Записки научных
семинаров ПОМИ
Том 436, 2015 г.

A. Alpeev

THE ENTROPY OF GIBBS MEASURES ON SOFIC GROUPS

Abstract. We show that for every local potential on a sofic group there exists a shift-invariant Gibbs measure. Under some condition we show that the sofic entropy of the corresponding shift action does not depend on a sofic approximation.

§1. Introduction

In the work [1], Lewis Bowen made a great progress in the isomorphism problem for Bernoulli shifts by defining so-called sofic entropy. This led to a great line of research. There is a very important question which is still far from being resolved. The value of the sofic entropy depends a priori on a sofic approximation. In the classical realm of amenable entropy, the dependence on the choice of a Følner sequence is eliminated by means of the Ornstein–Weiss covering argument. In the case of sofic entropy, we still do not have an analogous result. In numerous examples, the sofic entropy is indepedent of the choice of a sofic approximation. For Bernoulli shifts, it was proved in the very paper [1] (see also [11]); later, Hayes proved in [9] the entropy formula for a class of algebraic actions over sofic groups, which, in particular, implies that the sofic entropy for these actions is independent of the choice of a sofic approximation. There are also results in other directions. In the work [4], Carderi constructed examples of actions having different values of the sofic entropy for different sofic approximations. Nevertheless, it is still not known whether an action can have two different nonnegative values of the sofic entropy.

In this paper, we will work with so-called Gibbs measures on sofic groups. Gibbs measures in general are extensively studied, as well as some particular examples like the Ising model, Potts model, etc. We will prove the following theorem.

Key words and phrases: Gibbs measure, sofic entropy, Dobrushin uniqueness condition.

This research is supported by the RSF grant 14-11-00581.

Theorem 1. *Let G be a sofic group, (X, ρ, μ) be a compact measure-metric space, and φ be a continuous local potential on X^G. Then there exists a shift-invariant Gibbs measure on X^G for the potential φ.*

The proof involves a construction of a sequence of Gibbs measures on finite models coming from sofic approximations. It turns out that this sequence of measures comes very handy for the computation of the modified sofic entropy.

We would like to note that in [5, Sec. 5] Chung essentially considered a Gibbs measure in a very special case when the potential φ is such that $\varphi(y)$ depends only on the value of $y \in X^G$ in e, the group identity.

The remaining part of the paper is devoted to the proof of the following theorem.

Theorem 2. *Let G be a sofic group and (X, μ, ρ) be a finite metric space with the uniform probability measure. If a local potential ψ on X^G is such that $\beta\psi$ has a unique Gibbs measure for every $\beta \in [0, 1]$, then the modified sofic entropy of the corresponding shift action is the same for every sofic approximation.*

The definition of sofic entropy we employ in most parts of the paper is not the standard one. This modified sofic entropy was essentially defined by Bowen in [2]; in addition, he essentially proved that for ergodic actions this new entropy coincides with the standard one. For more details the reader is referred to Sec. 4 of [2], Sec. 7 of [3], and Sec. 4 of [9]. It is a standard fact from the theory of Gibbs measures that if a shift-invariant Gibbs measure is unique for a potential, then it is ergodic; see Sec. 3.6 for greater details. So we have that the original sofic entropy does not depend on a sofic approximation.

The conditions of Theorem 2 may sound too restrictive, but the famous Dobrushin uniqueness condition [6] shows that for a moderately large set of potentials this condition is satisfied. For more details, see Sec. 8.

Acknowledgments. I would like to thank my advisor Prof. Anatoly M. Vershik for valuable comments.

§2. The structure of the paper

In Sec. 3, we introduce standard definitions. In Sec. 4, we show that for every local potential on a sofic group there is at least one invariant Gibbs measure. In Sec. 5, we discuss some important applications of the uniqueness of a Gibbs measure. It mostly consists of technical lemmas showing that some maps are continuous and some limits are uniform. In Sec. 6, we define pressure and infer a formula for it. The definition of pressure a priori involves a sofic approximation, but the formula we prove does not, which clears the road for our main result. In Sec. 7, we express entropy in terms of pressure (in a way that does not involve a sofic approximation), which allows us to finish the proof of Theorem 2. In Sec. 8, we discuss the Dobrushin uniqueness condition.

§3. Definitions and conventions

For a metric compact space X, we will denote by $\mathcal{M}(X)$ the set of all Borel probability measures on X. This set will be endowed with the weak-* topology unless otherwise stated explicitly. We will denote by $\|\mu\|_V = \sup_{\|f\|_\infty=1}\{\int\limits_X f d\mu\}$ the total variation norm on the set of all Borel measures. The symbol $\Subset$ will stand for "a finite subset."

