Ergod. Th. & Dynam. Sys. (2005), **25**, 1173–1208
doi:10.1017/S0143385704000872
© *2005 Cambridge University Press*
Printed in the United Kingdom

Topological entropy for divergence points

CHEN ERCAI†, TASSILO KÜPPER‡ and SHU LIN†

† *Department of Mathematics, Nanjing Normal University, Nanjing 210097, People's Republic of China*
(e-mail: ecchen@njnu.edu.cn, shulin003@sohu.com)
‡ *Mathematical Institute, University of Cologne, D-50931 Cologne, Germany*
(e-mail: kuepper@mi.uni-koeln.de)

(*Received* 1 *May* 2004 *and accepted in revised form* 20 *September* 2004)

Abstract. Let X be a compact metric space, f a continuous transformation on X, and Y a vector space with linear compatible metric. Denote by $M(X)$ the collection of all the probability measures on X. For a positive integer n, define the nth empirical measure $L_n : X \mapsto M(X)$ as

$$L_n x = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{f^k x},$$

where δ_x denotes the Dirac measure at x. Suppose $\Xi : M(X) \mapsto Y$ is continuous and affine with respect to the weak topology on $M(X)$. We think of the composite

$$\Xi \circ L_n : X \xrightarrow{L_n} M(X) \xrightarrow{\Xi} Y$$

as a continuous and affine deformation of the empirical measure L_n. The set of divergence points of such a deformation is defined as

$$D(f, \Xi) = \{x \in X \mid \text{the limit of } \Xi L_n x \text{ does not exist}\}.$$

In this paper we show that for a continuous transformation satisfying the specification property, if $\Xi(M(X))$ is a singleton, then set of divergence points is empty, i.e. $D(f, \Xi) = \emptyset$, and if $\Xi(M(X))$ is not a singleton, then the set of divergence points has full topological entropy, i.e.

$$h_{\text{top}}(D(f, \Xi)) = h_{\text{top}}(f).$$

1. *Introduction and definitions*

Multifractal analysis studies the size of the local characters of dynamical systems or fractal measures, e.g. the Hausdorff dimension or the topological entropy of the following level sets,

$$K_\alpha = \left\{ x \in X \;\middle|\; \lim_{n\to\infty} \phi_n(x) = \alpha \right\},$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

where X is the underlying set of a dynamical system or the support of a fractal measure and $\{\phi_n : X \mapsto \mathbb{R}\}$ is a sequence of functions. Since the limit may not exist, we are interested in the structure of the set of points for which the above limit does not exist, the so-called set of divergence points. Our result is that although the set of divergence points is usually small in measure, it may have full topological entropy. We give the background of the our research motivation, notations and definitions in this section. Then we study the topological entropy of sup sets in §2 and study the topological entropy of divergence points in §3.

1.1. *Background of multifractal analysis.* Let X be a metric space (typically the underlying set of a dynamical system or the support of a fractal measure) and let $\phi : X \mapsto \mathbb{R}$ be a function associating some dynamical characteristics with each point $x \in X$. Multifractal analysis studies the level sets

$$K_\alpha = \{x \in X \mid \phi(x) = \alpha\}, \quad \alpha \in \mathbb{R},$$

of ϕ. Usually, one might study the Hausdorff dimension ($\dim_{\mathrm{H}}$) or the topological entropy (h_{top}) of the above set. $g_1(\alpha) = \dim_{\mathrm{H}}(K_\alpha)$, $g_2(\alpha) = h_{\mathrm{top}}(K_\alpha)$ and similarly defined functions are known as multifractal spectra. Since multifractal spectra of different local characteristics ϕ associated with a dynamical system or a fractal measure encode important information about the underlying dynamics and geometry, the study of multifractal spectra plays an important role in multifractal analysis.

Often the local characteristics of dynamical systems or fractal measures are defined in terms of limit processes, i.e. $\phi : X \mapsto \mathbb{R}$ has the form

$$\phi(x) = \lim_{n\to\infty} \phi_n(x)$$

for a sequence of functions $\phi_n : X \mapsto \mathbb{R}$. For example, if f is a continuous transformation of a compact metric space (X, d) and $\varphi : X \mapsto \mathbb{R}$ is a continuous function, then the ergodic average of φ at the point x is defined by the following limit

$$\phi(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k x).$$

As a consequence of the well-known Birkhoff ergodic theorem, if μ is ergodic with respect to f then $\phi(x) = \int \varphi \, d\mu$ for μ-almost every $x \in X$. However, the limits involved do not necessarily exist at all points. Points for which the limits do not exist are called divergence points. The collection of all such points

$$\left\{ x \in X \;\middle|\; \text{the limit } \lim_{n\to\infty} \phi_n(x) \text{ does not exist} \right\}$$

is defined as the set of divergence points. The set of divergence points has until very recently been considered of little interest in dynamical systems and geometric measure theory. As a result, they are always *a priori* discarded. However, recent work [**2, 4–6**] has changed such attitudes. They have shown that in many cases the set of divergence points has full dimension, i.e.

$$\dim_{\mathrm{H}} \left\{ x \in X \;\middle|\; \text{the limit } \lim_{n\to\infty} \phi_n(x) \text{ does not exist} \right\} = \dim_{\mathrm{H}} X.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

In [**7–9**], Olsen and Winter unified and extended many results in multifractal analysis of local characteristics of dynamical systems and fractal measures by introducing and developing a framework on subshifts of finite type which includes the notions of:

- divergence points;
- relative (or conditional) multifractal spectra;
- mixed multifractal spectra.

Olsen and Winter's framework is based on so-called deformations of empirical measures and the above spectra are obtained by considering different deformations, i.e. let $\sum$ be a subshift of finite type, which is associated with a conformal iterated function system $((V, E), (U_i)_{i\in V}, (X_i)_{i\in V}, (S_e)_{e\in E})$, let $M\left(\sum\right)$ denote the family of probability measures on $\sum$, and let $F : \sum \mapsto \sum$ be the shift map. For a positive integer n, the nth empirical measure $L_n : \sum \mapsto M\left(\sum\right)$ is defined by

$$L_n w = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{F^k w} \quad \forall w \in \sum,$$

where δ_x denotes the Dirac measure at x. Consider continuous deformations of L_n, i.e. pairs (Y, Ξ) where Y is a metric space and

$$\Xi : M\left(\sum\right) \mapsto Y$$

is continuous with respect to the weak topology. The composite

$$\Xi \circ L_n : \sum \xrightarrow{L_n} M\left(\sum\right) \xrightarrow{\Xi} Y$$

is called a continuous deformation of the empirical measure L_n. The multifractal spectrum of the deformation (Y, Ξ) is defined by

$$g(\alpha) = \dim\left\{w \in \sum \;\middle|\; \lim_{n\to\infty} \Xi L_n w = \alpha\right\}.$$

Olsen and Winter established a variational principle for the multifractal spectra of arbitrary affine and continuous deformations and studied the spectra of sup sets, eq sets, and sub sets, i.e. the collection of points such that $A(\Xi L_n w)$ (where $A(y_n)$ is the set of accumulation points of the sequence $(y_n)_n$) is contained in, is equal to, or contains a subset C of Y. The research on such sets provides a complete description of the set of divergence points.

We need to point out that in symbolic spaces, when we consider the Hausdorff dimension of the level sets we are in fact studying the topological entropy of them since in this case for every α one has $\dim_{\mathrm{H}}(K_\alpha) = |E|h_{\mathrm{top}}(K_\alpha)$, where $|E|$ is the number of elements in E. One can also see from the various works [**2, 4–9**] that for the systems they considered, there is a close relation between the Hausdorff dimension and the topological entropy of certain level sets.

Since the calculation of Hausdorff dimension always needs special information about the metric and the differential structure of the system, it seems that the study of the topological entropy multifractal spectra may be carried out in a more general setting than the study of Hausdorff dimension multifractal spectra. In [**10**], Takens and Verbitskiy studied the topological entropy of the level sets

$$K_\alpha = \left\{x \in X \;\middle|\; \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \varphi(f^i(x)) = \alpha\right\}, \quad \alpha \in \mathbb{R},$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

where f is a continuous transformation of a compact metric space (X, d) and φ is an arbitrary continuous function on X. They obtained the following variational principle for transformations satisfying the specification property:

$$h_{\text{top}}(f, K_\alpha) = \sup\left\{h_\mu(f) \;\middle|\; \mu \text{ is invariant and } \int \varphi \, d\mu = \alpha\right\}.$$

Observe that the system (X, f) is more general than $\left(\sum, F\right)$, where $\sum$ is a subshift of finite type and F is the shift map on $\sum$, or (J, f), where f is an expanding $C^{1+\delta}$-conformal topological mixing map and J is the corresponding repeller. We wondered what we can say about the divergence points of such a system (X, f). This becomes the beginning of our investigation.

1.2. *Topological entropy of non-compact sets.* Topological entropy is an important quantity describing the chaoticity of a dynamical system. It was first defined on compact metric spaces. In 1973, Bowen [**1**] generalized the topological entropy to non-compact or non-invariant sets by using separated sets or span sets. In this section, we use the following equivalent definition of topological entropy [**10**].

Definition 1.1. (Topological entropy) Let (X, d) be a compact metric space and $f : X \mapsto X$ a continuous transformation. For any $n \in \mathbb{N}$, we define a new metric d_n on X as follows:

$$d_n(x, y) = \max\{d(f^k(x), f^k(y)) \mid k = 0, 1, \ldots, n-1\}$$

and for every $\epsilon > 0$ we denote by $B_n(x, \epsilon)$ an open ball of radius ϵ in the metric d_n around x, i.e.

$$B_n(x, \epsilon) = \{y \in X \mid d_n(x, y) < \epsilon\}.$$

Let a subset $Z \subseteq X$ be given. Fix $\epsilon > 0$. We say that an at most countable collection of balls $\Gamma = \{B_{n_i}(x_i, \epsilon)\}_i$ covers Z if $Z \subseteq \bigcup_i B_{n_i}(x_i, \epsilon)$. For such a Γ, put $n(\Gamma) = \min_i n_i$. Let $s \geq 0$ and define

$$m(Z, s, N, \epsilon) = \inf_\Gamma \sum_i \exp(-s n_i),$$

where the infinum is taken over all collections $\Gamma = \{B_{n_i}(x_i, \epsilon)\}_i$ covering Z such that $n(\Gamma) \geq N$. The quantity $m(Z, s, N, \epsilon)$ does not decrease with N, hence the following limit exists

$$m(Z, s, \epsilon) = \lim_{N\to\infty} m(Z, s, N, \epsilon) = \sup_{N>0} m(Z, s, N, \epsilon).$$

It is easy to show that there exists a critical value of the parameter s, which we denote by $h_{\text{top}}(f, Z, \epsilon)$, where $m(Z, s, \epsilon)$ jumps from $+\infty$ to 0, i.e.

$$m(Z, s, \epsilon) = \begin{cases} +\infty, & s < h_{\text{top}}(f, Z, \epsilon) \\ 0, & s > h_{\text{top}}(f, Z, \epsilon). \end{cases}$$

Then the following limit exists

$$h_{\text{top}}(f, Z) = \lim_{\epsilon\to 0} h_{\text{top}}(f, Z, \epsilon).$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

We call $h_{\text{top}}(f, Z)$ the *topological entropy* of f restricted to Z, or simply, the topological entropy of Z, denoted by $h_{\text{top}}(Z)$ when there is no confusion about f.

We give some of the basic properties and results on topological entropy of non-compact sets that are needed in the following proofs. The reader can refer to [**10**] for the proofs and references therein.

THEOREM 1.2. *The topological entropy as defined above satisfies the following:*
(1) $h_{\text{top}}(f, Z_1) \le h_{\text{top}}(f, Z_2)$ *for any* $Z_1 \subseteq Z_2 \subseteq X$;
(2) $h_{\text{top}}(f, Z) = \sup_i h_{\text{top}}(f, Z_i)$, *where* $Z = \bigcup_{i=1}^{\infty} Z_i \subseteq X$.

THEOREM 1.3. *Let* $f : X \mapsto X$ *be a continuous transformation on a compact metric space and* ν *an ergodic invariant measure. For* $\epsilon > 0$, $\delta > 0$ *denote by* $N_f^{\nu}(\delta, \epsilon, n)$ *the minimal number of* ϵ*-balls in the* d_n*-metric which cover a set of measure at least* $1 - \delta$. *Then, for each* $\delta \in (0, 1)$ *we have*

$$h_{\nu}(f) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log N_f^{\nu}(\delta, \epsilon, n) = \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log N_f^{\nu}(\delta, \epsilon, n).$$

Remark 1.4. Suppose ν is ergodic and $Y \subseteq X$ is such that $\nu(Y) > 1 - \delta$. Denote by $S(Y, \epsilon, n)$ the maximal cardinality of an (n, ϵ)-separated set in Y. Then we have $S(Y, \epsilon, n) \ge N_f^{\nu}(\delta, \epsilon, n)$.

THEOREM 1.5. (Entropy distribution principle) *Let* $f : X \mapsto X$ *be a continuous transformation. Suppose a set* $Z \subseteq X$ *and a constant* $s \ge 0$ *are such that for any* $\epsilon > 0$ *one can find a Borel probability measure* $\mu = \mu_{\epsilon}$ *satisfying:*
(1) $\mu_{\epsilon}(Z) > 0$;
(2) $\mu_{\epsilon}(B_n(x, \epsilon)) \le C(\epsilon) e^{-ns}$ *for some constant* $C(\epsilon) > 0$ *and every ball* $B_n(x, \epsilon)$ *such that* $B_n(x, \epsilon) \cap Z \ne \varnothing$.
Then $h_{\text{top}}(f, Z) \ge s$.

1.3. *Specification property and deformation of measures.* From now on we consider a system that satisfies the specification property.

Definition 1.6. (Specification property) A continuous transformation $f : X \mapsto X$ satisfies specification property if for any $\epsilon > 0$ there exists an integer $m = m(\epsilon)$ such that for arbitrary finite intervals $I_j = [a_j, b_j] \subset \mathbb{N}$, $j = 1, 2, \ldots, k$, such that

$$\operatorname{dist}(I_i, I_j) \ge m(\epsilon), \quad i \ne j,$$

and any $x_1, \ldots, x_k$ in X, there exists a point $x \in X$ such that

$$d(f^{p+a_j}(x), f^p x_j) < \epsilon$$

for all $p = 0, 1, \ldots, b_j - a_j$ and every $j = 1, \ldots, k$.

Definition 1.7. (Linearly compatible metric) If Y is a vector space and d is a metric in Y, then we say that d is linearly compatible if:
(1) for all $x_1, x_2, y_1, y_2 \in Y$, we have $d(x_1 + x_2, y_1 + y_2) \le d(x_1, y_1) + d(x_2, y_2)$;
(2) for all $x, y \in Y$ and all $\lambda \in \mathbb{R}$ we have $d(\lambda x, \lambda y) \le |\lambda| d(x, y)$.

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

One can observe immediately that if d is induced by a norm, then d is linearly compatible.

We denote by $M(X)$ the collection of all the probability measures on X and $M(f, X)$ the collection of all the f invariant probability measures on X. Since X is a compact metric space, we have $M(X)$ is a compact metric space under the weak topology. Also the weak topology in $M(X)$ is induced by the following metric

$$d_P(\mu, \nu) = \sup_{g \in \mathrm{Lip}(1)} \left| \int g \, d\mu - \int g \, d\nu \right|,$$

where Lip(1) denotes the family of Lipschitz functions $g : X \mapsto \mathbb{R}$.

For a positive integer n, define the nth empirical measure $L_n : X \mapsto M(X)$ as

$$L_n x = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{f^k x},$$

where δ_x denotes the Dirac measure at x. Consider continuous and affine deformations of L_n, i.e. pairs (Y, Ξ) where Y is a vector space with linearly compatible metric and $\Xi : M(X) \mapsto Y$ is continuous and affine with respect to the weak topology on $M(X)$. We think of the composite

$$\Xi \circ L_n : X \xrightarrow{L_n} M(X) \xrightarrow{\Xi} Y$$

as a continuous and affine deformation of the empirical measure L_n. The set of divergence points of such a deformation is defined as

$$D(f, \Xi) = \{x \in X \mid \text{the limit of } \Xi L_n x \text{ does not exist}\}.$$

Since the set of divergence points is typically large, we do not restrict ourselves to the study of points $x \in X$ for which the limit $\lim_{n\to\infty} \Xi L_n x$ exists, but develop a more general framework which also includes divergence points. For a sequence $(y_n)_n$ in a metric space X, let $A(y_n)$ denote the set of accumulation points of the sequence $(y_n)_n$, i.e.,

$$A(y_n) = \{y \in Y \mid \text{there exists a subsequence } (y_{n_k})_k \text{ such that } y_{n_k} \to y\}.$$

For a subset C of Y, we consider the set, $\Delta(C)$, of points $x \in X$ for which $A(\Xi L_n x) \subseteq C$, i.e.

$$\Delta(C) = \left\{ x \in X \;\middle|\; \lim_{n\to\infty} \mathrm{dist}(\Xi L_n x, C) = 0 \right\} = \{x \in X \mid A(\Xi L_n x) \subseteq C\}.$$

$\Delta(C)$ is called a sup set, which we will study in §2.

