

# DOMINATION, ALMOST ADDITIVITY, AND THERMODYNAMIC FORMALISM FOR PLANAR MATRIX COCYCLES

BALÁZS BÁRÁNY, ANTTI KÄENMÄKI, AND IAN D. MORRIS

**ABSTRACT.** In topics such as the thermodynamic formalism of linear cocycles, the dimension theory of self-affine sets, and the theory of random matrix products, it has often been found useful to assume positivity of the matrix entries in order to simplify or make feasible certain types of calculation. It is natural to ask how positivity may be relaxed or generalised in a way which enables similar calculations to be made in more general contexts. On the one hand one may generalise by considering *almost additive* or *asymptotically additive potentials* which mimic the properties enjoyed by the logarithm of the norm of a positive matrix cocycle; on the other hand one may consider matrix cocycles which are *dominated*, a condition which includes positive matrix cocycles but is more general. In this article we explore the relationship between almost additivity and domination for planar cocycles. We show in particular that a locally constant linear cocycle in the plane is almost additive if and only if it is either conjugate to a cocycle of isometries, or satisfies a property slightly weaker than domination which is introduced in this paper. Applications to matrix thermodynamic formalism are presented.

## 1. INTRODUCTION

For the purposes of this article a *linear cocycle* over a dynamical system  $T: X \rightarrow X$  will be a skew-product

$$F: X \times \mathbb{R}^d \rightarrow X \times \mathbb{R}^d, \quad (x, p) \mapsto (Tx, A(x)p),$$

where  $A: X \rightarrow GL_d(\mathbb{R})$  is continuous and  $X$  is a compact metric space. Writing  $A_T^n(x) = A(T^{n-1}x) \cdots A(x)$ , we thus have  $F^n(x, p) = (T^n x, A_T^n(x)p)$  for all  $n \in \mathbb{N}$  and

$$A_T^{m+n}(x) = A_T^m(T^n x) A_T^n(x) \tag{1.1}$$

for all  $m, n \in \mathbb{N}$ . In numerous contexts it has been found useful to consider cocycles in which all of the matrices  $A(x)$  are *positive*: we note for example such diverse articles as [19, 20, 23, 31]. Under this assumption the cocycle satisfies the inequality

$$|\log \|A_T^{m+n}(x)\| - \log \|A_T^m(T^n x)\| - \log \|A_T^n(x)\|| \leq C$$

for some constant  $C > 0$  depending only on  $A$ . This has led some authors to extend results for positive linear cocycles by considering, instead of a linear cocycle, a sequence of continuous functions  $f_n: X \rightarrow \mathbb{R}$  satisfying the inequality

$$|f_{n+m}(x) - f_m(T^n x) - f_n(x)| \leq C$$

for all  $x \in X$  and  $n, m \geq 1$ . Such sequences of functions are referred to in the literature as *almost additive* and have been investigated in [4, 6, 10, 21, 33]. The condition of almost additivity implies trivially a further property, *asymptotic additivity* (see for example Feng and Huang [16, Proposition A.5]), which has been applied in [13, 16, 22]. In another category of works, positivity is replaced by the more general hypothesis of *domination*: under this hypothesis there exists a continuous splitting  $\mathbb{R}^d = \mathcal{U}(x) \oplus \mathcal{V}(x)$ , which is preserved by the cocycle, such that  $\|A_T^n(x)u\| \geq Ce^{n\varepsilon}\|A_T^n(x)v\|$  for all unit vectors  $u \in \mathcal{U}(x)$  and  $v \in \mathcal{V}(x)$ , for some constants  $C, \varepsilon > 0$  (see [7] and references therein). For linear cocycles the hypothesis of domination implies

---

*Date:* July 25, 2019.

*2000 Mathematics Subject Classification.* Primary 37D30, 37D35.

*Key words and phrases.* Products of matrices, dominated splitting, thermodynamic formalism, almost additivity, Gibbs states.

the hypothesis of almost additivity, but the converse is false, as can be seen trivially for the case of cocycles where all of the linear maps are isometries, or where all are equal to the identity. The purpose of this article is to explore precisely the relationship between domination and almost additivity in the context of locally constant two-dimensional linear cocycles over the shift. In this project we are motivated principally by applications to the topics of matrix thermodynamic formalism and the geometry of self-affine fractals.

