Ergod. Th. & Dynam. Sys. (2018), **38**, 540–565 doi:10.1017/etds.2016.45
© *Cambridge University Press, 2016*

Entropy and its variational principle for locally compact metrizable systems

ANDRÉ CALDAS and MAURO PATRÃO

Departamento de Matemática, Universidade de Brasília-DF, Brasil
(e-mail: mauropatrao@gmail.com)

(*Received* 6 *November* 2015 *and accepted in revised form* 12 *April* 2016)

Abstract. For a given topological dynamical system $T : X \to X$ over a compact set X with a metric d, the *variational principle* states that

$$\sup_{\mu} h_{\mu}(T) = h(T) = h_d(T),$$

where $h_{\mu}(T)$ is the Kolmogorov–Sinai entropy with the supremum taken over every T-invariant probability measure, $h_d(T)$ is the Bowen entropy and $h(T)$ is the topological entropy as defined by Adler, Konheim and McAndrew. In Patrão [Entropy and its variational principle for non-compact metric spaces. *Ergod. Th. & Dynam. Sys*. **30** (2010), 1529–1542], the concept of topological entropy was adapted for the case where T is a proper map and X is locally compact separable and metrizable, and the variational principle was extended to

$$\sup_{\mu} h_{\mu}(T) = h(T) = \min_{d} h_d(T),$$

where the minimum is taken over every distance compatible with the topology of X. In the present work, we drop the properness assumption and extend the above result for any continuous map T. We also apply our results to extend some previous formulae for the topological entropy of continuous endomorphisms of connected Lie groups that were proved in Caldas and Patrão [Dynamics of endomorphisms of Lie groups. *Discrete Contin. Dyn. Syst*. **33** (2013). 1351–1363]. In particular, we prove that any linear transformation $T : V \to V$ over a finite-dimensional vector space V has null topological entropy.

1. *Introduction*

In this paper, we extend the *variational principle* for entropies to continuous maps $T : X \to X$ defined over a metrizable locally compact separable space X. We have adapted the classical proof of the variational principle due to Misiurewicz (see [**12**, Theorem 8.6]).

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

The original *topological entropy* was defined by Adler *et al* (see [**1**]), by adapting the measure theoretic definition, which we call *Kolmogorov–Sinai entropy*. In this work, this will be called the *AKM entropy*. By assuming that the underlying space is equipped with a distance function, Bowen and Dinaburg (see [**3**, **6**]) have defined a different concept of entropy, which we shall call *Bowen entropy*. In the case of a compact metrizable space, this concept of entropy coincides with the *AKM entropy*. In particular, in the compact case, the value of the *Bowen entropy* does not depend on the particular distance chosen. For a given topological dynamical system $T : X \to X$ over a compact metric space (X, d), Dinaburg and Bowen showed that

$$\sup_{\mu} h_{\mu}(T) = h(T) = h_d(T),$$

where the supremum is taken over every T-invariant probability measure. Probably for this reason, Bowen himself also named his entropy *topological entropy*. However, for non-compact spaces, the Bowen concept of entropy gives different values for equivalent distances over the same space, while the *AKM entropy* would always be infinite. Later, Patrão realized that, even though the value of the *Bowen entropy* depends on the function of the particular distance chosen, by adapting the *AKM entropy* to what we are calling *topological entropy* in this paper, a variational principle relating the measure theoretic *Kolmogorov–Sinai entropy*, the *topological entropy* and the *Bowen entropy* could be demonstrated, as long as the dynamical system T could be extended to a metrizable one point compactification of X (see [**10**]).

Inspired by the work of Adler *et al* in [**1**] and the work of Patrão in [**10**], we extend the definition of topological entropy for arbitrary topological spaces. Inspired by the work of Bowen in [**3**], we simplify his definition of entropy $h_d(T)$ for non-compact sets. We call it *d-entropy* and denote it with a superscript by $h^d(T)$. Then we show, for the case where X is locally compact metrizable and separable, that

$$\sup_{\mu} h_{\mu}(T) = h(T) = \min_{d} h_d(T) = \min_{d} h^d(T),$$

where the minimum is attained for every metric d that can be extended to the one point compactification of X (see Theorem 3.1). Notice that while the metric d might come from the one point compactification, this does not mean that the dynamical system T itself needs to have an extension to the one point compactification. In fact, T can be any continuous system over a locally compact separable metrizable space X. This is a substantial improvement compared with Patrão's preview result in [**10**]. In achieving such a result, it was fundamental that we were able to restate the different definitions of entropy using a unified approach that allowed easier comparisons between them.

We also apply our results to extend some previous formulae for the topological entropy of continuous endomorphisms of connected Lie groups that were proved in [**5**]. In the case of a connected semisimple Lie group, without assuming the endomorphism to be surjective, we prove that its topological entropy always vanishes. In the case of a connected compact Lie group, a linear connected reductive Lie group or a connected nilpotent Lie group, we prove (without assuming the endomorphism to be surjective) that the topological entropy coincides with the topological entropy of the endomorphism's restriction to the maximal connected and compact subgroup of the center of the maximal subgroup where

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

the endomorphism is surjective. In particular, Proposition 4.5 implies that any linear transformation $T : V \to V$ over a finite-dimensional vector space V has null topological entropy. This extends the result of [**10**, Proposition 4.2].

The paper is organized in the following way. In §2, we recall some elementary definitions related to the different types of entropy, extend those definitions to a broader class of dynamical systems and prove some fundamental facts which are used in the subsequent work. In §3, we prove our main result. Finally, in §4, we apply the main result in order to determine the topological entropy for endomorphisms of some classes of Lie groups.

2. *Preliminaries*

This section is devoted to recalling some elementary definitions related to different types of entropy and to proving some fundamental facts which are used in the subsequent work. We also extend the *topological entropy*—originally defined for compact systems—to an arbitrary topological dynamical system.

A *topological dynamical system*—or simply a *dynamical system*—$T : X \to X$ is a continuous map T defined over a topological space X. A *measurable dynamical system* $T : X \to X$ is a measurable map T defined over a measurable space X. If we embed X with the Borel σ-algebra, a *topological dynamical system* becomes also a *measurable dynamical system*.

Recall that a family $\mathscr{A}$ of subsets of X is a *cover of* X, or simply a *cover*, when

$$X = \bigcup_{A \in \mathscr{A}} A.$$

If the sets in $\mathscr{A}$ are disjoint, then we say that $\mathscr{A}$ is a *partition of* X. A *subcover* of $\mathscr{A}$ is a family $\mathscr{B} \subset \mathscr{A}$ which is itself a cover of X. If $Y \subset X$ and $\mathscr{A}$ is a cover of X, then we denote by $Y \cap \mathscr{A}$ the cover of Y given by

$$Y \cap \mathscr{A} = \{A \cap Y \mid A \in \mathscr{A}\}.$$

We denote by $N(\mathscr{A})$ the least cardinality amongst the subcovers of $\mathscr{A}$. For $Y \subset X$, $N_Y(\mathscr{A})$ is a shorthand for $N(Y \cap \mathscr{A})$.

Given two covers $\mathscr{A}$ and $\mathscr{B}$ of an arbitrary set X, we say that $\mathscr{A}$ is *finer* than $\mathscr{B}$ or that $\mathscr{A}$ *refines* $\mathscr{B}$—and write $\mathscr{B} \prec \mathscr{A}$—when every element of $\mathscr{A}$ is a subset of some element of $\mathscr{B}$. We also say that $\mathscr{B}$ is *coarser* than $\mathscr{A}$. The relation $\prec$ is a *preorder*, and if we identify the *symmetric* covers (i.e. covers $\mathscr{A}$ and $\mathscr{B}$ such that $\mathscr{A} \prec \mathscr{B}$ and $\mathscr{B} \prec \mathscr{A}$), we have a *lattice*. As usual, $\mathscr{A} \vee \mathscr{B}$ denotes the representative of the coarsest covers of X that refines both $\mathscr{A}$ and $\mathscr{B}$, given by

$$\mathscr{A} \vee \mathscr{B} = \{A \cap B \mid A \in \mathscr{A},\ B \in \mathscr{B},\ A \cap B \neq \emptyset\}.$$

Given a dynamical system $T : X \to X$ and a cover $\mathscr{A}$, for each $n \in \mathbb{N}$ we define

$$\mathscr{A}^n = \mathscr{A} \vee T^{-1}(\mathscr{A}) \vee \cdots \vee T^{-(n-1)}(\mathscr{A}).$$

If we want to emphasise the dynamical system T, we write $\mathscr{A}^n_T$, instead.

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

2.1. *Compactification.* The variational principle has been demonstrated for compact dynamical systems. For our extended version, we shall treat the dynamical system $T : X \to X$ as a subsystem of a compact metrizable one (Definition 2.1).

For a topological space X to be contained as a subspace in its one point compactification $X \cup \{\infty\}$, and in order for this one point compactification to be metrizable, it is necessary and sufficient that X is a metrizable locally compact separable space. This is a consequence of the Urysohn metrization theorem (see, for example, [**4**, Theorems 9.10, 11.2 and 11.3]).

Definition 2.1. (Subsystem) We say that a (topological) dynamical system $T : X \to X$ is a *(topological) subsystem* of $S : Z \to Z$ when $X \subset Z$ has the induced topology and $T(x) = S(x)$ for every $x \in X$. We also say that S *extends* T to Z. If, instead, $T : X \to X$ and $S : Z \to Z$ are measurable dynamical systems and if X is a measurable subset of Z, we say that T is a *measurable subsystem* of S.