3.1. Shift spaces. Let G be a countable group. Let (X, ρ, μ) be a compact measure-metric space with a metric ρ and a measure μ. Consider the set X^G endowed with the product topology and a compatible metric r. We define the *shift action* of G on X^G by the formula $(gy)(h) = y(hg)$ for $y \in X^G$ and $g, h \in G$. If ν is a Borel probability measure on X^G, we will denote by $g_\sharp(\nu)$ the pushforward of the measure ν along the map g. For $x \in X^G$ and $A \subset G$, we will denote by $x|_A \in X^A$ the restriction of x to A. If A, B are disjoint subsets of G and $x \in X^A$, $y \in X^B$, we will denote by $x \sqcup y \in X^{A\sqcup B}$ their gluing.

3.2. The Kantorovich distance. The Kantorovich distance provides a very convenient language for working with the weak-* topology on the set of probability measures on a metric compact space. It was first defined in [10]; see also [12] for the history of the question. Consider the set $\mathcal{M}(X)$ of all Borel probability measures on X endowed with the weak-* topology. For every pair $\mu_1, \mu_2 \in \mathcal{M}(X)$, consider the set of all couplings, that is, the set of all measures $\xi \in \mathcal{M}(X \times X)$ such that ${\mathrm{pr}_1}_\sharp(\xi) = \mu_1$ and ${\mathrm{pr}_2}_\sharp(\xi) = \mu_2$ (where pr_i denotes the standard coordinate projection). The *Kantorovich* (or *mass transportation*) *distance* is defined as

$$\Delta(\mu_1, \mu_2) = \inf_\xi \left\{ \int\limits_{X\times X} r(x_1, x_2) d\xi(x, y) \right\}.$$

It is a standard fact that the Kantorovich distance is indeed a metric and that it is compatible with the weak-* topology. We also have that $\operatorname{diam}(\mathcal{M}(X,\rho),\Delta) = \operatorname{diam}(X,\rho)$. The Kantorovich distance enjoys a very nice convexity inequality:

$$\Delta\left(\sum_i \alpha_i \mu_{1,i}, \sum_i \alpha_i \mu_{2,i}\right) \leqslant \sum_i \alpha_i \Delta(\mu_{1,i}, \mu_{2,i})$$

for $\mu_{j,i} \in \mathcal{M}(X)$ and $\alpha_i > 0$, $\sum\limits_i \alpha_i = 1$. It is easy to prove the following estimate:

$$\Delta(\mu_1, \mu_2) \leqslant \frac{\operatorname{diam}(X)\|\mu_1 - \mu_2\|_V}{2}.$$

We will denote by $l(\nu_1, \nu_2)$ the Kantorovich distance between measures in $\mathcal{M}(X^G, r)$.

3.3. Sofic groups. We start with the definition of a sofic approximation. For a positive integer n, we denote $[n] = \{1, \ldots, n\}$. Let $\operatorname{Sym}(n)$ be a symmetric group considered with its standard action on $[n]$. We endow this group with the so-called *Hamming distance*:

$$d(p_1, p_2) = n^{-1} \left|\{k \in [n], \quad p_1(k) \neq p_2(k)\}\right|.$$

Let G be a countable group. Let $\Sigma = \{\sigma_i\}_{i=1}^{\infty}$ be a sequence of maps $\sigma_i : G \to \operatorname{Sym}(s_i)$ (not necessarily homomorphisms). We will say that it is a *sofic approximation* if

1) for any $g, h \in G$, $g \neq h$, we have $d_{s_i}(\sigma_i(g), \sigma_i(h)) \to 1$;
2) for any $g, h \in G$, we have $d_{s_i}(\sigma_i(gh), \sigma_i(g)\sigma_i(h)) \to 0$.

We will say that a group G is a *sofic group* if it has at least one sofic approximation. From now on, let G be a sofic group and fix some sofic approximation Σ.

Now we will introduce some useful notation. For a positive integer i and $k \in [s_i]$, denote by $\theta_{i,k} : X^{[s_i]} \to X^G$ the map defined by the formula $(\theta_{i,k}(\tau))(g) = \tau((\sigma_i(g))(k))$ for $g \in G$ and $\tau \in X^{[s_i]}$. We also define a map $\Theta_i : \mathcal{M}(X^{[s_i]}) \to \mathcal{M}(X^G)$ by the formula

$$\Theta_i(\eta) = \frac{1}{s_i} \sum_{k \in [s_i]} \theta_{i,k_\sharp}(\eta).$$

It is conjugate to the map $\Theta_i^* : C(X^G) \to C(X^{[s_i]})$ defined by the equation

$$(\Theta_i^*(\varphi))(\tau) = \frac{1}{s_i} \sum_{k \in [s_i]} \varphi(\theta_{i,k}(\tau));$$

it is obvious that

$$\int_{X^G} \varphi(x) d(\Theta_i(\eta))(x) = \int_{X^{s_i}} (\Theta_i^*(\varphi))(\tau) d\eta(\tau).$$

3.4. Sofic entropy. Let Ω be a finite or countable set. Let $\eta \in \mathcal{M}(\Omega)$. The *Shannon entropy* of η is defined as

$$H(\eta) = -\sum_{\omega \in \Omega} \eta(\{\omega\}) \log \eta(\{\omega\}),$$

with the usual convention that $0 \log 0 = 0$.

