C O L L O Q U I U M M A T H E M A T I C U M

VOL. 127 2012 NO. 1

# MAXIMAL ENTROPY MEASURES IN DIMENSION ZERO

BY

DAWID HUCZEK (Wrocław)

**Abstract.** We prove that an invertible zero-dimensional dynamical system has an invariant measure of maximal entropy if and only if it is an extension of an asymptotically $h$-expansive system of equal topological entropy.

**1. Introduction.** The variational principle is a result connecting the notions of measure-theoretic and topological entropy—it states that the topological entropy of a dynamical system is the supremum of the entropies of all invariant measures on that system. This supremum is not always attained and a general criterion for deciding whether a maximal entropy measure exists has been missing for many classes of systems. It has been known for some time that maximal entropy measures exist on asymptotically $h$-expansive systems (see [M])—in fact, in this case the entropy function is upper semicontinuous on the space of invariant measures (hence it attains its maximum). Trivially, any extension of an asymptotically $h$-expansive system that does not increase the topological entropy also has a maximal entropy measure. In this paper we prove that, in the zero-dimensional case, this sufficient condition is also necessary, i.e. that any zero-dimensional system that has a measure of maximal entropy is the extension of an asymptotically $h$-expansive system with equal topological entropy.

## 2. Preliminaries

**Dynamical systems.** Throughout this work a *dynamical system* will be a pair $(X, T)$, where $X$ is a compact metric space, and $T$ is a homeomorphism on $X$.

**Measure-theoretic entropy.** All entropy computations use logarithms to base 2. The facts we recall are very standard and their proofs can be found in most textbooks on entropy, e.g. [D].

2010 *Mathematics Subject Classification*: Primary 37A35; Secondary 37B40.
*Key words and phrases*: zero-dimensional system, topological entropy, asymptotically $h$-expansive system.

DOI: 10.4064/cm127-1-4  

For a measurable partition $\mathcal{A}$ we define the *entropy* of $\mathcal{A}$ with respect to an invariant measure $\mu$ by

$$H(\mu,\mathcal{A}) = -\sum_{A\in\mathcal{A}} \mu(A)\log\mu(A).$$

Define also

$$H_n(\mu,\mathcal{A}) = \frac{1}{n}H(\mu,\mathcal{A}^n),$$

where

$$\mathcal{A}^n = \bigvee_{j=0}^{n-1} T^{-j}(\mathcal{A}) = \Big\{\bigcap_{j=0}^{n-1} A_j : A_j \in T^{-j}(\mathcal{A})\Big\}.$$

The sequence $H_n(\mu,\mathcal{A})$ decreases and its limit is denoted by $h(\mu,\mathcal{A})$. Finally define the *measure-theoretical* entropy of $\mu$ as

$$h(\mu) = \sup_{\mathcal{A}} h(\mu,\mathcal{A}).$$

All the notions introduced above depend on the transformation $T$; the same measure and partition can have different entropies for various transformations. Throughout this paper, however, we always consider only one transformation on any given space, and all entropies are calculated with respect to this transformation, so we need not be concerned by this issue.

Given two measurable partitions of $X$, say $\mathcal{A}$ and $\mathcal{B}$, we denote the *conditional entropy* of $\mathcal{B}$ given $\mathcal{A}$ as

$$H(\mu,\mathcal{B}\,|\,\mathcal{A}) = \sum_{A\in\mathcal{A}} \mu(A)H(\mu_A,\mathcal{B}), \tag{2.1}$$

where $\mu_A$ is defined as $\mu_A(B) = \mu(A\cap B)/\mu(A)$. Define

$$H_n(\mu,\mathcal{B}\,|\,\mathcal{A}) = \frac{1}{n}H(\mu,\mathcal{B}^n\,|\,\mathcal{A}^n).$$

Again, this sequence decreases to a value called the *dynamical conditional entropy* of $\mathcal{B}$ given $\mathcal{A}$ and denoted by $h(\mu,\mathcal{B}\,|\,\mathcal{A})$. We say that $\mathcal{B}$ is a *refinement* of $\mathcal{A}$ (denoted by $\mathcal{A}\prec\mathcal{B}$) if each set from $\mathcal{B}$ is a subset of some set from $\mathcal{A}$. In this case we have

$$H(\mu,\mathcal{B}\,|\,\mathcal{A}) = H(\mu,\mathcal{B}) - H(\mu,\mathcal{A}).$$

The following estimate is completely trivial, but we will use it several times while proving the main result, so we mention it explicitly:

$$H(\mu,\mathcal{A}) \le \log\#\mathcal{A}, \tag{2.2}$$

where $\#\mathcal{A}$ is the number of elements of $\mathcal{A}$.

