RECURRENCE, DIMENSION AND ENTROPY

AI-HUA FAN, DE-JUN FENG AND JUN WU

Abstract

Let (Σ_A, T) be a topologically mixing subshift of finite type on an alphabet consisting of m symbols and let $\Phi: \Sigma_A \longrightarrow \mathbf{R}^d$ be a continuous function. Denote by $\sigma_\Phi(x)$ the ergodic limit $\lim_{n\to\infty} n^{-1} \sum_{j=0}^{n-1} \Phi(T^j x)$ when the limit exists. Possible ergodic limits are just mean values $\int \Phi d\mu$ for all T-invariant measures. For any possible ergodic limit α, the following variational formula is proved:

$$h_{\mathrm{top}}(\{x \in \Sigma_A : \sigma_\Phi(x) = \alpha\}) = \sup\left\{h_\mu : \int \Phi d\mu = \alpha\right\}$$

where h_μ denotes the entropy of μ and h_{top} denotes topological entropy. It is also proved that unless all points have the same ergodic limit, then the set of points whose ergodic limit does not exist has the same topological entropy as the whole space Σ_A.

1. *Introduction*

Let T be the shift map on $\Sigma = \{0, 1, \ldots, m-1\}^{\mathbf{N}}$ ($m \geqslant 2$ an integer). Given an $m \times m$ matrix A with entries 0 or 1, we consider the subshift of finite type (Σ_A, T) [**4**]. We shall always assume that A is primitive. That means the dynamical system (Σ_A, T) is topologically mixing. Now let Φ be a continuous function defined on Σ_A taking values in $\mathbf{R}^d$. We consider the ergodic limit, when it exists,

$$\sigma_\Phi(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \Phi(T^j x).$$

The quantity $\sigma_\Phi(x)$ is regarded as the *recurrence* of x relative to Φ (the term 'recurrence' takes its usual sense when $\Phi = (1_{B_1}, \ldots, 1_{B_d})$ where 1_B denotes the characteristic function of a set B).

Let L_Φ be the set of α such that $\alpha = \sigma_\Phi(x)$ for some $x \in \Sigma_A$. As a consequence of the Birkhoff ergodic theorem, L_Φ is a non-empty compact convex set. In this paper, we investigate the sizes of the sets with given recurrences:

$$E_\Phi(\alpha) = \{x \in \Sigma_A : \sigma_\Phi(x) = \alpha\} \quad (\alpha \in L_\Phi).$$

We also investigate the size of the set of points such that the limit defining $\sigma_\Phi(x)$ does not exist. The size of the sets in σ_A will be measured by their topological entropy. Notice that h_{top} is well-defined for non-compact invariant sets using Bowen's definition [**3**].

Let $\mathscr{M}_{\mathrm{inv}}$ be the set of all T-invariant Borel probability measures concentrated on Σ_A. The function $\Phi: \Sigma_A \longrightarrow \mathbf{R}^d$ induces a map $\Phi_*: \mathscr{M}_{\mathrm{inv}} \longrightarrow \mathbf{R}^d$, called the *projection map*, given by

$$\Phi*(\mu) = \int_{\Sigma_A} \Phi d\mu \quad (\mu \in \mathscr{M}_{\mathrm{inv}}).$$

We notice that $L_\Phi = \Phi_*(\mathscr{M}_{\mathrm{inv}})$ (thus L_Φ is non-empty, convex and compact). For $\alpha \in L_\Phi$, let

$$\mathscr{F}_\Phi(\alpha) = \{\mu \in \mathscr{M}_{\mathrm{inv}} : \Phi_*(\mu) = \alpha\}.$$

We call $\mathscr{F}_\Phi(\alpha)$ the *fibre of projection* α.

Received 5 July 1999; revised 25 September 2000.

2000 *Mathematics Subject Classification* 58F03 (primary), 28A78 (secondary).

J. London Math. Soc. (2) 64 (2001) 229–244.

The main results of the present paper are the following two theorems.

THEOREM A. *Suppose that* $\Phi:\Sigma_A \longrightarrow \mathbf{R}^d$ *is a continuous function. For any* $\alpha \in L_\Phi$, *we have the variational formula*

$$h_{\text{top}}(E_\Phi(\alpha)) = \max_{\mu \in \mathscr{F}_\Phi(\alpha)} h_\mu$$

where h_μ *is the entropy of* μ. *Moreover,* $h_{\text{top}}(E_\Phi(\alpha))$ *is an upper semi-continuous function of* α.

THEOREM B. *Suppose that* $\Phi:\Sigma_A \longrightarrow \mathbf{R}^d$ *is a continuous function. Then the set of points* x *such that the limit defining* $\sigma_\Phi(x)$ *does not exist is of the same topological entropy as that of* Σ_A *unless all points* $x \in \Sigma_A$ *have the same ergodic limit.*

The techniques we use in proving our two theorems are inspired by dimension theory. Recall that Σ is a metric space where a metric is defined by $d(x, y) = m^{-n}$ for $x = (x_j)_{j\geqslant 1}$ and $y = (y_j)_{\geqslant 1}$ where n is the largest value such that $x_j = y_j$ $(1 \leqslant j \leqslant n)$. Different notions of dimensions are then defined on Σ. We shall talk about the Hausdorff dimension $\dim_{\mathrm{H}}$, the packing dimension $\dim_{\mathrm{P}}$ and the upper box dimension $\dim_{\mathrm{B}}$ (see [**8**, **15**, **17**] for a general account of dimensions). We will exploit the fact that, in our purely symbolic setting, topological entropy is related to dimension by

$$\frac{1}{\log m} h_{\text{top}}(E_\Phi(\alpha)) = \dim_{\mathrm{H}}(E_\Phi(\alpha)) = \dim_{\mathrm{P}}(E_\Phi(\alpha)).$$

Theorem A is a variational principle, but it does not follow from the well-known variational principle of Walters [**22**], since the invariant set $E_\Phi(\alpha)$ is not compact. We emphasize that Theorem A holds for any continuous function Φ. If Φ has some regularity like Hölder continuity or summable variation, the result of Theorem A is part of the folklore in multifractal analysis (the thermodynamical formalism is used there but it does not work in our case because there is a lack of differentiability of pressure function and a lack of Gibbs property). Even for these regular functions, discussions of $E_\Phi(\alpha)$ for boundary points α of L_Φ are scarce, which is actually a subtle problem.

The invariant set studied by Theorem B may be called the divergence set. The Birkhoff ergodic theorem says that the divergence set is of zero measure with respect to any invariant measure. Theorem B states that it is either empty or large in the sense that it has full topological entropy. The result of full topological entropy was obtained for Hölder functions by Barreira and Schmeling [**1**] (see also [**5**, **18**]). They used the thermodynamical formalism which does not work for merely continuous functions.

The variational formula in Theorem A for those Φ depending only on finitely many coordinates was simplified in [**10**].

The key points in our proof are: the entropy h_μ is upper semi-continuous and can be approximated by the conditional entropies of μ; each conditional entropy is the entropy of a Markov measure; the recurrent set $E_\Phi(\alpha)$ can be approximated by homogeneous Moran sets.

The paper is organized as follows. §2 is devoted to preliminaries. The variational formula is proved in §3 and the divergence set is studied in §4. Finally §5 contains some remarks.

2. *Preliminaries*

Most of the material in this section is known. We recall it here for convenience, at the same time introducing notation.

For $k \geqslant 1$, $\Sigma_{A,k}$ denotes the set of finite sequences $\omega = (x_1, \ldots, x_k)$ such that $a_{x_i, x_{i+1}} = 1$ for all $1 \leqslant i < k$. These sequences ω are called (*admissible*) *words of length* $|\omega|$ $(= k)$. For $\omega = (a_1, \ldots, a_k) \in \Sigma_{A,k}$, the *k-cylinder* $[\omega]$ is defined by $\{x \in \Sigma_A : x_1 = a_1, \ldots, x_k = a_k\}$. There is a one-to-one correspondence between $\Sigma_{A,k}$ and the set of k-cylinders, so sometimes we shall use $\Sigma_{A,k}$ to denote the set of all k-cylinders. The prefix of length n of a point $x \in \Sigma$ will be written as $x|_n$.