Let ν be any invariant probability measure on X^G. For a positive integer i and $\delta > 0$, we denote by $\mathrm{Appr}_{i,\delta}(\nu)$ the set of all probability measures m on $X^{[s_i]}$ such that $l(\Theta_i(m), \nu) < \delta$. Then we define

$$h_{i,\delta}(\nu) = s_i^{-1} \sup\{H(m) | m \in \mathrm{Appr}_{i,\delta}(\nu)\}.$$

If $\mathrm{Appr}_{i,\delta}$ is empty, we set $h_{i,\delta}(\nu) = -\infty$. We define

$$h_\delta(\nu) = \limsup_{i \to \infty} h_{i,\delta}(\nu),$$
$$h(\nu) = \inf_{\delta > 0} h_\delta(\nu).$$

The latter quantity will be called the *modified sofic entropy.* It is always not smaller than the original sofic entropy (for the same sofic approximation), and it coincides with the original sofic entropy for ergodic actions (see Sec. 4 of [2], Sec. 7 of [3], and Sec. 4 of [9]).

3.5. Probability kernels. Let X and Y be two metric compact spaces. A *probability kernel* is a map $\pi : \mathcal{M}(X) \to \mathcal{M}(Y)$ defined by the identity

$$\int_Y \varphi(y) d(\pi(\nu))(y) = \int_X d\nu(x) \int_Y \varphi(y) d(\bar{\pi}(x))(y)$$

for some continuous map $\bar{\pi} : X \to \mathcal{M}(Y)$. We will denote by $\mathrm{Ker}(X, Y)$ the set of all probability kernels from X to Y. It is easy to see that the map π is continuous and affine. Now we will define two topologies on the set of probability kernels. The first one is the weak topology. Let Δ be the

Kantorovich distance on $\mathcal{M}(Y)$. Let $\pi_1, \pi_2 \in \operatorname{Ker}(X, Y)$. We define the weak distance by the formula

$$L_{\text{weak}}(\pi_1, \pi_2) = \sup_{x \in X} \Delta(\pi_1(\delta_x), \pi_2(\delta_x))$$

and the strong distance by the formula

$$L_{\text{strong}}(\pi_1, \pi_2) = \sup_{x \in X} \|(\pi_1(\delta_x), \pi_2(\delta_x)\| .$$

It is obvious that the strong distance induces a stronger topology than the weak distance does. It is not hard to prove the following estimate:

$$\Delta(\pi_1(\nu), \pi_2(\nu)) \leqslant L_{\text{weak}}(\pi_1, \pi_2)$$

for any $\nu \in \mathcal{M}(X)$ and any $\pi_1, \pi_2 \in \operatorname{Ker}(X, Y)$. Thus, it is easy to see that the map $\mathcal{M}(X) \times \operatorname{Ker}(X, Y) \to \mathcal{M}(Y)$ defined as $(\nu, \pi) \mapsto \pi(\nu)$ is continuous (here $\operatorname{Ker}(X, Y)$ is endowed with the weak topology).

3.6. Gibbs measures. For a detailed exposition, the reader is referred to [6–8]. A *potential* is a continuous function $\varphi : X^G \to \mathbb{R}$ that depends only on finitely many coordinates of X^G; that is, there exists a finite subset A of G such that for any $x, x' \in X^G$ coinciding outside of A we have $\varphi(x) = \varphi(x')$. We will denote by $\operatorname{supp} \varphi$ the minimal subset with this property. Now we would like to define the set of Gibbs measures correspoding to the potential φ. For every finite subset Λ of G, any $y \in X^{\Lambda^c}$ and $x \in X^{\Lambda}$, let us define the Hamiltonian

$$\mathcal{H}^y_{\varphi,\Lambda}(x) = \sum_{g \in G, (\operatorname{supp} \varphi) g \cap \Lambda \neq \varnothing} \varphi(g(x \sqcup y))$$

and the partition function

$$\mathcal{Z}^y_{\varphi,\Lambda} = \int\limits_{X^\Lambda} e^{-\mathcal{H}^y_{\varphi,\Lambda}(x)} d\mu^{\otimes \Lambda}(x).$$

Then for every $\Lambda \Subset G$ we introduce a probability kernel

$$\pi_{\varphi,\Lambda} : \mathcal{M}(X^G) \to \mathcal{M}(X^G)$$

by defining its values on δ-measures:

$$\pi_{\varphi,\Lambda} : \delta_x \mapsto (\mathcal{Z}^{x|_{\Lambda^c}}_{\varphi,\Lambda})^{-1} e^{-\mathcal{H}^{x|_\Lambda}_{\varphi,\Lambda}(t)} (d\mu(t)^{\otimes \Lambda} \otimes \delta_{x|_{\Lambda^c}}).$$