In particular, this gives us the following

LEMMA 2.1. *If $\mathcal{A}$ and $\mathcal{B}$ are two measurable partitions of the same space and every set of $\mathcal{A}$ intersects at most $L$ sets of $\mathcal{B}$, then for any measure $\mu$ we have $H(\mu, \mathcal{B} \mid \mathcal{A}) < \log L$.*

To prove this, simply apply estimate (2.2) to each term in the formula (2.1) (for each $A$, $H(\mu_A, \mathcal{B})$ is equivalent to the entropy of a partition consisting of at most $L$ elements).

We will also use a corollary of the Shannon–McMillan–Breiman Theorem:

COROLLARY 2.2. *For any ergodic measure $\mu$, any finite partition $\mathcal{A}$ and any positive $\varepsilon$ and $\delta$ there exists a number $N$ such that for any $n > N$ there exists a subset $\mathcal{C}$ of $\mathcal{A}^n$ of total measure $\mu$ at least $1-\delta$ (i.e. $\sum_{C\in\mathcal{C}} \mu(C) > (1-\delta)$) such that*

$$2^{-n(h(\mu,\mathcal{A})+\varepsilon)} < \mu(C) < 2^{-n(h(\mu,\mathcal{A})-\varepsilon)} \quad \text{for every } C \in \mathcal{C}.$$

**Topological and measure-theoretical entropy in topological dynamical systems.** A *topological dynamical system* is a pair $(X, T)$, where $X$ is a compact metric space and $T$ is a continuous map of $X$ into itself. Two such systems $(X, T)$ and $(Y, S)$ are *conjugate* if there exists a homeomorphism $\pi : X \to Y$ such that $\pi \circ T = S \circ \pi$.

Let $\mathcal{M}_T(X)$ denote the set of $T$-invariant measures on $X$. We can consider measure-theoretical entropy as a function on $\mathcal{M}_T(X)$ which takes positive values (including, possibly, infinity).

For any compact metric space $X$ a function $f : X \to \mathbb{R}$ is called *upper semicontinuous* if it is the infimum of a family of continuous functions. Equivalently, $f$ is upper semicontinuous if for any convergent sequence $x_n \to x$ we have $\overline{\lim} f(x_n) \le f(x)$. It is well-known (see e.g. [D]) that if $(X, T)$ is a subshift then the entropy function $\mu \mapsto h(\mu)$ is upper semicontinuous on $\mathcal{M}_T(X)$.

The *topological entropy* of $(X, T)$ can be defined as

$$\mathbf{h}(T) = \sup_{\mu \in \mathcal{M}_T(X)} h(\mu, T).$$

See [D] for the more usual definition and the proof of its equivalence with the one presented here.

**Entropy structure and asymptotic $h$-expansiveness in dimension zero.** Recall that a (two-sided) *full shift* is the space $\Lambda^{\mathbb{Z}}$ (where $\Lambda$ is some finite set with discrete topology) with the shift transformation, i.e. the action $(Tx)_n = x_{n+1}$, whereas a *subshift* (or a *symbolic dynamical system*) is any closed, $T$-invariant subset $X$ of $\Lambda^{\mathbb{Z}}$. It is well known that the entropy of any invariant measure $\mu$ on a subshift is attained as the entropy of this measure with respect to the partition into cylinders of length one, i.e. $h(\mu) = h(\mu, \mathcal{B})$, where $\mathcal{B}$ is the family of sets of the form $\{x \mid x_0 = l\}$ for $l \in \Lambda$.

The space of invariant measures on any full shift (and therefore also a subshift) with the weak-$*$ topology is metrizable, and the metric can be explicitly given as

$$d(\mu, \nu) = \sum_{i=1}^{\infty} \frac{1}{2^i} \sum_{B \in \mathcal{B}_i} |\mu(B) - \nu(B)|,$$

where $\mathcal{B}_i$ is the family of all cylinders of length $i$. An easy and useful corollary of this definition is that for any $\varepsilon > 0$ there exist $\delta > 0$ and $n \in \mathbb{N}$ such that if two measures differ by no more than $\delta$ on all cylinders of length up to $n$, then the distance between the measures is less than $\varepsilon$.