LEMMA 2.2. *Suppose that $T : X \to X$ is a measurable subsystem of $S : Z \to Z$. If $\mathscr{Z}$ is a covering of Z and $\mathscr{C} = X \cap \mathscr{Z}$, then*

$$\mathscr{C}_T^n = X \cap \mathscr{Z}_S^n.$$

Proof. The case $n = 1$ is trivial. Assume it is true for n. Since $T^{-n}(\mathscr{C}) = X \cap S^{-n}(\mathscr{Z})$,

$$\begin{aligned}\mathscr{C}_T^{n+1} &= \mathscr{C}_T^n \cap T^{-n}(\mathscr{C}) \\ &= (X \cap \mathscr{Z}_S^n) \cap (X \cap S^{-n}(\mathscr{Z})) \\ &= X \cap \mathscr{Z}_S^{n+1}.\end{aligned}$$

□

LEMMA 2.3. *Let X be a topological space with metrizable one point compactification $X^* = X \cup \{\infty\}$. Then any topological dynamical system $T : X \to X$ can be regarded as a subsystem of a dynamical system $S : Z \to Z$, with Z compact metrizable and such that the natural projection*

$$\pi : Z \to X^*$$
$$x \mapsto \pi(x) = \begin{cases} x & \text{if } x \in X, \\ \infty & \text{if } x \notin X, \end{cases}$$

is continuous.

Proof. Consider the compact and metrizable space

$$\widetilde{Z} = \prod_{n=0}^{\infty} X^*.$$

Identify X with a subset of $\widetilde{Z}$ by the injection

$$\iota : X \to Z$$
$$x \mapsto (T^n x)_{n=0}^{\infty}.$$

If we let $\pi_n : \widetilde{Z} \to X^*$ be the projection onto the nth coordinate, it is easy to see that ι is continuous, because $\pi_n \circ \iota = T^n$ is continuous for every $n \in \mathbb{N}$. Also, restricted to its

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

image, ι has continuous inverse π_0. So X is homeomorphic to $\iota(X)$. If we identify X and $\iota(X)$, T is just the restriction to X of the shift

$$\begin{aligned} S: \quad & \widetilde{Z} \to \widetilde{Z} \\ & (x_n)_{n=0}^{\infty} \mapsto (x_{n+1})_{n=0}^{\infty}. \end{aligned}$$

Now let Z be the closure of $\iota(X)$ in $\widetilde{Z}$. Since S is continuous and

$$S(\iota(X)) = \iota(T(X)) \subset \iota(X),$$

by taking the closure,

$$S(Z) = S(\overline{\iota(X)}) \subset \overline{S(\iota(X))} \subset \overline{\iota(X)} = Z.$$

That is, $S : Z \to Z$ is a dynamical system that can be regarded as an extension of T.

To see that π is continuous, we shall demonstrate that, in fact, $\pi = \pi_0$. It is evident that, for $x \in X$,

$$\pi(\iota(x)) = x$$

is exactly the projection in the first coordinate. On the other hand, for $a = (a_j)_{j\in\mathbb{N}} \in Z \setminus \iota(X)$, we just have to show that $a_0 = \infty$. But, in fact, since $Z = \overline{\iota(X)}$, there exists $x_n \in X$ such that $\iota(x_n) \to a$: that is, $T^j(x_n) \to a_j$. In particular, $x_n \to a_0$. Now, if $a_0 \neq \infty$, the continuity of T^j implies that

$$a_j = \lim_{n\to\infty} T^j(x_n) = T^j(a_0) :$$

that is, $a = \iota(a_0)$. □

The projection π, on Lemma 2.3, induces the pseudometric

$$\widetilde{d}(x, y) = d(\pi(x), \pi(y))$$

over Z. We denote the restriction of this distance to X by the same letter d. Since π is continuous, this pseudometric is such that the 'open balls' are, in fact, open. Let $X^c = Z \setminus X$. Then, these balls are also such that they either contain X^c or have empty intersection with it.

A consequence of Lemma 2.3 is that X is open in Z. In fact, $X^c = \pi^{-1}(\infty)$ is closed. It follows that $X \subset Z$ is a Borel set and the Borel sets of X are Borel sets of Z. We may restrict Borel measures over Z to the σ-algebra of Borel sets of X. On the other hand, if μ is a Borel measure over X, then we can extend it to Z by declaring $\mu(X^c) = 0$, or, equivalently, we might define $\mu(C) = \mu(C \cap X)$ for every Borel set $C \subset Z$. We shall use the same letter to denote the measure over Z and its restriction to X or any other Borel subset. If we want to make the distinction clear, we may write $\mu|_X$. According to the following lemma, when $T : X \to X$ is a subsystem of $S : Z \to Z$ and μ is an S-invariant finite measure, then μ is also T-invariant.

LEMMA 2.4. *When $T : X \to X$ is a measurable subsystem of $S : Z \to Z$ and μ is an S-invariant finite measure, then μ is also T-invariant.*

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Proof. Take a measurable set $A \subset X$. Since μ is S-invariant,

$$\begin{aligned}\mu(A) + \mu(X^c) &= \mu(A \cup X^c)\\ &= \mu(S^{-1}(A) \cup S^{-1}(X^c))\\ &\le \mu(S^{-1}(A) \cup X^c)\\ &= \mu(T^{-1}(A) \cup X^c)\\ &= \mu(T^{-1}(A)) + \mu(X^c).\end{aligned}$$

This means that

$$\mu(A) \le \mu(T^{-1}(A)).$$

But, on the other hand,

$$\mu(T^{-1}(A)) \le \mu(S^{-1}(A)) = \mu(A). \qquad \square$$

LEMMA 2.5. *Suppose $T : X \to X$ is a measurable subsystem of $S : Z \to Z$. Whenever μ is an S-invariant finite measure, for any Borel set $A \subset X$ and any $n \in \mathbb{N}$,*

$$\mu(X^c \cap S^{-n}(A)) = 0.$$

Proof. Lemma 2.4 implies that μ is also T-invariant. Since $T^{-n}(A) = X \cap S^{-n}(A)$,

$$\begin{aligned}\mu(X \cap S^{-n}(A)) &= \mu(T^{-n}(A))\\ &= \mu(A)\\ &= \mu(S^{-n}(A))\\ &= \mu(X \cap S^{-n}(A)) + \mu(X^c \cap S^{-n}(A)).\end{aligned}$$

Therefore $\mu(X^c \cap S^{-n}(A)) = 0$. $\square$

2.2. *Kolmogorov–Sinai entropy.* Consider the finite measure space $(X, \mathscr{B}, \mu)$ and a finite measurable partition $\mathscr{C}$. The *partition entropy* of $\mathscr{C}$ is

$$H_\mu(\mathscr{C}) = \sum_{C \in \mathscr{C}} \mu(C) \log \frac{1}{\mu(C)}.$$

For the measurable dynamical system $T : X \to X$, if μ is a T-invariant finite measure, then the *partition entropy of T with respect to $\mathscr{C}$* is

$$h_\mu(T, \mathscr{C}) = \lim_{n \to \infty} \frac{1}{n} H_\mu(\mathscr{C}^n)$$

and the *Kolmogorov–Sinai entropy* of T is

$$h_\mu(T) = \sup_{\substack{\mathscr{C}\text{: finite}\\ \text{measurable partition}}} h_\mu(T, \mathscr{C}).$$

The most basic properties of the partition entropy are a consequence of the concavity of the function

$$f(x) = -x \log x = x \log \frac{1}{x}.$$

The following lemma states properties that are standard for probability measures (see [**12**]). They are easily restated in terms of finite measures, but we shall not need them in this general case.

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

LEMMA 2.6. *If μ is a probability measure and $\mathscr{C}$ and $\mathscr{D}$ finite measurable partitions, then:*

(1) $H_\mu(\mathscr{C} \vee \mathscr{D}) \leq H_\mu(\mathscr{C}) + H_\mu(\mathscr{D})$;
(2) *if $\mu_n(C) \to \mu(C)$ for every $C \in \mathscr{C}$, then $H_{\mu_n}(\mathscr{C}) \to H_\mu(\mathscr{C})$; and*
(3) $H_\mu(\mathscr{C}) \leq \log N(\mathscr{C})$.

LEMMA 2.7. *If $\mu = \alpha\gamma + \beta\nu$ is a convex combination of finite measures and $\mathscr{C}$ is a finite measurable partition, then*

$$\alpha H_\gamma(\mathscr{C}) + \beta H_\nu(\mathscr{C}) \leq H_\mu(\mathscr{C}).$$

Proof. For each $C \in \mathscr{C}$, the concavity of $x \log (1/x)$ implies that

$$\alpha\left(\gamma(C) \log \frac{1}{\gamma(C)}\right) + \beta\left(\nu(C) \log \frac{1}{\nu(C)}\right) \leq \mu(C) \log \frac{1}{\mu(C)}.$$

Now one just has to sum up for all $C \in \mathscr{C}$. □

LEMMA 2.8. *If μ is a finite measure over X with $0 \leq \mu(X) \leq 1$ and if $Y \subset X$ is measurable, then*

$$H_\mu(\mathscr{C}) \leq H_\mu(Y \cap \mathscr{C}) + H_\mu(Y^c \cap \mathscr{C})$$

for any finite measurable partition $\mathscr{C}$.

Proof. Fix $C \in \mathscr{C}$. Let $p = \mu(Y \cap C)$ and $q = \mu(Y^c \cap C)$. If $p = 0$, then $\mu(C) = q$ and

$$\begin{aligned}\mu(C) \log \frac{1}{\mu(C)} &= q \log \frac{1}{q} \\ &= q \log \frac{1}{q} + p \log \frac{1}{p}.\end{aligned}$$

A similar conclusion follows when $q = 0$. Assume that $p, q \neq 0$. Then, since $p, q \leq \mu(C)$,

$$\begin{aligned}\mu(C) \log \frac{1}{\mu(C)} &= p \log \frac{1}{\mu(C)} + q \log \frac{1}{\mu(C)} \\ &\leq p \log \frac{1}{p} + q \log \frac{1}{q}.\end{aligned}$$

The lemma follows if we sum up for $C \in \mathscr{C}$. □

Notice that our definition of the Kolmogorov–Sinai entropy does not assume that μ is a probability measure. This will prove to be useful when X is metrizable locally compact (but not necessarily compact). In this case, the set of probability measures is not compact in the weak-$*$ topology, while the set of finite measures with $0 \leq \mu(X) \leq 1$ is indeed compact.

