Relative entropy, dimensions and large deviations for *g*-measures

2000 J. Phys. A: Math. Gen. 33 675

(http://iopscience.iop.org/0305-4470/33/4/302)

J. Phys. A: Math. Gen. **33** (2000) 675–689. Printed in the UK PII: S0305-4470(00)02332-5

Relative entropy, dimensions and large deviations for g-measures

Jean-René Chazottes† and Eric Olivier‡

† CPT-CNRS, Luminy case 907, F-13288 Marseille Cedex 9, France
‡ CMI, 39 rue F Joliot Curie, F-13543 Marseille Cedex 13, France

E-mail: jeanrene@cptsg2.univ-mrs.fr and olivier@gyptis.univ-mrs.fr

Received 4 March 1999, in final form 16 November 1999

Abstract. We prove large-deviation results for empirical measures, with respect to an arbitrary g-measure on a subshift of finite type, by using an ergodic theorem for relative entropy we establish. Extending some results of Cajar on Billingsley dimensions to g-measures, we relate the relative entropy to the dimensions of saturated sets. This allows a dimensional interpretation of the large-deviation results and the study of the set of non-generic points which are proved to be of full dimension.

1. Introduction and set-up

The aim of this paper is to investigate some ergodic or statistical properties of the so-called g-measures. Loosely speaking, g-measures form a class of invariant measures lying between Gibbs measures, whose ergodic properties are well understood, and general equilibrium states, whose ergodic properties are substantially more difficult to handle. The notion of a g-measure has a long history in probability theory under various names (a chain with complete connections [1] and a uniform martingale). We take a non-standard definition of g-measures which is equivalent to the more usual one in the case when the function g is continuous. (Ω, σ) denotes the full shift on, say, two symbols. (We give more information on this at the end of this section.) The function g will be called a g-function if g is a continuous function from Ω to $(0, 1)$ such that

$$\sum_{\omega':\sigma\omega'=\omega} g(\omega') = 1. \tag{1}$$

Write $\mathcal{G}$ for the set of all such functions. The measure μ is a g-measure if it is shift-invariant and satisfies

$$\lim_{n\to\infty} \frac{\mu[\pi_{n+1}\cdot\omega]}{\mu[\pi_n\cdot\sigma\omega]} = g(\omega) \qquad \text{for almost all } \omega \text{ w.r.t. } \mu. \tag{2}$$

For every $g \in \mathcal{G}$, there is always at least one g-measure by a standard ergodic theory argument. A g-measure is always non-atomic and of full support (for proofs of these facts and other background information, the reader is referred to [11]). In the case where g is Hölder continuous, the g-measure is known to be unique (this corresponds to Bowen Gibbs measures [2]). One of the weakest conditions known to ensure uniqueness of g-measures was provided by Berbee [1]. For a long time it was an open question whether or not for every continuous g-function there is a unique g-measure, but this was settled recently by Bramson and Kalikow [3] who constructed an example of a continuous g-function with two distinct g-measures (see also [15]). In section 2, we define topological pressure and information to characterize g-measures as equilibrium states.

In this paper, we follow two directions, namely large deviations and Billingsley dimensions. We generalize some previous results, relate them and give some consequences. Let us be more precise. We now describe the organization of the article.

Large deviations appear naturally in statistical mechanics (see [9]) in the equivalence of ensembles. In the Markovian or Gibbsian context, one can analyse large deviations precisely. A key quantity in this study is relative entropy. We generalize its definition when the reference measure is an arbitrary g-measure and prove a pointwise ergodic theorem expressed in terms of information on the measures involved (section 3). This enables us to establish the lower bound of large deviations for empirical measures, which is more delicate than the upper one that we derive using a standard argument, once the existence of the cumulant generating function is proved (section 4). In [10], large-deviation estimations are obtained in a more general setting, but uniqueness of equilibrium states is assumed there in order to use derivatives of the pressure. Moreover, our approach is extensively based on the properties of relative entropy.

Billingsley dimensions appeared naturally as an extension of the Hausdorff dimension in the context of sets of real numbers characterized by digit properties of their s-adic representations (see [5] and references therein). These sets are typical examples of saturated sets. In the context of dynamical systems, typical saturated sets are sets of generic points of invariant measures. Loosely speaking, one defines the Billingsley dimension by replacing the diameter of cylinders by some non-atomic measure in the Carathéodory construction. Taking any g-measure as such a reference measure, we derive some formula giving access to the dimension of the saturated sets. It generalizes the formula given in [5] for Markov measures to arbitrary g-measures (section 5).

In section 6, we give different consequences of the preceding sections. First of all, we establish the relationship between Billingsley dimensions and the relative entropy through a simple formula. In the next subsection, we perform the multifractal analysis of Hölder continuous functions by using the large-deviation results of section 4 and the formula we have just talked about. We refer to [14] for a general and unified presentation of the multifractal formalism. The last subsection is dedicated to the study of non-generic points, that is, using the language of multifractal analysis, of the points for which level sets are not defined. We show that such points are 'observable' in the sense that they have full Hausdorff dimension, though they form a null-measure set. We point out that our method for proving this makes use of the formula derived in section 5, computing the Billingsley dimension of saturated sets. We mention the related work in [17] where, roughly speaking, it is proved that the set for which lower and upper pointwise dimensions, of any Gibbs measure, are different has a strictly positive Hausdorff dimension and is dense.

We end this section by collecting together basic notations, definitions and some important basic facts.

Let Ω denote the full shift space $\{0, \ldots, k-1\}^{\mathbb{N}}$ which is endowed with the product topology and $\mathcal{M}_\sigma$ the set of Borel probability measures on Ω which are invariant under the shift map $\sigma : \Omega \to \Omega$, $(\sigma\omega)_n = \omega_{n+1}$. For any $\omega \in \Omega$ and $n \geqslant 1$, $[\pi_n \cdot \omega]$ is the nth cylinder about ω: $[\pi_n \cdot \omega] := \{\omega' : \pi_n \cdot \omega' = \pi_n \cdot \omega\}$ and $\Omega_n := \{0, \ldots, k-1\}^n = \pi_n(\Omega)$ is the set of words of length n. Cylinder sets generate the Borel sigma-algebra $\mathfrak{B}$. We use also the following notation: $\mathcal{M}$ for the set of probability measures on Ω, endowed with the weak* topology, $\mathcal{C}(\Omega)$ for the set of real continuous functions on Ω (we use the supremum norm

$\| \cdot \|_\infty$). The results of this paper remain valid with straightforward modifications in the case of any topologically mixing subshift of finite type, that is a subshift such that admissible words are given by a primitive matrix A (i.e. $\exists n_0 > 0,$ s.t. $\forall (i, j) \in \{0, 1, \ldots, k-1\}^2$, $(A^n)_{i,j} > 0$ for any $n \geqslant n_0$).

Let us point out a key property we shall use, namely the specification property, which ensures that for any admissible words u and v there exists a word w of length at most n_0, such that the concatenated word uwv is admissible. It is shown in [7] that such a property implies that ergodic measures are dense in $\mathcal{M}_\sigma$ in the weak* topology. This fact will be useful when we deal with large deviations, as well as when we study non-generic points.

2. Thermodynamic formalism of g-measures

The two main ingredients we need are topological pressure and information.

For all $a \in \Omega_n$, we denote by $\overline{a}$ the unique n-periodic point of $[a]$. Now define the pressure of any continuous potential (see [16] for instance):

Definition 2.1. *Let $\phi \in \mathcal{C}(\Omega)$. Then*

$$P(\phi) = \lim_{n \to +\infty} \frac{1}{n} \log \sum_{a \in \Omega_n} \exp\, S_n\phi(\overline{a}) \tag{3}$$

where $S_n\phi$ denotes the Birkhoff sum $\sum_{k=0}^{n-1} \phi \circ \sigma^k$ throughout this paper.