We consider cocycles in the simplest non-commutative setting, namely in the case of planar matrices. A cocycle is dominated if and only if there is a uniform exponential gap between singular values of its iterates. This is equivalent to the existence of a strongly invariant multicone in the projective space; see [1, 7]. Domination originates from [28, 29] and it is an important concept in differentiable dynamical systems; see [9, 11]. Our contribution in this article to this line of research is to show that a planar matrix cocycle is dominated if and only if matrices are proximal and the norms in the generated sub-semigroup satisfy a certain multiplicativity property; see Corollary 2.4. Higher dimensions are more difficult: [7, §4] show that the connected components of the multicone need not be convex.

Of the several motivations for studying almost additive potentials, this article is concerned principally with thermodynamic formalism. In Theorem 2.9 we will show that almost additive potentials arising from the norm potential of a two-dimensional locally-constant linear cocycle over the full shift can in almost all cases be studied simply by using the classical thermodynamic formalism. In fact, in our results, we are able to characterise all the properties of equilibrium states for these norm potentials by means of the properties of matrices. Theorem 2.8 gives a positive answer to [2, Question 7.4] in the two dimensional case. Furthermore, in Example 2.10, answering a folklore question, we show the existence of a quasi-Bernoulli equilibrium state which is not a Gibbs measure for any Hölder continuous potential.

## 2. PRELIMINARIES AND STATEMENTS OF RESULTS

For the remainder of this article we specialise to cocycles whose values are invertible two-dimensional real matrices. We take  $A \subset GL_2(\mathbb{R})$ , set  $X = A^{\mathbb{N}}$ , denote the left shift on  $X$  by  $T$ , and let  $A(x)$  be the first matrix in the infinite sequence  $x \in X$ . Let

$$F: X \times \mathbb{R}^d \rightarrow X \times \mathbb{R}^d, \quad (x, p) \mapsto (Tx, A(x)p)$$

be a linear cocycle over  $T$ . We see that  $A_T^n(x)$  is the product of  $n$  first matrices in  $x \in X$ , and the cocycle identity (1.1) clearly holds. Let  $\mathcal{S}(A)$  denote the sub-semigroup generated by  $A$ , that is,  $\mathcal{S}(A) = \{A_1 \cdots A_n : n \in \mathbb{N} \text{ and } A_i \in A \text{ for all } i \in \{1, \dots, n\}\}$ . So in particular,  $A_T^n(x) \in \mathcal{S}(A)$  for all  $x = (A_1, A_2, \dots) \in X$  and  $n \in \mathbb{N}$ .

**2.1. Domination.** Following [7] we say that a compact and nonempty subset  $A \subset GL_2(\mathbb{R})$  is *dominated* if there exist constants  $C > 0$  and  $0 < \tau < 1$  such that

$$\frac{|\det(A_1 \cdots A_n)|}{\|A_1 \cdots A_n\|^2} \leq C\tau^n$$

for all  $A_1, \dots, A_n \in A$ . We let  $\mathbb{RP}^1$  denote the real projective line, which is the set of all lines through the origin in  $\mathbb{R}^2$ . We call a proper subset  $\mathcal{C} \subset \mathbb{RP}^1$  a *multicone* if it is a finite union of closed projective intervals. We say that  $\mathcal{C} \subset \mathbb{RP}^1$  is a *strongly invariant multicone* for  $A \subset GL_2(\mathbb{R})$  if it is a multicone and  $A\mathcal{C} \subset \mathcal{C}^o$  for all  $A \in A$ . Here  $\mathcal{C}^o$  is the interior of  $\mathcal{C}$ . By [7, Theorem B], a compact set  $A \subset GL_2(\mathbb{R})$  has a strongly invariant multicone if and only if  $A$  is dominated. We say that  $\mathcal{C} \subset \mathbb{RP}^1$  is an *invariant multicone* for  $A \subset GL_2(\mathbb{R})$  if it is a multicone and  $A\mathcal{C} \subset \mathcal{C}$  for all  $A \in A$ .

