Counting preimages

MICHAŁ MISIUREWICZ† and ANA RODRIGUES‡

† *Department of Mathematical Sciences, IUPUI, 402 N. Blackford Street,*
Indianapolis, IN 46202, USA
(e-mail: mmisiure@math.iupui.edu)
‡ *Department of Mathematics, University of Exeter, Harrison Building,*
Streatham Campus, North Park Road, Exeter, EX4 4QF, UK
(e-mail: A.Rodrigues@exeter.ac.uk)

Abstract. For non-invertible maps, subshifts that are mainly of finite type and piecewise monotone interval maps, we investigate what happens if we follow backward trajectories, which are random in the sense that, at each step, every preimage can be chosen with equal probability. In particular, we ask what happens if we try to compute the entropy this way. It turns out that, instead of the topological entropy, we get the metric entropy of a special measure, which we call the *fair measure*. In general, this entropy (the *fair entropy*) is smaller than the topological entropy. In such a way, for the systems that we consider, we get a new natural measure and a new invariant of topological conjugacy.

1. *Introduction, motivation and questions*

When working on topological entropy of one-dimensional dynamical systems, one would like to have some simple method for computing it. The simplest class of such systems is the class of piecewise strictly monotone interval maps. Various computational methods are known (see, e.g., [**5**–**7**]), but none of them is really general and simple. Even for the Markov maps, before starting computations (which are relatively simple, because one computes only the spectral radius of a non-negative matrix (see, e.g., [**1**]) one has to identify the Markov structure and find the transition matrix. This structure may be complicated and the matrix can be large.

One of the obvious ideas for computing topological entropy is to count preimages of a given point.

Definition 1.1. The class $\mathcal{PMM}$ (piecewise monotone mixing) consists of all continuous maps $f : [0, 1] \to [0, 1]$ that are piecewise monotone (with a finite number of pieces) and topologically mixing.

Clearly, if a map f is in $\mathcal{PMM}$, then it is piecewise strictly monotone. It is also known (see, e.g., [**4**]) that it is *locally eventually onto*: that is, for every non-empty open set U there is n such that $f^n(U) = [0, 1]$.

We can compute the topological entropy $h(f)$ of $f \in \mathcal{PMM}$ by counting preimages of any point. This follows immediately from the well known properties of interval maps: characterization of the entropy via the number of laps of iterates and the theorem on entropy and horseshoes.

THEOREM 1.2. *If $f \in \mathcal{PMM}$ and $x \in [0, 1]$ then*

$$h(f) = \lim_{n \to \infty} \frac{1}{n} \log \operatorname{Card}(f^{-n}(x)). \tag{1}$$

While this is an interesting theoretical result, it is usually useless if we want to compute topological entropy using a computer. The number of preimages of x under f^n grows exponentially with n, and keeping track of all of them requires a lot of memory, while computing their preimages (when we pass from n to $n + 1$) requires a lot of time.

In some similar situations, the problem is solved by replacing the full tree of preimages by one branch, chosen randomly. This type of procedure is used, for instance, to draw an attractor for an iterated function system (see, e.g., [**3**]) or the Julia set for a rational map of the Riemann sphere to itself (see, e.g., [**8**]). In fact, in the latter case, the measures equidistributed along longer and longer pieces of a random backward trajectory converge in the weak-* topology to the measure with maximal entropy (see [**8**]).

Thus, our procedure will be the following. Denote by $c(x)$ the cardinality of the set $f^{-1}(x)$. We start with a point x_0 and proceed by induction. Given x_n, we choose x_{n+1} from the set $f^{-1}(x_n)$ randomly: that is, the probability of choosing any of those points is $1/c(x_n)$. Then we go to the limit with the geometric averages of $c(x_0), c(x_1), \ldots, c(x_n)$ as n goes to infinity. We hope that the limit exists and its logarithm is the topological entropy of f. Moreover, we hope that the measures equidistributed along longer and longer pieces of a random backward trajectory converge in the weak-* topology to the measure with maximal entropy.

Alas, a simple computer experiment where f is a unimodal map with the turning point periodic of period three suggests that the limit of geometric averages exists, but is approximately 1.5874 instead of the expected golden ratio (that is, approximately 1.618). This immediately modifies our expectations, by removing reference to the topological entropy from them. However, we want to know what is going on.

Question 1.3. What can we say about the convergence of the geometric averages of $c(x_0), c(x_1), \ldots, c(x_n)$ as n goes to infinity for a random choice of the backward trajectory? Does it exist for every initial point x_0? Or for almost every initial point in some sense? How much does the limit depend on the initial point?

Question 1.4. What can we say about the convergence of the measures equidistributed along longer and longer pieces of a random backward trajectory? Does it exist for every initial point x_0? Or for almost every initial point in some sense? How much does the limit depend on the initial point?

Question 1.5. What is the connection between the limits mentioned in the two preceding questions? In particular, is the logarithm of the first limit the entropy of the measure that is the second limit?

Of course, the questions can be asked not only for interval maps. We will give full answers for transitive subshifts of finite type and almost full answers for interval maps from the class $\mathcal{PMM}$. Let us also mention that the questions are meaningful only for non-invertible systems.

In general, we expect to get a measure and its entropy, and we hope they will be unique. Then we can speak of the *fair measure* and the *fair entropy*. Clearly, if two systems are conjugate, the conjugacy will carry the fair measure to the fair measure so, in particular, the fair entropies of those systems will be equal. In such a way, we get an invariant of topological conjugacy.

Let us note a big difference between the answers to our questions that apply to *every* point and to *almost every* point. We really aim at the 'every point' theorems. 'Almost every' would mean practically 'almost every with respect to the fair measure'. However, we do not know *a priori* the fair measure so, if the system has some natural measure, this may mean 'almost none with respect to the natural measure'. Fortunately, in the cases that we consider, even if we cannot get the 'every' results, we will be able to get something more than 'almost every'.

The paper is organized as follows. In §2, we provide the definitions and describe what can be done without too many conditions on the dynamical system. This is continued in §3, where we consider the natural extension of the system. In §§4 and 5, we investigate more closely the special classes of maps: that is, subshifts of finite type and piecewise monotone topologically mixing interval maps, respectively. We obtain, for these classes, stronger results than in the general case. In §6, we show that the strongest results that we obtained for subshifts of finite type apply also for the mixing Markov interval maps. Finally, in §7, we find necessary and sufficient conditions for the fair entropy to be equal to the topological entropy for subshifts of finite type.

2. *General case*

While, in the preceding section, we were talking mainly about single backward trajectories, when we want to prove something about them we need to have information about whole trees of preimages. Therefore we will be working in the following set-up.

Let X be a compact metric space and let $f : X \to X$ be a continuous map. We assume that f is a surjection and that there exists a partition $\mathcal{X}$ of X into Borel sets X_i, $i = 1, 2, \ldots s$ such that, for every i, the restriction of f to X_i is a homeomorphism onto its image (which is also a Borel set).

We define a partition $\mathcal{A}$ as the common refinement of the partitions $\{f(X_i), X \setminus f(X_i)\}$, $i = 1, 2, \ldots, s$. Clearly, $\mathcal{A}$ is a finite partition of X into Borel sets. For each $A \in \mathcal{A}$, denote by $p(A)$ the set of those numbers $i \in \{1, 2, \ldots, s\}$ for which $A \subset f(X_i)$. Thus, each $x \in A$ has a preimage in every X_i such that $i \in p(A)$ and no more preimages. In particular, the number $c(x)$ of preimages of x depends only on A (and therefore can be denoted by $c(A)$). Since f is a surjection, this number is always positive.

