aligned}
p_{n+1} - p_n &= (a_{n-1}L_1 + \ldots + a_3L_{n-3} + a_2L_{n-2} + a_1L_{n-1}) \\
&\quad -(a_{n-2}L_1 + \ldots + a_2L_{n-3} + a_1L_{n-2}) \\
&= a_{n-3}L_1 + \ldots + a_1L_{n-3} + L_{n-2} + 2L_{n-1} \\
&= p_{n-1} + L_{n-2} + 2L_{n-1} \\
&= p_{n-1} + p_{n-2} + q_{n-2} + 2(p_{n-1} + q_{n-1}) \\
&= 3p_{n-1} + p_{n-2} + 2q_{n-1} + q_{n-2}\,.
\end{aligned}$$

This proves (ii). To prove (iii), we note that

$$\begin{aligned}
2q_n + q_{n-1} &= 2\beta_{n-1} + 2\alpha_1\beta_{n-2} + 2\alpha_2\beta_{n-3} + \ldots \\
&\quad + 2\alpha_{n-2}\beta_1 + \beta_{n-2} + \alpha_1\beta_{n-3} + \ldots \\
&\quad + \alpha_{n-3}\beta_1 \\
&= 2\beta_{n-1} + 3\beta_{n-2} + \alpha_3\beta_{n-3} + \ldots \\
&\quad + \alpha_{n-1}\beta_1\,,
\end{aligned}$$

and

$$\begin{aligned}
q_{n+1} &= \beta_n + \beta_{n-1} + 2\beta_{n-2} + \alpha_3\beta_{n-3} + \ldots \\
&\quad + \alpha_{n-1}\beta_1\,.
\end{aligned}$$

Hence,

$$q_{n+1} - 2q_n - q_{n-1} = (\beta_n - \beta_{n-1} - \beta_{n-2})$$

as asserted. ■

Proposition 4.4. *Let λ_n be the quantity as given in (27b), then*

$$\lambda_n = \frac{L(\mathbf{v}^{(n)})}{s_n} = \frac{(p_{n+1} + p_n) + (q_{n+1} + q_n)}{s_n}\,.$$

Proof. It follows directly from Theorem (17b) and (34). ■

To evaluate $\limsup_{n\to\infty} \frac{(p_{n+1}+p_n)+(q_{n+1}+q_n)}{ns_n}$, we need the following proposition:

Proposition 4.5. *Let $\lambda = 1 + \sqrt{2}$. Then the following holds.*

(i) *λ, $-\frac{1}{\lambda}$ and -1 are the characteristic roots of γ_n. Here $\gamma_{n+1} = \gamma_n + 3\gamma_{n-1} + \gamma_{n-2}$ with $\gamma_2 = \gamma_1 = 1$ and $\gamma_2 = 0$.*

(ii) *There are constants c_s, d_s, c_α, d_α, c_γ, d_γ, e_γ for which*

$$s_n = c_s\lambda^n + d_s\left(-\frac{1}{\lambda}\right)^n,$$

$$\alpha_n = c_\alpha\lambda^n + d_\alpha\left(-\frac{1}{\lambda}\right)^n,$$

and

$$\gamma_n = c_\gamma\lambda^n + d_\gamma\left(-\frac{1}{\lambda}\right)^n + e_\gamma(-1)^n\,.$$

Here s_n and α_n are defined in (28) and (33), respectively.

Proposition 4.6. *The following limit exists.*

$$\lim_{n\to\infty} \frac{q_n}{\lambda^n} = c_\alpha\left(\frac{q_4}{\lambda^3} + \frac{q_3}{\lambda^4} + \sum_{i=4}^{\infty} \frac{k_i}{\lambda^i}\right) =: q^*$$

where $k_i = \beta_i - \beta_{i-1} - \beta_{i-2}$, and c_α is defined as in Proposition 4.5 (ii).