For a different choice of the deformation (Y, Ξ) we obtain different spectra and sets of divergence points.

- *The spectrum and the set of divergence points of ergodic averages.* Let $\varphi : X \mapsto \mathbb{R}$ be continuous and define $\Xi : M(X) \mapsto \mathbb{R}$ by

$$\Xi : \mu \mapsto \int \varphi \, d\mu.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

In this case we obtain for $C \subseteq \mathbb{R}$,

$$\Delta(C) = \left\{ x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}\left(\frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k x), C \right) = 0 \right\} \quad \text{and}$$

$$D(f, \Xi) = \left\{ x \in X \;\middle|\; \text{the limit } \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k x) \text{ does not exist} \right\}.$$

- *The spectrum and the set of divergence points of empirical measures.* Define $\Xi : M(X) \mapsto M(X)$ by

$$\Xi : \mu \mapsto \mu.$$

In this case we obtain for $C \subseteq M(X)$,

$$\Delta(C) = \left\{ x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}\left(\frac{1}{n} \sum_{k=0}^{n-1} \delta_{f^k x}, C \right) = 0 \right\} \quad \text{and}$$

$$D(f, \Xi) = \left\{ x \in X \;\middle|\; \text{the limit } \frac{1}{n} \sum_{k=0}^{n-1} \delta_{f^k x} \text{ does not exist} \right\}.$$

- *The spectrum and the set of divergence points of local Lyapunov exponents.* Let X be a differentiable manifold and $f : X \mapsto X$ be a C^1 map. The local Lyapunov exponent of f at the point x is defined by

$$\chi(x) = \lim_{n\to\infty} \frac{1}{n} \log |(Df^n)(x)|.$$

Define $\Xi : M(X) \mapsto \mathbb{R}$ by

$$\mu \to \int Df \, d\mu.$$

In this case we obtain for $C \subseteq \mathbb{R}$

$$\Delta(C) = \left\{ x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}\left(\frac{1}{n} \log |(Df^n)(x)|, C \right) = 0 \right\} \quad \text{and}$$

$$D(f, \Xi) = \left\{ x \in X \;\middle|\; \text{the limit } \frac{1}{n} \log |(Df^n)(x)| \text{ does not exist} \right\}.$$

- *The mixed spectrum and the set of divergence points of ergodic averages of arbitrary families of continuous functions.* Assume that the family of maps $(M(X) \mapsto \mathbb{R} : \mu \to \int \varphi_i \, d\mu)_{i\in I}$ is totally bounded. Define $\Xi : M(X) \mapsto l^\infty(I)$ by

$$\Xi : \mu \to \left(\int \varphi_i \, d\mu \right)_{i\in I}.$$

In this case we obtain for $C \subseteq l^\infty(I)$,

$$\Delta(C) = \left\{ x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}\left(\left(\frac{1}{n} \sum_{k=0}^{n-1} \varphi_i(f^k x) \right)_{i\in I}, C \right) = 0 \right\} \quad \text{and}$$

$$D(f, \Xi) = \left\{ x \in X \;\middle|\; \text{the limit } \left(\frac{1}{n} \sum_{k=0}^{n-1} \varphi_i(f^k x) \right)_{i\in I} \text{ does not exist} \right\}.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Definition 1.8. ($\Lambda(\Delta(C))$) For $C \subseteq Y$ and any $\delta > 0$ and $n \in \mathbb{N}$ put

$$P(C, \delta, n) = \{x \in X \mid \Xi L_n x \in B(C, \delta)\}.$$

Fix some $\epsilon > 0$ and let $N(C, \delta, n, \epsilon)$ be the minimal number of balls $B_n(x, \epsilon)$, which is necessary for covering the set $P(C, \delta, n)$ (if $P(C, \delta, n) = \varnothing$, then put $N(C, \delta, n, \epsilon) = 1$). Observe that $N(C, \delta, n, \epsilon)$ will not increase as δ decreases and will not decrease as ϵ decreases, so the following limit denoted by $\Lambda(\Delta(C))$ exists

$$\Lambda(\Delta(C)) = \lim_{\epsilon \to 0} \lim_{\delta \to 0} \liminf_{n \to \infty} \frac{1}{n} \log N(C, \delta, n, \epsilon).$$

Let $C(Y) = \{C \subseteq Y \mid C \text{ is closed and convex}\}$. We will always equip $C(Y)$ with the Hausdorff metric. Below are three observations that will frequently be used in the following sections.

OBSERVATION 1.9. *There exists a constant M such that*

$$d(0, \Xi\mu) \leq M, \quad \text{for all } \mu \in M(X).$$

Proof. This follows from the fact that $\Xi : M(X) \mapsto Y$ is continuous and $M(X)$ is compact. □

OBSERVATION 1.10. *For a positive integer $n \in \mathbb{N}$, given $x, x' \in X$, then $d_n(x, x') \leq \epsilon$ implies $d(L_n x, L_n x') \leq \epsilon$.*

Proof. This is indeed the case since

$$\begin{aligned} d(L_n x, L_n x') &= \sup_{g \in \mathrm{Lip}(1)} \left| \int g \, d(L_n x) - \int g \, d(L_n x') \right| \\ &= \sup_{g \in \mathrm{Lip}(1)} \left| \frac{1}{n} \sum_{i=0}^{n-1} g(f^i x) - \frac{1}{n} \sum_{i=0}^{n-1} g(f^i x') \right| \\ &\leq \frac{1}{n} \sum_{i=0}^{n-1} \sup_{g \in \mathrm{Lip}(1)} |g(f^i x) - g(f^i x')| \\ &\leq \frac{1}{n} \sum_{i=0}^{n-1} d(f^i x, f^i x') \\ &\leq \frac{1}{n} \sum_{i=0}^{n-1} \epsilon = \epsilon. \end{aligned}$$

□

OBSERVATION 1.11. *Let $V(\Xi, \epsilon) = \sup_{\mu, \nu \in M(X), d(\mu, \nu) \leq \epsilon} d(\Xi\mu, \Xi\nu)$. Then $V(\Xi, \epsilon)$ goes to zero as ϵ goes to zero.*

Proof. This observation follows from the fact that $\Xi : M(X) \mapsto Y$ is continuous and $M(X)$ is compact. □

2. *Topological entropy of sup sets*

In this section, we try to obtain information about the topological entropy of sup set $\Delta(C)$. Our main result is the following theorem.

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

THEOREM 2.1. *Suppose $f : X \mapsto X$ is a transformation of a compact metric space (X, d) which satisfies the specification property. Let Y be a metric space with linearly compatible metric, $\Xi : M(X) \mapsto Y$ be continuous and affine, and $L_n : X \mapsto M(X)$ be defined as before. Suppose C is a closed convex subset of Y and let*

$$\Delta(C) = \left\{ x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}(\Xi L_n x, C) = 0 \right\}.$$

With the notation

$$H(\Delta(C)) = \sup_{\mu \in M(X,f), \Xi\mu \in C} h_\mu(f),$$

we then have:

(1)
$$h_{\text{top}}(\Delta(C)) = H(\Delta(C)) = \Lambda(\Delta(C)),$$
where h_{top} denotes the topological entropy of a set, $M(X, f)$ is the invariant probability measure with respect to f, and $\Lambda(\Delta(C))$ is defined as before;

(2) *the map $\Lambda\Delta : C(Y) \mapsto \mathbb{R}$ as $C \to h_{\text{top}}(\Delta(C)) = \Lambda(\Delta(C))$ is upper semi-continuous. Furthermore, for $C_1, C_2 \in C(Y)$ with $\Delta(C_i) \neq \varnothing$ $(i = 1, 2)$,*
$$\Lambda(\Delta(s_1 C_1 + s_2 C_2)) \geq s_1 \Lambda(\Delta(C_1)) + s_2 \Lambda(\Delta(C_2))$$
holds for arbitrary $s_1, s_2 \geq 0$ with $s_1 + s_2 = 1$, here $sC = \{sx \mid x \in C\}$ and $C_1 + C_2 = \{x_1 + x_2 \mid x_i \in C_i\}$ for $C, C_1, C_2 \in C(Y)$ and $s \in \mathbb{R}$.

Proof of (1). The proof follows [**10**]. We divided the proof into two parts. The first is the proof of $h_{\text{top}}(\Delta(C)) \leq \Lambda(\Delta(C)) \leq H(\Delta(C))$ and then the opposite inequalities. Without loss of generality we assume $\Delta(C) \neq \varnothing$.

Part 1. Proof of $h_{\text{top}}(\Delta(C)) \leq \Lambda(\Delta(C)) \leq H(\Delta(C))$.
First we prove that $h_{\text{top}}(\Delta(C)) \leq \Lambda(\Delta(C))$. Let

$$G(C, \delta, k) = \bigcap_{n=k}^{+\infty} P(C, \delta, n) = \bigcap_{n=k}^{+\infty} \{x \in X \mid \Xi L_n x \in B(C, \delta)\}.$$

Then for each $\delta > 0$,

$$\Delta(C) = \left\{ x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}(\Xi L_n x, C) = 0 \right\} \subseteq \bigcup_{k=1}^{+\infty} G(C, \delta, k).$$

Since $h_{\text{top}}(f, \Delta(C), \epsilon) \leq \sup_k h_{\text{top}}(f, G(C, \delta, k), \epsilon)$, it is enough to show that $h_{\text{top}}(f, G(C, \delta, k), \epsilon) \leq \Lambda(\Delta(C))$ holds for any $k \geq 1$.

For arbitrary $k \geq 1$, $G(C, \delta, k)$ as a subset of $P(C, \delta, n)$ for $n \geq k$ can be covered by $N(C, \delta, n, \epsilon)$ balls $B_n(x, \epsilon)$ for all $n \geq k$. Therefore, for every $s \geq 0$,

$$m(G(C, \delta, k), s, \epsilon) \leq N(C, \delta, n, \epsilon) \exp(-ns).$$

Let $s > \Lambda(\Delta(C))$ and $\gamma = \frac{1}{2}(s - \Lambda(\Delta(C))) > 0$. Since

$$\Lambda(\Delta(C)) = \lim_{\epsilon\to 0} \lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log N(C, \delta, n, \epsilon),$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

for all sufficiently small $\epsilon > 0$ and $\delta > 0$, there is a monotonic sequence of integers $n_l \to \infty$ such that

$$N(C, \delta, n_l, \epsilon) \leq \exp(n_l(\Lambda(\Delta(C)) + \gamma)).$$

Without loss of generality we can assume that $n_l \geq k$. Then we can get

$$m(G(C, \delta, k), s, \epsilon) \leq N(C, \delta, n_l, \epsilon) \exp(-n_l s) \leq \exp(-n_l \gamma)$$

and hence $m(G(C, \delta, k), s, \epsilon) = 0$. Therefore $h_{\text{top}}(f, G(C, \delta, k), \epsilon) \leq s$ and

$$h_{\text{top}}(f, \Delta(C), \epsilon) \leq \sup_k h_{\text{top}}(f, G(C, \delta, k), \epsilon) \leq s.$$

So $h_{\text{top}}(\Delta(C)) = \lim_{\epsilon \to 0} h_{\text{top}}(f, \Delta(C), \epsilon) \leq s$. The conclusion $h_{\text{top}}(\Delta(C)) \leq \Lambda(\Delta(C))$ follows since $s > \Lambda(\Delta(C))$ is arbitrary.

Then we are going to show that $\Lambda(\Delta(C)) \leq H(\Delta(C))$. It suffices to show that for any $\gamma > 0$, there is a measure $\mu \in M(f, X, \Xi, C) = \{\mu \mid \Xi\mu \in C, \mu \in M(f, X)\}$ such that

$$h_\mu(f) \geq \Lambda(\Delta(C)) - \gamma.$$

Recall that

$$\Lambda(\Delta(C)) = \lim_{\epsilon \to 0} \lim_{\delta \to 0} \liminf_{n \to \infty} \frac{1}{n} \log N(C, \delta, n, \epsilon),$$

hence there is a sufficiently small $\epsilon_0 > 0$ such that for all $\epsilon < \epsilon_0$, the following holds

$$\lim_{\delta \to 0} \liminf_{n \to \infty} \frac{1}{n} \log N(C, \delta, n, \epsilon) > \Lambda(\Delta(C)) - \frac{\gamma}{3}.$$

Put $\epsilon_k = \epsilon_0/2^k$, $k \in \mathbb{N}$. For any $k \geq 1$, one can find a sufficiently small δ_k such that $\delta_k \to 0$ and

$$\liminf_{n \to \infty} \frac{1}{n} \log N(C, \delta_k, n, \epsilon_k) > \Lambda(\Delta(C)) - \frac{2}{3}\gamma.$$

Then for each k we can choose some $n_k \in \mathbb{N}$, $n_k \to \infty$ such that

$$N_k = N(C, \delta_k, n_k, \epsilon_k) \geq \exp(n_k(\Lambda(\Delta(C)) - \gamma)).$$

Let C_k be the centers of some minimal covering of $P(C, \delta_k, n_k)$ by balls $B_{n_k}(x, \epsilon_k)$. Obviously $\#(C_k) = N_k$ and $B_{n_k}(x, \epsilon_k) \cap P(C, \delta_k, n_k) \neq \varnothing$ for every $x \in C_k$. Define a probability measure

$$\mu_k = \frac{1}{N_k} \sum_{x \in C_k} L_{n_k} x$$

and let μ be some limit point for the sequence μ_k, then we have μ is an invariant measure. Also we have $\mu \in M(f, X, \Xi, C)$. In fact, for each $x \in C_k$, let $y(x)$ be a corresponding point such that $y(x) \in B_{n_k}(x, \epsilon) \cap P(C, \delta_k, n_k)$ and $y_1(x) \in C$ such that $\Xi L_{n_k} y(x) \in B(y_1(x), \delta_k)$, then since $\Xi : M(X) \mapsto Y$ is continuous and affine and Y is a

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

vector space with linear compatible metric, we have

$$\begin{aligned} d\left(\Xi\mu_k, \frac{1}{N_k}\sum_{x\in C_k} y_1(x)\right) &\le \frac{1}{N_k}\sum_{x\in C_k} d(\Xi L_{n_k}x, y_1(x)) \\ &\le \frac{1}{N_k}\sum_{x\in C_k} (d(\Xi L_{n_k}x, \Xi L_{n_k}y(x)) + d(\Xi L_{n_k}y(x), y_1(x))) \\ &\le \frac{1}{N_k}\sum_{x\in C_k} (V(\Xi,\epsilon_k) + \delta_k) \\ &\le V(\Xi,\epsilon_k) + \delta_k, \end{aligned}$$

where $V(\Xi,\epsilon) = \sup_{\mu,\nu\in M(X), d(\mu,\nu)\le\epsilon} d(\Xi\mu, \Xi\nu)$. The third inequality follows by Observation 1.10. By Observation 1.11, $V(\Xi,\epsilon_k)$ goes to zero as k goes to infinity. Also since C is a closed convex subset of Y, we conclude that $(1/N_k)\sum_{x\in C_k} y_1(x) \in C$. So we have

$$\operatorname{dist}(\Xi\mu_k, C) \le V(\Xi,\epsilon_k) + \delta_k \to 0$$

as k goes to infinity. Hence,

$$\operatorname{dist}(\Xi\mu, C) = \lim_{k\to\infty} \operatorname{dist}(\Xi\mu_k, C) = 0.$$

Repeating the second half of the proof of the classical variational principle [**11**] we conclude that

$$h_\mu(f) \ge \limsup_{k\to\infty} \frac{1}{n_k}\log N_k \ge \liminf_{k\to\infty} \frac{1}{n_k}\log N_k \ge \Lambda(\Delta(C)) - \gamma,$$

which finishes the proof.

Part 2. Proof of $h_{\text{top}}(\Delta(C)) \ge \Lambda(\Delta(C)) \ge H(\Delta(C))$.