Remark 2.2. *Pressure can be defined as a capacity using a Carathéodory construction. We refer to [14] for details.*

A useful characterization of g-measures due to Ledrappier uses information on shift-invariant measures. For $\nu \in \mathcal{M}_\sigma$ define the following Borel function from Ω to $[-\infty, +\infty]$:

$$I_\mu := \sum_{a \in \Omega_1} -\chi_{[a]} \log \mathbb{E}(\chi_{[a]} | \sigma^{-1}\mathfrak{B}) \tag{4}$$

where $\mathbb{E}$ denotes the conditional expectation with respect to $\sigma^{-1}\mathfrak{B}$ in $L^1(\mu)$. The information I_μ is μ-integrable and such that $\int I_\mu \, \mathrm{d}\mu = h(\mu)$. Define $\phi_g := \log g$ for any g-function g. We have the following:

Theorem 2.3 ([11]). *Let $g \in \mathcal{G}$. A shift-invariant measure μ on Ω is a g-measure if and only if $\phi_g = -I_\mu$, μ-a.e.*

A g-measure can be defined equivalently as the equilibrium state of the potential ϕ_g, that is as the supremum of $\int \phi_g \, \mathrm{d}\nu + h(\nu)$ over all shift-invariant measures, which is the topological pressure of ϕ_g, $P(\phi_g)$ ($h(\nu)$ is the entropy of ν). This is the variational principle. By theorem 2.3, it follows that $P(\phi_g) = 0$. Non-uniqueness means that the set of equilibrium states of ϕ_g, denoted by $\mathcal{I}_{\phi_g}$, is not a singleton (which can be interpreted as a phase transition when one thinks about Ω as a one-dimensional lattice).

Remark 2.4. *Condition (1), given in section 1, reads $\sum_{\omega' : \sigma\omega' = \omega} \mathrm{e}^{\phi_g(\omega')} = 1$ that is $\mathcal{L}_{\phi_g} 1 = 1$, where $\mathcal{L}_{\phi_g}$ is the transfer operator associated with ϕ_g. Moreover, we have that μ is a g-measure if, and only if, $\mathcal{L}^*_{\phi_g} \mu = \mu$, where $\mathcal{L}^*_{\phi_g}$ is the dual of $\mathcal{L}_{\phi_g}$ in the weak* sense. This is the third possible definition of g-measures equivalent to the two others. (See again [11].)*

Theorem 2.3 leads to the natural definition of canonical information:

Definition 2.5 (Canonical information). *Let μ be a g-measure associated with $g \in \mathcal{G}$. The function $I^c_\mu := -\phi_g$ is called the canonical information of μ.*

As a g-measure is necessarily supported by Ω (see [13]), one can consider for any integer n the real-valued function g_n^μ defined for $\omega \in \Omega$ by $g_n^\mu(\omega) := \mu[\pi_{n+1} \cdot \omega]/\mu[\pi_n \cdot \sigma\omega]$. Clearly this is a g-function depending only on the first $(n + 1)$ symbols. The unique g-measure μ_n of g_n^μ is by definition the n-step Markov approximation of μ and one has $I_{\mu_n}^c = -\log g_n^\mu$. One has the following lemma:

Lemma 2.6 (Markov approximation and information, [13]). *If μ is a g-measure whose n-step Markov approximation is μ_n, then $I_{\mu_n}^c$ converges uniformly to I_μ^c.*

This lemma will be used several times in the following. Before proceeding further, let us make the following simple remark:

Remark 2.7. *Any probability measure of full support can be obtained as a weak* limit of its Markov approximations. This is implied by the definition of weak* topology.*

Since $I_{\mu_n}^c$ uniformly converges to I_μ^c, there exists a sequence $(\varepsilon_n)_n$ of non-negative real numbers decreasing to 0 and such that for all $\omega \in \Omega$ and $n \geqslant 0$,

$$e^{-n\varepsilon_n} \leqslant \frac{\mu[\pi_n \cdot \omega]}{\exp\left(S_n I_\mu^c(\omega)\right)} \leqslant e^{n\varepsilon_n}. \tag{5}$$

To close this section, we mention that if $\phi : \Omega \to \mathbb{R}$ is a Hölder continuous† function, then there is a unique measure, μ_ϕ, for which there are constants $0 < \underline{c} \leqslant \overline{c}$ such that, for any $n \geqslant 1$ and any $\omega \in \Omega$, we have

$$\underline{c} \leqslant \frac{\mu_\phi[\pi_n \cdot \omega]}{\exp\left(-nP(\phi) + S_n\phi(\omega)\right)} \leqslant \overline{c}.$$

The measure μ_ϕ is a Gibbs measure. In fact (see [16]), on can always assume $P(\phi) = 0$ because any Hölder continuous potential can be normalized. Compare with (5).

3. Relative entropy

Let $\phi \in \mathcal{C}(\Omega)$ and η be a shift-invariant measure on Ω. Define $h(\eta|\mathcal{I}_\phi) := P(\phi) - \big(h(\eta) + \eta(\phi)\big)$ (where $\mathcal{I}_\phi$ denotes the set of equilibrium states associated with ϕ, which is never empty by expansivity of the shift map, see [16]). By the variational principle, $h(\cdot|\mathcal{I}_\phi)$ is a positive map on $\mathcal{M}_\sigma(\Omega)$ and $h(\eta|\mathcal{I}_\phi) = 0$ if and only if $\eta \in \mathcal{I}_\phi$. Moreover, entropy being an affine and upper semi-continuous map from $\mathcal{M}_\sigma$ to $\mathbb{R}^+$, it is clear that $h(\cdot|\mathcal{I}_\phi)$ is an affine lower semi-continuous map from $\mathcal{M}_\sigma$ to $\mathbb{R}^+$.

Let μ and η be shift-invariant measures on Ω. If μ is supported by Ω then,

$$H_n(\eta|\mu) := \int \log \frac{\eta[\pi_n \cdot \omega]}{\mu[\pi_n \cdot \omega]} \, d\eta(\omega) = \sum_{a \in \Omega_n} \eta[a] \log \frac{\eta[a]}{\mu[a]}. \tag{6}$$

Definition 3.1 (Relative entropy). *The relative entropy of η with respect to μ is the positive quantity*

$$h(\eta|\mu) := \limsup_{n\to+\infty} \frac{1}{n} H_n(\eta|\mu). \tag{7}$$

Clearly, $h(\mu|\mu) = 0$ for any μ, but in general $h(\eta|\mu) = 0$ does not imply $\eta = \mu$.

Define $\mathcal{J}_\sigma$ as the sigma-algebra of shift-invariant Borel sets of Ω.

† What follows holds when ϕ has summable variations.

