A MULTIFRACTAL MASS TRANSFERENCE PRINCIPLE FOR GIBBS MEASURES WITH APPLICATIONS TO DYNAMICAL DIOPHANTINE APPROXIMATION

AI-HUA FAN, JÖRG SCHMELING, AND SERGE TROUBETZKOY

ABSTRACT. Let μ be a Gibbs measure of the doubling map T of the circle. For a μ-generic point x and a given sequence $\{r_n\} \subset \mathbb{R}^+$, consider the intervals $(T^n x - r_n \pmod 1, T^n x + r_n \pmod 1)$. In analogy to the classical Dvoretzky covering of the circle we study the covering properties of this sequence of intervals. This study is closely related to the local entropy function of the Gibbs measure and to the hitting times for moving targets. A mass transference principle is obtained for Gibbs measures which are multifractal. Such a principle was proved by Beresnevich and Velani [BV] for mono-fractal measures. In the symbolic language we completely describe the combinatorial structure of a typical relatively short sequence, in particular we can describe the occurrence of "atypical" relatively long words. Our results have a direct and deep number-theoretical interpretation via inhomogeneous dyadic Diophantine approximation by numbers belonging to a given (dyadic) Diophantine class.

1. Introduction

Let (X, d) be a complete metric space. Given a sequence $\{x_n\}_{n\geq 1}$ of points in X and a sequence $\{r_n\}_{n\geq 1}$ of positive numbers we define

$$\begin{aligned} I(\{x_n\}, \{r_n\}) : &= \overline{\lim_{n\to\infty}} B(x_n, r_n), \\ F(\{x_n\}, \{r_n\}) : &= X \setminus I(\{x_n\}, \{r_n\}) \end{aligned}$$

where $B(x_n, r_n)$ denotes the ball of center x_n with radius r_n. We propose to study the sizes of the sets $I(\{x_n\}, \{r_n\})$ and $F(\{x_n\}, \{r_n\})$, including their Hausdorff dimensions, when $\{x_n\}$ is an orbit of a dynamical system $T : X \to X$.

In this paper we concentrate our treatment on the doubling map on the circle by keeping the technical details as low as possible. But we should emphasize that it is just a technical question to generalize the results and the method to piecewise expanding maps on the circle or more generally to conformal maps.

Our main tool in this article is an implicit multifractal mass transference principle (MMTP): the dimension of $I(\{x_n\}, \{r_n\})$ yields the dimension of $I(\{x_n\}, \{r_n^\alpha\})$ for a certain interval of parameter α under suitable conditions on the centers x_n and the radii r_n. A mass transference principle (MTP) was developed by Beresnevich and Velani [BV] in order to study the Duffin-Schaeffer problem in the metric number theory (see also [BDV]). A MTP is used by Jaffard [Jaf] in his study of lacunary wavelet series. The MTP states that under certain explicit assumptions on the set in question $I(\{x_n\},\{r_n\})$ one can completely understand the Hausdorff measure theory of the set $I(\{x_n\},\{r_n^\alpha\})$. However in concrete situations it is extremely hard, if not impossible, to verify the assumptions. In contrast we do not state an abstract theorem, but provide a large class of dynamically defined sets that exhibit an MMTP. To be more precise, in the obtained MTP the condition imposed on the centers x_n implies that they are extremely well distributed with respect to Lebesgue measure. Unlike the very mono-fractal Lebesgue measure we consider multifractal measures. The presence of a multifractal structure is the novelty of this paper and represents the main difficulty of proving the results. More precisely we work in a dynamical framework with an invariant Gibbs measure. We suppose that the sequence of centers x_n is the orbit of a typical point x with respect to the Gibbs measure. In particular we can apply our results to the case of Lebesgue measure, which is a Gibbs measure, to recuperate the result of [BV], when $\{x_n\} = \{2^n x\}$ is a Lebesgue typical orbit.

To prove the MMTP for dynamically defined sets, we take a symbolic representation of our dynamical system. The main ingredient is a careful analysis of the first occurrence of finite words in typical sequences. These results are much stronger than those obtainable by large deviation techniques: we show that certain very unlikely events must occur much earlier than large deviations predict, however which unlikely event occurs is completely random. We prove a dichotomy for a certain range of parameters while the unlikely events do occur, they form a negligible part of all events. Then after a phase transition they take over (Theorem 2·1). This theorem implies that we could have, similarly to [BV], posed a deterministic requirement on the centers x_n and the MTP holds for them. Since these conditions, even for Lebesgue measure, are hard to check we prefer stating our MMTP in terms of generic points.

By *Diophantine approximation* we mean the study of the sets $I(\{x_n\},\{r_n\})$ and $F(\{x_n\},\{r_n\})$. This terminology is motivated by fact that classic Diophantine approximation is a special case. Let $X = \mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ be the unit circle equipped with the metric

$$\|x - y\| = \inf_{k\in\mathbb{Z}} |(x-y) - k|.$$

Let $\{x_n\} = \{n\alpha \pmod 1\}$ be the orbit of the irrational rotation determined by an irrational number α. Then $0 \in I(\{x_n\},\{r_n\})$ means $\|\alpha n\| < r_n$ holds for an infinite number of n's. This is nothing but the *homogeneous Diophantine approximation* of α. More generally $y \in I(\{x_n\},\{r_n\})$ means $\|\alpha n - y\| < r_n$ holds for an infinite number of n's. This is what is called inhomogeneous Diophantine approximation. In [FS], based on the results in [ST, Bu], both $I(\{x_n\},\{r_n\})$ and $F(\{x_n\},\{r_n\})$ have been analyzed for an irrational number α when $r_n = n^{-\kappa}$. The case for general sequence $\{r_n\}$ has been studied in [FW2].

In this article instead of studying continued fraction approximations we study dyadic base approximations. This means that we replace the irrational rotation of the classical Diophantine problem by the circle doubling map $x \to 2x \pmod 1$. It describes the approximation by powers of 2 with the error of approximation being $n^{-\kappa}$. Our techniques allow us to treat other integer bases and subshifts of finite type without introducing new ideas, and even to treat other non linear piecewise expanding maps on the circle. However we stick to the dyadic case in order to avoid overcomplicated notation. Note that irrational circle rotations are uniquely ergodic all with the same invariant measure, Lebesgue measure. In contrast we have a fixed dynamical system but an infinite dimensional simplex of invariant measures.

Our problem is to study how well $2^n x \pmod 1$ approximates a point y. This depends on the point x, since the underlying dynamical system is not uniquely ergodic. More precisely we consider the set of y's with the same approximation rate for a given point x. This set depending on the point x can be considered as a random set with a fixed invariant Gibbs measure μ as the probability measure. Applying our mass transference principle, we show that the dimension of this set is constant for a μ-typical x and this enables us to calculate this constant in terms of the approximation rate. Moreover, the essential part of such a set with a given approximation rate does not depend on x for a certain interval of approximation rates, while outside this interval these sets are completely random.

Our problem is closely related to, but different from, the dynamical Borel-Cantelli lemma and the shrinking target problem. Consider a measure preserving map T. A shrinking target is a sequence of balls with decreasing radii and with centers fixed or moving (more generally, other forms than balls are also allowed). The question is to study the set of orbits $T^n x$ (or equivalently of the initial point x) which hit the target or equivalently which are well approximated by the target. See for example [BDV, CK, C, Do, HV, HV1, Ph]. In these articles one fixes there a sequence of shrinking targets C_n (or one asks which sequence of shrinking targets one can fix) and one studies the speed with which the target can decrease so that it is hit by almost all points infinitely often in the sense $T^n x \in C_n$. We need to go beyond the dynamical Borel-Cantelli lemma and the usual statistical tool of large deviations since they do not give satisfactory answers in our situation. We do not fix the center of the target, rather we define the center of the target in a dynamical way. We vary the speed and ask how many points will hit the target infinitely often. Many of the previously known results are proved in a more general framework than ours, however if one restricts them to our framework, then many of them can be derived from Theorem 2.2.

Our results can be interpreted in a more probabilistic way which is closely related to the classical Dvoretzky covering problem. Namely, consider an i.i.d. sequence $\{x_n\} \subset \mathbb{S}^1$ uniformly distributed on the unit circle $\mathbb{S}^1$ with respect to Lebesgue measure, a decreasing sequence of positive numbers $\{\ell_n\} \subset \mathbb{R}^+$ and the associated random intervals $(x_n - \ell_n/2 \pmod 1, x_n + \ell_n/2 \pmod 1)$ (i.e. $r_n = \ell_n/2$ in the above terminology). Since $\{x_n\}$ are independent and uniformly distributed, the Borel–Cantelli Lemma assures that almost surely (a.s. for short) we have $I(\{x_n\}, \{r_n\}) = \mathbb{S}^1$ except for a set of null Lebesgue measure, i.e. Lebesgue a.e. point in $\mathbb{S}^1$ is covered infinitely often by the intervals with probability one if and only if $\sum_{n=1}^{\infty} \ell_n = \infty$. Moreover $\sum_{n=1}^{\infty} \ell_n < \infty$ implies that Lebesgue a.e. point in $\mathbb{S}^1$ is covered finitely often with probability one. In 1956, Dvoretzky observed the possibility that *all* points in $\mathbb{S}^1$ are covered infinitely often with probability one for some slowly decreasing sequence $\{\ell_n\}$ [D]. In 1972, Shepp obtained a necessary and sufficient condition for all points in $\mathbb{S}^1$ to be covered infinitely often with probability one [Sh]:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \exp(\ell_1 + \cdots + \ell_n) = \infty.$$

This condition is satisfied for example by $\ell_n = \frac{1}{n}$. Important contributions were made by J.P. Kahane, P. Billard, P. Erdös, S. Orey, B. Mandelbrot et al. See Kahane's book [K] for a full history and a complete reference up to 1985 and see [BF, F1, F2, FK, FW1, JS] for more recent developments. In our case we choose the sequence $\{x_n\}$ dynamically, the stochasticity of our choice comes from choosing the initial point randomly with respect to an invariant Gibbs measure. Thus our process is stationary but not independent.

2. Rigorous statements of results

2.1. **Results on typical sequences.** In the present work, we consider the dynamics defined by the angle doubling map on the circle. We shall consider a generic orbit $\{x_n\} = \{T^n x\}$ of this map relative to a Gibbs measure ν_ϕ. Recall that the doubling map $T : \mathbb{S}^1 \to \mathbb{S}^1$ is defined by

$$Ts = 2s \pmod 1.$$

Divide the circle into 2^n dyadic intervals $C^p := C_n^p := [p/2^n, (p+1)/2^n)$. Let $E(t)$ be the dimension spectrum of ν_ϕ, which is defined by

$$E(t) := \dim_H \left\{ y : \lim_{r \to 0} \frac{\log \nu_\phi((y-r, y+r))}{\log r} = t \right\}. \tag{2·1}$$

Theorem 2.1. *Fix $\beta > 0$. Consider two sufficiently large n and L such that $n \ll L$, a ν_ϕ-typical point x and the set of dyadic interval C_n^p satisfying $\nu_\phi(C_n^p) \sim 2^{-\beta n}$ which the orbit segments $x, Tx, \dots, T^{L-1}x$ visits. Then*

$$\sharp\big\{C_n^p : \nu_\phi(C_n^p) \sim 2^{-\beta n}, \text{visited by } x \text{ up to time } L\big\} \sim 2^{n \min\{E(\beta),\ E(\beta) - \beta + (\log_2 L)/n\}}.$$

Here $a_n \sim b_n$ means that the ratio a/b is subexponential in n. A typical case of $n \ll L$ is $L = 2^{cn}$ for which we have $(\log_2 L)/n = c$.

Let h_{ν_ϕ} be the measure-theoretic entropy of ν_ϕ. If $\nu_\phi(C_n^p) \sim 2^{-\beta n}$ with $\beta > h_{\nu_\phi}$, we say that C_n^p is atypical with respect to ν_ϕ.

This result implies that there are two regimes depending on which of the two values the min in the exponent takes. Observe that the average visiting time of x to C_n^p with $\nu_\phi(C_n^p) \sim 2^{-\beta n}$ up to time L is

$$\mathbb{E}_{\nu_\phi} \sum_{n=0}^{L-1} 1_{C_n^p}(T^n x) \sim L 2^{-\beta n}.$$

For a fixed large L and $\beta > (\log_2 L)/n$, large deviation techniques tell us that the probability of visiting such an interval is extremely unlikely. However the theorem says there are still exponentially many $2^{(E(\beta)-\beta+(\log_2 L)/n)n}$ of them. However for fixed choices of such cylinders they will occur with 0 probability.

If $(\log_2 L)/n > h_{\nu_\phi}$ then the number of such intervals is exponentially smaller than the ones with measure at least $2^{-nE((\log_2 L)/n)}$. In the other regime, i.e., $\beta < (\log_2 L)/n$, the main contribution to the number of visited intervals scales like $2^{-h_{\nu_\phi} n}$.

There are two further phase transitions at $e^\pm$ where e^+ and e^- are respectively the maximal and minimal pointwise entropy of ν_ϕ. If $(\log_2 L)/n > e^+$, then all intervals C_n^p are visited. Between e^- and e^+, the occurrence of the intervals is directly connected with their local entropy, i.e. it is more likely to see intervals with large probability than with small probability. If $(\log_2 L)/n < e^-$, the intervals with the largest probability disappear and the likelihood of an event is directly correlated with the deviation for the typical event.

2.2. **Diophantine results.** We are interested in the solvability of infinitely many positive integers n in the inequation

$$\|T^n x - y\| = \|2^n x - y\| < r_n. \tag{2·2}$$

This is dyadic Diophantine approximation, homogeneous in the case $y = 0$ and inhomogeneous in the case $y \neq 0$. The sets $I(\{2^n x\}, \{r_n\})$ and $F(\{2^n x\}, \{r_n\})$ are respectively the sets of y which are well approximable or badly approximable with speed r_n. In other words I is the set of points obeying a Diophantine inequality with speed r_n. One of our theorems is similar to Jarnik type results in number theory. For $\kappa > 0$ consider the special sequence $r_n = \frac{1}{n^\kappa}$. Write

$$J_n^\kappa(s) = (T^n s - r_n \pmod 1, T^n s + r_n \pmod 1).$$

For $s \in \mathbb{S}^1$ let

$$I^\kappa(s) := \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} J_n^\kappa(s) = \left\{ t \in \mathbb{S}^1 \ : \ \sum_{n=0}^{\infty} \mathbf{1}_{J_n^\kappa(s)}(t) = \infty \right\},$$
$$F^\kappa(s) := \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} J_n^\kappa(s)^c = \left\{ t \in \mathbb{S}^1 \ : \ \sum_{n=0}^{\infty} \mathbf{1}_{J_n^\kappa(s)}(t) < \infty \right\}.$$

The following decomposition is obvious:

$$\mathbb{S}^1 = F^\kappa(s) \cup I^\kappa(s), \quad F^\kappa(s) \cap I^\kappa(s) = \varnothing.$$

If $\cup_{n=N}^{\infty} J_n^\kappa(s)$ is open and dense, then the set $I^\kappa(s)$ is a residual set in particular, $I^\kappa(s) \neq \varnothing$. It is the case for a typical point s relative to an ergodic measure with full support. However, as we will see, it is possible for $F^\kappa(s) = \varnothing$ for typical points. Let ν_ϕ, ν_ψ be two T-invariant probability Gibbs measures on $\mathbb{S}^1$ associated to normalized Hölder potentials ϕ and ψ (i.e. the pressures of ϕ and ψ are equal to zero). The measure ν_ϕ will be used to describe the randomness and the measure ν_ψ to describe sizes of sets.

Let

$$\begin{aligned} \kappa_{\phi,\psi,\mathbb{S}^1} : &= \sup \left\{ \kappa : \nu_\psi(I^\kappa(s)) = 1 \text{ for } \nu_\phi - a.e. \ s \right\}, \\ \kappa^F_{\phi,\mathbb{S}^1} : &= \sup \left\{ \kappa : F^\kappa(s) = \varnothing \text{ for } \nu_\phi - a.e. \ s \right\}. \end{aligned}$$

We are interested in the following questions:

(**Q1**) *How to determine the critical value $\kappa_{\phi,\psi,\mathbb{S}^1}$? More precisely when is $I^\kappa(s)$ of full ν_ψ-measure for ν_ϕ-almost every s?*

(**Q2**) *How to determine the critical value $\kappa^F_{\phi,\mathbb{S}^1}$? More precisely when is $I^\kappa(s)$ equal to $\mathbb{S}^1$ for ν_ϕ-almost every s ?*

(**Q3**) *What are the Hausdorff dimensions $\dim_H(F^\kappa(s))$, $\dim_H(I^\kappa(s))$ for ν_ϕ-almost every s ?*

Our answers to these questions are stated in the following theorems. Let

$$e^- := \inf_{\nu:\text{invariant}} \int (-\phi) d\nu, \tag{2·3}$$

$$e_{\max} := \int (-\phi) d\text{Leb}, \tag{2·4}$$

$$e^+ := \sup_{\nu:\text{invariant}} \int (-\phi) d\nu \tag{2·5}$$

where e_- and e_+ are respectively the minimal and maximal local entropy of ν_ϕ.

Theorem 2.2. *The critical value $\kappa_{\phi,\psi,\mathbb{S}^1}$ satisfies*

$$\kappa_{\phi,\psi,\mathbb{S}^1} = \frac{1}{\int (-\phi) d\nu_\psi}.$$

Notice that the integral $\int(-\phi)d\nu_\psi$ is nothing but the conditional entropy of ν_ϕ relative to ν_ψ. The theorem says that for ν_ϕ-a.e s the set $I^\kappa(s)$ supports the Gibbs measure ν_ψ if κ is small enough so that $\int(-\phi)d\nu_\psi < \frac{1}{\kappa}$. Also notice that for fixed s, the question whether $\nu_\psi(I^\kappa(s)) = 1$ is the shrinking target problem or dynamical Borel-Cantelli lemma (see [HV]), which formally corresponds to the case $\mu_\psi = \delta_s$, because the Dirac measure δ_s is not a Gibbs measure.

Usually the shrinking target problem concerns the set of x satisfying the Diophantine inequation (2·2). But the problem studied in this paper concerns the set of y satisfying the Diophantine inequation (2·2). The later set is actually a random set depending on x. This is a difference between the shrinking target problem and the problem studied in this paper.

The following theorem describes a new phenomenon in chaotic dynamical systems–it is possible to hit every shrinking target with a precise order of shrinking. A counterpart in the irrational rotation, which is not chaotic, is the theorem of Markov-Minkowski concerning the Diophantine inequation $\|n\alpha - y\| < 1/(\sqrt{5}n)$ ([Cas], p. 11 and p. 48).

Theorem 2.3. *The critical value $\kappa^F_{\phi,\mathbb{S}^1}$ satisfies*

$$\kappa^F_{\phi,\mathbb{S}^1} = \frac{1}{e_+}.$$

The theorem says that if κ is so small that $e^+ < \frac{1}{\kappa}$, then $I^\kappa(s) = \mathbb{S}^1$ or equivalently $F^\kappa(s) = \varnothing$ for ν_ϕ-a.e. s. This is the counterpart of the Kahane-Billard-Shepp condition for the random Dvoretzky covering. This result is generalized to the class of exponentially mixing dynamical systems in [FLL] where the authors use a different method inspired from [FK]. This technique originated from Salem-Zygmund's uniform estimate on random trigonometric polynomials (see [K]).

Theorem 2.4. *For ν_ϕ-a.e. s we have*

$$\dim_H F^\kappa(s) = \begin{cases} 1 & \text{if } \frac{1}{\kappa} \le e_{\max} \\ E(\frac{1}{\kappa}) & \text{if } \frac{1}{\kappa} > e_{\max} \end{cases}.$$

Theorem 2.5. *For ν_ϕ-a.e. s we have*

$$\dim_H I^\kappa(s) = \begin{cases} \frac{1}{\kappa} & \text{if } \frac{1}{\kappa} \le h_{\nu_\phi} \\ E(\frac{1}{\kappa}) & \text{if } h_{\nu_\phi} < \frac{1}{\kappa} < e_{\max} \\ 1 & \text{if } \frac{1}{\kappa} \ge e_{\max} \end{cases}.$$

We will transfer the problem to a similar one in a symbolic framework. As we shall see, our problem is closely related to hitting times and the later is related to local entropy.

The structure of the article is as follows. We start in Section 3 with background on ergodic theory, symbolic dynamics, decay of correlations, and multifractal analysis. In this section we prove higher order ϕ-mixing and a variational principle which are essential in the proofs of the main results. In Section 4 we transfer the covering problem to the symbolic setting and then relate covering properties to hitting time asymptotics, this transfer involves non-trivial work. In Section 5 we prove a first simple relation between hitting times and local entropy. This yields the proof of the Ornstein-Weiss return time theorem in the special case of Gibbs measures and also allows us to determine the critical exponent $\kappa_{\phi,\psi,\mathbb{S}^1}$. For the other exponents, more sophisticated estimates are needed. Sections 6 and 7 contain the core estimates on the probabilities of hitting time events. The fundamental tools relating hitting times to the entropy spectrum are developed. In Section 8 we study the structure of a short typical sequence. In particular we make a substantial improvement in the mass transference principle [BV] to multifractal Gibbs measures. Section 9 contains the results in the symbolic framework for the full shift while Section 10 generalizes these results to subshifts of finite type. Finally in Section 11 we prove the main theorems announced in Section 2 by transferring them from the shift space, again this transfer is not trivial.

3. Background

Convention. All logarithms and exponential functions in this article are taken to base 2. With this convention the notions of entropy and dimension coincide in our setup.

Ergodic theory. We need various standard definitions from ergodic theory: the metric entropy of an invariant measure ν denoted by h_ν, the notion of the Gibbs measure μ_ϕ with respect to a potential ϕ and the topological entropy for non compact sets E denoted by $h_{\text{top}}(E)$. The definitions of all these notions can be found in [P].

