THERMODYNAMIC FORMALISM FOR THE POSITIVE GEODESIC FLOW ON THE MODULAR SURFACE

GODOFREDO IOMMI

ABSTRACT. In this note we study the thermodynamic formalism for the positive geodesic flow on the modular surface. We define the pressure and prove the variational principle. We also establish conditions for the the pressure to be real analytic and for the potentials to have unique equilibrium states.

1. INTRODUCTION

This note is devoted to study ergodic properties of the positive geodesic flow on the modular surface. This flow was introduced by Svetlana Katok and Boris Gurevich in [GK] building up on previous work by Katok [K2]. It is an interesting flow because it has strong symbolic properties, as we will see at the end of this section. Our main purpose is to develop a thermodynamic formalism for this flow. We define the pressure, prove a variational principle and establish sufficient conditions for a potential to have equilibrium measures (see Section 4). This provides not only a strong set of tools to further study this flow, but it is has interest on its own right. Indeed, since the flow is defined on a non-compact set usual techniques do not hold. We overcome this major obstruction combining a symbolic representation of this flow obtained by Gurevich and Katok [GK] and a thermodynamic formalism for suspension flow over countable Markov shifts developed by Barreira and Iommi in [BI].

Let us begin with some basic definitions (for more details we refer to the book [K1] and the article [KU]). Denote by $\mathcal{H} = \{z \in \mathbb{C} : \text{Im } z > 0\}$ the upper half-plane endowed with the hyperbolic metric. Geodesics in $\mathcal{H}$ are either semi-circles which meet the boundary perpendicularly or vertical straight lines. The geodesic flow of $\mathcal{H}$, denote by $\{\overline{\psi}_t\}$, is the flow on the unit tangent bundle, $T^1\mathcal{H}$, of $\mathcal{H}$ which moves $\omega \in T^1\mathcal{H}$ along the geodesic it determines at unit speed.

The group of Möbius transformations acting on $\mathcal{H}$ by orientation preserving isometries can be identified with the group $\text{PSL}(2, \mathbb{Z})$. The *modular surface* is defined by $M = \text{PSL}(2, \mathbb{Z})\backslash\mathcal{H}$, which is a (non-compact) surface of constant negative curvature. Topologically, is a sphere with one cusp and two singularities. Note that $T^1\mathcal{H}$ can be identified with $\text{PSL}(2, \mathbb{R})$ by sending $\omega = (z, \zeta) \in T^1\mathcal{H}$ onto the unique $g \in \text{PSL}(2, \mathbb{R})$ such that $z = g(i)$ and $\zeta = g'(z)(\iota)$, where ι is the unit vector at the point i to the imaginary axis pointing up. In this coordinate system the geodesic flow takes the form

$$\overline{\psi}_t \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}.$$

The geodesic flow $\{\overline{\psi}_t\}$ on $\mathcal{H}$ descends to the geodesic flow $\{\psi_t\}$ on M via the projection $\pi : T^1\mathcal{H} \mapsto T^1M$. We will be interested in an invariant sub-system of $\{\psi_t\}$. In order to define it we need to consider the fundamental region for $\mathrm{PSL}(2,\mathbb{Z})$ given by

$$F = \{z \in \mathcal{H} : |z| \geq 1, |\mathrm{Re}\, z| \leq 1/2\},$$

whose sides are identified by the generators of $\mathrm{PSL}(2,\mathbb{Z})$, $T(z) = z+1$ and $S(z) = -1/z$ (see [K1, p.55]). Any geodesic can be represented by a series of segments in F. We will only be interested in oriented geodesic that do not go to the cusps of M in either direction.

Definition 1.1. *A geodesic γ is called* positive *if all segments comprising the geodesic γ in F are positively (clockwise) oriented. The set of vectors in T^1M tangent to positive geodesics is a non-compact invariant set of the geodesic flow on T^1M. We call the restriction of the geodesic flow to this set the* positive geodesic flow.

Positive geodesic have interesting coding properties. There are several ways to represent geodesics by symbolic sequences, for a detailed exposition on the subject see [KU]. Here we will be interested in two if them:

1.1. **The geometric code.** As we just saw, the sides of F are identified with the generators of $\mathrm{PSL}(2,\mathbb{Z})$, $T(z) = z+1$ and $S(z) = -1/z$. The *geometric code* (also known as Morse code) of a geodesic in F (that does not go to the cusp in either direction) is a bi-infinite sequence of integers. The idea is to code the geodesic by recording the sides of the region F that are cutted by the geodesic. The boundary of F has three sides, the left and right vertical sides are labeled by T and T^{-1}, respectively. The circular boundary is labeled by S. Any oriented geodesic, that does not go to the cusp, returns to the circular boudary of F infinitely often. The geometric code is obtained as follows. Choose an initial point on the circular boundary of F and count the number of times it hits the vertical boundary of F moving in the direction of the geodesic. We assign a positive integer to each block of hits of the right vertical side and a negative number to the left vertical side. if we move the initial point in the opposite direction we obtain a sequence of nonzero integers

$$[\gamma] = [\ldots, n_{-1}, n_0, n_1, \ldots]$$

that we denote *geometric code*.