It is not hard to see that $\pi_{\Lambda'} \circ \pi_\Lambda = \pi_\Lambda \circ \pi_{\Lambda'} = \pi_\Lambda$ for every $\Lambda' \subset \Lambda \Subset G$. We will say that ν is a *Gibbs measure* for the potential φ if $\pi_{\varphi,\Lambda}(\nu) = \nu$ for every $\Lambda \Subset G$. The set of all Gibbs measures for φ will be denoted by

$\mathcal{G}_\varphi(X,G)$, or simply by $\mathcal{G}_\varphi$. A simple compactness argument shows that there always exists at least one Gibbs measure. We will say that φ is a *unique Gibbs measure (UGM) potential* if there is only one Gibbs measure for φ.

Lemma 1. *If ν_φ is a unique Gibbs measure for a potential φ, then the corresponding shift action is ergodic.*

Proof. Consider the so-called tail σ-algebra, that is,

$$\mathcal{T} = \bigcap_{\Lambda \Subset G} \mathscr{B}_{\Lambda^c},$$

where $\mathscr{B}$ stands for the σ-subalgebra on X^G generated by the cylinder sets supported on Λ^c (σ-subalgebras and their intersections are considered in the ν_φ-mod 0 sense). It is not hard to see that this subalgebra is trivial ν_φ-mod 0. By Proposition 14.9 of [8], we have that the subalgebra of invariant sets is also trivial ν_φ-mod 0, which means exactly that the shift action is ergodic. □

§4. The existence of a shift-invariant measure

This section is devoted to the proof of Theorem 1.

Lemma 2. *For any $\varepsilon' > 0$ and any $g \in G$ there is a positive integer N such that for any $i > N$ we have*

$$|\{k \in [s_i] | r(g(\theta_{i,k}(\tau)), \theta_{i,(\sigma_i(g))(k)}(\tau)) < \varepsilon' \text{ for any } \tau \in X^{[s_i]}\}| > (1-\varepsilon')s_i.$$

Proof. Let us find a finite subset F of G such that $r(x,y) < \varepsilon'$ for any two points $x, y \in X^G$ such that $x|_F = y|_F$. By the definition of a sofic approximation, there is N such that for every $i > I$ the set

$$A_i = \{k \in [s_i] | (\sigma_i(hf))(k) = (\sigma_i(f) \circ \sigma_i(f))(k) \text{ for every } f \in F\}$$

has at least $(1-\varepsilon')s_i$ elements. It is now easy to see that for every $\tau \in X^{[s_i]}$, $k \in A_i$ and for every $f \in F$ we have $((g\theta_{i,k}(\tau))(f) = (\theta_{i,(\sigma_i(g))(k)}(\tau))(f)$, since the left-hand side equals $(\theta_{i,k}(\tau))(fg) = \tau((\sigma_i(fg))(k))$ and the right-hand side equals $\tau(((\sigma_i(f)) \circ (\sigma_i(g)))(k))$. Thus, we have

$$r(g(\theta_{i,k}(\tau)), \theta_{i,(\sigma_i(g))(k)}(\tau)) < \varepsilon'. \qquad \square$$

Lemma 3. *For any $\varepsilon > 0$ and any $g \in G$ there is a positive integer N such that for any $i > N$ and for any measure η on $X^{[s_i]}$ we have $l(g_\sharp(\Theta_i(\eta)), \Theta_i(\eta)) < \varepsilon$.*

Proof. Let us take $\varepsilon' > 0$ such that $\varepsilon < \varepsilon'(1+\operatorname{diam}(X^G, r))$. The application of the previous lemma gives us a number N. Let us fix any $i > N$. By the properties of the Kantorovich distance and the previous lemma, the size of the set of $k \in [s_i]$ such that $l(g_\sharp\theta_{i,k}(\eta)_\sharp, \theta_{i,(\sigma_i(g))(k)}) < \varepsilon'$ is greater than $(1-\varepsilon')s_i$. Denote this set by A_i. It is easy to see that

$$l(g_\sharp(\Theta_i(\eta)), \Theta_i(\eta)) = l(s_i^{-1} g_\sharp \sum_{k\in[s_i]} \theta_{i,k_\sharp}(\eta)), s_i^{-1} \sum_{k\in[s_i]} \theta_{i,k_\sharp}(\eta))$$

$$= l(s_i^{-1} \sum_{k\in[s_i]} g_\sharp(\theta_{i,k_\sharp}(\eta)), s_i^{-1} \sum_{k\in[s_i]} \theta_{i,(\sigma_i(g))(k)_\sharp}(\eta))$$

$$\leqslant s_i^{-1} \sum_{k\in[s_i]} l(g_\sharp(\theta_{i,k}(\eta)), \theta_{i,(\sigma_i(g))(k)}(\eta)) \leqslant \frac{|A|}{s_i}\varepsilon' + \frac{s_i - |A|}{s_i}\operatorname{diam}(X^G, r)$$

$$\leqslant \varepsilon'(1+\operatorname{diam}(X^G, r)) < \varepsilon. \quad \square$$