Recall that a matrix  $A$  is *proximal* if it has two real eigenvalues with unequal absolute values, *parabolic* if it has only one eigenspace, i.e. the single eigenvalue has geometric multiplicity one, and *conformal* if it has two eigenvalues with the same absolute values. In other words, a matrix  $A$  is conformal if and only if there exists an invertible matrix  $M$ , which we call a *conjugation matrix* of  $A$ , such that  $|\det(A)|^{-1/2}MAM^{-1} \in O(2)$ , where  $O(2)$  is the group of  $2 \times 2$  orthogonal matrices.

Furthermore, we say that a set  $A \subset GL_2(\mathbb{R})$  is *strongly conformal* if all the elements of  $A$  are conformal with respect to the same conjugation matrix. Strongly conformality is equivalent to the fact that all the elements in the generated semigroup are conformal.

For a proximal matrix  $A$ , let  $\lambda_u(A)$  and  $\lambda_s(A)$  be the largest and smallest eigenvalues of  $A$  in absolute value, respectively. If the eigenvalues are equal in absolute value, then the choice of  $\lambda_u(A)$  and  $\lambda_s(A)$  is arbitrary. Note that if  $A$  is diagonalisable, then there exist linearly independent subspaces  $u(A), s(A) \in \mathbb{RP}^1$  such that  $|\lambda_u(A)| = \|A|u(A)\|$  and  $|\lambda_s(A)| = \|A|s(A)\|$ . We call  $u(A) \in \mathbb{RP}^1$  the eigenspace of  $A$  corresponding to  $\lambda_u(A)$  and  $s(A) \in \mathbb{RP}^1$  the eigenspace corresponding to  $\lambda_s(A)$ . If  $A \subset GL_2(\mathbb{R})$ , then we define  $X_u(A)$  and  $X_s(A)$  to be the closures of the sets of all unstable and stable directions of proximal elements of  $\mathcal{S}(A)$ , i.e. the sets

$$X_u(A) = \overline{\{u(A) : A \in \mathcal{S}(A) \text{ is proximal}\}},$$

$$X_s(A) = \overline{\{s(A) : A \in \mathcal{S}(A) \text{ is proximal}\}},$$

respectively. Recall that  $\mathcal{S}(A)$  is the sub-semigroup of  $GL_2(\mathbb{R})$  generated by  $A$ , i.e. the set of all finite products formed by the elements of  $A$ . We say that  $A \subset GL_2(\mathbb{R})$  has an *unstable multicone*  $\mathcal{C}$  if  $\mathcal{S}(A)$  contains at least one proximal element and

- (1)  $\mathcal{C} \cap X_s(A) = \emptyset$ ,
- (2)  $\partial\mathcal{C} \cap X_u(A) = \emptyset$ ,
- (3) each connected component of  $\mathcal{C}$  intersects  $X_u(A)$ .

Finally, we say that a semigroup  $\mathcal{S} \subset GL_2(\mathbb{R})$  is *almost multiplicative* if there exists a constant  $\kappa > 0$  such that  $\|AB\| \geq \kappa\|A\|\|B\|$  for all  $A, B \in \mathcal{S}$ . We note that since clearly  $\|AB\| \leq \|A\|\|B\|$  for all  $A, B \in \mathcal{S}(A)$  for every  $A \subset GL_2(\mathbb{R})$ , the condition  $\|AB\| \geq \kappa\|A\|\|B\|$  for all  $A, B \in \mathcal{S}(A)$  is equivalent to the statement that every cocycle taking values in  $\mathcal{S}(A)$  is almost additive in the sense defined in the introduction.

Our main result for matrix cocycles is the following theorem.

**Theorem 2.1.** *Let  $A \subset GL_2(\mathbb{R})$ . If the sub-semigroup  $\mathcal{S}(A)$  is almost multiplicative, then exactly one of the two following conditions hold:*

- (1)  $A$  is strongly conformal,
- (2)  $A$  has an invariant unstable multicone and  $\mathcal{S}(A)$  does not contain parabolic elements.

The next two propositions show that if the proximal elements of  $A$  form a compact set, then the converse claim holds in Theorem 2.1.

**Proposition 2.2.** *Let  $A \subset GL_2(\mathbb{R})$  be such that  $A$  has an invariant unstable multicone and  $\mathcal{S}(A)$  does not contain parabolic elements. Let  $A_e$  be the collection of all conformal elements of  $A$ . Then*

- (1)  $A \setminus A_e$  is nonempty and contains only proximal elements,
- (2)  $A_e$  is strongly conformal and  $\mathcal{S}(\{| \det(A) |^{-1/2} A : A \in A_e\})$  is finite.