LEMMA 2.9. *Given a measurable dynamical system $T : X \to X$ and a finite T-invariant measure μ, then, for $\alpha \geq 0$,*

$$h_{\alpha\mu}(T) = \alpha h_\mu(T).$$

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Proof. We can assume that $\alpha \neq 0$, since $h_0(T) = 0$.

For any measurable finite partition $\mathscr{C}$,

$$\frac{1}{n}\sum_{C\in\mathscr{C}^n} \alpha\mu(C)\log\frac{1}{\alpha\mu(C)} = \frac{\alpha}{n}\sum_{C\in\mathscr{C}^n}\mu(C)\log\frac{1}{\alpha\mu(C)}$$
$$= \frac{\alpha}{n}\left(\sum_{C\in\mathscr{C}^n}\mu(C)\log\frac{1}{\mu(C)}\right) + \frac{\alpha}{n}\mu(X)\log\frac{1}{\alpha}.$$

Now we just have to take the limit for $n \to \infty$. □

The following lemma will be very important for reducing the proof of the variational principle to the compact case.

LEMMA 2.10. *Let $S : Z \to Z$ be a measurable dynamical system and let $T : X \to X$ be a subsystem with $X \subset Z$ measurable. If μ is an S-invariant measure and if*

$$\mathscr{Z} = \{Z_0, \ldots, Z_k\}$$

is a measurable partition of Z such that $Z_1, \ldots, Z_k \subset X$, then μ is T-invariant and

$$h_\mu(S, \mathscr{Z}) \leq h_\mu(T, \mathscr{C}),$$

where $\mathscr{C} = X \cap \mathscr{Z}$.

Proof. According to Lemma 2.4, μ is in fact T-invariant. According to Lemma 2.2,

$$\mathscr{C}_T^n = X \cap \mathscr{Z}_S^n.$$

Now Lemma 2.8 implies that

$$H_\mu(\mathscr{Z}_S^n) \leq H_\mu(X \cap \mathscr{Z}_S^n) + H_\mu(X^c \cap \mathscr{Z}_S^n)$$
$$= H_\mu(\mathscr{C}_T^n) + H_\mu(X^c \cap \mathscr{Z}_S^n).$$

Dividing by n and taking the limit as $n \to \infty$ will prove our claim as soon as we show that the second term on the right-hand side is constant.

Since elements $C \in \mathscr{Z}^n$ are of the form

$$C = C_0 \cap S^{-1}(C_1) \cap \cdots \cap S^{-n+1}(C_{n-1}),$$

with $C_j \in \mathscr{Z}$, Lemma 2.5 implies that the only $C \in X^c \cap \mathscr{Z}^n$ with non-null measure is

$$C = X^c \cap \bigcap_{j=0}^{n-1} S^{-1}(Z_0).$$

In particular, since $X^c \cap \mathscr{Z}$ is a finite partition of X^c,

$$\mu(C) = \mu(X^c).$$

Therefore

$$H_\mu(X^c \cap \mathscr{Z}_S^n) = \mu(X^c)\log\frac{1}{\mu(X^c)}$$

is constant. □

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

2.2.1. *Conditional entropy.* Given a probability measure μ and two finite measurable partitions $\mathscr{C}$ and $\mathscr{D}$, the *conditional entropy* is an important tool for relating $H_\mu(\mathscr{C})$ and $H_\mu(\mathscr{D})$. For a measurable set C with $\mu(C) > 0$, probability μ conditioned to C, $\mu(\cdot\,|C)$, is given by

$$\mu(B\,|C) = \frac{\mu(B \cap C)}{\mu(C)}.$$

For our purposes, when $\mu(C) = 0$, the conditional probability can be defined arbitrarily.

Definition 2.11. (Conditional entropy) Given a probability measure μ and two finite measurable partitions $\mathscr{C}$ and $\mathscr{D}$, the *conditional entropy* is defined as the expected value

$$H_\mu(\mathscr{D}\,|\mathscr{C}) = \sum_{C\in\mathscr{C}} \mu(C) H_{\mu(\cdot\,|C)}(\mathscr{D}).$$

Conditional entropy possesses the following properties.

LEMMA 2.12. *Let $T : X \to X$ be a measurable dynamical system with T-invariant probability measure μ. If $\mathscr{C}$ and $\mathscr{D}$ are two measurable finite partitions, then*

$$h_\mu(T, \mathscr{C}) \le h_\mu(T, \mathscr{D}) + H_\mu(\mathscr{C}\,|\mathscr{D}).$$

Proof. This is item (iv) of [**12**, Theorem 4.12]. □

2.3. *Topological entropy.* A purely topological concept of entropy for a compact system introduced by Adler *et al* in [**1**], analogous to the Kolmogorov–Sinai entropy, is the *topological entropy*. Dinaburg and Goodman proved (see [**6**, **7**]) the *variational principle*, which states that, for metrizable compact systems, the *topological entropy* is equal to the supremum of the Kolmogorov–Sinai entropies taken over all T-invariant probability measures. In [**10**], Patrão noticed that when the dynamical system admitted a one point compactification, the variational principle still holds as long as we adapt the original definition of topological entropy. In the present paper, we provide a definition of *topological entropy* that extends the previous definitions and allows us to prove the variational principle for metrizable separable locally compact systems.

Definition 2.13. (Cover entropy) Given a cover $\mathscr{A}$ of a set X, the *cover entropy of* $\mathscr{A}$ is

$$H(\mathscr{A}) = \log N(\mathscr{A}).$$

Motivated by the definition presented in [**10**], we shall restrict our attention to open covers of a certain type.

Definition 2.14. (Admissible cover) In a topological space X, an open cover $\mathscr{A}$ is said to be *admissible* when at least one of its elements have compact complement. If every set has compact complement, $\mathscr{A}$ is said to be *strongly admissible*, or *s-admissible* for short.

Remark 2.15. If $\widetilde{X}$ is a compactification of X, then, for any admissible cover $\mathscr{A}$ of X, there exists an open cover $\widetilde{\mathscr{A}}$ of $\widetilde{X}$, such that $\mathscr{A} = X \cap \widetilde{\mathscr{A}}$. In fact, we might simply take, for $\widetilde{\mathscr{A}}$, the family of all open sets $\widetilde{A} \subset \widetilde{X}$ such that $X \cap \widetilde{A} \in \mathscr{A}$. Notice that $\widetilde{\mathscr{A}}$ is sure to cover $\widetilde{X}$ for the simple fact that there is a compact set $K \subset X$ such that $X \setminus K \in \mathscr{A}$; therefore $\widetilde{X} \setminus K$ belongs to $\widetilde{\mathscr{A}}$. Since K is covered by elements in $\widetilde{\mathscr{A}}$, $\widetilde{\mathscr{A}}$ is a cover for $\widetilde{X}$.

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Definition 2.16. (Topological entropy) For a dynamical system $T : X \to X$ and a cover $\mathscr{A}$, the *topological entropy of T with respect to $\mathscr{A}$* is

$$h(T, \mathscr{A}) = \lim_{n\to\infty} \frac{1}{n} H(\mathscr{A}^n).$$

The *topological entropy of T* is

$$h(T) = \sup_{\mathscr{A}\text{: admissible}} h(T, \mathscr{A}).$$

Throughout this paper, the term *AKM entropy* refers to the original definition of entropy given by Adler, Konheim and McAndrew, while the term *topological entropy* refers to our modified definition.

Remark 2.17. Notice that, as in the case of the AKM entropy and the Kolmogorov–Sinai entropy, the limit in Definition 2.16 exists thanks to the inequality

$$N(\mathscr{A} \vee \mathscr{B}) \le N(\mathscr{A})N(\mathscr{B})$$

(see [**12**, Theorem 4.10]).

We now state some very basic properties satisfied by the topological entropy. Most of the arguments are consequences of the following simple lemma.

LEMMA 2.18. *Given a dynamical system $T : X \to X$, let $\mathscr{A}$ and $\mathscr{B}$ be covers of X such that $\mathscr{B} \prec \mathscr{A}$. Then, for all $k \in \mathbb{N}$ and every subset $Y \subset X$:*

(1) $\mathscr{B}^k \prec \mathscr{A}^k$;
(2) $N_Y(\mathscr{B}^k) \le N_Y(\mathscr{A}^k)$;
(3) $N_{T^{-1}(Y)}(T^{-1}(\mathscr{A})) \le N_Y(\mathscr{A})$; *and*
(4) $h(T, \mathscr{B}) \le h(T, \mathscr{A})$.

When the space is compact, it is a simple fact that

$$h(T^k) = kh(T).$$

For the non-compact case, only an inequality follows from a similar argument (see Remark 2.20).

LEMMA 2.19. *Consider the dynamical system $T : X \to X$ and let $k \in \mathbb{N}$. Then*

$$h(T^k) \le kh(T).$$

Proof. Let $\mathscr{A}$ be an admissible cover of X. Notice that $(\mathscr{A}_T^k)_{T^k}^n = \mathscr{A}_T^{kn}$. So

$$k\frac{1}{kn} H(\mathscr{A}_T^{kn}) = \frac{1}{n} H((\mathscr{A}_T^k)_{T^k}^n).$$

Taking the limit for $n \to \infty$,

$$kh(T, \mathscr{A}) = h(T^k, \mathscr{A}_T^k).$$

Since $\mathscr{A} \prec \mathscr{A}_T^k$,

$$h(T^k, \mathscr{A}) \le h(T^k, \mathscr{A}^k) = kh(T, \mathscr{A}).$$

Now we just have to take the supremum for every admissible cover $\mathscr{A}$ to conclude that

$$h(T^k) \le kh(T). \qquad \square$$

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Remark 2.20. The fact that $\mathscr{A}_T^k$ is not necessarily an admissible cover obstructed our way of showing that $h(T^k) = kh(T)$. In the case of Kolmogorov–Sinai entropy, for instance, the supremum of $h_\mu(T, \mathscr{C})$ is taken for every finite measurable partition $\mathscr{C}$. In this case, the demonstration of Lemma 2.19 is easily adapted to show that $h_\mu(T^k) = kh_\mu(T)$. One just has to notice that, since $\mathscr{C}_T^k$ is a finite measurable partition,

$$kh_\mu(T, \mathscr{C}) = h_\mu(T^k, \mathscr{C}_T^k) \leq h_\mu(T^k).$$

For locally compact separable metrizable systems, the equality $h(T^k) = kh(T)$ will follow from the variational principle (Corollary 3.4).