LEMMA 2.2. *We have* $\Lambda\Delta : C(Y) \mapsto \mathbb{R}$ *is upper semi-continuous.*

Proof. $C(Y) = \{A \subseteq Y \mid A \text{ is closed and convex}\}$. Endow $C(Y)$ with Hausdorff metric D. For all $A \in C(Y)$, by the definition of $\Lambda(\Delta(A))$, for all $\gamma > 0$, there is $\epsilon_0 > 0$ and $\delta_0 > 0$ such that for all $\epsilon < \epsilon_0$ and $\delta < \delta_0$,

$$\liminf_{n\to\infty} \frac{1}{n}\log N(A,\delta,n,\epsilon) < \Lambda(\Delta(A)) + \gamma.$$

Let $a_\gamma = \delta_0/2$, then for all $K \in C(Y)$ satisfying $D(K, A) < a_\gamma$, given $r < a_\gamma$, we have

$$B(K, r) \subseteq B(K, a_\gamma) \subseteq B(A, \delta_0).$$

Hence, we have

$$N(K, r, n, \epsilon) \le N(A, \delta_0, n, \epsilon).$$

So for all $\epsilon < \epsilon_0$,

$$\liminf_{n\to\infty} \frac{1}{n}\log N(K,r,n,\epsilon) \le \liminf_{n\to\infty} \frac{1}{n}\log N(A,\delta_0,n,\epsilon) < \Lambda(\Delta(A)) + \gamma.$$

Let r and ϵ go to zero, for all $K \in C(Y)$ satisfying $D(K, A) < a_\gamma$ we have

$$\Lambda(\Delta(K)) < \Lambda(\Delta(A)) + \gamma.$$

That is $\Lambda\Delta : C(Y) \mapsto \mathbb{R}$ is upper semi-continuous. □

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Proof of $\Lambda(\Delta(C)) \geq H(\Delta(C))$. It suffices to prove that for arbitrary $\gamma > 0$, for all $\mu \in M(f, X)$ such that $\Xi\mu \in C$, $\Lambda(\Delta(C)) \geq h_\mu(f) - 4\gamma$ holds.

For arbitrary $\gamma > 0$, let $\epsilon > 0$ and $\delta > 0$ be so small such that the following hold:

(1) $\gamma > \delta$;
(2) for all $\mu, \nu \in M(f, X)$ with $d(\mu, \nu) < \epsilon \Longrightarrow d(\Xi\mu, \Xi\nu) < \delta$;
(3) $\liminf_{n\to\infty}(1/n)\log N(C, 4\delta, n, \epsilon) < \Lambda(\Delta(C)) + \gamma$.

Also taking any $\mu \in M(f, X)$, $\Xi\mu \in C$, we can approximate μ by an invariant measure ν with the following properties:

(1) $\nu = \sum_{i=1}^{k} \lambda_i \nu_i$, where $\lambda_i > 0$ with $\sum_{i=1}^{k} \lambda_i = 1$, and $\nu_i (1 \leq i \leq k)$ is ergodic;
(2) $h_\nu(f) \geq h_\mu(f) - \gamma$;
(3) $d(\Xi\mu, \Xi\nu) < \delta$.

Since each ν_i is ergodic, for almost every point x, $L_n x \to \nu_i$. Let

$$Y_i(N) = \{x \in X \mid d(\Xi L_n x, \Xi\nu_i) < \delta,\ \forall n \geq N\}.$$

For $N \in \mathbb{N}$ large enough, $Y_i(N)$ has ν_i measure of at least $1 - \gamma$. By Theorem 1.3, for each $i \in \{1, 2, \ldots, k\}$, there is a N_i such that for all $n_i > N_i$ the minimal number of 4ϵ-balls in d_{n_i} metric, which is necessary to cover $Y_i(N)$ is greater than or equal to $\exp(n_i(h_{\nu_i}(f) - \gamma))$. This implies that the cardinality of a maximal $(n_i, 4\epsilon)$-separated set in $Y_i(N)$ is greater than or equal to $\exp(n_i(h_{\nu_i}(f) - \gamma))$. Choose N_0 large enough such that for all $n > N_0$, $n_i = [\lambda_i n] > \max\{\max_{1\leq i\leq k}\{N_i\}, N\}$. Let $C(n_i, 4\epsilon)$ be some maximal $(n_i, 4\epsilon)$-separated set of $Y_i(N)$. For each $(x_1, \ldots, x_k)$ such that $x_i \in C(n_i, 4\epsilon)$, by the specification property of f, there is a point $y = y(x_1, \ldots, x_k)$ satisfying $d_{n_i}(f^{\sum_{k=0}^{i-1}(n_k + m(\epsilon))} y, x_i) < \epsilon$. Let $m = m(\epsilon)$ and $\hat{n} = m(k-1) + \sum_{i=1}^{k} n_i$. We have that different $(x_1, \ldots, x_k) \in C_{n_1} \times C_{n_2} \times \cdots \times C_{n_k}$ correspond to different points $y(x_1, \ldots, x_k)$. In fact, take such two points $y = y(x_1, \ldots, x_k)$ and $y' = y'(x'_1, \ldots, x'_k)$, one can easily show that

$$d_{\hat{n}}(y, y') > 2\epsilon.$$

Now we are in a position to estimate the distance between $\Xi L_{\hat{n}}(y)$ and $\Xi\mu$. Constant M follows Observation 1.10:

$$\begin{aligned}
&d(\Xi L_{\hat{n}} y, \Xi\mu)\\
&\leq d(\Xi L_{\hat{n}} y, \Xi\nu) + d(\Xi\mu, \Xi\nu)\\
&\leq d(\Xi L_{\hat{n}} y, \Xi\nu) + \delta\\
&\leq d\left(\Xi L_{\hat{n}} y, \Xi \sum_{i=1}^{k} \lambda_i L_{n_i} x_i\right) + d\left(\Xi \sum_{i=1}^{k} \lambda_i L_{n_i} x_i, \Xi \sum_{i=1}^{k} \lambda_i \nu_i\right) + \delta\\
&= d\left(\Xi L_{\hat{n}} y, \Xi \sum_{i=1}^{k} \lambda_i L_{n_i} x_i\right) + \sum_{i=1}^{k} \lambda_i d(\Xi L_{n_i} x_i, \Xi\nu_i) + \delta\\
&\leq d\left(\Xi L_{\hat{n}} y, \Xi \sum_{i=1}^{k} \lambda_i L_{n_i} x_i\right) + 2\delta\\
&\leq d\left(\Xi\left(\frac{n_1}{\hat{n}} L_{n_1} y + \frac{m}{\hat{n}} L_m f^{n_i} y + \cdots + \frac{n_k}{\hat{n}} L_{n_k} f^{\sum_{j=0}^{k-1}(n_j + m)} y\right), \Xi \sum_{i=1}^{k} \lambda_i L_{n_i} x_i\right) + 2\delta
\end{aligned}$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

$$
\begin{aligned}
&\le d\Big(\Xi\sum_{i=1}^{k}\frac{n_i}{\hat n}L_{n_i}f^{\sum_{j=0}^{i-1}(n_j+m)}y,\ \Xi\sum_{i=1}^{k}\lambda_i L_{n_i}x_i\Big)\\
&\quad+\sum_{i=1}^{k-1}\frac{m}{\hat n}d\Big(\Xi L_m f^{\sum_{j=1}^{i}n_j+(i-1)m}y,\ 0\Big)+2\delta\\
&\le d\Big(\Xi\sum_{i=1}^{k}\frac{n_i}{\hat n}L_{n_i}f^{\sum_{j=0}^{i-1}(n_j+m)}y,\ \Xi\sum_{i=1}^{k}\lambda_i L_{n_i}x_i\Big)+\frac{m(k-1)}{\hat n}M+2\delta\\
&\le \sum_{i=1}^{k}\frac{n_i}{\hat n}d(\Xi L_{n_i}f^{\sum_{j=0}^{i-1}(n_j+m)}y,\ \Xi L_{n_i}x_i)\\
&\quad+\sum_{i=1}^{k}d\left(\Xi\frac{n_i}{\hat n}L_{n_i}x_i,\ \Xi\lambda_i L_{n_i}x_i\right)+\frac{m(k-1)}{\hat n}M+2\delta\\
&\le 3\delta+\frac{m(k-1)}{\hat n}M+\sum_{i=1}^{k}d\left(\Xi\frac{n_i}{\hat n}L_{n_i}x_i,\ \Xi\lambda_i L_{n_i}x_i\right)\\
&\le 3\delta+\frac{m(k-1)}{\hat n}M+\sum_{i=1}^{k}d(\Xi\lambda_i L_{n_i}x_i,\ \Xi\lambda_i L_{n_i}x_i)+\sum_{i=1}^{k}d\left(\left(\frac{n_i}{\hat n}-\lambda_i\right)\Xi L_{n_i}x_i,0\right)\\
&\le 3\delta+\frac{m(k-1)}{\hat n}M+\frac{m(k-1)}{\hat n}M\\
&= 3\delta+\frac{2m(k-1)}{\hat n}M.
\end{aligned}
$$

Since n_i can be chosen arbitrarily large we conclude that $y\in P(C,4\delta,\hat n)$. Since

$$
\begin{aligned}
\#C_{n_1}\times\cdots\times\#C_{n_k}&\ge\exp([\lambda_1 n](h_{\nu_1}(f)-\gamma)+\cdots+[\lambda_k n](h_{\nu_k}(f)-\gamma))\\
&\ge\exp(n(h_\nu(f)-2\gamma))\\
&\ge\exp(n(h_\mu(f)-3\gamma))
\end{aligned}
$$

and $d_{\hat n}(y,y')>2\epsilon$ for $y=(x_1,\dots,x_k)$, $y'=(x_1',\dots,x_k')$ corresponding to different $(x_1,\dots,x_k)$ and $(x_1',\dots,x_k')$, we have

$$
\liminf_{\hat n\to\infty}\frac{1}{\hat n}\log N(C,4\delta,\hat n,\epsilon)\ge\liminf_{n\to\infty}\frac{1}{\hat n}\log\#C_{n_1}\times\cdots\times\#C_{n_k}\ge h_\mu(f)-3\gamma.
$$

Due to the choice of ϵ and δ, we have $\Lambda(\Delta(C))+\gamma\ge h_\mu(f)-3\gamma$, which finishes the proof of $\Lambda(\Delta(C))\ge H(\Delta(C))$. □

Proof of $h_{\mathrm{top}}(\Delta(C))\ge\Lambda(\Delta(C))$. This is similar to [**10**]. Chose arbitrary $\gamma>0$ and let $s=\Lambda(\Delta(C))-\gamma$. The idea is to construct a Moran fractal $F\subseteq\Delta(C)$ such that $h_{\mathrm{top}}(F)\ge s$, then we can get

$$
h_{\mathrm{top}}(\Delta(C))\ge h_{\mathrm{top}}(F)\ge\Lambda(\Delta(C))-\gamma
$$

and finish the proof.

Let $\gamma>0$ be arbitrary. By the definition of $\Lambda(\Delta(C))$, there is $\epsilon>0$ small enough such that

$$
\lim_{\delta\to0}\liminf_{n\to\infty}\frac{1}{n}\log M(C,\delta,n,8\epsilon)\ge\Lambda(\Delta(C))-\frac{\gamma}{2}.
$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Since f satisfies the specification property, let $m = m(\epsilon)$ be as in the definition of the specification property and

$$m_k = m\left(\frac{\epsilon}{2^k}\right), \quad k \in \mathbb{N}.$$

Then we can choose a sequence $\delta_k \downarrow 0$ and a sequence $n_k \uparrow \infty$ such that

$$M_k = M(C, \delta_k, n_k, 8\epsilon) > \exp(n_k(\Lambda(\Delta(C)) - \gamma)), \quad n_k \geq 2^{m_k}.$$

Since $M(C, \delta_k, n_k, 8\epsilon)$ is the cardinality of a maximal $(n_k, 8\epsilon)$-separated set in $P(C, \delta_k, n_k)$. Let $C_k = \{x_i^k \mid i = 1, \ldots, M_k\}$ be one of the maximal $(n_k, 8\epsilon)$-separated sets.

Step 1. Construction of intermediate sets D_k.
We can choose some sequence of integers $\{N_k\}$ such that:
(1) $N_1 = 1$;
(2) $N_k \geq 2^{n_{k+1}+m_{k+1}}$, $k \geq 2$;
(3) $N_{k+1} \geq 2^{N_1 n_1 + \cdots + N_k(n_k+m_k)}$, $k \geq 1$.

Such a sequence grows very fast and satisfies

$$\lim_{k\to\infty} \frac{n_{k+1} + m_{k+1}}{N_k} = 0 \quad \text{and} \quad \lim_{k\to\infty} \frac{N_1 n_1 + \cdots + N_k(n_k + m_k)}{N_{k+1}} = 0.$$

For any N_k-tuple $(i_1, \ldots, i_{N_k}) \in \{1, \ldots, M_k\}^{N_k}$, let $y(i_1, \ldots, i_{N_k})$ be some point which shadows pieces of orbits $\{x_{i_j}^k, fx_{i_j}^k, \ldots, f^{n_k-1}x_{i_j}^k\}$, $j = 1, \ldots, N_k$, with a gap m_k, i.e.

$$d_{n_k}(x_{i_j}, f^{a_j^k} y(i_1, \ldots, i_{N_k})) < \frac{\epsilon}{2^k},$$

where $a_j^k = (n_k + m_k)(j-1)$, $j = 1, \ldots, N_k$. Such a point $y(i_1, \ldots, i_{N_k})$ exists because f satisfies the specification property. Collect all such points into the set

$$D_k = \{y(i_1, \ldots, i_{N_k}) \mid i_1, \ldots, i_{N_k} \in \{1, \ldots, M_k\}\}.$$

Let $t_k = N_k n_k + (N_k - 1)m_k$. One could easily verify that for $(i_1, \ldots, i_{N_k}) \neq (j_1, \ldots, j_{N_k})$,

$$d_{t_k}(y(i_1, \ldots, i_{N_k}), y(j_1, \ldots, j_{N_k})) > 6\epsilon.$$

Hence, $\#D_k = M_k^{N_k}$.

Step 2. Construction of L_k.
We construct L_k by induction.

Let $L_1 = D_1$ and put $l_1 = n_1$. Suppose L_k is defined. We come to construct L_{k+1}. Let

$$l_{k+1} = l_k + m_{k+1} + t_{k+1} = N_1 n_1 + N_2(n_2 + m_2) + \cdots + N_{k+1}(n_{k+1} + m_{k+1}).$$

For every $x \in L_k$ and $y \in D_{k+1}$, by definition of the specification property there is a point $z = z(x, y)$ such that

$$d_{l_k}(x, z) < \frac{\epsilon}{2^{k+1}}, \quad d_{t_{k+1}}(y, f^{l_k+m_{k+1}} z) < \frac{\epsilon}{2^{k+1}}.$$

Collect all such points into the set

$$L_{k+1} = \{z = z(x, y) \mid x \in L_k, y \in D_{k+1}\}.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

We can prove that for different pairs $(x, y), (x', y'), x, x' \in L_k, y, y' \in D_{k+1}$,

$$d_{l_{k+1}}(z(x, y), z'(x', y')) > 5\epsilon.$$

Indeed for such two points, if $y \neq y'$ then since $y, y' \in D_{k+1}$, we have $d_{t_{k+1}}(y, y') > 6\epsilon$. We know from the construction of L_{k+1} that

$$d_{t_{k+1}}(y, f^{l_k+m_{k+1}}z) < \frac{\epsilon}{2^{k+1}} \quad \text{and} \quad d_{t_{k+1}}(y', f^{l_k+m_{k+1}}z') < \frac{\epsilon}{2^{k+1}}.$$

Hence, we get

$$\begin{aligned} d_{l_{k+1}}(z, z') &\geq d_{t_{k+1}}(f^{l_k+m_{k+1}}z, f^{l_k+m_{k+1}}z') \\ &\geq d_{t_{k+1}}(y, y') - d_{t_{k+1}}(y, f^{l_k+m_{k+1}}z) - d_{t_{k+1}}(y', f^{l_k+m_{k+1}}z') \\ &\geq 6\epsilon - \frac{\epsilon}{2^k}. \end{aligned}$$

Otherwise, if $x \neq x'$, then since

$$d_{l_k}(x, z) < \frac{\epsilon}{2^{k+1}} \quad \text{and} \quad d_{l_k}(x', z') < \frac{\epsilon}{2^{k+1}},$$

we get

$$d_{l_{k+1}}(z, z') \geq d_{l_k}(z, z') \geq d_{l_k}(x, x') - \frac{\epsilon}{2^k}.$$

Then following the discussion we have

$$\begin{aligned} d_{l_{k+1}}(z, z') &\geq \min\left\{\min_{x\neq x', x, x' \in D_1} d_{n_1}(x, x'), 6\epsilon\right\} - \sum_{i=1}^{k} \frac{\epsilon}{2^i} \\ &\geq 6\epsilon - \epsilon \\ &\geq 5\epsilon. \end{aligned}$$

Step 3. The Moran fractal F.