Theorem 3.2 (Pointwise convergence to relative entropy). *Let μ be a g-measure and η a shift-invariant measure. Then the following statements hold:*

(i) $$\lim_{n\to+\infty} \frac{1}{n} \log \frac{\eta[\pi_n \cdot \omega]}{\mu[\pi_n \cdot \omega]} = \mathbb{E}\big(I_\mu^c - I_\eta | \mathcal{J}_\sigma\big)(\omega) \qquad \eta\text{-a.e. in } L^1(\eta)$$

(ii) $$h(\eta|\mu) = \int (I_\mu^c - I_\eta)\, \mathrm{d}\eta = h(\eta|\mathcal{I}_{\phi_g}).$$

Proof. (i) Let μ_n be the n-step Markov approximation of the g-measure μ. It follows from the definition of $I_{\mu_n}^c$ that $\log \mu[\pi_n \cdot \omega] = \sum_{k=0}^{n-1} I_{\mu_{n-k+1}}^c(\sigma^k \omega)$. Therefore,

$$\left| \frac{1}{n} S_n I_\mu^c(\omega) + \frac{1}{n} \log \mu[\pi_n \cdot \omega] \right| \leqslant \frac{1}{n} \sum_{k=0}^{n-1} \| I_\mu^c - I_{\mu_k}^c \|_\infty$$

and since $I_{\mu_k}^c$ goes to I_μ^c in the uniform norm on $\mathcal{C}(\Omega)$ (lemma 2.6), then the Cesaro lemma implies

$$\lim_{n\to+\infty} \left(\frac{1}{n} S_n I_\mu^c(\omega) + \frac{1}{n} \log \mu[\pi_n \cdot \omega] \right) = 0. \tag{8}$$

For any shift-invariant measure η, the Birkhoff ergodic theorem ensures that

$$\lim_{n\to+\infty} \frac{1}{n} S_n I_\mu^c(\omega) = \mathbb{E}(I_\mu^c | \mathcal{J}_\sigma)(\omega) \qquad \eta \text{ a.e. in } L^1(\eta). \tag{9}$$

Now, by the Shannon–McMillan–Breiman theorem:

$$\lim_{n\to+\infty} \frac{1}{n} \log \eta[\pi_n \cdot \omega] = \mathbb{E}(I_\eta | \mathcal{J}_\sigma)(\omega) \qquad \eta \text{ a.e. in } L^1(\eta). \tag{10}$$

Combining (8)–(10), one obtains assertion (i).

(ii) Definition of the relative entropy together with the pointwise convergence given in (ii) implies

$$\begin{aligned} h(\eta|\mu) &= \limsup_{n\to+\infty} \frac{1}{n} \int \log \frac{\eta[\pi_n \cdot \omega]}{\mu[\pi_n \cdot \omega]}\, \mathrm{d}\eta(\omega) \\ &= \int \mathbb{E}(I_\mu^c - I_\eta | \mathcal{J}_\sigma)\, \mathrm{d}\eta = \int (I_\mu^c - I_\eta)\, \mathrm{d}\eta. \end{aligned}$$

Since $I_\mu^c = -\phi_g$, $\int (I_\mu^c - I_\eta)\, \mathrm{d}\eta = -\eta(\phi_g) - h(\eta) = h(\eta|\mathcal{I}_{\phi_g})$. □

We end this section with a remark relating relative entropy to some capacities.

Remark 3.3. *Using a result in [14], appendix II, we can write*

$$h(\eta|\mathcal{I}_\phi) = P(\phi) - P_{G(\eta)}(\phi)$$

where η is any ergodic measure, $G(\eta)$ is the set of η-generic points (see section 6.1 for the precise definition) and ϕ is any continuous function.

4. Large deviations for empirical measures

We establish large-deviation results for the sequences $\big(T_n : \Omega \to \mathcal{M}\big)_n$ and $\big(\tilde{T}_n : \Omega \to \mathcal{M}\big)_n$ of empirical and cyclic empirical measures on the space (Ω, μ), when μ is a g-measure. By definition, $\tilde{T}_n(\omega) := T_n(\overline{\pi_n \cdot \omega})$, for any $n \geqslant 1$ and any $\omega \in \Omega$ and $T_n(\omega) := (1/n) \sum_{k=0}^{n-1} \delta_{\sigma^k \omega}$. (Recall that $\overline{\pi_n \cdot \omega}$ is the n-periodic point made up with the n first coordinates of ω.) Let us remark that in the case of a topological mixing subshift of finite type, cyclic empirical measures are well defined up to a subsequence using the specification property (see the introduction for the definition).

4.1. Upper bound

The logarithmic moment generating function Ψ, associated with a sequence $\big(Y_n : \Omega \to \mathcal{M}\big)_n$ of random variables on the space (Ω, μ), is the convex map from $\mathcal{C}(\Omega)$ to $\mathbb{R} \cup \{+\infty\}$, defined, when it does exist, by setting for all $\psi \in \mathcal{C}(\Omega)$:

$$\Psi(\psi) := \lim_{n\to+\infty} \frac{1}{n} \log \int \exp(\langle nY_n(\omega) \mid \psi\rangle)\, \mathrm{d}\mu(\omega).$$

In the general case, it is proved in [8] that if Ψ exists and is not identically equal to $+\infty$, then for every compact subset K of $\mathcal{M}$, one has the following upper bound:

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu\{\omega : Y_n(\omega) \in K\} \leqslant -\inf_{\eta\in K}\{\Psi^*(\eta)\} \tag{11}$$

where Ψ^* is the Legendre–Fenchel transform of Ψ, that is

$$\forall \mu \in \mathcal{M} \qquad \Psi^*(\mu) := \sup_{f\in\mathcal{C}(\Omega)} \{\langle\mu|f\rangle - \Psi(f)\}$$

or, in an equivalent way

$$\forall f \in \mathcal{C}(\Omega) \qquad \Psi(f) := \sup_{\mu\in\mathcal{M}} \{\langle\mu|f\rangle - \Psi^*(\mu)\}.$$

(Recall that $\langle\mu|f\rangle$ stands for $\int f\, \mathrm{d}\mu$ or $\mu(f)$.) One has the following lemma:

Lemma 4.1 (Moment generating function). *If μ is a g-measure on Ω, then the convex map $\boldsymbol{P}_g$ from $\mathcal{C}(\Omega)$ to $\mathbb{R}$ defined by $\boldsymbol{P}_g(\psi) = P(\phi_g+\psi)$ for all $\psi \in \mathcal{C}(\Omega)$, is the common logarithmic moment generating function of $(\tilde{T}_n)_n$ and $(T_n)_n$.*

Proof. Let us denote by $\mathrm{var}_n(\psi)$ the variations of ψ, on the n-cylinders of Ω. One has for any integer $n > 1$, $|S_n\psi(\omega) - S_n\psi(\overline{\pi_n \cdot \omega})| \leqslant \sum_{k=1}^{n} \mathrm{var}_k(\psi)$ and thus

$$\exp\bigg(-\sum_{k=1}^{n} \mathrm{var}_k(\psi)\bigg) \leqslant \frac{\int \exp\big(S_n\psi(\omega)\big)\,\mathrm{d}\mu(\omega)}{\int \exp\big(S_n\psi(\overline{\pi_n \cdot \omega})\big)\,\mathrm{d}\mu(\omega)} \leqslant \exp\bigg(\sum_{k=1}^{n} \mathrm{var}_k(\psi)\bigg).$$

Because $\mathrm{var}_n(\psi)$ goes to 0, as n goes to infinity, then the Cesaro lemma gives

$$\lim_{n\to+\infty}\left|\frac{1}{n}\log\int \exp\big(S_n\psi(\omega)\big)\,\mathrm{d}\mu(\omega) - \frac{1}{n}\log\sum_{a\in\Omega_n} \exp\big(S_n\psi(\overline{a})\big)\mu[a]\right| = 0. \tag{12}$$

Moreover, for any integer $n > 0$:

$$\sum_{a\in\Omega_n} \exp\big(S_n\psi(\overline{a})\big)\mu[a] = \sum_{a\in\Omega_n} \exp\big(S_n(\phi_g+\psi)(\overline{a})\big)\frac{\mu[a]}{\exp\ S_n\phi_g(\overline{a})}. \tag{13}$$

Since μ is a g-measure, one can easily deduce from (5) and (3) that

$$\lim_{n\to+\infty}\frac{1}{n}\log\sum_{a\in\Omega_n} \exp\big(S_n\psi(\overline{a})\big)\mu[a] = \lim_{n\to+\infty}\frac{1}{n}\log\sum_{a\in\Omega_n} \exp\ S_n(\phi_g+\psi)(\overline{a}) = P(\phi_g+\psi).$$