2.4. *Bowen entropy.* Bowen introduced, in [**3**], a definition of entropy that coincides with AKM's topological entropy when the dynamical system $T : X \to X$ is compact metrizable. We shall present Bowen's entropy in a different fashion, that is easier to compare with the topological entropy.

Choose a distance d in X and, given $\varepsilon > 0$, denote the family of balls of radius ε by

$$\mathscr{B}_d(\varepsilon) = \{B_d(x; \varepsilon) \mid x \in X\}.$$

Also, for $n \in \mathbb{N}$, we define

$$d_n(x, y) = \max_{0 \leq j < n} d(T^j x, T^j y).$$

In the literature, (n, ε)-spanning sets are usually defined as sets $E \subset X$ for which, given any $x \in X$, there exists $y \in E$ such that $d_n(x, y) \leq \varepsilon$. We adopt an equivalent definition, but in terms of covers. We do so in a way that is easier to relate to the (n, ε)-spanning sets with our enhanced definition of topological entropy.

Definition 2.21. ((n, ε)-spanning set) Let $T : X \to X$ be a dynamical system with a distance d. For a given $\varepsilon > 0$ and $n \in \mathbb{N}$, a set $E \subset X$ is a (n, ε)-*spanning set* when

$$X = \bigcup_{x \in E} B_{d_n}(x; \varepsilon).$$

That is, the family $\{B_{d_n}(x; \varepsilon) \mid x \in E\}$ is a cover for X.

Definition 2.22. (d-entropy) Let $T : X \to X$ be a dynamical system and let d be a distance for X. Given $\varepsilon > 0$ and $n \in \mathbb{N}$, define

$$h^d(T, \varepsilon) = \lim_{n \to \infty} \frac{1}{n} \log N(\mathscr{B}_{d_n}(\varepsilon))$$

and

$$h^d(T) = \sup_{\varepsilon > 0} h^d(T, \varepsilon).$$

We denote the usual Bowen entropy by

$$h_d(T) = \sup_{\substack{\varepsilon > 0 \\ K: \text{ compact}}} \limsup_{n \to \infty} \frac{1}{n} \log N_K(\mathscr{B}_{d_n}(\varepsilon)).$$

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Remark 2.23. Again, as in the case of the topological entropy and the Kolmogorov–Sinai entropy, the limit in Definition 2.22 exists thanks to the inequality

$$N(\mathscr{B}_{d_n}(\varepsilon)) \leq N(\mathscr{B}_{d_q}(\varepsilon))N(\mathscr{B}_{d_{n-q}}(\varepsilon))$$

for all $q \in \mathbb{N}$ such that $0 < q < n$ (see [**12**, Theorem 4.10]).

An alternative way to characterize the d-entropy (and Bowen's entropy) is by means of *separated sets*. Misiurewicz's proof of the variational principle needs this characterization.

Let $T : X \to X$ be a dynamical system with distance d. Given $\varepsilon > 0$ and $n \in \mathbb{N}$, we say that a set $S \subset X$ is (n, ε)-*separated* when, for all pairs of distinct points $x, y \in S$, $d_n(x, y) \geq \varepsilon$. For a subset $Y \subset X$ and $\varepsilon > 0$, let us write $s(n, \varepsilon, Y)$ for the supremum amongst the cardinalities of (n, ε)-separated subsets of Y.

LEMMA 2.24. *In a dynamical system $T : X \to X$ in which (X, d) is a metric space, given a subset $Y \subset X$ and $\varepsilon > 0$, for all $n \in \mathbb{N}$,*

$$N_Y(\mathscr{B}_{d_n}(\varepsilon)) \leq s(n, \varepsilon, Y) \leq N_Y\left(\mathscr{B}_{d_n}\left(\frac{\varepsilon}{2}\right)\right).$$

Proof. The first inequality arises from the following claim by the existence (through Zorn's lemma) of maximal (n, ε)-separated sets.

Claim. If $E \subset Y$ is a maximal (n, ε)-separated subset of Y, then

$$\mathscr{B}_E = \{Y \cap B_{d_n}(x; \varepsilon) \mid x \in E\}$$

is a cover of Y.

If $\mathscr{B}_E$ is not a cover, then, taking $y \in Y \setminus \bigcup_{x \in E} B_{d_n}(x; \varepsilon)$, the set $E \cup \{y\}$ is (n, ε)-separated, which infringes the maximality of the set E.

For the second inequality, if $S \subset Y$ is a (n, ε)-separated set and if $\mathscr{B} \subset \mathscr{B}_{d_n}(\varepsilon/2)$ is a covering of Y, then for each $s \in S$, there exists an $e(s) \in \mathscr{B}$ such that $s \in e(s)$. This mapping is injective. In fact, if $e(s_1) = e(s_2)$, then $d_n(s_1, s_2) < \varepsilon$. Also, since S is (n, ε)-separated, we must have $s_1 = s_2$. □

The following proposition characterizes the d-entropy in terms of separated sets and also shows that our formulation of Bowen's entropy (Definition 2.22) is equivalent to that of Bowen himself.

PROPOSITION 2.25. *For a dynamical system $T : X \to X$, where (X, d) is a metric space,*

$$h^d(T) = \sup_{\varepsilon > 0} \lim_{n \to \infty} \frac{1}{n} \log s(n, \varepsilon, X)$$

$$h_d(T) = \sup_{\substack{\varepsilon > 0 \\ K:\ compact}} \lim_{n \to \infty} \frac{1}{n} \log s(n, \varepsilon, K).$$

Proof. The proof is immediate, from Lemma 2.24. Just take the log, divide by n, take the limit for $n \to \infty$ and take the supremum for $\varepsilon > 0$ in

$$N(\mathscr{B}_{d_n}(\varepsilon)) \leq s(n, \varepsilon, X) \leq N\left(\mathscr{B}_{d_n}\left(\frac{\varepsilon}{2}\right)\right).$$

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Also, for the Bowen entropy, do the same and also take the supremum over compact $K \subset X$ for

$$N_K(\mathscr{B}_{d_n}(\varepsilon)) \leq s(n, \varepsilon, K) \leq N_K\left(\mathscr{B}_{d_n}\left(\frac{\varepsilon}{2}\right)\right). \qquad \square$$

When defining the d-entropy, we have used the families $\mathscr{B}_{d_n}(\varepsilon)$. Notice that those families are not the same as $[\mathscr{B}_d(\varepsilon)]^n$. The following lemma shows that the families $[\mathscr{B}_d(\varepsilon)]^n$ would work as well, making the d-entropy and the topological entropy much easier to compare.

LEMMA 2.26. *Let $T : X \to X$ be a dynamical system with distance d. Then*

$$h^d(T) = \sup_{\varepsilon>0} h(T, \mathscr{B}_d(\varepsilon))$$

$$h_d(T) = \sup_{\substack{\varepsilon>0 \\ K:\ compact}} \limsup_{n\to\infty} \frac{1}{n} \log N_K(\mathscr{B}_d(\varepsilon)^n).$$

Proof. It is enough to show that

$$[\mathscr{B}_d(\varepsilon)]^n \prec \mathscr{B}_{d_n}(\varepsilon) \prec \left[\mathscr{B}_d\left(\frac{\varepsilon}{2}\right)\right]^n.$$

Indeed, this would imply that, for any $Y \subset X$,

$$\frac{1}{n} \log N_Y([\mathscr{B}_d(\varepsilon)]^n) \leq \frac{1}{n} \log N_Y(\mathscr{B}_{d_n}(\varepsilon)) \leq \frac{1}{n} \log N_Y\left(\left[\mathscr{B}_d\left(\frac{\varepsilon}{2}\right)\right]^n\right).$$

Therefore the claim for the first equality will follow if we make $n \to \infty$ and take the supremum for $\varepsilon > 0$ and $Y = X$. For the second equality, instead of taking $Y = X$, we also have to take the supremum for every $Y \subset X$ compact.

It is immediate that

$$\mathscr{B}_{d_n}(\varepsilon) \subset [\mathscr{B}_d(\varepsilon)]^n,$$

since every ball in the distance d_n has the form

$$B_d(x; \varepsilon) \cap \cdots \cap T^{-(n-1)} B_d(T^{n-1}x; \varepsilon).$$

Therefore $[\mathscr{B}_d(\varepsilon)]^n \prec \mathscr{B}_{d_n}(\varepsilon)$. On the other hand, a set $A \in [\mathscr{B}_d(\varepsilon/2)]^n$ has the form

$$A = B_d\left(x_0; \frac{\varepsilon}{2}\right) \cap \cdots \cap T^{-(n-1)} B_d\left(x_{n-1}; \frac{\varepsilon}{2}\right).$$

In particular, if $A \neq \emptyset$, by taking $x \in A$, we have that, for $j = 0, \ldots, n-1$, $d(T^j x, x_j) < \varepsilon/2$. So

$$B_d\left(x_j; \frac{\varepsilon}{2}\right) \subset B_d(T^j x; \varepsilon).$$

Therefore

$$A \subset B_d(x; \varepsilon) \cap \cdots \cap T^{-(n-1)} B_d(T^{n-1}x; \varepsilon) \in \mathscr{B}_{d_n}(\varepsilon). \qquad \square$$

The following lemma states the existence of the *Lebesgue number* in a form that is easy to apply to the construction of refinements for a given cover.