For every k put

$$F_k = \bigcup_{x\in L_k} \overline{B_{l_k}}\left(x, \frac{\epsilon}{2^{k-1}}\right),$$

where $\overline{B_{l_k}}(x, \delta)$ is the closed ball around x of radius δ in the metric d_{l_k}, i.e.

$$\overline{B_{l_k}}(x, \delta) = \{y \in X \mid d_{l_k}(x, y) \leq \delta\}.$$

It is easy to show that for each k the following is satisfied:

(1) for any $x, x' \in L_k, x \neq x'$, the sets $\overline{B_{l_k}}(x, \epsilon/2^{k-1}), \overline{B_{l_k}}(x', \epsilon/2^{k-1})$ are disjoint;

(2) if $z \in L_{k+1}$ descends from $x \in L_k$, then

$$\overline{B_{l_{k+1}}}\left(z, \frac{\epsilon}{2^k}\right) \subseteq \overline{B_{l_k}}\left(x, \frac{\epsilon}{2^{k-1}}\right),$$

which means $F_{k+1} \subseteq F_k$.

Finally put

$$F = \bigcap_{k\geq 1} F_k.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Then F is a non-empty closed subset of X.

First we show that for every $x \in F$ one has

$$A(\Xi L_n x) \subseteq C.$$

Recall that $\Xi : M(X) \mapsto Y$ is affine and continuous, so we can extend the map to the set of finite measures on X. That is, given a finite measure μ on X,

$$\Xi\mu = \mu(X)\Xi\left(\frac{\mu}{\mu(X)}\right).$$

For each $n \in \mathbb{N}$ and $x \in X$, let $S_n x$ denote $\sum_{i=0}^{n-1} \delta_{f^i x}$. For arbitrary $y \in D_k$, by definition there is $(i_1, \ldots, i_{N_k}) \in \{i, \ldots, M_k\}^{N_k}$ such that $y = y(i_1, \ldots, i_{N_k})$ shadows pieces of orbits $\{x_{i_j}^k, f x_{i_j}^k, \ldots, f^{n_k-1} x_{i_j}^k\}$, $j = 1, \ldots, N_k$ with a gap m_k, i.e.

$$d_{n_k}(x_{i_j}, f^{a_j^k} y(i_1, \ldots, i_{N_k})) < \frac{\epsilon}{2^k},$$

where $a_j^k = (n_k + m_k)(j-1)$, $j = 1, \ldots, N_k$. Since $x_{i_j}^k \in C_k \subseteq P(C, \delta_k, n_k)$, there is a point $h_{i_j}^k \in C$ such that

$$d(\Xi S_{n_k} x_{i_j}^k, n_k h_{i_j}^k) \leq n_k \delta_k.$$

Hence, we have

$$\begin{aligned} d(\Xi S_{n_k} f^{a_j^k} y, n_k h_{i_j}^k) &\leq d(\Xi S_{n_k} f^{a_j^k} y, \Xi S_{n_k} x_{i_j}^k) + d(\Xi S_{n_k} x_{i_j}^k, n_k h_{i_j}^k) \\ &\leq n_k d(\Xi L_{n_k} f^{a_j^k} y, \Xi L_{n_k} x_{i_j}^k) + n_k \delta_k \\ &\leq n_k \left(V\left(\Xi, \frac{\epsilon}{2^k}\right) + \delta_k \right), \end{aligned}$$

where $V(\Xi, \epsilon/2^k)$ goes to zero as k goes to infinity by Observation 1.11.

Represent the interval $[0, t_k - 1]$ as follows:

$$[0, t_k - 1] = \bigcup_{j=0}^{N_k-1} [a_j^k, a_j^k + n_k - 1] \cup \bigcup_{j=0}^{N_k-2} [a_j^k + n_k, a_j^k + n_k + m_k - 1].$$

Hence,

$$\begin{aligned} &d\left(\Xi S_{t_k} y, \sum_{j=1}^{N_k} n_k h_{i_j}^k \right) \\ &\quad \leq d\left(\Xi S_{t_k} y, \sum_{j=1}^{N_k} \Xi S_{n_k} x_{i_j}^k \right) + d\left(\sum_{j=1}^{N_k} \Xi S_{n_k} x_{i_j}^k, \sum_{j=1}^{N_k} n_k h_{i_j}^k \right) \\ &\quad \leq d\left(\Xi S_{t_k} y, \sum_{j=1}^{N_k} \Xi S_{n_k} x_{i_j}^k \right) + N_k n_k \delta_k \\ &\quad \leq \sum_{j=1}^{N_k} d(\Xi S_{n_k} f^{a_j^k} y, \Xi S_{n_k} x_{i_j}^k) + \sum_{j=1}^{N_k-1} d(\Xi S_{m_k} f^{a_j^k + n_k} y, 0) + N_k n_k \delta_k \\ &\quad \leq N_k n_k V\left(\Xi, \frac{\epsilon}{2^k}\right) + \sum_{j=1}^{N_k-1} m_k d(\Xi L_{m_k} f^{a_j^k + n_k} y, 0) + N_k n_k \delta_k \end{aligned}$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

$$\leq N_k n_k V\left(\Xi, \frac{\epsilon}{2^k}\right) + (N_k - 1)m_k M + N_k n_k \delta_k$$
$$= N_k n_k \left(V\left(\Xi, \frac{\epsilon}{2^k}\right) + \delta_k\right) + (N_k - 1)m_k M,$$

where $M = \sup_{\mu \in M(X)}\{d(\Xi\mu, 0)\}$.

Since for each $x_{i_j}^k$, $k \in \mathbb{N}$, $j = 1, 2, \ldots, N_k$, there is a $h_{i_j}^k \in C$ such that

$$d(\Xi S_{n_k} x_{i_j}^k, n_k h_{i_j}^k) \leq n_k \delta_k.$$

Fix such points $h_{i_j}^k$ for each $x_{i_j}^k$. Hence, for each $z \in L_k$, there is a sequence of points

$$(h_{i_1}^1, \ldots, h_{i_{N_1}}^1, h_{i_1}^2, \ldots, h_{i_{N_2}}^2, \ldots, h_{i_1}^k, \ldots, h_{i_{N_k}}^k)$$

corresponding to it. Let

$$h(z) = \sum_{q=1}^{k} n_q \sum_{p=1}^{N_q} h_{i_p}^q$$

and

$$R_k = \max_{z \in L_k}\{d(\Xi S_{l_k} z, h(z))\}.$$

For each $y \in D_k$, write $h(y) = n_k \sum_{p=1}^{N_k} h_{i_p}^k$. We estimate R_k by induction.

When $k = 1$, $L_1 = D_1 = C_1 \subseteq P(C, \delta_1, n_1)$ and $l_1 = n_1$. We have

$$R_1 \leq l_1 \delta_1.$$

For $z \in L_{k+1}$, there is a unique $x \in L_k$, $y \in D_{k+1}$ such that

$$d_{l_k}(x, z) < \frac{\epsilon}{2^{k+1}}, \quad d_{t_{k+1}}(y, f^{l_k + m_{k+1}} z) < \frac{\epsilon}{2^{k+1}}.$$

Hence,

$$\begin{aligned} d(\Xi S_{l_{k+1}} z, h(z)) &\leq d(\Xi S_{l_k} z, h(x)) + d(\Xi S_{m_{k+1}} f^{l_k} x, 0) + d(\Xi S_{t_{k+1}} f^{l_k + m_{k+1}} z, h(y)) \\ &\leq d(\Xi S_{l_k} z, \Xi S_{l_k} x) + d(\Xi S_{l_k} x, h(x)) + d(\Xi S_{m_{k+1}} f^{l_k} x, 0) \\ &\quad + d(\Xi S_{t_{k+1}} f^{l_k + m_{k+1}} z, \Xi S_{t_{k+1}} y) + d(\Xi S_{t_{k+1}} y, h(y)) \\ &\leq l_k V\left(\Xi, \frac{\epsilon}{2^{k+1}}\right) + R_k + m_{k+1} M + t_{k+1} V\left(\Xi, \frac{\epsilon}{2^{k+1}}\right) \\ &\quad + N_{k+1} n_{k+1}\left(V\left(\Xi, \frac{\epsilon}{2^{k+1}}\right) + \delta_{k+1}\right) + (N_{k+1} - 1)m_{k+1} M \\ &= R_k + (l_k + N_{k+1} n_{k+1} + t_{k+1}) V\left(\Xi, \frac{\epsilon}{2^{k+1}}\right) + N_{k+1} m_{k+1} M \\ &\quad + N_{k+1} n_{k+1} \delta_{k+1}. \end{aligned}$$

Recall that $l_k = N_1 n_1 + \cdots + N_k(n_k + m_k)$ and $l_{k+1} = l_k + m_{k+1} + t_{k+1}$, we have

$$R_{k+1} \leq R_k + 2l_{k+1} V\left(\Xi, \frac{\epsilon}{2^{k+1}}\right) + l_{k+1}\delta_{k+1} + N_{k+1} m_{k+1} M.$$

By induction

$$R_k \leq 2\sum_{p=1}^{k} l_p \left(V\left(\Xi, \frac{\epsilon}{2^p}\right) + \delta_p + \frac{N_p m_p}{l_p} M\right).$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Write

$$c_p = V\left(\Xi, \frac{\epsilon}{2^p}\right) + \delta_p + \frac{N_p m_p}{l_p} M.$$

Since $V(\Xi, \epsilon/2^p) \to 0$, $\delta_k \downarrow 0$ as k goes to infinity, $l_k \geq N_k(n_k + m_k)$, the sequence $\{n_k\}$ is such that $n_k \to \infty$ as $k \to \infty$, and $n_k \geq 2^{m_k}$, we conclude that $(N_p m_p / l_p)M \to 0$ as k increases to ∞. Hence,

$$R_k \leq 2 \sum_{p=1}^{k} l_p c_p,$$

where $c_p \to 0$ as $p \to \infty$. By the choice of N_k, we have $l_k \geq 2^{l_{k-1}}$, hence for sufficiently large k one has

$$\frac{R_k}{l_k} \leq 2c_k + 2\frac{1}{k} \sum_{p=1}^{k-1} c_p \to 0 \quad \text{as } k \to \infty.$$

Given $x \in F$, $n \in \mathbb{N}$, and $n > l_1$. There exists a unique $k \geq 1$ such that

$$l_k < n \leq l_{k+1}.$$

Also there is a unique j, $0 \leq j \leq N_{k+1} - 1$, such that

$$l_k + j(n_{k+1} + m_{k+1}) < n \leq l_k + (j+1)(n_{k+1} + m_{k+1}).$$

Since $x \in F \subseteq F_{k+1}$, there exists $z \in L_{k+1}$ such that

$$d_{l_{k+1}}(x, z) < \frac{\epsilon}{2^k}.$$

Then for such a $z \in L_{k+1}$, there exists $\hat{x} \in L_k$ and $y \in D_{k+1}$ such that

$$d_{l_k}(\hat{x}, z) < \frac{\epsilon}{2^{k+1}}, \quad d_{t_{k+1}}(y, f^{l_k + m_{k+1}} z) < \frac{\epsilon}{2^{k+1}}.$$

Therefore,

$$d_{l_k}(\hat{x}, x) < \frac{\epsilon}{2^{k-1}}, \quad d_{t_{k+1}}(y, f^{l_k + m_{k+1}} x) < \frac{\epsilon}{2^{k-1}}.$$

Moreover, if $j > 0$, then by the definition of D_{k+1} there exist points $x_{i_1}^{k+1}, \ldots, x_{i_j}^{k+1} \in C_{k+1}$ such that

$$d_{n_{k+1}}(x_{i_t}^{k+1}, f^{a_t^{k+1}} y) < \frac{\epsilon}{2^{k+1}},$$

where $a_t^{k+1} = (n_{k+1} + m_{k+1})(t-1)$, $t = 1, 2, \ldots, j$, and hence

$$d_{n_{k+1}}(x_{i_t}^{k+1}, f^{l_k + m_{k+1} + a_t^{k+1}} x) < \frac{\epsilon}{2^{k-2}}.$$

Represent $[0, n-1]$ as

$$[0, n-1] = [0, l_k - 1] \cup \bigcup_{t=1}^{j} [l_k + (t-1)(m_{k+1} + n_{k+1}), l_k + t(m_{k+1} + n_{k+1}) - 1]$$
$$\cup [l_k + j(m_{k+1} + n_{k+1}), n-1].$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

We have

$$\begin{aligned} d(\Xi S_{l_k}x, h(\hat{x})) &\le d(\Xi S_{l_k}x, \Xi S_{l_k}\hat{x}) + d(\Xi S_{l_k}\hat{x}, h(\hat{x})) \\ &\le l_k V\left(\Xi, \frac{\epsilon}{2^{k-1}}\right) + R_k. \end{aligned}$$

On each of the intervals $[a_t^{k+1}, a_t^{k+1} + (m_{k+1} + n_{k+1}) - 1]$, where $a_t^{k+1} = l_k + (t-1)(m_{k+1} + n_{k+1})$

$$\begin{aligned} &d(\Xi S_{m_{k+1}+n_{k+1}} f^{a_t^{k+1}} x, n_{k+1} h_{i_t}^{k+1}) \\ &\quad \le d(\Xi S_{m_{k+1}} f^{a_t^{k+1}+n_{k+1}} x, 0) + d(\Xi S_{n_{k+1}} f^{a_t^{k+1}} x, \Xi S_{n_{k+1}} x_{i_t}^{k+1}) \\ &\qquad + d(\Xi S_{n_{k+1}} x_{i_t}^{k+1}, n_{k+1} h_{i_t}^{k+1}) \\ &\quad \le M m_{k+1} + n_{k+1} V\left(\Xi, \frac{\epsilon}{2^{k-2}}\right) + n_{k+1}\delta_{k+1}. \end{aligned}$$

Then we have the following estimate:

$$\begin{aligned} &d(\Xi S_n x, h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}) \\ &\quad \le d(\Xi S_{n-l_k-j(m_{k+1}+n_{k+1})} f^{l_k+j(m_{k+1}+n_{k+1})} x, 0) + d(\Xi S_{l_k} x, h(\hat{x})) \\ &\qquad + \sum_{t=1}^{j} d(\Xi S_{m_{k+1}+n_{k+1}} f^{a_t^{k+1}} x, n_{k+1} h_{i_t}^{k+1}) \\ &\quad \le (n - l_k - j(m_{k+1} + n_{k+1}))M + d(\Xi S_{l_k} x, h(\hat{x})) \\ &\qquad + j\left(m_{k+1} M + n_{k+1} V\left(\Xi, \frac{\epsilon}{2^{k-2}}\right) + n_{k+1}\delta_{k+1}\right) \\ &\quad \le (n - l_k - j(m_{k+1} + n_{k+1}))M + l_k V\left(\Xi, \frac{\epsilon}{2^{k-1}}\right) + R_k \\ &\qquad + j\left(m_{k+1} M + n_{k+1} V\left(\Xi, \frac{\epsilon}{2^{k-2}}\right) + n_{k+1}\delta_{k+1}\right) \\ &\quad \le R_k + (l_k + j n_{k+1}) V\left(\Xi, \frac{\epsilon}{2^{k-2}}\right) + j n_{k+1}\delta_{k+1} + (n_{k+1} + (j+1)m_{k+1})M. \end{aligned}$$

Hence we have

$$\begin{aligned} &d\left(\Xi L_n x, \frac{1}{n}\left(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\right)\right) \\ &\quad \le \frac{R_k}{n} + \frac{l_k + j n_{k+1}}{n} V\left(\Xi, \frac{\epsilon}{2^{k-2}}\right) + \frac{n_{k+1} j}{n}\delta_{k+1} + \left(\frac{n_{k+1} + m_{k+1}}{N_k} + \frac{m_{k+1}}{n_{k+1}}\right) M \end{aligned}$$

goes to zero as n increases to infinity. Let

$$N(x) = N_1 n_1 + \cdots + N_k n_k + j n_{k+1}.$$

Since C is a convex and closed subset of Y, we have that

$$\frac{1}{N(x)}\left(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\right) \in C.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

We estimate that

$$\begin{aligned}
&d\Big(\Xi L_n x, \frac{1}{N(x)}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big)\Big) \\
&\quad \le d\Big(\Xi L_n x, \frac{1}{n}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big)\Big) \\
&\qquad + d\Big(\frac{1}{n}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big), \frac{1}{N(x)}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big)\Big) \\
&\quad \le d\Big(\Xi L_n x, \frac{1}{n}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big)\Big) \\
&\qquad + d\Big(\Big(\frac{1}{n} - \frac{1}{N(x)}\Big)\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big), 0\Big) \\
&\qquad + d\Big(\frac{1}{N(x)}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big), \frac{1}{N(x)}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big)\Big) \\
&\quad = d\Big(\Xi L_n x, \frac{1}{n}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big)\Big) \\
&\qquad + d\Big(\Big(\frac{1}{n} - \frac{1}{N(x)}\Big)\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big), 0\Big) \\
&\quad \le d\Big(\Xi L_n x, \frac{1}{n}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big)\Big) + \frac{N_1 m_1 + \cdots + N_k m_k + j m_{k+1}}{n} M.
\end{aligned}$$

So we have

$$\lim_{n\to\infty} \operatorname{dist}(\Xi L_n x, C) \le \lim_{n\to\infty} d\Big(\Xi L_n x, \frac{1}{N(x)}\Big(h(\hat{x}) + \sum_{t=1}^{j} n_{k+1} h_{i_t}^{k+1}\Big)\Big) = 0.$$

Then similar to the discussion in [**10**], we can construct a measure μ on F which satisfies the request of the entropy distribution principle with $s = \Lambda(\Delta(C)) - \gamma$ and then we get the result

$$h_{\text{top}}(\Delta(C)) \ge h_{\text{top}}(F) \ge \Lambda(\Delta(C)) - \gamma.$$

The conclusion $h_{\text{top}}(\Delta(C)) \ge \Lambda(\Delta(C))$ follows since γ is arbitrary.