Let

$$Z_{i,\varphi} = \int\limits_{X^{[s_i]}} e^{-(\Theta_i^*(\varphi))(\tau)} d\mu^{\otimes[s_i]}(\tau).$$

Let

$$Z_{i,\varphi} = \int_{X^{[s_i]}} e^{-(\Theta_i^*(\varphi))(\tau)} d\mu^{\otimes[s_i]}(\tau).$$

Then let $\eta_{i,\varphi} = e^{-(\Theta_i^*(\varphi))(\tau)} Z_{i,\varphi}^{-1} d\mu^{\otimes[s_i]}(\tau)$. Let $\nu_{i,\varphi} = \Theta_i(\eta_{i,\varphi})$.

Lemma 4. *For any $\varepsilon' > 0$ and $\Lambda \Subset G$ there is a positive integer N such that for any $i > N$ the inequality*

$$|\{k | \pi_{\varphi,\Lambda}(\theta_{i,k_\sharp}(\eta_{i,\varphi})) = \theta_{i,k_\sharp}(\eta_{i,\varphi})\}| > (1-\varepsilon')s_i$$

holds.

Proof. For convenience, denote $T = \operatorname{supp} f$; also denote

$$D = (\Lambda \cup \Lambda^{-1} \cup T \cup T^{-1})^5.$$

For a fixed positive integer i, let A_i be the set of $k \in [s_i]$ such that the map $g \mapsto (\sigma_i(g))(k)$ is injective from D and that $(\sigma_i(g) \circ \sigma_i(h))(k) = (\sigma_i(gh))(k)$. It easily follows from the definition of a sofic approximation that $|A_i|/s_i \to 1$. It is a simple exercise to see that for $k \in A_i$ we have $\pi_{\varphi,\Lambda}(\theta_{i,k_\sharp}(\eta_{i,\varphi})) = \theta_{i,k_\sharp}(\eta_{i,\varphi})$. $\square$

Lemma 5. *For any $\Lambda \Subset G$ and $\varepsilon > 0$ there is N such that*

$$l(\pi_{\varphi,\Lambda}(\nu_{i,\varphi}), \nu_{i,\varphi}) < \varepsilon$$

for every $i > N$.

Proof. Let $\varepsilon' > 0$ be such that $\varepsilon < \varepsilon' \operatorname{diam}(X^G, r)$; then we apply the previous lemma. Let $i > N$. Let A be the set of $k \in [s_i]$ such that $\pi_{\varphi,\Lambda}(\theta_{i,k_\sharp}(\eta_{i,\varphi})) = \theta_{i,k_\sharp}(\eta_{i,\varphi})$. Its size is at least $(1-\varepsilon')s_i$, so

$$\begin{aligned} l(\pi_{\varphi,\Lambda}(\nu_{i,\varphi}), \nu_{i,\varphi}) &\leqslant s_i^{-1} \sum_{k \in [s_i]} l(\pi_{\varphi,\Lambda}(\theta_{i,k_\sharp}(\eta_{i,\varphi})), \theta_{i,k_\sharp}(\eta_{i,\varphi})) \\ &\leqslant \varepsilon' \operatorname{diam}(X^G, r) < \varepsilon. \end{aligned}$$ □

Let ν_φ be any accumulation point for $\nu_{i,\varphi}$. The following lemma completes the proof of Theorem 1.

Lemma 6. *The measure ν_φ is a shift-invariant Gibbs measure for the potential φ.*

In the case where φ is a UGM potential, we denote by ν_φ its unique (and, obviously, invariant) Gibbs measure.

§5. Some implications of the uniqueness of a Gibbs measure

We will denote by $\operatorname{Pot}(X, G)$ the set of local potentials on X^G. Let us introduce the following seminorm on the set $\operatorname{Pot}(X, G)$:

$$\|\varphi\|_{\mathrm{var}} = \sum_{g \in G} \sup\left\{ |\varphi(x) - \varphi(y)|, \quad x|_{G\setminus\{g\}} = y|_{G\setminus\{g\}} \right\}.$$

We will call it the *variational seminorm.*

Lemma 7. *The map $(\varphi, m) \mapsto \pi_{\varphi,\Lambda}(m)$ is continuous.*

Proof. It is easy to verify that the map $\operatorname{Pot}(X, G) \to \operatorname{Ker}(X, Y)$ defined as $\varphi \mapsto \pi_{\varphi,\Lambda}$ is continuous if $\operatorname{Ker}(X, Y)$ is endowed with the strong topology and $\operatorname{Pot}(X, G)$ is endowed with the topology induced by the variational seminorm. Consequently, it is continuous if $\operatorname{Ker}(X, Y)$ is endowed with the weak topology, which implies the required assertion. □

