Symbolic dynamics: entropy = dimension = complexity

Stephen G. Simpson
Department of Mathematics
Vanderbilt University
http://www.math.psu.edu/simpson
sgslogic@gmail.com

First draft: March 17, 2010
This draft: February 16, 2017

Abstract

Let d be a positive integer. Let G be the additive monoid $\mathbb{N}^d$ or the additive group $\mathbb{Z}^d$. Let A be a finite set of symbols. The shift action of G on A^G is given by $S^g(x)(h) = x(g+h)$ for all $g, h \in G$ and all $x \in A^G$. A G-subshift is defined to be a nonempty closed set $X \subseteq A^G$ such that $S^g(x) \in X$ for all $g \in G$ and all $x \in X$. Given a G-subshift X, the topological entropy $\mathrm{ent}(X)$ is defined as usual [31]. The standard metric on A^G is defined by $\rho(x, y) = 2^{-|F_n|}$ where n is as large as possible such that $x{\upharpoonright}F_n = y{\upharpoonright}F_n$. Here $F_n = \{0, 1, \ldots, n\}^d$ if $G = \mathbb{N}^d$, and $F_n = \{-n, \ldots, -1, 0, 1, \ldots, n\}^d$ if $G = \mathbb{Z}^d$. For any $X \subseteq A^G$ the Hausdorff dimension $\dim(X)$ and the effective Hausdorff dimension $\mathrm{effdim}(X)$ are defined as usual [15, 26, 27] with respect to the standard metric. It is well known that $\mathrm{effdim}(X) = \sup_{x\in X} \liminf_n \mathrm{K}(x{\upharpoonright}F_n)/|F_n|$ where K denotes Kolmogorov complexity [10]. If X is a G-subshift, we prove that $\mathrm{ent}(X) = \dim(X) = \mathrm{effdim}(X)$, and $\mathrm{ent}(X) \geq \limsup_n \mathrm{K}(x{\upharpoonright}F_n)/|F_n|$ for all $x \in X$, and $\mathrm{ent}(X) = \lim_n \mathrm{K}(x{\upharpoonright}F_n)/|F_n|$ for some $x \in X$.

Keywords: symbolic dynamics, entropy, Hausdorff dimension, Kolmogorov complexity.

MSC2010 Subject Classification: Primary 37B10, Secondary 37B40, 94A17, 68Q30.

We thank the anonymous referees for comments which led to improvements in this paper. In particular, the proof of Lemma 5.1 given below was suggested by the first referee and is much simpler than our original proof.

This paper has been published in *Theory of Computing Systems*, 56(3):2015, 527–543.

1 Introduction

The purpose of this paper is to elucidate a close relationship among three disparate concepts which are known to play a large role in three diverse branches of contemporary mathematics. The concepts are:

entropy, Hausdorff dimension, Kolmogorov complexity.

Some relationships among these concepts are well known; see for instance [5, 25, 39]. Nevertheless, it seems to us that the full depth of the relationships has been insufficiently appreciated. Below we prove that, in an important special case, all three concepts coincide.

Here is a brief overview of the above-mentioned concepts.

1. *Hausdorff dimension* is a basic concept in metric geometry. See for instance the original paper by Hausdorff [15] and the classic treatise by C. A. Rogers [29]. To any set X in a metric space one assigns a nonnegative real number $\dim(X)$ = the Hausdorff dimension of X. In the case of smooth sets such as algebraic curves and surfaces, the Hausdorff dimension is a nonnegative integer and coincides with other familiar notions of dimension from algebra, differential geometry, etc. For example, the Hausdorff dimension of a smooth surface in n-dimensional space is 2. On the other hand, Hausdorff dimension applies also to non-smooth sets with nonintegral dimension, e.g., fractals and Julia sets [11].

2. *Kolmogorov complexity* plays an important role in information theory [8, 32], theoretical computer science [19, 37], and recursion/computability theory [10, 22]. To each finite mathematical object τ one assigns a nonnegative integer $\mathrm{K}(\tau)$ = the complexity of τ. Roughly speaking, $\mathrm{K}(\tau)$ is

the length in bits of the shortest computer program which describes τ. In this sense $\mathrm{K}(\tau)$ measures the "amount of information" which is inherent in τ.

3. *Entropy* is an important concept in dynamical systems theory [9]. Classically, a dynamical system consists of a set X together with a mapping $T : X \to X$ and one studies the long-term behavior of the orbits $\langle T^n(x) \mid n = 0, 1, 2, \ldots \rangle$ for each $x \in X$. More generally, one considers an action T of a group or semigroup G on a set X, and then the orbit of $x \in X$ is $\langle T^g(x) \mid g \in G \rangle$. The entropy of the system X, T is a nonnegative real number which has a rather complicated definition but is intended to quantify the "exponential growth rate" of the system.

   An especially useful class of dynamical systems are the symbolic systems, a.k.a., subshifts [16, 20, 33, 34]. Given a finite set of symbols A, one defines the shift action of G on A^G as usual. A subshift is then defined to be a closed, shift-invariant subset of A^G. These symbolic systems play a large role in general dynamical systems theory, because for any dynamical system X, T one can consider partitions $\pi : X \to A$ and then the behavior of an orbit $\langle T^g(x) \mid g \in G \rangle$ is reflected by the behavior of its "symbolic trace," $\langle \pi(T^g(x)) \mid g \in G \rangle$, which is a point in A^G.

Our main results in this paper are Theorems 4.2 and 5.3 below. They say the following. Let d be a positive integer, let G be the additive monoid $\mathbb{N}^d$ or the additive group $\mathbb{Z}^d$, let A be a finite set of symbols, and let $X \subseteq A^G$ be a subshift. Then, the entropy of X is equal to the Hausdorff dimension of X with respect to the standard metric on A^G. Moreover, the entropy of X has a sharp characterization in terms of the Kolmogorov complexity of the finite configurations which occur in the orbits of X.

In connection with the characterization of entropy in terms of Kolmogorov complexity, it is interesting to note that both of these concepts originated with A. N. Kolmogorov, but in different contexts [35, 18].

2 How this paper came about

This paper is an outcome of my reading and collaboration over the past several years. Here are some personal comments on that process.

It began with my study of Bowen's alternative definition of topological entropy [3, pages 125–126]. Obviously Bowen's definition resembles the standard definition of Hausdorff dimension in a metric space, and this led me to consider the following question:

> Given a subshift X, what is the precise relationship between the topological entropy of X and the Hausdorff dimension of X?

Specifically, let A be a finite set of symbols. From [3, Proposition 1] it was clear to me that the topological entropy of a one-sided subshift $X \subseteq A^{\mathbb{N}}$ is equal to the Hausdorff dimension of X with respect to the standard metric. And eventually I learned that this result appears explicitly in Furstenberg 1967 [12, Proposition III.1]. But what about other kinds of subshifts on A? For instance, what about the two-sided case, i.e., subshifts in $A^{\mathbb{Z}}$ or $A^{\mathbb{Z}^d}$ or more generally A^G where G is a countable amenable group [38]? And what about the general one-sided case, i.e., subshifts in $A^{\mathbb{N}^d}$ or more generally A^G where G is countable amenable semigroup, whatever that may mean?

During February, March and April of 2010 I discussed these issues with several colleagues: John Clemens, Vaughn Climenhaga [7, Example 4.1], Manfred Denker [9], Michael Hochman [16], Anatole Katok [17], Daniel Mauldin, Yakov Pesin [25], Jan Reimann [26], Alexander Shen [37], Daniel Thompson, Jean-Paul Thouvenot [17]. All of these discussions were extremely helpful. In particular, Hochman and Mauldin provided several ideas which play an essential role in this paper.

3 Background

In this section we present some background material concerning symbolic dynamics, entropy, Hausdorff dimension, and Kolmogorov complexity. All of the concepts and results in this section are well known.