Let ξ_0 be the partition consisting of all 1-cylinders $[0], [1], \ldots, [m-1]$. Let ξ_n be the join of the partitions $T^{-j}\xi_0$ $(0 \leqslant j \leqslant n)$. Since ξ_0 is a generator, the entropy h_μ of an invariant measure $\mu \in \mathscr{M}_{\text{inv}}$ can be expressed as [**22**]

$$h_\mu = \lim_{n\to\infty} \frac{H_\mu(\xi_n)}{n} \quad \text{where } H_\mu(\xi_n) = -\sum_{A \in \xi_n} \mu(A) \log \mu(A).$$

The *nth conditional entropy of* μ, denoted by $h_\mu^{(n)}$, is defined by

$$h_\mu^{(0)} = H_\mu(\xi_0), \quad h_\mu^{(n)} = H_\mu(\xi_n) - H_\mu(\xi_{n-1}) \ \ (\forall\, n \geqslant 1).$$

Using elementary properties of the conditional entropy [**22**, p. 80], the following proposition may be proved.

PROPOSITION 1. *For each $\mu \in \mathscr{M}_{\text{inv}}$, we have*

$$h_\mu = \lim_{n\to\infty} h_\mu^{(n)} = \inf_n h_\mu^{(n)}.$$

The entropy h_μ is an upper semi-continuous functional defined on $\mathscr{M}_{\text{inv}}$ with respect to the weak topology.*

Markov measures on the full shift space were discussed in [**7**]. We present them here for subshifts of finite type. Markov measures form a special class of invariant measures and they are dense in $\mathscr{M}_{\text{inv}}$. A Borel probability measure μ on Σ_A is uniquely determined by its values on cylinders. On the other hand, any set function μ defined on cylinders satisfying the following conditions: for all $a \in \Sigma_{A,n}$, all $n \geqslant 1$

$$\sum_{a \in \Sigma_{A,n}} \mu([a]) = 1, \quad \sum_{\epsilon} \mu([a, \epsilon]) = \mu([a])$$

may be uniquely extended to a Borel probability measure on Σ_A. Such a measure μ is invariant if and only if for all $a \in \Sigma_{A,n}$, all $n \geqslant 1$

$$\sum_{\epsilon} \mu([\epsilon, a]) = \mu([a]).$$

These three conditions may be referred to as the normalization condition, the consistency condition and the invariance condition. Let $\ell \geqslant 1$ be an integer. By a *Markov measure of order* ℓ or simply ℓ*-Markov measure*, we mean a measure $\mu \in \mathscr{M}_{\text{inv}}$ having the Markov property

$$\mu([\epsilon_1, \ldots, \epsilon_n]) = \mu([\epsilon_1, \ldots, \epsilon_{n-1}]) \frac{\mu([\epsilon_{n-\ell}, \ldots, \epsilon_n])}{\mu([\epsilon_{n-\ell}, \ldots, \epsilon_{n-1}])} \, (n > \ell).$$

The proof of the following proposition is straightforward.

PROPOSITION 2. *Suppose that $\mu \in \mathscr{M}_{\mathrm{inv}}$ is an ℓ-Markov measure. The entropy of μ is*

$$h_\mu = -\sum_{\epsilon_1,\ldots,\epsilon_{\ell+1}} \mu([\epsilon_1,\ldots,\epsilon_{\ell+1}]) \log \frac{\mu([\epsilon_1,\ldots,\epsilon_{\ell+1}])}{\mu([\epsilon_1,\ldots,\epsilon_\ell])}.$$

Moreover, $h_\mu^{(n)} = h_\mu$ for $n \geqslant \ell+1$.

For $n \geqslant 1$, denote by Δ_n the set of probability vectors p defined on $\Sigma_{A,n}$ satisfying

$$\sum_\epsilon p(\epsilon, \epsilon_1, \ldots, \epsilon_{n-1}) = \sum_\epsilon p(\epsilon_1, \ldots, \epsilon_{n-1}, \epsilon).$$

The Markov property is equivalent to

$$\mu([\epsilon_1,\ldots,\epsilon_{n+\ell}]) = \mu([\epsilon_1,\ldots,\epsilon_{\ell+1}]) \prod_{j=2}^{n} \frac{\mu([\epsilon_j,\ldots,\epsilon_{j+\ell}])}{\mu([\epsilon_j,\ldots,\epsilon_{j+\ell-1}])} \quad (n \geqslant 1).$$

It follows that an ℓ-Markov measure is uniquely determined by the function p defined on $\Sigma_{A,\ell+1}$ by

$$p(x_1,\ldots,x_{\ell+1}) = \mu([x_1,\ldots,x_{\ell+1}])$$

which belongs to $\Delta_{\ell+1}$. Conversely, given $p \in \Delta_k$, we define for $n > k$

$$\mu([a_1,\ldots,a_n]) = p(a_1,\ldots,a_k) \prod_{j=2}^{n-k+1} \frac{p(a_j,\ldots,a_{j+k-1})}{\sum_\epsilon p(a_j,\ldots,a_{j+k-2},\epsilon)}.$$

This set function μ can be uniquely extended to a $(k-1)$-Markov measure. In this way, we get a one-to-one correspondence between Δ_k and the set of all $(k-1)$-Markov measures.

Let $p \in \Delta_k$ $(k \geqslant 1)$. If $p(x) > 0$ for $x \in \Sigma_{A,k}$, we write $p \in \Delta_k^+$. From the last expression of $\mu([a_1,\ldots,a_n])$ we get the following (see **[4]** for the definition of Gibbs measures).

PROPOSITION 3. *Suppose that $p \in \Delta_k^+$. Then the $(k-1)$-Markov measure corresponding to p is the Gibbs measure associated to the potential*

$$\psi(x) = \log p(x_1,\ldots,x_k) - \log \sum_\epsilon p(x_1,\ldots,x_{k-1},\epsilon).$$

Let us summarize. Given $\mu \in \mathscr{M}_{\mathrm{inv}}$, we have a sequence of Markov measures $\{\mu_k\}$ which correspond to the functions $\{p_k \in \Delta_k\}$ induced from μ by

$$p_k(a_1,\ldots,a_k) = \mu([a_1,\ldots,a_k]).$$

These Markov measures approach μ in the following sense (Propositions 1 and 2):

$$\mu = w^* - \lim_{k\to\infty} \mu_k, \quad h_\mu = \lim_{k\to\infty} h_{\mu_k}.$$

We call μ_k the *kth Markov approximation* of μ. If the support of μ is the whole space Σ_A, its Markov approximations are all Gibbs measures then ergodic (Proposition 3).

In our proof of Theorem A, we need to construct a kind of Cantor set, called a homogeneous Moran set. It is helpful to think of Σ as the interval $[0, 1]$ and cylinders as subintervals. Let $\{n_k\}_{k\geqslant 1}$ be a sequence of positive integers and $\{c_k\}_{k\geqslant 1}$ be a sequence of positive numbers satisfying $n_k \geqslant 2$, $0 < c_k < 1$, $n_1 c_1 \leqslant \delta$ and $n_k c_k \leqslant 1$ $(k \geqslant 2)$, where δ is some positive number. Let

$$D = \bigcup_{k\geqslant 0} D_k \text{ with } D_0 = \{\varnothing\}, D_k = \{(i_1, \ldots, i_k); 1 \leqslant i_j \leqslant n_j, 1 \leqslant j \leqslant k\}.$$

If $\sigma = (\sigma_1, \ldots, \sigma_k) \in D_k$, $\tau = (\tau_1, \ldots, \tau_m) \in D_m$, we define $\sigma * \tau = (\sigma_1, \ldots, \sigma_k, \tau_1, \ldots, \tau_m)$. Suppose that J is an interval of length δ. A collection $\mathscr{F} = \{J_\sigma : \sigma \in D\}$ of subintervals of J is said to have a *homogeneous Moran structure* if it satisfies:

(1) $J_\varnothing = J$.