1.2. **The arithmetic code.** An oriented geodesic $\gamma \in \mathcal{H}$ is called *reduced* if its endpoints u, v satisfiy $0 < u < 1$ and $w > 1$. Recall that a geodesic which does not go to the cusp in either direction is such that its end points are irrationals. Consider the minus continued fraction associated to u and v,

$$w = n_1 + \cfrac{1}{n_2 - \cfrac{1}{n_3 - \cfrac{1}{n_4 - \ldots}}}, \quad \frac{1}{u} = n_0 - \cfrac{1}{n_{-1} - \cfrac{1}{n_{-2} - \cfrac{1}{n_{-3} - \ldots}}}$$

The following code is given to γ,

$$(\gamma) = (\ldots, n_{-2}, n_{-1}, n_0, n_1, n_2, \ldots).$$

Now, it is possible to show that an arbitrary geodesic $\gamma \in M$ can also be represented by doubly infinite sequence. The idea is to construct a cross section for the geodesic flow on $T^1 M$. Every oriented geodesic can be represented as a bi-infinite sequence of segments σ_i between returns to the cross section. To each segment it corresponds a reduced geodesic γ_i. It turns out that (see [GK]) all these geodesics have the same arithmetic code, except for a shift. This sequence is the *arithmetic code* of γ,

$$(\gamma) = (\ldots, , n_{-2}, n_{-1}, n_0, n_1, n_2, \ldots).$$

1.3. **Equality of codes.** Positive geodesic behave very well with respect to the arithmetic and geometric codes. Indeed, Gurevich and Katok [GK] proved the following result which is a generalisation of earlier results by Katok [K2],

Theorem 1.1 (Gurevich-Katok). *A geodesic γ is positive if and only if the arithmetic and the geometric codes for γ coincide.*

In this note we study the thermodynamic formalism for the positive geodesic flow. These are ideas and techniques that come originally from statistical mechanics and were brought into dynamical systems by Ruelle and Sinai in the early seventies. This formalism provides procedures for the choice of invariant measures. Let us stress that the positive geodesic flow has many invariant measures, hence the problem of choosing relevant ones. A good understanding of the thermodynamic formalism could allow us, for instance, to develop a dimension theory (e.g multifractal formalism, Bowen formula) for this flow, see [Pe] for details.

Thermodynamic formalism was studied for the geodesic flow on compact negatively curved manifolds (more generally for Axiom A flows) by Bowen and Ruelle [BR]. They proved that, in that context, the geodesic flow can be coded by a suspension flow over a sub-shift of finite type defined over a finite alphabet (this is done using the Markov partitions obtained by Bowen [B1] and Ratner [R]). This relation allowed them to reduce the study of the geodesic flow to a suspension flow. They proved existence and uniqueness of equilibrium measures (see Section 4 for precise definitions) for a wide range of potentials.

It was shown by Gurevich and Katok (see Section 3 and [GK]) that the positive geodesic flow in the modular surface can be coded by a suspension flow over a countable Markov shift. The fact that the model is not compact is a major obstruction if ones wants to apply the results by Bowen and Ruelle. Nevertheless, thermodynamic formalism for suspension flows over countable Markov shifts was developed and studied by Barreira and Iommi in [BI]. In this note, making use of the results in [GK] and [BI], we provide a definition of pressure for the positive geodesic flow. Moreover, we obtain a variational principle and establish conditions under which a large class of potentials have equilibrium measures.

2. Preliminaries from Ergodic theory

In this section we review results form ergodic theory that will be used in the rest of the paper. First we recall the definition of pressure for countable Markov shifts and some of its properties. We then recall the suspension flow construction and study the relation between the invariant measures for the shift and the invariant measures for the suspension flow. This relation will be essential in what follows because it allow us to relate the study of the thermodynamic formalism for the flow with the study of the thermodynamic formalism for the shift.

2.1. **Thermodynamic formalism for countable Markov shifts.** Let B be a transition matrix defined on the alphabet of natural numbers. That is, the entries of the matrix $B = B(i,j)_{\mathbb{N}\cup\{0\}\times\mathbb{N}\cup\{0\}}$ are zeros and ones (with no row and no column made entirely of zeros). The countable Markov shift (Σ_B, σ) is the set

$$\Sigma_B := \left\{ (x_n)_{n\in\mathbb{N}\cup\{0\}} : B(x_n, x_{n+1}) = 1 \text{ for every } n \in \mathbb{N}\cup\{0\} \right\},$$

together with the shift map $\sigma : \Sigma_B \to \Sigma_B$ defined by $\sigma(x_0, x_1, x_2, x_3 \dots) = (x_1, x_2, x_3, \dots)$.

Remark 2.1. *Analogously, we can define a two-sided countable Markov shift by*

$$\Sigma_B^* := \{(x_n)_{n\in\mathbb{Z}} : B(x_n, x_{n+1}) = 1 \text{ for every } n \in \mathbb{Z}\},$$

together with the shift map $\sigma : \Sigma_B^* \to \Sigma_B^*$ defined by $(\sigma x)_n = x_{n+1}$.

Recall that the space Σ_B is equipped with the topology generated by the cylinder sets

$$C_{a_0\cdots a_n} = \{x \in \Sigma_B : x_i = a_i \text{ for } i = 0,\dots,n\}.$$

Given a function $\rho\colon \Sigma_B \to \mathbb{R}$ we define

$$V_n(\rho) := \sup\{|\rho(x) - \rho(y)| : x, y \in \Sigma_B, \ x_i = y_i \text{ for } i = 0,\dots,n-1\},$$

where $x = (x_0 x_1 \cdots)$ and $y = (y_0 y_1 \cdots)$. We say that ρ is *locally Hölder* if there exist constants $K > 0$ and $\theta \in (0,1)$ such that $V_n(\rho) \le K\theta^n$ for all $n \in \mathbb{N}$. Note that since nothing is required for $n = 0$ a locally Hölder function is not necessarily bounded.