Symbolic dynamics. We use various standard notions from symbolic dynamics. Let

$$\Sigma_2^+ := \{0,1\}^{\mathbb{N}}.$$

For $y = (y_i)_{i\geq 0} \in \Sigma_2^+$ let $\sigma y \in \Sigma_2^+$ be such that $(\sigma y)_i = y_{i+1}$ for all $i \geq 0$. The map $\sigma : \Sigma_2^+ \to \Sigma_2^+$ is called the shift map. We denote a cylinder set by

$$C_n(y) := [y_0, y_1, \cdots, y_{n-1}] := \{z \in \Sigma_2^+ : \ z_i = y_i \text{ for } i = 0, 1, \ldots, n-1\}.$$

We will denote the length of the cylinder by $|C_n(y)| = n$. We will denote by

$$\pi(y) = \sum_{i=0}^{\infty} \frac{y_i}{2^{i+1}}$$

the natural projection from Σ_2^+ to $[0,1) := \mathbb{S}^1 := \mathbb{R}/\mathbb{Z}$. We consider the $\frac{1}{2}$-metric on Σ_2^+, i.e. for $x, y \in \Sigma_2^+$ let $d(x,y) = \frac{1}{2^n}$ where n is the least integer such that $x_n \neq y_n$. The pull back of the circle metric $\rho(x,y) := |\sum_{i=0}^{\infty} \frac{(x_i - y_i)}{2^{i+1}}|$ is almost equivalent to d in the sense that for $x \in \Sigma_2^+$ the ratio $\operatorname{diam}_\rho(C_n(x))/\operatorname{diam}_d(C_n(x))$ is bounded from below and above uniformly in n and x. Thus Hausdorff dimensions do not change under the projection, for details see [S1]. We denote by $\mu_{\max}$ the measure of maximal entropy for the shift. The projection of $\mu_{\max}$ is the Lebesgue measure on the circle.

3.1. **Fast decay of correlations.**

One of the key tools in our study is fast decay of correlations. This is related to Ruelle's theorem on transfer operators. Recall that for a α-Hölder potential $\phi : \Sigma_2^+ \to \mathbb{R}$, i.e.

$$[\phi]_\alpha := \sup_{x\neq y} \frac{|\phi(x)-\phi(y)|}{d(x,y)^\alpha} < \infty,$$

the transfer operator associated to ϕ is defined as follows

$$L_\phi f(x) = \sum_{\sigma y = x} 2^{\phi(y)} f(y).$$

This operator acts on the space of continuous functions $C(\Sigma_2^+)$ equipped with the supremum norm $\|f\|_\infty$ and on the space of α-Hölder continuous functions $H_\alpha(\Sigma_2^+)$ equipped with the Hölder norm

$$|||f||| := \|f\|_\infty + [f]_\alpha.$$

The well known Ruelle theorem asserts that [Ru]

(i) The spectral radius $\lambda > 0$ of $L_\phi : H_\alpha \to H_\alpha$ is a simple eigenvalue with an strictly positive eigenfunction h and there is a probability eigenmeasure ν for the adjoint operator L_ϕ^*, i.e. $L_\phi^*\nu = \lambda\nu$.

(ii) Choose h such that $\langle h, \nu\rangle := \int h d\nu = 1$. There exist constants $c > 0$ and $0 < \beta < 1$ such that for any $f \in H_\alpha$ we have

$$\|\lambda^{-n} L_\phi^n f - \langle f, \nu\rangle h\| \le c\beta^n |||f|||. \tag{3·6}$$

Let $P(\phi) = \log\lambda$ and call it the pressure of ϕ. The measure $\mu := h\nu$, denoted by μ_ϕ, is the so-called Gibbs measure associated to ϕ. Assume that ϕ is normalized, that is to say $\lambda = 1$. The Gibbs measure μ has the Gibbs property: there exists a constant $\gamma > 1$ such that

$$\frac{1}{\gamma} 2^{S_n\phi(x)} \le \mu(C_n[x]) \le \gamma 2^{S_n\phi(x)} \tag{3·7}$$

holds for all $x \in \Sigma_2^+$ and all $n \ge 1$ where

$$S_n f(y) := \sum_{j=0}^{n-1} f(\sigma^j y).$$

The Gibbs property (3·7) implies the following quasi-Bernoulli property (also known as Rényi's property) of μ_ϕ: for any two cylinders A and B we have

$$\frac{1}{\gamma^3}\mu_\phi(A)\mu_\phi(B) \le \mu_\phi(A\cap\sigma^{-|A|}B) \le \gamma^3\mu_\phi(A)\mu_\phi(B). \tag{3·8}$$

For the first inequality take a point $x \in A \cap \sigma^{-|A|}B$. By using three times the Gibbs property we get

$$\mu_\phi(A \cap \sigma^{-|A|}B) \geq \frac{1}{\gamma} 2^{S_{|A|}\phi(x)+S_{|B|}(\sigma^{|A|}x)} \geq \frac{1}{\gamma^3}\mu_\phi(A)\mu_\phi(B).$$

The second inequality can be proved in the same way.

This quasi-Bernoulli property can be generalized in the following way.

Theorem 3.1 (Higher order ϕ-mixing)**.** *Let $\mu = \mu_\phi$ be the Gibbs measure associated to a Hölder potential function ϕ. Let $\omega > 1$ be a sufficiently large integer. For any cylinder D_0 and any finite number of cylinders $D_1, \ldots, D_k$ of length n we have*

$$\gamma^{-3}\left(1 - c\beta^n)\right)^k \leq \frac{\mu\left(D_0 \cap \bigcap_{j=1}^k \sigma^{-[n_0+j(n+d)]}D_j\right)}{\prod_{j=0}^k \mu(D_j)} \leq \gamma^3\left(1 + c\beta^n)\right)^k \tag{3·9}$$

where $n_0 \geq |D_0|$ and $d = d(n) := \omega n$.

Proof. First remark that

$$D_0 \cap \bigcap_{j=1}^k \sigma^{-[n_0+j(n+d)]}D_j = D_0 \cap \sigma^{-|D_0|}\mathcal{B}$$

where

$$\mathcal{B} = \bigcap_{j=1}^k \sigma^{-[n_0-|D_0|+j(n+d)]}D_j$$

is a finite union of disjoint cylinders, which we denote by B_i's. Applying the quasi-Bernoulli property (3·8) to $A = D_0$ and $B = B_i$ we get

$$\frac{1}{\gamma^3}\mu_\phi(D_0)\mu_\phi(B_i) \leq \mu_\phi(D_0 \cap \sigma^{-|D_0|}B_i) \leq \gamma^3\mu_\phi(D_0)\mu_\phi(B_i).$$

Sum over all B_i's and we get

$$\frac{1}{\gamma^3}\mu_\phi(D_0)\mu_\phi(\mathcal{B}) \leq \mu_\phi(D_0 \cap \sigma^{-|D_0|}\mathcal{B}) \leq \gamma^3\mu_\phi(D_0)\mu_\phi(\mathcal{B}). \tag{3·10}$$

Notice that the invariance of μ_ϕ implies

$$\mu_\phi(\mathcal{B}) = \mu_\phi\left(\bigcap_{j=1}^k \sigma^{-[(j-1)(n+d)]}D_j\right).$$

Combining this with the inequalities (3·10), it suffices to prove

$$\left(1 - c\beta^n)\right)^k \leq \frac{\mu\left(\bigcap_{j=1}^k \sigma^{-[(j-1)(n+d)]}D_j\right)}{\prod_{j=1}^k \mu(D_j)} \leq \left(1 + c\beta^n)\right)^k. \tag{3·11}$$

Actually we can prove a little more. For simplicity, we write

$$\mathbb{E}f = \int f d\mu, \quad \|f\|_1 = \|f\|_{L^1(\mu)}.$$

From the inequality

$$|\mathbb{E}(f\circ\sigma^n\cdot g)| = |\mathbb{E}(f\cdot L^n g)| \le \|L^n g\|_\infty \|f\|_1$$

(applied to $g-\mathbb{E}g$ instead of g) and Ruelle's theorem, we deduce that for non-negative Hölder functions g and f we have

$$\left(1 - c\frac{\beta^n |\!|\!| g-\mathbb{E}g |\!|\!|}{\mathbb{E}g}\right) \le \frac{\mathbb{E}(f\circ\sigma^n\cdot g)}{\mathbb{E}f\mathbb{E}g} \le \left(1 + c\frac{\beta^n |\!|\!| g-\mathbb{E}g |\!|\!|}{\mathbb{E}g}\right).$$

Inductively, for a finite number of non-negative functions $g_1,\cdots,g_k\in H_\alpha$ and for integers $0=n_1<n_2<\cdots<n_k$ we have

$$\prod_{j=1}^{k-1}\left(1 - c\frac{\beta^{n_{j+1}-n_j} |\!|\!| g_j-\mathbb{E}g_j |\!|\!|}{\mathbb{E}g_j}\right)$$

$$\le \frac{\mathbb{E}\prod_{j=1}^k g_j\circ\sigma^{n_j}}{\prod_{j=1}^k \mathbb{E}g_j} \le \prod_{j=1}^{k-1}\left(1 + c\frac{\beta^{n_{j+1}-n_j} |\!|\!| g_j-\mathbb{E}g_j |\!|\!|}{\mathbb{E}g_j}\right).$$

To get (3·11), we apply these inequalities to characteristic functions of cylinders $g_j = 1_{D_j}$. In fact, since all cylinders D_j have the same length n, we have

$$|\!|\!| g_j |\!|\!| = 1 + 2^{\alpha n}, \quad \frac{1}{\mathbb{E}g_j} = \frac{1}{\mu(D_j)} \le \gamma 2^{n\max_x(-\phi(x))}$$

(the inequality is a consequence of the Gibbs property). Take $d := \omega n$ with a sufficiently large integer ω so that $\beta^\omega 2^{\alpha+\max(-\phi)} < 1$. Take n_j such that $n_1 = 0$ and $n_{j+1}-n_j = n+d$ for $j\ge 2$ and the inequalities (3·11) follow. □

3.2. **Multifractal analysis.**

Furthermore we will use various facts from multifractal analysis which can be found in the reference [P]. The notion of Hausdorff dimension of a set will be denoted by $\dim_H$. For a point $y\in\Sigma_2^+$ and an invariant measure ν we denote the *lower and upper local entropies* of ν at y by

$$(3\cdot12)\qquad \underline{h}_\nu(y) := \varliminf_{n\to\infty} -\frac{1}{n}\log\nu(C_n(y)), \quad \overline{h}_\nu(y) := \varlimsup_{n\to\infty} -\frac{1}{n}\log\nu(C_n(y)).$$

We define the *local entropy* $h_\nu(y)$ if the limit exists. In our framework we have fixed a generating partition, but we could have used any other generating partition. For a function $f:\Sigma_2^+\to\mathbb{R}$ we denote the ergodic sum by

$$S_m f(y) := \sum_{j=0}^{m-1} f(\sigma^j y).$$

We denote a Gibbs measure with respect to a Hölder potential ϕ by μ_ϕ. Without loss of generality we may assume that the potential is normalized so that its pressure $P(\phi)=0$. Then

$$(3\cdot13)\qquad \underline{h}_{\mu_\phi}(y) = -\varliminf_{n\to\infty}\frac{1}{n}S_n\phi(y)$$

and $h_{\mu_\phi}(y)$ satisfies a similar relation when the limit exists. If ν is an ergodic invariant measure then for ν a.e. y

$$h_{\mu_\phi}(y) = -\int_{\Sigma_2^+} \phi \, d\nu.$$

Furthermore if ν is another Gibbs measure μ_ψ then for μ_ψ a.e. y

$$h_{\mu_\phi}(y) = -\frac{d}{dt} P(\psi + t\phi)|_{t=0}. \tag{3·14}$$

Multifractal analysis deals with the study of the entropy spectrum. By our convention of the base of logarithm, the dimension spectrum introduced in Equation (2·1) and the entropy spectrum coincide, i.e., for $\mu_\phi = \pi^{-1}\nu_\phi$

$$E(t) = h_{\text{top}} \left\{ y \in \Sigma_2^+ \, : \, h_{\mu_\phi}(y) = t \right\}.$$

Here the topological entropy of an non-compact set $Z \subset \Sigma_2^+$ is given by the following Bowen type definition:

$$h_{\text{top}}(Z) := \inf \left\{ s : \lim_{n\to\infty} \inf \left\{ \sum_i 2^{-s|C_i|} \, : \, Z \subset \cup_i C_i \text{ and } |C_i| > n \right\} < \infty \right\}.$$

The following conditional variational is well known ([BSS, FF, FFW]).

Theorem 3.2 (Variational principle I)**.** *Let ϕ be a Hölder function such that $P(\phi) = 0$. For any $t \in \mathbb{R}$, we have*

$$E(t) = \sup_{\nu:\ invariant} \left\{ h(\nu) : \int (-\phi) d\nu = t \right\}. \tag{3·15}$$

The sup is attained by a Gibbs measure defined by the following equation.

$$E(t(q)) = P(q\phi) - q\frac{d}{ds}P(s\phi)|_{s=q} = h_{\mu_{-P(q\phi)+q\phi}} \tag{3·16}$$

where $t(q) = -\frac{d}{ds}P(q\phi)|_{s=q}$. The range of the function $t(q)$ is an interval $[e^-, e^+]$, possibly degenerate to a singleton.

The constants $e^\pm$ were defined in equation (2·3).

Let us state some more useful facts concerning the variational principle. The pressure is analytic and convex, therefore its derivative is nonincreasing. Furthermore if the function ϕ is not cohomologous to a constant then the derivative is strictly increasing and $e^- < e^+$. Thus the function $t(q)$ is invertible on the interval $[e^-, e^+]$. If t is not in this interval, then there is no point $y \in \Sigma_2^+$ with local entropy equal to t. The entropy $E(t)$ attains its maximum at the value

$$e_{\max} = t(0) = \int_{\Sigma_2^+} (-\phi) d\mu_{\max}.$$

We have $t(q) \leq e_{\max}$ if and only if $q \geq 0$. Furthermore

$$e^+ = \max_{\mu:\text{invariant}} \int (-\phi) \, d\mu, \quad e^- = \min_{\mu:\text{invariant}} \int (-\phi) \, d\mu.$$

The entropy spectrum is concave and real analytic in the interval (e^-, e^+). Its graph lies below the diagonal. It touches the diagonal at the unique point h_{μ_ϕ} and hence $E'(t) < 1$ if $t > h_{\mu_\phi}$. Moreover the interval $[e^-, e^+]$ is degenerate if and only if ϕ is cohomologous to the constant $-h_{\text{top}}$, i.e. the measure μ_ϕ is the measure of maximal entropy. In the degenerate case we have $e^- = e^+ = h_{\text{top}}$ and $E(h_{\text{top}}) = h_{\text{top}}$. For typical potentials in the sense of Baire for the Hölder topology, $E(e^-) = E(e^+) = 0$.

We will need the following variational principle.

Theorem 3.3 (Variational principle II)**.** *Let ϕ be a Hölder function. For any $t \in \mathbb{R}$, we have*

$$h_{\text{top}} \left\{ \underline{h}_{\mu_\phi}(y) < t \right\} = h_{\text{top}} \left\{ \overline{h}_{\mu_\phi}(y) < t \right\} = \sup_{s<t} E(s),$$

$$h_{\text{top}} \left\{ \underline{h}_{\mu_\phi}(y) \geq t \right\} = h_{\text{top}} \left\{ \overline{h}_{\mu_\phi}(y) \geq t \right\} = \sup_{s\geq t} E(s).$$

Proof. Let us start with the proof of the first fact. From the trivial fact

$$\left\{ \underline{h}_{\mu_\phi}(y) < t \right\} \supset \left\{ \overline{h}_{\mu_\phi}(y) < t \right\} \supset \bigcup_{s<t} \{ h_{\mu_\phi}(y) = s \},$$

we get immediately the following inequalities

$$h_{\text{top}} \left\{ \underline{h}_{\mu_\phi}(y) < t \right\} \geq h_{\text{top}} \left\{ \overline{h}_{\mu_\phi}(y) < t \right\} \geq \sup_{s<t} E(s).$$

Since $\sup_{t<e_{\max}} E(t) = 1$ the converse inequalities are trivial in the case $t \geq e_{\max}$. It remains to consider the case $t < e_{\max}$. Notice that we have $E(t) = \sup_{s<t} E(s)$. Also notice that there exists a positive number $q(t) > 0$ such that

$$\min_{q\geq 0} (P(q\phi) + qt) = P(q(t)\phi) + q(t)t = E(t).$$

Now let y be any point such that $\underline{h}_{\mu_\phi}(y) < t$. For $q = q(t) > 0$ we can apply Equation (3·13) to yield

$$\begin{aligned} \underline{h}_{\mu_{-P(q\phi)+q\phi}}(y) &= \varliminf_{n\to\infty} -\frac{1}{n} S_n\big(- P(q\phi) + q\phi\big)(y) \\ &= P(q\phi) + q \left(\varliminf_{n\to\infty} -\frac{1}{n} S_n \phi(y) \right) \\ &\leq P(q\phi) + qt = E(t). \end{aligned}$$

Thus applying the mass distribution principle (see Theorem 7.2 of [P]) yields $h_{\text{top}} \left\{ \underline{h}_{\mu_\phi}(y) < t \right\} \leq E(t)$, which completes the proof of the first line.

The second fact may be similarly proved. We just point out the following differences that

$$\left\{ \overline{h}_{\mu_\phi}(y) \geq t \right\} \supset \left\{ \underline{h}_{\mu_\phi}(y) \geq t \right\} \supset \bigcup_{s\geq t} \{ h_{\mu_\phi}(y) = s \},$$

and that for $t > e_{\max}$ there exists a negative number $q(t) < 0$ such that $E(t) = P(q(t)\phi) + q(t)t$. □

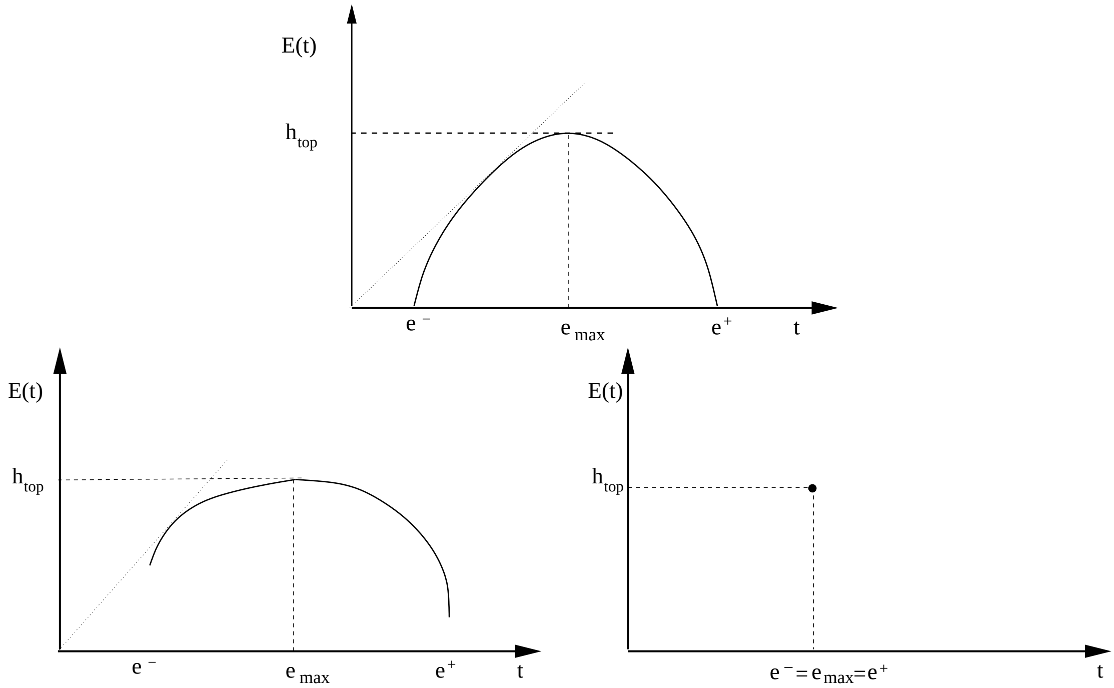

Figure 1. The entropy spectrum for typical, nontypical and degenerate potentials.

4. Covering questions are described by hitting times

It is well known that the doubling map is semi-conjugate to the shift map on Σ_2^+. As we shall see, the initial covering questions can be translated into similar questions concerning the shift map and these questions are described by the hitting time that we are going to define now. We will also see that hitting times are related to local entropy.

For $x \in \Sigma_2^+$ and C a cylinder, let

$$\tau(x, C) := \inf\{l \geq 1 : \sigma^l x \in C\}$$

be the *first hitting time* of C by x. For $x, y \in \Sigma_2^+$ let

$$\tau_n(x, y) := \tau(x, C_n(y))$$

$$\alpha(x, y) := \varliminf_{n \to \infty} \frac{1}{n} \log \tau_n(x, y). \tag{4·1}$$

Let $\lfloor a \rfloor$ denote the integer part of a real number a. Let

$$\begin{aligned} \mathcal{F}^\kappa(x) &:= \{y \in \Sigma_2^+ : \ y \notin \cap_{N=1}^\infty \cup_{n \geq N} C_{\lfloor \kappa \log n \rfloor}(\sigma^n x)\}, \\ \mathcal{I}^\kappa(x) &:= \{y \in \Sigma_2^+ : \ y \in \cap_{N=1}^\infty \cup_{n \geq N} C_{\lfloor \kappa \log n \rfloor}(\sigma^n x)\}. \end{aligned}$$

We have the following trivial decomposition

$$\Sigma_2^+ = \mathcal{F}^\kappa(x) \cup \mathcal{I}^\kappa(x), \qquad \mathcal{F}^\kappa(x) \cap \mathcal{I}^\kappa(x) = \varnothing.$$

Suppose that μ_ϕ, μ_ψ are σ-invariant probability Gibbs measures on Σ_2^+. Let

$$\kappa_{\phi,\psi,\Sigma_2^+} := \sup\{\kappa : \mu_\psi(\mathcal{I}^\kappa(x)) = 1 \text{ for } \mu_\phi - a.e.\ x\},$$
$$\kappa^F_{\phi,\Sigma_2^+} := \sup\{\kappa : \mathcal{F}^\kappa(x) = \varnothing \text{ for } \mu_\phi - a.e.\ x\}.$$

One of our goals is to determine the values of both critical exponents $\kappa_{\phi,\psi,\Sigma_2^+}$ and $\kappa^F_{\phi,\Sigma_2^+}$ and the other one is to compute the Hausdorff dimensions of $\mathcal{F}^\kappa(x)$ and $\mathcal{I}^\kappa(x)$. Let

$$\mathcal{O}(x) = \{\sigma^n x : n \geq 0\}, \quad \mathcal{O}^+(x) = \mathcal{O}(x) \setminus \{x\}.$$

Lemma 4.1. *There exists an integer $n_0 \geq 1$ such that $y = \sigma^{n_0} x$ (i.e. $y \in \mathcal{O}^+(x)$) if and only if the hitting time sequence $\tau_k(x, y)$ is bounded.*

Proof. If $y = \sigma^{n_0} x$ then it is obvious that $\tau_k(x, y) \leq n_0$ for all k. Conversely, suppose there is a positive constant such that $\tau_k(x, y) \leq K$. Fix an integer $1 \leq t \leq K$ such that $\tau_{k_i}(x, y) = t$ holds for an infinite subsequence k_i. Then $\sigma^t x \in C_{k_i}(y)$ for all i. Letting $i \to \infty$ we get $\sigma^t x = y$. □

Lemma 4.2.

$$\left\{y \in \Sigma_2^+ : \alpha(x,y) > \frac{1}{\kappa}\right\} \subset \mathcal{F}^\kappa(x) \subset \left\{y \in \Sigma_2^+ : \alpha(x,y) \geq \frac{1}{\kappa}\right\} \cup \mathcal{O}^+(x),$$
$$\left\{y \in \Sigma_2^+ : \alpha(x,y) < \frac{1}{\kappa}\right\} \setminus \mathcal{O}^+(x) \subset \mathcal{I}^\kappa(x) \subset \left\{y \in \Sigma_2^+ : \alpha(x,y) \leq \frac{1}{\kappa}\right\}.$$

Proof. The top left and bottom right inclusions imply one another. Let us prove the bottom right inclusion. Suppose $y \in \mathcal{I}^\kappa(x)$. Then $y \in C_{\lfloor \kappa \log n \rfloor}(\sigma^n x)$ or equivalently $\sigma^n x \in C_{\lfloor \kappa \log n \rfloor}(y)$ for infinitely many n. Thus $\tau_{\lfloor \kappa \log n \rfloor}(x, y) \leq n$ for infinitely many n, which implies $\alpha(x, y) \leq \kappa^{-1}$.