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

LEMMA 2.27. (Lebesgue number) *Suppose that (X, d) is a metric space which admits a compactification $(\widetilde{X}, \widetilde{d})$. Let $\mathscr{A} = X \cap \widetilde{\mathscr{A}}$, where $\widetilde{\mathscr{A}}$ is an open cover of $\widetilde{X}$. Then there exists $\varepsilon > 0$ such that*

$$\mathscr{A} \prec \mathscr{B}_d(\varepsilon).$$

Proof. Let

$$C_\varepsilon = \{x \in \widetilde{X} \mid \exists A \in \widetilde{\mathscr{A}},\ B_{\widetilde{d}}(x; \varepsilon) \subset A\}$$

be the set of all $x \in \widetilde{X}$ such that the balls centered in x with radius ε are contained in some element of $\widetilde{\mathscr{A}}$. Now we shall find $\varepsilon > 0$ such that $\widetilde{X} = C_\varepsilon$.

Claim. $\widetilde{X} = \bigcup_{\varepsilon>0} C_\varepsilon$.

For each $x \in \widetilde{X}$, there is a $A \in \widetilde{\mathscr{A}}$ containing x. This means that there exists $\varepsilon > 0$ with $B_{\widetilde{d}}(x; \varepsilon) \subset A$.

Claim. $C_\varepsilon \subset \operatorname{int}(C_{\varepsilon/2})$.

For $x \in C_\varepsilon$, take $A \in \mathscr{A}$ such that $B_{\widetilde{d}}(x; \varepsilon) \subset A$. Notice that

$$B_{\widetilde{d}}\left(x; \frac{\varepsilon}{2}\right) \subset C_{\varepsilon/2}.$$

In fact, for $y \in B_{\widetilde{d}}(x; \varepsilon/2)$, $B_{\widetilde{d}}(y; \varepsilon/2) \subset B_{\widetilde{d}}(x; \varepsilon)$: that is, $x \in \operatorname{int}(C_{\varepsilon/2})$.

Joining both claims gives

$$\widetilde{X} = \bigcup_{\varepsilon>0} \operatorname{int}(C_\varepsilon).$$

Since $\widetilde{X}$ is compact, there exists $\varepsilon > 0$ such that $\widetilde{X} = C_\varepsilon$. This means that $\widetilde{\mathscr{A}} \prec \mathscr{B}_{\widetilde{d}}(\varepsilon)$. Taking the intersection with X,

$$\mathscr{A} = X \cap \widetilde{\mathscr{A}} \prec X \cap \mathscr{B}_{\widetilde{d}}(\varepsilon).$$

Now we just have to observe that, since

$$B_d(x; \varepsilon) = X \cap B_{\widetilde{d}}(x; \varepsilon)$$

for any $x \in X$,

$$\mathscr{B}_d(\varepsilon) \subset X \cap \mathscr{B}_{\widetilde{d}}(\varepsilon).$$

Therefore

$$\mathscr{A} \prec \mathscr{B}_d(\varepsilon).$$ □

Finally, we relate the topological entropy and the d-entropy.

PROPOSITION 2.28. *Let $T : X \to X$ be a dynamical system and let $\widetilde{X}$ be a metrizable compactification of X. If d is the restriction to X of a distance $\widetilde{d}$ in $\widetilde{X}$, then*

$$h^d(T) = \sup_{\widetilde{\mathscr{A}}} h(T, X \cap \widetilde{\mathscr{A}}),$$

where the supremum is taken over all open covers $\widetilde{\mathscr{A}}$ of $\widetilde{X}$.

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Proof. The family $\mathscr{D} = \{B_{\widetilde{d}}(x; \varepsilon) \mid x \in X\}$ covers $\widetilde{X}$ and is such that $\mathscr{B}_d(\varepsilon) = X \cap \mathscr{D}$. This implies that

$$h^d(T) \leq \sup_{\widetilde{\mathscr{A}}\text{: open}} h(T, X \cap \widetilde{\mathscr{A}}).$$

Now Lemma 2.27 implies that, for any open cover of $\widetilde{X}$, $\widetilde{\mathscr{A}}$, there exists $\varepsilon > 0$ such that

$$X \cap \widetilde{\mathscr{A}} \prec \mathscr{B}_d(\varepsilon).$$

Also, this means that

$$h(T, X \cap \widetilde{\mathscr{A}}) \leq h(T, \mathscr{B}_d(\varepsilon)).$$

Taking the supremum in ε and applying the Lemma 2.26, we conclude that

$$h(T, X \cap \widetilde{\mathscr{A}}) \leq h^d(T).$$

Taking the supremum for all open covers $\widetilde{\mathscr{A}}$,

$$\sup_{\widetilde{\mathscr{A}}\text{: open}} h(T, X \cap \widetilde{\mathscr{A}}) \leq h^d(T). \qquad \square$$

When the dynamical system is compact, we know that the d-entropy does not depend on the distance d. The following corollary to Proposition 2.28 extends this result.

COROLLARY 2.29. *Let $T : X \to X$ be a dynamical system with X admitting a metrizable compactification $\widetilde{X}$. If d and c are the restriction to X of distances $\widetilde{d}$ and $\widetilde{c}$ in $\widetilde{X}$, then*

$$h^d(T) = h^c(T).$$

Starting with a technique similar to the one that we used in the demonstration of Lemma 2.27, we relate Bowen's entropy and the topological entropy.

PROPOSITION 2.30. *Let $T : X \to X$ be a dynamical system with a distance d. Then*

$$h(T) \leq h_d(T) \leq h^d(T).$$

Proof. The Bowen entropy will be calculated according to Lemma 2.26. It is evident, from the definitions, that $h_d(T) \leq h^d(T)$.

Given an admissible cover $\mathscr{A}$, take a compact $K \subset X$ such that $K^c \in \mathscr{A}$. Let

$$C_\varepsilon = \{x \in K \mid \exists A \in \mathscr{A},\ B_d(x; 2\varepsilon) \subset A\}$$

be the set of all $x \in K$ such that the ball centered at x with radius 2ε is a subset of some element in $\mathscr{A}$. Just as in the proof of Lemma 2.27, we can take $\varepsilon > 0$ such that $K = C_\varepsilon$.

Claim. $\mathscr{A} \prec \mathscr{B}_d(\varepsilon)$.

If $B \in \mathscr{B}_d(\varepsilon)$ does not intersect K, then $B \subset K^c \in \mathscr{A}$. Also, if B does intersect K, then, taking $x \in B \cap K$ from our choice of ε, there is an $A_x \in \mathscr{A}$ such that

$$B \subset B_d(x; 2\varepsilon) \subset A_x :$$

that is, $\mathscr{A} \prec \mathscr{B}_d(\varepsilon)$.

Therefore, defining

$$\mathscr{D} = \mathscr{B}_d(\varepsilon) \cup \{K^c\},$$

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

it follows that $\mathscr{A} \prec \mathscr{D}$. Let us partition X in $K_0, \ldots, K_{n-1}$ and $\widetilde{K}$ with

$$\widetilde{K} = K^c \cap T^{-1}(K^c) \cap \cdots \cap T^{-(n-1)}(K^c) \in \mathscr{D}^n$$

and, for $m = 0, \ldots, n-1$,

$$K_m = K^c \cap T^{-1}(K^c) \cap \cdots \cap T^{-(m-1)}(K^c) \cap T^{-m}(K).$$

Notice that, in fact, $X = \widetilde{K} \cup \bigcup_{m=0}^{n-1} K_m$, so

$$N(\mathscr{D}^n) \leq N_{\widetilde{K}}(\mathscr{D}^n) + \sum_{m=0}^{n-1} N_{K_m}(\mathscr{D}^n).$$

Since $\widetilde{K} \in \mathscr{D}^n$,

$$N_{\widetilde{K}}(\mathscr{D}^n) = 1.$$

For $m = 0, \ldots, n-1$, denote by $\mathscr{B}_m$ the family of sets of the form

$$A_0 \cap T^{-1}(A_1) \cap \cdots \cap T^{-(n-1)}(A_{n-1}),$$

where $A_j = K^c$ for $j = 0, \ldots, m-1$, and $A_j \in \mathscr{B}_d(\varepsilon)$ for $j \geq m$: that is,

$$\mathscr{B}_m = K^c \cap T^{-1}(K^c) \cap \cdots \cap T^{-(m-1)}(K^c) \cap T^{-m}(\mathscr{B}_d(\varepsilon)^{n-m}) \subset \mathscr{D}^n.$$

If $\mathscr{F} \subset \mathscr{B}_d(\varepsilon)^{n-m}$ is a cover of K, then

$$K^c \cap T^{-1}(K^c) \cap \cdots \cap T^{-(m-1)}(K^c) \cap T^{-m}(\mathscr{F}) \subset \mathscr{B}_m$$

is a cover of K_m. In this way,

$$\begin{aligned} N_{K_m}(\mathscr{D}^n) &\leq N_{K_m}(\mathscr{B}_m) \\ &\leq N_K(\mathscr{B}_d(\varepsilon)^{n-m}) \\ &\leq N_K(\mathscr{B}_d(\varepsilon)^n). \end{aligned}$$

Therefore

$$\begin{aligned} N(\mathscr{A}^n) &\leq N(\mathscr{D}^n) \\ &\leq N_{\widetilde{K}}(\mathscr{D}^n) + \sum_{m=0}^{n-1} N_{K_m}(\mathscr{D}^n) \\ &\leq 1 + \sum_{m=0}^{n-1} N_{K_m}(\mathscr{D}^n) \\ &\leq (n+1) N_K(\mathscr{B}_d(\varepsilon)^n). \end{aligned}$$

Taking the log, dividing by n and making $n \to \infty$, and using Lemma 2.26, we get that

$$\begin{aligned} h(T, \mathscr{A}) &\leq \limsup_{n\to\infty} \frac{1}{n}(\log(n+1) + \log N_K(\mathscr{B}_d(\varepsilon)^n)) \\ &= \limsup_{n\to\infty} \frac{1}{n} \log N_K(\mathscr{B}_d(\varepsilon)^n) \\ &\leq h_d(T), \end{aligned}$$

since $(1/n)\log(n+1) \to 0$. Taking the supremum in $\mathscr{A}$,

$$h(T) \leq h_d(T). \qquad \square$$

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Remark 2.31. When the families $\mathscr{B}_d(\varepsilon)$ are admissible covers, Lemma 2.26 warrants that

$$h_d(T) \leq h(T).$$

Nonetheless, even when d is the restriction of a distance given in some compactification of the system, $\mathscr{B}_d(\varepsilon)$ might not be admissible. As an example, just take $\mathscr{B}_d\left(\frac{1}{2}\right)$ in $(0, 1)$ with the Euclidean distance. No set in $\mathscr{B}_d\left(\frac{1}{2}\right)$ will ever have compact complement; therefore $\mathscr{B}_d\left(\frac{1}{2}\right)$ is not admissible.