We now prove the second conclusion of Theorem 2.1. By Lemma 2.2, we have that $\Lambda\Delta : C(Y) \mapsto \mathbb{R}$ is semi-upper continuous. We continue to show the other conclusion. For any $C_1, C_2 \in C(Y)$ such that $\Delta(C_i) \ne \varnothing$ and $s_1, s_2 \ge 0$ with $s_1 + s_2 = 1$, here $sC = \{sx \mid x \in C\}$ and $C_1 + C_2 = \{x_1 + x_2 \mid x_i \in C_i\}$ for $C, C_1, C_2 \in C(Y)$ and $s \in \mathbb{R}$. By Part 2 of the proof of Theorem 2.1,

$$\Lambda(\Delta(s_1 C_1 + s_2 C_2)) \ge \sup_{\mu \in M(f,X), \Xi\mu \in s_1 C_1 + s_2 C_2} h_\mu(f).$$

Take any $\mu_1, \mu_2 \in M(f, X)$ such that $\Xi\mu_1 \in C_1$ and $\Xi\mu_2 \in C_2$, then $\Xi(s_1\mu_1 + s_2\mu_2) \in s_1 C_1 + s_2 C_2$. Hence, we have

$$\Lambda(\Delta(s_1 C_1 + s_2 C_2)) \ge h_{s_1\mu_1 + s_2\mu_2}(f) = s_1 h_{\mu_1}(f) + s_2 h_{\mu_2}(f).$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Since μ_1, μ_2 are arbitrary, we get

$$\begin{aligned}\Lambda(\Delta(s_1C_1 + s_2C_2)) &\geq s_1 \sup_{\mu\in M(f,X),\,\Xi\mu\in C_1} h_\mu(f) + s_2 \sup_{\mu\in M(f,X),\,\Xi\mu\in C_2} h_\mu(f)\\ &= s_1\Lambda(\Delta(C_1)) + s_2\Lambda(\Delta(C_2)).\end{aligned}$$

□

COROLLARY 2.3. *Suppose that $f : X \mapsto X$ is a continuous transformation with specification property on a compact metric space X. Let $\varphi : X \mapsto \mathbb{R}$ be continuous. Then for any closed subinterval C of $\mathbb{R}$,*

$$h_{\text{top}}\left\{x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}\left(\frac{1}{n}\sum_{i=0}^{n-1}\varphi(f^i x), C\right) = 0\right\} = \sup_{\mu\in M(f,X),\,\Xi\mu\in C} h_\mu(f).$$

COROLLARY 2.4. *Suppose that $f : X \mapsto X$ is a continuous transformation with the specification property on a compact metric space X. Then for any closed subset C of $M(X)$,*

$$h_{\text{top}}\left\{x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}\left(\frac{1}{n}\sum_{i=0}^{n-1}\delta_{f^i x}, C\right) = 0\right\} = \sup_{\mu\in M(f,X),\,\mu\in C} h_\mu(f).$$

As an application of Theorem 2.1, we study the mixed multifractal spectra. Mixed multifractal spectra combine different local characteristics which depend on various different aspects of the underlying dynamical system. Let $(Y_i, \Xi_i)_i$ be a (possibly uncountable) family of deformations. We define the mixed multifractal spectrum of the family $(Y_i, \Xi_i)_i$ by

$$h_{\text{top}}(\Delta(C)) = h_{\text{top}}\left\{x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}((\Xi_i L_n x)_i, C) = 0\right\}.$$

The mixed multifractal spectrum of the family $(Y_i, \Xi_i)_i$ combine information provided by the different deformations $(Y_i, \Xi_i)_i$ and study how they depend on each other.

THEOREM 2.5. (Abstract mixed multifractal spectra) *Suppose that $f : X \mapsto X$ is a continuous transformation with specification property on a compact metric space X. Let $(Y_i, \Xi_i)_i$ be a (possibly uncountable) family of deformations and assume that Y_i is a normed vector space and that $\Xi_i : M(X) \mapsto Y_i$ is affine and continuous. Define the vector spaces $\mathbf{x}_i Y_i$ and $[\mathbf{x}_i Y_i]^\infty$ by*

$$\begin{aligned}\mathbf{x}_i Y_i &= \{(y_i)_i \mid y_i \in Y_i \text{ for all } i\},\\ [\mathbf{x}_i Y_i]^\infty &= \left\{(y_i)_i \in \mathbf{x}_i Y_i \;\middle|\; \sup_i \|y_i\| < \infty\right\},\end{aligned}$$

and equip $[\mathbf{x}_i Y_i]^\infty$ with the norm $\|(y_i)_i\| = \sup_i \|y_i\|$. Assume that $\sup_{\mu\in M(X),i} \|\Xi_i\mu\| < \infty$ and that the map

$$M(X) \mapsto [\mathbf{x}_i Y_i]^\infty : \mu \to (\Xi_i\mu)_i$$

is continuous. If $C \subseteq [\mathbf{x}_i Y_i]^\infty$ is closed and convex, then

$$h_{\text{top}}\left\{x \in X \;\middle|\; \lim_{n\to\infty} \operatorname{dist}((\Xi_i L_n x)_i, C) = 0\right\} = \sup_{\mu\in M(f,X),\,(\Xi_i\mu)_i\in C} h_\mu(f).$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

3. *Topological entropy of the set of divergence points*

In this section we discuss the topological entropy of the set of divergence points. We simplify $\Delta(\{y\})$ as $\Delta(y)$.

THEOREM 3.1. *Suppose that $f : X \mapsto X$ is a continuous transformation with the specification property. Y is a vector space with linearly compatible metric. $\Xi : M(X) \mapsto Y$ is affine and continuous. Let $L_n : X \mapsto M(X)$ be defined as before. Then either:*

(1) *there exists $y_0 \in Y$ such that for all $x \in X$, $A(\Xi L_n x) = \{y_0\}$; or*

(2) *the set $D(f, \Xi) = \{x \in X \mid \lim_{n\to\infty} \Xi L_n x$ does not exist$\}$ is of the same topological entropy as the whole space X.*

LEMMA 3.2. *For any $\delta > 0$, there exists $N = N(\delta)$ such that for any $k \geq 1$, $\epsilon > 0$ small enough, if $n \geq N$, then*

$$(N(y, \delta, n, 4\epsilon))^k \leq N(y, 4\delta, nk + (k-1)m(\epsilon), \epsilon),$$

where $m(\epsilon)$ is taken from the definition of the specification property.

Proof. It is a standard conclusion that

$$N(y, \delta, n, \epsilon) \leq M(y, \delta, n, \epsilon) \leq N(y, \delta, n, \tfrac{1}{2}\epsilon).$$

Hence, $N(y, \delta, n, 4\epsilon) \leq M(y, \delta, n, 4\epsilon)$. Let S be some $(n, 4\epsilon)$-separated set of $P(y, \delta, n)$ with maximal cardinality. Take arbitrary k points $x_1, \ldots, x_k \in S$. By the specification property, there is some $x = x(x_1, \ldots, x_k)$ which shadows pieces of orbits $\{x_i, \ldots, f^{n-1}x_i\}, i = 1, 2, \ldots, k$, with a gap $m(\epsilon)$. We claim that different choices of $(x_1, \ldots, x_k)$ produce different $x = x(x_1, \ldots, x_k)$. Indeed suppose $(x_1, \ldots, x_k) \neq (x'_1, \ldots, x'_k)$ then there exists $1 \leq t \leq k$, $x_t \neq x'_t$. Simplify $nk + (k-1)m(\epsilon) = \hat{n}$ and $m = m(\epsilon)$. We know from the construction of $x(x_1, \ldots, x_k), x'(x'_1, \ldots, x'_k)$ that

$$d_n(f^{(t-1)(n+m)}x(x_1, \ldots, x_k), x_t) < \epsilon; \quad d_n(f^{(t-1)(n+m)}x'(x'_1, \ldots, x'_k), x'_t) < \epsilon.$$

Since $x_t \neq x'_t$ and $x_t, x'_t \in M(y, \delta, n, 4\epsilon)$, we have $d_n(x_t, x'_t) > 4\epsilon$. Hence,

$$\begin{aligned} &d_{\hat{n}}(x(x_1, \ldots, x_k), x'(x'_1, \ldots, x'_k)) \\ &\quad \geq d_n(f^{(t-1)(n+m)}x(x_1, \ldots, x_k), f^{(t-1)(n+m)}x'(x'_1, \ldots, x'_k)) \\ &\quad \geq d_n(x_t, x'_t) - d_n(f^{(t-1)(n+m)}x(x_1, \ldots, x_k), x_t) \\ &\qquad - d_n(f^{(t-1)(n+m)}x'(x'_1, \ldots, x'_k), x'_t) \\ &\quad > 2\epsilon. \end{aligned}$$

Rewrite $\Xi L_{\hat{n}} x$ as follows:

$$\begin{aligned} \Xi L_{\hat{n}} x &= \frac{n}{\hat{n}}\Xi L_n x + \frac{m}{\hat{n}}\Xi L_m f^n x + \frac{n}{\hat{n}}\Xi L_n f^{n+m} x + \cdots + \frac{n}{\hat{n}}\Xi L_n f^{(k-1)(n+m)} x \\ &= \sum_{i=1}^{k} \frac{n}{\hat{n}}\Xi L_n f^{(i-1)(n+m)} x + \sum_{i=1}^{k-1} \frac{m}{\hat{n}}\Xi L_m f^{in+(i-1)m} x. \end{aligned}$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Then we have

$$
\begin{aligned}
&d(\Xi L_{\hat{n}}x, y)\\
&\le d\Big(\sum_{i=1}^{k}\frac{n}{\hat{n}}\Xi L_n f^{(i-1)(n+m)}x, y\Big) + d\Big(\sum_{i=1}^{k-1}\frac{m}{\hat{n}}\Xi L_m f^{in+(i-1)m}x, 0\Big)\\
&\le \sum_{i=1}^{k} d\left(\frac{n}{\hat{n}}\Xi L_n f^{(i-1)(n+m)}x, \frac{1}{k}y\right) + \sum_{i=1}^{k-1}\frac{m}{\hat{n}}d(\Xi L_m f^{in+(i-1)m}x, 0)\\
&\le \sum_{i=1}^{k}\frac{1}{k}d(\Xi L_n f^{(i-1)(n+m)}x, y) + \sum_{i=1}^{k}\left(\frac{1}{k}-\frac{n}{\hat{n}}\right)d(\Xi L_n f^{(i-1)(n+m)}x, 0) + \frac{(k-1)m}{\hat{n}}M\\
&\le \sum_{i=1}^{k}\frac{1}{k}(d(\Xi L_n f^{(i-1)(n+m)}x, \Xi L_n x_i) + d(\Xi L_n x_i, y)) + \frac{(k-1)m}{\hat{n}}M + \frac{(k-1)m}{\hat{n}}M\\
&\le \frac{2(k-1)m}{\hat{n}}M + \delta + \sum_{i=1}^{k}\frac{1}{k}d(\Xi L_n f^{(i-1)(n+m)}x, \Xi L_n x_i)\\
&\le \frac{2(k-1)m}{\hat{n}}M + \delta + V(\Xi, \epsilon).
\end{aligned}
$$

Let $\epsilon < \epsilon(\delta)$, where $\epsilon(\delta)$ is a constant such that $d(\mu, \nu) < \epsilon(\delta)$ implies $d(\Xi\mu, \Xi\nu) < \delta/2$, then there exists $N = N(\delta)$ such that for any $k \ge 1$, $\epsilon < \epsilon(\delta)$ small enough, and $n \ge N$,

$$(N(y, \delta, n, 4\epsilon))^k \le M(y, 4\delta, nk+(k-1)m(\epsilon), 2\epsilon) \le N(y, 4\delta, nk+(k-1)m(\epsilon), \epsilon). \ \square$$

LEMMA 3.3. *If f satisfies the specification property, then*

$$\Lambda(\Delta(y)) = \lim_{\epsilon\to 0}\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{n}\log N(y, \delta, n, \epsilon) = \lim_{\epsilon\to 0}\lim_{\delta\to 0}\liminf_{n\to\infty}\frac{1}{n}\log N(y, \delta, n, \epsilon).$$

Proof. Without loss of generality, we can assume $\Lambda(\Delta(y)) < +\infty$. Since

$$\Lambda(\Delta(y)) = \lim_{\epsilon\to 0}\lim_{\delta\to 0}\liminf_{n\to\infty}\frac{1}{n}\log N(y, \delta, n, \epsilon),$$

for any $\gamma > 0$, there exist $\epsilon_0 > 0$, $\delta_0 > 0$ such that for all $\epsilon < \epsilon_0$, $\delta < \delta_0$,

$$\liminf_{n\to\infty}\frac{1}{n}\log N(y, \delta, n, 4\epsilon) < \Lambda(\Delta(y)) + \frac{1}{2}\gamma.$$

Without loss of generality, we assume that $4\epsilon < \epsilon(\delta)$, where $\epsilon(\delta)$ denotes a constant such that $d(\mu, \nu) < \epsilon(\delta)$ implies $d(\Xi\mu, \Xi\nu) < \delta/2$. To show the equality in Lemma 3.3, it suffices to show that

$$\limsup_{n\to\infty}\frac{1}{n}\log N(y, 4\delta, n, \epsilon) < \Lambda(\Delta(y)) + 2\gamma.$$

Since

$$\liminf_{n\to\infty}\frac{1}{n}\log N(y, \delta, n, 4\epsilon) < \Lambda(\Delta(y)) + \frac{1}{2}\gamma,$$

we can take a sequence of integers $n_k \uparrow +\infty$ such that

$$N(y, \delta, n_k, 4\epsilon) < \exp(n_k(\Lambda(\Delta(y)) + \gamma)).$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

For each k, $n_k = (n+m)p_k - m + l_k$, $0 \le l_k < n+m$. Hence,

$$\begin{aligned} N(y,\delta,n,4\epsilon)^{p_k} &\le N(y,4\delta,(n+m)p_k - m,\epsilon) \\ &\le N(y,4\delta,n_k,\epsilon) \\ &< \exp(n_k(\Lambda(\Delta(y)) + \gamma)). \end{aligned}$$

That is, for n_k large enough,

$$N(y,\delta,n,4\epsilon) \le \exp\left(\frac{n_k}{p_k}(\Lambda(\Delta(y)) + \gamma)\right) \le \exp((n+m)(\Lambda(\Delta(y)) + 2\gamma)).$$

So we get

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log N(y,\delta,n,4\epsilon) &\le \Lambda(\Delta(y)) + 2\gamma + \lim_{n\to\infty} \frac{m}{n}(\Lambda(\Delta(y)) + 2\gamma) \\ &\le \Lambda(\Delta(y)) + 2\gamma. \end{aligned}$$

□

LEMMA 3.4. *Let* $\tilde{\Delta}(y) = \{x \in X \mid y \in A(\Xi L_n x)\}$. *Suppose* $\Delta(y) \neq \varnothing$. *Then* $h_{\text{top}}(\tilde{\Delta}(y)) = h_{\text{top}}(\Delta(y))$.