(2) For any $k \geqslant 0$ and $\sigma \in D_k$, $J_{\sigma*1}, J_{\sigma*2}, \ldots, J_{\sigma*n_{k+1}}$ are subintervals of J_σ and $J_{\sigma*i} \bigcap J_{\sigma*j} = \varnothing$ $(i \neq j)$.

(3) For any $k \geqslant 1$ and any $\sigma \in D_{k-1}$, $1 \leqslant j \leqslant n_k$, we have

$$\frac{|J_{\sigma*j}|}{|J_\sigma|} = c_k$$

where $|A|$ denotes the length of A.

If $\mathscr{F}$ is such a collection, $E := \bigcap_{k\geqslant 1} \bigcup_{\sigma \in D_k} J_\sigma$ is called a *homogeneous Moran set* determined by $\mathscr{F}$.

PROPOSITION 4 [**11**, **12**]. *For the homogeneous Moran set defined above, we have*

$$\dim_{\mathrm{H}} E \geqslant \liminf_{n\to\infty} \frac{\log n_1 n_2 \ldots n_k}{-\log c_1 c_2 \ldots c_{k+1} n_{k+1}}.$$

To end this section, we point out that the topological entropy of a subshift Σ_A is, up to a multiplicative factor $\log m$, equal to the Hausdorff dimension (and its packing) dimension.

3. *Proof of Theorem* A

Although the theorem is stated in terms of entropy, it will be more convenient to use dimension during the proof. Actually, we will prove that

$$\dim_{\mathrm{H}} E_\Phi(\alpha) = \dim_{\mathrm{P}} E_\Phi(\alpha) = \frac{1}{\log m} \max_{\mu \in \mathscr{F}_\Phi(\alpha)} h_\mu.$$

First we prove a formal formula of dimension. The matrix A being primitive, there is an integer $M \geqslant 1$ such that all the entries of A^M are strictly positive. Then for any $\omega = (x_j)_{j=1}^n \in \Sigma_{A,n}$ and any $0 \leqslant z \leqslant m-1$, there are $0 \leqslant y_1, \ldots, y_{M-1} \leqslant m-1$ such that

$$(x_1, \ldots, x_n, y_1, \ldots, y_{M-1}, z) \in \Sigma_{A, n+M}.$$

We call $\varpi = (x_1, \ldots, x_n, y_1, \ldots, y_{M-1})$ an *extension* of ω joining z.

Let $\alpha \in \mathbf{R}^d$, $n \geqslant 1$ and $\epsilon > 0$. Denote by $F(\alpha, n, \epsilon)$ the set of all n-cylinders each of which contains at least a point x such that

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} \Phi(T^j x) - \alpha \right| < \epsilon.$$

Let $f(\alpha, n, \epsilon)$ be the cardinality of $F(\alpha, n, \epsilon)$.

We shall use $S_n(\Phi, x)$ to denote the partial sum $\sum_{j=0}^{n-1} \Phi(T^j x)$ and use $A_n(\Phi, x)$ to denote the average $n^{-1} S_n(\Phi, x)$. We shall write $V_n(\Phi) = \sum_{j=1}^n \mathrm{var}_j(\Phi)$ where $\mathrm{var}_j(\Phi) = \sup_{x|_n = y|_n} |\Phi(x) - \Phi(y)|$ ($|\cdot|$ denoting the Euclidean norm).

PROPOSITION 5. *For $\alpha \in L_\Phi$, we have*

$$\lim_{\epsilon\to 0}\liminf_{n\to\infty}\frac{\log f(\alpha,n,\epsilon)}{\log m^n} = \lim_{\epsilon\to 0}\limsup_{n\to\infty}\frac{\log f(\alpha,n,\epsilon)}{\log m^n}\,(=:\Lambda_\Phi(\alpha)).$$

The function $\Lambda_\Phi : L_\Phi \longrightarrow [0,1]$ is concave and upper semi-continuous.

Proof. Without loss of generality, assume that $|\Phi(x)| \leqslant 1$ (for all x) where $|\cdot|$ denotes the Euclidean norm. We want to show that $\log f(\alpha,n,\epsilon)$, as a sequence of n, has a kind of subadditivity. More precisely, for any $\epsilon > 0$, there is an N such that

$$[f(\alpha,n,\epsilon)]^p \leqslant f(\alpha,(n+M)p,2\epsilon) \quad (\forall\, n \geqslant N, \forall\, p \geqslant 1).$$

In fact, suppose that $\{\omega_1,\ldots,\omega_p\} \subset F(\alpha,n,\epsilon)$. Let $\omega = \bar{\omega}_1\ldots\bar{\omega}_p$ where $\bar{\omega}_k$ is an extension of ω_k joining the leading alphabet of ω_{k+1} (with convention $\omega_{p+1}=\omega_1$). Let $x_k \in [\omega_k]$ $(1 \leqslant k \leqslant p)$ be a point such that

$$\left|\sum_{j=0}^{n-1}\Phi(T^j x_k) - n\alpha\right| < n\epsilon.$$

Let x be a point in $[\omega]$. We have

$$S_{(n+M)p}(\Phi,x) - (n+M)p\alpha = \sum_{k=0}^{p-1}\sum_{j=0}^{n+M-1}(\Phi(T^j z_k) - \alpha)$$

where z_k is a point in Σ_A with $z_k|_{n+M} = \bar{\omega}_k$ so that $z_k \in [\bar{\omega}_k] \subset [\omega_k]$. The above inner sum over j is bounded by

$$\left|\sum_{j=0}^{n+M-1}(\Phi(T^j x_k)-\alpha)\right| + \sum_{j=0}^{n+M-1}|\Phi(T^j z_k) - \Phi(T^j x_k)| \leqslant n\epsilon + 2M + V_n(\Phi) + 2M.$$

It follows that

$$|A_{n+M}(\Phi,x) - \alpha| \leqslant \frac{n}{n+M}\epsilon + \frac{4M}{n+M} + \frac{V_n(\Phi)}{n+M}.$$

Since Φ is continuous, $n^{-1}V_n(\Phi)$ tends to zero as $n\to\infty$. Therefore $|A_{n+M}(\Phi,x)-\alpha| \leqslant 2\epsilon$ for sufficiently large $n \geqslant N$ and for all $p \geqslant 1$. Then $[\omega]$, which contains x, is in $F(\alpha,(n+M)p,2\epsilon)$. Notice that different choices $\{\omega_1,\ldots,\omega_p\}$ give rise to different ω. Thus we get the desired subadditivity. By using this subadditivity, it is easy to get

$$\limsup_{n\to\infty}\frac{\log f(\alpha,n,\epsilon)}{\log m^n} \leqslant \liminf_{n\to\infty}\frac{\log f(\alpha,n,2\epsilon)}{\log m^n}$$

from which the equality of the two limits follows.