There are essentially two notions of (topological) pressure for a countable Markov shift. The first was proposed by Mauldin and Urbański [MU] and the second by Sarig [S1]. For the class of Markov shifts that we will be interested in this note both notions coincide.

Definition 2.1. *Let $\rho\colon \Sigma_B \to \mathbb{R}$ be a locally Hölder function. The* pressure *of ρ is defined by*

$$P_\sigma(\rho) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{x:\sigma^n x = x} \exp\left(\sum_{i=0}^{n-1} \rho(\sigma^i x)\right).$$

Remark 2.2. *Let (Σ_F, σ) be the full-shift on countably many symbols, that is*

$$\Sigma_F := \left\{ (x_n)_{n\in\mathbb{N}\cup\{0\}} : x_n \in \mathbb{N}\cup\{0\} \right\}.$$

If $\rho : \Sigma_F \to \mathbb{R}$ is a locally constant potential, that is $\rho|C_n := \log\lambda_n$, then there is a simple formula for the pressure (see, for example, [BI, Example 1])

$$(1) \qquad\qquad P_\sigma(\rho) = \log \sum_{n=0}^{\infty} \lambda_n.$$

This notion of pressure satisfies the variational principle (see [MU, S1]),

Theorem 2.1. *Let (Σ_B, σ) be a countable Markov shift and $\rho\colon \Sigma_B \to \mathbb{R}$ be a locally Hölder function, then*

$$P_\sigma(\rho) = \sup\left\{ h(\nu) + \int \rho \, d\nu : \nu \in \mathcal{M}_\sigma \ and \ -\int \rho \, d\nu < \infty \right\},$$

where $\mathcal{M}_\sigma$ denotes the set of $\sigma-$invariant probability measures and $h(\nu)$ denotes the entropy of the measure ν (for a precise definition see [W, Chapter 4]).

A measure $\nu \in \mathcal{M}_\sigma$ attaining the supremum, that is, $P_\sigma(\rho) = h(\nu) + \int \rho \, d\nu$ is called *equilibrium measure* for ρ.

We say that $\mu \in \mathcal{M}_\sigma$ is a *Gibbs measure* for the function $\rho \colon \Sigma_B \to \mathbb{R}$ if for some constants P, $C > 0$ and every $n \in \mathbb{N}$ and $x \in C_{a_0 \cdots a_n}$ we have

$$\frac{1}{C} \leq \frac{\mu(C_{a_0 \cdots a_n})}{\exp(-nP + \sum_{i=0}^n \phi(\sigma^k x))} \leq C.$$

There is a class of countable Markov shifts, introduced by Sarig [S2], that have simple combinatorics, similar to that of the full-shift. The thermodynamic formalism for Markov shifts belonging to this class is similar to the one of sub-shift of finite type defined on finite alphabets. We say that a countable Markov shift (Σ_B, σ), defined by the transition matrix $B(i,j)$ with $(i,j) \in \mathbb{N} \cup \{0\} \times \mathbb{N} \cup \{0\}$, satisfies the *BIP property* if and only if there exists $\{b_1, \ldots, b_n\} \in \mathbb{N} \cup \{0\}$ such that for every $a \in \mathbb{N} \cup \{0\}$ there exists $i, j \in \mathbb{N}$ with $B(b_i, a)B(a, b_j) = 1$. The following theorem sumarises results proven by Sarig in [S2, S3] and by Mauldin and Urbański [MU],

Theorem 2.2. *Let (Σ_B, σ) be a countable Markov shift satisfying the BIP property and $\rho : \Sigma_B \to \mathbb{R}$ a locally Hölder potential. Then, there exists $t^* > 0$ such that pressure function $t \to P(t\rho)$ has the following properties*

$$P_\sigma(t\rho) = \begin{cases} \infty & \text{if } t \leq t^* \\ \text{real analytic} & \text{if } t > t^*. \end{cases}$$

Moreover, if $t > t^*$, there exists a unique equilibrium measure for $t\rho$. If $\sum_{n=1}^\infty V_n(\rho) < \infty$ and $P_\sigma(\rho) < \infty$ then there exists a Gibbs measure for ρ.

2.2. **Suspension flows and invariant measures.** Let (Σ_B, σ) be a countable Markov shift and $\tau \colon \Sigma_B \to \mathbb{R}^+$ be a positive continuous function. Consider the space

$$(2) \qquad Y = \{(x,t) \in \Sigma_B \times \mathbb{R} \colon 0 \leq t \leq \tau(x)\},$$

with the points $(x, \tau(x))$ and $(\sigma(x), 0)$ identified for each $x \in \Sigma_B$. The *suspension semiflow* over σ with *height function* τ is the semiflow $\Phi = (\varphi_t)_{t \geq 0}$ on Y defined by

$$\varphi_t(x,s) = (x, s+t) \text{ whenever } s+t \in [0, \tau(x)].$$

In the case of two-sided Markov shifts we can define a suspension flow $(\varphi_t)_{t \in \mathbb{R}}$ in a similar manner.