Let $(X, T)$ be a zero-dimensional dynamical system (i.e. one where $X$ is a zero-dimensional space). Any such system can be expressed as an inverse limit

$$(X, T) = \varprojlim_{n \to \infty} (X_k, T_k),$$

where each $(X_k, T_k)$ is a subshift. To see this, consider a sequence of partitions $\mathcal{A}_k$ of $X$ such that $\mathcal{A}_k$ is a partition into clopen sets of diameter at most $1/k$, and $\mathcal{A}_k \prec \mathcal{A}_{k+1}$. Each $\mathcal{A}_k$, as a finite set, can be endowed with discrete topology and used as the alphabet of $X_k$. The system $X_k$ itself will be the set of points $x \in \mathcal{A}_k^{\mathbb{Z}}$ such that there exists $x' \in X$ with $T^n x' \in x_n$ for all $n \in \mathbb{Z}$ (easily seen to be closed and invariant under the shift transformation). The condition $\mathcal{A}_k \prec \mathcal{A}_{k+1}$ implies the existence of a natural factor map $\pi_k : X_{k+1} \to X_k$ by requiring that $\pi_k(x)_n$ be the set from $\mathcal{A}_k$ which contains $x_n$ (which is a set of $\mathcal{A}_{k+1}$. The inverse limit $Y = \varprojlim_{n \to \infty} (X_k, T_k)$ (where each $T_k$ denotes the shift transformation) is naturally a factor of $X$ (since each system $X_k$ is such a factor), but on the other hand different points in $X$ will necessarily have different images in $Y$, since the diameters of the sets in $\mathcal{A}_k$ decrease. Therefore we have a one-to-one continuous mapping between $X$ and $Y$ which commutes with the transformations, and thus the systems are conjugate.

By the *entropy structure* of $X$ we will understand the sequence of functions on $\mathcal{M}_T(X)$ (the set of $T$-invariant Borel probability measures on $X$) given by

$$\mu \mapsto h_k(\mu) = h(\mu_k),$$

where $\mu_k$ is the projection of $\mu$ onto $X_k$.

We will make use of two basic facts regarding the entropy structure. Firstly, its elements converge to the dynamical entropy of the measure:

$$\lim_{k \to \infty} h_k(\mu) = h(\mu).$$

Secondly, the notion of asymptotic $h$-expansiveness can be expressed in terms of the entropy structure, which we will in fact adopt here as the definition of asymptotic $h$-expansiveness:

Definition 2.3. A system $(X,T)$ with entropy structure $(h_k)$ is called *asymptotically $h$-expansive* if the sequence of functions $h_k$ converges to $h$ uniformly on $\mathcal{M}_T(X)$.

See [B-D] for proof that this definition is equivalent to the other commonly used ways of defining asymptotic $h$-expansiveness.

The fact below, proved originally by M. Misiurewicz [M], follows immediately from this definition (which is not to say that it is trivial in itself, the difficulties are simply moved into proving the equivalence of the two definitions of asymptotic $h$-expansiveness):

Theorem 2.4. *If $(X,T)$ is an asymptotically $h$-expansive system, then there exists a measure $\mu^*$ on $X$ such that $h(\mu^*) = \mathbf{h}(T)$.*

Indeed, $h$ is the uniform limit of the functions $h_k$, which are all upper semicontinuous. Therefore $h$ is also upper semicontinuous, and since it is a function on a compact set, it attains its maximum.

**3. The result.** Before we formulate and prove the main theorem, we introduce the following lemma which will be used in the proof:

Lemma 3.1. *Let $Z_n$ be the symbolic dynamical system over the alphabet $\{0,1\}$ in which every occurrence of the symbol $1$ is within a block of at least $N$ consecutive $1$'s. Then the topological entropy of $Z_n$ tends to $0$ as $n$ tends to infinity.*

*Proof.* For $n \in \mathbb{N}$ define the sequence $z^{(n)} = (z_k^{(n)})_{k\in\mathbb{Z}}$ as follows: $z_k^{(n)} = 0$ if $1 \le k \le n$ and $1$ otherwise. In other words, $z^{(n)}$ is the sequence consisting almost entirely of ones, with just one block of zeros, which has length $n$. Let $Z$ be the closure of the union of the orbits of $z^{(n)}$ over $n \in \mathbb{N}$. It is easy to see that the topological entropy of $Z$ is $0$ (the system $Z$ has only two ergodic measures, each concentrated on one of its two fixed points). Since $Z$ and the $Z_n$'s are all subsystems of the full binary shift, their invariant measures are all elements of the same metric space (the set of all invariant measures on the full shift), and thus the notion of proximity of measures on different systems is well-defined. Since the measure-theoretic entropy is an upper semicontinuous function on the set of invariant measures on a symbolic dynamical system, this means that for any $\varepsilon > 0$ there is some $\delta > 0$ such that any measure within the $\delta$-neighborhood of $\mathcal{M}_T(Z)$ has entropy less than $\varepsilon$.