Let $\mathfrak{M}$ be the space of all probability Borel measures on X. We will define an operator Φ from $\mathfrak{M}$ to itself. If $\mu \in \mathfrak{M}$, then we chop μ into pieces $\mu|_A$, $A \in \mathcal{A}$, divide each piece by $c(A)$ and push via $(f|_{X_i})^{-1}$ to each X_i with $i \in p(A)$. That is, we set

$$\Phi(\mu) = \sum_{A \in \mathcal{A}} \sum_{i \in p(A)} (f|_{X_i})^{-1}_* \left(\frac{\mu|_A}{c(A)} \right).$$

In other words, if $B \subset X$ is a Borel set then

$$\Phi(\mu)(B) = \sum_{A \in \mathcal{A}} \sum_{i \in p(A)} \frac{\mu(f(X_i \cap B) \cap A)}{c(A)}$$

$$= \sum_{A \in \mathcal{A}} \sum_{i=1}^{s} \frac{\mu(f(X_i \cap B) \cap A)}{c(A)} = \sum_{i=1}^{s} \sum_{A \in \mathcal{A}} \frac{\mu(f(X_i \cap B) \cap A)}{c(A)}.$$

In particular, if $B \subset X_i$ and $f(B) \subset A \in \mathcal{A}$, then $\mu(f(B)) = \Phi(\mu)(B) \cdot c(A)$.

Observe that if we replace the partition $\mathcal{A}$ by a finer one, the operator Φ will not change. Therefore this operator does not depend on a particular choice of the partition $\mathcal{X}$, as long as this partition satisfies the assumptions we made. Indeed, for two different partitions, we can take their common refinement, and that will lead to the common refinement of the corresponding partitions $\mathcal{A}$.

The operator Φ is a partial inverse of the operator f_* that carries the measure forward. Namely, it follows immediately from the definition of Φ that

$$f_* \circ \Phi = \mathrm{id}. \tag{2}$$

We will be looking for fixed points of Φ. If μ is a fixed point of Φ, then, for a set B contained in an element A of $\mathcal{A}$, its preimage consists of $c(A)$ sets B_j, each of them contained in one of the sets X_j, and $\mu(B_j) = \mu(B)/c(A)$ for every j. Thus, if we go backward, the chance of choosing any of the js is the same. In other words, our procedure for choosing a preimage is fair. This motivates the following name.

Definition 2.1. A measure $\mu \in \mathfrak{M}$ will be called a *fair measure* if $\Phi(\mu) = \mu$.

From the definition of Φ and the definition of the measure-theoretic Jacobian (see, e.g., [13]), we immediately get the following lemma.

LEMMA 2.2. *A measure $\mu \in \mathfrak{M}$ is fair if and only if its measure-theoretic Jacobian is $x \mapsto c(f(x))$.*

Another point of view on the Jacobian is to speak about the *g-measures* (see, e.g., [10, 11, 14]). The condition that the Jacobian of μ is $x \mapsto c(f(x))$ is equivalent to μ being a g-measure, where g is the reciprocal of the Jacobian. Therefore another way of stating Lemma 2.2 is as follows.

LEMMA 2.3. *A measure $\mu \in \mathfrak{M}$ is fair if and only if it is a g-measure for the function g given by $g(x) = 1/c(f(x))$.*

Continuing the translation into the language of g-measures, our operator Φ is dual to the Ruelle operator for the function $\xi = \log g$, that is, the function given by $\xi(x) = -\log c(f(x))$. However, for interval maps, this operator is usually not continuous.

Moreover, the function ξ usually does not satisfy the inequality $\sup \xi - \inf \xi < h_{\text{top}}(f)$, which is often assumed when the properties of g-measures are considered (see, e.g., [**9**]).

We will later need the following lemma.

LEMMA 2.4. *If a measure $\mu \in \mathfrak{M}$ is fair, so is its almost every ergodic component.*

Proof. Let $\mu \in \mathfrak{M}$ be fair. Ergodic components of μ are conditional measures μ_y for a measurable partition η into invariant subsets, and there is the corresponding measure $\varkappa$ on the factor space Y such that, for every measurable set $B \subset X$,

$$\mu(B) = \int_Y \mu_y(B) \, d\varkappa(y).$$

By Lemma 2.2, the Jacobian of μ is $x \mapsto c(f(x))$. This must be also the Jacobian of almost every ergodic component of μ. Indeed, if this is not the case, then there is a measurable set $A \subset X$ with $\mu(A) > 0$ on which the Jacobian of every ergodic component of μ is larger than the Jacobian of μ (or maybe we have to replace 'larger' by 'smaller', but the proof stays the same). We may assume that f is one-to-one on A. However, then (with the natural notation for Jacobians)

$$\mu(f(A)) = \int_Y \mu_y(f(A)) \, d\varkappa(y) = \int_Y \left(\int_A J_y(z) \, d\mu_y(z) \right) d\varkappa(y)$$

$$> \int_Y \left(\int_A J(z) \, d\mu_y(z) \right) d\varkappa(y) = \int_A J(z) \, d\mu(z) = \mu(f(A)),$$

which is a contradiction. Thus, again by Lemma 2.2, almost every ergodic component of μ is fair. □

The natural procedure for finding a fixed point of Φ would be to start with an arbitrary measure $\nu \in \mathfrak{M}$. Consider the sequence of averages

$$\left(\frac{1}{n} \sum_{i=0}^{n-1} \Phi^i(\nu) \right)_{n=1}^{\infty} \tag{3}$$

and take a weak-* limit of a subsequence of this sequence. The problem is that, in general, the operator Φ is not continuous. Therefore there is no guarantee that this limit will be a fixed point of Φ. Even for piecewise monotone interval maps, one can easily see that Φ can be discontinuous.

Example 2.5. Consider a full 3-horseshoe $f : [0, 1] \to [0, 1]$, that is, the 'connect the dots' map with 'dots' given by $f(0) = 0$, $f(1/3) = 1$, $f(2/3) = 0$ and $f(1) = 1$ (see Figure 1). If δ_x is the Dirac delta measure at x, then $\Phi(\delta_0) = (\delta_0 + \delta_{2/3})/2$. However, if instead of 0 we take a point ε close to 0, then $\Phi(\delta_\varepsilon) = (\delta_{\varepsilon/3} + \delta_{(2-\varepsilon)/3} + \delta_{(2+\varepsilon)/3})/3$. While $\delta_\varepsilon \to \delta_0$ as $\varepsilon \to 0$, the measures $\Phi(\delta_\varepsilon)$ converge to $(\delta_0 + 2\delta_{2/3})/3$ instead of $\Phi(\delta_0)$.

Nevertheless, we can say something about the accumulation points of the sequence (3).