It is evident that $0 \leqslant \Lambda_\Phi(\alpha) \leqslant 1$. Let $\alpha,\beta \in L_\Phi$. Let p, q be two positive integers. By subadditivity, for large n we have

$$[f(\alpha,n,\epsilon)^p[f(\beta,n,\epsilon)]^q \leqslant f(\alpha,(n+M)p,2\epsilon)f(\beta,(n+M)q,2\epsilon).$$

Let $u \in F(\alpha,(n+M)p,2\epsilon)$ and $v \in F(\beta,(n+M)q,2\epsilon)$. Take a point $x \in [uwv]$ where $w \in \Sigma_{A,M-1}$ such that uw is an extension of u joining the leading alphabet of v. As above, we can get

$$\begin{aligned}|S_{(p+q)(n+M)+M}&(\Phi,x) - (n+M)p\alpha - (n+M)q\beta| \\ &\leqslant 2\epsilon(n+M)(p+q) + V_{(n+M)p}(\Phi) + V_{(n+M)q}(\Phi) + M.\end{aligned}$$

It follows that if n is sufficiently large, $uwv \in F((p\alpha+q\beta)/(p+q),\ (n+M)(p+q)+M, 3\epsilon)$. Consequently, for large n we have

$$f(\alpha, np, 2\epsilon) f(\beta, nq, 2\epsilon) \leqslant f\left(\frac{p\alpha+q\beta}{p+q}, n(p+q), 3\epsilon\right).$$

By the equality of the two limits that we have already proved, we can get

$$\frac{p}{p+q}\Lambda_\Phi(\alpha)+\frac{q}{p+q}\Lambda_\Phi(\beta) \leqslant \Lambda_\Phi\left(\frac{p}{p+q}\alpha+\frac{q}{p+q}\beta\right).$$

This gives the rational concavity of the (bounded) function Λ_Φ. However, the concavity of Λ_Φ is a consequence of its rational concavity and its upper semi-continuity that we prove below.

Given $\alpha \in L_\Phi$. For any $\eta > 0$, there is $\epsilon > 0$ such that

$$\liminf_{n\to\infty} \frac{\log f(\alpha, n, \epsilon)}{\log m^n} < \Lambda_\Phi(\alpha)+\eta.$$

As above, it can be proved that for $\beta \in L_\Phi$ with $|\beta-\alpha| < \epsilon/3$ we have

$$F(\beta, n, \epsilon/3) \subset F(\alpha, n, \epsilon)$$

when n is sufficiently large. It follows that $f(\beta, n, \epsilon/3) \leqslant f(\alpha, n, \epsilon)$. Therefore

$$\Lambda_\Phi(\beta) \leqslant \liminf_{n\to\infty}\frac{\log f(\beta, n, \epsilon/3)}{\log m^n} \leqslant \liminf_{n\to\infty}\frac{\log f(\alpha, n, \epsilon)}{\log m^n} \leqslant \Lambda_\Phi(\alpha)+\eta.$$

This establishes the upper semi-continuity of Λ_Φ at α. □

PROPOSITION 6. *For $\alpha \in L_\Phi$, we have*

$$\dim_H E_\Phi(\alpha) = \dim_P E_\Phi(\alpha) = \Lambda_\Phi(\alpha).$$

Proof. *Step* 1: For $\alpha \in L_\Phi$, we have $\dim_P E_\Phi(\alpha) \leqslant \Lambda_\Phi(\alpha)$.
Let

$$G(\alpha, k, \epsilon) = \bigcap_{n=k}^{\infty} \{x \in \Sigma : |A_n(\Phi, x)-\alpha| < \epsilon\}.$$

It is clear that for any $\epsilon > 0$,

$$E_\Phi(\alpha) \subset \bigcup_{k=1}^{\infty} G(\alpha, k, \epsilon).$$

Notice that if $n \geqslant k$, $G(\alpha, k, \epsilon)$ is covered by the union of all cylinders $[\omega]$ with $\omega \in F(\alpha, n, \epsilon)$ whose total number is $f(\alpha, n, \epsilon)$. Therefore we have the following estimate

$$\overline{\dim}_B G(\alpha, k, \epsilon) \leqslant \limsup_{n\to\infty}\frac{\log f(\alpha, n, \epsilon)}{\log m^n} \quad (\forall \epsilon > 0, \forall k \geqslant 1).$$

On the other hand, by using the σ-stability of the packing dimension, we have

$$\dim_P E_\Phi(\alpha) \leqslant \dim_P\left(\bigcup_{k=1}^{\infty} G(\alpha, k, \epsilon)\right) \leqslant \sup_k \dim_P G(\alpha, k, \epsilon) \leqslant \sup_k \overline{\dim}_B G(\alpha, k, \epsilon).$$

This, together with Proposition 5, leads to the desired result.

Step 2: For $\alpha\in L_\Phi$, we have $\dim_{\mathrm{H}} E_\Phi(\alpha) \geqslant \Lambda_\Phi(\alpha)$.

Given $\delta > 0$. By Proposition 5, there are $\ell_j \uparrow \infty$ and $\epsilon_j \downarrow 0$ such that

$$f(\alpha, \ell_j, \epsilon_j) > m^{\ell_j(\Lambda_\Phi(\alpha)-\delta/2)}.$$

Write simply $F_{\ell_j} = F(\alpha, \ell_j, \epsilon_j)$ and $f_{\ell_j} = f(\alpha, \ell_j, \epsilon_j)$. Define a new sequence $\{\ell_j^*\}$ in the following manner

$$\underbrace{\ell_1, \ldots, \ell_1}_{N_1};\ \underbrace{\ell_2, \ldots, \ell_2; \ldots;}_{N_2}\ \underbrace{\ell_j, \ldots, \ell_j}_{N_j}; \ldots$$

where N_j is defined recursively by

$$N_j = 2^{\ell_{j+1}+N_{j-1}}\ \ (j \geqslant 2); \quad N_1 = 1.$$

Denote $n_j = f_{\ell_j^*}$ and $c_j = m^{-\ell_j^*}$. Define

$$\Theta^* = \prod_{j=1}^{\infty} F_{\ell_j^*}.$$

Observe that Θ^* is a homogeneous Moran set in Σ. More precisely Θ^* is constructed as follows. At level 0, we have only the initial cylinder Σ. In step j, cut a cylinder of level $j-1$ into $m^{\ell_j^*}$ cylinders and pick up n_j ones. By Proposition 4, we have

$$\begin{aligned}
\dim_{\mathrm{H}} \Theta^* &\geqslant \liminf_{k\to\infty} \frac{\log(n_1 \ldots n_k)}{-\log(c_1 \ldots c_k\, c_{k+1}\, n_{k+1})} \\
&\geqslant \liminf_{k\to\infty} \frac{\log(f_{\ell_1^*} \ldots f_{\ell_k^*})}{\log(2^{\ell_1^*+\cdots+\ell_k^*+\ell_{k+1}^*})} \\
&= \liminf_{k\to\infty} \frac{\log(f_{\ell_1^*} \ldots f_{\ell_k^*})}{\log(2^{\ell_1^*+\cdots+\ell_k^*})} \\
&\geqslant \Lambda_\Phi(\alpha) - \delta.
\end{aligned}$$

However Θ^* is a set in Σ, not necessarily in Σ_A. Based on Θ^*, we are going to construct a set Θ^{**} which is not only in Σ_A but also in $E_\Phi(\alpha)$, and which is of the same Hausdorff dimension as Θ^*.