The top right and bottom left inclusions imply one another. So, it remains to prove the bottom left inclusion. Suppose $\alpha := \alpha(x, y) < \kappa^{-1}$ and $y \notin \mathcal{O}^+(x)$. Take $\varepsilon > 0$ such that $\kappa < \frac{1}{\alpha+\varepsilon}$. By the definition of $\alpha := \alpha(x, y)$, there is a subsequence k_i such that $\log \tau_{k_i}(x, y) \leq (\alpha + \varepsilon) k_i$, i.e. $k_i \geq \frac{\log \tau_{k_i}(x,y)}{\alpha+\varepsilon}$. The definition of $\tau_{k_i}(x, y)$ implies that

$$\sigma^{\tau_{k_i}} x \in C_{k_i}(y) \subset C_{\left\lfloor \frac{\log \tau_{k_i}}{\alpha+\varepsilon} \right\rfloor}(y) \subset C_{\lfloor \kappa \log \tau_{k_i} \rfloor}(y).$$

Since $y \notin \mathcal{O}^+(x)$ the previous lemma yields that τ_{k_i} is not bounded. Thus $\sigma^n x \in C_{\lfloor \kappa \log n \rfloor}(y)$ or equivalently $y \in C_{\lfloor \kappa \log n \rfloor}(\sigma^n x)$ for infinitely many $n = \tau_{k_i}$.

□

We should point out that points y on the orbit $\mathcal{O}^+(x)$ have the property that $\alpha(x, y) = 0 < 1/\kappa$, but they are not necessarily contained in $\mathcal{I}^\kappa(x)$. For example, if x is an eventually periodic point but not periodic and if y is on the orbit $\mathcal{O}^+(x)$ but not in the cycle of x, then $y \notin \mathcal{I}^\kappa(x)$. However, for μ_ϕ-almost all x, we have the following situation:

Lemma 4.3. *For μ_ϕ a.e. x, we have $\mathcal{O}(x) \subset \mathcal{I}^\kappa(x)$ if $\frac{1}{\kappa} > h_{\mu_\phi}$ and $\mathcal{O}(x) \subset \mathcal{F}^\kappa(x)$ if $\frac{1}{\kappa} < h_{\mu_\phi}$.*

Proof. Let $y \in \mathcal{O}(x)$ where x is not eventually periodic. Then there exists a unique integer $n_0 \geq 0$ such that $y = \sigma^{n_0} x$. Define the hitting time after n_0 by

$$\tau_n^{(n_0)}(x, y) := \inf\{k > n_0 : \sigma^k x \in C_n(y)\} = \tau_n(\sigma^{n_0} x, y) + n_0.$$

Since $y \notin \mathcal{O}^+(\sigma^{n_0} x)$ Lemma 4.1 implies that $\tau_n^{(n_0)}(x, y) \to \infty$ as $n \to \infty$. Let

$$\alpha^{(n_0)}(x, y) = \varliminf_{n\to\infty} \frac{1}{n} \log \tau_n^{(n_0)}(x, y). \tag{4·2}$$

Hence

$$y \in \mathcal{I}^\kappa(x) \ \text{ if } \ \alpha^{(n_0)}(x, y) < \frac{1}{\kappa}, \quad \text{and} \quad y \in \mathcal{F}^\kappa(x) \ \text{ if } \ \alpha^{(n_0)}(x, y) > \frac{1}{\kappa}.$$

Now

$$\alpha^{(n_0)}(x, y) = \alpha(y, y) = \alpha(\sigma^{n_0} x, \sigma^{n_0} x).$$

Thus applying the Ornstein-Weiss return time theorem [OW] yields that $\alpha(x, x) = h_{\mu_\phi}$ for μ_ϕ-a.e. x. Finally the invariance of μ implies that $\alpha(\sigma^n x, \sigma^n x) = h_{\mu_\phi}$ for μ_ϕ a.e. x and for all n. □

5. Hitting time and local entropy: basic relation

As Lemmas 4.2 and 4.3 show, we have to study the hitting time $\alpha(x, y)$ of the Gibbs measure μ_ϕ. We will show that the hitting time is related to local entropy. Local entropy have been well studied in the literature [BPS, FFW, OW, P, PW, S2].

In this section, we start with a basic relation between hitting times and local entropy. This allows us to compute the critical value $\kappa_{\phi,\psi,\Sigma_2^+}$.

Let us first introduce a generalized notion of local entropy. Let $\{C_n\}$ be a sequence of (arbitrary) cylinders with length $|C_n| = n$. We define the *lower local entropy of the sequence* $\{C_n\}$ by

$$\underline{h}_{\mu_\phi}(\{C_n\}) := \varliminf_{n\to\infty} -\frac{\log \mu_\phi(C_n)}{n}. \tag{5·1}$$

5.1. **The Basic relation.** We have the following basic relation between local entropy and the hitting times.

Theorem 5.1. *Suppose that μ_ϕ is a Gibbs measure associated to a Hölder potential ϕ and that $\{C_n\}$ is a sequence of (arbitrary) cylinders of length n. Then for μ_ϕ a.e. x we have*

$$\varliminf_{n\to\infty} \frac{\log \tau(x, C_n)}{n} = \underline{h}_{\mu_\phi}(\{C_n\}) \tag{5·2}$$

Proof. A special case of this theorem was proven by Chazottes [C]. The proof follows the idea of Chazottes closely. We include it for completeness.

Let $\tau_n(x) := \tau(x, C_n)$. For any $\varepsilon > 0$ for $n = n(\varepsilon)$ large enough $\max_x \mu_\phi(C_n(x)) < e^{-(e^- - \varepsilon)n}$, thus $\mu_\phi(C_n) \to 0$. Fix $\varepsilon > 0$ and let

$$A_n := \{x \in \Sigma_2^+ : \ \tau_n(x)\mu_\phi(C_n) < 2^{-\varepsilon n}\},$$
$$B_n := \{x \in \Sigma_2^+ : \ \tau_n(x)\mu_\phi(C_n) > 2^{\varepsilon n}\}.$$

We will prove that

$$\sum \mu_\phi(A_n \cup B_n) \le \sum \mu_\phi(A_n) + \sum \mu_\phi(B_n) < \infty.$$

Once we have shown this we apply the first part of the Borel-Cantelli lemma to conclude the proof.

First consider the series $\sum \mu_\phi(A_n)$, which is simpler to handle. We have

$$A_n \subset A_n^0 \cup \cdots \cup A_n^m$$

where

$$A_n^i := \{x \in \Sigma_2^+ : \ \sigma^i x \in C_n\}, \quad m = \lfloor 2^{-\varepsilon n}/\mu_\phi(C_n)\rfloor.$$

Since $\mu_\phi(A_n^i) = \mu_\phi(A_n^j) = \mu_\phi(C_n)$, this yields

$$\mu(A_n) \le \left(\frac{2^{-\varepsilon n}}{\mu_\phi(C_n)} + 2\right)\mu_\phi(C_n) \le 2^{-\epsilon n} + 2\mu_\phi(C_n).$$

Now we distinguish two cases: $\underline{h}_{\mu_\phi}(\{C_n\}) > 0$ and $\underline{h}_{\mu_\phi}(\{C_n\}) = 0$. In the first case, $\mu_\phi(C_n)$ decays exponentially fast, so that $\sum \mu_\phi(C_n) < \infty$, then $\sum \mu_\phi(A_n) < \infty$. In the second case, since $\mu_\phi(C_n) \to 0$, we can find some subsequence n_k such that $\sum_k \mu_\phi(C_{n_k}) < \infty$ so that $\sum_k \mu_\phi(A_{n_k}) < \infty$. So

$$\varliminf_{n\to\infty} \frac{\log \tau(x, C_n)}{n} \le \varliminf_{k\to\infty} \frac{\log \tau(x, C_{n_k})}{n_k} = 0.$$

Now we turn to the analysis of the series $\sum \mu_\phi(B_n)$. Choose a large integer $\omega > 1$ as in Theorem 3.1 and $d := d(n) := \omega n$. Let

$$B_n^i := \{x : \sigma^{i(n+d)}x \notin C_n\}, \quad m := \lfloor 2^{\varepsilon n}/\mu_\phi(C_n)(n+d)\rfloor - 1.$$

Thus

$$B_n \subset B_n^0 \cap \cdots \cap B_n^m = \bigcup_{D_0,\ldots,D_m} D_0 \cap \sigma^{-(n+d)}D_1 \cap \cdots \cap \sigma^{-m(n+d)}D_m$$

where the D_i are cylinders (not necessarily distinct) of length n disjoint from C_n. Thus, by the higher order ϕ-mixing 3·9, we get

$$\begin{aligned}
\mu_\phi(B_n) &\leq \sum_{D_0,\ldots,D_m} \mu_\phi(D_0 \cap \sigma^{-(n+d)} D_1 \cap \cdots \cap \sigma^{-m(n+d)} D_m) \\
&\leq (1 + c\beta^d)^m \sum_{D_0,\ldots,D_m} \prod_{i=0}^{m} \mu_\phi(\sigma^{-i(n+d)} D_i) \\
&\leq [(1 + c\beta^d)(1 - \mu_\phi(C_n))]^{m+1} \\
&\leq \left(1 - \frac{\mu_\phi(C_n)}{2}\right)^{m+1} \\
&\leq e^{-(m+1)\mu_\phi(C_n)/2} \\
&\leq e^{-2^{\epsilon n - 1}/(n+d)}.
\end{aligned}$$

□

Corollary 5.2. *For any $y \in \Sigma_2^+$ and for μ_ϕ a.e. x*

$$\alpha(x, y) = \underline{h}_{\mu_\phi}(y).$$

An application of Fubini's Theorem yields

Corollary 5.3. *Let ν be a probability measure on Σ_2^+. Then for $\mu_\phi \times \nu$ a.e. (x, y) we have*

$$\alpha(x, y) = \underline{h}_{\mu_\phi}(y).$$

The hitting time $\alpha(x, x)$ is what we called the return time. The following result due to Ornstein and Weiss [OW] concerning the return time is well known and holds for all ergodic invariant measures. For Gibbs measures, it can be similarly proved as the above theorem.

Corollary 5.4. *For μ_ϕ a.e. x we have*

$$\alpha(x, x) = \alpha(\sigma^k x, \sigma^k x) = \underline{h}_{\mu_\phi}(x) = h_{\mu_\phi} \qquad (\forall k \geq 1).$$

5.2. **Determination of $\kappa_{\phi,\psi,\Sigma_2^+}$.**

Recall that $-\int \phi d\mu_\psi$ is nothing but the conditional entropy of μ_ϕ relative to μ_ψ. As a direct consequence of Lemma 4.2 and Chazottes' theorem, we get immediately the following critical value.

Theorem 5.5. *Let ϕ and ψ be Hölder functions on Σ_2^+. We have*

$$\kappa_{\phi,\psi,\Sigma_2^+} = \frac{1}{-\int_{\Sigma_2^+} \phi \, d\mu_\psi} = -\frac{1}{\frac{d}{dt} P(\psi + t\phi)|_{t=0}}.$$

Proof. Suppose that μ_ϕ and μ_ψ are ergodic Gibbs measures with $P(\phi) = P(\psi) = 0$. Corollary 5.3 implies that for $\mu_\phi \times \mu_\psi$ a.e. (x, y)

$$\alpha(x, y) = h_{\mu_\phi}(y) = -\int_{\Sigma_2^+} \phi \, d\mu_\psi = -\frac{d}{dt} P(\psi + t\phi)|_{t=0}.$$

Thus applying Lemma 4.2 yields the assertion of the theorem. □

6. Big hitting probability and Study of $\mathcal{F}^\kappa(x)$

We will give answers to question (**Q2**) and to the part of question (**Q3**) concerning $\mathcal{F}^\kappa(x)$.

6.1. Big hitting probability.

Heuristically points of small local entropy (i.e. large "local measure") are hit with big probability. More precisely we have

Lemma 6.1 (Big hitting probability)**.** *Fix $a > 0$ and let γ be a positive numbers less that a. Let $K := 2^{an}$. Fix L cylinders $C_1, \cdots, C_L$ of length n satisfying $\mu_\phi(C_i) \geq 2^{-(a-\gamma)n}$. Then*

$$\mu_\phi\{x : \exists C \in \{C_i\} \text{ such that } \tau_n(x, C) > K\} \leq 2^{-\lambda n}$$

for any positive λ for sufficiently large n.

Proof. Since $e^+ = \max_{\mu \text{ inv}} - \int \phi \, d\mu < \max -\phi < \alpha + \max -\phi$ the ω chosen in Theorem 3.1 satisfies $\beta^\omega < 2^{e^+}$. Let $d := d(n) := \omega n$. We have L possibilities for the cylinder C. Let $m := \lfloor K/(1+\omega)n \rfloor - 1$. Fix a choice C from these L cylinders and let $D_0, \ldots, D_m$ denote any cylinders of length n (possibly with repetition), which are disjoint from C.

For a fixed C, let G_C be the set of points in Σ_2^+ in which the chosen cylinder C, considered as a word, does not appear up to time K. In particular, it does not appear at times $n + d, \cdots, m(n + d)$. Thus

$$\mu_\phi(G_C) \leq \sum_{D_0,\ldots,D_m} \mu_\phi(D_0 \cap \sigma^{-n+d} D_1 \cap \cdots \cap \sigma^{-m(n+d)} D_m).$$

By the higher order ϕ-mixing 3·9, we get

$$\begin{aligned}
\mu_\phi(G_C) &\leq (1 + c\beta^d)^{m+1} \sum_{D_0,\ldots,D_m} \prod_{i=0}^{m} \mu_\phi(\sigma^{-i(n+d)} D_i) \\
&= \left[(1 + c\beta^d)(1 - \min_{C_i} \mu_\phi(C_i)) \right]^m \\
&\leq \left(1 - \frac{1}{2} \min_{C_i} \mu_\phi(C_i) \right)^m .
\end{aligned}$$

Summing over all the $L(\leq 2^n)$ possible cylinders C yields

$$\begin{aligned}
&\mu_\phi\{x : \exists C \in \{C_i\} \text{ such that } \tau_n(x, C) > K\} \\
&\quad \leq \sum_C \mu_\phi(G_C) \\
&\quad \leq L\left(1 - \frac{1}{2}\min_{C_i} \mu_\phi(C_i))\right)^m \\
&\quad \leq L\left(1 - \frac{1}{2}\min_{C_i} \mu_\phi(C_i))\right)^{2^{\gamma n}/(\min_{C_i} \mu_\phi(C_i)(1+\omega)n)} \\
&\quad \leq \text{const} \cdot 2^n \cdot \left(e^{-1/2}\right)^{2^{\gamma n}/(1+\omega)n} \\
&\quad \leq 2^{-\lambda n}
\end{aligned}$$

for any positive λ and sufficiently large n. □

6.2. **The set of late hits.**

Let us recall that $\{y \in \Sigma_2^+ : \alpha(x, y) \geq t\}$ is random but $\{y \in \Sigma_2^+ : \underline{h}_{\mu_\phi}(y) \geq t\}$ is deterministic (i.e. independent of x). The following theorem is deduced from Lemma 6.1 (big hitting probability) and Corollary 5.3 (Ornstein-Weiss type theorem on return times).

Theorem 6.2. *For any $t \geq 0$ and for μ_ϕ a.e. x we have*

$$\{y \in \Sigma_2^+ : \alpha(x, y) \geq t\} \subset \{y \in \Sigma_2^+ : \underline{h}_{\mu_\phi}(y) \geq t\}. \tag{6·1}$$

Moreover if ν is any probability measure on Σ_2^+, then for μ_ϕ a.e. x we have

$$\nu\left(\{y \in \Sigma_2^+ : \alpha(x, y) \geq t\}\Delta\{y \in \Sigma_2^+ : \underline{h}_{\mu_\phi}(y) \geq t\}\right) = 0.$$

Proof. The case $t = 0$ is trivial. Assume $t > 0$. Let

$$H_{\geq t}(x) = \{y \in \Sigma_2^+ : \alpha(x, y) \geq t\}, \quad E_{\geq t} = \{y \in \Sigma_2^+ : \underline{h}_{\mu_\phi}(y) \geq t\}.$$

By definition, we have

$$H_{\geq t}(x) = \bigcap_{\varepsilon > 0} \underline{\lim}_{n \to \infty} H_{n,\varepsilon}(x)$$

with $H_{n,\varepsilon}(x) = \{y : \tau_n(x, y) \geq 2^{(t-\varepsilon)n}\}$, and

$$E_{\geq t} = \bigcap_{\varepsilon > 0} \underline{\lim}_{n \to \infty} E_{n,\varepsilon}$$

with $E_{n,\varepsilon}(x) = \{y : \mu_\phi(C_n(y)) \leq 2^{-(t-2\varepsilon)n}\}$. Thus it remains to prove that for μ_ϕ-a.e. x there exists $n(x) > 0$ such that

$$H_{n,\varepsilon}(x) \subset E_{n,\varepsilon} \quad \forall n \geq n(x).$$

Equivalently

$$E_{n,\varepsilon}^c \subset H_{n,\varepsilon}^c(x) \quad \forall n \geq n(x).$$

Notice that $E^c_{n,\varepsilon}$ is the union of all n-cylinders C such that $\mu_\phi(C) > 2^{-(t-2\varepsilon)n}$. Let $\mathcal{C}_{n,\varepsilon}$ be the set of all these cylinders. Applying Lemma 6.1 to $\{C_1, \cdots, C_L\} := \mathcal{C}_{n,\varepsilon}$ leads to

$$\sum_n \mu_\phi\{x \in \Sigma_2^+ : \exists C \in \mathcal{C}_{n,\varepsilon} \text{ s.t. } \tau_n(x, C) \geq 2^{(t-\varepsilon)n}\} < \infty.$$

So, by the Borel-Cantelli lemma, for μ_ϕ-a.e. x, for large n and for all $C \in \mathcal{C}_{n,\varepsilon}$ we have $\tau_n(x, C) < 2^{(t-\varepsilon)n}$, i.e. $C \subset H^c_{n,\varepsilon}(x)$. This proves the first assertion.

To prove the second assertion, it suffices to show that for μ_ϕ-a.e. x we have

$$\nu\{y \in \Sigma_2^+ : \underline{h}_{\mu_\phi}(y) \geq t, \alpha(x, y) < t\} = 0.$$

Let

$$E = \{(x, y) : \alpha(x, y) = \underline{h}_{\mu_\phi}(y)\}, \quad E_x = \{y : \alpha(x, y) = \underline{h}_{\mu_\phi}(y)\}.$$

By Corollary 5.3, we have $\mu_\phi \times \nu(E) = 1$. Then Fubini's theorem asserts that for μ_ϕ-a.e. x we have $\nu(E_x) = 1$, i.e.

$$\nu(E^c_x) = \nu\{y : \alpha(x, y) \neq \underline{h}_{\mu_\phi}(y)\} = 0.$$

We conclude by noticing

$$\{y : \underline{h}_{\mu_\phi}(y) \geq t, \alpha(x, y) < t\} \subset E^c_x.$$

□

We should point out that (6·1) is equivalent to

$$\{y \in \Sigma_2^+ : \underline{h}_{\mu_\phi}(y) < t\} \subset \{y \in \Sigma_2^+ : \alpha(x, y) < t\}. \tag{6·2}$$

This justifies our heuristics that points of small local entropy are hit early. We point out that the inverse inclusion of (6·2) does not hold. Actually for $t < e^-$, the deterministic set $\{y \in \Sigma_2^+ : \underline{h}_{\mu_\phi}(y) < t\}$ is empty, but if $1/\kappa < t$, the random set $\{y \in \Sigma_2^+ : \alpha(x, y) < t\}$ contains $I^\kappa(x)$ which is a residual set.