Combining statements (12) and (13), the proof is complete □

Given $g \in \mathcal{G}$, one defines the following convex map:

$$\boldsymbol{H}_g : \mathcal{M} \to \mathbb{R}\cup\{+\infty\} \qquad \text{defined by} \quad \boldsymbol{H}_g(\eta) = \begin{cases} h(\eta|\mathcal{I}_{\phi_g}) & \text{if} \quad \eta \in \mathcal{M}_\sigma \\ +\infty & \text{if} \quad \eta \notin \mathcal{M}_\sigma. \end{cases}$$

Lemma 4.2. *If $g \in \mathcal{G}$, then the convex maps $\boldsymbol{P}_g$ and $\boldsymbol{H}_g$ are conjugate by Legendre–Fenchel transformation.*

Proof. Since $\boldsymbol{H}_g(\eta) = +\infty$ when $\eta \notin \mathcal{M}_\sigma$, then the variational principle gives

$$\sup_{\eta\in\mathcal{M}}\{\eta(\psi) - \boldsymbol{H}_g(\eta)\} = \sup_{\eta\in\mathcal{M}_\sigma}\{\eta(\psi) - \boldsymbol{H}_g(\eta)\} = \sup_{\eta\in\mathcal{M}_\sigma}\{\eta(\psi) + h(\eta) + \eta(\phi_g)\} = P(\phi_g + \psi).$$

□

So lemma 4.1 ensures the validity of the upper bound (11) and lemma 4.2 expresses the relationship between the relative entropy and the logarithmic moment generating function associated to any g-function. In the next subsection, we proceed to establish the more delicate lower bound.

4.2. Lower bound

Introduce the metric d^* on $\mathcal{M}$ by setting

$$\forall\mu \quad \eta \in \mathcal{M} \qquad d^*(\mu, \eta) = \sum_{n=1}^{\infty} \frac{1}{2^n} \sup_{a\in\Omega_n} |\mu[a] - \eta[a]|.$$

It is easy to check that d^* gives the weak* topology on $\mathcal{M}$. We now state the large-deviation lower bound for empirical measures.

Theorem 4.3 (Lower bound). *Let μ be a g-measure on Ω and H an open subset of $\mathcal{M}$. Then,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu\{\omega : T_n(\omega) \in H\} \geqslant -\inf\{\boldsymbol{H}_g(\eta) \mid \eta \in H\}. \tag{14}$$

Proof. We first prove the statement for cyclic empirical measures. If $H \cap \mathcal{M}_\sigma = \emptyset$ then $\inf\{\boldsymbol{H}_g(\eta) : \eta \in H\} = +\infty$ and (14) is trivial. We now suppose that H intersects $\mathcal{M}_\sigma$. Since ergodic measures are dense in weak* topology (see [7]) and $\boldsymbol{H}_g$ is lower semi-continuous, let ν be an ergodic measure such that $\boldsymbol{H}(\nu) \leqslant \inf\{\boldsymbol{H}_g(\eta) : \eta \in H\} + \varepsilon$, where $\varepsilon > 0$ is arbitrary. If one denotes by $\tilde{E}_n(H) := \{\omega : \tilde{T}_n(\omega) \in H\}$ then, using the definition of the cyclic empirical measure, one can write

$$\mu\big(\tilde{E}_n(H)\big) = \sum_{\substack{a\in\Omega_n \\ [a]\cap\tilde{E}_n(H)\neq\emptyset}} \mu[a] = \sum_{\substack{a\in\Omega_n \\ [a]\cap\tilde{E}_n(H)\neq\emptyset}} \frac{\mu[a]}{\nu[a]}\nu[a]$$
$$= \int \chi_{\tilde{E}_n(H)}(\omega) \frac{\mu[\pi_n \cdot \omega]}{\nu[\pi_n \cdot \omega]} \,\mathrm{d}\nu(\omega). \tag{15}$$

For $\varepsilon > 0$, introduce the set

$$\tilde{E}_n^\varepsilon(H) := \tilde{E}_n(H) \cap \left\{\frac{\mu[\pi_n \cdot \omega]}{\nu[\pi_n \cdot \omega]} \geqslant \mathrm{e}^{-n(\boldsymbol{H}_g(\nu)+\varepsilon)}\right\}.$$

Then it follows that

$$\int \chi_{\tilde{E}_n(H)}(\omega) \frac{\mu[\pi_n \cdot \omega]}{\nu[\pi_n \cdot \omega]} \,\mathrm{d}\nu(\omega) \geqslant \nu\big(\tilde{E}_n^\varepsilon(H)\big)\, \mathrm{e}^{-n(\boldsymbol{H}_g(\nu)+\varepsilon)}. \tag{16}$$

Using successively statements (15) and (16),

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log\mu\big(\tilde{E}_n(H)\big) &\geqslant \liminf_{n\to\infty}\frac{1}{n}\log\int\chi_{\tilde{E}_n(H)}(\omega)\frac{\mu[\pi_n\cdot\omega]}{\nu[\pi_n\cdot\omega]}\,\mathrm{d}\nu(\omega)\\ &\geqslant \liminf_{n\to\infty}\frac{1}{n}\log\left\{\nu\big(\tilde{E}_n^{\varepsilon}(H)\big)\,\mathrm{e}^{-n(\boldsymbol{H}_g(\nu)+\varepsilon)}\right\}\\ &\geqslant \liminf_{n\to\infty}\frac{1}{n}\log\nu\big(\tilde{E}_n^{\varepsilon}(H)\big)-\boldsymbol{H}_g(\nu)-\varepsilon.\end{aligned}$$

By part (ii) of theorem 3.2,

$$\lim_{n\to\infty}\frac{1}{n}\log\frac{\nu[\pi_n\cdot\omega]}{\mu[\pi_n\cdot\omega]}=\boldsymbol{H}_g(\nu)\qquad \nu\text{ a.e.}$$

and $\tilde{T}_n(\omega)$ converges (in the weak* sense) to ν for ν-almost all ω by the Birkhoff ergodic theorem. This gives

$$\liminf_{n\to\infty}\frac{1}{n}\log\nu\big(\tilde{E}_n^{\varepsilon}(H)\big)=0\qquad\text{i.e.}\quad \liminf_{n\to\infty}\frac{1}{n}\log\nu\big(\tilde{E}_n(H)\big)\geqslant\boldsymbol{H}_g(\nu)+\varepsilon.$$

Statement (14) holds because ε is arbitrary.

We consider again the ergodic measure $\nu\in H$ such that $\boldsymbol{H}_g(\nu)\leqslant\inf\{\boldsymbol{H}_g(\eta)\mid\eta\in H\}+\varepsilon$ and H being open. There exists $\rho>0$ such that $B^{\circ}(\nu,\rho)\subset H$. A straightforward computation gives a constant $c>0$ such that

$$\forall n>0\quad\forall\omega\in\Omega\qquad d^*\big(\tilde{T}_n(\omega),T_n(\omega)\big)\leqslant\frac{c}{n}. \tag{17}$$

From this we obtain the following sequence of inclusions, for n sufficiently large:

$$\tilde{E}_n\big(B^{\circ}(\nu,\rho/2)\big)\subset E_n\big(B^{\circ}(\nu,\rho)\big)\subset E_n(H)$$

where $E_n(L):=\{\omega:\ T_n(\omega)\in L\}$ for any subset L of $\mathcal{M}$. Using the result for cyclic empirical measures and these inclusions, it follows that

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log\mu\big(E_n(H)\big) &\geqslant \liminf_{n\to\infty}\frac{1}{n}\log\mu\left\{E_n\big(B^{\circ}(\nu,\rho)\big)\right\}\\ &\geqslant \liminf_{n\to\infty}\frac{1}{n}\log\mu\left\{\tilde{E}_n\big(B^{\circ}(\nu,\rho/2)\big)\right\}\\ &\geqslant -\boldsymbol{H}_g(\nu)\geqslant-\inf\{\boldsymbol{H}_g(\eta)\mid\eta\in H\}-\varepsilon\end{aligned}$$

Since ε is arbitrary, the proof is complete. □

Let us emphasize that the specification property is crucial to ensure the density of ergodic measures in $\mathcal{M}_\sigma$ (see the introduction).