We extend words in $F_{\ell_j^*}$ in the following manner. Let

$$A = \{a : x = x|_{\ell_j^*-1}\, a, x \in F_{\ell_j^*}\}, \quad B = \{b = x|_1, x \in F_{\ell_{j+1}^*}\}.$$

That is, A is the set of the last alphabets of the words in $F_{\ell_j^*}$ and B is the set of the first alphabets of the words in $F_{\ell_{j+1}^*}$. For each pair $(a, b)\in A\times B$, take one (and only one) word $w\in\Sigma_{A,M}$ such that awb is admissible. We call w a bridge word. For $y = u_1 u_2 \ldots \in \Theta^*$, define

$$y^* = u_1 w_1 u_2 w_2 \ldots u_j w_j u_{j+1} w_{j+1} \cdots$$

where w_j is the bridge word as defined above. We define Θ^{**} as the set of sequences y^*. By considering the map sending y to y^* which is nearly bi-Lipschitz, we have $\dim_{\mathrm{H}} \Theta^{**} = \dim_{\mathrm{H}} \Theta^*$. On the other hand, $\Theta^{**} \subset E_\Phi(\alpha)$. In fact, let $\ell_j^{**} = \ell_j^* + M$. For large n $(> \ell_1)$, there is a unique integer $J(n)$ such that

$$\sum_{i=1}^{j(n)} \ell_i^{**} \leqslant n < \sum_{i=1}^{J(n)+1} \ell_i^{**}.$$

The choice of N_j implies that $\ell_{k+1} = o(N_k)$. It follows that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{J(n)} \ell_i^{**}}{\sum_{i=1}^{J(n)+1} \ell_i^{**}} = 1.$$

Let $x = (x_n) \in \Theta^{**}$. Let Λ_j be the set of integers between $\ell_1^{**} + \ldots + \ell_{j-1}^{**} + 1$ and $\ell_1^{**} + \ldots + \ell_j^{**}$. We have

$$\begin{aligned} S_n(\Phi, x) &= \sum_{j=1}^{J(n)} \sum_{k\in\Lambda_j} \Phi(T^k x) + O(\ell_{J(n)+1}^{**}) \\ &= \sum_{j=1}^{J(n)} \ell_j^{**}(\alpha + o(1)) + o(n) \\ &= \alpha n + o(n). \end{aligned}$$

It follows that $x \in E_\Phi(\alpha)$. $\square$

We now prove the lower estimate

$$\dim_{\mathrm{H}} E_\Phi(\alpha) \geqslant \sup_{\mu\in\mathscr{F}_\Phi(\alpha)} \frac{h_\mu}{\log m}.$$

Since the entropy h_μ is upper semi-continuous as a function of μ (see Proposition 1) and $\mathscr{F}_\Phi(\alpha)$ is compact, the above supremum is attained. Let μ_0 be a measure in $\mathscr{F}_\Phi(\alpha)$ which attains the supremum. Take an invariant measure ν with support Σ_A. For any $\epsilon > 0$, consider $\mu_\epsilon = (1-\epsilon)\mu_0 + \epsilon h_\nu$ as an approximation of μ_0. Let $\alpha_\epsilon = \int \Phi d\mu_\epsilon$. We have

$$|\alpha_\epsilon - \alpha| \leqslant 2\epsilon \|\Phi\|$$

where $\|\Phi\| = \sup_x |\Phi(x)|$. Since the support of the invariant measure μ_ϵ is the whole space Σ_A, we can find a sequence of ergodic Markov measures $\mu_\epsilon^{(k)}$ such that $\mu_\epsilon^{(k)}$ tends to μ_ϵ in the weak* topology and $h_{\mu_\epsilon^{(k)}}$ tends to h_{μ_ϵ} as $k \to \infty$ (see Propositions 2 and 3). Let $\alpha_\epsilon^{(k)} = \int \Phi d\mu_\epsilon^{(k)}$. We have

$$|\alpha_\epsilon^{(k)} - \alpha_\epsilon| \leqslant \epsilon \ \ (k \geqslant k(\epsilon)).$$

Since $\Lambda_\Phi(\cdot)$ is upper semi-continuous (Proposition 5), for any η, when ϵ is sufficiently small and k is sufficiently large, we have

$$\Lambda_\Phi(\alpha) \geqslant \Lambda_\Phi(\alpha_\epsilon^{(k)}) - \eta \geqslant \frac{h_{\mu_\epsilon^{(k)}}}{\log m} - \eta$$

where for the second inequality we used the fact that $\Lambda_\Phi(\alpha) \geqslant h_\mu/\log m$ for any ergodic measure $\mu \in \mathscr{F}_\Phi(\alpha)$. In fact, the ergodicity implies that $\mu(E_\Phi(\alpha)) = 1$. It follows that $\dim_{\mathrm{H}} \mu \leqslant \dim_{\mathrm{H}} E_\Phi(\alpha)$. However $\dim \mu = h_\mu/\log m$ by the Shannon–McMillan–Breiman theorem [**19**, p. 261] (see [**9**] for $\dim_{\mathrm{H}} \mu$). Letting $k \to \infty$ gives

$$\Lambda_\Phi(\alpha) \geqslant \frac{h_{\mu_\epsilon}}{\log m} - \eta = \frac{(1-\epsilon)h_{\mu_0} + \epsilon h_\nu}{\log m} - \eta.$$

Now let $\epsilon \to 0$ and then $\eta \to 0$. We get $\Lambda_\Phi(\alpha) \geqslant h_{\mu_0}/\log m$.

We now prove the upper estimate

$$\dim_{\mathrm{P}} E_\Phi(\alpha) \leqslant \sup_{\mu\in\mathscr{F}_\Phi(\alpha)} \frac{h_\mu}{\log m}.$$

Given any $\epsilon > 0$. If $k \geqslant 1$ is sufficient large, we can find a function Φ_k depending only on the first k coordinates of $x \in \Sigma_A$ such that

$$|\Phi_k(x) - \Phi(x)| < \frac{\epsilon}{2} \quad (\forall\, x).$$

Thus, for $x \in E_\Phi(\alpha)$ we have

$$\limsup_{n\to\infty} \left| \frac{1}{n} S_n(\Phi_k, x) - \alpha \right| < \epsilon.$$

It follows that

$$E_\Phi(\alpha) \subset \bigcup_{j=1}^{\infty} \bigcap_{n=j}^{\infty} B(n, k, \epsilon) \ \ (\forall\, k = k(\epsilon) \text{ large and } \epsilon > 0)$$

where

$$B(n, k, \epsilon) = \{x \in \Sigma : |n^{-1} S_n(\Phi_k, x) - \alpha| < \epsilon\}.$$

In the following, we will show that for each $j \in \mathbf{N}$

$$\overline{\dim}_{\mathrm{B}} \bigcap_{n=j}^{\infty} B(n, k, \epsilon) \leqslant \sup_{|\Phi_*(\mu) - \alpha| < 3\epsilon} \frac{h_\mu}{\log m}.$$

Take all cylinders of the form $[x_1, \ldots, x_{n+k-1}]$ which intersect $B(n, k, \epsilon)$. They constitute a cover of $(n+k-1)$-cylinders of $B(n, k, \epsilon)$. Denote by $T(n, k, \epsilon)$ the number of cylinders contained in this cover. We are going to estimate this number.

For $\omega \in \Sigma_{A,\, n+k-1}$, denote by $N_\omega(\epsilon_1, \ldots, \epsilon_k)$ the number of words of the form $\epsilon_1 \ldots \epsilon_k$ appearing in ω. The family $\{N_\omega(\epsilon_1, \ldots, \epsilon_k)\}$ with varying $\epsilon_1, \ldots, \epsilon_k$ will be called the k-distribution of ω. Let $D = \{n(\epsilon_1, \ldots, \epsilon_k)\}$ be the k-distribution of some ω such that $[\omega]$ intersects $B(n, k, \epsilon)$. This means that

$$n(\epsilon_1, \ldots, \epsilon_k) = N_\omega(\epsilon_1, \ldots, \epsilon_k)$$

$$\left| \frac{1}{n} \sum_{\epsilon_1, \ldots, \epsilon_k} n(\epsilon_1, \ldots, \epsilon_k)\, \Phi_k(\epsilon_1, \ldots, \epsilon_k) - \alpha \right| < \epsilon.$$

We denote by $\mathscr{T}(D)$ the collection of all such cylinders $[\omega]$ and by $\Gamma(D)$ its cardinality. Let $\mathscr{P}_k$ be the set of all possible such distributions D. The cover is then decomposed into disjoint families $\mathscr{T}(D)$ with $D \in \mathscr{P}_k$. It is clear that the cardinality of $\mathscr{P}_k$, the number of families, is at most n^{m^k}. Then we have

$$T(n, k, \epsilon) = \sum_{D \in \mathscr{P}_k} \Gamma(D) \leqslant n^{m^k} \max_{D \in \mathscr{P}_k} \Gamma(D).$$

Thus

$$\frac{\log T(n, k, \epsilon)}{\log m^n} \leqslant \max_{D \in \mathscr{P}_k} \frac{\log \Gamma(D)}{\log m^n} + O\left(\frac{\log n}{n} \right).$$

(The constant in 'O' depends upon k.)