LEMMA 2.6. *Any accumulation point of the sequence (3) is an f-invariant measure.*

Proof. Let $(n_j)_{j=1}^{\infty}$ be an increasing sequence of positive integers and assume that

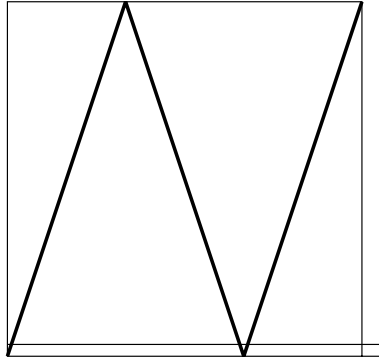$$\lim_{j \to \infty} \frac{1}{n_j} \sum_{i=0}^{n_j - 1} \Phi^i(\nu) = \mu.$$

FIGURE 1. An example where Φ is discontinuous.

in the weak-* topology. If $\varphi : X \to \mathbb{R}$ is a continuous function, then, by (2),

$$\left| f_*\left(\frac{1}{n_j} \sum_{i=0}^{n_j-1} \Phi^i(\nu) \right)(\varphi) - \left(\frac{1}{n_j} \sum_{i=0}^{n_j-1} \Phi^i(\nu) \right)(\varphi) \right|$$

$$= \frac{1}{n_j} \left| \left(f_*(\nu) + \sum_{i=0}^{n_j-2} \Phi^i(\nu) \right)(\varphi) - \left(\sum_{i=0}^{n_j-1} \Phi^i(\nu) \right)(\varphi) \right|$$

$$= \frac{1}{n_j} |(f_*(\nu)(\varphi) - \Phi^{n_j-1}(\nu))(\varphi)| \le \frac{2\|\varphi\|}{n_j},$$

and this converges to zero as $j \to \infty$. Therefore, by the continuity of f_*, $f_*(\mu) = \mu$. This means that μ is f-invariant. □

Similarly, we get the following result.

LEMMA 2.7. *Set*

$$\nu_n = \frac{1}{n} \sum_{k=0}^{n-1} \Phi^k(\nu) \tag{4}$$

and let $(n_j)_{j=1}^{\infty}$ be an increasing sequence of positive integers such that $\lim_{j\to\infty} \nu_{n_j} = \mu$ in the weak- topology. Then also $\lim_{j\to\infty} \Phi(\nu_{n_j}) = \mu$ in the weak-* topology.*

Proof. If $\varphi : X \to \mathbb{R}$ is a continuous function, then

$$|\Phi(\nu_{n_j})(\varphi) - \nu_{n_j}(\varphi)| = \frac{1}{n_j} \left| \left(\sum_{i=1}^{n_j} \Phi^i(\nu) \right)(\varphi) - \left(\sum_{i=0}^{n_j-1} \Phi^i(\nu) \right)(\varphi) \right|$$

$$= \frac{1}{n_j} |\Phi^{n_j}(\nu)(\varphi) - \nu(\varphi)| \le \frac{2\|\varphi\|}{n_j},$$

and this converges to zero as $j \to \infty$. Therefore, if the sequence $(\nu_{n_j})_{j=1}^{\infty}$ converges, then the sequence $(\Phi(\nu_{n_j}))_{j=1}^{\infty}$ converges to the same limit. □

3. Natural extension

Let us try to make connections between various approaches that we employed. Suppose that we constructed a fair measure μ. Consider the natural extension $(\hat{X}, \hat{f}, \hat{\mu})$ of the system (X, f, μ) and let $\pi : \hat{X} \to X$ be the natural projection. Then the partition of $\hat{X}$ into preimages of points under π is a measurable partition. Therefore there is a canonical system of conditional measures $(\mu_x)_{x \in X}$, where μ_x is a probability measure in $\pi^{-1}(x)$. Since $\pi_*(\hat{\mu}) = \mu$, this means that, for every measurable set $B \subset \hat{X}$,

$$\hat{\mu}(B) = \int_X \mu_x(B \cap \pi^{-1}(x)) \, d\mu(x). \tag{5}$$

If $x \in X$, then $\pi^{-1}(x)$ is the set of (infinite) branches of the *preimage tree rooted at x*. To build this tree, we start at x and draw edges from x to its preimages under f. Then we continue by induction: from every vertex y we draw edges to the elements of $f^{-1}(y)$. Since we want to get a tree, if x is periodic, then the same point can represent a countable number of vertices (so, formally, a vertex is of the form (y, n), where $f^n(y) = x$; this vertex is in the *nth generation of vertices*).

In the space $\pi^{-1}(x)$, we can also consider a probability measure ν_x given by the Jacobian of μ. To define this measure, it is enough to specify the measures of sets $[y, n]_x$, where $[y, n]_x$ is the set of all branches of the tree that pass through the vertex y of the nth generation. If J is the measure-theoretic Jacobian of f for the measure μ, then we set

$$\nu_x([y, n]_x) = \frac{1}{J(y) \cdot J(f(y)) \cdot \dots \cdot J(f^{n-1}(y))} = \frac{1}{c(f(y)) \cdot c(f^2(y)) \cdot \dots \cdot c(f^n(y))}.$$

Clearly, this choice of measures is consistent: that is,

$$\sum_{z \in f^{-1}(y)} \nu_x([z, n+1]_x) = \nu_x([y, n]_x).$$

Therefore the definition of ν_x is correct.

The measures ν_x are exactly the measures that we use for our 'random choice' of a backward trajectory. Note that their definition is really independent of the measure μ, because the function $x \mapsto c(f(x))$ does not depend on μ. Instead of the natural extension, we can use the inverse limit in the sense of topological dynamics and we get the same $\hat{X}$, $\hat{f}$ and π. Thus, even if we do not know anything about the existence of a fair measure (as in our questions), we can use the following terminology.

Definition 3.1. We will say that some condition is *satisfied for a random backward trajectory* of $x \in X$ if it is satisfied for ν_x almost every element of $\pi^{-1}(x)$.

The following lemma provides a link between our questions and fair measures.

LEMMA 3.2. *If μ is a fair measure, then $\mu_x = \nu_x$ for μ-almost every x.*

Proof. We need to check that (5) is satisfied if we replace μ_x with ν_x. Every measurable subset of $\hat{X}$ can be approximated (in the sense of measure $\hat{\mu}$) by finite unions of the sets B of the following form. There exists a measurable set $C \subset X$ and $A_1, A_2, \dots, A_n \in \mathcal{A}$ such that $f^i(C) \subset A_i$ for $i = 1, 2, \dots, n$ and

$$B = \{(\dots, x_2, x_1, x_0) : x_n \in C\}.$$

Thus, we can assume that the set B in (5) is of such form. Then

$$\nu_x(B \cap \pi^{-1}(x)) = \frac{1}{c(A_1) \cdot c(A_2) \cdot \cdots \cdot c(A_n)}$$

if $x \in f^n(C)$, and $\nu_x(B \cap \pi^{-1}(x)) = 0$, otherwise. Thus,

$$\int_X \nu_x(B \cap \pi^{-1}(x)) \, d\mu(x) = \frac{\mu(f^n(C))}{c(f(y)) \cdot c(f^2(y)) \cdot \cdots \cdot c(f^n(y))} = \mu(C) = \hat{\mu}(B).$$

This proves that ν_x satisfies (5), and therefore $\nu_x = \mu_x$ μ-almost everywhere. $\square$

Remark 3.3. With more work, we can remove from Lemma 3.2 the assumption that the measure μ is fair and define ν_x using only the Jacobian of μ (without the numbers $c(x_i)$). However, we do not need here such general result.

Now, if μ is a fair ergodic measure, we can get a partial answer to our questions.

THEOREM 3.4. *Let μ be a fair ergodic measure. Then:*
(a) *for μ-almost every point $x_0 \in X$, for a random choice of the backward trajectory of x_0 the geometric averages of $c(x_0), c(x_1), \ldots, c(x_n)$ converge to $\exp(\int \log c(x) \, d\mu(x))$ as n goes to infinity;*
(b) *for μ-almost every point $x_0 \in X$, for a random choice of the backward trajectory of x_0 the measures $(1/n) \sum_{k=0}^{n-1} \delta_{x_k}$ converge in the weak-* topology to μ; and*
(c) *if, additionally, f has a one-sided generator, then the limit in* (a) *is the exponential of the entropy of μ.*

Proof. Use the Birkhoff ergodic theorem for the system $(\hat{X}, \hat{f}^{-1}, \hat{\mu})$ and the function $x \mapsto \log c(\pi(x))$. The phrase 'for $\hat{\mu}$-almost every $y \in \hat{X}$' means the same as 'for μ-almost every $x \in X$ and μ_x-almost every $y \in \pi^{-1}(x)$'. Thus, by Lemma 3.2 and in view of Definition 3.1, averages of $\log c(x)$ converge to

$$\int \log c(\pi(y)) \, d\hat{\mu}(y) = \int \log c(x) \, d\mu(x).$$

This proves (a).