We write

$$\mathbb{N} = \{0, 1, 2, \ldots\} = \{\text{the nonnegative integers}\}$$

and

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} = \{\text{the integers}\}.$$

Throughout this paper, let G be the additive monoid $\mathbb{N}^d$ or the additive group $\mathbb{Z}^d$ where d is a fixed positive integer. An *action* of G on a set X is a mapping $T : G \times X \to X$ such that $T^e(x) = x$ and $T^g(T^h(x)) = T^{g+h}(x)$ for all $g, h \in G$ and all $x \in X$. Here e is the identity element of G. It is useful to write G in a specific[1] way as the union of a sequence of finite sets, namely $G = \bigcup_{n=0}^{\infty} F_n$ where $F_n = \{0, 1, \ldots, n\}^d$ if $G = \mathbb{N}^d$, and $F_n = \{-n, \ldots, -1, 0, 1, \ldots, n\}^d$ if $G = \mathbb{Z}^d$. In particular we have $F_0 = \{0\}^d = \{e\}$. We also write $F_{-1} = \emptyset =$ the empty set. For any finite set F we write $|F| =$ the cardinality of F. For any function Φ we write $\mathrm{dom}(\Phi) =$ the domain of Φ, and $\mathrm{rng}(\Phi) =$ the range of Φ, and

$$\Phi :\subseteq X \to Y$$

meaning that Φ is a function with $\mathrm{dom}(\Phi) \subseteq X$ and $\mathrm{rng}(\Phi) \subseteq Y$. Apart from this, all of our set-theoretic notation is standard.

[1] In particular, the sequence F_n with $n = 0, 1, 2, \ldots$ is a Følner sequence for G.

3.1 Topological entropy

We endow G with the discrete topology. Let X be a nonempty compact set in a topological space, and let $T : G \times X \to X$ be a continuous action of G on X. The ordered pair X, T is called a *compact dynamical system*. We now define the topological entropy of X, T.

An *open cover of* X is a set $\mathcal{U}$ of open sets such that $X \subseteq \bigcup \mathcal{U}$. In this case we write

$$C(X,\mathcal{U}) = \min\{|\mathcal{F}| \mid \mathcal{F} \subseteq \mathcal{U}, X \subseteq \bigcup \mathcal{F}\}.$$

Note that $C(X,\mathcal{U})$ is a positive integer. If $\mathcal{U}$ and $\mathcal{V}$ are open covers of X, then

$$\sup(\mathcal{U},\mathcal{V}) = \{U \cap V \mid U \in \mathcal{U}, V \in \mathcal{V}\}$$

is again an open cover of X, and

$$C(X, \sup(\mathcal{U},\mathcal{V})) \leq C(X,\mathcal{U})C(X,\mathcal{V}).$$

For each $g \in G$ and each open cover $\mathcal{U}$ of X, we have another open cover $\mathcal{U}^g = (T^g)^{-1}(\mathcal{U}) = \{(T^g)^{-1}(U) \mid U \in \mathcal{U}\}$. Hence, for each finite set $F \subset G$ we have an open cover $\mathcal{U}^F = \sup\{\mathcal{U}^g \mid g \in F\}$. Let us write $C(X,T,\mathcal{U},F) = C(X,\mathcal{U}^F)$. Note that $C(X,\mathcal{U}^g) \leq C(X,\mathcal{U})$, hence $C(X,\mathcal{U}^F) \leq C(X,\mathcal{U})^{|F|}$, hence $\log_2 C(X,T,\mathcal{U},F) \leq |F| \log_2 C(X,\mathcal{U})$. We define

$$\mathrm{ent}(X,T,\mathcal{U}) = \lim_{n\to\infty} \frac{\log_2 C(X,T,\mathcal{U},F_n)}{|F_n|} \tag{1}$$

and

$$\mathrm{ent}(X,T) = \sup\{\mathrm{ent}(X,T,\mathcal{U}) \mid \mathcal{U} \text{ is an open cover of } X\}.$$

The nonnegative real number $\mathrm{ent}(X,T)$ is known as the *topological entropy*[2] of X, T. It measures what might be called the "asymptotic exponential growth rate" of X, T. See for instance [9, 21, 31].

Lemma 3.1. The limit in equation (1) exists.

Proof. Let us write $C_n = C(X,T,\mathcal{U},F_n)$. Clearly $C_m \leq C_n$ whenever $m \leq n$. Moreover, it is easy to see that $C_{nk} \leq C_n^{k^d}$ for all positive integers k. We are trying to prove that $\log_2 C_n/|F_n|$ approaches a limit as $n \to \infty$. Assume $G = \mathbb{Z}^d$, so that $|F_n| = (2n+1)^d$. (The case $G = \mathbb{N}^d$ is similar, with $|F_n| = (n+1)^d$.)

Fix a positive integer m. Given $n \geq m$, let k be a positive integer such that $mk \leq n < m(k+1)$. We have $|F_n| \geq |F_{mk}|$ and

$$\frac{|F_{mk}|}{k^d|F_m|} = \left(\frac{2mk+1}{2mk+k}\right)^d > \left(\frac{2m}{2m+1}\right)^d$$

[2] Instead of $\log_2$ we could use $\log_b$ for any fixed $b > 1$, for instance $b = e$ or $b = 10$. The base $b = 2$ is convenient for information theory, where entropy is measured in bits.

and $\log_2 C_n \le \log_2 C_{m(k+1)} \le (k+1)^d \log_2 C_m$, hence

$$\frac{\log_2 C_n}{|F_n|} \le \frac{(k+1)^d \log_2 C_m}{|F_{mk}|} \le \frac{(k+1)^d \log_2 C_m}{k^d |F_m|} \left(\frac{2m+1}{2m}\right)^d .$$

As $n \to \infty$ we have $k \to \infty$, hence

$$\limsup_{n\to\infty} \frac{\log_2 C_n}{|F_n|} \le \frac{\log_2 C_m}{|F_m|} \left(\frac{2m+1}{2m}\right)^d ,$$

and this holds for all m, hence

$$\limsup_{n\to\infty} \frac{\log_2 C_n}{|F_n|} \le \liminf_{m\to\infty} \frac{\log_2 C_m}{|F_m|} .$$

In other words, $\lim_{n\to\infty} \log_2 C_n/|F_n|$ exists, Q.E.D. $\square$

Let $\mathcal{U}$ and $\mathcal{V}$ be open covers of X. We say that $\mathcal{U}$ *refines* $\mathcal{V}$ if each $U \in \mathcal{U}$ is included in some $V \in \mathcal{V}$. Obviously this implies $C(X,\mathcal{U}) \ge C(X,\mathcal{V})$, and it is also easy to see that $\mathrm{ent}(X,T,\mathcal{U}) \ge \mathrm{ent}(X,T,\mathcal{V})$.

Lemma 3.2. For each m we have $\mathrm{ent}(X,T,\mathcal{U}^{F_m}) = \mathrm{ent}(X,T,\mathcal{U})$.

Proof. Clearly $\mathcal{U}^{F_m}$ refines $\mathcal{U}$, hence $\mathrm{ent}(X,T,\mathcal{U}^{F_m}) \ge \mathrm{ent}(X,T,\mathcal{U})$. For all n we have $F_{m+n} = F_m + F_n$, hence $\mathcal{U}^{F_{m+n}} \subseteq (\mathcal{U}^{F_m})^{F_n}$, hence $\mathcal{U}^{F_{m+n}}$ refines $(\mathcal{U}^{F_m})^{F_n}$, hence $C(X,\mathcal{U}^{F_{m+n}}) \ge C(X,(\mathcal{U}^{F_m})^{F_n})$, hence $C(X,T,\mathcal{U},F_{m+n}) \ge C(X,T,\mathcal{U}^{F_m},F_n)$, hence

$$\frac{\log_2 C(X,T,\mathcal{U},F_{m+n})}{|F_n|} \ge \frac{\log_2 C(X,T,\mathcal{U}^{F_m},F_n)}{|F_n|} ,$$

hence

$$\frac{\log_2 C(X,T,\mathcal{U},F_{m+n})}{|F_{m+n}|} \cdot \frac{|F_{m+n}|}{|F_n|} \ge \frac{\log_2 C(X,T,\mathcal{U}^{F_m},F_n)}{|F_n|} .$$

Taking the limit as $n \to \infty$ and noting that

$$\lim_{n\to\infty} \frac{|F_{m+n}|}{|F_n|} = 1 ,$$

we see that $\mathrm{ent}(X,T,\mathcal{U}) \ge \mathrm{ent}(X,T,\mathcal{U}^{F_m})$. This completes the proof. $\square$

Let $\mathcal{U}$ be an open cover of X. We say that $\mathcal{U}$ is a *topological generator* if for each open cover $\mathcal{V}$ of X there exists m such that $\mathcal{U}^{F_m}$ refines $\mathcal{V}$. The following theorem says that we can use a topological generator to compute $\mathrm{ent}(X,T)$.