Now, observe that the system $Z$ is in a sense the "limit" of the $Z_n$'s—for large enough $n$ the whole of $Z_n$ is in an arbitrarily small neighborhood of $Z$ (because all blocks of $Z_n$ up to a certain length, which increases with $n$, are also blocks of $Z$). There exists a length $k$ and a number $\eta$ such that if two measures differ by no more than $\eta$ on all blocks of length less than $k$ (by measure of a block we mean the measure of the cylinder defined by that block), then the distance between them is less than $\delta$. There also exists a length $N$ such that for any block $B$ of length $N$ in $Z$ there exists an invariant measure $\mu$ on $Z$ such that any block $A$ of length less than $k$ occurs in $B$ with frequency differing by no more than $\eta/3$ from $\mu(A)$. As stated before, for large enough $n$ every point in $Z_n$ is can be expressed as a concatenation of blocks of length $N$ that all occur in $Z$. It follows that if $B$ now denotes a long enough block from $Z_n$, then there exists an invariant measure $\mu$ on $Z$ such that any block $A$ of length up to $k$ occurs in $B$ with frequency differing by no more than $2\eta/3$ from $\mu(A)$. This, however, means that for any invariant measure $\nu$ on $Z_n$ there exists a measure $\mu$ on $Z$ such that for any block $A$ of length up to $k$, $\nu(A)$ differs from $\mu(A)$ by less than $\eta$. In other words, $\nu$ is in the $\delta$-neighborhood of $\mathcal{M}_T(Z)$, and thus its entropy is less than $\varepsilon$, which concludes the proof. □

Main Theorem 3.2. *Let $(X,T)$ be a zero-dimensional dynamical system. The following two statements are equivalent:*

(1) *There exists a measure $\mu^* \in \mathcal{M}_T(X)$ such that $h(\mu^*,T) = \mathbf{h}(T)$.*
(2) *There exists an asymptotically $h$-expansive system $(Y,S)$, a factor of $(X,T)$, such that $\mathbf{h}(T) = \mathbf{h}(S)$.*

*Proof.* (2)⇒(1). This implication is trivial: $(Y,S)$, being an asymptotically $h$-expansive system, has a measure $\nu^*$ such that $h(\nu^*) = \mathbf{h}(S)$. If we set $\mu^*$ to be any preimage of $\nu^*$ on $X$, we see that $\mathbf{h}(T) = \mathbf{h}(S) = h(\nu^*) \le h(\mu^*) \le \mathbf{h}(T)$, so $\mu^*$ is a measure of maximal entropy on $X$.

(1)⇒(2). Let $\mu^*$ be the maximal entropy measure on $X$. As before, we represent $X$ in the form of an inverse limit

$$(X,T) = \varprojlim_{k\to\infty} (X_k, T_k),$$

where each system $X_k$ is a subshift over a finite alphabet. Denote the projection from $X_k$ onto $X_{k-1}$ by $\pi_k$. Let $\mu_k^*$ be the projection of $\mu^*$ onto $X_k$ and let $h_k^* = h(\mu_k^*)$.

For each $k$ we will define a subshift $Y_k$ such that each $Y_k$ is both a factor of $X_k$ (denote the factor map by $\psi_k$) and an extension of $Y_{k-1}$ (by a map we will call $\rho_k$), and the diagram

$$\begin{array}{ccccccc} X_1 & \xleftarrow{\pi_2} & X_2 & \xleftarrow{\pi_3} & X_3 & \longleftarrow & \cdots \\ {\scriptstyle\psi_1}\downarrow & & {\scriptstyle\psi_2}\downarrow & & {\scriptstyle\psi_3}\downarrow & & \\ Y_1 & \xleftarrow{\rho_2} & Y_2 & \xleftarrow{\rho_3} & Y_3 & \longleftarrow & \cdots \end{array}$$

commutes. Once $Y_k$ is defined, we denote by $\nu_k^*$ the projection of $\mu^*$ under $\psi_k$ from $X_k$.

Let $\varepsilon_k$ be a fixed summable sequence of positive numbers. The systems and maps mentioned above will be constructed inductively, by setting $Y_1 = X_1$ (with the map $\psi_1$ being the identity) and describing how to construct $Y_k$ given $Y_{k-1}$. Furthermore, we will construct our systems in such a way that $h(\nu_k^*) > h(\mu_k^*) - \varepsilon_{k+1}$.