Lemma 8. *Assume that $\varphi + \beta\psi$ is a unique Gibbs measure potential for all $\beta \in [0, 1]$. Then $\nu_{\varphi+\beta\psi}$ depends weakly continuously on β.*

Proof. Consider the space $[0,1] \times \mathcal{M}(X^G)$. Consider its subsets

$$\mathfrak{M}_\Lambda = \{(\beta, m) | \pi_{\Lambda,\varphi+\beta\psi}(m) = m\}$$

for every $\Lambda \Subset G$. By the previous lemma, all of them are closed. So, their intersection is closed, too. But it is exactly the graph of the function $\beta \mapsto \nu_{\varphi+\beta\psi}$. Since both $[0,1]$ and $\mathcal{M}(X^G)$ are Hausdorff compact, it follows that this function is continuous. $\square$

Lemma 9. *Under the assumptions of the previous lemma, $\nu_{i,\varphi+\beta\psi}$ tends to $\nu_{\varphi+\beta\psi}$ uniformly in β.*

Proof. Assuming the contrary, we have that there is $\varepsilon > 0$, a sequence (n_i) of positive integers, and a sequence β_i of numbers from $[0,1]$ such that $l(\nu_{n_i,\varphi+\beta_i\psi}, \nu_{\varphi+\beta_i\psi}) \geqslant \varepsilon$. Passing to a subsequence, we may assume that β_i tends to some β' and $\nu_{n_i,\varphi+\beta_i\psi}$ tends to some measure ν'. It is easy to see that ν' is a Gibbs measure for the potential $\varphi + \beta'\psi$ and that $l(\nu', \nu_{\varphi+\beta'\psi}) \geqslant \varepsilon$, which contradicts the uniqueness of a Gibbs measure for the potential $\varphi + \beta'\psi$. $\square$

§6. Pressure

Denote $P_{i,\varphi} = s_i^{-1} \log Z_{i,\varphi}$.

Definition 1. *The pressure of a UGM potential φ is defined as*

$$P_\varphi = \lim_{i\to\infty} P_{i,\varphi}$$

if this limit exists.

An important example: if $\varphi = 0$, then $P_\varphi = 0$.

Theorem 3. *Assume that $\varphi + \beta\psi$ is a UGM potential for every $\beta \in [0,1]$ and that P_φ exists. Then $\varphi + \psi$ has a pressure, and it can be expressed by the formula*

$$P_{\varphi+\psi} = P_\varphi - \int_0^1 d\beta \int_{X^G} \psi(\omega) d\nu_{\varphi+\beta\psi}.$$

For the proof we will need the following lemma.

Lemma 10.
$$\frac{dP_{i,\varphi+\beta\psi}}{d\beta} = -\int_{X^G} \psi(y)\, d\nu_{i,\varphi+\beta\psi}.$$

Proof. By definition,

$$\begin{aligned}\frac{dP_{i,\varphi+\beta\psi}}{d\beta} &= \frac{d}{d\beta}\left(s_i^{-1}\log Z_{i,\varphi+\beta\psi}\right) = \frac{1}{s_i Z_{i,\varphi+\beta\psi}}\frac{dZ_{i,\varphi+\beta\psi}}{d\beta} \\ &= \frac{1}{s_i Z_{i,\varphi+\beta\psi}}\int\limits_{X^{[s_i]}}\frac{d}{d\beta}\left(e^{-s_i(\Theta_i^*(\varphi+\beta\psi))(\tau)}\right)d\mu^{\otimes[s_i]} \\ &= \frac{1}{Z_{i,\varphi+\beta\psi}}\int\limits_{X^{[s_i]}}(\Theta_i^*(\psi))(\tau)e^{-s_i(\Theta_i^*(\varphi+\beta\psi))(\tau)}d\mu^{\otimes[s_i]} \\ &= \int\limits_{X^{[s_i]}}(\Theta(\psi))(\tau)d\eta_{i,\varphi+\beta\psi} = \int\limits_{X^G}\varphi(\omega)d\nu_{i,\varphi+\beta\psi}(\omega). \qquad \square\end{aligned}$$

Proof of Theorem 3. By the previous lemma, we obviously have

$$P_{i,\varphi+\psi} = P_\varphi - \int\limits_0^1 d\beta\int\limits_{X^G}\psi(\omega)d\nu_{i,\varphi+\beta\psi}.$$

Now, since $\nu_{i,\beta\varphi}$ tends to $\nu_{\beta\varphi}$ uniformly in β, we have

$$P_{\varphi+\psi} = \lim_{i\to\infty}P_{i,\varphi+\psi} = P_\varphi - \int\limits_0^1 d\beta\int\limits_{X^G}\psi(\omega)d\nu_{\varphi+\beta\psi}. \qquad \square$$

§7. The entropy formula

In this section, X will be a finite space with the discrete metric and the uniform probability measure. We will prove a nice formula for the sofic entropy of the shift action for the Gibbs measure of a UGM potential, which will lead to the proof of Theorem 2.