Theorem 3.3. If $\mathcal{U}$ is a topological generator, then $\mathrm{ent}(X,T) = \mathrm{ent}(X,T,\mathcal{U})$.

Proof. Let $\mathcal{V}$ be an open cover of X. Since $\mathcal{U}$ is a topological generator, let m be such that $\mathcal{U}^{F_m}$ refines $\mathcal{V}$. Then $\mathrm{ent}(X,T,\mathcal{U}^{F_m}) \ge \mathrm{ent}(X,T,\mathcal{V})$, so by Lemma 3.2 we have $\mathrm{ent}(X,T,\mathcal{U}) \ge \mathrm{ent}(X,T,\mathcal{V})$. Thus $\mathrm{ent}(X,T,\mathcal{U}) = \mathrm{ent}(X,T)$. $\square$

3.2 Symbolic dynamics

An important class of dynamical systems are the *symbolic* dynamical systems, also known as *subshifts*. We now present some background material on subshifts. See also [20, §13.10] and [4, 16, 33, 34].

As before, let d be a positive integer, and let G be the additive monoid $\mathbb{N}^d$ or the additive group $\mathbb{Z}^d$. Let A be a nonempty finite set of symbols. We endow A with the discrete topology. Let $A^G = \{x \mid x : G \to A\}$. We endow A^G with the product topology. Note that each $x \in A^G$ is a function from G to A. For each finite set $F \subset G$ and each $x \in A^G$ let $x{\upharpoonright}F$ be the restriction of x to F. Thus $A^F = \{x{\upharpoonright}F \mid x \in A^G\}$. For each $\sigma \in A^F$ we write $\mathrm{dom}(\sigma) = F$ and $|\sigma| = |F|$ and $[\![\sigma]\!] = \{x \in A^G \mid x{\upharpoonright}F = \sigma\}$. Note that $[\![\sigma]\!]$ is a nonempty clopen set in A^G, and $\{[\![\sigma]\!] \mid \sigma \in A^F\}$ is a pairwise disjoint covering of A^G. Let $A^* = \bigcup_{n=0}^{\infty} A^{F_n}$ and note that $\{[\![\sigma]\!] \mid \sigma \in A^*\}$ is a basis for the topology of A^G. For any $T \subseteq A^*$ we write $[\![T]\!] = \bigcup_{\sigma \in T} [\![\sigma]\!]$. Thus $[\![T]\!]$ is an open set in A^G.

The *shift action* of G on A^G is the mapping $S : G \times A^G \to A^G$ given by $S^g(x)(h) = x(g+h)$ for all $g, h \in G$ and all $x \in A^G$. Thus A^G, S is a compact dynamical system, known as the *full shift*. Since $F_0 = \{0\}^d$ is a singleton set, there is an obvious one-to-one correspondence between A^{F_0} and A, so we identify A^{F_0} with A. The *canonical open covering* of A^G is $\mathcal{U} = \mathcal{U}(A, G) = \{[\![a]\!] \mid a \in A\}$. For each finite set $F \subset G$ we have $\mathcal{U}^F = \{[\![\sigma]\!] \mid \sigma \in A^F\}$. By compactness of A^G it follows that $\mathcal{U}$ is a topological generator. Moreover $C(A^G, S, \mathcal{U}, F) = C(A^G, \mathcal{U}^F) = |\mathcal{U}^F| = |A^F| = |A|^{|F|}$, so by Theorem 3.3 we have $\mathrm{ent}(A^G, S) = \mathrm{ent}(A^G, S, \mathcal{U}) = \log_2 |A|$.

A set $X \subseteq A^G$ is said to be *shift-invariant* if $S^g(x) \in X$ for all $g \in G$ and all $x \in X$. A *subshift* is a nonempty, closed, shift-invariant subset of A^G. Each subshift $X \subseteq A^G$ gives rise to a compact dynamical system $X, S{\upharpoonright}G \times X$. We write $\mathrm{ent}(X) = \mathrm{ent}(X, S{\upharpoonright}G \times X)$, etc. Since $\mathcal{U}^F$ is a pairwise disjoint covering of A^G, we have $C(X, \mathcal{U}, F) = C(X, \mathcal{U}^F) = |X{\upharpoonright}F|$ where $X{\upharpoonright}F = \{x{\upharpoonright}F \mid x \in X\}$. Since $\mathcal{U}$ is a topological generator, it follows by Theorem 3.3 that

$$\mathrm{ent}(X) = \lim_{n\to\infty} \frac{\log_2 |X{\upharpoonright}F_n|}{|F_n|} . \tag{2}$$

Lemma 3.4. We have

$$\lim_{n\to\infty} |X{\upharpoonright}F_n| 2^{-s|F_n|} = \begin{cases} 0 & \text{if } s > \mathrm{ent}(X)\,, \\ \infty & \text{if } s < \mathrm{ent}(X)\,. \end{cases}$$

Proof. First suppose $s > \mathrm{ent}(X)$. Fix $\epsilon > 0$ such that $s - \epsilon > \mathrm{ent}(X)$. Equation (2) implies that for all sufficiently large n we have $(s-\epsilon)|F_n| > \log_2 |X{\upharpoonright}F_n|$, hence $|X{\upharpoonright}F_n| 2^{-|F_n|s} < 2^{-\epsilon|F_n|}$. Letting $n \to \infty$ we have $|F_n| \to \infty$, hence $\lim_{n\to\infty} |X{\upharpoonright}F_n| 2^{-s|F_n|} = 0$.

Next, suppose $s < \mathrm{ent}(X)$. Fix $\epsilon > 0$ such that $s + \epsilon < \mathrm{ent}(X)$. Equation (2) implies that for all sufficiently large n we have $(s+\epsilon)|F_n| < \log_2 |X{\upharpoonright}F_n|$, hence $|X{\upharpoonright}F_n| 2^{-|F_n|s} > 2^{\epsilon|F_n|}$. Letting $n \to \infty$ we have $|F_n| \to \infty$, hence $\lim_{n\to\infty} |X{\upharpoonright}F_n| 2^{-s|F_n|} = \infty$. □

3.3 Hausdorff dimension

Let X be a set in a metric space. The *s-dimensional Hausdorff measure* of X is defined as

$$\mu_s(X) \;=\; \lim_{\epsilon\to 0}\,\inf_{\mathcal{E}}\,\sum_{E\in\mathcal{E}}\mathrm{diam}(E)^s$$

where $\mathrm{diam}(E)$ is the diameter of E. Here $\mathcal{E}$ ranges over coverings of X with the property that $\mathrm{diam}(E)\le\epsilon$ for all $E\in\mathcal{E}$. The *Hausdorff dimension* of X is

$$\dim(X)=\inf\{s\mid \mu_s(X)=0\}.$$

Hausdorff measures and Hausdorff dimension have been widely studied, e.g., in connection with the geometry of fractals [11, 15, 29].

We now define what we mean by the Hausdorff dimension of a subshift. The *standard metric* on A^G is given by $\rho(x,y)=2^{-|F_n|}$ where $n=-1,0,1,2,\ldots$ is as large as possible such that $x{\restriction}F_n=y{\restriction}F_n$. (Recall that $F_{-1}=\emptyset$.) Clearly the standard metric on A^G induces the product topology on A^G. Moreover, the standard metric is an *ultrametric*, i.e., $\rho(x,y)\le\max(\rho(x,z),\rho(y,z))$ for all x,y,z. For any set $X\subseteq A^G$ we define $\dim(X)=$ the Hausdorff dimension of X with respect to the standard metric on A^G.