3. *Variational principle*

The preparations made in §2 allow us to use Misiurewicz's demonstration of the variational principle almost verbatim. Misiurewicz's original article is [**9**]. A didactic presentation can be found in, for example, [**12**].

We are concerned about the supremum of the Kolmogorov–Sinai entropies over every T-invariant Radon probability measure. If there is no such a probability measure, we agree that

$$\sup_{\mu} h_\mu(T) = 0.$$

According to Lemma 2.9, this is the same as taking the supremum over all T-invariant Radon measures μ with $0 \leq \mu(X) \leq 1$. In this case, there is always an invariant measure, namely, $\mu = 0$.

The following theorem is our main result. The proof will be provided after a few auxiliary results.

THEOREM 3.1. *Let $T : X \to X$ be a metrizable locally compact separable dynamical system. Then*

$$\sup_{\mu} h_\mu(T) = h(T) = \min_{d} h_d(T).$$

The infimum is attained when d is any distance that can be extended to the one point compactification of X.

The following lemma is the part of the variational principle that is valid for every topological dynamical system. The lemma remarkably generalizes [**8**, Proposition 1.4], since the only hypothesis we impose to the system $T : X \to X$ is that T is continuous. Notice how the concept of *admissible cover* allowed us to generalize [**8**, Proposition 1.4] without even appealing to advanced results such as the *ergodic decomposition*, the *Shannon–McMillan–Breiman theorem* or the *Birkhoff ergodic theorem*.

LEMMA 3.2. *Let $T : X \to X$ be a dynamical system and let μ be a T-invariant Radon probability measure. Then*

$$h_\mu(T) \leq h(T).$$

If X is metrizable with distance d, then

$$h_\mu(T) \leq h(T) \leq h_d(T).$$

Proof. The last statement follows from Proposition 2.30. The first part, the details of which we will present here, has actually been demonstrated by Misiurewicz. We start by

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

noticing that, for every Radon measure, the measure of a set A can be approximated from inside by a compact $K \subset A$.

Let μ be any T-invariant Radon probability measure. We shall show that, for any μ-invariant dynamical system $S : X \to X$,

$$h_\mu(S) \le h(S) + 2 + \log 2. \tag{1}$$

In particular, this is valid for T^n for any $n \in \mathbb{N}$. Then Lemma 2.19 implies that

$$\begin{aligned} h_\mu(T) &= \frac{1}{n} h_\mu(T^n) \\ &\le \frac{1}{n} h(T^n) + \frac{2 + \log 2}{n} \\ &\le h(T) + \frac{2 + \log 2}{n} \to h(T). \end{aligned}$$

This will finish the demonstration.

In order to show the validity of equation (1) for any given $S : X \to X$, take a finite measurable partition $\mathscr{C}$ such that $h_\mu(S) \le h_\mu(S, \mathscr{C}) + 1$. Choose, for each $C \in \mathscr{C}$, $D_C \subset C$ such that

$$\mu(C \setminus D_C) \le \frac{1}{N(\mathscr{C}) \log N(\mathscr{C})}.$$

Let $D^* = \bigcup_{C \in \mathscr{C}} (C \setminus D_C)$. Then

$$\mu(D^*) \le \frac{1}{\log N(\mathscr{C})}.$$

Now define a finite measurable partition $\mathscr{D}$ and a strongly admissible covering $\mathscr{A}$ by

$$\begin{aligned} \mathscr{D} &= \{D_C \mid C \in \mathscr{C}\} \cup \{D^*\} \\ \mathscr{A} &= \{D_C \cup D^* \mid C \in \mathscr{C}\}. \end{aligned}$$

Claim. $H_\mu(\mathscr{C} \mid \mathscr{D}) \le 1$.

First, notice that, for every $C \in \mathscr{C}$, $\mu(C \mid D_C) = 1$. So

$$H_{\mu(\cdot \mid D_C)}(\mathscr{C}) = 0.$$

Therefore, using Lemma 2.6,

$$\begin{aligned} H_\mu(\mathscr{C} \mid \mathscr{D}) &= \mu(D^*) H_{\mu(\cdot \mid D^*)}(\mathscr{C}) \\ &\le \mu(D^*) \log N(\mathscr{C}) \le 1. \end{aligned}$$

Claim. $h_\mu(S, \mathscr{D}) \le h(S, \mathscr{A}) + \log 2$.

Let us show that

$$N(\mathscr{D}^n) \le 2^n N(\mathscr{A}^n).$$

Taking the logarithm, dividing by n and taking the limit for $n \to \infty$ gives the claim.

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Let $\Lambda \subset \{1, \ldots, k\}^n$ with $\#\Lambda = N(\mathscr{A}^n)$, such that

$$X = \bigcup_{\lambda \in \Lambda} (K_0 \cup K_{\lambda_0}) \cap \cdots \cap T^{-(n-1)}(K_0 \cup K_{\lambda_{n-1}}),$$

where $\lambda = (\lambda_0, \ldots, \lambda_{n-1})$. Consider the mapping

$$f : \Lambda \times \{0, 1\}^n \to \mathscr{D}^n,$$

that maps (λ, x) to $Y_0 \cap \cdots \cap Y_{n-1}$, where

$$Y_j = \begin{cases} T^{-j} K_{\lambda_j} & \text{if } x_j = 1, \\ T^{-j} K_0 & \text{if } x_j = 0. \end{cases}$$

Since

$$\begin{aligned} X &= \bigcup_{\lambda \in \Lambda} T^0(K_{\lambda_0} \cup K_0) \cap \cdots \cap T^{-(n-1)}(K_{\lambda_{n-1}} \cup K_0) \\ &= \bigcup f(\Lambda \times \{0, 1\}^n), \end{aligned}$$

and since $\mathscr{D}^n$ is a partition, the image of f contains every non-empty element of $\mathscr{D}^n$.

From the previous claims and Lemma 2.12,

$$\begin{aligned} h_\mu(S) &\le h_\mu(S, \mathscr{C}) + 1 \\ &\le h_\mu(S, \mathscr{D}) + 2 \\ &\le h(S, \mathscr{A}) + 2 + \log 2 \\ &\le h(S) + 2 + \log 2, \end{aligned}$$

which demonstrates equation (1) and concludes the proof. □

Remark 3.3. In Lemma 3.2, we have actually shown that

$$h_\mu(T) \le \sup_{\mathscr{A}\text{: s-admissible}} h(T, \mathscr{A}) \le h(T).$$

Proof of Theorem 3.1. Using Lemma 3.2, we know that

$$\sup_\mu h_\mu(T) \le h(T) \le \inf_d h_d(T) \le \inf_d h^d(T),$$

where the infimum is taken over every distance d compatible with the topology in X. We just have to show that if r is the restriction to X of some distance in its one point compactification, then

$$h^r(T) \le \sup_\mu h_\mu(T).$$

In fact, this implies that

$$h^r(T) \le \sup_\mu h_\mu(T) \le h(T) \le \inf_d h_d(T) \le \inf_d h^d(T) \le h^r(T).$$

Using Proposition 2.25, it remains to show that, for each fixed $\varepsilon > 0$ and each sequence of (n, ε)-separated sets E_n, there is a Radon measure μ, which is T-invariant and has total

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

measure lower then or equal to one, and that there is a finite measurable partition $\mathscr{C}$ such that, for every $n \in \mathbb{N}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \#E_n \leq h_\mu(T, \mathscr{C}).$$

We shall take $S : Z \to Z$, as in Lemma 2.3, find a probability measure μ over the Borel sets of Z and a partition $\mathscr{Z}$, as in Lemma 2.10, and show that

$$\limsup_{n\to\infty} \frac{1}{n} \log \#E_n \leq h_\mu(S, \mathscr{Z}).$$

Then Lemma 2.10 will imply the desired conclusion.