We denote by $\mathcal{M}_\Phi$ the space of Φ-invariant probability measures on Y. Recall that a measure μ on Y is Φ-*invariant* if $\mu(\varphi_t^{-1} A) = \mu(A)$ for every $t \geq 0$ and every measurable set $A \subset Y$. We also consider the space $\mathcal{M}_\sigma$ of σ-invariant probability measures on Σ_B. There is a strong relation between this two spaces of invariant measures. Consider the space of $\sigma-$invariant measures for which τ is integrable,

$$(3) \qquad \mathcal{M}_\sigma(\tau) := \left\{ \mu \in \mathcal{M}_\sigma : \int \tau \, d\mu < \infty \right\}.$$

Denote by m the one dimensional Lebesgue measure and let $\mu \in \mathcal{M}_\sigma(\tau)$ then it follows directly from classical results by Ambrose and Kakutani [AK] that

$$(\mu \times m)|_Y / (\mu \times m)(Y) \in \mathcal{M}_\Phi.$$

The behaviour of the map $R \colon \mathcal{M}_\sigma \to \mathcal{M}_\Phi$, defined by

$$(4) \qquad R(\mu) = (\mu \times m)|_Y / (\mu \times m)(Y)$$

is closely related to the ergodic properties of the flow. Indeed, in the compact setting the map $R \colon \mathcal{M}_\sigma \to \mathcal{M}_\Phi$ is a bijection. This fact was used by Bowen and Ruelle [BR] to study and develop the thermodynamic formalism for Axiom A flows (these flows admit a compact symbolic representation). In the general (non-compact) setting there are several difficulties that can arise. For instance, the height function τ need not to be bounded above. It is, therefore, possible for a measure $\nu \in \mathcal{M}_\sigma$ to be such that $\int \tau \, d\nu = \infty$. In this situation the measure $\nu \times m$ is an infinite sigma-invariant measure for Φ. Hence, the map $R(\cdot)$ is not well defined and this makes it harder to reduce the study of the thermodynamic formalism of the flow Φ to the one of the shift σ. Another possible complication occurs if the height function τ is not bounded away from zero. Then it is possible that for an infinite (sigma-finite) invariant measure ν we have $\int \tau \, d\nu < \infty$. In this case the measure $(\nu \times m)|_Y / (\nu \times m)(Y) \in \mathcal{M}_\Phi$. In such a situation, the map R is not surjective. Again, the fact that R is not a bijective map makes it hard to translate problems from the flow to the shift.

The following can be obtained directly from the results by Ambrose and Kakutani [AK],

Lemma 2.1. *If $\tau \colon \Sigma_B \to \mathbb{R}$ is bounded away from zero, the map $R \colon \mathcal{M}_\sigma(\tau) \to \mathcal{M}_\Phi$ defined by*

$$R(\mu) = (\mu \times m)|_Y / (\mu \times m)(Y)$$

is bijective.

Given a continuous function $F \colon Y \to \mathbb{R}$ we define the function $\Delta_F \colon \Sigma \to \mathbb{R}$ by

$$\Delta_F(x) = \int_0^{\tau(x)} F(x, t) \, dt.$$

The function Δ_F is also continuous, moreover

$$(5) \qquad \int_Y F \, dR(\nu) = \frac{\int_\Sigma \Delta_F \, d\nu}{\int_\Sigma \tau \, d\nu}.$$

Remark 2.3 (Extension of potentials defined on the base)**.** *Let $\rho \colon \Sigma_B \to \mathbb{R}$ be a locally Hölder potential. It is shown in [BRW] that there exists a continuous function $F \colon Y \to \mathbb{R}$ such that $\Delta_F = \rho$. This provides a tool to construct examples.*

2.3. **Abramov's formula.** In this short subsection we recall a classical result by Abramov. The entropy of a flow with respect to an invariant measure can be defined by the entropy of the corresponding time one map. For the definition of entropy in the context of maps see [W, Chapter 4]. In 1959 Abramov [A] showed that in the context of suspension semiflows (among others) the following relation holds,

Proposition 2.1 (Abramov)**.** *Let $\mu \in \mathcal{M}_\Phi$ be such that $\mu = (\nu \times m)|_Y / (\nu \times m)(Y)$, where $\nu \in \mathcal{M}_\sigma$ then*

$$(6) \qquad h_\Phi(\mu) = \frac{h_\sigma(\nu)}{\int \tau \, d\nu}.$$

It follows from Abramov's result that

Lemma 2.2. *Let $\mu \in \mathcal{M}_\Phi$ be such that $\mu = (\nu \times m)|_Y / (\nu \times m)(Y)$, we have that $h_\Phi(\mu) = \infty$ if and only if $h_\sigma(\nu) = \infty$.*

When the phase space is non-compact there are several different notions of topological entropy of a flow, we will consider the following,

Definition 2.2. *The topological entropy of the suspension flow* (Y, Φ) *denoted by* $h(\Phi)$ *is defined by*

$$h(\Phi) = \sup \left\{ h_\Phi(\mu) : \mu \in \mathcal{M}_\Phi \right\}.$$

3. Suspension flow representation for the positive geodesic flow

It was shown by Gurevich and Katok in [GK] that the positive geodesic flow on the modular surface has a symbolic representation as a suspension flow. In this section we recall this construction and describe some of its properties.

Consider the alphabet $\mathcal{A} = \{3, 4, 5, 6, \dots\}$. Let A be the transition matrix defined by $A(i, j) = 1$ for every pair $(i, j) \in \mathcal{A} \times \mathcal{A}$ except for the pairs $(i, j) \in \{(3, 3), (3, 4), (3, 5), (4, 3), (5, 3)\}$, where $A(i, j) = 0$. Note that the pairs for which the matrix has an entry equal to zero correspond to the five Platonic bodies. Denote by (Σ_A, σ) the corresponding (non-compact) countable Markov shift. The proof of the following result is straight forward.