$X_k$ is an extension of $Y_{k-1}$ (via the map $\psi_{k-1} \circ \pi_k$), so it is conjugate to a subsystem of the product $Y_{k-1} \times X_k$ (with the product action) consisting of points of the form $(\psi_{k-1} \circ \pi_k(x), x)$. For our purposes this subsystem is best viewed as a two-row subshift that has sequences from $Y_{k-1}$ in the first row and their preimages from $X_k$ in the second row.

Let $\mathcal{A}_k$ denote the partition into cylinders (of length 1) determined by symbols in the first row only, and $\mathcal{B}_k$ the partition into cylinders of length 1 determined by symbols in both rows. Observe that $h(\mu_k^*, \mathcal{B}_k) = h(\mu_k^*) = h_k^*$ and $h(\mu_k^*, \mathcal{A}_k) = h(\nu_{k-1}^*)$.

Let $\Lambda_k$ be the alphabet of $X_k$ (in its new representation as a two-row subshift) with one additional symbol, which we will label 0, in the second row—we can naturally view $X_k$ as a subshift over this extended alphabet (a subshift in which 0 never occurs in the second row). Let $\delta_k = \varepsilon_{k+1}/\log \#\Lambda_k$. Apply Corollary 2.2, with the constants $\varepsilon_k$ and $\delta_k$, twice: first to the system $(X_k, \mathcal{A}_k, \mu_k^*)$, and then to $(X_k, \mathcal{B}_k, \mu_k^*)$. That way we obtain a number $N_k$ and two collections of cylinders of length $N_k$: a collection $\mathcal{A}'_k \subset \mathcal{A}_k^{N_k}$ and a collection $\mathcal{B}'_k \subset \mathcal{B}_k^{N_k}$, such that the total measure $\mu_k^*$ of both $\mathcal{A}'_k$ and $\mathcal{B}'_k$ is at least $1 - \delta_k/2$, and:

1. For any $B \in \mathcal{B}'_k$,
$$2^{-N_k(h_k^* + \varepsilon_k)} < \mu_k^*(B) < 2^{-N_k(h_k^* - \varepsilon_k)}.$$
2. For any $A \in \mathcal{A}'_k$,
$$2^{-N_k(h(\nu_{k-1}^*) + \varepsilon_k)} < \mu_k^*(A) < 2^{-N_k(h(\nu_{k-1}^*) - \varepsilon_k)}.$$

Using the fact that $h_{k-1}^* > h(\nu_{k-1}^*) > h_{k-1}^* - \varepsilon_k$, we can rewrite the latter estimate as follows:
$$2^{-N_k(h_{k-1}^* + \varepsilon_k)} < \mu_k^*(A) < 2^{-N_k(h_{k-1}^* - 2\varepsilon_k)}.$$
Note that the cylinders $A \in \mathcal{A}^{N_k}$ correspond to blocks of length $N_k$ occurring in the first row of sequences from $X_k$, and similarly cylinders $B \in \mathcal{B}^{N_k}$ correspond to two-row blocks of length $N_k$.

We can also require that $N_k$ be so large that the topological entropy of the system $Z_{N_k}$ (defined in Lemma 3.1) is less than $\varepsilon_k$, and also that $(\log N_k)/N_k < \varepsilon_k$.

Now, pick any $A$ in $\mathcal{A}'_k$ and let $\mathcal{C}_k(A)$ denote the collection of cylinders $B \in \mathcal{B}'_k$ that are subsets of $A$, i.e. two-row blocks from $\mathcal{B}'_k$ that have $A$ in the first row. Any set from $\mathcal{C}_k(A)$ has measure at least $2^{-N_k(h_k^* + \varepsilon_k)}$. They are all disjoint subsets of $A$, which has measure at most $2^{-N_k(h^*_{k-1}-2\varepsilon_k)}$, therefore $\mathcal{C}_k(A)$ has at most $L_k = 2^{N_k(h^*_k - h^*_{k-1}+3\varepsilon_k)}$ elements. Now, let $\mathcal{C}_k = \bigcup_{A\in\mathcal{A}'_k} C_k(A)$. Observe that $\mathcal{C}_k$ contains at most $L_k$ different cylinders that share the first row. Also observe that the total measure of $\mathcal{C}_k$ is at least $1-\delta_k$.

Now, define the map $\psi_k$ on $X_k$ as follows: In any element $x$ of $X_k$, at every position that is not covered by an occurrence of a block from $\mathcal{C}_k$, replace the symbol in the second row with the symbol 0, otherwise make no change. In other words, if for $n \in \mathbb{Z}$ there exist $i \le n \le j$ such that $x_{[i,j]} \in \mathcal{C}_k$, then $(\psi_k(x))_n = x_n$, otherwise $(\psi_k(x))_n = 0$.