Proof. Obviously $\Delta(y) \subseteq \tilde{\Delta}(y)$. Hence, $h_{\text{top}}(\Delta(y)) \le h_{\text{top}}(\tilde{\Delta}(y))$. To prove the other inequality, it suffices to show that for arbitrary $t > \Lambda(\Delta(y))$,

$$h_{\text{top}}(\tilde{\Delta}(y)) \le t.$$

Put $t_1 = \frac{1}{2}(t - \Lambda(\Delta(y)))$. Recall that by Lemma 3.3,

$$\Lambda(\Delta(y)) = \lim_{\epsilon\to 0} \lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log N(y,\delta,n,\epsilon) = \lim_{\epsilon\to 0} \lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N(y,\delta,n,\epsilon),$$

where $N(y,\delta,n,\epsilon)$ is the minimal number of (n,ϵ)-balls to cover

$$P(y,\delta,n) = \{x \in X \mid \Xi L_n x \in B(y,\delta)\}.$$

Hence, there exist $\epsilon_0 > 0$ $\delta_0 > 0$ such that for all $\epsilon < \epsilon_0$, $\delta < \delta_0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log N(y,\delta,n,\epsilon) < \Lambda(\Delta(y)) + \frac{t_1}{2}.$$

This implies that there is $N_0 \in \mathbb{N}$ such that $N(y,\delta,n,\epsilon) < \exp(n(t - t_1))$ for $n \ge N_0$. Since for all $n \ge N_0$,

$$\begin{aligned} \tilde{\Delta}(y) &\subseteq \bigcap_{k=0}^{+\infty} \bigcup_{n'\ge k} \{x \in X \mid \Xi L_{n'} x \in B(y,\delta)\} \\ &\subseteq \bigcap_{k=n}^{+\infty} \bigcup_{n'\ge k} \{x \in X \mid \Xi L_{n'} x \in B(y,\delta)\} \end{aligned}$$

we have

$$\begin{aligned} m(\tilde{\Delta}(y),t,n,\epsilon) &= \inf_{\Gamma, n(\Gamma)\ge n} \sum_i \exp(-tn_i) \\ &\le \sum_{k=n}^{+\infty} N(y,\delta,k,\epsilon)\exp(-tk) \\ &\le \sum_{k=n}^{+\infty} \exp(-t_1 k) < +\infty. \end{aligned}$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Hence, we have $h_{\mathrm{top}}(f, \tilde{\Delta}(y), \epsilon) < t$. The conclusion $h_{\mathrm{top}}(\tilde{\Delta}(y)) \le t$ follows by the definition of topological entropy. Due to the arbitrary choice of t, we get $h_{\mathrm{top}}(\tilde{\Delta}(y)) \le \Lambda(\Delta(y))$. We are done. □

LEMMA 3.5. *Given $y \in Y$, $\delta > 0$, $k \in \mathbb{N}$. Let $\epsilon(\delta)$ be a constant such that*

$$d(\mu, \nu) \le \epsilon(\delta) \Longrightarrow d(\Xi\mu, \Xi\nu) \le \frac{\delta}{2}.$$

Then for any $\epsilon' < \epsilon(\delta)$, let $m(\epsilon')$ be as in the definition of specification property and $a > 8m(\epsilon')M/\delta$, there exists $N \in \mathbb{N}$ such that for all $n \ge N$, let $x_1, \ldots, x_n \in X$ satisfying $\Xi L_{n_j} x_j \in B(y, \delta)$ with $n_j \ge a$, then for arbitrary $x_0 \in X$, if $z \in X$ such that z ϵ' shadows pieces of orbits $\{x_0, fx_0, \ldots, f^{b-1}x_0\}$, $(b \le k)$, $\{x_j, \ldots, f^{n_j-1}x_j\}$, $j = 1, 2, \ldots, n$, with a gap $m(\epsilon')$, we have

$$\Xi L_{b+nm(\epsilon')+\sum_{i=1}^{n} n_i} z \in B(y, 2\delta).$$

Proof. Simplify $\tilde{n} = b + nm(\epsilon') + \sum_{i=1}^{n} n_i$, $n' = \sum_{i=1}^{n} n_i$, and $m(\epsilon') = m$. We have

$$\begin{aligned} L_{\tilde{n}} z &= \frac{b}{\tilde{n}} L_b z + \frac{m}{\tilde{n}} L_m f^b z + \frac{n_1}{\tilde{n}} L_{n_1} f^{b+m} z + \cdots + \frac{n_n}{\tilde{n}} L_{n_n} f^{b+nm+\sum_{j=1}^{n-1} n_j} z \\ &= \frac{b}{\tilde{n}} L_b z + \sum_{i=1}^{n} \frac{m}{\tilde{n}} L_m f^{b+\sum_{j=1}^{i-1}(n_j+m)} z + \sum_{i=1}^{n} \frac{n_i}{\tilde{n}} L_{n_i} f^{b+im+\sum_{j=1}^{i-1} n_j} z. \end{aligned}$$

Hence, by the property of linearly compatible metric,

$$\begin{aligned} d(\Xi L_{\tilde{n}} z, y) &\le \frac{b}{\tilde{n}} d(\Xi L_b z, 0) + d\left(\sum_{i=1}^{n} \frac{m}{\tilde{n}} L_m f^{b+\sum_{j=1}^{i-1}(n_j+m)} z, 0 \right) \\ &\quad + d\left(\sum_{i=1}^{n} \frac{n_i}{\tilde{n}} \Xi L_{n_i} f^{b+im+\sum_{j=1}^{i-1} n_j} z, y \right) \\ &\le \frac{b+nm}{\tilde{n}} M + d\left(\sum_{i=1}^{n} \frac{n_i}{\tilde{n}} \Xi L_{n_i} f^{b+im+\sum_{j=1}^{i-1} n_j} z, \sum_{i=1}^{n} \frac{n_i}{n'} \Xi L_{n_i} x_i \right) \\ &\quad + d\left(\sum_{i=1}^{n} \frac{n_i}{n'} \Xi L_{n_i} x_i, \sum_{i=1}^{n} \frac{n_i}{n'} y \right) \\ &\le \frac{b+nm}{\tilde{n}} M + d\left(\sum_{i=1}^{n} \frac{n_i}{\tilde{n}} \Xi L_{n_i} f^{b+im+\sum_{j=1}^{i-1} n_j} z, \sum_{i=1}^{n} \frac{n_i}{n'} \Xi L_{n_i} x_i \right) \\ &\quad + \sum_{i=1}^{n} \frac{n_i}{n'} d(\Xi L_{n_i} x_i, y) \\ &\le \frac{b+nm}{\tilde{n}} M + \delta + \sum_{i=1}^{n} d\left(\frac{n_i}{\tilde{n}} \Xi L_{n_i} f^{b+im+\sum_{j=1}^{i-1} n_j} z, \frac{n_i}{n'} \Xi L_{n_i} x_i \right) \\ &\le \frac{b+nm}{\tilde{n}} M + \delta + \sum_{i=1}^{n} \frac{n_i}{n'} d\left(\frac{n'}{\tilde{n}} \Xi L_{n_i} f^{b+im+\sum_{j=1}^{i-1} n_j} z, \Xi L_{n_i} x_i \right) \\ &\le \frac{b+nm}{\tilde{n}} M + \delta + \sum_{i=1}^{n} \frac{n_i}{n'} d(\Xi L_{n_i} f^{b+im+\sum_{j=1}^{i-1} n_j} z, \Xi L_{n_i} x_i) \\ &\quad + \sum_{i=1}^{n} \frac{n_i}{n'} d\left(\left(1 - \frac{n'}{\tilde{n}}\right) \Xi L_{n_i} f^{b+im+\sum_{j=1}^{i-1} n_j} z, 0 \right) \end{aligned}$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

$$\leq \frac{b+nm}{\tilde{n}}M + \delta + \frac{\delta}{2} + \left(1 - \frac{n'}{\tilde{n}}\right)M$$
$$= \frac{3}{2}\delta + \frac{2(b+nm)}{\tilde{n}}M$$
$$\leq \frac{3}{2}\delta + \frac{2bM}{na} + \frac{2m}{a}M.$$

Hence, there is $N \in \mathbb{N}$, large enough such that for all $n \geq N$, $\Xi L_{\tilde{n}} z \in B(y, 2\delta)$. □

LEMMA 3.6. *Let different y, y' such that $\Delta(y)$, $\Delta(y') \neq \varnothing$ be given. Let $\tilde{\Delta}(\{y, y'\}) = \tilde{\Delta}(y) \cap \tilde{\Delta}(y')$. We have*

$$h_{\text{top}}(\tilde{\Delta}(\{y, y'\})) = \min\{\Lambda(\Delta(y)), \Lambda(\Delta(y'))\}.$$

Proof. Obviously $h_{\text{top}}(\tilde{\Delta}(\{y, y'\})) \leq \min\{\Lambda(\Delta(y)), \Lambda(\Delta(y'))\}$. On the other hand, for any $s < \min\{\Lambda(\Delta(y)), \Lambda(\Delta(y'))\}$, let $\gamma = \min\{\Lambda(\Delta(y)), \Lambda(\Delta(y'))\} - s$. Recall that

$$\Lambda(\Delta(y)) = \lim_{\epsilon\to 0} \lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log N(y, \delta, n, \epsilon),$$

where $N(y, \delta, n, \epsilon)$ is the minimal number of (n, ϵ)-balls to cover $P(y, \delta, n) = \{x \in X \mid \Xi L_n x \in B(y, \delta)\}$. Choose a sufficiently small $\epsilon > 0$ such that

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log M(y, \delta, n, 8\epsilon) > s + \frac{2}{3}\gamma; \quad \lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log M(y', \delta, n, 8\epsilon) > s + \frac{2}{3}\gamma.$$

Hence, we can choose a decreasing sequence δ_k such that

$$\liminf_{n\to\infty} \frac{1}{n} \log M(y, \delta_{2k-1}, n, 8\epsilon) > s + \frac{1}{3}\gamma; \quad \liminf_{n\to\infty} \frac{1}{n} \log M(y', \delta_{2k-1}, n, 8\epsilon) > s + \frac{1}{3}\gamma.$$

Let $\{y_k\} = \{y, y', y, y', \dots\}$. For each δ_k there is a $r_k = r(\delta_k)$ such that for all $x, y \in X$ with $d_n(x, y) < r_k$, $d(\Xi L_n x, \Xi L_n y) < \delta_k/2$. From such a sequence r_k, we can construct $\epsilon_k \downarrow 0$ such that

$$\epsilon_1 < \min\{\epsilon, r_1\}; \quad \epsilon_k < r_k; \quad \epsilon_{k+1} < \tfrac{1}{2}\epsilon_k.$$

By the definition of specification property, for each ϵ_k and arbitrary pieces of orbits, there exists a point that belongs to X which can shadow the pieces of orbits with a gap $m_k = m(\epsilon_k)$. Hence we can choose a sequence $n_k \uparrow +\infty$ such that

$$M_k : M(y_k, \delta_k, n_k, 8\epsilon) > \exp(n_k s), \quad n_k > \max\left\{2^{m_k}, \frac{8m_k M}{\delta_k}\right\}.$$

By definition, M_k is the cardinality of a maximal $(n_k, 8\epsilon)$-separated set in $P(y_k, \delta_k, n_k)$. Denote by $C_k = \{x_i^k \mid i = 1, 2, \dots, M_k\}$ one of these maximal $(n_k, 8\epsilon)$-separated sets.

To begin the construction of F, we first construct a sequence of integers $\{Q_k\}$.

$k = 1$, Q_1 is an integer such that we take $x_{i_1}^1, \dots, x_{i_{Q_1}}^1 \in C_1$, if $z \in X$ ϵ_1 shadows pieces of orbits $\{x_{i_j}^1, \dots, f^{n_1-1}x_{i_j}^1\}$, $j = 1, 2, \dots, Q_1$ with a gap m_1, then

$$\Xi L_{(Q_1-1)m_1+Q_1 n_1} z \in B(y, 2\delta_1).$$

Suppose we have already defined Q_k, we now present a construction of Q_{k+1}. Let

$$t_k = \sum_{j=1}^{k} Q_j(m_j + n_j) - m_1.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

By Lemma 3.5, there is a $Q_{k+1} \in \mathbb{N}$ such that let $x_{i_1}^{k+1}, \ldots, x_{i_{Q_{k+1}}}^{k+1} \in C_{k+1}$ with $n_{k+1} > 8m_{k+1}M/\delta_{k+1}$, then for arbitrary $x_0 \in X$, if $z \in X$ ϵ_{k+1} shadows pieces of orbits

$$\{x_0, \ldots, f^{t_k-1}x_0\}, \ldots, \{x_{i_j}^k, \ldots, f^{n_{k+1}-1}x_{i_j}^k\}, \quad j = 1, 2, \ldots, Q_{k+1}$$

with a gap m_{k+1}, then we have

$$\Xi L_{t_k+Q_{k+1}(n_{k+1}+m_{k+1})}z = \Xi L_{t_{k+1}}z \in B(y_{k+1}, 2\delta_{k+1}).$$

We can definitely require that $\{Q_k\}$ satisfies:

(1) $Q_k \geq 2^{n_{k+1}+m_{k+1}}, k \geq 1$;

(2) $Q_{k+1} \geq 2^{t_k}, k \geq 1$.

These restrictions would mean that $\{Q_k\}$ is increasing very fast and, in particular,

$$\lim_{k\to\infty} \frac{n_{k+1}+m_{k+1}}{Q_k} = 0 \quad \text{and} \quad \lim_{k\to\infty} \frac{t_k}{Q_{k+1}} = 0.$$

Second we construct inductively a sequence of finite sets L_k.

$k = 1$, for any Q_1-tuple $(i_1, \ldots, i_{Q_1}) \in \{1, 2, \ldots, M_1\}^{Q_1}$, let $y = y(i_1, \ldots, i_{Q_1})$ be some point which shadows pieces of orbits $\{x_{i_j}^1, fx_{i_j}^1, \ldots, f^{n_1-1}x_{i_j}^1\}$, $j = 1, 2, \ldots, Q_1$, with a gap m_1, i.e.

$$d_{n_1}(x_{i_j}^1, f^{a_j^1}y(i_1, \ldots, i_{Q_1})) < \epsilon_1$$

where $a_j^1 = (j-1)(n_1+m_1)$, $j = 1, 2, \ldots, Q_1$. Such a point $y = y(i_1, \ldots, i_{Q_1})$ exists because f satisfies the specification property. Collect all such points into a set

$$L_1 = \{y(i_1, \ldots, i_{Q_1}) \mid i_1, \ldots, i_{Q_1} \in \{1, 2, \ldots, M_1\}\}.$$

We claim that different tuples $(i_1, \ldots, i_{Q_1})$ produce different points $y(i_1, \ldots, i_{Q_1})$. Actually if $(i_1, \ldots, i_{Q_1}) \neq (j_1, \ldots, j_{Q_1})$, then it is easy to show that

$$d_{t_1}(y(i_1, \ldots, i_{Q_1}), y(j_1, \ldots, j_{Q_1})) > 6\epsilon.$$

This implies that $\#L_1 = M_1^{Q_1}$.

Suppose that we have already defined a set L_k, we construct L_{k+1} as follows.

For every $x \in L_k$ and $(i_1, \ldots, i_{Q_{k+1}}) \in M_{k+1}^{Q_{k+1}} = D_{k+1}$, let $z = z(x, (i_1, \ldots, i_{Q_{k+1}}))$ be some point that ϵ_{k+1} shadows pieces of orbits

$$\{x, fx, \ldots, f^{t_k-1}x\}, \quad \{x_{i_j}^{k+1}, \ldots, f^{n_{k+1}-1}x_{i_j}^{k+1}\}, \quad j = 1, 2, \ldots, Q_{k+1}$$

with a gap m_{k+1}, i.e.

$$d_{t_k}(x, z) < \epsilon_{k+1}, \quad d_{n_{k+1}}(x_{i_j}^{k+1}, f^{t_k+m_{k+1}+a_j^{k+1}}z) < \epsilon_{k+1},$$

where $a_j^{k+1} = (n_{k+1}+m_{k+1})(j-1)$. Such a point exists due to the specification property of f. Denote by $(i_1, \ldots, i_{Q_{k+1}}) = \mathbf{i}$. Collect all these points into the set

$$L_{k+1} = \{z = z(x, \mathbf{i}) \mid x \in L_k, \mathbf{i} \in D_{k+1}\}.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

We can show that different pairs $(x, \mathbf{i})$ produce different points $z = z(x, \mathbf{i})$. In fact, we have that if $(x, \mathbf{i}) \neq (x', \mathbf{i}')$ with $x, x' \in L_k$, $\mathbf{i}, \mathbf{i}' \in D_{k+1}$, then

$$d_{t_{k+1}}(z, z') > 5\epsilon.$$

This is indeed the case if one follows step (2) of Theorem 2.1 and notices that $\epsilon_k < \epsilon/2^{k-1}$.

Now for every k put $F_k = \bigcup_{x \in L_k} \overline{B_{t_k}}(x, \epsilon_k)$, where $\overline{B_l}(x, \delta)$ is the closed ball around x of radius δ in the metric d_l, i.e.

$$\overline{B_l}(x, \delta) = \{y \in X \mid d_l(x, y) \leq \delta\}.$$

Similar to Theorem 2.1, one can show that for every k the following are satisfied:

(1) for any $x, x' \in L_k$, $x \neq x'$, $\overline{B_{t_k}}(x, \epsilon_k)$, $\overline{B_{t_k}}(x', \epsilon_k)$ are disjoint;

(2) if $z \in L_{k+1}$ descends from $x \in L_k$, then

$$\overline{B_{t_{k+1}}}(z, \epsilon_{k+1}) \subseteq \overline{B_{t_k}}(x, \epsilon_k).$$

Hence, $F_{k+1} \subseteq F_k$. Finally we put $F = \bigcap_{k \geq 1} F_k$. F is a non-empty closed subset of X. Before we come to estimate the topological entropy of F, we first show the following lemma.