Lemma 3.1. *The countable Markov shift* (Σ_A, σ) *satisfies the BIP property.*

Remark 3.1. *Note that the full-shift on the alphabet* $\mathcal{A}' = \{6, 7, 8, \dots\}$, *that we will denote by* (Σ_F, σ), *is a sub-shift of* (Σ_A, σ).

Constructing an appropriate cross section and computing the first return time function of the geodesic flow to it, Gurevich an Katok [GK] showed that the positive geodesic flow on the modular surface can be represented by the suspension flow (Y^*, Φ^*). This flow is defined over the countable (two-sided) Markov shift (Σ_A^*, σ), where

$$\Sigma_A^* := \left\{ (x_n)_{n \in \mathbb{Z}} : A(x_n, x_{n+1}) = 1 \text{ for every } n \in \mathbb{Z} \right\},$$

and it has height function $\tau^*(x) = 2 \log w(x)$, where

$$w(x) = n_1 + \cfrac{1}{n_2 - \cfrac{1}{n_3 - \cfrac{1}{n_4 - \dots}}},$$

for $x = (\dots n_{-1}, n_0, n_1, n_2, n_3 \dots)$. Note that the height function only depends on the *future* coordinates of $x \in \Sigma_A^*$.

The result in [GK] is that there exits a continuous bijection, $\pi : Y^* \to T^1 M^+$, where $T^1 M^+$ is the space of unit vectors tangent to positive geodesics. With the property that for every $t \in \mathbb{R}$ we have $\pi \circ \phi_t(x) = \psi_t \circ \pi(x)$. Sumarising,

Theorem 3.1 (Gurevich-Katok). *The suspension flow* (Y^*, Φ^*) *is a symbolic representation of the positive geodesic flow on the modular surface.*

Since every potential $\overline{\rho} : \Sigma_A^* \to \mathbb{R}$ is cohomologous to a potential $\rho : \Sigma_A^* \to \mathbb{R}$ that only depends on *future* coordinates, we can reduce the study of the suspension flow to that of the suspension semi-flow. This is a standard procedure, see for example [PP, p.93]. Therefore, in order to study the ergodic theory of the positive geodesic flow it is enough to understand the ergodic theory of the suspension semi-flow (Y, Φ) defined over the countable Markov shift (Σ_A, σ) with height function $\tau : \Sigma_A \to \mathbb{R}$ defined by $\tau(x) = 2 \log w(x)$.

Remark 3.2. *Let* $c = (3 + \sqrt{5})/6$. *If* $x = (\dots n_{-1}, n_0, n_1, n_2, n_3 \dots)$ *then*

$$2 \log(cn_1) \leq \tau(x) \leq 2 \log n_1.$$

4. Thermodynamic formalism for the positive geodesic flow

The main purpose of this note is to define the pressure function for the positive geodesic flow on the modular surface and to study its properties. In order to do so, we will use the symbolic representation explained in the previous section together with results of Barreira and Iommi [BI] on thermodynamic formalism for suspension flows over countable Markov shifts.

As we have seen in Section 3 it is enough to study the suspension semi-flow (Y, Φ). We start by defining the class of potentials that we will consider. Denote by

$$(7) \qquad \mathcal{P} := \{F : Y \to \mathbb{R} : \text{the potential } \Delta_F \text{ is locally Hölder}\}.$$

Remark 4.1. *Note that if $F \in \mathcal{P}$ then there exists sequences $(s_n)_n$ and $(S_n)_n$ such that $s_n \leq \Delta_F | C_n \leq S_n$.*

Following the strategy developed by Barreira and Iommi in [BI] we define the pressure.

Definition 4.1. *Let $F \in \mathcal{P}$, the pressure of F with respect to the semi-flow (Y, τ) is defined by*

$$(8) \qquad P_\Phi(F) = \inf\{t \in \mathbb{R} : P_\sigma(\Delta_F - t\tau) \leq 0\}.$$

We assume the convention that $P_\Phi(F) = \infty$ when the infimum is taken over the empty set.

Of particular interest is the case in which the potential F is the null potential. Indeed, in that case the pressure is equal to the entropy of the flow $P_\Phi(0) = h(\Phi)$. This case was studied by Polyakov [Po] and by Gurevich and Katok [GK]. They obtained estimates for the entropy of the flow. In [GK] the following bound was obtained

$$(9) \qquad 0.7771 \leq h(\Phi) \leq 0.8161.$$

Note that the entropy of the geodesic flow on the modular surface has entropy equal to one.

Let us make a few comments on Definition 4.1. First note that there are potentials for which the equation $P_\sigma(\Delta_F - t\tau) = 0$ does not have a root.

Example 4.1. *Let $\rho : \Sigma_A \to \mathbb{R}$ be the locally constant potential defined by $\rho | C_n = \log(\log n)^{-2}$. In virtue of Remark 2.3 there exists a potential $F \in \mathcal{P}$ such that $\Delta_F = \rho$. Let us extend ρ to the full-shift on $\mathbb{N}$ defining it to be equal to zero on the cylinders where it remains to be defined. By Remark 2.2 we can bound the pressure by*

$$P_\sigma(\Delta_F - t\tau) = P_\sigma(\log(\log n)^{-2} - t \log n^{-2}) < \log\left(\sum_{n=1}^{\infty} \frac{1}{n^{-2t}(\log n)^2}\right).$$

Therefore $P_\Phi(F) = 1/2$ and $P_\sigma(\Delta_F - P_\Phi(F)\tau) < 0$.