5. Billingsley dimensions of saturated sets

We define Billingsley dimensions by means of a Carathéodory construction. (Refer to [14] for a modern account on such a construction.) We mention that Billingsley dimensions are not defined in [14]. Given an arbitrary non-atomic probability measure μ on Ω, a Borel set $M\subset\Omega$ and a strictly positive number ε, $\mathcal{R}_\varepsilon(M)$ is by definition the set of all covers of M with cylinders of μ-measure less than ε. For $\beta\in[0,1]$, define

$$\boldsymbol{C}_\varepsilon^{\beta}(M;\mu):=\inf_{R\in\mathcal{R}_\varepsilon(M)}\sum_{c\in R}\mu(c)^{\beta}.$$

Since the map $\varepsilon \mapsto C_\varepsilon^\beta(M; \mu)$ is monotonic, we set

$$C^\beta(M; \mu) := \lim_{\varepsilon \to 0^+} C_\varepsilon^\beta(M; \mu).$$

It is easy to check that by construction

$$\beta < \beta' \quad \text{and} \quad C^\beta(M; \mu) < \infty \quad \Rightarrow \quad C^{\beta'}(M; \mu) = 0$$
$$\beta < \beta' \quad \text{and} \quad C^{\beta'}(M; \mu) > 0 \quad \Rightarrow \quad C^\beta(M; \mu) = \infty.$$

This leads to the following definition.

Definition 5.1 (Billingsley dimensions). *Let μ be a non-atomic probability measure on Ω, M an arbitrary Borel subset of Ω and $C^\beta(M; \mu)$ defined above. Then the Billingsley dimension of M with respect to μ is*

$$\dim_\mu M := \inf\{\beta \in [0, 1] : C^\beta(M; \mu) = 0\} = \sup\{\beta \in [0, 1] : C^\beta(M; \mu) = \infty\}.$$

If λ denotes the Parry measure (the unique measure of maximal entropy), then $\dim_\lambda$ is nothing but the usual Hausdorff dimension $\dim_H$ when Ω is equipped with the usual distance

$$d(\omega, \omega') = k^{-\nu(\omega, \omega')}$$

where

$$\nu(\omega, \omega') := \begin{cases} \max\{n \in \mathbb{N} : \pi_n \cdot \omega = \pi_n \cdot \omega'\} & \text{if} \quad \omega \neq \omega' \\ +\infty & \text{if} \quad \omega = \omega'. \end{cases}$$

(Notice that $h(\lambda) = h_{top} = \log k$ and recall that k is the cardinal of the alphabet generating the space Ω.)

Given two probability measures μ and η on Ω, the singularity function of η with respect to μ is defined in [5] by setting

$$\forall \omega \in \Omega \qquad \frac{\eta}{\mu}(\omega) := \liminf_{n \to +\infty} \frac{\log \eta[\pi_n \cdot \omega]}{\log \mu[\pi_n \cdot \omega]} \quad (\geqslant 0)$$

with the classical conventions for the undetermined ratios of the form $\log(x)/\log(y)$. If μ is non-atomic, then it is proved in [5] that for every subset M of Ω,

$$\dim_\mu M = \inf_\eta \sup \left\{ \frac{\eta}{\mu}(\omega) : \omega \in M \right\}$$

where the infimum is taken over all probability measures η on Ω. Cajar defines a quasi-metric q on $\mathcal{M}_\sigma$ by setting

$$q(\mu, \eta) := \sup_{\omega \in \Omega} \max \left\{ \left| \log \frac{\mu}{\eta}(\omega) \right|, \left| \log \frac{\eta}{\mu}(\omega) \right| \right\}.$$

This quasi-metric is constructed in such a way that for any Borel subset M of Ω, if μ is a q-limit of the sequence $(\mu_n)_n$ of non-atomic probability measures (i.e. $\lim_n q(\mu, \mu_n) = 0$), then

$$\lim_{n \to +\infty} \dim_{\mu_n} M = \dim_\mu M. \tag{18}$$

Note that the topology induced by q is not in general the weak* topology (see remark 2.7). Moreover, this quasi-metric is distinct from the $\bar{d}$-distance recently used to evaluate the speed of convergence of Markov approximations for some sufficiently mixing g-measures [4].

We can state the following lemma:

Lemma 5.2 (q-limits and g-measures). *Any g-measure μ on Ω is the q-limit of the sequence $(\mu_n)_n$ of its Markov approximations.*

Proof. Pick $\omega \in \Omega$ and let $\theta : \mathbb{N} \to \mathbb{N}$ be a strictly increasing function such that

$$\frac{\eta}{\nu}(\omega) = \lim_{n\to\infty} \frac{\log\ \eta[\pi_{\theta(n)}\omega]}{\log\ \nu[\pi_{\theta(n)}\omega]}$$

where η and ν are probability measures. By weak* compacity of the set $\mathcal{M}$, one can suppose that

$$\frac{1}{\theta(n)} \sum_{k=0}^{\theta(n)-1} \delta_{\sigma^k\omega}$$

converges weakly to a probability measure ρ_ω. If η and ν are g-measures, respectively, associated with the g-functions g and g', one can deduce from lemma 2.6:

$$\frac{\eta}{\nu}(\omega) = \lim_{n\to\infty} \frac{\log\ \eta[\pi_{\theta(n)}\omega]}{\log\ \nu[\pi_{\theta(n)}\omega]} = \frac{\rho_\omega(I_\eta^c)}{\rho_\omega(I_\nu^c)}. \tag{19}$$

Now, if μ is a g-measure, there exists a compact interval $J \subset (0, +\infty)$ which contains both $\rho(I_\mu^c)$ and $\rho(I_{\mu_n}^c)$ for any probability measure ρ and any $n \in \mathbb{N}$. This is due to the fact that $I_{\mu_n}^c$ uniformly converges to I_μ^c (by lemma 2.6) and the information function of a g-measure is strictly positive and continuous. From (19) we deduce that

$$\left|\log\ \frac{\mu}{\mu_n}(\omega)\right| = \left|\log\ \frac{\rho_\omega(I_\mu^c)}{\rho_\omega(I_{\mu_n}^c)}\right| \leqslant K|\rho_\omega(I_\mu^c - I_{\mu_n}^c)| \leqslant K\ \|\ I_\mu^c - I_{\mu_n}^c \|_\infty$$

where K is a positive constant given by a classical inequality of calculus. We thus obtain

$$q(\mu, \mu_n) \leqslant K\ \|\ I_\mu^c - I_{\mu_n}^c \|_\infty$$

which concludes the proof by again applying lemma 2.6 □

For any $\omega \in \Omega$, let $\Delta(\omega)$ be the set of accumulation points (with respect to the weak* topology) of the sequence of empirical measures $(T_n(\omega))_n$. $\Delta(\omega)$ is a non-empty compact connected subset of $\mathcal{M}_\sigma$ (see, for instance, [7]).