Lemma 11. *Let φ be a UGM potential. The measure $\eta_{i,\varphi}$ is such that it maximizes the quantity*

$$H(m)/s_i - \int\limits_{X^G}\varphi(y)d(\Theta_i(m))(y)$$

among all the probability measures on $X^{[s_i]}$.

Proof. A simple application of the Lagrange multipliers method. $\square$

Lemma 12. *Let φ be a UGM potential. Then*

$$h(\nu_\varphi) = \lim_{i\to\infty} s_i^{-1} H(\eta_{i,\varphi})$$

if the latter limit exists.

Proof. Let us first prove that $h(\nu_\varphi) \geqslant \lim_{i\to\infty} s_i^{-1} H(\eta_{i,\varphi})$. It suffices to prove that $h_\delta(\nu_\varphi) \geqslant \lim_{i\to\infty} s_i^{-1} H(\eta_{i,\varphi})$ for every δ. But this is obvious, since for sufficiently large i we have $l(\nu_\varphi, \Theta_i(\eta_{i,\varphi})) < \delta$.

Let us now prove that for every $\varepsilon > 0$ there exist $\delta > 0$ and a positive integer i_0 such that for every $i > i_0$ we have

$$H(m)/s_i \leqslant \lim_{j\to u} s_j^{-1} H(\eta_{j,\varphi}) + \varepsilon$$

for any $m \in \mathrm{Appr}_{i,\delta}(\nu_\varphi)$. In order to do this, we pick $\delta > 0$ such that for any $\kappa', \kappa'' \in \mathcal{M}(X^{[s_i]})$ with $l(\kappa', \kappa'') \leqslant 2\delta$ the inequality

$$\left| \int\limits_{X^G} \varphi(y) d\kappa' - \int\limits_{X^G} \varphi(y) d\kappa'' \right| \leqslant \varepsilon/2$$

holds. Then we pick i_0 such that for every $i > i_0$ we have

$$H(\eta_{i,\varphi})/s_i \leqslant \lim_{j\to\infty} H(\eta_{j,\varphi})/s_j + \varepsilon/2$$

and $l(\nu_{i,\varphi}, \nu_\varphi) < \delta$. Let us now fix any $i > i_0$. By the previous lemma,

$$\begin{aligned} H(m)/s_i - \int\limits_{X^G} \varphi(y) d(\Theta_i(m))(y) \\ \leqslant H(\eta_{i,\varphi})/s_i - \int\limits_{X^G} \varphi(y) d(\Theta_i(\eta_{i,\varphi}))(y) \\ \leqslant \lim_{j\to\infty} H(\eta_{j,\varphi})/s_j + \varepsilon/2 - \int\limits_{X^G} \varphi(y) d\nu_{i,\varphi}(y), \end{aligned}$$

and since $l(\nu_{i,\varphi}, m) \leqslant 2\delta$, it follows that

$$
\begin{aligned}
&H(m)/s_i \\
&\quad\leqslant \int\limits_{X^G} \varphi(y) d(\Theta(m))(y) - \int\limits_{X^G} \varphi(y) d\nu_{i,\varphi}(y) + \varepsilon/2 + \lim_{j\to\infty} H(\eta_{j,\varphi})/s_j \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \leqslant \lim_{j\to\infty} H(\eta_{j,\varphi})/s_j + \varepsilon,
\end{aligned}
$$

which implies that

$$h_\delta(\nu_\varphi) \leqslant \lim_{j\to\infty} H(\eta_{j,\varphi})/s_j + \varepsilon,$$

so

$$h(\nu_\varphi) \leqslant \lim_{j\to\infty} H(\eta_{j,\varphi})/s_j + \varepsilon.$$

Since ε can be taken arbitrarily small, we have

$$h(\nu_\varphi) \leqslant \lim_{j\to\infty} H(\eta_{j,\varphi})/s_j. \qquad \square$$

Theorem 4. *If φ is a UGM potential and has a pressure, then the modified sofic entropy can be expressed by the formula*

$$h(\nu_\varphi) = P_\varphi + \int\limits_{X^G} \varphi(y) d\nu_\varphi(y) + \log|X|.$$

Proof. For any positive integer i and for any element $\tau \in X^{[s_i]}$, by the construction of $\eta_{i,\varphi}$ we have

$$\eta_{i,\varphi}(\{\tau\}) = \frac{e^{-\sum\limits_{k\in[s_i]} \varphi(\theta_{i,k}(\tau))}}{|X|^{s_i} Z_{i.\varphi}}.$$