When the equation defining the pressure has root, regularity properties of the pressure can be obtained.

Theorem 4.1 (Regularity)**.** *Let $F \in \mathcal{P}$ with the property that there exists $\epsilon > 0$ such that if $q \in (1 - \epsilon, 1 + \epsilon)$ then the equation*

$$P_\sigma(\Delta_{qF} - t\tau) = 0$$

has a unique root. Then, then the function $q \mapsto P_\Phi(qF)$ is real analytic in the interval $(1 - \epsilon, 1 + \epsilon)$.

Proof. Since Σ_A satisfies the BIP property and if $q \in (1-\epsilon, 1+\epsilon)$ then Δ_{qF} is locally Hölder, the function $t \mapsto P_\sigma(\Delta_{qF} - t\tau)$, when finite, is real analytic (see subsection 2.1 or [S3, Corollary 4]). The result now follows from the implicit function theorem: $P_\sigma(\Delta_{qF} - P_\Phi(qF)\tau) = 0$, and in order to verify the nondegeneracy condition note that (see [PU, Chapter 4])

$$\frac{\partial}{\partial t} P_\sigma(\Delta_{qF} - t\tau)\Big|_{t=s} = -\int_\Sigma \tau \, d\mu < 0,$$

where μ denotes the equilibrium (Gibbs) measure of $\Delta_{qF} - s\tau$ (the existence of such measure was proved in [S3, Theorem 1], see subsection 2.1). $\qquad\square$

The next two results are direct consequence of more general statements obtained in [BI, Theorem 1 and Theorem 2], the proofs are the same. Both are fundamental results in thermodynamic formalism. In particular they show that the notion of pressure we have introduced is a *correct* definition.

Theorem 4.2 (Approximation property). *If $F \in \mathcal{P}$ then*

$$P_\Phi(F) = \sup \left\{ P_{\Phi|K}(F) : K \subset Y \text{ compact and } \Phi\text{-invariant} \right\},$$

where $P_{\Phi|K}(\cdot)$ is the pressure on compact spaces (see [W, Chapter 9]).

Theorem 4.3 (Variational principle). *If $F \in \mathcal{P}$ then*

$$(10) \qquad P_\Phi(F) = \sup \left\{ h_\mu(\Phi) + \int_Y F \, d\mu : \mu \in \mathcal{M}_\Phi \text{ and } -\int_Y F \, d\mu < \infty \right\}.$$

A measure $\mu \in \mathcal{M}_\Phi$ such that $P_\Phi(F) = h_\mu(\Phi) + \int_Y F \, d\mu$ is called *equilibrium measure* for F.

Theorem 4.4 (Equilibrium measures). *Let $F \in \mathcal{P}$ the following properties are equivalent:*

  (1) *there is an equilibrium measure $\mu_F \in \mathcal{M}_\Phi$ for F;*
  (2) *$P_\sigma(\Delta_F - P_\Phi(F)\tau) = 0$ and there is an equilibrium measure $\nu_F \in \mathcal{M}_\sigma(\tau)$ for $\Delta_F - P_\Phi(F)\tau$.*

Proof. We will assume that $P_\Phi(F) < \infty$, otherwise there is no equilibrium measure. Recall that from Definition 2.1 of topological pressure we have that $P_\sigma(\Delta_F - P_\Phi(F)\tau) \leq 0$.

Let us first consider the case in which $P_\sigma(\Delta_F - P_\Phi(F)\tau) < 0$. Given $\nu \in \mathcal{M}_\sigma(\tau)$, by the variational principle for the pressure $P_\sigma(\cdot)$ (see subsection 2.1 or [S1, Theorem 3]) we have

$$h_\nu(\sigma) + \int_\Sigma \Delta_F \, d\nu - P_\Phi(F) \int_\Sigma \tau \, d\nu < 0.$$

Since the space $\mathcal{M}_\sigma(\tau)$ can be identified with $\mathcal{M}_\Phi$, using Abramov's formula (see subsection 2.3) we obtain that for every $\mu \in \mathcal{M}_\Phi$,

$$h_\mu(\Phi) + \int_Y F \, d\mu < P_\Phi(F).$$

Therefore, there are no equilibrium measures in this case.

Assume now that $P_\sigma(\Delta_F - P_\Phi(F)\tau) = 0$, and let $\nu_F \in \mathcal{M}_\sigma(\tau)$ be an equilibrium measure for the potential $\Delta_F - P_\Phi(F)\tau$. Then

$$P_\sigma(\Delta_F - P_\Phi(F)\tau) = h_{\nu_F}(\sigma) + \int_\Sigma (\Delta_F - P_\Phi(F)\tau)\ d\nu_F = 0.$$

Set $\mu_F = R(\nu_F)$. Since $\nu_F \in \mathcal{M}_\sigma(\tau)$ we have $\int_\Sigma \tau\ d\nu_F < \infty$, and thus

$$P_\Phi(F) = \frac{h_{\nu_F}(\sigma)}{\int_\Sigma \tau\ d\nu_F} + \frac{\int_\Sigma \Delta_F\ d\nu_F}{\int_\Sigma \tau\ d\nu_F} = h_{\mu_F}(\Phi) + \int_Y F\ d\mu_F.$$