Definition 5.3 (Saturated sets). *A subset M of Ω is said to be saturated if it is saturated in class for the following equivalence relation:*

$$\forall(\omega, \omega') \in \Omega \times \Omega \qquad \omega \sim \omega' \iff \Delta(\omega) = \Delta(\omega'). \tag{20}$$

For every $H \subset \mathcal{M}_\sigma(\Omega)$, let us denote by $\nabla(H) := \{\omega : \Delta(\omega) = H\}$. Note that $\nabla(H)$ is either an empty or a saturated set. The so-called smallest saturated sets are of the form $\nabla\Delta(\omega)$ for some $\omega \in \Omega$, i.e. the equivalence class of ω. Stated in our notation, Cajar proves that for a finite step Markov measure μ on Ω, one has

$$\forall\omega \in \Omega \qquad \dim_\mu \nabla\Delta(\omega) = \inf\left\{\frac{h(\eta)}{\eta(I_\mu^c)} : \eta \in \Delta(\omega)\right\}. \tag{21}$$

We now give the natural generalization of (21) in the case of g-measures.

Theorem 5.4 (Billingsley dimensions of saturated sets). *If μ is a g-measure on Ω, then*

$$\forall\omega \in \Omega \qquad \dim_\mu \nabla\Delta(\omega) = \inf\left\{\frac{h(\eta)}{\eta(I_\mu^c)} : \eta \in \Delta(\omega)\right\}.$$

Proof. Let μ_n be the n-step Markov approximation of the g-measure μ. A simple reformulation of (21) gives for all integers n,

$$\dim_{\mu_n} \nabla\Delta(\omega) = \inf_{\eta\in\Delta(\omega)} \{\gamma(\eta|\mu_n)\} \tag{22}$$

with $\gamma(\eta|\mu_n) := h(\eta)/\eta(I^c_\mu)$ to ease the notation. By lemma 5.2, the g-measure μ is the q-limit of the Markov measures μ_n. It follows from (18) that

$$\dim_\mu \nabla\Delta(\omega) = \lim_{n\to+\infty} \dim_{\mu_n} \nabla\Delta(\omega) = \lim_{n\to+\infty} \inf_{\eta\in\Delta(\omega)} \{\gamma(\eta|\mu_n)\}. \tag{23}$$

In addition, there exists a constant K such that for any $\eta \in \mathcal{M}_\sigma$,

$$|\gamma_n(\eta|\mu) - \gamma(\eta|\mu_n)| = h(\eta)\left|\frac{1}{\eta(I^c_\mu)} - \frac{1}{\eta(I^c_{\mu_n})}\right| \leqslant h_{top}\, K\, \|I^c_\mu - I^c_{\mu_n}\|_\infty$$

which implies that $\gamma(\cdot|\mu_n)$ converges uniformly on $\mathcal{M}_\sigma$ to $\gamma(\cdot|\mu)$ by lemma 2.6. Therefore, the following commutation of symbols arises:

$$\lim_{n\to+\infty} \inf_{\eta\in\Delta(\omega)} \{\gamma(\eta|\mu_n)\} = \inf_{\eta\in\Delta(\omega)} \left\{ \lim_{n\to+\infty} \gamma(\eta|\mu_n) \right\} = \inf_{\eta\in\Delta(\omega)} \{\gamma(\eta|\mu)\}$$

and by comparison with statement (23), the proof is complete. □

6. Some applications and consequences

6.1. The relationship between relative entropy and Billingsley dimensions

For any shift-invariant measure η, let

$$G(\eta) := \left\{\omega :\ \forall f \in \mathcal{C}(\Omega),\ \lim_{n\to\infty} (1/n) S_n f(\omega) = \eta(f)\right\}$$

be the set of η-generic points. It is clear that for any $\omega \in G(\eta)$, $\Delta(\omega) = \{\eta\}$ and $\nabla\{\eta\} = G(\eta)$, which proves that $G(\eta)$ is a saturated set. Let us recall that if η and ν are two distinct ergodic measures, then $\nu(G(\eta)) = 0$. One has $\nu(G(\nu)) = 1$ if and only if ν is ergodic. (See [7] for details.) It is possible to give the following expression which only makes use of entropy and relative entropy:

Proposition 6.1 (Billingsley dimensions and relative entropy). *Let μ be a g-measure and η a shift-invariant measure. Then*

$$\dim_\mu G(\eta) = \frac{h(\eta)}{h(\eta) + h(\eta|\mu)}. \tag{24}$$

This makes an explicit link between Billingsley dimensions and relative entropy. When μ coincides with the Parry measure λ on Ω (which is the unique measure of maximal entropy), assertion (24) reduces to

$$h(\eta|\lambda) = h_{top}\big(1 - \dim_H G(\eta)\big) \tag{25}$$

since $P(0) = h_{top}$.

6.2. Multifractal spectrum for continuous functions

We now apply the large-deviation results of section 4 to some closed d^*-balls centred on an ergodic measure η. Define $B(\eta,\varepsilon) := \{\eta' \in \mathcal{M} \mid d^*(\eta,\eta') \leqslant \varepsilon\}$.

Proposition 6.2. *Let μ be a g-measure on Ω. For any ergodic measure η on Ω*

$$\lim_{\varepsilon\to 0^+}\lim_{n\to\infty}\frac{1}{n}\log\mu\{\omega : T_n(\omega)\in B(\eta,\varepsilon)\} = -h(\eta|\mu).$$

Proof. One has for any $\varepsilon > 0$, since $\overline{B(\eta,\varepsilon)^\circ} = B(\eta,\varepsilon)$:

$$\lim_{n\to\infty}\frac{1}{n}\log\mu\{\omega : T_n(\omega)\in B(\eta,\varepsilon)\} = -\inf\{\boldsymbol{H}_g(\eta') : \eta'\in B(\eta,\varepsilon)\} = -\boldsymbol{H}_g(\eta_\varepsilon)$$

for some element $\eta_\varepsilon \in B(\eta,\varepsilon)$ ($\boldsymbol{H}_g$ is lower semi-continuous on $\mathcal{M}_\sigma$ and $B(\eta,\varepsilon)$ is compact). Moreover, we have the following inequality:

$$-\boldsymbol{H}_g(\eta_\varepsilon) = \lim_{n\to\infty}\frac{1}{n}\log\mu\{\omega : T_n(\omega)\in B(\eta,\varepsilon)\} \geqslant -\boldsymbol{H}_g(\eta).$$

As the metric d^* is compatible with the weak* topology, η_ε tends to η in the weak* sense as ε goes to 0. By lower semi-continuity, $\boldsymbol{H}_g(\eta) \leqslant \liminf_{\varepsilon\to 0^+}\boldsymbol{H}_g(\eta_\varepsilon)$ and finally

$$\lim_{\varepsilon\to 0^+}\lim_{n\to\infty}\frac{1}{n}\log\mu\{\omega : T_n(\omega)\in B(\eta,\varepsilon)\} = -\boldsymbol{H}_g(\eta) = -h(\eta|\mu). \qquad \square$$

One can deduce large-deviation results for a time average of continuous functions by applying the so-called contraction principle (see, for instance, [8]). The trick is to note that for any continuous function ψ and any $\omega\in\Omega$, $\int\psi\,\mathrm{d}T_n(\omega) = (1/n)S_n\psi(\omega)$. Thus, defining the map from $\mathcal{M}$ to $\mathbb{R}$ such that $\nu\mapsto\int\psi\,\mathrm{d}\nu$ which is obviously continuous (with respect to weak* topology), one constructs by pull-back a deviation function $\tilde{\boldsymbol{H}}_g$ on $\mathbb{R}$ in the following way:

$$\forall\alpha\in\mathbb{R} \qquad \tilde{\boldsymbol{H}}_g(\alpha) = \begin{cases} \inf\{h(\eta|\mathcal{I}_{\phi_g}) : \eta(\psi)=\alpha\} & \text{if} \quad \alpha\in\,]\underline{\alpha}^\psi,\overline{\alpha}^\psi[\\ +\infty & \text{if} \quad \alpha\notin\,]\underline{\alpha}^\psi,\overline{\alpha}^\psi[. \end{cases}$$

The next lemma defines the real constants $\underline{\alpha}^\psi, \overline{\alpha}^\psi$. Let $\tilde{\boldsymbol{P}}_g : \mathbb{R}\to\mathbb{R}$, $\tilde{\boldsymbol{P}}_g(\beta) := P(\phi_g+\beta\psi)$ which is a convex function. Then we have the following lemma.