Lemma 3.5. For all subshifts $X\subseteq A^G$ we have $\mathrm{ent}(X)\ge\dim(X)$.

Proof. For each $E\subseteq A^G$ we have $\mathrm{diam}(E)\le 2^{-|F_n|}$ if and only if $E\subseteq[\![\sigma]\!]$ for some $\sigma\in A^{F_n}$. Therefore, in the definition of $\mu_s(X)$ and $\dim(X)$ for an arbitrary set $X\subseteq A^G$, we may safely assume that each E is a basic open set, i.e., $E=[\![\sigma]\!]$ for some $\sigma\in A^*$. Moreover, for each $\sigma\in A^*$ we have $\mathrm{diam}([\![\sigma]\!])=2^{-|\sigma|}$.

Assume now that X is a subshift, and suppose $s>\mathrm{ent}(X)$. By Lemma 3.4 we have

$$\lim_{n\to\infty}|X{\restriction}F_n|2^{-|F_n|s}=0\,. \tag{3}$$

But for each n we have $X\subseteq\bigcup_{x\in X}[\![x{\restriction}F_n]\!]$ and $\mathrm{diam}([\![x{\restriction}F_n]\!])=2^{-|F_n|}$, so (3) implies that $\mu_s(X)=0$, hence $s\ge\dim(X)$. Since this holds for all $s>\mathrm{ent}(X)$, it follows that $\mathrm{ent}(X)\ge\dim(X)$. □

Remark 3.6. In §4 we shall prove that for all subshifts $X\subseteq A^G$, $\mathrm{ent}(X)=\dim(X)$. In other words, the topological entropy of a subshift is equal to its Hausdorff dimension with respect to the standard metric. While the special case $G=\mathbb{N}$ is due to Furstenberg [12, Proposition III.1], the general result for $G=\mathbb{N}^d$ or $G=\mathbb{Z}^d$ appears to be new.

3.4 Kolmogorov complexity

We now present some background material on Kolmogorov complexity.

As in §3.2 let $A^*=\bigcup_{n=0}^\infty A^{F_n}$. In addition let $\{0,1\}^*$ be the set of finite sequences of 0's and 1's. For each Turing machine M and each finite sequence $\alpha\in\{0,1\}^*$, let $M(\alpha)$ be the run of M with input α. A function

$$\Phi :\subseteq \{0,1\}^* \to A^*$$

is said to be *partial computable* if there exists a Turing machine M such that for all $\alpha \in \{0,1\}^*$, $\alpha \in \mathrm{dom}(\Phi)$ if and only if $M(\alpha)$ eventually halts, in which case it halts with output $\Phi(\alpha)$. For each such Φ and each $\xi \in A^*$ let

$$\mathrm{K}_\Phi(\xi) = \min(\{|\alpha| \mid \Phi(\alpha) = \xi\} \cup \{\infty\}).$$

A partial computable function $\Psi :\subseteq \{0,1\}^* \to A^*$ is said to be *universal* if for each partial computable function $\Phi :\subseteq \{0,1\}^* \to A^*$ there exists a constant c such that for all $\xi \in A^*$ we have $\mathrm{K}_\Psi(\xi) \leq \mathrm{K}_\Phi(\xi) + c$. The existence of such a universal function is easily proved. Fix such a universal function Ψ. For each $\xi \in A^*$ we define the *Kolmogorov complexity* of ξ to be $\mathrm{K}(\xi) = \mathrm{K}_\Psi(\xi)$. Note that $\mathrm{K}(\xi)$ is well defined up to an additive constant, i.e., up to $\pm\, O(1)$. Here "well defined" means that $\mathrm{K}(\xi)$ is independent of the choice of Ψ.

Remark 3.7. Actually the complexity notion K defined above is only one of several variant notions, denoted in [37] as KP, KS, KM, KA, KD. These variants are useful in many contexts [10]. However, for our purposes in this paper, the differences among them are immaterial.

3.5 Effective Hausdorff dimension

We now present some background material concerning the effective or computable variant of Hausdorff dimension. Throughout this paper the words "effective" and "computable" refer to Turing's theory of computability and unsolvability [30, 36].

A *Polish space* is a complete separable metric space. An *effectively presented Polish space* consists of a Polish space Z, ρ together with a mapping $\Phi : \mathbb{N} \to Z$ such that $\mathrm{rng}(\Phi)$ is dense in Z, ρ and the real-valued function $(m, n) \mapsto \rho(\Phi(m), \Phi(n)) : \mathbb{N} \times \mathbb{N} \to [0, \infty)$ is computable. In this case we define the *basic open sets* of Z, ρ, Φ to be those of the form

$$B(n, r) = \{x \in Z \mid \rho(\Phi(n), x) < r\}$$

where r is a positive rational number and $n \in \mathbb{N}$. A sequence of basic open sets B_i, $i = 1, 2, \ldots$ is said to be *computable* if there exist computable sequences n_i, r_i, $i = 1, 2, \ldots$ such that $B_i = B(n_i, r_i)$ for all i. A set $X \subseteq Z$ is said to be *effectively closed* if its complement $Z \setminus X$ is *effectively open*, i.e., $Z \setminus X = \emptyset$ or $Z \setminus X = \bigcup_{i=1}^\infty B_i$ where B_i, $i = 1, 2, \ldots$ is a computable sequence of basic open sets. We say that X is *effectively compact* if it is effectively closed and *effectively totally bounded*, i.e., there exists a computable function $f : \mathbb{N} \to \mathbb{N}$ such that $X \subseteq \bigcup_{n=1}^{f(i)} B(n, 2^{-i})$ for each i.

Let s be a positive real number. We say that X is *effectively s-null* if there exists a computable double sequence of basic open sets B_{ij}, $i, j = 1, 2, \ldots$, such that $X \subseteq \bigcup_{j=1}^\infty B_{ij}$ and $\sum_{j=1}^\infty \mathrm{diam}(B_{ij})^s \leq 2^{-i}$ for each i. The *effective Hausdorff dimension of* X is defined as

$$\mathrm{effdim}(X) \;=\; \inf\{s \mid X \text{ is effectively } s\text{-null}\}.$$

Note that, although the Hausdorff dimension of a singleton point $\{x\}$ is always 0, there may be no computable way to "observe" this, so the effective Hausdorff dimension of a noncomputable point may be > 0. In fact, for any set X one has

$$\text{effdim}(X) \;=\; \sup_{x\in X} \text{effdim}(\{x\})\,. \tag{4}$$

On the other hand, it is known that $\text{effdim}(X) = \dim(X)$ provided X is effectively compact. See for instance [10, Chapter 13] and [26, 27, 28].

The above definitions and remarks apply to the effectively compact, effectively presented[3] Polish space A^G with the standard metric as defined in §3.3. In particular we have $\text{effdim}(X) = \dim(X)$ for all effectively closed sets $X \subseteq A^G$. In §5 below we shall prove that $\text{effdim}(X) = \dim(X)$ for all subshifts $X \subseteq A^G$. This result holds even if X is not effectively closed.

For arbitrary subsets of A^G, the following theorem exhibits a relationship between effective Hausdorff dimension and Kolmogorov complexity. We shall see in Theorem 5.3 that the relationship is even closer when X is a subshift.

Theorem 3.8 (Mayordomo's Theorem)**.** For any set $X \subseteq A^G$ we have

$$\text{effdim}(X) \;=\; \sup_{x\in X} \liminf_{n\to\infty} \frac{\mathrm{K}(x{\restriction}F_n)}{|F_n|}\,.$$

Proof. This follows from (4) together with [10, Theorem 13.3.4]. □

3.6 Measure-theoretic entropy

We now present some background material on measure-theoretic entropy. We state two important theorems without proof but with references to the literature.