If, additionally, f has a one-sided generator, then the entropy is the integral of the logarithm of the Jacobian, so $h_\mu(f) = \int \log c(f(x)) \, d\mu(x) = \int \log c(x) \, d\mu(x)$ (the second equality holds because μ is invariant), and this proves (c).

Finally, the proof of (b) is almost the same as the proof of (a), except that, instead of the ergodic theorem, we use the fact that almost every point of $\hat{X}$ is generic for $\hat{\mu}$. $\square$

Of course, since we get the answers not for every point of X, but only for almost every one (with respect to some special measure), those are not full answers. Moreover, they are answers only when we know that a fair measure exists and is ergodic. In fact, we would like a unique fair measure, because, otherwise, the limits can heavily depend on the point (one limit for almost every point for one measure; another limit for almost every point for another measure). We will address the existence and uniqueness of the fair measure in some cases in the next sections.

4. *Subshifts of finite type*

Suppose that (X, f) is a transitive one-sided subshift of finite type with the 0–1 transition matrix $M = (m_{ij})$. Then $c(x)$ depends only on the cylinder of length 1 to which x belongs and, if it is the jth cylinder, then it is equal to

$$c_j = \sum_i m_{ij}. \tag{6}$$

We can choose as X_i the ith cylinder of length 1 and $f(X_i)$ to be a union of a finite number of cylinders of length 1. Therefore both the sets X_i and the elements of the partition $\mathcal{A}$ are clopen (close and open) sets. This implies that the operator Φ is continuous.

In this situation, we can use, for instance, the results of [**11**] to deduce the existence and uniqueness of the fair measure as well as its strong ergodic properties. Alternatively, we can use the method mentioned in §2. However, we can do better: namely, we can construct a fair measure explicitly. This way, we get also an explicit formula for its entropy, which is the limit that we are looking for.

Write

$$q_{ij} = \frac{m_{ij}}{c_j}$$

and consider the matrix $Q = (q_{ij})$. For each j, $\sum_i q_{ij} = 1$, so Q has a positive left eigenvector corresponding to the eigenvalue 1. This implies that the spectral radius of Q is 1, so there is a unique probability vector $p = (p_i)$ that is a right eigenvector of Q corresponding to the eigenvalue 1. We can normalize it so that $\sum_i p_i = 1$. Moreover, each p_i is strictly positive.

Now we set

$$p_{ij} = \frac{q_{ij} p_j}{p_i} = \frac{m_{ij} p_j}{c_j p_i}.$$

It is easy to check that this gives us a Markov measure μ with probabilities of the states p_i and transition probabilities p_{ij}. The measure-theoretic Jacobian of μ is constant on cylinders of length 2 and is equal to c_j and on the ijth cylinder, provided this cylinder is non-empty, that is, $m_{ij} = 1$. Thus, by Lemma 2.2, the measure μ is fair.

Whenever $m_{ij} > 0$, also $p_{ij} > 0$. Thus, since our subshift of finite type was transitive, μ is ergodic. Therefore we can use Theorem 3.4. Moreover, in our specific situation we can strengthen it.

THEOREM 4.1. *Let (X, f) be a transitive subshift of finite type. Then for the fair measure μ constructed above:*

(a) *for every point $x_0 \in X$, for a random choice of the backward trajectory of x_0, the geometric averages of $c(x_0), c(x_1), \ldots, c(x_n)$ converge to $\exp(h_\mu(f))$ as n goes to infinity; and*

(b) *for every point $x_0 \in X$, for a random choice of the backward trajectory of x_0, the measures $(1/n) \sum_{k=0}^{n-1} \delta_{x_k}$ converge in the weak-* topology to μ.*

Moreover, μ is the only fair measure for f.

Remark 4.2. This is a kind of theorem about the fair measure that we would like to have in every situation. Therefore we will call statements (a), (b) and the uniqueness of the fair

measure *the Fair Theorem*. Thus, Theorem 4.1 can be restated as: *the Fair Theorem holds for transitive subshifts of finite type*.

Proof of Theorem 4.1. If the zeroth coordinates of the points x_0 and y_0 are equal, then the backward trees for those points look the same, that is, there is a natural bijection between them. This bijection sends the measure ν_{x_0} to ν_{y_0} and, if the branch (y_n) corresponds to (x_n), then the distance between x_n and y_n goes to zero as $n \to \infty$. Therefore, if the choice of the backward trajectory of x_0 gives the convergence described in Theorem 3.4(a) and (b), then the choice of the corresponding backward trajectory of y_0 gives the same convergence. This means that this convergence occurs for every y_0. Moreover, in our case, $\mathcal{X}$ is a generator, so we can use Theorem 3.4(c). This completes the proof of (a) and (b).

If there is another fair measure for f, then, by Lemma 2.4, there is another ergodic fair measure and, by Theorem 3.4, this contradicts (b). □

Now we can look back at the results of the computer experiment described in §1. The interval map considered there was Markov with the 2×2 transition matrix (m_{ij}), where $m_{12} = m_{21} = m_{22} = 1$ and $m_{11} = 0$. Therefore $c_1 = 1$ and $c_2 = 2$. This gives

$$Q = \begin{bmatrix} 0 & 1/2 \\ 1 & 1/2 \end{bmatrix},$$

and thus $p_1 = 1/3$ and $p_2 = 2/3$. Hence $p_{11} = 0$, $p_{12} = 1$ and $p_{21} = p_{22} = 1/2$. Now we can compute the entropy of μ

$$h_\mu(f) = -\sum_{i,j} p_i p_{ij} \log p_{ij} = -\frac{2}{3} \log \frac{1}{2} = \log(2^{2/3}) \approx \log 1.5874.$$

Thus the number that we got in the experiment was (as it should be) the exponential of the entropy of the unique fair measure for the corresponding subshift of finite type. We will elaborate on the connection between the interval Markov maps and subshifts of finite type in §6.

5. *Piecewise monotone interval maps*

We will show that, in spite of the discontinuity of Φ, the procedure mentioned in §2 works for piecewise monotone interval maps.

Thus, in this section, our space is $X = [0, 1]$, a map $f : X \to X$ is continuous, piecewise monotone and topologically mixing (that is, it belongs to $\mathcal{PMM}$) and X_i, $i = 1, 2, \ldots, s$, are intervals of monotonicity of f. We really want $\mathcal{X}$ to be a partition, so we assign each turning point of f to only one of the intervals X_i, X_{i+1}. Note that then some elements of the partition $\mathcal{A}$ are singletons consisting of the image of a turning point.