Proof. Let $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$. Using (35c), we get that

$$\begin{pmatrix} q_{n+1} \\ q_n \end{pmatrix} = \mathbf{A}\begin{pmatrix} q_n \\ q_{n-1} \end{pmatrix} + \begin{pmatrix} k_n \\ 0 \end{pmatrix}$$

with initial conditions $\begin{pmatrix} q_4 \\ q_3 \end{pmatrix}$. Note that

$$\mathbf{A}^n = \begin{pmatrix} \alpha_{n+1} & \alpha_n \\ \alpha_n & \alpha_{n-1} \end{pmatrix},$$

where we assume $\alpha_0 = 0$. Using the variation of constant formula, we then obtain that

$$\begin{pmatrix} q_n \\ q_{n-1} \end{pmatrix} = \mathbf{A}^{n-4}\begin{pmatrix} q_4 \\ q_3 \end{pmatrix} + \sum_{i=4}^{n-1} \mathbf{A}^{n-1-i}\begin{pmatrix} k_i \\ 0 \end{pmatrix}.$$

Hence,

$$q_n = (\alpha_{n-3}q_4 + \alpha_{n-4}q_3) + \sum_{i=4}^{n-1} \alpha_{n-i}k_i\,.$$

Applying the ratio tests, we conclude that

$$\sum_{i=4}^{\infty} \frac{k_i}{\lambda^i}$$

converges. The proof of the proposition is thus complete. ■

Remark 4.1. We note that $k_i > 0$ for all $i \geq 4$. Hence the partial sum $q_n^* := c_\alpha \left(\frac{q_4}{\lambda^3} + \frac{q_3}{\lambda^4} + \sum_{i=4}^{n} \frac{k_i}{\lambda^i} \right)$ converges upward to q^*.

Corollary 4.1

(i) *Given* $\varepsilon > 0$, *there exists* $N \in \mathbb{N}$ *such that*

$$(q^* - \varepsilon)\lambda^n < q_n < (q^* + \varepsilon)\lambda^n \tag{37}$$

whenever $n \geq N$.

(ii) $\lim_{n\to\infty} \frac{q_n}{ns_n} = 0$ *and, hence,*

$$\lim_{n\to\infty} \frac{L(\mathbf{v}^{(n)})}{ns_n} = \lim_{n\to\infty} \frac{p_{n+1} + p_n}{ns_n}\,.$$

Proposition 4.7.

$$\lim_{n\to\infty} \frac{p_n}{n\lambda^n} = q^* c_\gamma (2\lambda^{-1} + \lambda^{-2})\,,$$

where c_λ *is defined as in Proposition (4.5) (ii).*

Proof. Let $g_n = 2q_{n-1} + q_{n-2}$ and $\mathbf{B} = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. We see via the induction that $\mathbf{B}^n = \begin{pmatrix} \gamma_{n+1} & \cdot & \cdot \\ \gamma_n & \cdot & \cdot \\ \gamma_{n-1} & \cdot & \cdot \end{pmatrix}$. Then

$$\begin{pmatrix} p_{n+1} \\ p_n \\ p_{n-1} \end{pmatrix} = \mathbf{B} \begin{pmatrix} p_n \\ p_{n-1} \\ p_{n-2} \end{pmatrix} + \begin{pmatrix} g_n \\ 0 \\ 0 \end{pmatrix}.$$

Now we let $\varepsilon > 0$ be fixed and $N = N(\varepsilon) > 0$ be given as in Corollary , then for $n > N$,

$$\begin{pmatrix} p_n \\ p_{n-1} \\ p_{n-2} \end{pmatrix} = \mathbf{B}^{n-N} \begin{pmatrix} p_N \\ p_{N-1} \\ p_{N-2} \end{pmatrix} + \sum_{i=N}^{n-1} \begin{pmatrix} \gamma_{n+1} & \cdot & \cdot \\ \gamma_n & \cdot & \cdot \\ \gamma_{n-1} & \cdot & \cdot \end{pmatrix} g_i\,, \tag{38}$$

where we set $\gamma_{-1} = 0$. Using (38), we obtain that

$$p_n = O(\lambda^n) + c_\gamma \sum_{i=N}^{n-1} \lambda^{n-i} g_i\,.$$

It follows from (37) and (38) that, we have

$$\begin{aligned} &c_\gamma \sum_{i=N}^{n-1} \lambda^{n-i}(q^* - \varepsilon)(2\lambda^{-1} + \lambda^{-2})\lambda^i + O(\lambda^n) \\ &\leq p_n \leq c_\gamma \sum_{i=N}^{n-1} \lambda^{n-i}(q^* + \varepsilon)(2\lambda^{-1} + \lambda^{-2})\lambda^i \\ &\quad + O(\lambda^n)\,. \end{aligned}$$

Hence,

$$\begin{aligned} &(q^* - \varepsilon)c_\gamma(2\lambda^{-1} + \lambda^{-2}) \\ &\leq \lim_{n\to\infty} \frac{p_n}{n\lambda^n} \leq (q^* + \varepsilon)c_\gamma(2\lambda^{-1} + \lambda^{-2}) \end{aligned}$$

Since ε is arbitrary, the assertion of the proposition holds. ■

We are ready to state the main result of this paper.