Since $n(\epsilon_1, \ldots, \epsilon_k) = N_\omega(\epsilon_1, \ldots, \epsilon_k)$, $\{n(\epsilon_1, \ldots, \epsilon_k)/n\}$ defines a probability vector. For any $\eta > 0$, there is a number $N = N(\eta)$ such that when $n > N$ we can find $p \in \Delta_k^+$ such that (see [**10**, Lemma 3])

$$\left| \frac{n(\epsilon_1, \ldots, \epsilon_k)}{n} - p(\epsilon_1, \ldots, \epsilon_k) \right| < \eta, \quad p(\epsilon_1, \ldots, \epsilon_k) > \frac{\eta}{m^{k+1}}.$$

Consider the Markov measure v_p corresponding to p. For any cylinder $[\omega]\in\mathscr{T}(D)$ with $\omega=(x_i)_{i=1}^{n+k-1}$, we have $N_\omega(\epsilon_1,\ldots,\epsilon_k)=n(\epsilon_1,\ldots,\epsilon_k)$, so we have

$$v_p([\omega]=\frac{p(x_1,\ldots,x_k)}{t(x_1,\ldots,x_k)}\prod_{\epsilon_1,\ldots,\epsilon_k}t(\epsilon_1,\ldots,\epsilon_k)^{n(\epsilon_1,\ldots,\epsilon_k)}$$

$$\geqslant\frac{\eta}{m^{k+1}}\prod_{\epsilon_1,\ldots,\epsilon_k}t(\epsilon_1,\ldots,\epsilon_k)^{n(\epsilon_1,\ldots,\epsilon_k)}$$

where

$$t(a_1,\ldots,a_k)=\frac{p(a_1,\ldots,a_k)}{\sum_\epsilon p(a_1,\ldots,a_{k-1},\epsilon)}.$$

Let $a=a(D)$ denote the right-hand side of the above inequality. Then

$$a\Gamma(D)\leqslant v_p\Bigg(\bigcup_{[\omega]\in\mathscr{T}(D)}[\omega]\Bigg)\leqslant 1.$$

Combining the last two expressions gives

$$\Gamma(D)\leqslant\frac{1}{a}\leqslant\frac{m^{k+1}}{\eta}\prod_{\epsilon_1,\ldots,\epsilon_k}t(\epsilon_1,\ldots,\epsilon_k)^{-n(\epsilon_1,\ldots,\epsilon_k)}.$$

Then

$$\frac{\log\Gamma(D)}{\log m^n}\leqslant O\left(\frac{|\log\eta|}{n}\right)-\sum_{\epsilon_1,\ldots,\epsilon_k}\frac{n(\epsilon_1,\ldots,\epsilon_k)}{n}\log_m t(\epsilon_1,\ldots,\epsilon_k)$$

$$\leqslant O\left(\frac{|\log\eta|}{n}\right)-\sum_{\epsilon_1,\ldots,\epsilon_k}p(\epsilon_1,\ldots,\epsilon_k)\log_m t(\epsilon_1,\ldots,\epsilon_k)+O(\eta|\log\eta|).$$

Note that $\Phi_*(v_p)$ is near α in the sense that

$$|\Phi_*(v_p)-\alpha|\leqslant\left|\sum_{\epsilon_1,\ldots,\epsilon_k}p(\epsilon_1,\ldots,\epsilon_k)\,\Phi_k(\epsilon_1,\ldots,\epsilon_k)-\alpha\right|+\epsilon$$

$$\leqslant\left|\sum_{\epsilon_1,\ldots,\epsilon_k}\frac{n(\epsilon_1,\ldots,\epsilon_k)}{n}\Phi_k(\epsilon_1,\ldots,\epsilon_k)-\alpha\right|+m^k\eta\|\Phi_k\|+\epsilon$$

$$\leqslant m^k\eta\|\Phi_k\|+2\epsilon.$$

Now we can conclude that

$$\frac{\log\Gamma(D)}{\log m^n}\leqslant O\left(\frac{|\log\eta|}{n}\right)+O(\eta|\log\eta|)+\sup_{\mu\in\Delta_k,\,|\Phi_*(\mu)-\alpha|<m^k\eta\|\Phi_k\|+2\epsilon}\frac{h_\mu}{\log m}.$$

Since the right-hand side is independent of D, we get

$$\frac{\log T(n,k,\epsilon)}{\log m^n}\leqslant O\left(\frac{|\log\eta|+\log n}{n}\right)+O(\eta|\log\eta|)+\sup_{|\Phi_*(\mu)-\alpha|<m^k\eta\|\Phi_k\|+2\epsilon}\frac{h_\mu}{\log m}.$$

Let $n\to\infty$ then let $\eta\to 0$, we get

$$\overline{\dim}_{\mathrm{B}}\bigcap_{n=j}^{\infty}B(n,k,\epsilon)\leqslant\limsup_{n\to\infty}\frac{\log T(n,k,\epsilon)}{\log m^n}\leqslant\sup_{|\Phi_*(\mu)-\alpha|\leqslant 3\epsilon}\frac{h_\mu}{\log m}.$$

By the σ-stability of the packing dimension and the inequality $\dim_{\mathrm{P}} \leqslant \overline{\dim}_{\mathrm{B}}$ (see [17]), we have

$$\dim_{\mathrm{P}}(E_\Phi(\alpha)) \leqslant \dim_{\mathrm{P}}\left(\bigcup_{j=1}^{\infty}\bigcap_{n=j}^{\infty} B(n,k,\epsilon)\right)$$
$$\leqslant \sup_j \overline{\dim}_{\mathrm{B}} \bigcap_{n=j}^{\infty} B(n,k,\epsilon) \leqslant \sup_{|\Phi_*(\mu)-\alpha|\leqslant 3\epsilon} \frac{h_\mu}{\log m}.$$

Note that the set of invariant measures μ such that $|\Phi_*(\mu)-\alpha| \leqslant 3\epsilon$ is compact and that h_μ is upper semi-continuous. For any $\ell \geqslant 1$, there is an invariant measure μ_ℓ such that

$$h_{\mu_\ell} = \sup_{|\Phi_*(\mu)-\alpha|\leqslant 1/\ell} h_\mu.$$

We can also assume that μ_ℓ converges to some measure μ_∞. It is clear that $\mu_\infty \in \mathscr{F}_\Phi(\alpha)$. Thus, by the upper semi-continuity of the entropy,

$$\dim_{\mathrm{P}} E_\Phi(\alpha) \leqslant \liminf_{\epsilon\to 0} \sup_{|\Phi_*(\mu)-\alpha|\leqslant 3\epsilon} \frac{h_\mu}{\log m} = \frac{h_{\mu_\infty}}{\log m}. \qquad \square$$

4. *Proof of Theorem* B

As in the proof of Theorem A, we will use Hausdorff dimension instead of topological entropy.

We first prove the fact that if the limit defining $\sigma_\Phi(x)$ does not exist for some point x, it does not exist for points in a set of Hausdorff dimension equal to that of Σ_A. Such points are called *divergent points*.

We start with a method to construct new divergent points by using a given divergent point. Take a subset of words $W^* \subset \Sigma_{A,M-1}$ such that for any $t,s \in \{0,1,\ldots, m-1\}$ there is one and only one word $w \in W^*$ such that $tws \in \Sigma_{A,M+1}$ (maybe there are several choices for a given pair (t,s), but we choose only one). Assume that Φ takes real values. Without loss of generality, we assume that $|\Phi(x)| \leqslant 1$ (for all $x \in \Sigma_A$). Denote by $\underline{\sigma}_\Phi(x)$ and $\overline{\sigma}_\Phi(x)$ the $\liminf$ and $\limsup$ of $n^{-1}\sum_{j=0}^{n-1}\Phi(T^jx)$.