6.3. Computation of $\dim_H\{y : \alpha(x, y) \geq t\}$ and $\dim_H \mathcal{F}^\kappa(x)$.

Theorem 6.3. *For μ_ϕ-a.e. x, we have*

$$\dim_H \big\{y : \alpha(x, y) \geq t\big\} = \dim_H \big\{y : \underline{h}_{\mu_\phi} \geq t\big\}.$$

Proof. By the second variational principle (Theorem 3.3), there exists an $s \geq t$ such that

$$\dim_H\{y : \underline{h}_{\mu_\phi} \geq t\} = \dim_H \mu_{-P(q(s)\phi)+q(s)\phi}. \tag{6·3}$$

Applying Corollary 5.3 (with $\nu = \mu_{-P(q(s)\phi)+q(s)\phi}$) implies that

$$\mu_{-P(q(s)\phi)+q(s)\phi}(\{y : \underline{h}_{\mu_\phi}(y) = \alpha(x, y) = s\}) = 1 \quad \text{for } \mu_\phi - \text{a.e. } x.$$

It follows that for μ_ϕ-a.e. x we have

$$\begin{aligned} \dim_H\{y : \alpha(x, y) \geq t\} &\geq \dim_H\{y : \underline{h}_{\mu_\phi}(y) = \alpha(x, y) = s\} \\ &\geq \dim \mu_{-P(q(s)\phi)+q(s)\phi}. \end{aligned}$$

This, together with (6·3), implies

$$\dim_H \left\{y : \ \alpha(x,y) \geq t\right\} \geq \dim_H \left\{y : \ \underline{h}_{\mu_\phi} \geq t\right\} \quad \mu_\phi\text{-a.e.}$$

Now we turn to the reverse inequality. Observe the following decomposition

$$\{y : \alpha(x,y) \geq t\} = \{\alpha(x,y) \geq t, \underline{h}_{\mu_\phi}(y) < t\} \cup \{\alpha(x,y) \geq t, \underline{h}_{\mu_\phi}(y) \geq t\}.$$

Since

$$\dim_H\{\underline{h}_{\mu_\phi}(y) \geq t, \alpha(x,y) \geq t\} \leq \dim_H\{\underline{h}_{\mu_\phi}(y) \geq t\},$$

it suffices to remark that $\{y : \ \underline{h}_{\mu_\phi}(y) < t, \alpha(x,y) \geq t\} = \varnothing$ for μ_ϕ a.e. x. □

By this theorem, Lemmas 4.2 and 4.3, the second variational principle (Theorem 3.3) and the fact that there are no points with local entropy larger than e^+ [S2] we get

Theorem 6.4. *For μ_ϕ-a.e. x we have*

$$\begin{aligned} h_{\text{top}}(\mathcal{F}^\kappa(x)) &= 1 \qquad\quad \text{for } \frac{1}{\kappa} \leq e_{\max}, \\ h_{\text{top}}(\mathcal{F}^\kappa(x)) &= h_{\mu_{q(\kappa)\phi}} \quad \text{for} \quad e_{\max} \leq \frac{1}{\kappa} < \ e_+ \end{aligned}$$

where $q(\kappa)$ is chosen such that $h_{\mu_\phi}(y) = \frac{1}{\kappa}$ for $\mu_{q(\kappa)\phi}$ a.e. y. For μ_ϕ-a.e. x we also have

$$\begin{aligned} \mathcal{F}^\kappa(x) &= \varnothing \text{ (or equivalently } \mathcal{I}^\kappa(x) = \mathbb{S}^1) \text{ if } \ \frac{1}{\kappa} > e_+, \\ \mathcal{F}^\kappa(x) &\neq \varnothing \text{ (or equivalently } \mathcal{I}^\kappa(x) \neq \mathbb{S}^1) \text{ if } \ \frac{1}{\kappa} < e_+. \end{aligned}$$

Remark that the case $\frac{1}{\kappa} = e^+$ is not covered by the theorem because $E(t)$ is not continuous at $t = e^+$. We have the upper bound $\dim_H \mathcal{F}^{1/e_+} \leq E(e^+)$. A result due to Kahane for the random covering shows that a strict inequality may occur ([K], p.160).

7. Small hitting probability and upper bound of $\dim_H\{y : \alpha(x,y) \leq s\}$

7.1. **Small hitting probability.**

Lemma 7.1 (Small hitting probability)**.** *Fix $a > 0, b > 0, c > 0$ and $\gamma > \max(b-c, 0)$. Let $K := \lfloor 2^{an} \rfloor, L := \lfloor 2^{bn} \rfloor, N := \lfloor 2^{cn} \rfloor$. Fix L different cylinders $C_1, \cdots C_L$ of length n satisfying*

$$\mu_\phi(C_i) \leq 2^{-(a+\gamma)n}.$$

Then for any positive λ and sufficiently large n we have

$$\mu_\phi\{x : \ \tau_n(x, C_i) \leq K \text{ for } N \text{ different cylinders among the } C_i\} \leq 2^{-\lambda n}.$$

Proof. Let S be the set in question. That $x \in S$ means there exist times $\ell_1 < \ell_2 < \cdots < \ell_N \leq K$ and different cylinders $C_{i_1}, C_{i_2}, \cdots, C_{i_N}$ such that

$$\sigma^{\ell_1} x \in C_{i_1}, \quad \sigma^{\ell_2} x \in C_{i_2}, \quad \cdots, \quad \sigma^{\ell_N} x \in C_{i_N}.$$

In this sequence $\{\ell_k\}$ of length N there is a subsequence of $N/(3n+d)$ terms (with $d = \omega n$ as in Theorem 3.1), denoted (τ_j) such that $\tau_j - \tau_{j-1} \geq 3n+d$. For example, we may take $\tau_j = \ell_{(3n+d)j}$. Thus $x \in S$ implies

$$\sigma^{\tau_1} x \in C_{j_1}, \quad \sigma^{\tau_2} x \in C_{j_2}, \quad \cdots, \quad \sigma^{\tau_{N'}} x \in C_{j_{N'}}$$

for $N' := \lfloor N/(3n+d) \rfloor$ different cylinders taken from the list $C_1, C_2, \cdots, C_L$. Thus to each $x \in S$ we can associate the sequences (τ_j) and (C_{j_k}). Thus

$$x \in C(x) := \bigcap \sigma^{-\tau_i}(C_{j_i})$$

and S is covered by the union of $C(x)$. The higher order ϕ-mixing 3·9 implies that the measure of $C(x)$ is bounded by

$$\max_{1 \leq i \leq L} \mu_\phi(C_i)^{N'} (1 + c\beta^d)^{N'}.$$

Now, we have to estimate the number of different sets $C(x)$. First we have $\binom{L}{N'}$ choices for the N' different cylinders from the list of L words. Then we can choose $\binom{K}{N'}$ places (i.e. we fix the sequence τ_j) to put the chosen words in order to determine $C(x)$. Finally we have $N'!$ ways to arrange words into these N' (now fixed) places.

Thus the measure of the set in question can be majorized by

$$\binom{L}{N'} \binom{K}{N'} \cdot N'! \cdot \max_{C_i} \mu_\phi(C_i)^{N'} \cdot (1 + c\beta^d)^{N'}.$$

This is equal to

$$\frac{L!}{(L-N')!} \cdot \frac{K!}{(K-N')!N'!} \cdot (\max_{C_i} \mu_\phi(C_i))^{N'} \cdot (1 + c\beta^d)^{N'}.$$

Next using the estimates

$$\frac{L!}{(L-N')!} \leq L^{N'}, \quad \frac{K!}{(K-N')!N'!} \leq \text{const} \cdot K^{N'} \cdot \frac{e^{N'}}{N'^{N'}}$$

(the second one is implied by Stirling's formula), we conclude that the measure is majorized by

$$\begin{aligned} &\text{const} \cdot L^{N'} \cdot K^{N'} \cdot e^{N'} \cdot N'^{-N'} \cdot \left(2^{-(a+\gamma)n}\right)^{N'} \cdot (1 + c\beta)^{N'} \\ &= \text{const} \cdot \left(2^{bn} \cdot 2^{an} \cdot e \cdot 2^{-cn} \cdot 2^{-(a+\gamma)n} \cdot (1 + c\beta^d)\right)^{N'} \\ &\leq \text{const} \left(e \cdot (1 + c\beta^d) \cdot 2^{(b-c-\gamma)n}\right)^{N'}. \end{aligned}$$

Provided $\gamma > b - c$, this is less that $2^{\lambda n}$ for any positive λ and sufficiently large n. □

7.2. **Upper bound of** $\dim_H\{y : \alpha(x,y) \le s\}$**.**

Theorem 7.2. *If* $h_{\mu_\phi} < s < e_{\max}$ *then for* μ_ϕ*-a.e.* x *we have*

$$h_{\text{top}}\,\{y\ :\ \alpha(x,y) \le s\} \le E(s). \tag{7·1}$$

If $0 < s \le h_{\mu_\phi}$ *then for* all x *we have*

$$h_{\text{top}}\,\{y\ :\ \alpha(x,y) \le s\} \le s. \tag{7·2}$$

Proof. Let

$$\mathcal{A}_x(s) = \{y : \alpha(x,y) \le s\}.$$

The case $s \le h_{\mu_\phi}$ is simple. In fact, if $a > s$, $\mathcal{A}_x(s)$ is contained in

$$\overline{\lim_{n\to\infty}}\{y : \tau_n(x,y) \le 2^{an}\} = \overline{\lim_{n\to\infty}} \bigcup_{k=1}^{2^{an}} C_n(\sigma^k x) = \bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty}\bigcup_{k=1}^{2^{an}} C_n(\sigma^k x).$$

So, for $N \ge 1$, $\{C_n(\sigma^k x)\}_{n\ge N, 1\le k\le 2^{an}}$ is a cover of $\mathcal{A}_x(s)$. However $\sum_{n=N}^{\infty} 2^{an}\cdot 2^{-nt} < \infty$ for any $t > a$. This proves $h_{\text{top}}(\mathcal{A}_x(s)) \le a$ for $a > s$. We conclude by letting $a \downarrow s$.

Now let us look at the case $h_{\mu_\phi} < s < e_{\max}$. Write

$$\mathcal{A}_x(s) = \Big(\mathcal{A}_x(s) \cap \{y : \underline{h}_{\mu_\phi}(y) \le s\}\Big) \bigcup \Big(\mathcal{A}_x(s) \cap \{y : \underline{h}_{\mu_\phi}(y) > s\}\Big).$$

Since $h_{\text{top}}\{y : \underline{h}_{\mu_\phi}(y) \le s\} \le E(s)$, it suffices to show:

$$h_{\text{top}}\left(\mathcal{A}_x(s) \cap \left\{y\ :\ \underline{h}_{\mu_\phi}(y) > s\right\}\right) \le E(s) \quad \mu_\phi\text{–a.e. } x. \tag{7·3}$$

Let

$$\mathcal{H}(h', h'') = \left\{y\ :\ h' \le \underline{h}_{\mu_\phi}(y) \le h''\right\}.$$

It is clear that in order to show the inequation (7·3), it suffices to show

$$h_{\text{top}}\left(\mathcal{A}_x(s) \cap \mathcal{H}(h', h'')\right) \le E(s) \quad \mu_\phi\text{–a.e. } x \tag{7·4}$$

for all choices $s < h' < h''$.

Assume that $s < h^{'} < h^{''}$. In the following we fix an arbitrary number a such that

$$s < a < \min\{h', e_{\max}\}.$$

Since $s < e_{\max}$, such a's exist. Then we fix a number δ_0 such that

$$0 < \delta_0 < \min\{\frac{(1-E'(s))\,(h^{'}-a)}{2},\ \frac{(1-E'(s))\,(e_{\max}-a)}{2}\}.$$

Remark that $E'(s) < 1$ because $s > h_{\mu_\phi}$.

We have already seen that

$$\mathcal{A}_x(s) \subset \bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty}\bigcup_{|C|=n, \tau(x,C)\le 2^{an}} C.$$

This means if $z \in \mathcal{A}_x(s)$, then there exist infinitely many n's such that

$$\tau_n\,(x,z)) \le 2^{an}.$$

If furthermore $z \in \mathcal{H}(h^{'}, h^{''})$, then for n large enough

$$\mu_\phi \left(C_n(z)\right) \leq 2^{-h^{'}n}.$$

Let

$$\begin{aligned} A &= \{z : \tau_n\left(x, z\right) \leq 2^{an}, 2^{-(e_{\max}-\delta_0)n} < \mu_\phi(C_n(z)) \leq 2^{-h^{'}n} \ \text{ i.o.}\} \\ B &= \{z : \tau_n\left(x, z\right) \leq 2^{an}, 2^{-(e_{\max}+\delta_0)n} < \mu_\phi(C_n(z)) \leq 2^{-(e_{\max}-\delta_0)n} \ \text{ i.o.}\} \\ C &= \{z : \tau_n\left(x, z\right) \leq 2^{an}, \mu_\phi(C_n(z)) \leq 2^{-(e_{\max}+\delta_0)n} \ \text{ i.o.}\} \end{aligned}$$

where i.o. means the facts hold for infinitely many n's. What we have just seen shows

$$\mathcal{A}_x(s) \cap \mathcal{H}(h^{'}, h^{''}) \subset A \cup B \cup C.$$

We are going to show that the entropy of each of A, B and C is bounded by $E(a)$.

Before estimating the entropy, We make a remark. For $n \geq 1$ and $0 < h_1 < h_2$, let

$$\mathfrak{L}_n(h_1, h_2) = \{C : |C| = n, 2^{-h_2 n} \leq \mu_\phi(C) \leq 2^{-h_1 n}.$$

Then for $\epsilon > 0$ and for n sufficiently large (depending on h_1, h_2 and ϵ) we have

$$\text{(7·5)} \qquad \text{Card } \mathfrak{L}_n(h_1, h_2) \leq 2^{n[E(h_2)+\epsilon]} \ \text{ if } h_2 < e_{\max}$$

$$\text{(7·6)} \qquad \text{Card } \mathfrak{L}_n(h_1, h_2) \leq 2^{n[E(h_1)+\epsilon]} \ \text{ if } h_1 > e_{\max}.$$

In fact, assume $h_2 < e_{\max}$ (the other case may be similarly proved), there exists a positive number q such that $E(h_2) = P(q) + h_2 q$. Then

$$2^{-qh_2 n} \text{ Card } \mathfrak{L}_n(h_1, h_2) \leq \sum_{C \in \mathfrak{L}_n(h_1, h_2)} \mu_\phi(C)^q \leq 2^{n(P(q)+\epsilon)}.$$

Proof of $h_{\text{top}}(A) \leq E(a)$. Assume that $h' \leq e_{\max} - \delta_0$. Otherwise $A = \varnothing$. The function $E(t)$ is continuous then uniformly continuous on the interval $[h^{'}, e_{\max} - \delta_0]$. Then for any $\epsilon > 0$, there exists $\delta > 0$ such that $0 \leq E(h_2) - E(h_1) < \epsilon$ whenever $h^{'} \leq h_1 < h_2 \leq e_{\max} - \delta_0$ and $h_2 - h_1 < \delta_0$. We assume that $\epsilon < \frac{1}{2}\left(1 - E'(s)\right)(h^{'} - a)$ and $\delta < \delta_0$. We divide the interval $[h^{'}, e_{\max} - \delta_0]$ into intervals of the form $[h_1, h_2]$. It is clear that we have only to show

$$\text{(7·7)} \qquad h_{\text{top}} A(h_1, h_2) \leq E(a)$$

where

$$A(h_1, h_2) = \{z : \tau_n(x, z) \leq 2^{an}, 2^{-h_2 n} \leq \mu_\phi(C_n) \leq 2^{-h_1 n} \ \text{ i.o.}\}.$$

Let $\mathcal{L}_n(x; a, h_1, h_2) = \mathfrak{L}_n(h_1, h_2) \cap \{C_n : \tau(x, C_n) \leq 2^{an}\}$, i.e.

$$\mathcal{L}_n(x; a, h_1, h_2) = \{C_n : \tau(x, C_n) \leq 2^{an}, \quad 2^{-h_2 n} \leq \mu_\phi(C_n) \leq 2^{-h_1 n}\}$$

Notice that $\bigcup_{n\geq N} \mathcal{L}_n(x;a,h_1,h_2)$ is a cover of $A(h_1,h_2)$. We are then led to estimate the number

$$N_n(x;a,h_1,h_2) := \mathrm{Card}\mathcal{L}_n(x;a,h_1,h_2).$$

We apply Small Hitting Probability Lemma to

$$b = E(h_2)+\epsilon, \quad c = E(a), \quad \gamma = h_1 - a.$$

Since $E'(t)$ is decreasing and $E'(s) < 1$, we have

$$\begin{aligned} b-c &= E(h_2)+\epsilon - E(a) = (E(h_2)-E(h_1)) + \epsilon + (E(h_1)-E(a)) \\ &\leq E^{'}(s)(h_2-h_1) + \epsilon + E^{'}(s)(h_1-a) \\ &\leq h_2 - h_1 + \epsilon + E^{'}(s)(h_1-a) < h_1 - a = \gamma \end{aligned}$$

where we have used the facts

$$h_2 - h_1 < \delta_0 \leq \frac{(1-E'(s))\,(h_1-a)}{2}, \quad \epsilon < \frac{(1-E'(s))(h'-a)}{2}.$$

So, using (7·5), we can really apply Small Hitting Probability Lemma to get

$$\sum_n \mu_\phi\{x : N_n(x;a,h_1,h_2) > 2^{nE(a)}\} < \infty$$

By the Borel-Cantelli Lemma, for μ_ϕ-a.e. x, we have $N_n(x;a,h_1,h_2) \leq 2^{nE(a)}$ for $n \geq n(x)$. So, if $N \geq n(x)$, for any $\eta > 0$ we have

$$\sum_{n\geq N} \sum_{C\in\mathcal{L}_n(x;a,h_1,h_2)} (\mathrm{diam}C)^{E(a)+\eta} \leq \sum_{n\geq N} 2^{-n(E(a)+\eta)} \cdot 2^{nE(a)} < \infty.$$

Therefore (7·7) is proved.

Proof of $h_{\mathrm{top}}(B) \leq E(a)$. It is easier. Formally $B = A(h_1,h_2)$ with $h_1 = e_{\max} - \delta_0$, $h_2 = e_{\max} + \delta_0$. Remark that

$$\mathrm{Card}\{C_n : 2^{-h_2 n} \leq \mu_\phi(C_n) \leq 2^{-h_1 n}\} \leq 2^n,$$

because there are 2^n cylinders of order n. Let

$$b = E(e_{\max}) = 1, \quad c = E(a), \quad \gamma = h_1 - a.$$

Since $\delta_0 < \frac{(1-E'(s))(e_{\max}-a)}{2}$, we have

$$b - c = E(e_{\max}) - E(a) < E^{'}(s)(e_{\max}-a) < e_{\max} - a - 2\delta_0 < \gamma.$$

So we can also apply Small Hitting Probability Lemma to get the same estimate for $N(x;a,h_1,h_2)$ as above, which leads to the desired result.

Proof of $h_{\mathrm{top}}(C) \leq E(a)$ This case is also easy. First notice that for any $\eta > 0$ and for n large enough we have

$$\mathrm{Card}\{C_n : \mu_\phi(C_n) \leq 2^{-(e_{\max}+\delta_0)n}\} \leq 2^{n(E(e_{\max}+\delta_0)+\eta)}.$$

This can be similarly proved as (7·6). Let

$$b = E(e_{\max}+\delta_0)+\eta, \quad c = E(a)+\eta, \quad \gamma = e_{\max}+\delta_0 - a.$$

Since $E'(t) < 1$ for $t > h_{\mu_\phi}$, we have

$$b - c = E(e_{\max} + \delta_0) - E(a) < e_{\max} + \delta_0 - a = \gamma.$$

Small Probability Lemma implies that for μ-a.e. x and for large n we have

$$\text{Card}\{C_n : \tau(x, z) \leq 2^{an}, \mu_\phi(C_n) \leq 2^{-(e_{\max}+\delta_0)n}\} \leq 2^{n(E(a)+\eta)}$$

which implies $h_{\text{top}} C \leq E(a) + \eta$. □

Theorem 7.3. *If $h_{\mu_\phi} < s < e_{\max}$ then for μ_ϕ-a.e. x we have*

$$h_{\text{top}} \{y \,:\, \alpha(x, y) \leq s\} = E(s).$$

Proof. We simply need to prove the reverse inequality of (7·1) in Theorem 7.2. By Theorem 3.2 there is a Gibbs measure with entropy $E(s)$ supported on $\{y : h_{\mu_\phi}(y) = s\}$. Then Corollary 5.3 implies the result. □

For $0 < s < h_{\mu_\phi}$, the opposite inequality of (7·2):

$$h_{\text{top}} \{y \,:\, \alpha(x, y) \leq s\} \geq s$$

also holds. But its proof is much more involved. It can not be deduced from the mass transference principle as stated in [BV] since μ_ϕ has nontrivial entropy spectrum. In the next section we make a substantial improvement in the mass transference principle to multifractal Gibbs states. In order to prove it, we need to undertake a full investigation of the structure of typical sequences.

8. Typical sequences and Lower bound of $\dim_H\{y : \alpha(x, y) \leq c\}$

Recall that μ_ϕ is a Gibbs measure associated to a normalized Hölder potential ϕ. A cylinder C of length n is said to be a (n, ε)-cylinder if

$$2^{-(h+\varepsilon)n} \leq \mu_\phi(C) \leq 2^{-(h-\varepsilon)n}$$

where $h = h_\phi$ denotes the entropy of μ_ϕ. We denote by $\mathcal{C}_{n,\varepsilon}$ the set of all (n, ε)-cylinders. Sometimes we will say that a (n, ε)-cylinder is a good cylinder or the word determining a (n, ε)-cylinder is a good word. As we shall prove, a relatively short typical word contains plenty of good subwords of a fixed length and they are even different.

The following notations will be used. If C and D are cylinders, we denote by $C \star D$ the cylinder $C \cap \sigma^{-|C|} D$. If we read C and D as words, $C \star D$ is nothing but the concatenation of the words C and D. Let $d \geq 1$ be an integer, by $C \star_d D$ we mean $C \cap \sigma^{-(|C|+d)} D$, i.e.

$$C \star_d D = \bigcup_{G:|G|=d} C \star G \star D.$$

For a set S, $\sharp S$ will denote the cardinality of S.

8.1. **Frequency of good words in a typical orbit.**

Lemma 8.1. *Let μ_ϕ be a Gibbs measure with entropy $h := h_{\mu_\phi} > 0$. For any $\varepsilon > 0$, there exist an integer $n(\varepsilon) \geq 1$ and a Borel set $\mathcal{G}_\varepsilon$ with $\mu_\phi(\mathcal{G}_\varepsilon) > 1-\varepsilon$ such that for any $x \in \mathcal{G}_\varepsilon$ and any $n \geq n(\varepsilon)$, the cylinder $C = C_n(x)$ is a (n, ε)-cylinder. Consequently, if $n \geq n(\varepsilon)$, we have*

$$(1-\varepsilon)2^{(h-\varepsilon)n} \leq \sharp\mathcal{C}_{n,\varepsilon} \leq 2^{(h+\varepsilon)n}.$$

Proof. By the Shannon McMillan Breiman theorem, for μ_ϕ-a.e. x we have

$$\lim_{n\to\infty} -\frac{\log \mu_\phi(C_n(x))}{n} = h_{\mu_\phi}.$$

Then by Egorov's theorem, there is a number $n(\varepsilon) \geq 1$ such that the set

$$\mathcal{G}_\varepsilon := \left\{ y \in \Sigma_2^+ \ : \ -\frac{1}{n} \log \mu_\phi(C_n(y)) \in [h_{\mu_\phi} - \varepsilon, h_{\mu_\phi} + \varepsilon], \quad \forall n > n(\varepsilon) \right\}$$

has measure $\mu_\phi(\mathcal{G}_\varepsilon) > 1 - \varepsilon$.