Theorem 4.1. *Let* $\mathbf{H} = \mathbf{V} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, *then*

$$\begin{aligned} h(\Sigma_{\mathbf{H},\mathbf{V}}) &\geq q^* c_\gamma (2\lambda^{-1} + \lambda^{-2}) \frac{1}{c_s} (\lambda + 1) \\ &\geq q_n^* c_\gamma (2\lambda^{-1} + \lambda^{-2}) \frac{1}{c_s} (\lambda + 1) =: h_n\,. \end{aligned}$$

The second inequality holds for all $n \geq 4$.

We remark that the known lower and upper bounds of $h(\Sigma_{\mathbf{H},\mathbf{V}})$, where $\mathbf{H} = \mathbf{V} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, are $\log \frac{1+\sqrt{5}}{2} (\approx 0.481212)$ and $\frac{1}{2} \log 2 (\approx 0.346574)$, respectively. Our estimate in (39) gives

$$h(\Sigma_{\mathbf{H},\mathbf{V}}) \geq h_{5000} \approx 0.404089\,.$$

References

Baxter, R. J. [1982] *Exactly Solved Models in Statistical Mechanics* (Academic Press, NY–London).

Bowen, R. [1971] "Entropy of group endomorphism and homogeneous spaces," *Trans. Amer. Math. Soc.* **14**, 23–42.

Bowen, R. [1988] "On axion a diffeomorphism (conference board math. science)," *Trans. Amer. Math. Soc. Circuits Syst.* **35**, 1273–1290.

Chow, S. N., Mallet-Paret, J. & Van Vleck, E. S. [1996] "Pattern formation and spatial chaos in spatially discrete evolution equations," *Rand. Comp. Dyn.* **4**, 109–178.

Chua, L. O. & Yang, L. [1988a] "Cellular neural networks: Applications," *IEEE Trans. Circuits Syst.* **35**, 1272–1273.

Chua, L. O. & Yang, L. [1988b] "Cellular neural networks: Theory," *IEEE Trans. Circuits Syst.* **35**, 1257–1290.

Juang, J. & Lin, S. S. [2000] "Cellular neural networks: Mosaic patterns and spatial chaos," *SIAM J. Appl. Math.* **60**, 891–915.

Juang, J., Lin, S. S. & Shieh, S. F. [2000] "The spatial entropy of two dimensional Markov shifts of finite type," *Int. J. Bifurcation and Chaos* **10**, 2129–2140.

Juang, J., Shieh, S. F. & Turyn. L. [2002] "Cellular neural networks: Space-dependent template, mosaic patterns and spatial chaos," *Int. J. Bifurcation and Chaos* **12**, 1717–1730.

Kasteleyn, P. W. [1961] "The statistics of dimers on lattice," *I. Physica* **27**, 1209–1225.

Lieb, E. H. [1967] "Residual entropy of square ice," *Phys. Rev.* **162**, 162–172.

Mane, R. *Ergodic Theory and Differential Dynamics* (Spring-Verlag).

Markely, N. G. & Paul, M. E. [1981a] "Matrix subshifts for $\mathbf{Z}^{\nu}$ symbolic dynamics," *Proc. London Math. Soc.* **43**, 251–272.

Markely, N. G. & Paul, M. E. [1981b] "Maximal measures and entropy for $\mathbb{Z}^{\nu}$ subshifts of finite type," *Collection: Classical Mechanics and Dynamical Systems*, Medford, MA, 1979 (Marcel Dekker, NY), pp. 136–157.

Misiurewicz, M. [1976] "A short proof of the variational principle for a $\mathbf{Z}_{+}^{N}$ action on a compact space," *Astérisque* **40**, 147–187.

Parry, J. [1964] "Intrinsic Markov chains," *Trans. Amer. Math. Soc.* **112**, 55–66.

Robinson, C. [1993] *Dynamical Systems, Stability, Symbolic Dynamics, and Chaos* (CRC Press, London).

Schmidt, K. [1990] *Algebraic Ideas in Ergodic Theory*, Conference Board Math. Science. Trans. Amer. Math. Soc. Providence, RI.

Temperley, H. N. V. & Lieb, E. H. [1971] "Relation between the 'percolation' and 'colouring' problem and other graph-theoretical problems assoicated with regular planner lattices: Some exact results for the 'percolation' problem," *Proc. Roy. Soc. London Series* **A322**, 251–280.