LEMMA 3.7. *For each $x \in F$, we have $\{y, y'\} \subseteq A(\Xi L_n x)$.*

Proof. Since $x \in F$, for each k, we have $x \in F_k$, which means that there is a $\hat{x}_k \in L_k$ such that $x \in \overline{B_{t_k}}(\hat{x}_k, \epsilon_k)$. Hence, we have

$$\begin{aligned} d(\Xi L_{t_k} x, y_k) &\leq d(\Xi L_{t_k} x, \Xi L_{t_k} \hat{x}_k) + d(\Xi L_{t_k} \hat{x}_k, y_k) \\ &\leq d(\Xi L_{t_k} x, \Xi L_{t_k} \hat{x}_k) + 2\delta_k \\ &\leq \tfrac{1}{2}\delta_k + 2\delta_k = \tfrac{5}{2}\delta_k. \end{aligned}$$

So

$$\lim_{k \to \infty} \Xi L_{t_{2k-1}} x = y \quad \text{and} \quad \lim_{k \to \infty} \Xi L_{t_{2k}} x = y',$$

which implies $\{y, y'\} \subseteq A(\Xi L_n x)$. □

Now we want to show that there is a measure μ on F such that for every sufficiently large n and every point $x \in X$ such that $B_n(x, \frac{1}{2}\epsilon) \cap F \neq \varnothing$, one has

$$\mu(B_n(x, \tfrac{1}{2}\epsilon)) \leq \exp(-ns).$$

For every $k \geq 1$, define an atomic probability measure μ_k as follows

$$\mu_k(\{z\}) = \frac{1}{\#(L_k)} \quad \text{for every } z \in L_k.$$

Obviously we have $\mu_k(F_k) = 1$. Also similar to proof in [**10**] we have that the sequence of probability measure $\{\mu_k\}$ converges in the weak topology. Denote the limiting measure by μ, we also have $\mu(F) = 1$.

For every point $x \in X$ such that $B_n(x, \epsilon/2) \cap F \neq \varnothing$, suppose $n \geq t_1$, since $B_n(x, \epsilon)$ is an open set, thus since $\mu_k \to \mu$, we have

$$\mu(B_n(x, \epsilon)) \leq \lim_{k \to \infty} \mu_k(B_n(x, \epsilon)) = \lim_{k \to \infty} \frac{1}{\#L_k} \#\{z \in L_k \cap B_n(x, \epsilon)\}.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

For such a n, there exist unique $k \geq 1$ and $j \in \{0, 1, \ldots, Q_{k+1} - 1\}$ such that

$$t_k < n \leq t_{k+1}, \quad t_k + (n_{k+1} + m_{k+1})j < n \leq t_k + (n_{k+1} + m_{k+1})(j+1).$$

First of all we have $\#(B_n(x,\epsilon) \cap L_k) \leq 1$ and thus $\mu_k(B_n(x,\epsilon)) \leq \#(L_k)^{-1}$. Actually suppose there are two different points $z_1, z_2 \in L_k \cap B_n(x,\epsilon)$. This means $d_n(z_1, z_2) < 2\epsilon$. On the other hand we have $d_{t_k}(z_1, z_2) > 5\epsilon$. Hence, we get

$$2\epsilon > d_n(z_1, z_2) \geq d_{t_k}(z_1, z_2) > 5\epsilon,$$

which is a contradiction. We continue to show $\mu_{k+1}(B_n(x,\epsilon)) \leq (\#(L_k) \times M_{k+1}^j)^{-1}$. Suppose that $z_1, z_2 \in L_{k+1} \cap B_n(x,\epsilon)$. Then there exist $x_1, x_2 \in L_k$ and $\mathbf{i}, \mathbf{i}' \in D_{k+1} = \{(i_1, \ldots, i_{Q_{k+1}}) \mid i_j \in \{1, 2, \ldots, M_{k+1}\}$ such that

$$z_1 = z(x_1, \mathbf{i}), \quad z_2 = z(x_2, \mathbf{i}').$$
$$\mathbf{i} = (i_1, \ldots, i_{Q_{k+1}}), \quad \mathbf{i}' = (i'_1, \ldots, i'_{Q_{k+1}}).$$

We claim that $x_1 = x_2$ and $(i_1, \ldots, i_j) = (i'_1, \ldots, i'_j)$. Indeed if $x_1 \neq x_2$ then

$$\begin{aligned} d_{t_k}(x_1, x_2) &\leq d_{t_k}(x_1, z_1) + d_{t_k}(z_1, x) + d_{t_k}(x, z_2) + d_{t_k}(z_2, x_2) \\ &\leq \epsilon_k + \epsilon + \epsilon + \epsilon_k \\ &\leq \frac{\epsilon}{2^k} + \epsilon + \epsilon + \frac{\epsilon}{2^k} < 5\epsilon, \end{aligned}$$

which is in contradiction with the fact that $d_{t_k}(x_1, x_2) > 5\epsilon$. Also $(i_1, \ldots, i_j) = (i'_1, \ldots, i'_j)$. If $j = 0$, there is nothing to prove. Suppose $j > 0$ and there is $1 \leq t \leq j$ such that $i_t \neq i'_t$. Since

$$\begin{aligned} d_{n_{k+1}}(x_{i_t}^{k+1}, f^{t_k + m_{k+1} + a_t^{k+1}} z_1) &\leq \epsilon_{k+1} < \frac{\epsilon}{2^{k+1}}, \\ d_{n_{k+1}}(x_{i'_t}^{k+1}, f^{t_k + m_{k+1} + a_t^{k+1}} z_2) &\leq \epsilon_{k+1} < \frac{\epsilon}{2^{k+1}}, \end{aligned}$$

where $a_t^{k+1} = (n_{k+1} + m_{k+1})(t-1)$. We have

$$\begin{aligned} d_{n_{k+1}}(x_{i_t}^{k+1}, x_{i'_t}^{k+1}) &\leq d_{n_{k+1}}(x_{i_t}^{k+1}, f^{t_k + m_{k+1} + a_t^{k+1}} z_1) + d_{n_{k+1}}(x_{i'_t}^{k+1}, f^{t_k + m_{k+1} + a_t^{k+1}} z_2) \\ &\quad + d_{n_{k+1}}(f^{t_k + m_{k+1} + a_t^{k+1}} z_1, f^{t_k + m_{k+1} + a_t^{k+1}} z_2) \\ &\leq \epsilon_{k+1} + \epsilon_{k+1} + d_n(z_1, z_2) \\ &\leq \epsilon_{k+1} + \epsilon_{k+1} + 2\epsilon \leq 4\epsilon, \end{aligned}$$

which contradicts the fact that $d_{n_{k+1}}(x_{i_t}^{k+1}, x_{i'_t}^{k+1}) > 8\epsilon$ since $x_{i_t}^{k+1}, x_{i'_t}^{k+1}$ are different points in the $(n_{k+1}, 8\epsilon)$-separated set C_{k+1}. Now since for every two points $z_1, z_2 \in L_{k+1} \cap B_n(x,\epsilon)$, let $z_1 = z(x_1, \mathbf{i})$, $z_2 = z(x_2, \mathbf{i}')$, $\mathbf{i} = (i_1, \ldots, i_{Q_{k+1}})$ and $\mathbf{i}' = (i'_1, \ldots, i'_{Q_{k+1}})$, then we have $x_1 = x_2$ and $(i_1, \ldots, i_j) = (i'_1, \ldots, i'_j)$. Hence,

$$\mu_{k+1}(B_n(x,\epsilon)) \leq \frac{1}{\#(L_k) M_{k+1}^{Q_{k+1}}} M_{k+1}^{Q_{k+1}-j} = \frac{1}{\#(L_k) M_{k+1}^j}.$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

We can continue to show that for any $p > 1$, one has

$$\mu_{k+p}\left(B_n\left(x, \frac{\epsilon}{2}\right)\right) < \frac{1}{\#(L_k)M_{k+1}^j}.$$

The conclusion is true because every point $z \in L_{k+p} \cap B_n(x, \frac{1}{2}\epsilon)$ descends from point of L_{k+1} which is in $B_n(x, \epsilon)$. Otherwise there exist $z_1 \in L_{k+1}$, $z_2 \in L_{k+p} \cap B_n(x, \frac{1}{2}\epsilon)$, z_2 descends from z_1 such that $d_n(z_1, x) > \epsilon$. Hence, we have

$$d(z_1, z_2) \geq d_n(z_1, x) - d_n(x, z_2) > \tfrac{1}{2}\epsilon.$$

On the other hand we have

$$\begin{aligned} d_n(z_1, z_2) \leq d_{t_{k+1}}(z_1, z_2) &\leq \epsilon_{k+2} + \cdots + \epsilon_{k+p} \\ &\leq \frac{\epsilon}{2^{k+2}} + \cdots + \frac{\epsilon}{2^{k+p}} \\ &\leq \frac{\epsilon}{2^{k+1}} \leq \frac{\epsilon}{2}, \end{aligned}$$

which is in contradiction with the above discussion. Hence, there are at most $M_{k+2}^{Q_{k+1}} \times \cdots \times M_{k+p}^{Q_{k+p}}$ points in $L_{k+p} \cap B_n(x, \frac{1}{2}\epsilon)$, which descends from a given point in $L_{k+1} \cap B_n(x, \epsilon)$. Hence,

$$\mu_{k+p}\left(B_n\left(x, \frac{1}{2}\epsilon\right)\right) \leq \frac{M_{k+1}^{Q_{k+1}-j} M_{k+2}^{Q_{k+2}} \cdots M_{k+p}^{Q_{k+p}}}{\#(L_k)M_{k+1}^{Q_{k+1}} M_{k+2}^{Q_{k+2}} \cdots M_{k+p}^{Q_{k+p}}} = \frac{1}{\#(L_k)M_{k+1}^j}.$$

Recall that

$$n - t_k - j(n_{k+1} + m_{k+1}) \leq n_{k+1} + m_{k+1};$$

$$t_k = \sum_{i=1}^{k} Q_i(n_i + m_i) - m_1;$$

$$\frac{n - t_k - j(n_{k+1} + m_{k+1})}{t_k + j(n_{k+1} + m_{k+1})} \leq \frac{n_{k+1} + m_{k+1}}{Q_k} \to 0 \quad \text{as } k \to \infty;$$

$$\lim_{k\to\infty} \frac{m_k}{n_k} = 0;$$

and $M_k \geq \exp(n_k s)$, we obtain for k large enough

$$\begin{aligned} \#(L_k)M_{k+1}^j &= M_1^{Q_1} M_2^{Q_2} \cdots M_k^{Q_k} M_{k+1}^j \\ &\geq \exp(s(Q_1 n_1 + \cdots + Q_k n_k + j n_{k+1})) \\ &\geq \exp((s - \tfrac{1}{2}\gamma)(Q_1(n_1 + m_1) + \cdots + Q_k(n_k + m_k) + j(n_{k+1} + m_{k+1})) \\ &\geq \exp((s - \gamma)n). \end{aligned}$$

Observe that $k \uparrow +\infty$ as $n \uparrow +\infty$. We have for all sufficiently large n

$$\mu(B_n(x, \tfrac{1}{2}\epsilon)) \leq \lim_{k\to+\infty} \mu_k(B_n(x, \tfrac{1}{2}\epsilon)) \leq \exp(-n(s - \gamma))$$

holds for every x such that $B_n(x, \frac{1}{2}\epsilon) \cap F \neq \varnothing$. Since γ is arbitrary, by the entropy distribution principle we have $h_{\text{top}}(\{y, y'\}) \geq h_{\text{top}}(F) \geq s$. Also since s is arbitrary, we get $h_{\text{top}}(\tilde{\Delta}\{y, y'\}) \geq h_{\text{top}}(F) \geq \min\{\Lambda(\Delta(y)), \Lambda(\Delta(y'))\}$. Thus finish the proof of Lemma 3.6. □

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

LEMMA 3.8. *Let M be the constant such that $d(0, \Xi\mu) \le M$ for all $\mu \in M(X)$. Then $d(\Xi L_{n+1}x, \Xi L_n x) \le 2M/(n+1)$ for all $n \in \mathbb{N}$ and $x \in X$.*

Proof. For all $n \in \mathbb{N}$, we have

$$\begin{aligned}
d(\Xi L_{n+1}x, \Xi L_n x) &= d\left(\frac{1}{n+1}\sum_{i=0}^{n}\Xi\delta_{f^i x}, \frac{1}{n}\sum_{i=0}^{n-1}\Xi\delta_{f^i x}\right)\\
&\le \sum_{i=0}^{n-1}\frac{1}{n}d\left(\frac{n}{n+1}\Xi\delta_{f^i x}, \Xi\delta_{f^i x}\right) + \frac{1}{n+1}d(\Xi\delta_{f^n x}, 0)\\
&\le \sum_{i=0}^{n-1}\frac{1}{n}\left(d(\Xi\delta_{f^i x}, \Xi\delta_{f^i x}) + d\left(\left(1-\frac{n}{n+1}\right)\Xi\delta_{f^i x}, 0\right)\right)\\
&\quad + \frac{1}{n+1}d(\Xi\delta_{f^n x}, 0)\\
&\le \sum_{i=0}^{n-1}\frac{1}{n}\left(1-\frac{n}{n+1}\right)d(\Xi\delta_{f^i x}, 0) + \frac{1}{n+1}d(\Xi\delta_{f^n x}, 0)\\
&\le \sum_{i=0}^{n-1}\frac{1}{n}\frac{1}{n+1}d(\Xi\delta_{f^i x}, 0) + \frac{1}{n+1}d(\Xi\delta_{f^n x}, 0)\\
&\le \frac{2M}{n+1}.
\end{aligned}$$

□

LEMMA 3.9. *For all $x \in X$, $n_1, n_2 \in \mathbb{N}$ and $y_1, y_2 \in Y$,*

$$d(\Xi L_{n_1+n_2}x, y_2) \le d(\Xi L_{n_1}x, y_1) + d(\Xi L_{n_2}f^{n_1}x, y_2) + d(y_1, y_2).$$

Proof. Since

$$\Xi L_{n_1+n_2}x = \frac{n_1}{n_1+n_2}\Xi L_{n_1}x + \frac{n_2}{n_1+n_2}\Xi L_{n_2}f^{n_1}x,$$

we obtain

$$\begin{aligned}
d(\Xi L_{n_1+n_2}x, y_2) &\le d\left(\frac{n_1}{n_1+n_2}\Xi L_{n_1}x + \frac{n_2}{n_1+n_2}\Xi L_{n_2}f^{n_1}x, \frac{n_1}{n_1+n_2}y_1 + \frac{n_2}{n_1+n_2}y_2\right)\\
&\quad + d\left(\frac{n_1}{n_1+n_2}y_1 + \frac{n_2}{n_1+n_2}y_2, \frac{n_1}{n_1+n_2}y_2 + \frac{n_2}{n_1+n_2}y_2\right)\\
&\le \frac{n_1}{n_1+n_2}d(\Xi L_{n_1}x, y_1) + \frac{n_2}{n_1+n_2}d(\Xi L_{n_2}f^{n_1}x, y_2)\\
&\quad + \frac{n_1}{n_1+n_2}d(y_1, y_2)\\
&\le d(\Xi L_{n_1}x, y_1) + d(\Xi L_{n_2}f^{n_1}x, y_2) + d(y_1, y_2).
\end{aligned}$$

□

LEMMA 3.10. *If $y \in A(\Xi L_n x)$ for some $x \in X$, then there is some $x' \in X$ such that $A(\Xi L_n x') = \{y\}$.*

Proof. Suppose $\lim_{k\to+\infty}\Xi L_{n_k}x = y$. Choose a sequence $\delta_k \downarrow 0$. For each δ_k, there is a $r_k = r(\delta_k)$ such that for all $x, z \in X$ with $d_n(x, z) < r_k$, we have $d(\Xi L_n x, \Xi L_n y) < \delta_k/2$. Choose $\epsilon_k \downarrow 0$ such that $\epsilon_k < r_k$, $\epsilon_{k+1} < \frac{1}{2}\epsilon_k$. By the definition of the specification

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

property, for each ϵ_k, there is $m_k = m(\epsilon_k)$ such that for arbitrary pieces of orbits, there exists a point that belongs to X, which ϵ_k shadows the pieces of orbits with a gap m_k. Hence, without loss of generality we can assume that n_k satisfies

$$\Xi L_{n_k} x \in B(y, \delta_k) \quad \text{and} \quad n_k > \max\left\{2^{m_k}, \frac{8m_k M}{\delta_k}\right\}.$$

We construct a sequence of integers $\{Q_k\}$ and a sequence of points $\{x_k\}$ such that $x_k \to x'$ and $A(\Xi L_n x') = \{y\}$.