Moreover, if  $A \setminus A_e$  is compact, then  $A \setminus A_e$  has a strongly invariant multicone  $\mathcal{C}$  such that  $AC = \mathcal{C}$  for all  $A \in A_e$ .

**Proposition 2.3.** *Let  $A_e, A_h \subset GL_2(\mathbb{R})$  be such that*

- (1)  $A_h$  is nonempty, compact, and has a strongly invariant multicone  $\mathcal{C}$ ,
- (2)  $A_e$  is strongly conformal and  $AC = \mathcal{C}$  for all  $A \in A_e$ .

Then  $\mathcal{S}(A_e \cup A_h)$  is almost multiplicative.

The previous three statements have two immediate corollaries. The first one studies the case where  $A$  contains only proximal elements. The second one is for finite collections.

**Corollary 2.4.** *If  $A \subset GL_2(\mathbb{R})$  is compact, then the following two statements are equivalent:*

- (1)  $A$  has a strongly invariant multicone,
- (2)  $A$  contains only proximal elements and  $\mathcal{S}(A)$  is almost multiplicative.

**Corollary 2.5.** *If  $A \subset GL_2(\mathbb{R})$  is finite, then the following two statements are equivalent:*

- (1) the sub-semigroup  $\mathcal{S}(A)$  is almost multiplicative,
- (2)  $A$  can be decomposed into two sets  $A_e$  and  $A_h$  such that  $A_e$  is strongly conformal and if  $A_h \neq \emptyset$ , then  $A_h$  has a strongly invariant multicone  $\mathcal{C}$  such that  $AC = \mathcal{C}$  for all  $A \in A_e$ .

**2.2. Thermodynamic formalism.** If the set  $A \subset GL_2(\mathbb{R})$  is finite, then it makes sense to consider thermodynamic formalism for matrix cocycles. In this context, it is rather standard practise to use separate alphabet to index the elements in the sub-semigroup.

Let  $N \geq 2$  be an integer and  $\Sigma = \{1, \dots, N\}^{\mathbb{N}}$  be the collection of all infinite words obtained from integers  $\{1, \dots, N\}$ . We denote the left shift operator by  $\sigma$  and equip  $\Sigma$  with the product discrete topology. The *shift space*  $\Sigma$  is clearly compact. If  $\mathbf{i} = i_1 i_2 \dots \in \Sigma$ , then we define  $\mathbf{i}|_n = i_1 \dots i_n$  for all  $n \in \mathbb{N}$ . The empty word  $\mathbf{i}|_0$  is denoted by  $\emptyset$ . Define  $\Sigma_n = \{\mathbf{i}|_n : \mathbf{i} \in \Sigma\}$  for all  $n \in \mathbb{N}$  and  $\Sigma_* = \bigcup_{n \in \mathbb{N}} \Sigma_n \cup \{\emptyset\}$ . Thus  $\Sigma_*$  is the collection of all finite words. The length of  $\mathbf{i} \in \Sigma_* \cup \Sigma$  is denoted by  $|\mathbf{i}|$ . If  $\mathbf{i} \in \Sigma_n$  for some  $n$ , then we set  $[\mathbf{i}] = \{\mathbf{j} \in \Sigma : \mathbf{j}|_n = \mathbf{i}\}$ . The set  $[\mathbf{i}]$  is called a *cylinder set*. Cylinder sets are open and closed and they generate the Borel  $\sigma$ -algebra.

The longest common prefix of  $\mathbf{i}, \mathbf{j} \in \Sigma_* \cup \Sigma$  is denoted by  $\mathbf{i} \wedge \mathbf{j}$ . The concatenation of two words  $\mathbf{i} \in \Sigma_*$  and  $\mathbf{j} \in \Sigma_* \cup \Sigma$  is denoted by  $\mathbf{ij}$ . If  $A \subset \Sigma$  and  $\mathbf{i} \in \Sigma_*$ , then  $\mathbf{i}A = \{\mathbf{ij} : \mathbf{j} \in A\}$ . For example, if  $\mathbf{i}, \mathbf{j} \in \Sigma_*$ , then  $[\mathbf{ij}] = \mathbf{i}[\mathbf{j}] = \mathbf{ij}\Sigma$ . If  $\mathbf{i} \in \Sigma_*$  and  $n \in \mathbb{N}$ , then by  $\mathbf{i}^n$  we mean the concatenation  $\mathbf{i} \dots \mathbf{i}$  where  $\mathbf{i}$  is repeated  $n$  times. Finally, denote by  $\#_k \mathbf{i}$  the number of appearances of the symbol  $k \in \{1, \dots, N\}$  in  $\mathbf{i} \in \Sigma_*$ , i.e.  $\#_k \mathbf{i} = \#\{n : i_n = k \text{ for } n \in \{1, \dots, |\mathbf{i}|\}\}$ .