The upper estimate $\sharp\mathcal{C}_{n,\varepsilon} \leq 2^{(h_{\mu_\phi}+\varepsilon)n}$ follows from

$$2^{-(h_{\mu_\phi}+\varepsilon)n}\sharp\mathcal{C}_{n,\varepsilon} \leq \sum_{C\in\mathcal{C}_{n,\varepsilon}} \mu_\phi(C) \leq 1.$$

The lower estimate $(1-\varepsilon)2^{(h_{\mu_\phi}-\varepsilon)n} \leq \sharp\mathcal{C}_{n,\varepsilon}$ follows from $\mathcal{G}_\varepsilon \subset \bigcup_{C\in\mathcal{C}_{n,\varepsilon}} C$ and

$$1 - \varepsilon \leq \mu_\phi(\mathcal{G}_\varepsilon) \leq \sum_{C\in\mathcal{C}_{n,\varepsilon}} \mu_\phi(C) \leq 2^{-(h_{\mu_\phi}-\varepsilon)n}\sharp\mathcal{C}_{n,\varepsilon}.$$

□

We call the set $\mathcal{G}_\varepsilon$ the set of ε-good points. By the definition of $\mathcal{G}_\varepsilon$, we have

$$\mathcal{G}_\varepsilon = \bigcap_{n=n(\varepsilon)}^{\infty} \bigcup_{C\in\mathcal{C}_{n,\varepsilon}} C.$$

Hence it is a G_δ set. We will write it as a decreasing limit of open sets in the following manner

$$\mathcal{G}_\varepsilon = \bigcap_{N=n(\varepsilon)}^{\infty} \bigcap_{n=n(\varepsilon)}^{N} \bigcup_{C\in\mathcal{C}_{n,\varepsilon}} C.$$

This representation of $\mathcal{G}_\varepsilon$ is useful in the proof of the following lemma.

Lemma 8.2. *Let $0 < \varepsilon < 1/2$ and let $L' \geq 1$ be an arbitrary integer. For any cylinder D of length L', we have*

$$\mu_\phi(D \cap \sigma^{-|D|}\mathcal{G}_\varepsilon) \geq \frac{1}{2\gamma^4} 2^{-L'\|\phi\|_\infty}$$

where $\gamma > 1$ is the constant involved in the Gibbs property of μ_ϕ (3·7).

Proof. We first recall the following quasi-Bernoulli property of μ_ϕ (3·8): for any two cylinders A and B we have

$$\mu_\phi(A \cap \sigma^{-|A|}B) \geq \frac{1}{\gamma^3}\mu_\phi(A)\mu_\phi(B).$$

Let us prove the lemma. The set $\mathcal{G}_\varepsilon$ is the decreasing limit of the open sets

$$\mathcal{G}_{N,\varepsilon} = \bigcap_{n=n(\varepsilon)}^{N} \bigcup_{C \in \mathcal{C}_{n,\varepsilon}} C.$$

Observe that $\mathcal{G}_{N,\varepsilon}$ is a union of cylinders of length N. Thus we have

$$\mu_\phi(D \cap \sigma^{-|D|}\mathcal{G}_\varepsilon) = \lim_{N\to\infty} \mu_\phi(D \cap \sigma^{-|D|}\mathcal{G}_{N,\varepsilon}) = \lim_{N\to\infty} \sum_C \mu_\phi(D \cap \sigma^{-|D|}C)$$

where C varies over all N-cylinders contained in $\mathcal{G}_{N,\varepsilon}$. First applying the quasi-Bernoulli property and then using the fact that $\mu_\phi(\mathcal{G}_{N,\varepsilon}) \geq 1 - \varepsilon > 1/2$, yields

$$\sum_C \mu_\phi(D \cap \sigma^{-|D|}C) \geq \frac{\mu_\phi(D)}{\gamma^3} \sum_C \mu_\phi(C) = \frac{\mu_\phi(D)}{\gamma^3}\mu_\phi(\mathcal{G}_{N,\varepsilon}) \geq \frac{\mu_\phi(D)}{2\gamma^3}.$$

To conclude, it suffices to remark that

$$\mu_\phi(D) \geq \frac{1}{\gamma}2^{-|D|\ \|\phi\|_\infty}$$

which is assured by the Gibbs property of μ_ϕ. □

The next theorem essentially says that a typical word of length $2^{cL''}$ contains many good subwords of length n with an arbitrary but fixed prefix D of length L'. We keep the notations $n(\epsilon)$ and $\mathcal{G}_\epsilon$ appearing in Lemma 8.1.

Theorem 8.3. *Let $c > 0$ be fixed. Let $0 < \varepsilon < \min(\frac{1}{2}, c)$, $0 < \eta < \frac{1}{2}$ and $L' \geq 1$. There exist an integer $n(\varepsilon, \eta, L') \geq L' + n(\varepsilon)$ and a Borel set $\mathcal{E}(\varepsilon, \eta, L')$ with $\mu_\phi(\mathcal{E}(\varepsilon, \eta, L')) > 1 - \eta$ such that if $x \in \mathcal{E}(\varepsilon, \eta, L')$ and $L'' > n(\varepsilon, \eta, L')$, for each L'-cylinder D there are at least $2^{(c-\varepsilon)L''}$ points of the finite orbit $\sigma^j x$ $(2^{L'} + 1 \leq j \leq 2^{cL''})$, which fall into $D \cap \sigma^{-L'}\mathcal{G}_\varepsilon$.*

Proof. Let

$$m(L') := \frac{1}{2\gamma^4}2^{-L'\|\phi\|_\infty}$$

be the lower bound which appeared in the last lemma. For $x \in \Sigma_2^+$, define

$$n_{D,L',\varepsilon}(x) := \inf\left\{ n \in \mathbb{N} \,:\, \frac{1}{N}\sum_{j=2^{L'}+1}^{2^{L'}+N} \mathbf{1}_{D\cap\sigma^{-L'}\mathcal{G}_\varepsilon}(\sigma^j x) > \frac{1}{2}m(L'), \forall N \geq n \right\}$$

and

$$n_{L',\varepsilon}(x) = \max_D n_{D,L',\varepsilon}(x).$$

By Lemma 8.2 and Birkhoff's ergodic theorem we have

$$\mu_\phi(x \in \Sigma_2^+ \,:\, n_{L',\varepsilon}(x) < \infty) = 1.$$

So, for any $\eta > 0$, there exists an integer $\widehat{n}(L', \varepsilon, \eta)$ such that the Borel set

$$\mathcal{E}(L', \varepsilon, \eta) := \{x \in \Sigma_2^+ : \ n_{L',\varepsilon}(x) \le \widehat{n}(L', \varepsilon, \eta)\}$$

satisfies

$$\mu_\phi(\mathcal{E}(L', \varepsilon, \eta)) > 1 - \eta.$$

Fix $n(L', \varepsilon, \eta) \ge 1$ sufficiently large so that

$$\frac{1}{2} m(L')[2^{\varepsilon n(L',\varepsilon,\eta)} - 2^{L'}] \ge 1,$$

$$n(L', \varepsilon, \eta) - L' \ge n(\varepsilon),$$

$$2^{cn(L',\varepsilon,\eta)} - 2^{L'} \ge \widehat{n}(L', \varepsilon, \eta).$$

Assume $x \in \mathcal{E}(L', \varepsilon, \eta)$ and $L'' \ge n(L', \varepsilon, \eta)$. Since $N := 2^{cL''} - 2^{L'} \ge \widehat{n}(L', \varepsilon, \eta)$, we have

$$\begin{aligned} \sum_{j=2^{L'}+1}^{2^{cL''}} \mathbf{1}_{D \cap \mathcal{G}_\varepsilon}(\sigma^j x) &\ge \frac{1}{2} m(L')[2^{cL''} - 2^{L'}] \\ &\ge \frac{1}{2} m(L')[2^{\varepsilon n(L',\varepsilon,\eta)} - 2^{L'}] \cdot 2^{(c-\varepsilon)L''} \\ &\ge 2^{(c-\varepsilon)L''}. \end{aligned}$$

□

Let C be a cylinder of length n. If $C_n(\sigma^j x) = C$, we say that the cylinder C is seen in x at time j. Let $\varepsilon > 0$, $L' < L''$ and let D be a cylinder of length L'. For any $x \in \Sigma_2^+$, we define a finite tree, denoted $\mathcal{T}(x, D, L', L'', \varepsilon)$, as follows:

- the nodes of the finite tree $\mathcal{T}(x, D, L', L'', \varepsilon)$ are all those cylinders $D \star G'$, where G' is a $(\ell - L', \varepsilon)$-cylinder with $L' + n(\varepsilon) \le \ell \le L''$, each of which contains at least one $(L'', 2\varepsilon)$-cylinder seen in x at a moment between the time $2^{L'} + 1$ and the time $2^{cL''}$;
- an ℓ-cylinder $D \star G' \in \mathcal{T}(x, D, L', L'', \varepsilon)$ is the parent of an $(\ell+1)$-cylinder $D \star G'' \in \mathcal{T}(x, D, L', L'', \varepsilon)$ if and only if $G'' \subset G'$.

Fix $L' < L''$. For $L' + n(\varepsilon) \le \ell \le L''$, denote

$$T(x, D, \ell, \varepsilon) := \sharp\{D \star G' \in \mathcal{T}(x, D, L', L'', \varepsilon) : |D \star G'| = \ell\}.$$

Theorem 8.3 implies that if L'' satisfies the condition of Theorem 8.3 and if $x \in \mathcal{E}(L', \varepsilon, \eta)$, then in between the times $2^{L'} + 1$ and $2^{cL''}$, for each L'-cylinder D we can see at least $2^{(c-\varepsilon)L''}$ cylinders of length L'' in x of the form

$$D \star G' \qquad (G' \in \mathcal{C}_{L''-L',\varepsilon}). \tag{8·1}$$

By the quasi-Bernoulli property (3·8), it is easy to see that if L'' is sufficiently larger than L' then the cylinders $D \star G'$ are good in the sense

$$G := D \star G' \in \mathcal{C}_{L'',2\varepsilon}. \tag{8·2}$$

Thus we have

$$T(x, D, L'', \varepsilon) \geq 2^{(c-\varepsilon)L''}.$$

Next we will prove that with big probability, for all $L' + n(\varepsilon) \leq \ell \leq L''$

$$T(x, D, \ell, \varepsilon) \geq 2^{(c-2\varepsilon)\ell}.$$

8.2. **Trees associated to a typical orbit.**

Assume that $L'' \geq n(L', \varepsilon, \eta)$. Let $L' + n(\varepsilon) \leq \ell \leq L''$, $x \in \mathcal{E}(L', \varepsilon, \eta)$, and D be a L'-cylinder. By definition $T(x, D, \ell, \varepsilon)$ is the number of *different cylinders* of the form

$$D \star G' \quad \text{with } G' \in \mathcal{C}_{\ell - L', \varepsilon}$$

each of which contains at least one $(L'', 2\varepsilon)$-cylinder belonging to the list $C_{L''}(\sigma^j x)$, $2^{L'} + 1 \leq j \leq 2^{cL''}$.

Theorem 8.4. *There exists $n_0(\varepsilon)$ such that for sufficiently large L'' and for $L' + n_0(\varepsilon) \leq \ell \leq L''$ we have*

$$\mu_\phi \left\{ x \in \mathcal{E}(L', \varepsilon, \eta) : T(x, D, \ell, \varepsilon) \leq 2^{(c-2\varepsilon)(\ell - L')} \right\} \leq 2^{-2^{(c-2\varepsilon)L''}}.$$

In the rest of this subsection and the next two subsections we prepare for the proof of this theorem, which will be presented in the subsection 8.5. We need to estimate the measures

$$\mu_\phi \{ x \in \mathcal{E}(L', \varepsilon, \eta) : T(x, D, \ell, \varepsilon) = K \}$$

for $K \leq 2^{(c-2\varepsilon)(\ell - L')}$. We will do that in the following.

For $1 \leq t \leq L'' + d$ (where $d := \omega L''$ with ω defined in Theorem 3.1), let

$$\Lambda_t = \left\{ 2^{L'} + k(L'' + d) + t : 0 \leq k \leq \frac{2^{cL''} - 2^{L'}}{L'' + d} \right\}.$$

Fix K cylinders $C_1, \cdots, C_K \in \mathcal{C}_{\ell - L', \varepsilon}$. Let

$$\Upsilon_t(x; C_1, C_2, \cdots, C_K) = \sharp \left\{ j \in \Lambda_t : C_{L''}(\sigma^j x) \in \mathcal{C}_{L'', 2\varepsilon} \text{ implies } C_{L''}(\sigma^j x) \subset D \star \widetilde{C} \right\}$$

where

$$D \star \widetilde{C} := \bigcup_{i=1}^{K} D \star C_i.$$

$T(x, D, \ell, \varepsilon) = K$ means there exist exactly K different $(\ell - L', \varepsilon)$-cylinders, say $C_1, C_2, \cdots, C_K$ such that all $(L'', 2\varepsilon)$-cylinders seen in x in between the times $2^{L'} + 1$ and $2^{cL''}$ are contained in some of the $D \star C_i$'s, i.e. contained in $D \star \widetilde{C}$. On the other hand, by Theorem 8.3, there are at least $2^{(c-\varepsilon)L''}$ of the $(L'', 2\varepsilon)$-cylinders seen in x in between the times $2^{L'} + 1$ and $2^{cL''}$.

So, for at least one t the number of the $(L'', 2\varepsilon)$-cylinders seen at moments belonging to Λ_t and contained in $D \star \widetilde{C}$ is at least $\dfrac{2^{(c-\varepsilon)L''}}{L''+d}$. Thus we get

$$\{x \in \mathcal{E}(L', \varepsilon, \eta) : T(x, D, \ell, \varepsilon) = K\} \subset \bigcup_{t=1}^{L''+d} \bigcup_{C_1, \cdots, C_K} E_t(C_1, \cdots, C_K)$$

where the second union is taken over all possible collections $C_1, \cdots, C_K$ of $(\ell - L', \varepsilon)$-cylinders, and where

$$E_t(C_1, \cdots, C_K) = \left\{ x \in \mathcal{E}(L', \varepsilon, \eta) : \Upsilon_t(x; C_1, C_2, \cdots, C_K) \geq \frac{2^{(c-\varepsilon)L''}}{L''+d} \right\}.$$

Therefore, using the fact that the number of $(\ell - L', \varepsilon)$-cylinders is at most $2^{(h+\varepsilon)(\ell-L')}$, we have proved

Lemma 8.5.

$$\begin{aligned} & \mu_\phi(x \in \mathcal{E}(L', \varepsilon, \eta) : T(x, D, \ell, \varepsilon) = K) \\ \leq \ & (L''+d) \binom{2^{(h+\varepsilon)(\ell-L')}}{K} \sup_{t; C_1, \cdots, C_K} \mu_\phi(E_t(C_1, \cdots, C_K)). \end{aligned}$$

8.3. **Generalized quasi Bernoulli property.**

In order to estimate the measure $\mu_\phi(E_t(C_1, \cdots, C_K))$, we need the following generalized quasi Bernoulli property.

Let A be any cylinder and $L \geq 1$ be any integer. For $x \in A$, we define $k_A := \inf\{k \geq 0 : C_L(\sigma^{|A|+k(L+d(L))}x) \in \mathcal{C}_{L,\varepsilon}\}$ and let

$$\iota_A(x) = |A| + k_A(L + d(L))$$

where $d(L) = \omega L$ for ω defined in Theorem 3.1.

Lemma 8.6 (Generalized quasi Bernoulli property)**.** *Let A be any cylinder, $G \in \mathcal{C}_{L,\varepsilon}$ and ι_A be defined as above. Then*

$$\mu_\phi(x \in A : C_L(\sigma^{\iota_A(x)}x) = G) \leq \frac{\gamma^3}{1-2\varepsilon} \mu_\phi(A)\mu_\phi(G).$$

Proof. Notice that

$$\{x \in A : C_L(\sigma^{\iota_A(x)}x) = G\} = \bigcup_{i=0}^{\infty} A_i$$

where

$$A_i = \{x \in A : C_L(\sigma^{\iota_A(x)}x) = G, \iota_A(x) = |A| + i(L+d)\}.$$

For $i = 0$, we have

$$A_0 = A \star G.$$

So, by the Gibbs property (3·7) we get

$$\mu_\phi(A_0) \leq \gamma^3 \mu_\phi(A)\mu_\phi(G).$$

For $i \geq 1$, we have

$$A_i \subset \bigcup_{B_1,\cdots,B_i \notin \mathcal{C}_{L,\varepsilon}} A \star B_1 \star_d \cdots \star_d B_i \star_d G.$$

So, by the higher order ϕ-mixing (3·9) we get

$$\mu_\phi(A_i) \leq \gamma^3 (1+\beta^d)^i \mu_\phi(A)\mu_\phi(G) \left(\sum_{B \notin \mathcal{C}_{L,\varepsilon}} \mu_\phi(B) \right)^i.$$

Since $\sum_{B \notin \mathcal{C}_{L,\varepsilon}} \mu_\phi(B) \leq \mu_\phi(\mathcal{G}^c_{L,\varepsilon}) \leq \varepsilon$, we get

$$\mu_\phi(A_i) \leq \gamma^3 (\varepsilon(1+\beta^d))^i \mu_\phi(A)\mu_\phi(G).$$

Thus

$$\begin{aligned} \mu_\phi(x \in A : C_L(\sigma^{\iota_A(x)}x) = G) &\leq \gamma^3 \mu_\phi(A)\mu_\phi(G) \sum_{i=0}^{\infty} (\varepsilon(1+\beta^d))^i \\ &= \frac{\gamma^3}{1-\varepsilon(1+\beta^d)} \mu_\phi(A)\mu_\phi(G). \end{aligned}$$

We finish the proof by observing that $\beta < 1$. □

8.4. **Estimation of** $\mu_\phi(E_t(C_1, \cdots, C_K))$**.**

Let t be fixed. We define inductively

$$\begin{aligned} \iota_1(x) &= \inf\{j \in \Lambda_t : C_{L''}(\sigma^j x) \in \mathcal{C}_{L'',2\varepsilon}\}; \\ \iota_{k+1}(x) &= \inf\{j \in \Lambda_t : j > \iota_k(x); C_{L''}(\sigma^j x) \in \mathcal{C}_{L'',2\varepsilon}\}. \end{aligned}$$

Let

$$\widetilde{n} := \frac{2^{(c-\varepsilon)L''}}{L''+d}. \tag{8·3}$$

We have

$$\iota_i(x) < \infty \quad \text{if } x \in E_t(C_1, \cdots, C_K), \text{and} \ \text{ if } i \leq \widetilde{n}.$$

Then

$$\mu_\phi(E_t(C_1, \cdots, C_K)) \leq \sum \mu_\phi\left(x : \sigma^{\iota_i(x)}x \in F_i, 1 \leq \forall i \leq \widetilde{n}\right) \tag{8·4}$$

where the sum is taken over all F_i's with the property

$$F_i \in \mathcal{C}_{L'',2\varepsilon}, \quad F_i \subset D \star \widetilde{C} \quad (1 \leq \forall i \leq \widetilde{n}).$$

Lemma 8.7. *Let $n \geq 1$ and let $F_i \in \mathcal{C}_{L'',2\varepsilon}$ with $1 \leq i \leq n$. We have*

$$\mu_\phi\left(x : C_{L''}(\sigma^{\iota_i(x)}x) = F_i; i = 1, 2, \cdots, n\right) \leq \left(\frac{\gamma^3}{1-4\varepsilon}\right)^n \prod_{i=1}^n \mu_\phi(F_i).$$

Proof. We prove it by induction on n. Let

$$\mathcal{Q}_n = \{x : C_{L''}(\sigma^{\iota_i(x)}x) = F_i; i = 1, 2, \cdots, n\}.$$

Write

$$\mathcal{Q}_{n+1} = \mathcal{Q}_n \cap \{x : C_{L''}(\sigma^{\iota_{n+1}(x)}x) = F_{n+1}\}.$$

Notice that $\mathcal{Q}_n$ is a disjoint union of cylinders, say

$$\mathcal{Q}_n = \bigcup A_j.$$

Furthermore if $x \in A_j$ we have

$$C_{L''}(\sigma^{\iota_{n+1}(x)}x) = F_{n+1} \iff C_{L''}(\sigma^{\iota_{A_j}(x)}x) = F_{n+1}.$$

Thus, using the generalized Bernoulli property (Lemma 8.6), we have

$$\begin{aligned}
\mu_\phi(\mathcal{Q}_{n+1}) &= \sum_j \mu_\phi(x \in A_j, C_{L''}(\sigma^{\iota_{A_j}(x)}x) = F_{n+1}) \\
&\le \frac{\gamma^3}{1-4\varepsilon} \sum_j \mu_\phi(A_j)\mu_\phi(F_{n+1}) \\
&= \frac{\gamma^3}{1-4\varepsilon} \mu_\phi(\mathcal{Q}_n)\mu_\phi(F_{n+1}).
\end{aligned}$$

□

Lemma 8.8.

$$\mu_\phi(E_t(C_1, \cdots, C_K)) \le \left(2\gamma^6 K 2^{(-h+\varepsilon)(\ell-L')}\right)^{\frac{2^{(c-\varepsilon)L''}}{L''+d}}.$$

Proof. By the last lemma, we have

$$\mu_\phi(E_t(C_1, \cdots, C_K)) \le \left(\frac{\gamma^3}{1-4\varepsilon}\right)^{\tilde{n}} \sum_{F_1, \cdots, F_{\tilde{n}}} \prod_{i=1}^{\tilde{n}} \mu_\phi(F_i)$$

where the sum is taken over all collections $F_1, \ldots, F_n$ consisting of different $(L'', 2\varepsilon)$-cylinders contained in $D \star \widetilde{C}$. Recall that $\tilde{n}$ is defined in (8·3).