Let X, μ be a probability space. An action $T : G \times X \to X$ is said to be *measure-preserving* if $\mu((T^g)^{-1}(P)) = \mu(P)$ for each $g \in G$ and each μ-measurable set $P \subseteq X$. In this case the ordered triple X, T, μ is called a *measure-theoretic dynamical system*. We now proceed to define the measure-theoretic entropy of X, T, μ.

A *measurable partition* of X is a finite set $\mathcal{P}$ of pairwise disjoint μ-measurable subsets of X such that $X = \bigcup \mathcal{P}$. In this case we write

$$H(X,\mu,\mathcal{P}) = -\sum_{P\in\mathcal{P}} \mu(P)\log_2 \mu(P)\,.$$

If $\mathcal{P}$ and $\mathcal{Q}$ are measurable partitions of X, then

$$\sup(\mathcal{P},\mathcal{Q}) = \{P \cap Q \mid P \in \mathcal{P}, Q \in \mathcal{Q}\}$$

[3]Our Φ for A^G is obtained as follows. Let $\# : A^* \to \mathbb{N}$ be a *standard Gödel numbering* of A^*. In other words, for each $\sigma \in A^*$ let $\#(\sigma)$ be a numerical code for σ from which σ can be effectively recovered. Let a be a fixed symbol in A. Define $\Phi : \mathbb{N} \to A^G$ by letting $\Phi(\#(\sigma)) = x_\sigma \in A^G$ where $x_\sigma \in [\![\sigma]\!]$ and $x_\sigma(g) = a$ for all $g \in G \setminus \text{dom}(\sigma)$.

is again a measurable partition of X, and it can be shown [9, 10.4(d)] that

$$H(X, \mu, \sup(\mathcal{P}, \mathcal{Q})) \leq H(X, \mu, \mathcal{P}) + H(X, \mu, \mathcal{Q})\,. \tag{5}$$

For each $g \in G$ and each measurable partition $\mathcal{P}$ of X, we have another measurable partition $\mathcal{P}^g = (T^g)^{-1}(\mathcal{P}) = \{(T^g)^{-1}(P) \mid P \in \mathcal{P}\}$. Hence, for each finite set $F \subset G$ we have a measurable partition $\mathcal{P}^F = \sup\{\mathcal{P}^g \mid g \in F\}$. Let us write $H(X, T, \mu, \mathcal{P}, F) = H(X, \mu, \mathcal{P}^F)$. It follows from (5) that $H(X, T, \mu, \mathcal{P}, F) \leq |F| H(X, \mu, \mathcal{P})$. We define

$$\mathrm{ent}(X, T, \mu, \mathcal{P}) = \lim_{n\to\infty} \frac{H(X, T, \mu, \mathcal{P}, F_n)}{|F_n|} \tag{6}$$

and

$$\mathrm{ent}(X, T, \mu) = \sup\{\mathrm{ent}(X, T, \mu, \mathcal{P}) \mid \mathcal{P} \text{ is a measurable partition of } X\}.$$

It can be proved that the limit in (6) exists. The nonnegative real number $\mathrm{ent}(X, T, \mu)$ is known as the *measure-theoretic entropy* of X, T, μ. It plays an important role in ergodic theory. See for instance [9, 21, 24].

Let X, T, μ be a measure-theoretic dynamical system. A set $P \subseteq X$ is said to be *G-invariant* if $(T^g)^{-1}(P) \subseteq P$ for all $g \in G$. The system X, T, μ is said to be *ergodic* if for every G-invariant μ-measurable set $P \subseteq X$ we have $\mu(P) = 0$ or $\mu(P) = 1$.

Now let d be a positive integer, let $G = \mathbb{N}^d$ or $\mathbb{Z}^d$, let A be a nonempty finite set of symbols, and let $X \subseteq A^G$ be a subshift. A Borel probability measure μ on X is said to be *shift-invariant* if $\mu((S^g)^{-1}(P)) = \mu(P)$ for each $g \in G$ and each Borel set $P \subseteq X$. In this case X, S, μ is a measure-theoretic dynamical system, and we write $H(X, \mu, \mathcal{P}) = H(X, S, \mu, \mathcal{P})$, $\mathrm{ent}(X, \mu) = \mathrm{ent}(X, S, \mu)$, etc. As in §3.2 it can be shown that $\mathrm{ent}(X, \mu) = \mathrm{ent}(X, \mu, \mathcal{P})$ where $\mathcal{P}$ is the *canonical measurable partition* of X, namely $\mathcal{P} = \{[\![a]\!] \cap X \mid a \in A\}$.

In the case of an ergodic subshift, there is the following suggestive characterization of measure-theoretic entropy.

Theorem 3.9 (Shannon/McMillan/Breiman)**.** Let $X \subseteq A^G$ be a subshift, and let μ be an ergodic, shift-invariant, probability measure on X. Then for μ-almost all $x \in X$ we have

$$\mathrm{ent}(X, \mu) = \lim_{n\to\infty} \frac{\log_2 \mu([\![x{\upharpoonright}F_n]\!])}{-|F_n|}\,.$$

Proof. See [23]. □

We end this section by noting a significant relationship between topological entropy and measure-theoretic entropy.

Theorem 3.10 (Variational Principle)**.** For any subshift $X \subseteq A^G$ we have

$$\mathrm{ent}(X) = \max_{\mu} \mathrm{ent}(X, \mu)$$

where μ ranges over ergodic, shift-invariant, probability measures on X.

Proof. See [21] and [9, §§16–20]. □

4 Entropy = dimension

As in §3 let d be a positive integer, let $G = \mathbb{N}^d$ or $G = \mathbb{Z}^d$, let A be a finite set of symbols, and let $X \subseteq A^G$ be a subshift. The purpose of this section is to prove that $\mathrm{ent}(X) = \dim(X)$. The special case $G = \mathbb{N}$ is due to Furstenberg [12, Proposition III.1]. However, the general result for $G = \mathbb{N}^d$ or $G = \mathbb{Z}^d$ appears to be new.

As a warm-up for our proof of the general result, we first present Furstenberg's proof of the special case $G = \mathbb{N}$.

Theorem 4.1 (Furstenberg 1967)**.** Let $X \subseteq A^{\mathbb{N}}$ be a one-sided subshift. Then $\mathrm{ent}(X) = \dim(X)$.

Proof. By Lemma 3.5 we have $\mathrm{ent}(X) \geq \dim(X)$. To prove $\mathrm{ent}(X) \leq \dim(X)$ it suffices to prove $\mathrm{ent}(X) \leq s$ for all s such that $\mu_s(X) = 0$. Since $\mu_s(X) = 0$ let $\mathcal{E}$ be such that $X \subseteq \bigcup \mathcal{E}$ and $\sum_{E \in \mathcal{E}} \mathrm{diam}(E)^s < 1$. As noted in §3.3, we may safely assume that each $E \in \mathcal{E}$ is of the form $E = [\![\sigma]\!]$ where $\sigma \in A^*$, so that $\mathrm{diam}(E) = 2^{-|\sigma|}$. By compactness we may assume that $\mathcal{E}$ is finite. Let us write $\mathcal{E} = \{[\![\sigma]\!] \mid \sigma \in I\}$ where $I \subset A^*$ is finite. Let $m = \max\{|\sigma| \mid \sigma \in I\}$. From

$$\sum_{\sigma \in I} 2^{-|\sigma| s} = \sum_{E \in \mathcal{E}} \mathrm{diam}(E)^s < 1$$

it follows that

$$\sum_{\sigma_1, \ldots, \sigma_k} 2^{-(|\sigma_1| + \cdots + |\sigma_k|)s} = \sum_{k=1}^{\infty} \left(\sum_{\sigma \in I} 2^{-|\sigma| s} \right)^k = M < \infty$$

where the first sum is taken over all nonempty finite sequences $\sigma_1, \ldots, \sigma_k \in I$.