Lemma 6.3 (Interval of the spectrum). *Let g be an element of $\mathcal{G}$, ψ a continuous function on Ω and $\tilde{\boldsymbol{P}}_g$ defined as above. Then*

(i)
$$\frac{\mathrm{d}^+\tilde{\boldsymbol{P}}_g}{\mathrm{d}\beta}(\beta) = \sup\{\eta(\psi) : \eta\in\mathcal{I}_{\phi_g+\beta\psi}\}$$
$$\frac{\mathrm{d}^-\tilde{\boldsymbol{P}}_g}{\mathrm{d}\beta}(\beta) = \inf\{\eta(\psi) : \eta\in\mathcal{I}_{\phi_g+\beta\psi}\}$$

(ii)
$$\lim_{\beta\to+\infty}\frac{\mathrm{d}^+\tilde{\boldsymbol{P}}_g}{\mathrm{d}\beta}(\beta) = \sup_{\eta\in\mathcal{M}_\sigma}\{\eta(\psi)\} =: \overline{\alpha}^\psi$$
$$\lim_{\beta\to-\infty}\frac{\mathrm{d}^-\tilde{\boldsymbol{P}}_g}{\mathrm{d}\beta}(\beta) = \inf_{\eta\in\mathcal{M}_\sigma}\{\eta(\psi)\} =: \underline{\alpha}^\psi.$$

Proof. The proof follows easily from the characterization of equilibrium states as tangent functional to the pressure (that is, we have $\mu \in \mathcal{I}_\phi \iff \forall \xi \in \mathcal{C}(\Omega),\ P(\phi+\xi)-P(\phi) \geqslant \mu(\phi)$, where ϕ is a given potential; see [16]). □

It is readily checked that $\tilde{\boldsymbol{P}}_g$ and $\tilde{\boldsymbol{H}}_g$ are conjugate by the Legendre–Fenchel transformation. One obtains the following easy corollary of proposition 6.2 and lemma 6.3:

Corollary 6.4. *Let g be a g-function on Ω, μ an associated g-measure and ψ a potential on Ω. Then, for any $\alpha \in]\underline{\alpha}^\psi, \overline{\alpha}^\psi[$,*

$$\lim_{\varepsilon\to 0^+}\lim_{n\to+\infty}\frac{1}{n}\log\mu\left\{\omega : \frac{1}{n}S_n\psi(\omega) \in [\alpha-\varepsilon, \alpha+\varepsilon]\right\} = \tilde{H}_g(\alpha).$$

Now suppose that for a given $\beta \in \mathbb{R}$, the simplex $\mathcal{I}_{\phi_g+\beta\psi}$ is reduced to a single point denoted by μ_β which is necessarily ergodic. This means that for the underlying thermodynamic formalism, there are no phase transitions at inverse temperature β. By lemma 6.3, $\mu_\beta(\psi) := \alpha_\beta$ for some $\alpha_\beta \in]\underline{\alpha}^\psi, \overline{\alpha}^\psi[$. For any shift-invariant measure η on Ω such that $\eta(\psi) = \alpha_\beta$, it follows from the variational principle that

$$h(\eta) + \eta(\phi_g+\beta\psi) \geqslant h(\mu_\beta) + \mu_\beta(\phi_g+\beta\psi) \qquad \text{i.e.} \quad h(\mu_\beta|\mu) \leqslant h(\eta|\mu). \tag{26}$$

By definition of $\tilde{\boldsymbol{H}}$, this implies that $h(\mu_\beta|\mu) = \tilde{\boldsymbol{H}}(\alpha_\beta)$ and corollary 6.4 gives

$$\lim_{\varepsilon\to 0^+}\lim_{n\to+\infty}\frac{1}{n}\log\mu\left\{\omega : \frac{1}{n}S_n\psi(\omega) \in [\alpha_\beta-\varepsilon, \alpha_\beta+\varepsilon]\right\} = h(\mu_\beta|\mu). \tag{27}$$

Now define, for any $\alpha \in]\underline{\alpha}^\psi, \overline{\alpha}^\psi[$, the level sets

$$E_\alpha(\psi) := \left\{\omega\in\Omega : \lim_{n\to+\infty}\frac{1}{n}S_n\psi(\omega) = \alpha\right\}. \tag{28}$$

Our aim is to compute $\dim_H E_\alpha(\psi)$. We have the following proposition:

Proposition 6.5. *Let ψ be a Hölder continuous function. Then, for any $\alpha \in]\underline{\alpha}^\psi, \overline{\alpha}^\psi[$,*

$$\dim_H E_\alpha(\psi) = 1 + \frac{1}{h_{top}}\lim_{\varepsilon\to 0^+}\lim_{n\to+\infty}\frac{1}{n}\log\lambda\left\{\omega : \frac{1}{n}S_n\psi(\omega) \in [\alpha-\varepsilon, \alpha+\varepsilon]\right\}.$$

Proof. The constant function $g_0 = 1/k$ is trivially a g-function on Ω and the unique associated g-measure is the Parry measure λ. Assume that ψ is normalized (which is not a restriction). The potentials $\phi_{g_0} + \beta\psi$ and $\beta\psi$ are cohomologous for all $\beta \in \mathbb{R}$, so they admit a unique common equilibrium state μ_β which is a Bowen Gibbs measure. If ψ is not cohomologous to 0, then $\beta \mapsto \mu_\beta(\psi)$ is an analytic strictly increasing map from $\mathbb{R}$ onto $]\underline{\alpha}^\psi, \overline{\alpha}^\psi[$ and thus there exists an analytic one-to-one map b from $]\underline{\alpha}^\psi, \overline{\alpha}^\psi[$ onto $\mathbb{R}$, such that $\mu_{b(\alpha)}(\psi) = \alpha$, for all $\alpha \in]\underline{\alpha}^\psi, \overline{\alpha}^\psi[$. In this case, for any given $\alpha \in]\underline{\alpha}^\psi, \overline{\alpha}^\psi[$, statement (27) becomes

$$\lim_{\varepsilon\to 0^+}\lim_{n\to+\infty}\frac{1}{n}\log\lambda\left\{\omega : \frac{1}{n}S_n\psi(\omega) \in [\alpha_\beta-\varepsilon, \alpha_\beta+\varepsilon]\right\} = h(\mu_{b(\alpha)}|\lambda) \tag{29}$$

that is, by using (25),

$$\dim_H G(\mu_{b(\alpha)}) = 1 + \frac{1}{h_{top}}\lim_{\varepsilon\to 0^+}\lim_{n\to+\infty}\frac{1}{n}\log\lambda\left\{\omega : \frac{1}{n}S_n\psi(\omega) \in [\alpha-\varepsilon, \alpha+\varepsilon]\right\}. \tag{30}$$

It is clear that

$$G(\mu_{b(\alpha)}) \subset E_\alpha(\mu_\psi) = \bigcap_{\varepsilon>0}\bigcup_{m>0}\bigcap_{n\geqslant m}\left\{\omega : \frac{1}{n}S_n\psi(\omega) \in [\alpha-\varepsilon, \alpha+\varepsilon]\right\}. \tag{31}$$

A study of such multifractal problems can be found in [12] and it is proved that, in fact, $\dim_H E_\alpha(\psi) = \dim_H G(\mu_{b(\alpha)})$. This concludes the proof □

6.3. Hausdorff dimension of sets of non-generic points

Non-generic points are related to the irregular part of the multifractal spectrum, that is the part of the spectrum for which level sets are not defined.