Since $\mu_\phi(D \star C_i) \le \gamma^3 \mu_\phi(D)\mu_\phi(C_i)$ and $\mu_\phi(C_i) \le 2^{(-h+\varepsilon)(\ell-L')}$, we have

$$\sum_{F \in \mathcal{C}_{L'',2\varepsilon}, F \subset D \star \widetilde{C}} \mu_\phi(F) \le \mu_\phi(D \star \widetilde{C}) \le K\gamma^3 2^{(-h+\varepsilon)(\ell-L')}.$$

So,

$$\mu_\phi(E_t(C_1, \cdots, C_K)) \le \left(\frac{\gamma^6}{1-4\varepsilon} K 2^{(-h+\varepsilon)(\ell-L')}\right)^{\tilde{n}}.$$

□

8.5. **Number of branches of a tree: Proof of Theorem 8.4.**
By Lemmas 8.5 and 8.8, we have

$$\mu_\phi\Big(x \in \mathcal{E}(L', \varepsilon, \eta) : T(x, D, \ell, \varepsilon) = K\Big)$$

$$\leq \quad (L'' + d) \begin{pmatrix} 2^{(h+\varepsilon)(\ell - L')} \\ K \end{pmatrix} \Big(2\gamma^6 K 2^{-(h-\varepsilon)(\ell-L')}\Big)^{\frac{2^{(c-\varepsilon)L''}}{L''+d}}. \tag{8·5}$$

For $K \leq 2^{(c-2\varepsilon)(\ell - L')}$ and for $\ell \leq L''$, we have on one hand

$$\begin{pmatrix} 2^{(h+\varepsilon)(\ell - L')} \\ K \end{pmatrix} \leq 2^{(h+\varepsilon)(\ell-L')K} \leq 2^{(h+\varepsilon)L'' 2^{(c-2\varepsilon)L''}}; \tag{8·6}$$

and on the other hand

$$K 2^{-(h-\varepsilon)(\ell - L')} \leq 2^{(c-h-\varepsilon)(\ell-L')},$$

which implies that there exists an integer $n_0(\varepsilon)$ such that if $\ell - L' \geq n_0(\varepsilon)$ we have

$$2\gamma^6 K 2^{-(h-\varepsilon)(\ell-L')} \leq \frac{1}{2}, \quad \text{i.e.} \;\; 2\gamma^6 2^{-(h-c-\varepsilon)(\ell-L')} \leq \frac{1}{2}. \tag{8·7}$$

So, from (8·5), (8·6) and (8·7) we get

$$\mu_\phi\Big(x \in \mathcal{E}(L', \varepsilon, \eta) : T(x, D, \ell, \varepsilon) = K\Big)$$

$$\leq \quad (L'' + d) \cdot 2^{(h+\varepsilon)L'' 2^{(c-2\varepsilon)L''} - \frac{2^{(c-\varepsilon)L''}}{L''+d}}. \tag{8·8}$$

Choose L'' sufficiently large so that

$$(h + \varepsilon)L'' 2^{(c-2\varepsilon)L''} \leq \frac{1}{2} \cdot \frac{2^{(c-\varepsilon)L''}}{L'' + d}. \tag{8·9}$$

From (8·8) and (8·9), we get

$$\mu_\phi(x \in \mathcal{E}(L', \varepsilon, \eta) : T(x, D, \ell, \varepsilon) = K) \leq (L'' + d) \cdot 2^{-\frac{2^{(c-\varepsilon)L''}}{2(L''+d)}}.$$

Summing over all $K \leq 2^{(c-2\varepsilon)(\ell-L')}$, we obtain

$$\mu_\phi\left\{x \in \mathcal{E}(L', \varepsilon, \eta) : T(x, D, \ell, \varepsilon) \leq 2^{(c-2\varepsilon)(\ell-L')}\right\}$$

$$\leq \quad (L'' + d) \cdot 2^{(c-2\varepsilon)(\ell-L')} \cdot 2^{-\frac{2^{(c-\varepsilon)L''}}{2(L''+d)}} \leq 2^{-2^{(c-2\varepsilon)L''}}$$

for large L'', because $2^{-\frac{2^{(c-\varepsilon)L''}}{2(L''+d)}}$ tends to zero superexponentially fast.

8.6. **Homogeneity of trees.** We have just studied the finite random tree $\mathcal{T}(x, D, L', L'', \varepsilon)$. Recall that D is an L'-cylinder and that the nodes of the tree are those cylinders contained in D each of which contains some $(L'', 2\varepsilon)$-cylinder visited by some $\sigma^j x$ for $2^{L'} + 1 \leq j \leq 2^{cL''}$. It has been proved that with a big probability the tree has at least $2^{(c-2\varepsilon)(L''-L')}$ branches (i.e. different visited $(L'', 2\varepsilon)$-cylinders). More precise, Theorem 8.4 implies

$$\mu_\phi(A(\varepsilon, L', L'', D)|\mathcal{E}(L', \varepsilon, \eta)) \geq 1 - \frac{2^{-2^{(c-2\varepsilon)L''}}}{\mu_\phi(\mathcal{E}(L', \varepsilon, \eta))} \tag{8·10}$$

if L'' is sufficiently large, where

$$A(\varepsilon, L', L'', D) = \left\{x \in \Sigma_2^+ : T(x, D, L', L'', \varepsilon) \geq 2^{(c-2\varepsilon)(L''-L')}\right\}.$$

Recall $T(x, D, L', L'', \varepsilon)$ is the number of branches of the tree $\mathcal{T}(x, D, L', L'', \varepsilon)$. We are going to show that this tree share a kind of homogeneity with big probability, as stated in the following proposition.

For an ℓ-cylinder D' contained in D ($L' \leq \ell \leq L''$), we denote

$$T(x, D, L', L'', \varepsilon; D') = \sharp\{\text{branches of } \mathcal{T}(x, D, L', L'', \varepsilon) \text{ passing by } D'\}.$$

Fix an large integer $\omega > 1$ appearing in Theorem 3.1 and an integer $\Delta > 1$ such that

$$h2^{(c-3\varepsilon)\Delta} \geq 4$$

where h is the entropy of the Gibbs measure μ_ϕ. Define $H(D, L', L'', \varepsilon)$ to be the set of $x \in \Sigma_2^+$ such that

$$T(x, D, L', L'', \varepsilon; D') \leq (1+\omega)L''2^{(c-3\varepsilon)(L''-\ell)}$$

for all D' with $L' \leq |D'| = \ell \leq L'' - \Delta$.

Proposition 8.9. *Let $\omega > 1$ be a fixed sufficiently large number depending on the Gibbs measure μ_ϕ. Let $\Delta > 1$ such that $h2^{(c-3\varepsilon)\Delta} \geq 4$. Then*

$$\mu_\phi(H(D, L', L'', \varepsilon)) \geq 1 - 2^{-(L''-L')} \tag{8·11}$$

when $L'' - L'$ are sufficiently large.

Proof. Let $d = \omega L''$ and Λ_t be the same as in the last subsection, where $1 \leq t \leq L'' + d$. We consider the subtree $\mathcal{T}_t(x, D, L', L'', \varepsilon)$ of $\mathcal{T}(x, D, L', L'', \varepsilon)$ consisting of those $(L'', 2\varepsilon)$-cylinders visited by $\sigma^j x$ with $j \in \Lambda_t$. This is clear that

$$H(D, L', L'', \varepsilon)^c \subset \bigcup_{t=1}^{(1+\omega)L''} B_t$$

where B_t is the set of x such that on the subtree $\mathcal{T}_t(x, D, L', L'', \varepsilon)$ at least one node represented by a ℓ-cylinder contains at least $2^{c(L''-\ell)}$ branches.

In the following we fix D, L', L'', ε and t. For $k \geq 1$ and $N \geq 2^{(c-3\varepsilon)(L''-\ell)}$, define $B(\ell, k, N)$ to be the set of $x \in \Sigma_2^+$ such that there exactly k cylinders of length ℓ contained in D each of which contains exactly N different $(L'', 2\varepsilon)$-cylinders visited by $\sigma^j x$ for $j \in \Lambda_t$. Then

$$B_t \subset \bigcup_{\ell=L'}^{L''-\Delta} \bigcup_{k=1}^{2^{L''-\ell}} \bigcup_{N \geq 2^{(c-3\varepsilon)(L''-\ell)}} B(\ell, k, N).$$

If $x \in B(\ell, k, N)$, then there are kN different $(L'', 2\varepsilon)$-cylinders visited at different times belonging to Λ_t. So, We have

$$\mu_\phi(B(\ell, k, N)) \leq \left(\begin{array}{c} 2^{\ell - L'} \\ k \end{array} \right) 2^{-(h-2\varepsilon)(L''-L')kN}.$$

Here we have used the fact that there are $2^{\ell - L'}$ ℓ-cylinders contained in D and the higher order ϕ-mixing property of the Gibbs measure μ_ϕ. It follows that

$$\begin{aligned} \sum_{N \geq 2^{(c-3\varepsilon)(L''-\ell)}} \sum_{k=1}^{2^{\ell-L'}} \mu(B(\ell, k, N)) &\leq e \sum_{N \geq 2^{(c-3\varepsilon)(L''-\ell)}} 2^{(\ell-L')-(h-2\varepsilon)(L''-L')N} \\ &\leq e 2^{(\ell-L')-(h-2\varepsilon)(L''-L')2^{(c-3\varepsilon)(L''-\ell)}}. \end{aligned}$$

For the first inequality, we have used the following simple fact

$$\sum_{k=1}^{m} \left(\begin{array}{c} m \\ k \end{array} \right) a^k = (1+a)^m - 1 \leq e^{am} - 1 \leq eam \quad (\text{if } am \leq 1).$$

Therefore, since $\ell \leq L'' - \Delta$ implies $h 2^{(c-3\varepsilon)(L''-\ell)} \geq 4$, we have

$$\sum_{\ell = L'}^{L''-\Delta} \sum_{N \geq 2^{(c-3\varepsilon)(L''-\ell)}} \sum_{k=1}^{2^{\ell-L'}} \mu(B(\ell, k, N)) \leq e(L'' - L' - \Delta) 2^{-2(L''-L')}.$$

Observe that this estimate is independent of t. So, we get

$\mu(H(D, L', L'', \varepsilon)^c) \leq (1+\omega) L'' \cdot e(L'' - L' - \Delta) 2^{-2(L''-L')} \leq \ 2^{-(L''-L')}$.

if $L'' - L'$ is sufficiently large. □

8.7. The Cantor set and lower bound of $\dim_H\{y : \alpha(x, y) \leq c\}$.

The next theorem is an improvement of the mass transference principle [BV] to the multifractal measure μ_ϕ.

Theorem 8.10. (Multifractal mass transference principle) *For $0 < c < h_{\mu_\phi}$, and for μ_ϕ-a.e. x we have*

$$h_{\text{top}} \{y \,:\, \alpha(x, y) \leq c\} = c.$$

We have only to show the lower bound of the entropy.

The set $A(\varepsilon, \eta, L', L'', D)$ was defined for each L'-cylinder D in the last subsection. Define now

$$A(\varepsilon, \eta, L', L'') = \bigcap_{D : |D| = L'} A(\varepsilon, \eta, L', L'', D).$$

The following corollary is a consequence of (8·10) and of (8·11). Recall that (8·10) is implied by Theorem 8.4

Corollary 8.11. *There exists an integer $n_0(\varepsilon)$ such that for $L'' \geq n(\varepsilon, \eta, L')$ we have*

$$\mu_\phi(\mathcal{E}(L', \varepsilon, \eta) \cap A(\varepsilon, \eta, L', L'') \cap H(\varepsilon, L', L'')) \geq 1 - \eta - 2^{L' - 2^{(c-2\varepsilon)L''}} - 2^{-L''}.$$

Proof. Remark that if $A_1, A_2, \cdots, A_n$ are events on a probability space such that $P(A_j) \geq 1 - a_j$ for all $1 \leq j \leq n$, then

$$P\Big(\bigcap_{j=1}^{n} A_j\Big) \geq 1 - \sum_{j=1}^{n} a_j. \tag{8·12}$$

Applying this remark to the conditional probability $\mu_\phi(\cdot|\mathcal{E}(L', \varepsilon, \eta))$, from (8·10) we deduce that

$$\mu_\phi(A(\varepsilon, \eta, L', L'')|\mathcal{E}(L', \varepsilon, \eta)) \geq 1 - 2^{L'} \times \frac{2^{-2^{(c-2\varepsilon)L''}}}{\mu_\phi(\mathcal{E}(L', \varepsilon, \eta))}.$$

In other words,

$$\mu_\phi(\mathcal{E}(L', \varepsilon, \eta) \cap A(\varepsilon, \eta, L', L'')) \geq \mu_\phi(\mathcal{E}(L', \varepsilon, \eta)) - 2^{L' - 2^{(c-2\varepsilon)L''}}.$$

Then apply once again the remark to this estimation and (8·11). We finish the proof by noting that $\mu_\phi(\mathcal{E}(L', \varepsilon, \eta)) \geq 1 - \eta$ (see Theorem 8.3). □

Proof of Theorem 8.10. Let $\varepsilon > 0$ be an arbitrary small number. We can find an increasing sequence of integers $(L_k)_{k \geq 0}$ such that

$$L_0 = 0, \quad 2^{-L_k} + 2^{-2^{(c-2\varepsilon)L_k}} \leq \frac{\varepsilon}{2^{k+1}}. \tag{8·13}$$

and that for each $k \geq 1$, the pair $(L', L'') = (L_{k-1}, L_k)$ satisfies the condition of Theorem 8.4 and of Corollary 8.11. Apply Corollary 8.11. to $L' = L_{k-1}, L'' = L_k$ and $\eta = \frac{\varepsilon}{2^{k+1}}$. in order to get reads as

$$\mu_\phi(\mathcal{E}_k(\varepsilon)) > 1 - \frac{\varepsilon}{2^k} \tag{8·14}$$

where

$$\mathcal{E}_k(\varepsilon) := \mathcal{E}(L_{k-1}, \varepsilon, \eta_k) \cap A(L_{k-1}, L_k, \varepsilon) \cap A(L_{k-1}, L_k, \varepsilon).$$

Define

$$\mathcal{E}^*(\varepsilon) = \bigcap_{k=1}^{\infty} \mathcal{E}_k(\varepsilon).$$

From (8·14) we get $\mu_\phi(\mathcal{E}_k^*(\varepsilon))) \geq 1 - \frac{\varepsilon}{2^k}$ and

$$\mu_\phi(\mathcal{E}^*(\varepsilon)) \geq 1 - \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = 1 - \varepsilon. \tag{8·15}$$

Now, for each $x \in \mathcal{E}^*(\varepsilon)$, we construct a Cantor set as follows.

First step: for $1 \leq \ell \leq L_1$, consider the family $\mathfrak{C}_\ell(x)$ of (ℓ, ε)-cylinders which contain at least one $(L_1, 2\varepsilon)$-cylinder seen in x between the times 1 and 2^{cL_1}. This yields a tree $\mathfrak{T}_{L_1}(x)$ of height L_1. The nodes of the tree $\mathfrak{T}_{L_1}(x)$ are the (ℓ, ε)-cylinders, with $1 \leq \ell \leq L_1$, belonging to $\mathfrak{C}_\ell(x)$. The edges are defined by the containment relation. We will extend this tree inductively.

Second step: Let $k \geq 2$. Suppose that we have constructed a tree $\mathfrak{T}_{L_{k-1}}(x)$ of height L_{k-1}. We will construct a tree of height L_k. Let

$$L' = L_{k-1}, \quad L'' = L_k.$$

Fix a L'-cylinder D seen in x before time $2^{cL'}$, which is the label of a node of the tree $\mathfrak{T}_{L_{k-1}}(x)$ at level L_{k-1}. For $L'+1 \leq \ell \leq L''$, take all (ℓ, ε)-cylinders that contain at least one $(L'', 2\varepsilon)$-cylinder of the form $D \star G$ seen in x between the times $2^{L'}+1$ and $2^{cL''}$. As before we denote this family by $\mathfrak{C}_\ell(x)$ (both D and G varying). The tree $\mathfrak{T}_{L_k}(x)$ is obtained from $\mathfrak{T}_{L_{k-1}}(x)$ by adding branches to each D. That is to say, by splitting D into (ℓ, ε)-cylinders belonging to $\mathfrak{C}_\ell(x)$.

We define

$$C_\infty(x) = \bigcap_{k=1}^{\infty} \bigcap_{\ell=L_{k-1}+n_0(\varepsilon)}^{L_k} \bigcup_{C \in \mathfrak{C}_\ell(x)} C.$$

We have $C_\infty(x) \subset \{y : \alpha(x,y) \leq c\}$, since for any $y \in C_\infty(x)$ and for all $k \geq 1$

$$y \in \bigcup_{C \in \mathfrak{C}_{L_k}(x)} C,$$

i.e. $y \in C_{L_k}(\sigma^j x)$ for some

$$2^{L_{k-1}} + 1 \leq j \leq 2^{cL_k}.$$

We claim that $\dim_H C_\infty(x) \geq c - 2\varepsilon$. Define a probability measure ν on $C_\infty(x)$ as follows. First attribute an equal mass to each L_1-cylinder D_1 in $\mathfrak{C}_{L_1}(x)$. Then redistribute equally the total mass of D_1 to all its sub-L_2-cylinders in $\mathfrak{C}_{L_2}(x)$, and so on.

Observe that

$$\log_2 \sharp\mathfrak{C}_{L_k}(x) \geq \sum_{j=1}^{k} (c-2\varepsilon)(L_j - L_{j-1}) \geq (c-2\varepsilon)L_k \tag{8·16}$$

It follows that for $C \in \mathfrak{C}_{L_k}(x)$ we have

$$\nu(C) \leq \frac{1}{\sharp\mathfrak{C}_{L_k}(x)} \leq 2^{-(c-2\varepsilon)L_k}. \tag{8·17}$$

For any $L_k \leq \ell \leq L_{k+1} - \Delta$ and any $C \in \mathfrak{C}_\ell(x)$, we have

$$\begin{aligned} \nu(C) &\leq 2^{-(c-2\varepsilon)L_k} \cdot (1+\omega)L_{k+1}2^{(c-3\varepsilon)(L_{k+1}-\ell)} \cdot 2^{-(c-2\varepsilon)(L_{k+1}-L_k)} \\ &= (1+\omega)L_{k+1}2^{-\varepsilon(L_{k+1}-\ell)}2^{-(c-2\varepsilon)\ell}. \end{aligned}$$

If, furthermore, $\ell \leq L_{k+1}/2$, then $L_{k+1} \leq 2(L_{k+1}-\ell) \leq e^{\varepsilon(L_{k+1}-\ell)}$ so that

$$\nu(C) \leq (1+\omega)2^{-(c-2\varepsilon)\ell}. \tag{8·18}$$

If $\ell \geq L_{k+1}/2$, then $L_{k+1} \leq 2\ell \leq 2^{\varepsilon\ell}$ so that

$$\nu(C) \leq 2^{-(c-2\varepsilon)L_k} \cdot (1+\omega) \leq 2^{-(c-3\varepsilon)\ell}. \tag{8·19}$$

Thus, we have proved that with probability bigger than $1-\varepsilon$ the estimations (8·17), (8·18) and (8·19) holds, which implies

$$\dim_H\{y : \alpha(x,y) \le c\} \ge c - 3\varepsilon.$$

□

Remark: The proofs in this section can be used to obtain a more precise estimate on the growth rate of the tree, however this estimate is not necessary for our purpose. Namely one can show that $L_{k-1} \ll l \le \frac{c}{h} L_k$ then

$$T(x, D, l, \varepsilon) \ge 2^{(h-3\varepsilon)l}.$$

This implies that the upper box counting dimension of the corresponding Cantor set is $h - 3\varepsilon$ while the lower box dimension equals the Hausdorff dimension equals $c - 2\varepsilon$.

9. Results for the full shift

Our strategy is to prove all the theorems in the symbolic framework and then transfer them to the circle. Let us get together the already obtained results in the symbolic framework.

Corollary 9.1. *For $0 < \kappa < \infty$ we have μ_ϕ-a.e.*

$$\sup\{E(t) : \frac{1}{t} \le \kappa\} \ge \dim_H \mathcal{F}^\kappa(x) \ge \sup\{E(t) : \frac{1}{t} < \kappa\}.$$

For $\kappa \le 1/h_{\mu_\phi}$ (i.e. $1/\kappa \ge h_{\mu_\phi}$) we have μ_ϕ-a.e.

$$\sup\{E(t) : \frac{1}{t} \ge \kappa\} \ge \dim_H \mathcal{I}^\kappa(x) \ge \sup\{E(t) : \frac{1}{t} > \kappa\},$$

and for $\kappa > 1/h_{\mu_\phi}$ (i.e. $1/\kappa < h_{\mu_\phi}$) we have μ_ϕ-a.e.

$$\dim_H \mathcal{I}^\kappa(x) = 1/\kappa.$$

Proof. The first line is a consequence of Lemma 4.2, Theorem 6.3 and Theorem 3.3.

The second line is a consequence of Lemma 4.2, Theorem 7.2 and Theorem 3.3.

The third line is a direct consequence of Lemma 4.2, Theorems 7.2 and 8.10. □

Corollary 9.2. *Let $1/\kappa \in (e^-, e^+)$. Then for μ_ϕ a.e. x*

$$\dim_H \mathcal{F}^\kappa(x) = \max_{\nu-ergodic}\{h_\nu : \alpha(x,y) \le \frac{1}{\kappa}\ \nu - a.e.y\}.$$

For $1/\kappa \in (h_{\mu_\phi}, e^+)$ and μ_ϕ a.e. x

$$\dim_H \mathcal{I}^\kappa(x) = \max_{\nu-ergodic}\{h_\nu : \alpha(x,y) \ge \frac{1}{\kappa}\ \nu - a.e.y\}.$$

The properties of the entropy spectrum which were stated in the background section immediately imply the following corollary.

Corollary 9.3. *For $1/\kappa \in (e^-, e^+)$ and μ_ϕ a.e. x we have*

$$\sup_{-P'(q)\geq\frac{1}{\kappa}} [P(q\phi) - P'(q\phi)q] \geq \dim_H \mathcal{F}^\kappa(x) \geq \sup_{-P'(q)>\frac{1}{\kappa}} [P(q\phi) - P'(q\phi)q] .$$

For $1/\kappa \in (h_{\mu_\phi}, e^+)$ and μ_ϕ a.e. x we have

$$\sup_{-P'(q)\leq\frac{1}{\kappa}} [P(q\phi) - P'(q\phi)q] \geq \dim_H \mathcal{I}^\kappa(x) \geq \sup_{-P'(q)<\frac{1}{\kappa}} [P(q\phi) - P'(q\phi)q] .$$

If we consider a typical potential in the Baire sense for the Hölder topology, then the function $E(t)$ is continuous on the nontrivial interval (e^-, e^+), equals 0 on the endpoints (see [S2]). Hence the right hand side and left hand side inequalities in Corollary 9.1 and Corollary 9.3 are equal. Since the maximum value of $E(t)$ is attained at the value $t = -\int_{\Sigma_2^+} \phi \, d\mu_{\max}$ and equals $h_{\rm top}(\Sigma_2^+) = 1$ we have the following corollary.