We say that the sequence  $\Phi = (\phi_n)_{n \in \mathbb{N}}$  of functions  $\phi_n : \Sigma \rightarrow \mathbb{R}$  is *sub-additive* if there exists  $C_1 \geq 0$  such that

$$\phi_{n+m}(\mathbf{i}) \leq \phi_n(\mathbf{i}) + \phi_m(\sigma^n \mathbf{i}) + C_1$$

for all  $n, m \in \mathbb{N}$  and  $\mathbf{i} \in \Sigma$ . A sub-additive sequence  $\Phi = (\phi_n)_{n \in \mathbb{N}}$  is *almost-additive* if there exists  $C_2 \geq 0$  such that

$$\phi_{n+m}(\mathbf{i}) \geq \phi_n(\mathbf{i}) + \phi_m(\sigma^n \mathbf{i}) - C_2$$

for all  $n, m \in \mathbb{N}$  and  $\mathbf{i} \in \Sigma$ . Finally, we say that an almost-additive sequence  $\Phi$  is *additive* if the constants  $C_1$  and  $C_2$  in the above inequalities can be chosen to 0. For example, if  $\phi : \Sigma \rightarrow \mathbb{R}$  is a function, then  $(\sum_{k=0}^{n-1} \phi \circ \sigma^k)_{n \in \mathbb{N}}$  is additive. In this context, the function  $\phi$  is called a *potential*. We say that a potential  $\phi$  is *Hölder continuous*, if there exist  $C > 0$  and  $0 < \tau < 1$  such that

$$|\phi(\mathbf{i}) - \phi(\mathbf{j})| \leq C\tau^{|\mathbf{i} \wedge \mathbf{j}|}.$$

for all  $\mathbf{i}, \mathbf{j} \in \Sigma$ .

If  $\Phi = (\phi_n)_{n \in \mathbb{N}}$  is sub-additive, then the *pressure* of  $\Phi$  is defined by

$$P(\Phi) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{\mathbf{i} \in \Sigma_n} \exp \max_{\mathbf{j} \in [\mathbf{i}]} \phi_n(\mathbf{j}). \quad (2.1)$$

The limit above exists by the standard properties of sub-additive sequences. Let  $\mu$  be a  $\sigma$ -invariant probability measure on  $\Sigma$  and recall that the *Kolmogorov-Sinai entropy* of  $\mu$  is

$$h_\mu = - \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\mathbf{i} \in \Sigma_n} \mu([\mathbf{i}]) \log \mu([\mathbf{i}]).$$

In addition, if  $\Phi = (\phi_n)_{n \in \mathbb{N}}$  is a sub-additive sequence, then we set

$$\Lambda_\mu(\Phi) = \lim_{n \rightarrow \infty} \frac{1}{n} \int_\Sigma \phi_n(\mathbf{i}) d\mu(\mathbf{i}).$$

It is easy to see that

$$P(\Phi) \geq h_\mu + \Lambda_\mu(\Phi)$$

for all  $\sigma$ -invariant probability measures  $\mu$ . The variational principle

$$P(\Phi) = \sup \{h_\mu + \Lambda_\mu(\Phi) : \mu \text{ is } \sigma\text{-invariant and } \Lambda_\mu(\Phi) \neq -\infty\}$$

is proved in [14]. For matrix cocycles this was obtained earlier in [24]. A  $\sigma$ -invariant measure  $\mu$  satisfying

$$P(\Phi) = h_\mu + \Lambda_\mu(\Phi)$$

is called an *equilibrium state* for  $\Phi$ . Such a measure always exists in the context of matrix cocycles, but it is not known if a general sub-additive sequence has an equilibrium state; see [5].