LEMMA 5.1. *Let K be a subinterval of X and let $\varkappa \in \mathfrak{M}$. Then*

$$\Phi(\varkappa)(K) \geq \frac{1}{s} \varkappa(f(K)). \tag{7}$$

Proof. We can partition K into subintervals (some of them perhaps degenerate) by intersecting with the preimages of the elements of $\mathcal{A}$. For each of those subintervals, (7) holds (with this subinterval replacing K), by the definition of Φ. Taking the sum of

those inequalities, we get (7) for K. In fact, at the right-hand side, we may get even more than $(1/s)\varkappa(f(K))$, because the images of our subintervals need not be disjoint, but the inequality goes in the right direction. □

LEMMA 5.2. *Let $\mu \in \mathfrak{M}$ be an accumulation point of the sequence (3) and let K be a subinterval of X such that μ-measures of the boundary of K and of the boundary of $f(K)$ are zero. Then*

$$\mu(K) \geq \frac{1}{s}\mu(f(K)). \tag{8}$$

Proof. Define ν_n as in (4) and assume that $(n_j)_{j=1}^{\infty}$ is an increasing sequence of positive integers such that $\lim_{j\to\infty} \nu_{n_j} = \mu$ in the weak-* topology. Let us recall that the weak-* convergence of measures implies convergence of measures of sets whose boundaries have limit measure zero. Therefore, by Lemmas 2.7 and 5.1,

$$\mu(K) = \lim_{j\to\infty} \Phi(\nu_{n_j})(K) \geq \frac{1}{s} \lim_{j\to\infty} \nu_{n_j}(f(K)) = \frac{1}{s}\mu(f(K)). \qquad \square$$

LEMMA 5.3. *Let $\mu \in \mathfrak{M}$ be an accumulation point of the sequence (3). Then μ is non-atomic.*

Proof. Suppose that μ has an atom at some point $x \in X$. Since μ is invariant by Lemma 2.6, $\mu(\{f^n(x)\}) \geq \mu(\{x\})$ for every $n > 0$. Therefore x has to be periodic and if P is the periodic orbit of x, then $\mu(\{y\}) = \mu(\{x\})$ for every $y \in P$. It follows that if $f^k(z) = x$ and $z \notin P$, then $\mu(\{z\}) = 0$.

As we mentioned, f is topologically exact (locally eventually onto). Thus there exists an interval $\tilde{L} \subset X$ disjoint from P and a number $k > 0$ such that $f^k(\tilde{L}) = X$. Let K be an interval containing x, whose boundary has measure μ zero. By [1, Lemma 1.2.1], there is a subinterval $\hat{L} \subset \tilde{L}$ such that $f^k(\hat{L}) = K$. Applying this lemma once more, we get a subinterval $L \subset \hat{L}$ such that $f^k(L) = K$ and the endpoints of L are mapped to the endpoints of K. Since the boundary of K has measure μ zero and μ is invariant, we see that the boundary of $f^i(L)$ for $i = 0, 1, \ldots, k-1$ has also measure μ zero.

A minor problem can occur because, in Lemma 5.2, the boundary is understood as the boundary relative to the interval X, so an endpoint of K may not belong to the boundary of K if it is an endpoint of X. However, then we can extend slightly L so that its image stays the same, but the singleton of the image under f^k of the corresponding endpoint of L has measure zero. This cannot be done if the endpoint of L where we want to enlarge L is an endpoint of X, but then this endpoint does not belong to the boundary of L, so the problem disappears.

As a result, we get a situation where we can apply Lemma 5.2 k times. Taking a descending sequence of intervals K_n with intersection $\{x\}$, we get, by induction, a descending sequence of intervals L_n. By Lemma 5.2, $\mu(L_n) \geq (1/s^k)\mu(K_n) \geq \mu(\{x\})$ for every n. The intersection of the intervals L_n cannot be a non-degenerate interval because its image under f^k is $\{x\}$ and f is mixing. Thus it is a singleton of some point z. Since L was disjoint from P, $z \notin P$. Also $\mu(\{z\}) > 0$, which contradicts what we proved earlier. □

We say that a measure has *full support* if every open non-empty set has positive measure.

LEMMA 5.4. *Let $\mu \in \mathfrak{M}$ be an accumulation point of the sequence (3). Then μ has full support.*

Proof. By Lemmas 5.2 and 5.3, the inequality (8) holds for every interval K. If K is a non-degenerate interval, then there is s such that $f^s(K) = X$. Then, by (8), $\mu(K) > 0$. □

Now we are ready to prove the existence of a fair measure.

THEOREM 5.5. *Every accumulation point of the sequence (3) is a fair measure.*

Proof. Let $\mu \in \mathfrak{M}$ be an accumulation point of the sequence (3). We will show that if K is a subinterval of X such that $f(K)$ is contained in some element $A \in \mathcal{A}$, then

$$\mu(K) = \frac{1}{c(A)}\mu(f(K)). \tag{9}$$

This means that the Jacobian of μ is $x \mapsto c(f(x))$ so, by Lemma 2.2, μ is fair.

Define ν_n as in (4) and assume that $(n_j)_{j=1}^{\infty}$ is an increasing sequence of positive integers such that $\lim_{j \to \infty} \nu_{n_j} = \mu$ in the weak-* topology. By the definition of Φ,

$$\Phi(\nu_n)(K) = \frac{1}{c(A)}\nu_n(f(K))$$

for every n. By Lemma 5.3, both K and $f(K)$ have boundaries of measure μ zero. Therefore, by Lemma 2.7,

$$\mu(K) = \lim_{j \to \infty} \Phi(\nu_{n_j})(K) = \frac{1}{c(A)} \lim_{j \to \infty} \nu_{n_j}(f(K)) = \frac{1}{c(A)}\mu(f(K)). \qquad \square$$

Remark 5.6. If μ is a fair measure, then we can use μ as the starting measure ν. Then the sequence (3) is constant, so μ satisfies all lemmas about the accumulations points of this sequence. In particular, μ is f-invariant, non-atomic and has full support.

Now we prove uniqueness of the fair measure. We start with a lemma which is probably well known, but it is simpler to prove it than to find it in the literature.

LEMMA 5.7. *Let μ and ν be mutually singular non-atomic probability measures with full support on X. Then, for μ-almost every $x \in X$,*

$$\lim_{\varepsilon, \delta \to 0} \frac{\nu([x - \delta, x + \varepsilon])}{\mu([x - \delta, x + \varepsilon])} = 0. \tag{10}$$

Proof. Since the measure $\mu + \nu$ is non-atomic and has full support, we may assume that it is the Lebesgue measure (we apply a homeomorphism of X that carries it to the Lebesgue measure). Since μ and ν are mutually singular, there is a Borel set $A \subset X$ with $\mu(A) = 1$ and $\nu(A) = 0$. By the (one-sided) Lebesgue density theorem, for μ-almost every $x \in A$ (and therefore for μ-almost every $x \in X$),

$$\lim_{\varepsilon, \delta \to 0} \frac{\mu([x - \delta, x + \varepsilon])}{\delta + \varepsilon} = \lim_{\varepsilon, \delta \to 0} \frac{(\mu + \nu)([x - \delta, x + \varepsilon] \cap A)}{\delta + \varepsilon} = 1. \tag{11}$$

Since

$$\delta + \varepsilon = (\mu + \nu)([x - \delta, x + \varepsilon]) = \mu([x - \delta, x + \varepsilon]) + \nu([x - \delta, x + \varepsilon]),$$

from (11),

$$\lim_{\varepsilon, \delta \to 0} \frac{\nu([x - \delta, x + \varepsilon])}{\delta + \varepsilon} = 0. \tag{12}$$

Now (10) follows from (11) and (12). ▢

The above lemma shows that, on a small scale, the mutually singular measures are very incompatible. Now we show, that on a large scale, the situation is quite different.