Suppose there are two points $u,v \in \Sigma_A$ (maybe $u=v$) such that $\overline{\sigma}_\Phi(u) > \underline{\sigma}_\Phi(v)$. Take two numbers a,b such that

$$\overline{\sigma}_\Phi(u) > a > b > \underline{\sigma}_\Phi(v).$$

For any $k \geqslant 1$, there exists $\ell_k \geqslant 1$ such that

$$|\Phi(x)-\Phi(y)| \leqslant \frac{a-b}{k} \text{ if } x_i = y_i \quad (1 \leqslant i \leqslant \ell_k).$$

Take a rapidly increasing sequence $\{n_i\}_{i\geqslant 1}$ satisfying

$$\frac{1}{n_{2k-1}}\sum_{j=0}^{n_{2k-1}-1}\Phi(T^j(u)) > a, \quad \frac{1}{n_{2k}}\sum_{j=0}^{n_{2k}-1}\Phi(T^j(v)) < b,$$
$$k = o(n_k), \quad \ell_k = o(n_k), \quad 2(M-1)k \leqslant \sum_{j=1}^{k} n_j, \quad \sum_{j=1}^{k-1} n_j = o(n_k).$$

Fix an integer $q \geqslant 1$. Let $x = x_1x_2\ldots \in \Sigma_A$. We cut x into words of lengths $\{qn_k\}$ where $\{n_k\}$ is as above. This means that

$$x = \bar{x}_1\bar{x}_2\ldots\bar{x}_k\ldots \text{ with } |\bar{x}_k| = qn_k.$$

Define $\bar{u}_k$ to be the prefix of u with $|\bar{u}_k| = n_{2k-1}$ and $\bar{v}_k$ to be the prefix of v with $|\bar{v}_k| = n_{2k}$. Denote

$$B_{2k-1} = \bar{u}_k, \quad B_{2k} = \bar{v}_k.$$

Now construct a new point $x^* \in \Sigma_A$ as follows

$$x^* = \bar{x}_1 W_1' B_1 W_1'' \bar{x}_2 W_2' B_2 W_2'' \dots \bar{x}_k W_k' B_k W_k'' \dots$$

where W_k' and W_k'' are admissible words in W^*.

PROPOSITION 7. *Suppose that $\bar{\sigma}_\Phi(u) > \underline{\sigma}_\Phi(v)$ for some $u, v \in \Sigma_A$ (maybe $u = v$). If $\underline{\sigma}_\Phi(x) = \bar{\sigma}_\Phi(x)$, then $\bar{\sigma}_\Phi(x^*) > \underline{\sigma}_\Phi(x^*)$ where x^* is constructed as above from x.*

Proof. Notice that the word $\bar{x}_k W_k' B_k W_k''$ is of length

$$L_k = qn_k + (M-1) + n_k + (M-1) = (q+1)\, n_k + 2(M-1).$$

Let $N_k = \sum_{j=1}^{k} L_j = (q+1) \sum_{j=1}^{k} n_j + 2k(M-1)$ and let $M_k = q \sum_{j=1}^{k} n_j$. Consider the partial sum

$$\sum_{j=0}^{N_{2k}-1} \Phi(T^j x^*) = \sum_{j=0}^{N_{2k-1}-1} \Phi(T^j x^*) + \sum_{j=N_{2k-1}}^{N_{2k}-1} \Phi(T^j x^*).$$

The first sum on the right is bounded by N_{2k-1} (since 1 is an upper bound of Φ). For the second sum, we have

$$\begin{aligned}
\sum_{j=N_{2k-1}}^{N_{2k}-1} \Phi(T^j x^*) &= \sum_{j=0}^{L_{2k}-1} \Phi(T^j(\bar{x}_{2k} W_{2k}' B_{2k} W_{2k}'' \dots)) \\
&\leqslant \sum_{j=M_{2k-1}}^{M_{2k}-\ell_k-1} \Phi(T^j x) + (qn_{2k} - \ell_k) \cdot \frac{a-b}{k} + (\ell_k + M - 1) \\
&\quad + \sum_{j=0}^{n_{2k}-\ell_k-1} \Phi(T^j v) + (n_{2k} - \ell_k) \cdot \frac{a-b}{k} + (\ell_k + M - 1) \\
&= \sum_{j=M_{2k-1}}^{M_{2k}-\ell_k-1} \Phi(T^j x) + \sum_{j=0}^{n_{2k}-\ell_k-1} \Phi(T^j v) \\
&\quad + \frac{(a-b)\,((q+1)\, n_{2k} - 2\ell_k)}{k} + 2(\ell_k + M - 1) \\
&=: A_k + B_k + C_k + D_k
\end{aligned}$$

where the inequality is obtained by cutting the sum into four sums: the first one is taken over the first $qn_{2k} - \ell_k$ terms whose sum is bounded by $\sum_{j=0}^{qn_{2k}-\ell_k+1} \Phi(T^j x) + (qn_{2k} - \ell_k)(a-b)/k$; the second one is taken over the next $\ell_k + M - 1$ terms each of which is controlled by 1; the third sum taken over the next $n_{2k} - \ell_k$ terms can be written as

$$\sum_{j=0}^{n_{2k}-\ell_k-1} \Phi(T^j(B_{2k} W_{2k}'' \dots))$$

which is bounded by $\sum_{j=0}^{n_{2k}-\ell_k-1} \Phi(T^j v) + (n_{2k} - \ell_k) \cdot (a-b)/k$; the last sum can be estimated like the second sum. Thus we get

$$\frac{1}{N_{2k}} \sum_{j=0}^{N_{2k}-1} \Phi(T^j x^*) \leqslant \frac{N_{2k-1}}{N_{2k}} + \frac{A_k}{N_{2k}} + \frac{B_k}{N_{2k}} + \frac{C_k}{N_{2k}} + \frac{D_k}{N_{2k}}.$$

By the properties of n_k, the quantities N_{2k-1}, C_k and D_k are all $o(N_{2k})$ and

$$\lim_{k\to\infty}\frac{A_k}{N_{2k}}=\frac{q}{q+1}\sigma_\Phi(x),\quad \liminf_{k\to\infty}\frac{B_k}{N_{2k}}\leqslant\frac{b}{q+1}.$$

It follows that $\underline{\sigma}_\Phi(x^*)\leqslant(q/(q+1))\sigma_\Phi(x)+b/(q+1)$. In the same way, we can prove that $\bar{\sigma}_\Phi(x^*)\geqslant(q/(q+1))\sigma_\Phi(x)+a/(q+1)$, so that $\bar{\sigma}_\Phi(x^*)-\underline{\sigma}_\Phi(x^*)>(a-b)/(q+1)>0$. □

Let G_Φ be the set of (good) points x such that $\underline{\sigma}_\Phi(x)=\bar{\sigma}_\Phi(x)$. When points u and v, the corresponding sequence $\{n_k\}$ and the number $q\geqslant 1$ are fixed as above, to any point $x\in G_\Phi$ we have associated a (unique) point x^* in Σ_A (the uniqueness comes from the unique choice in the definition of W^*). If we have a subset E of G_Φ, in this way we get a (bad) subset E_q of Σ_A.

PROPOSITION 8. *Keep the same notation as above. For any $E\subset G_\Phi$, we have*

$$\dim_{\mathrm{H}} E_q\geqslant\frac{q}{q+2}\dim_{\mathrm{H}} E.$$

Proof. Consider the bijective map $f\colon E\longrightarrow E_q$ defined by $f(x)=x^*$. It suffices to show that f^{-1} is $q/(q+2)$-Hölder [**15**, p. 139], or equivalently

$$\rho(f(x),f(y))\geqslant\rho(x,y)^{1+2/q}.$$

In fact, suppose that $\rho(x,y)=m^{-r}$.

(a) If $r\leqslant qn_1$, we have $\rho(f(x),f(y))=\rho(x,y)$.