Let $Y_k$ be the image of $X_k$ under $\psi_k$. Note that the projection $\rho_k$ of $Y_k$ onto the first row is a factor map onto $Y_{k-1}$, and since $\psi_k$ acts as identity on the first row, the diagram

$$\begin{array}{ccc} X_{k-1} & \xleftarrow{\ \pi_k\ } & X_k \\ {\scriptstyle\psi_{k-1}}\downarrow & & \downarrow{\scriptstyle\psi_k} \\ Y_{k-1} & \xleftarrow{\ \rho_k\ } & Y_k \end{array}$$

commutes. Since $Y_k$ and $X_k$ are symbolic systems over the same alphabet $\Lambda_k$, the partitions $\mathcal{A}_k$ and $\mathcal{B}_k$ defined above can apply to either of them, and we will not make a distinction in notation (it will be obvious from the context).

Let $\mathcal{C}'_k$ be the partition of $X_k$ obtained from $\mathcal{B}_k^{N_k}$ by replacing all sets which do not belong to $\mathcal{C}_k$ with their union, denoted as $C_k$ (recall that $\mu^*_k(C_k) < \delta_k$). Observe that if two points belong to the same cylinder from $\mathcal{C}_k$, then $\psi_k$ changes nothing on their first $N_k$ coordinates, so their images belong to the same cylinder from $\mathcal{B}_k^{N_k}$. On the other hand, the image of any point from $C_k$ is also in $C_k$. Hence

$$\mathcal{C}'_k \prec \psi_k^{-1}(\mathcal{B}_k^{N_k}),$$

and the two partitions are identical (and identical to $\mathcal{B}_k^{N_k}$) on the set $X_k \setminus C_k$. As a result we have the following calculation:

$$\begin{aligned} h(\mu^*_k) - h(\nu^*_k) &= h(\mu^*_k, \mathcal{B}_k) - h(\nu^*_k, \mathcal{B}_k) \\ &= h(\mu^*_k, \mathcal{B}_k) - h(\mu^*_k, \psi_k^{-1}(\mathcal{B}_k)) \\ &\le h(\mu^*_k, \mathcal{B}_k \,|\, \psi_k^{-1}(\mathcal{B}_k)) \le \frac{1}{N_k} H(\mu^*_k, \mathcal{B}_k^{N_k} \,|\, \psi_k^{-1}(\mathcal{B}_k^{N_k})) \\ &\le \frac{1}{N_k} H(\mu^*_k, \mathcal{B}_k^{N_k} \,|\, \mathcal{C}'_k). \end{aligned}$$

Since the partitions $\mathcal{B}_k^{N_k}$ and $\mathcal{C}'_k$ are identical outside of $C_k$, all terms in the conditional entropy formula (2.1) are zero, except the one corresponding to $C_k$. Since $\mathcal{B}_k^{N_k}$ is a partition consisting of at most $(\#\Lambda_k)^{N_k}$ elements, by applying estimate (2.2) we have

$$\begin{aligned}\frac{1}{N_k}H(\mu_k^*,\mathcal{B}_k^{N_k}\,|\,\mathcal{C}_k') &= \frac{1}{N_k}\mu_k^*(C_k)H((\mu_k^*)_{C_k},\mathcal{B}_k^{N_k})\\ &\le \frac{1}{N_k}\mu_k^*(C_k)\log{(\#\Lambda_k)}^{N_k}\\ &= \mu_k^*(C_k)\log\#\Lambda_k = \delta_k\log\#\Lambda_k < \varepsilon_{k+1}.\end{aligned}$$

Ultimately, we see that

$$h(\nu_k^*) > h(\mu_k^*) - \varepsilon_{k+1}.$$

We now set

$$Y = \varprojlim_{k\to\infty} Y_k.$$

It is obviously a factor of $X$; we need to show that it is asymptotically $h$-expansive and has topological entropy equal to that of $X$. Let $\nu^*$ be the projection of $\mu^*$ onto $Y$; observe that $\nu_k^*$ is the projection of $\nu^*$ onto $Y_k$. Since $h(\mu_k^*) \ge h(\nu_k^*) > h(\mu_k^*) - \varepsilon_{k+1}$, by passing to the limit ($h(\mu_k^*)$ and $h(\nu_k^*)$ are functions from the entropy structure of $X$ and $Y$ respectively, and thus they converge to entropies of their preimages in $X$ and $Y$, i.e. to $h(\mu^*)$ and $h(\nu^*)$) we obtain $h(\nu^*) = h(\mu^*) = \mathbf{h}(T)$, therefore $\mathbf{h}(S) = \mathbf{h}(T)$.