LEMMA 5.8. *Let μ and ν be non-atomic probability measures with full support on X. Then, for each $\varepsilon > 0$, there exists $\delta > 0$ such that if $K \subset X$ is an interval and $\mu(K) \geq \varepsilon$, then $\nu(K) \geq \delta$.*

Proof. Suppose that this is not true. Then there exists a sequence of intervals $[a_n, b_n] \subset X$ such that $\mu([a_n, b_n]) \geq \varepsilon$ and $\nu([a_n, b_n]) < 1/n$. By compactness, we may assume that $a_n \to a$ and $b_n \to b$ as $n \to \infty$. Since μ is non-atomic, $\mu([a, b]) \geq \varepsilon$, and therefore $b > a$. Similarly, $\nu([a, b]) = 0$. Since ν has full support, this implies that $b = a$ and we get a contradiction. This concludes the proof. ▢

Finally, we need to have tools to pass from a small to a large scale. The following lemma is well known, and is basically a special case of [**12**, Lemma 4.1]. Notice that, although formally in [**12**] the assumption is that the slope is constant, what is really used in the proof is only that it is larger than 2. Note also that, in our terminology, 'piecewise' means piecewise with a finite number of pieces.

LEMMA 5.9. *Let $g : X \to X$ be a piecewise linear map with slopes on each piece larger than 2. Then there exists $\varepsilon > 0$ such that, for Lebesgue almost every $x \in X$, there exists an increasing sequence (n_k) of positive integers such that, for every k, there exists an interval K_k containing x such that g^{n_k} is linear on K_k and the length of $g^{n_k}(K_k)$ is at least ε.*

Observe that, under the assumptions of this lemma, the length of K_k is at most 2^{-n_k}, so it goes to zero as k goes to infinity.

Now we are prepared to prove the result that is the main step for proving uniqueness of the fair measure.

PROPOSITION 5.10. *Let $\mu, \nu \in \mathfrak{M}$ be fair measures. Then they cannot be mutually singular.*

Proof. Since f is mixing, there is n such that f^n is at least two-to-one except perhaps one point. Since, by Remark 5.6, μ is non-atomic and has full support, we may assume that it is the Lebesgue measure. Then the slope of f^n on each lap is at least 2, so the slope of f^{2n} on each slope is at least 4. By Lemma 5.9, there is ε such that, for μ-almost every $x \in X$, there exists an increasing sequence (n_k) of positive integers such that, for every k, there exists an interval K_k containing x such that f^{nn_k} is linear on K_k and the length of $f^{nn_k}(K_k)$ is at least ε. By Lemma 5.8, there is $\delta > 0$ (depending only on ε) such that $\nu(f^{nn_k}(K_k)) \geq \delta$. Since μ and ν are fixed points of Φ,

$$\frac{\nu(K_k)}{\mu(K_k)} = \frac{\nu(f^{nn_k}(K_k))}{\mu(f^{nn_k}(K_k))} \geq \delta.$$

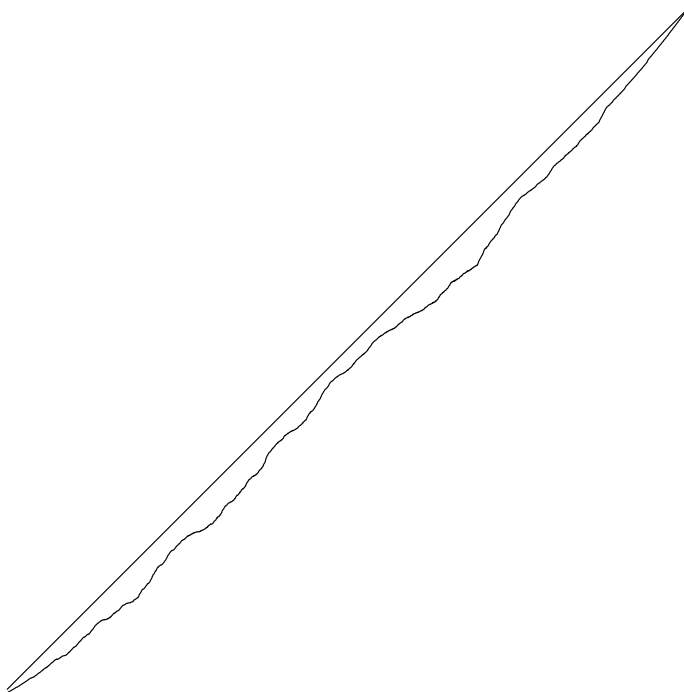Thus, by Lemma 5.7, μ and ν cannot be mutually singular. ▢

FIGURE 2. The graphs of $\exp(h_{\mu_{\text{fair}}}(f))$ and $\exp(h_{\text{top}}(f))$ for tent maps f with slopes from 1.42 to 2.

Now we are ready to prove uniqueness of the fair measure.

THEOREM 5.11. *The map f has only one fair measure.*

Proof. If there are two distinct fair measures, by Lemma 2.4, there are two distinct ergodic fair measures. Then they are mutually singular, which contradicts Proposition 5.10. □

Now let us look at the broader picture. We can start with any measure $\nu \in \mathfrak{M}$, take an accumulation point of the sequence (3) and always get the same fair measure. We will denote this measure by μ_{fair}.

THEOREM 5.12. *For every $\nu \in \mathfrak{M}$ the sequence (3) converges in the weak-* topology to μ_{fair}.*

Proof. Since the space $\mathfrak{M}$ is compact, every subsequence of (3) has a convergent subsequence and, by Theorems 5.5 and 5.11, the limit is μ_{fair}. Therefore the whole sequence (3) converges to μ_{fair}. □

Let us look at the entropy of the fair measure for the tent maps and compare it with the topological entropy (see Figure 2). The maps are parametrized by the slope, so the graph of the exponential of the topological entropy is a straight line. Except for the slope 2 (and the slope $\sqrt{2}$, which is just to the left of the beginning of the graphs; the map is transitive
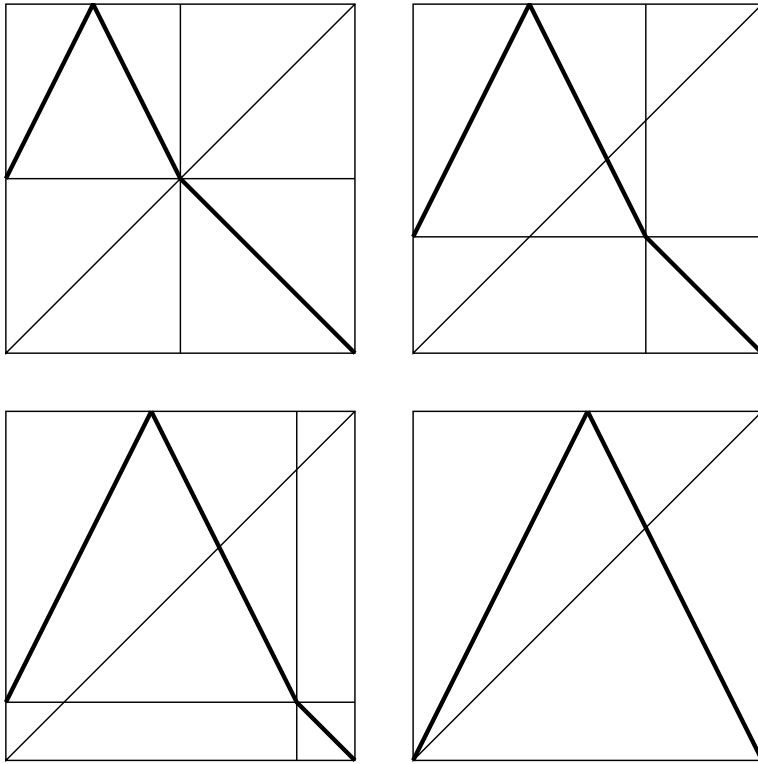
FIGURE 3. The maps g_b.

but not mixing then) the entropy of the fair measure is smaller than the topological entropy. However, the difference is not large. The entropy of the fair measure as the function of the slope looks continuous and increasing. We will prove that, in fact, this is the case.