(b) If $q\sum_{i=1}^{2k-1}n_i<r\leqslant q\sum_{i=1}^{2k}n_i$ for some $k\geqslant 1$, we have

$$\rho(f(x),f(y))=m^{-r-\sum_{i=1}^{2k-1}n_i-2(M-1)(2k-1)}\geqslant m^{-r-(r/q)-(r/q)}=m^{-r(1+(2/q))}.$$

(c) If $q\sum_{i=1}^{2k}n_i<r\leqslant q\sum_{i=1}^{2k+1}n_i$ for some $k\geqslant 1$, we also have

$$\rho(f(x),f(y))\geqslant m^{-r(1+(2/q))}.$$ □

To complete the proof of Theorem B we use the fact that $\dim_{\mathrm{H}} G_\Phi=\dim_{\mathrm{H}}\Sigma_A$. To see this, it suffices to notice that the measure of maximal entropy of (Σ_A,T) is the Parry measure and that G_Φ has full Parry measure [**22**, p. 194]. □

5. *Remarks*

(1) Let $C(\Sigma_A)$ be the set of real-valued continuous functions. Two functions $f_1,f_1\in C(\Sigma_A)$ are said to be *cohomologous* if $f_1-f_2=\psi\circ T-\psi+c$ for some $\psi\in C(\Sigma_A)$ and some constant c. This is an equivalence relation. Denote by $\tilde{C}(\Sigma_A)$ the quotient space relative to this relation. Suppose that $\Phi=(\Phi_1,\ldots,\Phi_d)$ is of summable variation. It may be proved that the ergodic limit $\sigma_\Phi(x)$ is a constant function on Σ_A if and only if Φ_j considered as function in the quotient space are of rank zero. More generally, it may also be proved that the dimension of L_Φ is the rank of $\{\Phi_j\}_{1\leqslant j\leqslant d}$ considered as elements in the quotient space.

(2) According to Theorem A, the dimension $\dim E_\Phi(\alpha)$ (or equivalently the topological entropy $h_{\mathrm{top}}(E_\Phi(\alpha))$) is concave and varies upper semi-continuously on L_Φ, then is continuous in the interior of L_Φ [**20**, Section 10]. We wonder if there is always

continuity on the boundary of L_Φ. The thermodynamical formalism shows differentiability in the interior for a function Φ of summable variation, and analyticity in the interior for a Hölder function Φ [**21**].

(3) When Φ is of summable variation, the variational formula (Theorem A), together with the thermodynamical formalism, implies the relation

$$P^*(\alpha) = -\max_{\mu \in \mathscr{F}_\Phi(\alpha)} h_\mu$$

where $P^*(\alpha)$ is the Legendre transform of the pressure function $P(\beta)$ (that is, the pressure of the potential $\langle \beta, \Phi \rangle$). It follows that

$$P(\beta) = \inf_{\alpha \in L_\Phi} \left[\langle \alpha, \beta \rangle + \max_{\mu \in \mathscr{F}_\Phi(\alpha)} h_\mu \right].$$

Then we might get good information on the pressure function whenever the fibres of projection are known. Therefore the fibres of projection are worthy of study. Some related works are by T. Bousch [**2**], S. Bullett and P. Sentenac [**6**] and O. Jenkinson [**13**, **14**] (see also [**16**]).

(4) Generalization to conformal expanding dynamical systems is easy to guess. Generalization to other systems is worthy of study and Liapunov exponents would be involved.

Acknowledgements. The paper was finished when the first author visited FIM at ETH-Zurich in Switzerland. He would like to thank Professors O. Lanford III and A.-S. Sznitman for their warm hospitality. The authors thank T. Bousch, F. Gamboa, O. Jenkinson, A. K. Manning and Z. Y. Wen for useful information and enlightening discussions. They also thank an anonymous referee for his (or her) suggestions from which the present version has benefited. D. J. Feng was supported by Tsinghua University Funding Program and J. Wu was supported by NSFC Grant #19971064. The second and third authors were supported by the Special Funds for Major State Basic Research Projects of China.

References

1. L. Barreira and J. Schmeling, 'Invariant sets with zero measure and full Hausdorff dimension', *AMS Electronic Research Announcements* 3 (1997) 114–118.
2. T. Bousch, 'Le poisson n'a pas d'arêtes', *Ann. Inst. H. Poincaré*, to appear.
3. R. Bowen, 'Topological entropy for noncompact sets', *Trans. Amer. Math. Soc.* 184 (1973) 125–136.
4. R. Bowen, 'Equilibrium states and the ergodic theory of Anosov diffeomorphisms', Lecture Notes in Mathematics 470 (Springer, 1975).
5. M. Brin and A. Katok, 'On local entropy', *Geometric Dynamics*, *Rio de Janeiro*, 1981, Lecture Notes in Mathematics 1007 (Springer, Berlin, 1983) 30–38.
6. S. Bullett and P. Sentenac, 'Ordered orbits of the shift, square roots, and the devil's staircase', *Math. Proc. Cambridge Philos. Soc.* 115 (1994) 451–481.
7. J. L. Doob, *Stochastic processes* (John Wiley, New York, 1953).
8. K. J. Falconer, *Fractal geometry: mathematical foundation and applications* (John Wiley, New York, 1990).
9. A. H. Fan, 'Sur les dimensions de mesures', *Studia Math.* 111 (1994) 1–17.
10. A. H. Fan and D. J. Feng, 'On the distribution of long-term time average on the symbolic space', *J. Statist. Phys.* 99 (2000) 813–856 (see also 'Analyse multifractale de la récurrence sur l'espace symbolique', *C. R. Acad. Sci. Paris Sér.* I *Math.* (1998) 629–632).
11. D. J. Feng, H. Rao and J. Wu, 'The net measure properties of symmetric Cantor sets and their applications', *Progr. Natur. Sci.* 7 (1997) 172–178.
12. D. J. Feng, Z. Y. Wen and J. Wu, 'Some dimensional results for homogeneous Moran sets', *Sci. China Ser.* A 40 (1997) 475–482.
13. O. Jenkinson, 'Conjugacy rigidity, cohomological triviality and barycentres of invariant measures', PhD Thesis, Warwick University, 1996.

14. O. Jenkinson, 'Geometric barycentres of invariant measures for circle maps', *Ergodic Theory Dynam. Systems*, to appear.

15. J. P. Kahane, *Some series of random functions* (Cambridge University Press, 1985).

16. R. Mañé, 'Generic properties and problems of minimizing measures of Lagrangian systems', *Nonlinearity* 9 (1996) 273–310.

17. P. Mattila, *Geometry of sets and measures in Euclidean spaces, fractals and rectifiability* (Cambridge University Press, 1995).

18. Y. Pesin and B. Pitskel, 'Topological pressure and the variational principle for non-compact sets', *Funct. Anal. Appl.* 18 (1984) 307–318.

19. K. Peterson, *Ergodic theory* (Cambridge University Press, 1983).

20. R. T. Rockafellar, *Convex analysis* (Princeton University Press, 1970).

21. D. Ruelle, 'Thermodynamic formalism: the mathematical structures of classical equilibrium statistical mechanics', *Encyclopedia of Mathematics and its Applications – Vol.* 5 (Addison Wesley, 1978).

22. P. Walters, *An introduction to ergodic theory* (Springer, Berlin, 1982).

Ai-Hua Fan
Département de Mathématiques
Université de Picardie Jules Verne
33 rue Saint Leu
80039 Amiens
France

ai-hua.fan@mathinfo.u-picardie.fr

Jun Wu
Department of Mathematics
Wuhan University
Wuhan 430072
China

wujunyu@public.wh.hb.cn

De-Jun Feng
Department of Applied Mathematics
Tsinghua University
Beijing 100084
China

Center for Advanced Study
Tsinghua University
Beijing
China

dfeng@math.tsinghua.edu.cn