It remains to show that $Y$ is asymptotically $h$-expansive, i.e. the function sequence $h_k$ converges uniformly. To this end we will show that the functions $h_k - h_{k-1}$ are bounded on $\mathcal{M}_S(Y)$ by terms of a summable series. Let $k > 0$, $\nu \in \mathcal{M}_S(Y)$, and let $\nu_k$ be the projection of $\nu$ onto $Y_k$. Recalling the two-row representation (and notation) described above, we see that $h_k(\nu) = h(\nu_k, \mathcal{B}_k)$, $h_{k-1}(\nu) = h(\nu_k, \mathcal{A}_k)$, and $h_k(\nu) - h_{k-1}(\nu) = h(\nu_k, \mathcal{B}_k\,|\,\mathcal{A}_k)$. In other words, when estimating $h_k - h_{k-1}$ we can restrict ourselves to the system $Y_k$ with these two partitions.

Fix $M > 0$. Assuming we know the first row of a block $C$ of length $M$, we must calculate how many ways there are of filling the second row. The second row contains two classes of symbols, zeros and non-zeros, the latter constituting second rows of blocks from the collection $\mathcal{C}_k$. We will first assume we know the positions of zeros and count the number of possible versions of the second row of $C$ given these positions. Consider an interval of length $N_k$ in $C$, which we shall label $D$.

Suppose the second row of $D$ contains no zeros. This means that $D$ (a block of length $N_k$) is entirely covered by blocks from $\mathcal{C}_k$ (also of length $N_k$). Let $D_1$ be the rightmost such block that ends within $D$ and let $D_2$ be the leftmost block that begins within $D$. Together, $D_1$ and $D_2$ cover all of $D$, so if we know their second rows, we know the second row of $D$. We know the first row of both $D_1$ and $D_2$, so we are left with at most $L_k$ possibilities of completing the second row of each, and therefore no more than $L_k^2$ possibilities of completing the second row of $D$ as a whole. By looking at the first row alone, we can determine neither where exactly $D_1$ ends, nor where $D_2$ begins, but there are only $N_k$ possible positions for each of them. Therefore we can simply multiply the $L_k^2$ by $N_k^2$ to end up with the maximum number of possible versions of the second row of $D$, which is $L_k^2 N_k^2$.

Suppose, on the other hand, that $D$ contains zeros. In this case, since non-zeros occur in groups of at least $N_k$, $D$ must begin with a sequence of non-zeros (possibly empty), then have any number of zeros, and then end with another (again, possibly empty) sequence of non-zeros. The two sequences of non-zeros form (parts of) second rows of blocks from $\mathcal{C}_k$, and we know their first rows, so we have at most $L_k$ ways of filling each sequence, and thus at most $L_k^2$ ways of filling the second row of $D$.

In either case, we see that if we know the first row of $C$ and the positions of zeros in the second row, then any subinterval of length $N_k$ in $C$ admits at most $L_k^2 N_k^2$ versions of the second row. By dividing $C$ into such disjoint intervals (of which there are fewer than $M/N_k + 1$), we have the following upper estimate on the number of possibilities in the second row:

$$\mathcal{N}'_{k,M} = (L_k^2 N_k^2)^{M/N_k+1}.$$

To estimate the total number of possible versions of the second row, we need to multiply the above by the number of possible patterns of zeros, which we will denote by $\mathcal{N}''_{k,M}$—observe that, since non-zero symbols occur in groups of length at least $N_k$, the number of possible patterns of zeros equals the number of blocks of length $M$ in the system $Z_{N_k}$ (recall that $Z_{N_k}$ was defined as the symbolic dynamical system in which the symbol 1 occurred in groups of length at least $N_k$). Since the topological entropy of $Z_{N_k}$ is strictly smaller than $\varepsilon_k$, for large enough $M$ we can assume that $\mathcal{N}''_{k,M} < 2^{M\varepsilon_k}$.