For each continuous function f on Ω, define the following set:

$$Ir(f) := \left\{\omega : \frac{S_n f(\omega)}{n} \text{ does not converge}\right\}.$$

Take the union of such sets over all continuous functions:

$$\mathfrak{I} = \bigcup_{f\in\mathcal{C}(\Omega)} Ir(f).$$

This is an invariant but not compact set which has measure zero for any ergodic measure (thanks to the Birkhoff ergodic theorem and separability of $\mathcal{C}(\Omega)$). By definition,

$$\Omega = \mathfrak{G} \cup \mathfrak{I} \qquad \text{with} \quad \mathcal{G} \cap \mathfrak{I} = \emptyset$$

where $\mathfrak{G}$ is the set of all generic points in Ω, i.e.

$$\mathfrak{G} = \bigcup_{\mu\in\mathcal{M}_\sigma} G(\mu) = \{\omega : \forall f \in \mathcal{C}(\Omega),\ (1/n)S_n f(\omega) \text{ converges}\}.$$

Now we can state the following proposition:

Proposition 6.6 (Non-generic points have full Hausdorff dimension). *Let $\mathfrak{I}$ the set defined above. Then*

$$\dim_H \mathfrak{I} = 1.$$

Proof. Let $(\mu_n)_n$ be a sequence of non-atomic measures converging weakly to λ and such that $h(\mu_n)$ converges to $h(\lambda) = h_{top}$. Then consider, for any $n \geqslant 1$, the segment $[\mu_n, \lambda] \subset \mathcal{M}_\sigma$, that is the measures of the form $s\mu_n + (1-s)\lambda,\ s \in [0,1]$. This segment is clearly a non-empty closed connected subset of $\mathcal{M}_\sigma$, so from [7] we know there exists an $\omega \in \Omega$ such that $\Delta(\omega) = [\mu_n, \lambda]$ for any $n \geqslant 1$. Moreover, the set of such ω is dense in Ω and clearly $\nabla[\mu_n, \mu] \subset \mathfrak{I}$. Thus

$$\dim_H \nabla[\mu_n, \lambda] \leqslant \dim_H \mathfrak{I} \qquad \text{for any} \quad n \geqslant 1.$$

Now, by theorem 5.4 and since the map $\nu \mapsto h(\nu)$ is affine on $\mathcal{M}_\sigma$, it follows that

$$\dim_H \nabla[\mu_n, \lambda] = \frac{\min(h(\mu_n), h_{top})}{h_{top}}.$$

By virtue of the variational principle, $h(\mu_n) < h_{top}$, which implies that

$$\dim_H \nabla[\mu_n, \lambda] = \frac{h(\mu_n)}{h_{top}}.$$

However, for any $\varepsilon > 0$, there exists $N = N(\varepsilon)$ such that $h(\mu_n)/h_{top} > 1 - \varepsilon$ for any $n \geqslant N$, so we obtain $\dim_H \mathfrak{I} \geqslant 1 - \varepsilon$. Since ε is arbitrary, the proof is complete. □

6.4. Final comments

Let ψ be Hölder continuous as above. It can be readily checked that

$$\lim_{n\to\infty} \frac{\log \mu_\psi[\pi_n \cdot \omega]}{\log \lambda[\pi_n \cdot \omega]} = \alpha \tag{32}$$

if and only if $\lim_{n\to\infty} (1/n) \log \mu_\psi[\pi_n \cdot \omega] = \alpha \cdot h_{top}$. One can replace $\log \mu_\psi[\pi_n \cdot \omega]$ by $S_n \psi(\omega)$ when ψ is normalized. The quantity (32) is the pointwise dimension of the measure μ_ψ. Thus, multifractal analysis of pointwise dimensions of Gibbs measures and Birkhoff averages of Hölder continuous functions are the same.

A natural question is to perform the multifractal analysis of pointwise Billingsley dimensions defined by replacing in the preceding formula μ_ψ by any ergodic measure η, and λ by any g-measure. Theorem 3.2 shows that this quantity is well defined almost everywhere and formula (24) gives $\dim_\mu G(\eta)$ for its almost-everywhere value. The problem is to obtain the Billingsley dimension of level sets of this pointwise dimension and to study irregular points which are conjectured to be of full dimension.

With the choice of the metric d (see section 5), it is well known that $h_{top}(\sigma|A) = \dim_H A \cdot h_{top}$, for any A. Thus proposition 6.6 reads $h_{top}(\sigma|\mathfrak{I}) = h_{top}$. Moreover, one obtains immediately from (6.5), up to a multiplicative factor, the topological entropy of the level sets $E_\alpha(\psi)$.

References

[1] Berbee H 1987 Chains with infinite connections: uniqueness and Markov representation *Prob. Theor. Rel. Fields* **76** 243–53

[2] Bowen R *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Mathematics vol 470)* (Berlin: Springer)

[3] Bramson M and Kalikow S A 1993 Nonuniqueness in g-functions *Israel J. Math.* **84** 153–60

[4] Bressaud X, Fernández R and Galves A 1999 Speed of $\overline{d}$-convergence for Markov approximations of chains with complete connections. A coupling approach *Electron. J. Probab.* **4**

[5] Cajar H 1981 *Billingsley Dimension in Probability Spaces (Lecture Notes in Mathematics vol 892)* (Berlin: Springer)

[6] Chazottes J R, Floriani E and Lima R 1998 Relative entropy and identifications of Gibbs measures in dynamical systems *J. Stat. Phys.* **90** 697–725

[7] Denker M, Grillenberger C and Sigmund K 1976 *Ergodic Theory on Compact Spaces (Lecture Notes in Mathematics vol 527)* (Berlin: Springer)

[8] Deuschel J D and Strook D W 1989 *Large Deviations* (New York: Academic)

[9] Ellis R S 1995 *Entropy, Large Deviations and Statistical Mechanics* (Berlin: Springer)

[10] Kifer Y 1990 Large deviations results for dynamical systems and stochastic processes *Trans. Am. Math. Soc.* **321** 505–24

[11] Ledrappier F 1974 Principe variationnel et systèmes dynamiques symboliques *Z. Wahr. Verw. Geb.* **30** 185–202

[12] Olivier E 1998 Analyse multifractale de fonctions continues *C. R. Acad. Sci. (Paris)* **326** 117–4

[13] Palmer M R, Parry W and Walters P 1977 *Large Sets of Endomorphis and of g-Measures (Lecture Notes in Mathematics vol 668)* pp 191–210

[14] Pesin Ya B 1997 *Dimension Theory in Dynamical Systems (Chicago Lectures in Mathematics)* (Chicago, IL: University of Chicago Press)

[15] Quas A 1996 Non ergodicity of C^1 expanding maps and g-measures *Ergod. Theory Dynam. Syst.* **16** 531–5

[16] Ruelle D 1978 Thermodynamic formalism: the mathematical structure of classical equilibrium statistical mechanics *Encyclopedia of Mathematics and its Applications* vol 5 (Reading, MA: Addison-Wesley)

[17] Shereshevesky M A 1991 A complement to Young's theorem on measure dimension: the difference between lower and upper pointwise dimensions *Nonlinearity* **4** 15–25