The previous paragraph applies to any subshift. We now bring in the special assumption $G = \mathbb{N}$. Because $G = \mathbb{N}$ and $F_n = \{0, 1, \ldots, n\}$, each $x \in A^G$ is an infinite sequence of symbols in A, and each $\sigma \in A^* = \bigcup_{n=0}^{\infty} A^{F_n}$ is a nonempty finite sequence of symbols in A. Thus, given $x \in X$, we can recursively define an infinite sequence $\sigma_1, \ldots, \sigma_k, \ldots \in I$ such that $S^{|\sigma_1| + \cdots + |\sigma_{k-1}|}(x) \in [\![\sigma_k]\!]$ for all k, and then $x = \sigma_1{}^\frown \cdots {}^\frown \sigma_k{}^\frown \cdots$ where ${}^\frown$ denotes concatenation of finite sequences. Now, given $n \geq 0$, let k be as small as possible such that $x{\restriction}F_n \subseteq \sigma_1{}^\frown \cdots {}^\frown \sigma_k$. We then have

$$|F_n| \leq |\sigma_1| + \cdots + |\sigma_k| < |F_n| + m \tag{7}$$

and $[\![x{\restriction}F_n]\!] \supseteq [\![\sigma_1{}^\frown \cdots {}^\frown \sigma_k]\!]$. Since the sets $[\![\xi]\!]$ for $\xi \in X{\restriction}F_n$ are pairwise disjoint, it follows that $|X{\restriction}F_n|$ is less than or equal to the number of finite sequences $\sigma_1, \ldots, \sigma_k \in I$ such that (7) holds. For each such finite sequence we have $2^{-|F_n|s} < 2^{ms} 2^{-(|\sigma_1| + \cdots + |\sigma_k|)s}$, so by summing over all such finite sequences we obtain $|X{\restriction}F_n| 2^{-|F_n|s} < 2^{ms} M$. Thus $|X{\restriction}F_n| 2^{-|F_n|s}$ is bounded as $n \to \infty$. It follows by Lemma 3.4 that $\mathrm{ent}(X) \leq s$, Q.E.D. □

We now generalize Furstenberg's result.

Theorem 4.2. Let $G = \mathbb{N}^d$ or $G = \mathbb{Z}^d$ where d is a positive integer. Let A be a finite nonempty set of symbols, and let $X \subseteq A^G$ be a subshift. Then $\mathrm{ent}(X) = \dim(X)$.

Proof. By Lemma 3.5 we have $\mathrm{ent}(X) \geq \dim(X)$. To prove $\mathrm{ent}(X) \leq \dim(X)$ it suffices to prove $\mathrm{ent}(X) \leq s$ for all s such that $\mu_s(X) = 0$. Using $\mu_s(X) = 0$ and the compactness of X, we can find finite sets $I_l \subset A^*$ for $l = 1, 2, \ldots$ such that $X \subseteq \bigcup_{\sigma \in I_l} [\![\sigma]\!]$ and $\sum_{\sigma \in I_l} 2^{-|\sigma| s} < 2^{-l}$ and $|\sigma| << |\tau|$ for all $\sigma \in I_l$ and all $\tau \in I_{l+1}$. Let $I_\infty = \bigcup_{l=1}^\infty I_l$. We have

$$\sum_{\sigma \in I_\infty} 2^{-|\sigma| s} < \sum_{l=1}^{\infty} 2^{-l} = 1$$

hence

$$\sum_{\sigma_1, \ldots, \sigma_k} 2^{-(|\sigma_1| + \cdots + |\sigma_k|)s} = \sum_{k=1}^{\infty} \left(\sum_{\sigma \in I_\infty} 2^{-|\sigma| s} \right)^k = M < \infty$$

where the first sum is taken over all nonempty finite sequences $\sigma_1, \ldots, \sigma_k \in I_\infty$.

For all $\sigma \in A^*$ and all $g \in G$, let $\sigma^g =$ the *g-translate* of σ, i.e., $\mathrm{dom}(\sigma^g) = \{g + h \mid h \in \mathrm{dom}(\sigma)\}$ and $\sigma^g(g+h) = \sigma(h)$ for all $h \in \mathrm{dom}(\sigma)$. Note that $|\sigma^g| = |\sigma|$ and $[\![\sigma^g]\!] = [\![\sigma]\!]^g = (S^g)^{-1}([\![\sigma]\!])$. Since X is a subshift and $X \subseteq \bigcup_{\sigma \in I_l} [\![\sigma]\!]$ for all l, we have

$$\forall l \, (\forall g \in G) \, (\forall x \in X) \, (\exists \sigma \in I_l) \, (x \in [\![\sigma^g]\!]).$$

Let $J_\infty = \bigcup_{l=1}^\infty J_l$ where $J_l = \{\sigma^g \mid \sigma \in I_l, g \in G\}$.

Lemma 4.3. Let $\epsilon > 0$ be given. For all sufficiently large n and each $x \in X$, we can find a pairwise disjoint set $L \subset J_\infty$ such that $\bigcup L \subseteq x{\upharpoonright}F_n$ and $|\bigcup L| > (1 - \epsilon)|F_n|$ and $|L| < \epsilon |F_n|$.

Proof. The proof may be viewed as a discrete analog of the classical proof of the Vitali Covering Lemma. Given an "extremely large" configuration $x{\upharpoonright}F_n$, we begin by filling in as much of $x{\upharpoonright}F_n$ as possible with pairwise disjoint "very very large" configurations from J_∞. After that, we fill in the gaps with pairwise disjoint "very large" configurations from J_∞. After that, we fill in the remaining gaps with pairwise disjoint "large" configurations from J_∞. Et cetera.

Specifically, let l be so large that $(1 - (1/4)^d)^l < \epsilon$ and $1 < \epsilon |\sigma|$ for all $\sigma \in I_l$, and let n be so large that $n >> |\sigma|$ for all $\sigma \in I_{2l-1}$. Given $x \in X$, let $\xi = x{\upharpoonright}F_n$ and let $K_1 = \{\tau \in J_{2l-1} \mid \tau \subset \xi\}$. Note that $|\bigcup K_1| \geq (3/4)^d |\xi|$, because $|\tau| << n$ for all $\tau \in J_{2l-1}$. Let $L_1 \subseteq K_1$ be pairwise disjoint[4] such that $|\bigcup L_1| \geq |\bigcup K_1|/3^d$. It follows that $|\bigcup L_1| \geq |\xi|/4^d$, hence $|\xi \setminus \bigcup L_1| \leq (1 - (1/4)^d)|\xi|$. If $|\xi \setminus \bigcup L_1| \leq (1 - (1/4)^d)^2 |\xi|$, let $L_2 = K_2 = \emptyset$. Otherwise, let $K_2 = \{\tau \in J_{2l-2} \mid \tau \subset \xi \setminus \bigcup L_1\}$ and note that $|\bigcup K_2| \geq (3/4)^d |\xi \setminus \bigcup L_1|$, because $|\tau| << |\upsilon|$ for all

[4]Here are the details. Define $L_1 = \{\upsilon_j \mid j = 1, 2, \ldots\}$ where $\upsilon_j \in K_1$ is chosen inductively so that $\upsilon_i \cap \upsilon_j = \emptyset$ for all $i < j$ and $|\upsilon_j|$ is as large as possible. Then for all $\tau \in K_1$ there exists $\upsilon \in L_1$ such that $\tau \cap \upsilon \neq \emptyset$ and $|\tau| \leq |\upsilon|$. From this it follows that $|\bigcup L_1| \geq |\bigcup K_1|/3^d$.