THEOREM 5.13. *For the family of the tent maps, parametrized by the slope from $(\sqrt{2}, 2]$, the entropy of the fair measure is continuous and strictly increasing as the function of the parameter.*

Proof. Denote by f_a the tent map with the slope a and by g_b the unimodal map defined as the 'connect the dots' map on $[0, 1]$ with the dots $(0, 1 - b)$, $(b/2, 1)$, $(b, 1 - b)$ and $(1, 0)$. Then the Jacobian of g_b for the Lebesgue measure is two on $[0, b]$ and one on $(b, 1]$, while g_b is two-to-one on $g_b([0, b])$ and one-to-one on $g_b((b, 1])$ (see Figure 3). Therefore the Lebesgue measure is the fair measure. Its entropy is the integral of the logarithm of the Jacobian so it is equal to $b \log 2$. If we allow also $a = \sqrt{2}$ and $b = 1/2$ then, since each g_b is transitive, there exists exactly one $a = \xi(b)$ for which g_b is conjugate to f_a (see, e.g., [**2**]). The topological entropy is a continuous function of a unimodal map with positive entropy (see [**1**]), so ξ is a continuous function. All maps g_b for $b \in (1/2, 1]$ are topologically mixing, so $\xi(b) \in (\sqrt{2}, 2]$. Moreover, $\xi(1/2) = \sqrt{2}$ and $\xi(1) = 2$. If ξ is not strictly increasing, then there exist two maps g_b and $g_{b'}$ which are conjugate to the same

f_a, so they are conjugate to each other. This contradicts the uniqueness of the fair measure. Thus ξ is a homeomorphism, so ξ^{-1} is well defined, continuous and strictly increasing. $\square$

6. *Interval Markov maps*

In this section, we show that, although the answers to our questions are not full for general mixing piecewise monotone interval maps, the situation is better for Markov maps, where we can use the results of §4.

THEOREM 6.1. *The Fair Theorem (see Theorem 4.1 and Remark 4.2) holds for Markov maps from $\mathcal{PMM}$.*

Proof. The uniqueness of the fair measure follows from Theorem 5.11.

Since f is Markov, there is a corresponding subshift of finite type $g : Y \to Y$. The only obstruction to the systems (X, f) and (Y, g) being conjugate is that the intervals of the Markov partition are closed, so the endpoints of those intervals different from 0 and 1 belong to two intervals (so, strictly speaking, it is not a partition).

Let P be the set of endpoints of the intervals of the Markov partition. There are only a countable number of points whose forward or backward trajectories pass through P, and the fair measures for f and g are non-atomic, so those systems with their fair measures are isomorphic. Moreover, if the preimage tree of a point $x_0 \in X$ does not contain a point of P, then its preimage tree is the same (including the measure) as the preimage tree of its counterpart for g, and (a) and (b) follow from Theorem 4.1.

Thus we can restrict our attention to the points whose backward trees pass through P. Since P is invariant, it follows that the root x_0 of such a tree belongs to P. Except for the branches that pass through P an infinite number of times, the rest of the tree, after removing the vertices from P, splits into a countable number of subtrees rooted at a vertex not belonging to P. For those subtrees, (a) and (b) hold, and this implies, for the corresponding parts of the tree rooted at x_0, that also (a) and (b) hold. If a branch passes through P an infinite number of times, then, since P is invariant and finite, it is periodic. Therefore there are only a finite number of such branches and, since we are talking about a random choice of the backward trajectory, we can ignore them (they constitute a set of measure zero for the measure ν_{x_0}). This completes the proof. $\square$

7. *Maximal entropy*

Let us recall that our motivation was to compute the topological entropy of a map. As we saw, we get instead the fair entropy. Thus we can ask about the conditions for fair entropy and the topological entropy to coincide. Let us consider this question for topologically mixing subshifts of finite type.

Thus let $f : X \to X$ be a topologically mixing subshift of finite type with the transition matrix $M = (m_{ij})$. Then it has the unique fair measure μ_{fair} and the unique measure of maximal entropy μ_{max}.

The spectral radius of M is λ, where $\log \lambda$ is the topological entropy of f. For the eigenvalue λ, M has positive eigenvectors: a left one (ℓ_i) and a right one (r_i). According to the Parry's formulas, μ_{max} is a Markov measure with measures of 1-cylinders $[i]$ equal to

$\ell_i r_i$ (where we normalize eigenvectors in such a way that $\sum_i \ell_i r_i = 1$), and the transition probability from i to j is

$$\frac{m_{ij} r_j}{\lambda r_i}.$$

LEMMA 7.1. *If $m_{ij} = 1$, then the Jacobian of $\mu_{\max}$ on the 2-cylinder $[ij]$ is $\lambda \ell_j / \ell_i$.*

Proof. The Jacobian on this cylinder is equal to the measure of the 1-cylinder $[j]$ divided by the measure of the 2-cylinder $[ij]$. Thus it is

$$\frac{\ell_j r_j}{\ell_i r_i \frac{m_{ij} r_j}{\lambda r_i}} = \frac{\ell_j r_j \lambda r_i}{\ell_i r_i m_{ij} r_j} = \frac{\lambda \ell_j}{\ell_i}. \qquad \square$$

To state a theorem, we need some definitions. We say that functions $\varphi, \psi : X \to \mathbb{R}$ are *cohomologous* if there is a function $v : X \to \mathbb{R}$ such that $\varphi - \psi = v \circ f - v$. In our case, the functions φ and ψ will be constant on 1-cylinders so, in this definition, we will assume that u is also constant on 1-cylinders. Consider the case when the functions are $\varphi = \log c \circ f$ and ψ is constant. Then we can write $\psi = \log K$ and, in the case when φ and ψ are cohomologous, can take $u = e^v$. Then φ and ψ are cohomologous if and only if there exists a vector (u_i) such that if $m_{ij} = 1$, then

$$\log c_j - \log K = \log u_j - \log u_i,$$

that is,

$$u_i = \frac{K u_j}{c_j}. \tag{13}$$

We will use the standard representation of our subshift of finite type by a graph. If L is a loop in this graph, then we will denote by $C(L)$ the product of the values of c along this loop and by $\Lambda(L)$ the product of λs along the loop, that is, $\Lambda(L) = \lambda^{|L|}$, where $|L|$ is the length of L. We will call a loop *simple* if it passes through every vertex of the graph at most once.

Now we are ready to state the theorem.

THEOREM 7.2. *Let f be a topologically mixing subshift of finite type. With the notation explained above, the following conditions are equivalent.*
(i)   *The fair entropy of f is equal to its topological entropy.*
(ii)   *The function $\log c \circ f$ is cohomologous to a constant.*
(iii) *$\mu_{\text{fair}} = \mu_{\max}$.*
(iv)   *For every i, j such that $m_{ij} = 1$, $\ell_i = \lambda \ell_j / c_j$.*
(v)   *For every loop L in the graph of f, $C(L) = \Lambda(L)$.*
(vi)   *For every simple loop L in the graph of f, $C(L) = \Lambda(L)$.*

Proof. (i) $\Rightarrow$ (iii) follows from the uniqueness of the measure of maximal entropy; (iii) $\Rightarrow$ (i) is immediate.

(iii) $\Leftrightarrow$ (iv) follows from Lemmas 7.1 and 2.2.