The demonstration is very similar to that of [**12**, Theorem 8.6]; the main difference is that we shall use the pseudometric $\widetilde{r}$ induced by r over Z (see the comment after Lemma 2.3). Let us first build up the Radon probability measure μ and show that it is S-invariant. Then Lemma 2.4 implies that μ is T-invariant. Define

$$\sigma_n = \frac{1}{\#E_n} \sum_{x\in E_n} \delta_x,$$

where δ_x is the Dirac measure with support in x. Also define

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} \sigma_n \circ S^{-j},$$

and notice that, restricted to X,

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} \sigma_n \circ T^{-j}.$$

In fact, since $S^{-j}(Z \setminus X) \subset Z \setminus X$,

$$\mu_n(Z \setminus X) = \frac{1}{n} \sum_{j=0}^{n-1} \sigma_n(S^{-j}(Z \setminus X)) = 0.$$

Claim. There is a subsequence n_k and a Radon probability measure μ such that $\mu_{n_k} \to \mu$ and such that

$$\lim_{k\to\infty} \frac{1}{n_k} \log \#E_{n_k} = \limsup_{n\to\infty} \frac{1}{n} \log \#E_n.$$

Also, for any measurable $C \subset Z$ with $\mu(\partial C) = 0$,

$$\lim \mu_{n_k}(C) = \mu(C).$$

In the weak-$*$ topology, the set of Radon probability measures μ over Z is easily seen to be closed. The fact that Z is separable implies that the set of Radon measures is separable. The Alaoglu theorem [**11**, Theorem 2.5.2] implies that it is compact and, therefore, sequentially compact.

It is clear that there is a subsequence n_k such that

$$\lim_{k\to\infty} \frac{1}{n_k} \log \#E_{n_k} = \limsup_{n\to\infty} \frac{1}{n} \log \#E_n.$$

From the sequential compactness, we can assume that n_k is such that μ_{n_k} converges to some μ. The last assertion is a consequence of the *portmanteau theorem* and can be found in [**2**], Theorem 2.1, item (v).

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Claim. The measure μ is S-invariant.

It is clear that $\mu_{n_k} \circ S^{-1} \to \mu \circ S^{-1}$. In fact, for any continuous $\phi : Z \to \mathbb{R}$, since R is continuous,

$$\begin{aligned}\int \phi \, d(\mu_{n_k} \circ S^{-1}) &= \int \phi \circ S \, d\mu_{n_k} \\ &\to \int \phi \circ S \, d\mu \\ &= \int \phi \, d(\mu \circ S^{-1}).\end{aligned}$$

On the other hand,

$$\begin{aligned}\left| \int \phi \, d(\mu_{n_k} - \mu_{n_k} \circ S^{-1}) \right| &= \left| \int \phi \frac{1}{n_k} \, d(\sigma_{n_k} - \sigma_{n_k} \circ S^{-n}) \right| \\ &= \frac{1}{n_k} \left| \int (\phi - \phi \circ S^{n_k}) \, d\sigma_{n_k} \right| \\ &\leq \frac{1}{n_k} \int \|\phi - \phi \circ S^{n_k}\|_\infty \, d\sigma_{n_k} \\ &\leq \frac{1}{n_k} \int 2\|\phi\|_\infty \, d\sigma_{n_k} \\ &= \frac{1}{n_k} 2\|\phi\|_\infty \to 0.\end{aligned}$$

This implies that

$$\mu = \lim \mu_{n_k} = \lim \mu_{n_k} \circ S^{-1} = \mu \circ S^{-1}.$$

Now, we choose a measurable partition $\mathscr{Z}$. For each $z \in Z$, there exists a ball $B_z = B(z; \varepsilon_z)$ with $\varepsilon_z < \varepsilon/2$ such that $\mu(\partial B_z) = 0$. Such a B_z exists because, since the border of such balls are all disjoint, there is at most a countable number of reals $\varepsilon > 0$ such that $B(z; \varepsilon)$ has border with non-null measure. Since Z is compact and the balls are open, there is a finite number of such balls, $B_0, \ldots, B_n$ covering Z. We can assume that $\{B_0, \ldots, B_n\}$ has no proper subcover. Let

$$Z_j = B_j \setminus (B_1 \cup \cdots \cup B_{j-1}).$$

Then $\mathscr{Z} = \{Z_0, \ldots, Z_k\}$ is a measurable partition. We can also assume that $Z \setminus X \subset B_0 = Z_0$: that is, $\mathscr{Z}$ satisfies the conditions of Lemma 2.10.

Claim. For each $C \in \mathscr{Z}^n$, $\mu(\partial C) = 0$.

Notice that, since S is continuous, the border operator ∂ possesses the following properties:
(1) $\partial A = \partial A^c$;
(2) $\partial(A_1 \cap \cdots \cap A_k) \subset \partial A_1 \cup \cdots \cup \partial A_k$; and
(3) $\partial S^{-1}(A) \subset S^{-1}(\partial A)$.
From items (1) and (2), each $Z_j = B_j \cap {B_1}^c \cap \cdots \cap {B_{j-1}}^c$ in $\mathscr{Z}$ has border with null measure. Also, from items (2) and (3), the same is true for the sets in $\mathscr{Z}^n$.

In addition, each $C \in \mathscr{Z}^n$ has diameter less then ε in the pseudometric $\widetilde{r}_n$, just as in the compact case.

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Claim. $\log \#E_n = H_{\sigma_n}(\mathscr{Z}^n)$.

Let $C \in \mathscr{Z}^n$. If $x, y \in C$, then there exist $C_0, \ldots, C_{n-1} \in \mathscr{Z}$ such that $T^j x, T^j y \in C_j$ for $j = 0, \ldots, n-1$. Since each element of $\mathscr{Z}$ has diameter less than ε, $d_n(x, y) < \varepsilon$. So C can contain at most one element of E_n: that is, $\sigma_n(C) = 0$ or $\sigma_n(C) = 1/\#E_n$. Therefore

$$H_{\sigma_n}(\mathscr{Z}^n) = \log \#E_n.$$

Passing from σ_n to μ_n is the same procedure as in the compact case, as we shall now detail. Notice that, for any measurable finite partition $\mathscr{D}$, Lemma 2.7 implies that

$$\sum_{j=0}^{n-1} \frac{1}{n} H_{\sigma_n \circ S^{-j}}(\mathscr{D}) \le H_{\mu_n}(\mathscr{D}).$$

For $n, q \in \mathbb{N}$ with $1 < q < n$, take an integer m such that $mq \ge n > m(q-1)$. Then, for every $j = 0, \ldots, q-1$,

$$\begin{aligned}\mathscr{Z}^n &\prec \mathscr{Z}^j \vee S^{-j}(\mathscr{Z}^{qm}) \\ &= \mathscr{Z}^j \vee S^{-j}(\mathscr{Z}^q) \vee S^{-(j+q)}(\mathscr{Z}^q) \vee \cdots \vee S^{-(j+(m-1)q)}(\mathscr{Z}^q).\end{aligned}$$

Therefore, using Lemma 2.6,

$$\begin{aligned}H_{\sigma_n}(\mathscr{Z}^n) &\le H_{\sigma_n}(\mathscr{Z}^j) + H_{\sigma_n \circ S^{-(j+0q)}}(\mathscr{Z}^q) + \cdots + H_{\sigma_n \circ S^{-(j+(m-1)q)}}(\mathscr{Z}^q) \\ &\le H_{\sigma_n}(\mathscr{Z}^q) + H_{\sigma_n \circ S^{-(j+0q)}}(\mathscr{Z}^q) + \cdots + H_{\sigma_n \circ S^{-(j+(m-1)q)}}(\mathscr{Z}^q) \\ &\le \log \#\mathscr{Z}^q + H_{\sigma_n \circ S^{-(j+0q)}}(\mathscr{Z}^q) + \cdots + H_{\sigma_n \circ S^{-(j+(m-1)q)}}(\mathscr{Z}^q).\end{aligned}$$

Summing up in $j = 0, \ldots, q-1$,

$$\begin{aligned}q \log \#E_n &= q H_{\sigma_n}(\mathscr{Z}^n) \\ &\le q \log \#\mathscr{Z}^q + \sum_{j=0}^{q-1} \sum_{a=0}^{m-1} H_{\sigma_n \circ S^{-(j+aq)}}(\mathscr{Z}^q) \\ &= q \log \#\mathscr{Z}^q + \sum_{p=0}^{n-1} H_{\sigma_n \circ S^{-p}}(\mathscr{Z}^q) + \sum_{p=n}^{mq-1} H_{\sigma_n \circ S^{-p}}(\mathscr{Z}^q) \\ &\le 2q \log \#\mathscr{Z}^q + n \sum_{p=0}^{n-1} \frac{1}{n} H_{\sigma_n \circ S^{-p}}(\mathscr{Z}^q) \\ &\le 2q \log \#\mathscr{Z}^q + n H_{\mu_n}(\mathscr{Z}^q).\end{aligned}$$

Since each element $C \in \mathscr{Z}^q$ has border with null measure,

$$\lim_{k \to \infty} \mu_{n_k}(C) = \mu(C).$$

Therefore, dividing by qn and making $k \to \infty$,

$$\begin{aligned}\limsup_{n \to \infty} \frac{1}{n} \log \#E_n &= \lim_{k \to \infty} \frac{1}{n_k} \log \#E_{n_k} \\ &\le 0 + \frac{1}{q} \lim_{k \to \infty} H_{\mu_{n_k}}(\mathscr{Z}^q) \\ &= \frac{1}{q} H_{\mu}(\mathscr{Z}^q)\end{aligned}$$

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

for every q. Now one just has to take the limit with $q \to \infty$ to get

$$\limsup_{n\to\infty} \frac{1}{n} \log \#E_n \leq h_\mu(S, \mathscr{X}). \qquad \square$$

The inequality at Lemma 2.19 was used to demonstrate the variational principle. Now, in its turn, the variational principle allows us to go a bit further.

COROLLARY 3.4. *Consider the metrizable locally compact separable dynamical system $T : X \to X$ and let $k \in \mathbb{N}$. Then*

$$h(T^k) = kh(T).$$

Proof. Equality holds for the Kolmogorov–Sinai entropy (Remark 2.20). Therefore the variational principle implies that it is valid for the topological entropy as well. $\square$

4. *Application: Lie group endomorphisms*

We finish by considering the entropy of continuous endomorphisms of Lie groups. We extend some results of [**5**] to endomorphisms that are not necessarily surjective. For a given Lie group G, its toral component $T(G)$ is the maximal connected and compact subgroup of the center of G. We aim to demonstrate that, in certain cases (compact, semisimple, linear reducible and nilpotent), the entropy of an endomorphism of a Lie group G is the entropy of the endomorphism restricted to the toral component $T(H)$, where H is the maximal subgroup of G where the restriction of the endomorphism is surjective.