Therefore we have the following upper estimate for the number of possible versions of the second row, given the first:

$$\mathcal{N}_{k,M} \le \mathcal{N}'_{k,M}\mathcal{N}''_{k,M}.$$

Passing to logarithms we have

$$\log \mathcal{N}_{k,M} \le \log \mathcal{N}'_{k,M} + \log \mathcal{N}''_{k,M} \le 2\left(\frac{M}{N_k} + 1\right)(\log L_k + \log N_k) + M\varepsilon_k.$$

Recall that

$$L_k = 2^{N_k(h_k^* - h_{k-1}^* + 3\varepsilon_k)},$$

so

$$\log L_k = N_k(h_k^* - h_{k-1}^* + 3\varepsilon_k).$$

It follows that

$$\frac{1}{M}\log \mathcal{N}_{k,M} \le \left(\frac{2}{N_k}+\frac{1}{M}\right)(N_k(h_k^*-h_{k-1}^*+3\varepsilon_k)+\log N_k)+\varepsilon_k.$$

The observation that every version of the first row admits at most $\mathcal{N}_{k,M}$ versions of the second row is equivalent to stating that any set from $\mathcal{A}_k^M$ intersects at most $\mathcal{N}_{k,M}$ sets from $\mathcal{B}_k^M$. Applying Lemma 2.1, we see that for any measure $\nu_k$ on $Y_k$ we have

$$\begin{aligned}\frac{1}{M}H(\nu_k,\mathcal{B}_k^M \mid \mathcal{A}_k^M) &< \frac{1}{M}\log \mathcal{N}_{k,M}\\ &\le \left(\frac{2}{N_k}+\frac{1}{M}\right)(N_k(h_k^*-h_{k-1}^*+3\varepsilon_k)+\log N_k)+\varepsilon_k,\end{aligned}$$

and thus in the limit

$$\begin{aligned}h(\nu_k,\mathcal{B}_k \mid \mathcal{A}_k) &= \lim_{M\to\infty}\frac{1}{M}H(\nu_k,\mathcal{B}_k^M \mid \mathcal{A}_k^M)\\ &\le 2(h_k^*-h_{k-1}^*+3\varepsilon_k)+\frac{2\log N_k}{N_k}+\varepsilon_k\\ &< 2(h_k^*-h_{k-1}^*)+9\varepsilon_k.\end{aligned}$$

Therefore

$$\begin{aligned}h_k(\nu)-h_{k-1}(\nu) = h(\nu_k,\mathcal{B}_k)-h(\nu_k,\mathcal{A}_k) &= h(\nu_k,\mathcal{B}_k \mid \mathcal{A}_k)\\ &< 2(h_k^*-h_{k-1}^*)+9\varepsilon_k.\end{aligned}$$

Since the series $h_k^*-h_{k-1}^*$ and $\varepsilon_k$ are both summable, we conclude that the sequence $h_k(\cdot)$ converges uniformly on $\mathcal{M}_S(Y)$ and thus the system $Y$ is asymptotically $h$-expansive. ■

On a final note, Theorem 3.2 also provides a method (albeit slightly unwieldy) of investigating whether a general dynamical system has a maximal entropy measure. By the result of [D-H], any dynamical system $Y$ has a zero-dimensional extension $X$ which is principal, i.e. the conditional entropy of any measure on $Y$ with respect to $X$ is 0. Since $X$ is zero-dimensional, Theorem 3.2 applies, and if we can use it to establish that $X$ has a maximal entropy measure $\mu^*$, then the image of $\mu^*$ on $X$ will be a maximal entropy measure as well. This leaves us with

THEOREM 3.3. *A dynamical system $X$ has a maximal entropy measure if and only if its principal extension $Y$ has an asymptotically $h$-expansive factor of equal topological entropy.*

Note that Theorem 3.2 itself does not apply to non-zero-dimensional systems. Intuitively, this is because they generally do not have as many factors. It is the zero-dimensionality (assumed in Theorem 3.2 and achieved by the principal extension in the general case) that causes the factor structure to be rich enough to ensure the existence of an asymptotically $h$-expansive factor of equal entropy.

**Acknowledgements.** The author would like to thank his PhD advisor, Prof. Tomasz Downarowicz, for all the help, insight and observations which made writing this paper possible.

This research is supported by MENII grant N N201 394537, Poland, for the years 2009–2012.

## *REFERENCES*

[B-D] M. Boyle and T. Downarowicz, *The entropy theory of symbolic extensions*, Invent. Math. 156 (2004), 119–161.

[D] T. Downarowicz, *Entropy in Dynamical Systems*, New Math. Monogr. 18, Cambridge Univ. Press, Cambridge, 2011.

[D-H] T. Downarowicz and D. Huczek, *Zero-dimensional principal extensions*, Acta Appl. Math., to appear.

[M] M. Misiurewicz, *Topological conditional entropy*, Studia Math. 55 (1976), 175–200.

Dawid Huczek
Institute of Mathematics and Computer Science
Wrocław University of Technology
50-372 Wrocław, Poland
E-mail: dawid.huczek@pwr.wroc.pl

*Received 12 September 2011;*
*revised 12 April 2012* (5543)