Corollary 9.4. *For a typical potential ϕ and μ_ϕ a.e. x we have*

$$\dim_H \mathcal{F}^\kappa(x) = h_{\rm top}(\Sigma_2^+) = 1 \quad \text{for } \kappa \geq \frac{1}{-\int \phi \, d\mu_{\max}.},$$

$$\dim_H \mathcal{I}^\kappa(x) = h_{\rm top}(\Sigma_2^+) = 1 \quad \text{for } \kappa \leq \frac{1}{-\int \phi \, d\mu_{\max}.}.$$

Let q_κ be the number such that $P'(q_\kappa\phi) = -\frac{1}{\kappa}$. Then

$$\dim_H \mathcal{F}^\kappa(x) = E\left(\tfrac{1}{\kappa}\right) = P\left(q_\kappa\phi\right) + \tfrac{1}{\kappa} q_\kappa \qquad \text{for } \kappa < \frac{1}{-\int \phi \, d\mu_{\max}.},$$

$$\dim_H \mathcal{I}^\kappa(x) = E\left(\tfrac{1}{\kappa}\right) = P\left(q_\kappa\phi\right) + \tfrac{1}{\kappa} t q_\kappa \qquad \text{for } \frac{1}{h_{\mu_\phi}} \geq \kappa > \frac{1}{-\int \phi \, d\mu_{\max}.}.$$

Finally we come to the answer of the symbolic version of question (**Q2**).

Corollary 9.5. *For μ_ϕ a.e. x we have $\kappa^F_{\phi,\Sigma_2^+} = \frac{1}{e^+}$. In particular,*

$$\mathcal{F}^\kappa(x) = \varnothing \text{ for } \kappa < \frac{1}{e^+} := \frac{1}{\max_{\mu \text{ ergodic}} \int (-\phi) \, d\mu}.$$

Proof. From multifractal analysis, it is well known that

$$e^+ = \max_\nu \int (-\phi) \, d\nu = \max_{y\in\Sigma_2^+} h_{\mu_\phi}(y).$$

Therefore

$$\mathcal{F}^\kappa(x) \subset \{y : \alpha(x,y) \geq 1/\kappa \text{ and } \underline{h}_{\mu_\phi}(y) \leq e^+ < 1/\kappa\} = \varnothing$$

by Lemma 4.2, Lemma 4.3 and Theorem 6.2. □

Using the techniques developed in the previous sections we can conclude a strong theorem on the structure of typical sequences. The subword structure of a typical sequence up to time L is completely determined by the entropy spectrum of the measure. Theorem 2.1 is re-stated as the following corollary.

Corollary 9.6. *Fix $\beta > 0$. Consider $n \ll L$ sufficiently large, a typical point x and the set of cylinders C_n^p of length n satisfying $\mu_\phi(C_n^p) \sim 2^{-\beta n}$ which are subwords of the cylinder $C_L(x)$, i.e. the orbit of x hits the cylinder C_n^p before time L. Then*

$$\sharp\big(C_n^p\big) \sim 2^{\min\big(E(\beta),\ E(\beta)-\beta+(\log L)/n\big)n}.$$

Here $a_n \sim b_n$ means that the ratio a_n/b_n is subexponential in the sense that

$$\log \frac{a_n}{b_n} = o(n).$$

Recall that from the multifractal analysis, it is well known that there are $2^{nE(\beta)}$ cylinders C_n^p such that

$$\mu_\phi(C_n^p) \sim 2^{-\beta n}.$$

Proof. We sketch the proof. If $\beta < (\log L)/n$ then Lemma 6.1 shows that for μ_ϕ-a.e. x, all $2^{E(\beta)n}$ cylinders occur in the orbit of x. If $\beta > (\log L)n$ then Lemma 7.1 implies the upper bound.

To prove the lower bound the idea is to use a symbolic version of the Katok horseshoe approximation and apply some facts discussed in Section 10 (note that the results of Section 10 do not depend on this lower bound). We also use the following strengthened version of Theorem 3.1 from [FS1]. In fact, the proof given in this article completely proves this statement.

Theorem 3.1′ [FS1] *Let $e^- < \beta < e^+$. There exists a sequence of subshifts of finite type $\big\{\Sigma_{A_k}^+\big\}_{k\geq 1}$ such that for all k and for all $x \in \Sigma_{A_k}^+$*

$$\beta - \frac{1}{k} \leq -\varliminf_{m\to\infty} \frac{\log \mu_\phi(C_m(x))}{m} \leq -\varlimsup_{m\to\infty} \frac{\log \mu_\phi(C_m(x))}{m} \leq \beta + \frac{1}{k}. \tag{9·1}$$

and $0 \leq E(\beta) - h_{\text{top}}(\Sigma_{A_k}^+) \leq \frac{1}{k}$.

Thus we have

$$\Sigma_{A_k}^+ \subset \bigcup_{N=1} \bigcap_{m\geq N} \{\bigcup_p C_m^p : \mu_\phi(C_m^p) \sim 2^{-(\beta\pm\frac{1}{k})m}\}$$

for each $k \geq 1$. So, it suffices to show that we can find cylinders among those intersecting $\Sigma_{A_k}^+$ where with $k := k(\epsilon)$ is sufficiently large. By Equation (9·1),

$$\phi_k := \phi_{A_k} := \phi|_{\Sigma_{A_k}^+} - P_{A_k}(\phi|_{\Sigma_{A_k}^+})$$

has an almost degenerated entropy specturm over $\Sigma_{A_k}^+$, i.e.

$$e_{\phi_k}^+ - e_{\phi_k}^- \leq \frac{2}{k}, \qquad E(\beta) - h_{\mu_{\phi_k}} < \frac{3}{k}.$$

In particular it is almost cohomologous to a constant on $\Sigma_{A_k}^+$. Note that

$$P_{A_k}(\phi|_{\Sigma_{A_k}^+}) = h_{\mu_{\phi_k}} + \int_{A_k} \phi_k \, d\mu_{\phi_k} = h_{\text{top}}(\Sigma_{A_k}^+) - \beta + o(1) \to E(\beta) - \beta$$

as $k \to \infty$. Then apply the first result concerning $\mathcal{I}^\kappa_A(x)$ in Theorem 10.2. Or more conveniently apply the estimate (8·16) which holds even when L_k is replaced by an ℓ such that $|\ell - L_k| \geq \Delta$. Thus we get that the number of n-cylinders in Σ_{A_k} which are hit before time L is bounded from below by 2^{cn} with

$$c := \frac{\log L}{n} + P_{A_n}(\phi|_{\Sigma^+_{A_n}}) - \frac{3}{n}$$

which approaches $E(\beta) - \beta + (\log L)/n$. This finishes the proof.

□

10. Extensions to subshifts of finite type

The previous results can be extended in a canonical way to subshifts of finite type: Σ_2^+ is replaced by a subshift space Σ_A and μ_ϕ and μ_ψ by two Gibbs measures of the subsystem $\sigma : \Sigma_A \to \Sigma_A$. Extensions to symbolic spaces of several symbols are also obvious.

Here we consider another kind of extension. Given a compact subset K in Σ_2^+. What can we say about $K \cap \mathcal{I}^\kappa(x)$ and $K \cap \mathcal{F}^\kappa(x)$?

We assume that the reference measures μ_ϕ and μ_ψ are Gibbs measure of the *full shift* $\sigma : \Sigma_2^+ \to \Sigma_2^+$ (or of a subshift of finite type). We can answer this question when $K = \Sigma_A$ is a subshift of finite type. The proofs are still slight modifications of those for the full shift, thus we only sketch them briefly here. We will emphasize the differences.

Let $\Sigma_A \subset \Sigma_2^+$ be a subshift of finite type. We are interested in the following two sets:

$$\mathcal{F}^\kappa_A(x) := \mathcal{F}^\kappa(x) \cap \Sigma_A \quad \text{and} \quad \mathcal{I}^\kappa_A(x) := \mathcal{I}^\kappa(x) \cap \Sigma_A.$$

Recall that $\mu_\phi(\Sigma_A) = 0$ if $\Sigma_A \neq \Sigma_2^+$ because Σ_A is a closed invariant set ($\sigma\Sigma_A \subset \Sigma_A$) and μ_ϕ is of full support and ergodic.

The analysis of these sets is related to the determination of the following restricted entropy spectrum.

10.1. **Restricted entropy spectrum.**

By the restricted entropy spectrum we means

$$E_A(\alpha) := \dim_H\{y \in \Sigma_A : \ h_{\mu_\phi}(y) = \alpha\}.$$

Recall that μ_ϕ is a Gibbs measure on the full shift space Σ_2^+, but in the above definition we are only interested in the entropy function h_{μ_ϕ} restricted on the subshift Σ_A.

Clearly the restriction $\phi|_{\Sigma_A}$ is a Hölder potential on Σ_A. Let $P_A(\phi|_{\Sigma_A})$ be the pressure of the potential $\phi|_{\Sigma_A} : \ \Sigma_A \to \mathbb{R}$ related to the subsystem $\sigma : \Sigma_A \to \Sigma_A$. Instead of $\phi|_{\Sigma_A}$, it will be more convenient to consider

$$\phi_A(x) := \phi|_{\Sigma_A} - P_A(\phi|_{\Sigma_A}).$$

We catch reader's attention to the notation: ϕ_A, ϕ_{Σ_A}. We denote by μ_{ϕ_A} the Gibbs measure on Σ_A associated to ϕ_A. Its entropy $h_{\mu_{\phi_A}}(x)$ is well defined by

$$h_{\mu_{\phi_A}}(x) = \lim_{n\to\infty} \frac{S_n(-\phi_A)(x)}{n}$$

if the limit exists. The multifractal spectrum of μ_{ϕ_A}:

$$\widetilde{E}_A(\alpha) := \dim_H\{y \in \Sigma_A : \ h_{\mu_{\phi_A}}(y) = \alpha\}.$$

is well known. We denote by $\tilde{e}_A^+, \tilde{e}_A^-$ the maximal and minimal entropy of $h_{\mu_{\phi_A}}$. That is

$$\tilde{e}_A^+ = \sup_{\text{supp}\mu\subset\Sigma_A} \int(-\phi_A)d\mu = \sup_{\text{supp}\mu\subset\Sigma_A} \int(-\phi)d\mu + P_A(\phi|_{\Sigma_A})$$
$$\tilde{e}_A^- = \inf_{\text{supp}\mu\subset\Sigma_A} \int(-\phi_A)d\mu = \inf_{\text{supp}\mu\subset\Sigma_A} \int(-\phi)d\mu + P_A(\phi|_{\Sigma_A}).$$

Define

$$e_A^- := \inf_{\text{supp}\mu\subset\Sigma_A} \int(-\phi)d\mu, \qquad e_A^+ := \sup_{\text{supp}\mu\subset\Sigma_A} \int(-\phi)d\mu.$$

Then we have

$$e_A^- = \tilde{e}_A^- - P_A(\phi|_{\Sigma_A}), \qquad e_A^+ = \tilde{e}_A^+ - P_A(\phi|_{\Sigma_A}).$$

The following lemma establishes the relation between the two spectra $E_A(\cdot)$ and $\tilde{E}_A(\cdot)$: The graph of $E_A(\cdot)$ is that of $\tilde{E}_A(\cdot)$ shifted to the right by a distance $-P_A(\phi|_{\Sigma_A})$. In the statement of the lemma, $\approx$ means that the ratio is bounded between two constants.

Lemma 10.1. *We have*

(1) $P_A(\phi|_{\Sigma_A}) \le 0$.
(2) $\phi_A(x)$ *is normalized in the sense that* $P_A(\phi_A) = 0$.
(3) *The two Gibbs measures* μ_{ϕ_A} *and* μ_ϕ *are related by*

$$\mu_{\phi_A}(C_n(x)) \approx 2^{-nP_A(\phi|_{\Sigma_A})}\mu_\phi(C_n(x)), \quad (\forall x \in \Sigma_A, \forall n \ge 1).$$

(4) *If one of* $h_{\mu_{\phi_A}}(x)$ *and* $h_{\mu_\phi}(x)$ *is well defined, they are related by*

$$h_{\mu_{\phi_A}}(x) = h_{\mu_\phi}(x) + P_A(\phi|_{\Sigma_A}), \ x \in \Sigma_A.$$

(5) *We have*

$$E_A(\alpha) = \tilde{E}_A(\alpha + P_\alpha(\phi|_{\Sigma_A})).$$

(6) *The set* $\{y \in \Sigma_A : \ h_{\mu_\phi}(y) = \alpha\}$ *is empty unless*

$$\tilde{e}_A^- \le \alpha + P_A(\phi) \le \tilde{e}_A^+$$

(7) $E_A(\alpha) \le E(\alpha)$.

Proof. (1) a consequence of the variational principle:

$$\begin{aligned} P_A(\phi|_{\Sigma_A}) &= \max_{\mu \text{ inv on } \Sigma_A}(h_\mu + \int_{\Sigma_A} \phi\, d\mu) \\ &\le \max_{\mu \text{ inv on } \Sigma}(h_\mu + \int_{\Sigma} \phi\, d\mu) = P(\phi) = 0. \end{aligned}$$

(2) is well known. (3) is a consequence of the Gibbs property of μ_{ϕ_A} and μ_ϕ:

$$\mu_{\phi_A}(C_n(x)) \approx 2^{S_n\phi_A(x)} = 2^{S_n\phi(x)-nP_A(\phi|_{\Sigma_A})} \approx 2^{-nP_A(\phi|_{\Sigma_A}(\phi)}\mu_\phi(C_n(x))$$

(4) is a consequence of the fact that each entropy is a limit of a Birkhoff mean. (5) is a consequence of (4) which implies

$$\{y \in \Sigma_A : \ h_{\mu_\phi}(y) = \alpha\} = \{y \in \Sigma_A : \ h_{\mu_{\phi_A}}(y) = \alpha + P_A(\phi)\}.$$

The above equation also implies (6), because the set $\{y \in \Sigma_A : \ h_{\mu_{\phi_A}}(y) = \alpha + P_A(\phi_{\Sigma_A})\}$ is empty unless $\tilde{e}_A^- \leq \alpha + P_A(\phi) \leq \tilde{e}_A^+$. The inequality in (7) follows from (5) and

$$\begin{aligned}\widetilde{E}_A(\alpha + P_A(\phi)) = & \sup_{\substack{\mathrm{supp}\mu\subset\Sigma_A \\ \int(-\phi_A)d\mu=\alpha+P_A(\phi)}} h_\mu = \sup_{\substack{\mathrm{supp}\mu\subset\Sigma_A \\ \int(-\phi)d\mu=\alpha}} h_\mu \\ \leq & \sup_{\int(-\phi)d\nu=\alpha} h_\nu = E(\alpha).\end{aligned}$$

□

Let $e_A^{\max}$ be the unique value at which $E_A(\alpha)$ attains its maximum. We have

$$e_A^{\max} = -\int_{\Sigma_A} \phi d\mu_{\mathrm{Parry}}, \qquad E_A(e_A^{\max}) = \dim_H(\Sigma_A)$$

where μ_{Parry} is the Parry measure on Σ_A. Recall that

$$e_A^- \leq \alpha_A^{\max} \leq e_A^+.$$

Now we can state our results concerning $\mathcal{I}_A^\kappa(x)$ and $\mathcal{F}_A^\kappa(x)$. Remark that unlike the full shift case, $\mathcal{I}_A^\kappa(x)$ is empty for large κ.

Theorem 10.2. *For $\mathcal{F}_A^\kappa(x)$, μ_ϕ-a.e. x we have*

$$\dim_H \mathcal{F}_A^\kappa(x) = \begin{cases} \dim_H(\Sigma_A) & \text{if } \frac{1}{\kappa} \leq e_A^{\max}, \\ E_A(\frac{1}{\kappa}) & \text{if } \frac{1}{\kappa} > e_A^{\max}, \end{cases}$$

$$\mathcal{F}_A^\kappa(x) = \varnothing \qquad \text{if } \frac{1}{\kappa} > e_A^+.$$

For $\mathcal{I}_A^\kappa(x)$, μ_ϕ-a.e. x we have

$$\dim_H \mathcal{I}_A^\kappa(x) = \begin{cases} \frac{1}{\kappa} + P_A(\phi|_{\Sigma_A}) & \text{if } -P_A(\phi|_{\Sigma_A}) \leq \frac{1}{\kappa} \leq h_{\mu_{\phi_A}} - P_A(\phi|_{\Sigma_A}) \\ E_A(\frac{1}{\kappa}) & \text{if } h_{\mu_{\phi_A}} - P_A(\phi|_{\Sigma_A}) \leq \frac{1}{\kappa} \leq e_A^{\max} \\ \dim_H(\Sigma_A) & \text{if } \frac{1}{\kappa} \geq e_A^{\max} \end{cases}$$

$$\mathcal{I}_A^\kappa(x) = \varnothing \quad \text{if } \frac{1}{\kappa} < -P_A(\phi|_{\Sigma_A}).$$

Proof. The proof is the same as before. Recall that Lemma 4.2 transforms the problem to the study of the return rate $\alpha(x, y)$, with y restricted onto Σ_A and that the study of $\alpha(x, y)$ is reduced to that of the entropy $h_{\mu_\phi}(y)$ which is related to the entropy $h_{\mu_{\phi_A}}(y)$, by Lemma 10.1.

The only statement, the last statement, which differs from the full shift case is that $\mathcal{I}_A^\kappa(x)$ may be empty. Let us prove this statement. Fix an $\varepsilon > 0$ such that

$$\frac{1}{\kappa} + \varepsilon < -P_A(\phi|_{\Sigma_A}).$$

First, by Lemma 10.1 (6), we have

$$h_{\mu_\phi}(y) \geq \tilde{e}_A^- - P_A(\phi|_{\Sigma_A}), \quad \text{i.e.} \quad h_{\mu_{\phi_A}}(y) \geq \tilde{e}_A^-$$

for all $y \in \Sigma_A$. Then, by Lemma 4.2,

$$\begin{aligned} \mathcal{I}_A^\kappa(x) &\subset \left\{ y \in \Sigma_A : \ \alpha(x, y) < \frac{1}{\kappa} + \epsilon \right\} \\ &= \{ y \in \Sigma_A : \ \alpha(x, y) < a, \ h_{\mu_{\phi_A}}(y) \geq \tilde{e}_A^- \} = \bigcup_{j=0}^{\infty} S_j \end{aligned}$$

where $a = -P_A(\phi|_{\Sigma_A})$ and

$$S_j = \{ y \in \Sigma_A : \alpha(x, y) < a, \ h_{\mu_{\phi_A}}(y) \in \big[j\varepsilon, (j+1)\varepsilon \big) + \tilde{e}_A^- \}.$$

Thus, Lemma 7.1 applied to

$$K = 2^{an}, \quad L = \max\{ 2^{nE_A(\tilde{e}_A^- + j\varepsilon)}, 2^{nE_A(\tilde{e}_A^- + (j+1)\varepsilon)} \}, \quad N = 1$$

implies that each set S_j is empty for μ_ϕ-a.e. x. □

11. Transferring to the circle

In this section we show that the results of the section 9 hold for the doubling map of the circle, i.e. replacing $\mathcal{F}^\kappa(x), \mathcal{I}^\kappa(x)$ by $F^\kappa(s), I^\kappa(s)$. Recall that the projection $\pi : \Sigma_2^+ \to \mathbb{S}$ was defined in the section 3. For $y \in \Sigma_2^+$ let $C_n^-(y)$ denote the cylinder of length n preceding $C_n(y)$ in the lexicographical order and $C_n^+(y)$ denote the immediate successor whenever they are defined. Note that $C_n^-(0^\infty)$and $C_n^+(1^\infty)$ are non defined for all n, while for any other point $y \in \Sigma_2^+$ $C_n^\pm(y)$ is well defined for sufficiently large n. Let

$$C_n^*(y) := C_n^-(y) \cup C_n(y) \cup C_n^+(y)$$

when it is defined.

Theorem 11.1. *For μ_ϕ a.e. x we have*

$$\dim_H(F^\kappa \pi(x)) = \dim_H \pi(\mathcal{F}^\kappa(x))$$

$$\dim_H(I^\kappa \pi(x)) = \dim_H \pi(\mathcal{I}^\kappa(x)).$$

Proof. For $x \in \Sigma_2^+$ with $x \neq 1^\infty, 0^\infty$, the projection of each of the cylinders $C_n^-(x)$, $C_n(x)$,$C_n^+(x)$ to $\mathbb{S}^1$ is an interval around $\pi(x)$ (for n sufficiently large). Moreover we have

$$\pi(C_{\lfloor\kappa\log n\rfloor+1}(x)) \subset \left(\pi(x) - \frac{1}{n^\kappa}, \pi(x) + \frac{1}{n^\kappa}\right) \subset \pi(C^*_{\lfloor\kappa\log n\rfloor}(x)). \tag{11·1}$$

Applying the left inclusion, it follows that

$$F^\kappa(\pi(x)) \subset \pi(\mathcal{F}^\kappa(x)).$$

Hence

$$\dim_H(F^\kappa\pi(x)) \leq \dim_H \pi(\mathcal{F}^\kappa(x)),$$

and similarly

$$\dim_H(I^\kappa\pi(x)) \geq \dim_H \pi(\mathcal{I}^\kappa(x)).$$

We turn to the reverse inequalities. For this we define

$$\tau_n^*(x,y) := \inf\{l \geq 1 \,:\, \sigma^l x \in C_n^*(y)\},$$

$$\tau_n^-(x,y) := \inf\{l \geq 1 \,:\, \sigma^l x \in C_n^-(y)\}$$

and

$$\tau_n^+(x,y) := \inf\{l \geq 1 \,:\, \sigma^l x \in C_n^+(y)\}$$

then

$$\tau_n^*(x,y) = \min\{\tau_n^-(x,y), \tau_n(x,y), \tau_n^+(x,y)\}$$

and

$$\alpha^*(x,y) = \min\{\alpha^-(x,y), \alpha(x,y), \alpha^+(x,y)\}$$

where $\alpha^*, \alpha^-, \alpha^+$ are defined in the corresponding way. Therefore in analogy to Lemma 4.2

$$\left\{\pi(y) \,:\, \alpha^*(x,y) > \frac{1}{\kappa}\right\} \subset F^\kappa(\pi(x)) \tag{11·2}$$

and

$$I^\kappa(\pi(x)) \subset \left\{\pi(y) \,:\, \alpha^*(x,y) \leq \frac{1}{\kappa}\right\}. \tag{11·3}$$

Next we need the following lemma to prove the reverse inequalities.

Lemma 11.2. *For any $x \in \Sigma_2^+$ and ν an ergodic Borel probability measure different from δ_{0^∞} and δ_{1^∞} we have*

$$\alpha^*(x,y) = \alpha(x,y) \qquad \nu - a.e.$$

Proof. We will prove that $\alpha^+(x,y) \geq \alpha(x,y)$ almost everywhere. The proof for $\alpha^-(x,y) \geq \alpha(x,y)$ a.e. is similar. Since

$$\alpha^*(x,y) = \min\{\alpha^-(x,y), \alpha(x,y), \alpha^+(x,y)\},$$

this will imply the lemma.