In order to reduce the general case to the cases covered in [**5**], a key point is the following lemma, which depends on the *variational principle* that we have developed.

LEMMA 4.1. *Let G be a connected Lie group and let $\phi : G \to G$ be a continuous endomorphism. There exists a natural number n such that ϕ restricted to $H = \phi^n(G)$ is surjective. In particular, H is the maximal subgroup of G where the restriction of the endomorphism is surjective and*

$$h(\phi) = h(\phi|_H).$$

Proof. For the first claim, we first consider the induced Lie algebra endomorphism $\phi' : \mathfrak{g} \to \mathfrak{g}$ given by the differential of ϕ at the identity of G. Since ϕ' is a linear map, there is a natural number n such that $(\phi')^n\mathfrak{g} = (\phi')^{n+1}\mathfrak{g}$. Putting $H = \phi^n(G)$, its Lie algebra is given by $\mathfrak{h} = (\phi')^n\mathfrak{g}$. Therefore $\phi(H) = H$, since both are connected subgroups and the Lie algebra of $\phi(H)$ is $\phi'\mathfrak{h} = \mathfrak{h}$. Thus ϕ restricted to H is surjective. If L is a subgroup of G where the restriction of ϕ is surjective, then $L = \phi^n(L) \subset \phi^n(G) = H$, which shows that H is the maximal subgroup of G where the restriction of ϕ is surjective.

For the second claim, notice that, as Lie groups, H and G satisfy the conditions of Theorem 3.1. Also, since H is a countable union of compact subsets and compact sets of H are also compact sets of G, H is a measurable subset of G. Every ϕ-invariant measure μ is such that

$$\begin{aligned} \mu(H) &= \mu(\phi^{-n}(H)) \\ &= \mu(G). \end{aligned}$$

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Therefore

$$\begin{aligned} h(\phi) &= \sup_{\mu} h_{\mu}(\phi) \\ &= \sup_{\mu} h_{\mu}(\phi|_H) \\ &= h(\phi|_H). \end{aligned}$$ □

Using the conclusion and notation of Lemma 4.1, we show that the results in [**5**] imply that, in the cases we are considering,

$$h(\phi) = h(\phi|_H) = h(\phi|_{T(H)}).$$

PROPOSITION 4.2. *Let G be a connected and compact Lie group and consider $\phi : G \to G$ to be a continuous endomorphism. Then*

$$h(\phi) = h(\phi|_{T(H)}),$$

where H is the maximal subgroup of G for which the restriction of ϕ is surjective.

Proof. By Lemma 4.1, $H = \phi^n(G)$. We know that H is a connected and compact Lie group. Using Lemma 4.1 and applying [**5**, Theorem 6.2] for ϕ restricted to H, we get that

$$h(\phi) = h(\phi|_H) = h(\phi|_{T(H)}).$$ □

In the proof of Proposition 4.2, it was evident that $H = \phi^n(G)$ is a compact group. It is also evident that H is nilpotent when G is nilpotent. However, for the semisimple and reductive cases, we need the following lemma.

LEMMA 4.3. *Let G be a semisimple, a reductive, or a nilpotent Lie group. If $\phi : G \to H$ is a surjective Lie group endomorphism, then H is, respectively, semisimple, reductive or nilpotent.*

Proof. The nilpotent case is trivial. Let $\mathfrak{g}$ denote the Lie algebra of G and let $\mathfrak{h}$ be the Lie algebra of H.

If G is reductive, then

$$\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{s},$$

where $\mathfrak{a}$ is the center of $\mathfrak{g}$, and

$$\mathfrak{s} = \mathfrak{s}_1 \oplus \cdots \oplus \mathfrak{s}_n,$$

where $\mathfrak{s}_j$ are simple ideals of $\mathfrak{g}$.

Claim. The Lie algebra $\phi'(\mathfrak{a})$ is in the center of $\mathfrak{h}$.

Since ϕ' is surjective,

$$[\mathfrak{g}, \phi'(\mathfrak{a})] = [\phi'(\mathfrak{g}), \phi'(\mathfrak{a})] = \phi'([\mathfrak{g}, \mathfrak{a}]) = 0.$$

Claim. The Lie algebra $\phi'(\mathfrak{s})$ is semisimple.

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Since

$$\phi'(\mathfrak{s}) = \phi'(\mathfrak{s}_1) + \cdots + \phi'(\mathfrak{s}_n),$$

we just have to show that, for each $j = 1, \ldots, n$, $\phi'(\mathfrak{s}_j)$ is simple and, for each $j, k = 1, \ldots, n$, either $\phi'(\mathfrak{s}_j) = \phi'(\mathfrak{s}_k)$ or $\phi'(\mathfrak{s}_j) \cap \phi'(\mathfrak{s}_k) = 0$.

Let $I \subset \phi'(\mathfrak{s}_j)$ be a non-null ideal. Then $J = \mathfrak{s}_j \cap \phi'^{-1}(I)$ is a non-null ideal of $\mathfrak{s}_j$. Since $\mathfrak{s}_j$ is simple, $J = \mathfrak{s}_j$: that is,

$$\phi'(\mathfrak{s}_j) = \phi'(J) \subset I \subset \phi'(\mathfrak{s}_j).$$

Therefore $\phi'(\mathfrak{s}_j)$ is simple whenever it is non-null.

Now let $I = \phi'(\mathfrak{s}_j) \cap \phi'(\mathfrak{s}_k)$. Then I is an ideal of both $\phi'(\mathfrak{s}_j)$ and $\phi'(\mathfrak{s}_k)$. Therefore either $I = 0$ or $\phi'(\mathfrak{s}_j) = \phi'(\mathfrak{s}_k)$.

Denoting by $\mathfrak{h}$ the Lie algebra of $\phi(G)$, we know that

$$\mathfrak{h} = \phi'(\mathfrak{g}) = \phi'(\mathfrak{a}) + \phi'(\mathfrak{s}).$$

But, since the intersection of an abelian Lie algebra and a semisimple Lie algebra is zero, the sum is, in fact, a direct sum. This proves that $\phi(G)$ is reductive. If G is semisimple, then $\mathfrak{a} = 0$ and, therefore, $\mathfrak{h} = \phi'(\mathfrak{s})$ is semisimple. □

PROPOSITION 4.4. *Let G be a connected semisimple Lie group and let $\phi : G \to G$ be a continuous endomorphism. Then*

$$h(\phi) = 0.$$

Proof. Consider $H = \phi^n(G)$, given by Lemma 4.1. Then Lemma 4.3 implies that H is a connected semisimple Lie group and that

$$h(\phi) = h(\phi|_H) = 0,$$

where the last equality follows from [**5**, Theorem 5.2]. □

PROPOSITION 4.5. *Let G be a connected linear reductive or a connected nilpotent Lie group and let $\phi : G \to G$ be a continuous endomorphism. Then*

$$h(\phi) = h(\phi|_{T(H)}),$$

where H is the maximal subgroup of G in which the restriction of ϕ is surjective. In particular, any linear transformation $T : V \to V$ over a finite-dimensional vector space V has null topological entropy.

Proof. By Lemma 4.1, $H = \phi^n(G)$. Then Lemma 4.3 implies that H is a connected linear reductive or a connected nilpotent.

Using Lemma 4.1 and applying Theorem 4.1 (for the nilpotent case) of [**5**] or Corollary 2 (for the linear reductive case) of [**5**] for ϕ restricted to H, we get that

$$h(\phi) = h(\phi|_H) = h(\phi|_{T(H)}).$$

The last claim follows from the fact that $(V, +)$, as well as H, is an abelian (a *fortiori*, nilpotent) Lie group with trivial toral component. □

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press

Acknowledgements. The first author was supported by CNPq grant no. 140888/11-0. The second author was supported by CNPq grant no. 310790/09-3.

REFERENCES

[**1**] R. Adler, A. Konheim and M. McAndrew. Topological entropy. *Trans. Amer. Math. Soc.* **114**(2) (1965), 309–319.
[**2**] P. Billingsley. *Convergence of Probability Measures (Wiley Series in Probability and Statistics)*. Wiley, Billingsley, Chicago, IL, 1999.
[**3**] R. Bowen. Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.* **153** (1971), 401–414.
[**4**] G. Bredon. *Topology and Geometry*. Springer, New York, 1993.
[**5**] A. Caldas and M. Patrão. Dynamics of endomorphisms of Lie groups. *Discrete Contin. Dyn. Syst.* **33** (2013), 1351–1363.
[**6**] E. Dinaburg. The relation between topological entropy and metric entropy. *Soviet Math. Dokl.* **11** (1969), 13–16.
[**7**] T. Goodman. Relating topological entropy to measure entropy. *Bull. Lond. Math. Soc.* **3** (1971), 176–180.
[**8**] M. Handel, B. Kitchens and D. Rudolph. Metrics and entropy for non-compact spaces. *Israel J. Math.* **91** (1995), 253–271.
[**9**] M. Misiurewicz. A short proof of the variational principle for a $\mathbb{Z}_+^{\mathbb{N}}$ action on a compact space. *Astérisque* **40** (1976), 147–157.
[**10**] M. Patrão. Entropy and its variational principle for non-compact metric spaces. *Ergod. Th. & Dynam. Sys.* **30** (2010), 1529–1542.
[**11**] G. Pedersen. *Analysis Now (Graduate Texts in Mathematics, 118)*. Springer, New York, 1989.
[**12**] P. Walters. *An Introduction to Ergodic Theory*. Springer, Berlin, 2000.

https://doi.org/10.1017/etds.2016.45 Published online by Cambridge University Press