(iv) $\Rightarrow$ (ii) is immediate. To prove (ii) $\Rightarrow$ (iv), for every j, use (13) and (6) to get

$$\sum_i m_{ij} u_i = \frac{K u_j}{c_j} \sum_i m_{ij} = K u_j.$$
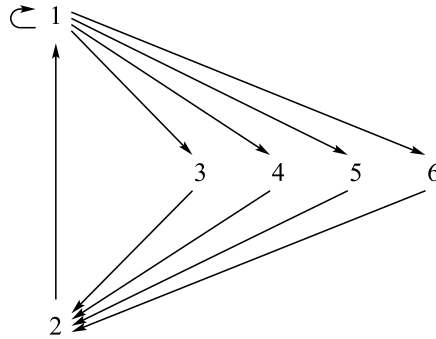
FIGURE 4. The graph of the subshift of finite type from the example.

This means that the positive vector (u_i) is a left eigenvector of M for a positive eigenvalue K. Therefore $K = \lambda$ and $u_i = t\ell_i$ for some positive constant t. This proves (iv).

(iv) $\Rightarrow$ (v) follows by multiplication of $\ell_i = \lambda\ell_j/c_j$ along the loop. To prove (v) $\Rightarrow$ (iv), set $k_1 = \ell_1$ and define the other k_i by induction along the arrows in the graph of f using the formula $k_i = \lambda k_j/c_j$ whenever $m_{ij} = 1$. The assumption that $C(L) = \Lambda(L)$ for every loop L guarantees that this definition is consistent: when we get back to i for which k_i is already defined, we get the same value of k_i. Now for every j

$$\sum_i k_i m_{ij} = \frac{\lambda k_j}{c_j} m_{ij} = \lambda k_j,$$

so (k_i) is a left eigenvector of M for the eigenvalue λ. Since $k_1 = \ell_1$, this vector is equal to (ℓ_i). This proves (iv).

(v) $\Rightarrow$ (vi) is immediate; (vi) $\Rightarrow$ (v) follows from the fact that every loop is a concatenation of simple loops. $\qquad\square$

Let us comment on the conditions in Theorem 7.2. We want to characterize mixing subshifts of finite type for which (i) (or (iii)) is true. From the theoretical point of view, (ii) is the main characterization. However, from the practical point of view, (vi) is the most important, because it is the simplest condition to check. Conditions (iv) and (v) are more technical and are included to make the proofs simpler.

The experiment illustrated in Figure 2 suggests that the fair entropy is equal to the topological entropy only if the function c is constant. We will give an example of a mixing subshift of finite type when this is not true, and then explain why Figure 2 suggests the opposite.

Example 7.3. Let f be the subshift of finite type with the graph as in Figure 4. That is, the transition matrix of f is

$$M = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

For every $n \geq 3$ there is a path from every vertex to every vertex in the graph, so f is mixing. We have $c_1 = 2$, $c_2 = 4$, and $c_i = 1$ for $i = 3, 4, 5, 6$. There are five simple loops: $1 \to 1$, and $1 \to i \to 2 \to 1$ for $i = 3, 4, 5, 6$. For the first simple loop, L_1, $C(L_1) = 2$ and $|L_1| = 1$. For any other simple loop, L, $C(L) = 8$ and $|L| = 3$. The vector $(2\ 2\ 1\ 1\ 1\ 1)$ is a left eigenvector of M for the eigenvalue 2, so $\lambda = 2$. Thus, condition (vi) of Theorem 7.2 is satisfied, so all other conditions of that theorem are also satisfied. Nevertheless, the function c is not constant.

Now we look again at Figure 2 and try to explain it, at least for Markov maps.

THEOREM 7.4. *For Markov tent maps f_a with slope $a \in (\sqrt{2}, 2]$, the fair entropy is equal to the topological entropy if and only if $a = 2$.*

Proof. For $a = 2$, the Lebesgue measure is fair, and both the fair entropy and the topological entropy are equal to log 2.

Assume now that $a < 2$. The corresponding subshift of finite type $g : Y \to Y$ (as in the proof of Theorem 6.1) with the fair measure and with the measure of maximal entropy is isomorphic to f_a with its fair measure and with its measure of maximal entropy, respectively, because all those measures are non-atomic. Therefore it is enough to prove that one of the conditions of Theorem 7.2 is violated for g.

The map f_a has a fixed point x, and x has two preimages. If x belongs to the interior of one of the intervals of the Markov partition, then, for the corresponding index j for g, there is a loop $j \to j$ and $c_j = 2$. Call this loop L. Then $C(L) = 2$ and $\Lambda(L) = a$. This violates condition (vi).

If x does not belong to the interior of one of the intervals of the Markov partition, then it is a common endpoint of two of them, with indices j and k. The map f_a in a neighborhood of x reverses orientation, so there is a loop $j \to k \to j$. Again, call this loop L. Then $C(L) = 4$ and $\Lambda(L) = a^2$ and, as before, this violates condition (vi). This completes the proof. □

Conjecture 7.5. Theorem 7.4 holds for all tent maps f_a with slope $a \in (\sqrt{2}, 2]$, which are not necessarily Markov.

While we do not know how to prove this conjecture, the numerical evidence (Figure 2) is strong. Moreover, the computations at the end of §4 together with the considerations in §6 show that, when a is the golden ratio, the fair entropy of f_a is log $2^{2/3}$. Hence, by Theorem 5.13, if $2^{2/3} \leq a < (\sqrt{5} + 1)/2$, then the fair entropy of f_a is less than log $2^{2/3}$. This proves that if $a \in [2^{2/3}, (\sqrt{5} + 1)/2]$, then the fair entropy of f_a is strictly smaller than its topological entropy.

REFERENCES

[1] Ll. Alsedà, J. Llibre and M. Misiurewicz. *Combinatorial Dynamics and Entropy in Dimension One (Advanced Series in Nonlinear Dynamics, 5)*, 2nd edn. World Scientific, Singapore, 2000.
[2] Ll. Alsedà and M. Misiurewicz. Semiconjugacy to a map of a constant slope. *Discrete Contin. Dyn. Syst. Ser. B* **20** (2015), 3403–3413.

[3] M. F. Barnsley. *Fractals Everywhere*, 2nd edn. Academic Press Professional, Boston, MA, 1993.

[4] L. Block and W. Coppel. *Dynamics in One Dimension (Lecture Notes in Mathematics, 1513)*. Springer, Berlin, 1992.

[5] L. Block, J. Keesling, S. H. Li and K. Peterson. An improved algorithm for computing topological entropy. *J. Statist. Phys.* **55** (1989), 929–939.

[6] H. Bruin. An algorithm to compute the topological entropy of a unimodal map. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **9** (1999), 1881–1882.

[7] R. Dilao and J. Amigo. Computing the topological entropy of unimodal maps. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **22** (2012), 1250152, 14pp.

[8] J. Hawkins and M. Taylor. Maximal entropy measure for rational maps and a random iteration algorithm. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **13** (2003), 1442–1447.

[9] F. Hofbauer and G. Keller. Equilibrium states for piecewise monotonic transformations. *Ergod. Th. & Dynam. Sys.* **2**(1) (1982), 23–43.

[10] M. Keane. Strongly mixing g-measures. *Invent. Math.* **16** (1972), 309–324.

[11] F. Ledrappier. Principe variationnel et systmes dynamiques symboliques. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **30** (1974), 185–202.

[12] M. Misiurewicz and S. Roth. No semiconjugacy to a map of constant slope. *Ergod. Th. & Dynam. Sys.* **36** (2016), 875–889.

[13] W. Parry. *Entropy and Generators in Ergodic Theory*. W. A. Benjamin, Inc., New York, 1969.

[14] P. Walters. Ruelle's operator theorem and g-measures. *Trans. Amer. Math. Soc.* **214** (1975), 375–387.