This shows that μ_F is an equilibrium measure for F. On the other hand, if we start with an equilibrium measure μ_F for F, then

$$P_\Phi(F) = h_{\mu_F}(\Phi) + \int_Y F\ d\mu_F.$$

The measure μ_F is obtained from a product measure $\nu_F \times m$ for some $\nu_F \in \mathcal{M}_\sigma(\tau)$. Therefore, using Abramov's formula,

$$0 = P_\sigma(u_F) \geq h_{\nu_F}(\sigma) + \int_\Sigma (\Delta_F - P_\Phi(F)\tau)\ d\nu_F = 0.$$

In particular, ν_F is an equilibrium measure for u_F. This completes the proof. $\square$

Remark 4.2. *Note that every potential $\Delta_F - P_\Phi(F)\tau$ has a Gibbs measure $\mu \in \mathcal{M}_\sigma$ (see [S3, Theorem 1]). Nevertheless, it is possible that this measure is not an equilibrium measure. Indeed, if $h_\sigma(\mu) = \infty$ and $\int(\Delta_F - P_\Phi(F)\tau)d\mu = -\infty$ then μ is not an equilibrium measure.*

The following are examples in which the conditions in part (2) of Theorem 4.4 fails.

Example 4.2 (No equilibrium measure 1). *Let $F : Y \to \mathbb{R}$ be the potential defined in Example* 4.1 *then*

$$P_\sigma(\Delta_F - P_\Phi(F)\tau) < 0.$$

Therefore, the potential F does not have an equilibrium measure.

Example 4.3 (No equilibrium measure 2). *Let $F : Y \to \mathbb{R}$ be a potential such that $\Delta_F(x) = -\log(n(\log n)^2)$. Recall that the existence of such potential is guaranteed by Remark* 2.3. *By Remark* 2.2 *we have that*

$$P((\Delta_F - t\tau)|\Sigma_F) = \log \sum_{n=6}^\infty \frac{1}{n^{1+2t}(\log n)^2}.$$

Therefore, we have

$$P((\Delta_F - P_\Phi(F)\tau) = 0.$$

Moreover $P_\Phi(F) < 1$, indeed note that for $t = 0$ the pressure function on the full-shift on $\mathbb{N}$, denoted by $\Sigma_\mathbb{N}$, is equal to zero. The result follows since $\Sigma_A \subset \Sigma_\mathbb{N}$.

The Gibbs measure $\mu \in \mathcal{M}_\sigma$ corresponding to the potential $\Delta_F - P_\Phi(F)\tau$, satisfies the following relation for a constant $C > 0$ and for every natural number $n \geq 6$,

$$\frac{1}{C} \leq \frac{\mu(C_n)}{n^{2P_\Phi(F)+1}(\log n)^2} \leq C.$$

Note that

$$\int \tau\ d\mu_\phi \geq \sum_{n=6}^\infty n^2 \mu(C_n) \geq \frac{1}{C} \sum_{n=6}^\infty \frac{n^2}{n^{2P_\Phi(F)+1}(\log n)^2} = \frac{1}{C} \sum_{n=6}^\infty \frac{1}{n^{2P_\Phi(F)-1}(\log n)^2}$$

Since $2P_\Phi(F) - 1 < 1$ we have that $\int \tau \, d\mu_\phi = \infty$.

In the next Theorem we establish conditions under which a potential has an equilibrium measure. This condition is a *small oscillation* type of condition.

Theorem 4.5. *Let $F \in \mathcal{P}$ be a bounded potential with $P_\Phi(F) < \infty$. If*

$$\tag{11} \sup F - \inf F < h_{top}(\Phi) - \frac{1}{2},$$

then F has an equilibrium measure $\mu_F \in \mathcal{M}_\Phi$.

Proof. Let us first show that $P_\sigma(\Delta_F - P_\Phi(g)\tau) = 0$. Denote by $s = \inf F$ and by $S = \sup F$. We have that $s\tau \leq \Delta_F \leq S\tau$. Moreover,

$$P_\sigma((s - t)\tau) \leq P_\sigma(\Delta_F - t\tau) \leq P_\sigma((S - t)\tau).$$

Let $t^* \in (1/2 + S, h_{top}(\Phi) + s)$ (by equation (11) this interval is non degenerate). We have that

$$0 < P_\sigma((s - t^*)\tau) \leq P_\sigma(\Delta_F - t^*\tau) \leq P_\sigma((S - t^*)\tau) < \infty.$$

Since $P_\Phi(F) < \infty$ and by the continuity of the function $t \to P_\sigma(\Delta_F - t\tau)$ we have that $P_\sigma(\Delta_F - P_\Phi(g)\tau) = 0$.