Now we will compute the Shannon entropy of $\eta_{i,\varphi}$:

$$
\begin{aligned}
H(\eta_{i,\varphi}) &= -\sum_{\tau\in X^{[s_i]}} \eta_{i,\varphi}(\{\tau\})\log(\eta_{i,\varphi}(\{\tau\})) \\
&= \sum_{\tau\in X^{[s_i]}} \frac{e^{-\sum\limits_{k\in[s_i]}\varphi(\theta_{i,k}(\tau))}}{|X|^{s_i} Z_{i.\varphi}} \left(\sum_{k\in[s_i]}\varphi(\theta_{i,k}(\tau)) + s_i\log|X| + \log Z_{i,\varphi}\right) \\
&= s_i\log|X| + \log Z_{i,\varphi} + \sum_{\tau\in X^{[s_i]}} \frac{e^{-\sum\limits_{k\in[s_i]}\varphi(\theta_{i,k}(\tau))}}{|X|^{s_i} Z_{i.\varphi}} \sum_{k\in[s_i]}\varphi(\theta_{i,k}(\tau)) \\
&= s_i\log|X| + \log Z_{i,\varphi} + s_i\int\limits_{X^G}\varphi(y)d\nu_{i,\varphi}.
\end{aligned}
$$

It is now easy to see that

$$
h(\nu_\varphi) = \lim_{i\to u}\frac{H(\eta)}{s_i} = P_\varphi + \int\limits_{X^G}\varphi(y)d\nu_\varphi(y) + \log|X|. \qquad \square
$$

Proof of Theorem 2. We will apply Theorem 3 to the case $\varphi = 0$. It was noted at the beginning of Sec. 6 that $P_\varphi = 0$. The formula from Theorem 3 does not involve a sofic approximation, so P_ψ does not depend on a sofic approximation. Then we apply Theorem 4, and again it contains nothing involving a sofic approximation, so $h(\nu_\psi)$ does not depend on a sofic approximation. $\square$

§8. A concluding remark

The conditions of Theorem 2, which are seemingly too restrictive, are justified by the Dobrushin uniqueness theorem. We will establish an adaptation of this theorem to our setting.

Let $\mathrm{pr} : X^G \to X$ be the map defined by the equation $\mathrm{pr}(x) = x(e)$. Let

$$
b_{g,\varphi} = \sup\Big\{\|\mathrm{pr}_\sharp(\pi_{\varphi,\{e\}}(\delta_{x_1})) - \mathrm{pr}_\sharp(\pi_{\varphi,\{e\}}(\delta_{x_2}))\|_V, \\
x_1, x_2 \in X^G, x_1|_{G\setminus\{g\}} = x_2|_{G\setminus\{g\}}\Big\}
$$

for $g \in G\setminus\{e\}$. Denote

$$
b_\varphi = \sum_{g\in G\setminus\{e\}} b_g.
$$

Theorem 5 (Dobrushin [6]). *Let $b_\varphi < 1$. Then φ has a unique Gibbs measure.*

References

1. L. Bowen, *Measure conjugacy invariants for actions of countable sofic groups.* — J. Amer. Math. Soc. **23** (2010), 217–245.
2. L. Bowen, *Entropy for expansive algebraic actions of residually finite groups.* — Ergodic Theory Dynam. Systems **31**, No. 3 (2011), 703–718.
3. L. Bowen, H. Li, *Harmonic models and spanning forests of residually finite groups.* — J. Funct. Anal. **263**, No. 7 (2012), 1769–1808.
4. A. Carderi, *Ultraproducts, weak equivalence and sofic entropy*, `arXiv:1509.03189` (2015).
5. N.-P. Chung, *Topological pressure and the variational principle for actions of sofic groups.* — Ergodic Theory Dynam. Systems **33**, No. 5 (2013), 1363–1390.
6. R. L. Dobrushin, *Description of a random field by its conditional probabilities and its regularity conditions.* — Teor. Veroyatnost. Primenen. **13** (1968), 201–229.
7. F. Rassoul-Agha, T. Seppäläinen, *A Course on Large Deviations with an Introduction to Gibbs Measures.* Amer. Math. Soc., Providence, RI, 2015.
8. H.-O. Georgii, *Gibbs Measures and Phase Transitions.* Walter de Gruyter, Berlin, 2011.
9. B. Hayes, *Fuglede–Kadison determinants and sofic entropy*, `arXiv:1402.1135` (2014).
10. L. V. Kantorovich, *On the translocation of masses.* — Dokl. Akad. Nauk SSSR **37**, Nos. 7–8 (1942), 227–229.
11. D. Kerr, *Sofic measure entropy via finite partitions.* — Groups Geom. Dyn. **7** (2013), 617–632.
12. A. M. Vershik, *The Kantorovich metric: initial history and little-known applications.* — Zap. Nauchn. Semin. POMI **312** (2004), 69–85.

Chebyshev Laboratory,
St. Petersburg State University,
14th Line 29b,
199178 St. Petersburg, Russia
E-mail: `alpeevandrey@gmail.com`

Поступило 15 сентября 2015 г.