$k = 1$, by Lemma 3.5 there is an integer Q_1 such that if $x_1 \in X$ ϵ_1 shadows Q_1 pieces of orbits

$$\underbrace{(x, fx, \ldots, f^{n_1-1}x), \ldots, (x, \ldots, f^{n_1-1}x)}_{Q_1}$$

then

$$\Xi L_{(Q_1-1)m_1+Q_1 n_1} x_1 \in B(y, 2\delta_1).$$

Fix Q_1, such x_1 exists due to the specification property. Fix x_1. Suppose we have already defined Q_k and x_k, now we come to construct Q_{k+1} and x_{k+1}. Let

$$t_k = \sum_{i=1}^{k} Q_i(m_i + n_i) - m_1.$$

By Lemma 3.5, there is Q_{k+1} such that if $x'' \in X$ ϵ_{k+1} shadows pieces of orbits

$$\{x_k, \ldots, f^{t_k-1}x_k\}, \underbrace{\{x, \ldots, f^{n_{k+1}-1}x\}, \ldots, \{x, \ldots, f^{n_{k+1}-1}x\}}_{Q_{k+1}}$$

with a gap m_{k+1}, then we have

$$\Xi L_{t_k+Q_{k+1}(n_{k+1}+m_{k+1})} x'' = \Xi L_{t_{k+1}} x'' \in B(y, 2\delta_{k+1}).$$

Fix Q_{k+1} and take such a x'' as x_{k+1}. Without loss of generality we can require that $\{Q_k\}$ satisfies

$$\lim_{k\to\infty} \frac{n_{k+1} + m_{k+1}}{Q_k} = 0 \quad \text{and} \quad \lim_{k\to\infty} \frac{t_k}{Q_{k+1}} = 0.$$

Finally let $x' \in \bigcap_k \overline{B_{t_k}}(x_k, \epsilon_k)$, we have that such a point exists.

We show $A(\Xi L_n x') = \{y\}$. Take a large integer n, there are unique $k \geq 1$ and $j \in \{0, 1, \ldots, Q_{k+1} - 1\}$ such that

$$t_k < n \leq t_{k+1}, \quad \hat{n} = t_k + (n_{k+1} + m_{k+1})j < n \leq t_k + (n_{k+1} + m_{k+1})(j+1).$$

We come to estimate $d(\Xi L_n x', y)$.

$$d(\Xi L_n x', y) \leq d(\Xi L_n x', \Xi L_{\hat{n}} x') + d(\Xi L_{\hat{n}} x', y).$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

By Lemma 3.8, we have

$$\begin{aligned}
&d(\Xi L_n x', \Xi L_{\hat{n}} x') \\
&\quad \le d(\Xi L_{\hat{n}} x', \Xi L_{\hat{n}+1} x') + d(\Xi L_{\hat{n}+1} x', \Xi L_{\hat{n}+2} x') + \cdots + d(\Xi L_{n-1} x', \Xi L_n x') \\
&\quad = \frac{2M}{\hat{n}+1} + \frac{2M}{\hat{n}+2} + \cdots + \frac{2M}{n} \\
&\quad \le \frac{2M(n-\hat{n})}{\hat{n}} \\
&\quad \le 2M \frac{n_{k+1}+m_{k+1}}{t_k + j(n_{k+1}+m_{k+1})} \\
&\quad = 2M \frac{n_{k+1}+m_{k+1}}{\sum_{i=1}^{k} Q_i(n_i+m_i) - m_1 + j(n_{k+1}+m_{k+1})} \to 0
\end{aligned}$$

as k goes to infinity. We go on to estimate $d(\Xi L_{\hat{n}} x', y)$. By Lemma 3.9,

$$\begin{aligned}
d(\Xi L_{\hat{n}} x', y) &\le d(\Xi L_{t_k} x', y) + d(\Xi L_{\hat{n}-t_k} f^{t_k} x', y) \\
&\le d(\Xi L_{t_k} x', \Xi L_{t_k} x_k) + d(\Xi L_{t_k} x_k, y) + d(\Xi L_{\hat{n}-t_k} f^{t_k} x', y) \\
&\le \tfrac{1}{2}\delta_k + 2\delta_k + d(\Xi L_{\hat{n}-t_k} f^{t_k} x', y) \\
&\le \tfrac{5}{2}\delta_k + d(\Xi L_{\hat{n}-t_k} f^{t_k} x', \Xi L_{\hat{n}-t_k} f^{t_k} x_{k+1}) + d(\Xi L_{\hat{n}-t_k} f^{t_k} x_{k+1}, y) \\
&\le \tfrac{5}{2}\delta_k + \tfrac{1}{2}\delta_{k+1} + d(\Xi L_{\hat{n}-t_k} f^{t_k} x_{k+1}, y).
\end{aligned}$$

Since $x_{k+1} \in L_{k+1}$, therefore x_{k+1} ϵ_{k+1} shadows pieces of orbits

$$\{x_k, \ldots, f^{t_k-1} x_k\}, \underbrace{\{x, \ldots, f^{N_{k+1}-1} x\}, \ldots, \{x, \ldots, f^{N_{k+1}-1} x\}}_{Q_{k+1}}$$

with a gap m_{k+1} and $\Xi L_{n_{k+1}} x \in B(y, \delta_{k+1})$, we have

$$\begin{aligned}
d(\Xi L_{\hat{n}-t_k} f^{t_k} x_{k+1}, y) &\le d(\Xi L_{\hat{n}-t_k} f^{t_k} x_{k+1}, \Xi L_{n_{k+1}} x) + d(\Xi L_{n_{k+1}} x, y) \\
&\le d(\Xi L_{\hat{n}-t_k} f^{t_k} x_{k+1}, \Xi L_{n_{k+1}} x) + \delta_{k+1}.
\end{aligned}$$

Rewrite $L_{\hat{n}-t_k} f^{t_k} x_{k+1}$ as follows:

$$\begin{aligned}
L_{\hat{n}-t_k} f^{t_k} x_{k+1} &= \frac{m_{k+1}}{\hat{n}-t_k} L_{m_{k+1}} f^{t_k} x_{k+1} + \frac{n_{k+1}}{\hat{n}-t_k} L_{n_{k+1}} f^{t_k+m_{k+1}} x_{k+1} + \cdots \\
&\quad + \frac{n_{k+1}}{\hat{n}-t_k} L_{n_{k+1}} f^{t_k+(j-1)(n_{k+1}+m_{k+1})+m_{k+1}} x_{k+1} \\
&= \sum_{i=1}^{j} \frac{n_{k+1}}{\hat{n}-t_k} L_{n_{k+1}} f^{t_k+m_{k+1}+(i-1)(n_{k+1}+m_{k+1})} x_{k+1} \\
&\quad + \sum_{i=1}^{j} \frac{m_{k+1}}{\hat{n}-t_k} L_{m_{k+1}} f^{t_k+(i-1)(n_{k+1}+m_{k+1})} x_{k+1}.
\end{aligned}$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Since $\Xi : M(X) \mapsto Y$ is affine and Y is a vector space with linearly compatible metric, we have

$$\begin{aligned}
&d(\Xi L_{\hat{n}-t_k} f^{t_k} x_{k+1}, \Xi L_{n_{k+1}} x) \\
&\quad \le d\left(\sum_{i=1}^{j} \frac{n_{k+1}}{\hat{n}-t_k} \Xi L_{n_{k+1}} f^{t_k+m_{k+1}+(i-1)(n_{k+1}+m_{k+1})} x_{k+1}, \Xi L_{n_{k+1}} x \right) \\
&\qquad + \sum_{i=1}^{j} \frac{m_{k+1}}{\hat{n}-t_k} d(\Xi L_{m_{k+1}} f^{t_k+(i-1)(n_{k+1}+m_{k+1})} x_{k+1}, 0) \\
&\quad \le \sum_{i=1}^{j} d\left(\frac{n_{k+1}}{\hat{n}-t_k} \Xi L_{n_{k+1}} f^{t_k+m_{k+1}+(i-1)(n_{k+1}+m_{k+1})} x_{k+1}, \frac{1}{j} \Xi L_{n_{k+1}} x \right) \\
&\qquad + \frac{m_{k+1}}{n_{k+1}+m_{k+1}} M \\
&\quad \le \frac{m_{k+1}}{n_{k+1}+m_{k+1}} M + \sum_{i=1}^{j} \frac{1}{j} d(\Xi L_{n_{k+1}} f^{t_k+m_{k+1}+(i-1)(n_{k+1}+m_{k+1})} x_{k+1}, \Xi L_{n_{k+1}} x) \\
&\qquad + \sum_{i=1}^{j} d\left(\left(\frac{n_{k+1}}{\hat{n}-t_k} - \frac{1}{j} \right) \Xi L_{n_{k+1}} f^{t_k+m_{k+1}+(i-1)(n_{k+1}+m_{k+1})} x_{k+1}, 0 \right) \\
&\quad \le \frac{m_{k+1}}{n_{k+1}+m_{k+1}} M + \delta_{k+1} + \sum_{i=1}^{j} \left| \frac{n_{k+1}}{\hat{n}-t_k} - \frac{1}{j} \right| M \\
&\quad \le \frac{m_{k+1}}{n_{k+1}+m_{k+1}} M + \delta_{k+1} + \sum_{i=1}^{j} \frac{m_{k+1}}{j(n_{k+1}+m_{k+1})} M \\
&\quad = \frac{2m_{k+1}}{n_{k+1}+m_{k+1}} M + \delta_{k+1} \to 0.
\end{aligned}$$

Hence, we have $\lim_{n\to\infty} d(\Xi L_n x', y) = 0$, which implies that we have found a point $x' \in X$ such that $A(\Xi L_n x') = \{y\}$. □

Proof of Theorem 3.1. Suppose Theorem 3.1(1) does not hold, then $\bigcup_{x\in X} A(\Xi L_n x)$ contains at least two points. It suffices to show $h_{\text{top}}(D(f, \Xi)) \ge h_{\text{top}}(f, X)$. Without loss of generality we can assume $h_{\text{top}}(f, X) > 0$. By the classical variational principle, we have

$$\begin{aligned}
h_{\text{top}}(f, X) &= \sup_{\mu\in M(X,f)} h_\mu(f) \\
&= \sup_{y\in \Xi M(X,f)} \sup_{\mu\in M(X,f), \Xi\mu=y} \{h_\mu(f)\} \\
&= \sup_{y\in \Xi M(X,f)} \Lambda(\Delta(y)).
\end{aligned}$$

Since $\Lambda\Delta : C(Y) \mapsto \mathbb{R}$ is upper semi-continuous and $\Xi M(X, f)$ is compact, there exists a point y_1 such that

$$h_{\text{top}}(f, X) = \Lambda(\Delta(y_1)).$$

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Since $\bigcup_{x\in X} A(\Xi L_n x)$ contains at least two points, by Lemma 3.10 there is a $y \in Y$ which differs from y_1 such that $\Delta(y) \neq \varnothing$. By Theorem 2.1,

$$\Lambda(\Delta((1-\delta)y_1 + \delta y)) \geq (1-\delta)\Lambda(\Delta(y_1)) + \delta\Lambda(\Delta(y)).$$

Hence, for $\delta < 1$, $\Delta((1-\delta)y_1 + \delta y) \neq \varnothing$. For any $\delta > 0$ consider the set $\tilde{\Delta}(\{y_1, (1-\delta)y_1 + \delta y\})$. By Lemma 3.6,

$$h_{\text{top}}(\tilde{\Delta}(\{y_1, (1-\delta)y_1 + \delta y\})) = \min\{\Lambda(\Delta(y_1)), \Lambda(\Delta((1-\delta)y_1 + \delta y))\}.$$

Letting $\delta \to 0$, we observe that

$$\tilde{\Delta}(\{y_1, (1-\delta)y_1 + \delta y\}) \subseteq D(f, \Xi) = \{x \in X \mid \text{the limit } \Xi L_n x \text{ does not exist}\}$$

and we get

$$h_{\text{top}}(D(f, \Xi)) \geq \lim_{\delta\to 0} h_{\text{top}}(\tilde{\Delta}(\{y_1, (1-\delta)y_1 + \delta y\})) \geq \Lambda(\Delta(y_1)).$$

Hence,

$$h_{\text{top}}(D(f, \Xi)) \geq \Lambda(\Delta(y_1)) = h_{\text{top}}(f, X).$$

This completes the proof of Theorem 3.1. □

The following are applications of Theorem 3.1.

COROLLARY 3.11. *Suppose $f : X \mapsto X$ is a continuous transformation with specification property on a compact metric space X. Let $\varphi : X \mapsto \mathbb{R}$ be continuous. Then:*

(1) *either all the points have the same ergodic average; or*
(2) $h_{\text{top}}\{x \in X \mid \text{the limit } (1/n)\sum_{i=0}^{n-1} \varphi(f^i x) \text{ does not exist}\} = h_{\text{top}}(f, X)$.

COROLLARY 3.12. *Suppose $f : X \mapsto X$ is a continuous transformation with specification property on a compact metric space X. Then*

(1) *either all the points have the same limit, i.e.* $\lim_{n\to\infty}(1/n)\sum_{i=0}^{n-1} \delta_{f^i x} = y_0$*; or*
(2) $h_{\text{top}}\{x \in X \mid \text{the limit } (1/n)\sum_{i=0}^{n-1} \delta_{f^i x} \text{ does not exist}\} = h_{\text{top}}(f, X)$.

THEOREM 3.13. *Suppose $f : X \mapsto X$ is a continuous transformation with specification property on a compact metric space X. Let $(Y_i, \Xi_i)_i$ be (a possibly uncountable) family of deformations and assume that Y_i is a normed vector space and that $\Xi_i : M(X) \mapsto Y_i$ is affine and continuous. Define the vector spaces $\mathbf{x}_i Y_i$ and $[\mathbf{x}_i Y_i]^\infty$ by*

$$\mathbf{x}_i Y_i = \{(y_i)_i \mid y_i \in Y_i \text{ for all } i\},$$
$$[\mathbf{x}_i Y_i]^\infty = \{(y_i)_i \in \mathbf{x}_i Y_i \mid \sup_i \|y_i\| < \infty\},$$

and equip $[\mathbf{x}_i Y_i]^\infty$ with the norm $\|(y_i)_i\| = \sup_i \|y_i\|$. Assume that $\sup_{\mu\in M(X),i} \|\Xi_i \mu\| < \infty$ and that the map

$$M(X) \to [\mathbf{x}_i Y_i]^\infty : \mu \to (\Xi_i \mu)_i$$

is continuous. Then:

(1) *either all points $x \in X$ have the same limit, i.e.* $\lim_{n\to\infty}(\Xi_i L_n x)_i = y_0 \in [\mathbf{x}_i Y_i]^\infty$ *or;*
(2) $h_{\text{top}}\{x \in X \mid \text{the limit } (\Xi_i L_n x)_i \text{ does not exist}\} = h_{\text{top}}(f, X)$.

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press

Acknowledgements. This work was partially carried out during the first author's stay at the University of Cologne in Germany. The author would like to thank Professor Küpper at Cologne university for his invitation to join his research group for cooperation. The first author is grateful to the Mathematical Institute of the University of Cologne for its hospitality. The first author thanks the VW-Foundation in Germany for support during the author's stay in Germany. This work was partially supported by the National Natural Science Foundation of China (10271057) and the Special Funds for Major State Basic Research Projects in China.

REFERENCES

[1] R. Bowen. Topological entropy for noncompact set. *Trans. Amer. Math. Soc.* **184** (1973), 125–136.
[2] E. C. Chen and J. C. Xiong. The pointwise dimension of self-similar measures. *Chin. Sci. Bull.* **44** (1999), 2136–2140.
[3] R. Ellis. *Entropy, Large Deviations, and Statistical Mechanics (Grundlehren der Mathmaticschen Wissenschaften, 271)*. Springer, New York, 1985.
[4] A. H. Fan and D. J. Feng. On the distribution of long-term time averages on symbolic space. *J. Statist. Phys.* **99** (2000), 813–856.
[5] A. H. Fan, D. J. Feng and J. Wu. Recurrence, dimension and entropy. *J. London Math. Soc.* **64** (2001), 229–244.
[6] D. J. Feng, K. S. Lau and J. Wu. Ergodic limits on the conformal repellers. *Adv. Math.* **169** (2002), 58–91.
[7] L. Olsen. Multifractal analysis of divergence points of deformed measure theoretical Birkhoff averages. *J. Math. Pures Appl.* **82** (2003), 1591–1649.
[8] L. Olsen. Multifractal analysis of divergence points of deformed measure theoretical Birkhoff averages III. *Preprint*, 2001.
[9] L. Olsen and S. Winter. Multifractal analysis of divergence points of deformed measure theoretical Birkhoff averages II. *Preprint*, 2001.
[10] F. Takens and E. Verbitskiy. On the variational principle for the topological entropy of certain non-compact sets. *Ergod. Th. & Dynam. Sys.* **23** (2003), 317–348.
[11] P. Walters. *An Introduction to Ergodic Theory*. Springer, Berlin, 1982.

https://doi.org/10.1017/S0143385704000872 Published online by Cambridge University Press