Fix $\epsilon > 0$. Let $\mathbf{1}_n$ be the characteristic function of the cylinder set consisting of n 1's. Since ν is not concentrated on 1^∞ we can find an n_ϵ sufficiently large that

$$\int \mathbf{1}_n(x)\, d\nu(x) < \epsilon \qquad (\forall n > n_\epsilon).$$

Now let y be a generic point for ν. Then there is an $n_0 = n_0(y) > n_\epsilon$ such that the Birkhoff sum

$$\frac{1}{m} S_m \mathbf{1}_n(y) < \epsilon \qquad (\forall m > n_0).$$

Let us consider the structure of $C_m^+(y)$.

$$\begin{array}{rlll} C_m^+(y) = & [y_1 \cdots y_{m-1} 1] & \text{if} & y = y_1 \cdots\cdots\cdots y_{m-1} 0 y_{m+1} \cdots \\ C_m^+(y) = & [y_1 \cdots y_{k-1} 100 \cdots 0] & \text{if} & y = y_1 \cdots y_{k-1} 011 \cdots 1 y_{m+1} \cdots . \end{array}$$

It follows that

$$C_m^+(y) \subset C_{k-1}(y)$$

where $k = k(y, m)$ is characterized by $y_k = 0$ and $y_j = 1$ $(\forall k < j \leq m\})$. Thus

$$\tau_m^+(x, y) \geq \tau_{k-1}(x, y). \tag{11·4}$$

For a given x, the more 1's at the end of $C_m(y)$ is the only way to enlarge the difference of x's hitting times of $C_m^+(y)$ and $C_m(y)$. Let $n > n_0$, $m > n - l - 1$ and assume that we have a block of $n + l$ ones at the end $(l > n)$. Then (11·4) becomes

$$\tau_m^+(x, y) \geq \tau_{m-n-l-1}(x, y).$$

These sequence differ most at the moment where long blocks of 1's in y. The worst situation is when this block occurs very early. We are going to estimate this first occurrence. First we observe that

$$\epsilon > \frac{1}{m} S_m \mathbf{1}_n(y) \geq \frac{l}{m}.$$

This implies that the first occurrence of the block in question is not earlier than

$$m - n - l - 1 \geq m - 2l > m(1 - 2\epsilon).$$

Therefore

$$\alpha^+(x, y) = \varliminf_{m\to\infty} \frac{\log \tau_m^+(x, y)}{m} \geq \varliminf_{m\to\infty} \frac{\log \tau_{m-n-l-1}(x, y)}{m} \geq (1 - 2\epsilon)\alpha(x, y).$$

Letting $\epsilon \to 0$ we obtain the result. □

We continue with the proof of the theorem. For any Borel set A we have

$$\dim_H \pi A = h_{\text{top}}(A)$$

since $\operatorname{diam} \pi(C) = 2^{-|C|}$ for any cylinder set C. Thus applying Theorem 6.3 yields

$$\dim_H \pi(\mathcal{F}^\kappa(x)) = h_{\text{top}}(\mathcal{F}^\kappa(x)) = h_{\mu_{q(\kappa)}\phi}.$$

Let $t(\kappa) = q(\kappa)$ if $\frac{1}{\kappa} \geq e_{\max}$ and $t(\kappa) = 0$ otherwise. Suppose $\varepsilon > 0$. By continuity of the multifractal spectrum we have

$$\lim_{\varepsilon \to 0} h_{t(\kappa-\varepsilon)\phi} = h_{t(\kappa)\phi}$$

and

$$h_{\mu_{t(\kappa-\varepsilon)\phi}}(y) = \frac{1}{\kappa - \varepsilon} > \frac{1}{\kappa} \quad \mu_{t(\kappa-\varepsilon)\phi}\text{−a.e. } y.$$

By Corollary 5.3 for $\mu_\phi \times \mu_{q(\kappa-\varepsilon)\phi}$ for a.e. (x, y) we have

$$\alpha^*(x, y) = \alpha(x, y) = h_{\mu_\phi}(y) > \frac{1}{\kappa}.$$

Thus $\pi(y) \in F^\kappa(\pi(x))$ for $\mu_{qt\kappa-\varepsilon)\phi}$ a.e. y and $\dim_H F_\kappa \geq h_{\mu_{t(\kappa-\varepsilon)\phi}}$. Taking the limit $\varepsilon \to 0$ shows

$$\dim_H F^\kappa(\pi(x)) \geq h_{\mu_{t(\kappa)\phi}} = \dim_H \pi(\mathcal{F}^\kappa(x)).$$

This completes the proof for the set F^κ.

It remains to show that $\dim_H I^\kappa(\pi(x)) \leq \dim_H \pi(\mathcal{I}^\kappa(x))$. If $\frac{1}{\kappa} \geq e_{\max}$ then this is trivial since $\dim_H \pi(\mathcal{I}^\kappa(x)) = 1$. Observe that for any κ we have $\dim_H I^\kappa \pi(x) \leq \frac{1}{\kappa}$. To see this consider the natural covering $(T^n\pi(x) - \frac{1}{n^\kappa}, T^n\pi(x) + \frac{1}{n^\kappa})$ of $I^\kappa(\pi(x))$. The s-covering sum is $\sum \frac{1}{n^{\kappa s}} < \infty$ if $s > \frac{1}{\kappa}$. Therefore, if $0 < \frac{1}{\kappa} \leq h_{\mu_\phi}$, we have $\dim I^\kappa(\pi(x)) \leq \frac{1}{\kappa} = \dim_H(\mathcal{I}^\kappa(x))$. Finally if $h_{\mu_\phi} \leq \frac{1}{\kappa} < e_{\max}$ the for any Hölder function $\hat{\phi} \in H^\alpha(\mathbb{S}^1)$ let $\phi = \hat{\phi} \circ \pi$. We have $\phi \in H^\alpha(\Sigma_2^+)$ and $\phi(x_1, \ldots, x_n 01^\infty) = \phi(x_1, \ldots, x_n, 10^\infty)$ thus by the Gibbs property we have

$$\lim_{n \to \infty} \frac{\log \mu_\phi(C_n^\pm(x))}{\log \mu_\phi(C_n(x))} = 1.$$

Hence $h^*_{\mu_\phi}(y) = h_{\mu_\phi}(y)$ for all $y \in \Sigma_2^+$.

Consider the set

$$\begin{aligned} I^\kappa(\pi(x)) \backslash \pi(\mathcal{I}^\kappa(x)) &= \{y : \pi(y) \in I^\kappa(\pi(x)), y \in \mathcal{F}^\kappa(x)\} \\ &\subset \{y : \alpha^*(x, y) < \frac{1}{\kappa}, \alpha(x, y) \geq \frac{1}{\kappa}\}. \end{aligned}$$

By Theorem 6.2 for μ_ϕ-a.e. x we have that the last set is contained in

$$\{y : \alpha^*(x, y) < \frac{1}{\kappa}, h_{\mu_\phi}(y) \geq \frac{1}{\kappa}\}.$$

Thus Lemma 7.1 implies that for any $\varepsilon > 0$ there are at most $C(\varepsilon) \cdot 2E(\frac{1}{\kappa})n$ cylinders of length n needed to cover $\{y : \alpha^*(x, y) < \frac{1}{\kappa}, h_{\mu_\phi}(y) \geq \frac{1}{\kappa} + \varepsilon\}$. Hence

$$\dim_H(I^\kappa(\pi(x)) \backslash \pi(\mathcal{I}^\kappa(x)) \leq E(\frac{1}{\kappa}) = \dim_H \pi(\mathcal{F}^\kappa(x)).$$

□

Corollary 11.3. *For μ_ϕ a.e. x we have $F^\kappa(\pi(x)) = \varnothing$ if $\frac{1}{\kappa} > e_+$.*

In Theorem 3.3 and Corollary 5.3 we can ignore the delta measure on fixed points since they have zero entropy and therefore do not give any contribution. This transfer procedure allows us to conclude the following Theorems and Corollaries from the analogous results of the section 9. These results contain more information than those stated in the introduction, thus we reformulate them. We set $\nu_\phi = \mu_\phi \circ \pi^{-1}$.

Theorem 11.4. (Theorem 2.2) *For ν_ϕ-a.e. s we have*

$$\kappa_{\phi,\psi,\mathbb{S}^1} = \frac{1}{-\int_{\mathbb{S}^1} \phi \, d\nu_\psi} == -\frac{1}{\frac{d}{dt}P(\phi + t\psi)|_{t=0}}.$$

Lemma 11.5. *For ν_ϕ a.e. s we have*

$$\sup\{E(t) : \frac{1}{t} \leq \kappa\} \geq \dim_H F^\kappa(s) \geq \sup\{E(t) : \frac{1}{t} < \kappa\}.$$

For ν_ϕ a.e. s and $\kappa < 1/h_{\nu_\phi}$ we have

$$\sup\{E(t) : \frac{1}{t} \geq \kappa\} \geq \dim_H I^\kappa(s) \geq \sup\{E(t) : \frac{1}{t} > \kappa\}.$$

Corollary 11.6. *For ν_ϕ a.e. s*

$$\sup_{-P'(q) \geq \frac{1}{\kappa}} [P(q\phi) - P'(q\phi)q] \geq \dim_H F^\kappa(s) \geq \sup_{-P'(q) > \frac{1}{\kappa}} [P(q\phi) - P'(q\phi)q] \, .$$

For ν_ϕ a.e. s and for $\kappa < 1/h_{\nu_\phi}$

$$\sup_{-P'(q) \leq \frac{1}{\kappa}} [P(q\phi) - P'(q\phi)q] \geq \dim_H I^\kappa(s) \geq \sup_{-P'(q) < \frac{1}{\kappa}} [P(q\phi) - P'(q\phi)q] \, .$$

Corollary 11.7. (Theorems 2.4 and 2.5) *For a typical potential ϕ and ν_ϕ a.e. s we have*

$$\dim_H F^\kappa(s) = \dim_H(\mathbb{S}^1) = h_{\text{top}}(\mathbb{S}^1) = 1 \quad \text{for } 1/\kappa \leq -\int \phi \, dLeb.,$$

$$\dim_H I^\kappa(s) = \dim_H(\mathbb{S}^1) = h_{\text{top}}(\mathbb{S}^1) = 1 \quad \text{for } 1/\kappa \geq -\int \phi \, dLeb..$$

Let q_κ be the number such that $P'(q_\kappa\phi) = -\frac{1}{\kappa}$. Then

$$\dim_H F^\kappa(s) = E\left(\tfrac{1}{\kappa}\right) = P\left(q_\kappa\phi\right) + \tfrac{1}{\kappa}q_\kappa \qquad \text{for } 1/\kappa > -\int \phi \, dLeb.,$$

$$\dim_H I^\kappa(s) = E\left(\tfrac{1}{\kappa}\right) = P\left(q_\kappa\phi\right) + \tfrac{1}{\kappa}q_\kappa \qquad \text{for } h_{\nu_\phi} \leq 1/\kappa < -\int \phi \, dLeb.,$$

$$\dim_H I^\kappa(s) = \tfrac{1}{\kappa} \qquad \text{for } 1/\kappa < h_{\nu_\phi}.$$

Remark: 1) If $\kappa > 1$ then $\sum l_n < \infty$ and we can not cover Lebesgue almost all points infinitely often no matter which orbit we consider. Thus it is likely that the dimension of $I^\kappa(s)$ is less than 1. In the degenerate case this is clear. To see this in the nondegenerate case note that since the graph of the entropy spectrum is below the diagonal we have $1 = h_{\text{top}} =$

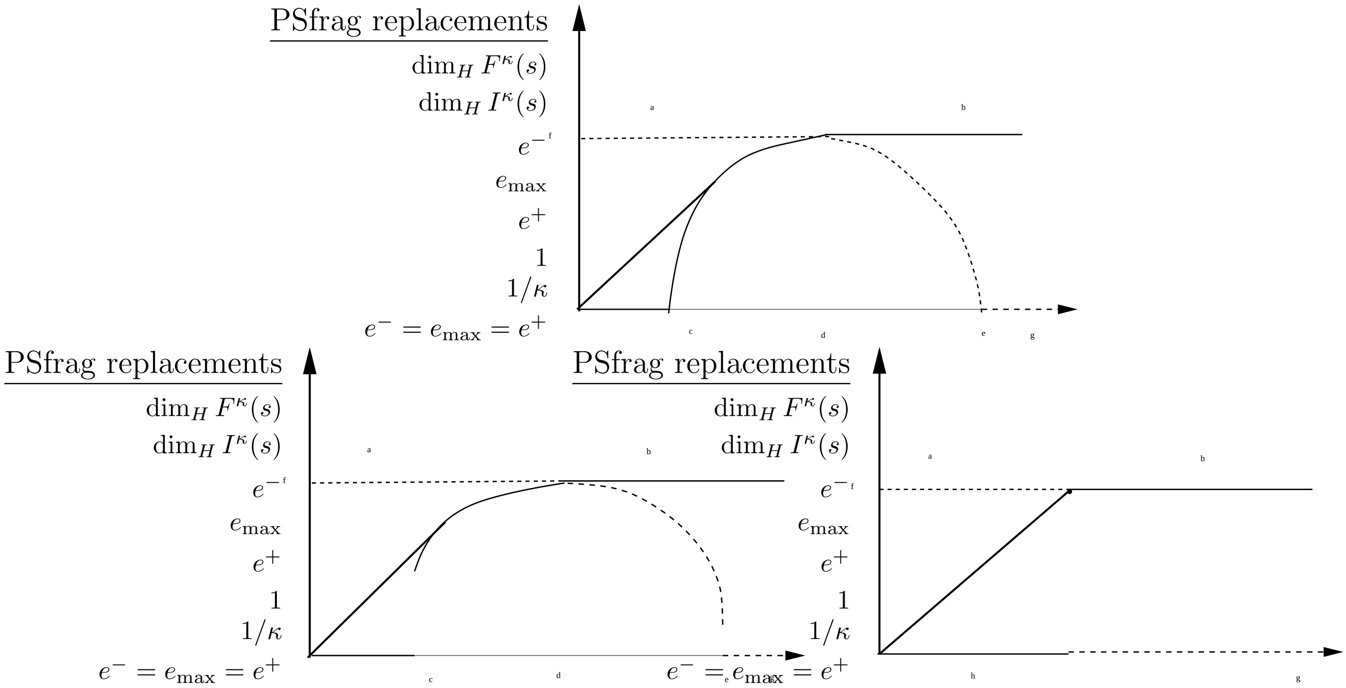

FIGURE 2. The dimension graphs in the typical, nontypical and degenerate cases. The graph of $\dim_H F^\kappa(s)$ is dotted and the graph of $\dim_H I^\kappa(s)$ is solid.

$E(e_{\max}) < e_{\max}$. Therefore the maximum dimension (i.e. 1) is attained for $\kappa < 1$.

2) For a non typical potential we have possibly discontinuities of the function $E(t)$ at $e^\pm$. At these points the upper and lower estimates of Corollary 11.6 do not coincide. This indicates that the question about infinite versus finite covering can not be completely answered in terms of the exponent κ. At this point the answer might depend on a constant c where $l_n = \frac{c}{n^\nu}$. This is in particular the case for the i.i.d. case mentioned in the introduction. The dynamical analog is Lebesgue measure whose entropy spectrum is degenerate. Therefore we can not get any information about the sequence $\frac{c}{n}$ which resembles the i.i.d. case.

Theorem 11.8. (Theorem 2.3) *For ν_ϕ a.e. s we have*

$$F^\kappa(s) = \varnothing \text{ for } \kappa < \frac{1}{-\inf_{\mu \text{ ergodic}} \int \phi \, d\mu} = \kappa^F_{\phi,\mathbb{S}^1}.$$

These results are summarized in Figure 2.

Remark: The result of Corollary 9.6 can also be transferred to the circle. The interpretation of this result is as follows. The distribution of a typical orbit up to time L is completely determined by the entropy spectrum of the measure.

Acknowledgment The authors thank Yue Li YU, Ling Min LIAO and Qin Long LI for their useful remarks.

References

[BF] J. Barral and A. H. Fan, *Covering numbers of different points in Dvoretzky and Poisson coverings*, Bull. Sci. Math. Fr. 129 (2005) 275–317.

[BPS] L. Barreira, Ya. Pesin, and J. Schmeling, *On a general concept of multifractality: multifractal spectra for dimensions, entropies, and Lyapunov exponents. Multifractal rigidity*, Chaos 7 (1997) 27–38.

[BSS] L. Barreira, B. Saussol and J. Schmeling, *Higher-dimensional multifractal analysis*, J. Math. Pures Appl. 81 (2002) 67–91.

[BDV] V. Beresnevich, D. Dickinson, and S. Velani, *Measure theoretic laws for lim sup sets*, Mem. Amer. Math. Soc. 179 (2006).

[BV] V. Beresnevich and S. Velani, *A mass transference principle and the Duffin-Schaeffer conjecture for Hausdorff measures*, Ann. Math. 164 (2006) 971–992.

[B] R. Bowen, *Entropy for non-compact sets*, Trans. Amer. Math. Soc. 184 (1973) 125–136.

[Bu] Y. Bugeaud, *A note on inhomogeneous Diophantine approximation*, Glasgow Math. J. 45 (2003) 105–110.

[Cas] J. W. S. Cassels, *An introduction to diophantine approximation*, Hafner Publishing Company, New York, 1972.

[C] J. R. Chazottes, *Dimensions and waiting times for Gibbs measures*, Journal of Statistical Physics 98 (2000) 305–320.

[CK] N. Chernov and D. Kleinbock, *Dynamical Borel-Cantelli lemmas for Gibbs measures*, Israel J. Math. 122 (2001) 1–27.

[Do] D. Dolgopyat, *Limit theorems for partially hyperbolic systems*, Trans. Amer. Math. Soc. 356 (2003) 1637–1689.

[D] A. Dvoretzky, *On covering a circle by randomly placed arcs,* Proc. Nat. Acad. Sci. USA 42 (1956) 199–203.

[F1] A. H. Fan, *How many intervals cover a point in Dvoretzky Covering?*, Israel J. Math. 131 (2002) 157–184.

[F2] A. H. Fan, *Some topics in the theory of multiplicative chaos*, in Progress in Probability, No 57, Ch. Bandt, M. Zaehle and U. Mosco (Ed.), Birkhauser, 2004, pp 119-134.

[FF] A. H. Fan and D. J. Feng, *On the distribution of long-term time averages on symbolic space*, J. Stat. Phys. 99 (2000) 813–856.

[FFW] A. H. Fan, D. J. Feng and J. Wu, *Recurrence, dimension and entropy*, J. London Math. Soc. 64 (2001) 229–244.

[FK] A. H. Fan and J. P. Kahane, *Rareté des intervalles recouvrant un point dans un recouvrement aleatoire*, Ann. Inst. Henri Poincaré 3 (1993) 453–466.

[FLL] A. H. Fan, B. Li and T. Langlet, *Quantitative uniform hitting in exponentially mixing systems*, pp 251-266 in Recent Developments in Fractals and Related Fields, Ed. J. Barral and S. Seuret, Series Applied and Numerical Harmonic Analysis, Birkhäuser, 2010.

[FS] A. H. Fan and J. Schmeling, *Coverings of the circle driven by rotations*, in Dynamical systems from number theory to probability II, A. Khrennikov (Ed.), Mathematical Modelling in Physics, Engineering and Cognitive Science, vol 6, Växjö University Press, 2003, pp. 7-15.

[FS1] A. H. Fan and J. Schmeling, *On fast Birkhoff averaging*, Math. Proc. Cambridge Philos. Soc. 135 (2003) 443–467.

[FW1] A. H. Fan and J. Wu, *On the covering by small random Intervals*, Ann. Inst. Henri Poincar, 40 (2004) 125-131.

[FW2] A. H. Fan and J. Wu, *A note on inhomogeneous diophantine approximation with a general error function*, Glasgow Math. J. 48 (2006) 187-191.

[Jaf] S. Jaffard, *On lacunary wavelet series,* Ann. Appl. Probab., 10 (2000), no. 1, 313-329.

[JS] J. Jonasson and J. Steif, *Dynamical models for circle covering: Brownian motion and Poisson updating,* Annals of Prob. Volume 36, Number 2 (2008), 739-764.

[HV] R. Hill and S. Velani, *The ergodic theory of shrinking targets*, Invent. Math. 119 (1995) 175–198.

[HV1] R. Hill and S. Velani, *The shrinking target problem for matrix transformations of tori*, J. London Math. Soc. (2) 60 (1999) 381–398.

[K] J. P. Kahane, *Some random series of functions*, Cambridge University Press, Cambridge, 1985.

[OW] D. Ornstein and B. Weiss, *Entropy and data compression schemes*, IEEE Trans. Inform. Theory 39 (1993) 78–83.

[P] Ya. Pesin, *Dimension theory in dynamical systems,* University of Chicago Press, Chicago, 1997.

[Ph] W. Philipp, *Some metrical theorems in number theory*, Pacific J. Math. 20 (1967) 109–127.

[PW] Ya. Pesin and H. Weiss, *The multifractal analysis of Gibbs measures: motivation, mathematical foundation, and examples,* Chaos 7 (1997) 89–106.

[Ru] D. Ruelle, *Statistical mechanics of a one-dimensional lattice gas*, Comm. Math. Phys. 9 (1968) 267–278.

[ST] J. Schmeling and S. Troubetzkoy, *Inhomogeneous Diophantine approximation and angular recurrence for polygonal billiards,* Mat. Sbornik 194 (2003) 295–309.

[S1] J. Schmeling, *Entropy preservation under Markov coding*, J. Stat. Phys. 104 (2001) 799–815.

[S2] J. Schmeling, *On the completeness of multifractal spectra,* Ergodic Theory Dynam. Systems 19 (1999) 1595–1616.

[Sh] L. Shepp, *Covering the circle with random arcs,* Israel J. Math. 11 (1972) 328–345.

A. H. Fan: LAMFA, UMR 6140, CNRS, University of Picardie, 33 Rue Saint Leu, 80039 Amiens, France

E-mail address: ai-hua.fan@u-picardie.fr

URL: http://www.mathinfo.u-picardie.fr/fan/

J. Schmeling: Mathematics Centre for Mathematical Sciences, Lund Institute of Technology, Lund University Box 118 SE-221 00 Lund, Sweden

E-mail address: joerg@maths.lth.se

URL: http://www.maths.lth.se/matematiklth/personal/joerg/

S. Troubetzkoy: Centre de physique théorique, Fédération de recherches des unités de mathématique de Marseille, Institut de mathématiques de Luminy and, Université de la Méditerranéen Luminy, Case 907, F-13288 Marseille Cedex 9, France

E-mail address: troubetz@iml.univ-mrs.fr

URL: http://iml.univ-mrs.fr/~troubetz/