It remains to show that the potential $\Delta_F - P_\Phi(F)\tau$ has an equilibrium measure. Recall that there exists a Gibbs measure μ associated to this potential. In order to show that this measure is an equilibrium measure it suffices to prove that

$$\tag{12} \int (\Delta_F - P_\Phi(F)\tau) \, d\mu < \infty.$$

But note that [PU, Chapter 4]

$$\frac{d}{dt}P_\sigma(\Delta_F - t\tau)\Big|_{t=P_\Phi(F)} = \int (\Delta_F - P_\Phi(F)\tau) \, d\mu.$$

Note that we have proved that there exists an interval of the form $[P_\phi(F) - \epsilon, P_\phi(F) + \epsilon]$ where the function $t \to P_\sigma(\Delta_F - t\tau)$ is finite. The result now follows, because when finite the function $t \to P_\sigma(\Delta_F - t\tau)$ is real analytic. $\square$

It is interesting to remark that in a wide range of different contexts similar small oscillation conditions on the potential have been imposed in order to prove existence and uniqueness of equilibrium measures. For instance, Hofbauer [H] originally in a symbolic setting and later Hofbauer and Keller [HK] in the context of the angle doubling map on the circle, gave examples which shows that this type of conditions was essential in their setting in order to have quasi-compactness of the transfer operator, and hence good equilibrium measures. Later, Denker and Urbański [DU], for rational maps, used similar conditions to prove uniqueness of equilibrium measures. Oliveira [O] proves the existence of equilibrium measures for potentials satisfying similar conditions for non-uniformly expanding maps. Recently, Bruin and Todd [BT] in the context of multimodal maps also made use of a similar condition to prove uniqueness of equilibrium measures.

References

[A] L. M. Abramov. *On the entropy of a flow.* Dokl. Akad. Nauk SSSR 128 (1959) 873–875.

[AK] W. Ambrose and S. Kakutani. *Structure and continuity of measurable flows.* Duke Math. J. 9, (1942). 25–42.

[BI] L. Barreira and G. Iommi. *Suspension flows over countable Markov shifts* J. Stat. Phys. 124 (2006), no. 1, 207–230.

[BRW] L. Barreira, L. Radu and C. Wolf. *Dimension of measures for suspension flows*, Dyn. Syst. **19** (2004), 89–107.

[B1] R. Bowen *Symbolic dynamics for hyperbolic flows*, Amer. J. Math. **95** (1973), 429–460.

[BR] R. Bowen and D. Ruelle. *The ergodic theory of Axiom A flows.* Invent. Math. 29 (1975), no. 3, 181–202

[BT] H. Bruin and M. Todd. *Equilibrium states for interval maps: potentials with* $\sup \phi - \inf \phi <$ $h_{\text{top}}(f)$. Comm. Math. Phys. 283 (2008), no. 3, 579–611.

[DU] M. Denker and M. Urbański. *Ergodic theory of equilibrium states for rational maps.* Nonlinearity 4 (1991), no. 1, 103–134.

[GK] B. Gurevich and S. Katok. *Arithmetic coding and entropy for the positive geodesic flow on the modular surface. Dedicated to the memory of I. G. Petrovskii on the occasion of his 100th anniversary.* Mosc. Math. J. 1 (2001), no. 4, 569–582, 645.

[H] F. Hofbauer. *Examples for the nonuniqueness of the equilibrium state.* Trans. Amer. Math. Soc. 228 (1977), no. 223–241.

[HK] F. Hofbauer and G. Keller. *Equilibrium states for piecewise monotonic transformations.* Ergodic Theory Dynam. Systems 2 (1982), no. 1, 23–43.

[K1] S. Katok. *Fuchsian groups.* Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, (1992). x+175 pp

[K2] S. Katok. *Coding of closed geodesics after Gauss and Morse.* Geom. Dedicata 63 (1996), no. 2, 123–145.

[KU] S. Katok and I. Ugarcovici. *Symbolic dynamics for the modular surface and beyond.* Bull. Amer. Math. Soc. (N.S.) 44 (2007), no. 1, 87–132

[MU] R.D. Mauldin and M. Urbański. *Dimensions and measures in infinite iterated function systems.* Proc. London Math. Soc. (3) 73 (1996), no. 1, 105–154.

[O] K. Oliveira. *Equilibrium states for non-uniformly expanding maps.* Ergodic Theory Dynam. Systems 23 (2003), no. 6, 1891–1905.

[PP] W. Parry and M. Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics.* Astrisque No. 187-188 (1990), 268 pp.

[Pe] Y. Pesin. *Dimension Theory in Dynamical Systems: Contemporary Views and Applications*, Chicago Lectures in Mathematics, Chicago University Press, (1997).

[Po] A.B. Polyakov *On a measure with maximal entropy for a special flow over a local perturbation of a countable topological Bernoulli scheme.* (Russian) Mat. Sb. 192 (2001), no. 7, 73–96; translation in Sb. Math. 192 (2001), no. 7-8, 1001–1024

[PU] F. Przytycki and M. Urbański. *Fractals in the Plane, Ergodic Theory Methods,* to appear in Cambridge University Press. Available at http://www.math.unt.edu/∼urbanski/book1.html.

[R] M. Ratner. *Markov partitions for Anosov flows on n-dimensional manifolds*, Israel J. Math. **15** (1973), 92–114.

[S1] O. Sarig. *Thermodynamic formalism for countable Markov shifts,* Ergodic Theory Dynam. Systems **19** (1999) 1565–1593.

[S2] O. Sarig. *Phase transitions for countable Markov shifts,* Comm. Math. Phys. **217** (2001) 555–577.

[S3] O. Sarig. *Existence of Gibbs measures for countable Markov shifts,* Proc. Amer. Math. Soc. **131** (2003) 1751–1758

[Sav] S. Savchenko. *Special flows constructed from countable topological Markov chains*, Funct. Anal. Appl. **32** (1998), 32–41.

[W] P. Walters. *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics 79, Springer, (1981).

Facultad de Matemáticas, Pontificia Universidad Católica de Chile (PUC), Avenida Vicuña Mackenna 4860, Santiago, Chile

E-mail address: giommi@mat.puc.cl

URL: http://www.mat.puc.cl/~giommi/