$\tau \in J_{2l-2}$ and all $v \in L_1$. As before let $L_2 \subseteq K_2$ be pairwise disjoint such that $|\bigcup L_2| \geq |\bigcup K_2|/3^d$. It follows as before that $|\xi \setminus \bigcup(L_1 \cup L_2)| \leq (1-(1/4)^d)^2|\xi|$. Continuing in this fashion for l steps, we obtain $L_1 \subseteq J_{2l-1}$ and $L_2 \subseteq J_{2l-2}$ and $\ldots$ and $L_l \subseteq J_l$ such that $|\xi \setminus \bigcup(L_1 \cup \cdots \cup L_l)| \leq (1-(1/4)^d)^l|\xi|$. Finally let $L = L_1 \cup \cdots \cup L_l$. By construction L is pairwise disjoint and $\bigcup L \subseteq \xi$. Moreover $|\xi \setminus \bigcup L| \leq (1-(1/4)^d)^l|\xi| < \epsilon|\xi| = \epsilon|F_n|$, hence $|\bigcup L| > (1-\epsilon)|\xi| = (1-\epsilon)|F_n|$. For each $\tau \in L$ we have $1 < \epsilon|\tau|$, hence $|L| < \epsilon|\bigcup L| \leq \epsilon|F_n|$. This proves Lemma 4.3. □

Lemma 4.4. Let ϵ and n be as in Lemma 4.3. Then $|X{\restriction}F_n|$ is less than or equal to $(|A|+1)^{2\epsilon|F_n|}$ times the number of sequences $\sigma_1, \ldots, \sigma_k \in I_\infty$ such that $|\sigma_k| + \cdots + |\sigma_k| \leq |F_n|$.

Proof. The idea of the proof is that, by Lemma 4.3, each $x{\restriction}F_n \in X{\restriction}F_n$ is almost entirely covered by a finite sequence of pairwise disjoint translates of elements of I_∞. These elements of I_∞ can be used to give a concise description of $x{\restriction}F_n$.

Given $x \in X$ let $L = \{\tau_1, \ldots, \tau_k\}$ be as in the conclusion of Lemma 4.3. For each $i = 1, \ldots, k$ let $\sigma_i \in I_\infty$ be such that $\tau_i = \sigma_i^g$ for some $g \in G$. Since $\tau_1, \ldots, \tau_k$ are pairwise disjoint and $\bigcup_{i=1}^k \tau_i = \bigcup L \subseteq x{\restriction}F_n$, we have $|\sigma_1| + \cdots + |\sigma_k| = |\tau_1| + \cdots + |\tau_k| \leq |F_n|$. Let $<_{\mathrm{lex}}$ be the lexicographical ordering of F_n. For each $i = 1, \ldots, k$ let $g_i =$ the least element of $\mathrm{dom}(\tau_i) \subseteq F_n$ with respect to $<_{\mathrm{lex}}$. Reordering $\tau_1, \ldots, \tau_k$ as necessary, we may assume that $g_1 <_{\mathrm{lex}} \cdots <_{\mathrm{lex}} g_k$. Let $U = F_n \setminus \bigcup_{i=1}^k \mathrm{dom}(\tau_i)$, and let $V = U \cup \{g_1, \ldots, g_k\}$. By Lemma 4.3 we have $|U| = |F_n| - |\bigcup L| < \epsilon|F_n|$ and $k = |L| < \epsilon|F_n|$, hence $|V| = |U| + k \leq m$ where $m = 2\lfloor\epsilon|F_n|\rfloor$. For each $j = 1, \ldots, m$ define $a_j \in A \cup \{0\}$ as follows. If $j \leq |V|$ let g be the jth element of V with respect to $<_{\mathrm{lex}}$. If $g \in U$, let $a_j = x(g)$. Otherwise, let $a_j = 0$. Clearly $x{\restriction}F_n$ can be recovered from the pair of sequences $a_1, \ldots, a_m$ and $\sigma_1, \ldots, \sigma_k$. This proves Lemma 4.4. □

To prove Theorem 4.2, let ϵ and n be as in Lemmas 4.3 and 4.4. Because $|\sigma_1| + \cdots + |\sigma_k| \leq |F_n|$ implies $2^{-|F_n|s} \leq 2^{-(|\sigma_1| + \cdots + |\sigma_k|)s}$, it follows from Lemma 4.4 and the definition of M that

$$|X{\restriction}F_n| 2^{-|F_n|s} < (|A|+1)^{2\epsilon|F_n|} M\,,$$

i.e.,

$$|X{\restriction}F_n| 2^{-|F_n|(s+2\epsilon\log_2(|A|+1))} < M\,.$$

Thus $|X{\restriction}F_n| 2^{-|F_n|(s+2\epsilon\log_2(|A|+1))}$ is bounded as n goes to infinity, so by Lemma 3.4 we have $\mathrm{ent}(X) \leq s + 2\epsilon\log_2(|A|+1)$. And this holds for all $\epsilon > 0$, so $\mathrm{ent}(X) \leq s$. The proof of Theorem 4.2 is now complete. □

5 Dimension = complexity

As before let d be a positive integer, let $G = \mathbb{N}^d$ or $G = \mathbb{Z}^d$, let A be a finite set of symbols, and let $X \subseteq A^G$ be a subshift. In this section we prove that the Hausdorff dimension of X is equal to the effective Hausdorff dimension of

X. In addition we obtain a sharp characterization of $\dim(X)$ in terms of the Kolmogorov complexity of finite pieces of the individual orbits of X, i.e., in terms of $\mathrm{K}(x{\restriction}F_n)$ for $x \in X$ and $n = 1, 2, \ldots$. Our results apply even when X is not effectively closed.

Lemma 5.1. For all $x \in X$ we have

$$\limsup_{n\to\infty} \frac{\mathrm{K}(x{\restriction}F_n)}{|F_n|} \;\leq\; \mathrm{ent}(X)\,. \tag{8}$$

Proof. Fix a positive integer m. Given $n \geq m$, let k be a positive integer such that $mk \leq n < m(k+1)$. Partitioning $F_{m(k+1)}$ into $(k+1)^d$ blocks of size $|F_m|$, we see that $|X{\restriction}F_n| \leq (k+1)^d |X{\restriction}F_m|$ and there is a constant c independent of n such that $\mathrm{K}(x{\restriction}F_n) \leq (k+1)^d \log_2 |X{\restriction}F_m| + 2\log_2 n + c$ for all $x \in X$. Thus

$$\frac{\mathrm{K}(x{\restriction}F_n)}{|F_n|} \leq \frac{(k+1)^d \log_2 |X{\restriction}F_m| + 2\log_2 n + c}{k^d |F_m|} \to \frac{\log_2 |X{\restriction}F_m|}{|F_m|}$$

as $n \to \infty$. Since this holds for all m, we now see that (8) follows from (2). $\square$

Lemma 5.2. For some $x \in X$ we have

$$\lim_{n\to\infty} \frac{\mathrm{K}(x{\restriction}F_n)}{|F_n|} \;=\; \mathrm{ent}(X)\,. \tag{9}$$

Proof. By the Variational Principle 3.10 let μ be an ergodic, shift-invariant, probability measure on X such that $\mathrm{ent}(X,\mu) = \mathrm{ent}(X)$. Fix $s < \mathrm{ent}(X)$. Let

$$D_n = \left\{\xi \in A^{F_n} \;\middle|\; \mathrm{K}(\xi) < |F_n| s\right\}\,.$$

Clearly $|D_n| \leq 2^{|F_n|s}$. Fix $\epsilon > 0$ such that $s + \epsilon < \mathrm{ent}(X)$, and let

$$T_n = \{\xi \in A^{F_n} \mid \mu([\![\xi]\!]) < 2^{-|F_n|(s+\epsilon)}\}\,.$$

The Shannon/McMillan/Breiman Theorem 3.9 tell us that for μ-almost all $x \in X$ and all sufficiently large n we have

$$\frac{\log_2 \mu([\![x{\restriction}F_n]\!])}{-|F_n|} > s + \epsilon\,,$$

i.e., $x{\restriction}F_n \in T_n$, i.e., $x \in [\![T_n]\!]$. On the other hand, for each n we have

$$\mu([\![D_n]\!] \cap [\![T_n]\!]) = \mu([\![D_n \cap T_n]\!]) \leq 2^{|F_n|s} 2^{-|F_n|(s+\epsilon)} = 2^{-|F_n|\epsilon}$$

and so

$$\sum_{n=1}^{\infty} \mu([\![D_n]\!] \cap [\![T_n]\!]) < \infty\,.$$

Thus the Borel/Cantelli Lemma tells us that, for μ-almost all x and all sufficiently large n, $x \notin [\![D_n]\!] \cap [\![T_n]\!]$. But then it follows that, for μ-almost all x and all sufficiently large n, $x \notin [\![D_n]\!]$, i.e., $x{\restriction}F_n \notin D_n$, i.e., $\mathrm{K}(x{\restriction}F_n) \geq |F_n| s$. Since this holds for all $s < \mathrm{ent}(X)$, we now see that (9) holds for μ-almost all $x \in X$. This completes the proof. $\square$

Theorem 5.3. Let $G = \mathbb{N}^d$ or $G = \mathbb{Z}^d$ where d is a positive integer. Let A be a finite set of symbols, and let $X \subseteq A^G$ be a subshift. Then

$$\mathrm{ent}(X) = \dim(X) = \mathrm{effdim}(X)\,.$$

Moreover

$$\dim(X) \geq \limsup_{n\to\infty} \frac{\mathrm{K}(x{\restriction}F_n)}{|F_n|}$$

for all $x \in X$, and

$$\dim(X) = \lim_{n\to\infty} \frac{\mathrm{K}(x{\restriction}F_n)}{|F_n|}$$

for some $x \in X$.

Proof. This follows from Theorems 3.8 and 4.2 and Lemmas 5.1 and 5.2. □

Questions 5.4.

1. Can we find an "elementary" or "direct" proof of Lemma 5.2? I.e., a proof which does not use measure-theoretic entropy?

2. Is it possible to generalize Theorems 4.2 and 5.3 so as to apply to wider classes of groups or semigroups? For example, do Theorems 4.2 and 5.3 continue to hold if G is an amenable group [38]?

3. Is it possible to generalize Theorems 4.2 and 5.3 so as to apply to scaled entropy and scaled Hausdorff dimension? For example, what about

$$\liminf_{n\to\infty} \frac{\mathrm{K}(x{\restriction}F_n)}{\sqrt{|F_n|}}\ ?$$

References

[1] M. Baaz, S.-D. Friedman, and J. Krajíček, editors. *Logic Colloquium '01: Proceedings of the Annual European Summer Meeting of the Association for Symbolic Logic, held in Vienna, Austria, August 6–11, 2001.* Number 20 in Lecture Notes in Logic. Association for Symbolic Logic, 2005. VIII + 486 pages.

[2] S. Bezuglyi and S. Kolyada, editors. *Topics in Dynamics and Ergodic Theory.* Number 310 in London Mathematical Society Lecture Note Series. Cambridge University Press, 2003. VIII + 262 pages.

[3] Rufus Bowen. Topological entropy for noncompact sets. *Transactions of the American Mathematical Society*, 184:125–136, 1973.

[4] Mike Boyle. Open problems in symbolic dynamics. In [6], pages 69–118, 2008.

[5] A. A. Brudno. Entropy and the complexity of the trajectories of a dynamical system. *Transactions of the Moscow Mathematical Society*, 2:127–151, 1983.

[6] K. Burns, D. Dolgopyat, and Y. Pesin, editors. *Geometric and Probabilistic Structures in Dynamics*. Number 469 in Contemporary Mathematics. American Mathematical Society, 2008. XVI + 340 pages.

[7] Vaughn Climenhaga. Bowen's equation in the non-uniform setting. *Ergodic Theory and Dynamical Systems*, 31(4):1163–1182, 2011.

[8] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 2nd edition, 2006. XXIII + 748 pages.

[9] Manfred Denker, Christian Grillenberger, and Karl Sigmund. *Ergodic Theory on Compact Spaces*. Number 527 in Lecture Notes in Mathematics. Springer-Verlag, 1976. IV + 360 pages.

[10] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic Randomness and Complexity*. Theory and Applications of Computability. Springer, 2010. XXVIII + 855 pages.

[11] Kenneth Falconer. *Fractal Geometry*. Wiley, 2nd edition, 2003. XXVIII + 337 pages.

[12] Harry Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Mathematical Systems Theory*, 1(1):1–49, 1967.

[13] G. Gallavotti, M. Keane, and D. Ruelle, editors. *International Conference on Dynamical Systems in Mathematical Physics*. Number 40 in Astérisque. Société Mathématique de France, 1976. IV + 192 pages.

[14] S. S. Goncharov, R. Downey, and H. Ono, editors. *Mathematical Logic in Asia: Proceedings of the 9th Asian Logic Conference, Novosibirsk*. World Scientific Publishing Company, Ltd., 2006. VIII + 328 pages.

[15] Felix Hausdorff. Dimension und äußeres Maß. *Mathematische Annalen*, 79(1–2):157–179, 1919.

[16] Michael Hochman and Tom Meyerovitch. A characterization of the entropies of multidimensional shifts of finite type. *Annals of Mathematics*, 171(3):2011–2038, 2010.

[17] Anatole Katok and Jean-Paul Thouvenot. Slow entropy type invariants and smooth realization of commuting measure-preserving transformations. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 33(3):323–338, 1997.

[18] A. N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information and Transmission*, 1(1):1–7, 1965.

[19] Ming Li and Paul Vitányi. *An Introduction to Kolmogorov Complexity and its Applications*. Graduate Texts in Computer Science. Springer-Verlag, 2nd edition, 1997. XX + 637 pages.

[20] Douglas Lind and Brian Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, 1995. XVI + 495 pages.

[21] Michał Misiurewicz. A short proof of the variational principle for a $\mathbb{Z}_+^N$-action on a compact space. In [13], pages 147–157, 1976.

[22] André Nies. *Computability and Randomness*. Oxford University Press, 2009. XV + 433 pages.

[23] Donald Ornstein and Benjamin Weiss. The Shannon-McMillan-Breiman theorem for a class of amenable groups. *Israel Journal of Mathematics*, 44(1):53–60, 1983.

[24] Donald Ornstein and Benjamin Weiss. Entropy and isomorphism theorems for actions of amenable groups. *Journal d'Analyse Mathématique*, 48(1):1–141, 1987.

[25] Yakov B. Pesin. *Dimension Theory in Dynamical Systems*. Chicago Lectures in Mathematics. University of Chicago Press, 1997. XI + 304 pages.

[26] Jan Reimann. *Computability and Fractal Dimension*. PhD thesis, University of Heidelberg, 2004. XI + 130 pages.

[27] Jan Reimann and Frank Stephan. Effective Hausdorff dimension. In [1], pages 369–385, 2005.

[28] Jan Reimann and Frank Stephan. On hierarchies of randomness tests. In [14], pages 215–232, 2006.

[29] C. A. Rogers. *Hausdorff Measures*. Cambridge University Press, 1970. VIII + 179 pages.

[30] Hartley Rogers, Jr. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, 1967. XIX + 482 pages.

[31] David Ruelle. Statistical mechanics on a compact set with $\mathbb{Z}^\nu$ action satisfying expansiveness and specification. *Transactions of the American Mathematical Society*, 185:237–251, 1973.

[32] Claude E. Shannon and Warren Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, 1949. V + 117 pages.

[33] Paul C. Shields. *The Ergodic Theory of Discrete Sample Paths*. Number 13 in Graduate Studies in Mathematics. American Mathematical Society, 1996. XI + 249 pages.

[34] Stephen G. Simpson. Medvedev degrees of 2-dimensional subshifts of finite type. *Ergodic Theory and Dynamical Systems*, 34(2):665–674, 2014. http://dx.doi.org/10.1017/etds.2012.152.

[35] Yakov Sinai. Kolmogorov-Sinai entropy. *Scholarpedia*, 4(3):2034, 2009. doi:10.4249/scholarpedia.2034.

[36] A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42:230–265, 1936.

[37] V. A. Uspensky and A. Shen. Relations between varieties of Kolmogorov complexity. *Mathematical Systems Theory*, 29(3):271–292, 1996.

[38] Benjamin Weiss. Actions of amenable groups. In [2], pages 226–262, 2003.

[39] Homer S. White. Algorithmic complexity of points in dynamical systems. *Ergodic Theory and Dynamical Systems*, 13(4):807–830, 1993.