We say that a probability measure  $\mu$  on  $\Sigma$  is *quasi-Bernoulli* if there exists a constant  $C \geq 1$  such that

$$C^{-1}\mu([i])\mu([j]) \leq \mu([ij]) \leq C\mu([i])\mu([j])$$

for all  $i, j \in \Sigma_*$ . If the constant  $C$  above can be chosen to 1, then  $\mu$  is a *Bernoulli* measure. In other words, a probability measure  $\mu$  is Bernoulli if there exist a probability vector  $(p_1, \dots, p_N)$  such that

$$\mu([i]) = p_{i_1} \cdots p_{i_n}$$

for all  $i = i_1 \cdots i_n \in \Sigma_n$  and  $n \in \mathbb{N}$ .

Let  $\phi: \Sigma \rightarrow \mathbb{R}$  be a continuous potential and  $\Phi = (\sum_{k=0}^{n-1} \phi \circ \sigma^k)_{n \in \mathbb{N}}$ . We say that a Borel probability measure  $\mu$  on  $\Sigma$  is a *Gibbs measure* for  $\phi$  if there exists a constant  $C \geq 1$  such that

$$C^{-1} \exp\left(-nP(\Phi) + \sum_{k=0}^{n-1} \phi(\sigma^k(j))\right) \leq \mu([i]) \leq C \exp\left(-nP(\Phi) + \sum_{k=0}^{n-1} \phi(\sigma^k(j))\right) \quad (2.2)$$

for all  $i \in \Sigma_n$ ,  $j \in [i]$ , and  $n \in \mathbb{N}$ . For example, the Bernoulli measure obtained from a probability vector  $(p_1, \dots, p_N)$  is a Gibbs measure for the potential  $i \mapsto \log p_{i_1}$ . If  $\phi$  is Hölder continuous, then there is unique  $\sigma$ -invariant Gibbs measure which also is unique equilibrium state; see [12, Theorems 1.4 and 1.22].

Similarly, if  $\Phi = (\phi_n)_{n \in \mathbb{N}}$  is sub-additive, then a Borel probability measure  $\mu$  on  $\Sigma$  is a *Gibbs-type measure* for  $\Phi$  if there exists a constant  $C \geq 1$  such that

$$C^{-1} \exp\left(-nP(\Phi) + \phi_n(j)\right) \leq \mu([i]) \leq C \exp\left(-nP(\Phi) + \phi_n(j)\right) \quad (2.3)$$

for all  $i \in \Sigma_n$ ,  $j \in [i]$ , and  $n \in \mathbb{N}$ . It is easy to see that a  $\sigma$ -invariant Gibbs-type measure is ergodic and hence the unique equilibrium state; see [26, §3.2]. If  $\Phi$  is almost-additive, then, similarly as with continuous potentials, there exist conditions to guarantee the existence of a  $\sigma$ -invariant Gibbs-type; see [5, §4.2].

Our main objective is to study thermodynamic formalism in the setting of matrix cocycles. Let  $A = (A_1, \dots, A_N) \in GL_2(\mathbb{R})^N$ ,  $s > 0$ , and define  $\phi_n^s: \Sigma \rightarrow \mathbb{R}$  for all  $n \in \mathbb{N}$  by setting  $\phi_n^s(i) = \log \|A_{i|n}\|^s$ , where  $A_i = A_{i_1} \cdots A_{i_n}$  for all  $i = i_1 \cdots i_n \in \Sigma_n$  and  $n \in \mathbb{N}$ . Then the sequence  $\Phi^s = (\phi_n^s)_{n \in \mathbb{N}}$  parametrised by  $s > 0$  is sub-additive. By [24, Theorems 2.6 and 4.1], for every choice of the matrix tuple  $A$ , there exists an ergodic equilibrium state for  $\Phi^s$ . The structure of the set of all equilibrium states for  $\Phi^s$  is well known. We say that  $A = (A_1, \dots, A_N) \in GL_2(\mathbb{R})^N$  is *irreducible* if there does not exist 1-dimensional linear subspace  $V$  such that  $A_i V = V$  for all  $i \in \{1, \dots, N\}$ ; otherwise  $A$  is *reducible*. In a reducible tuple  $A$ , all the matrices are simultaneously upper triangular in some basis. If  $A$  is irreducible, then there is unique equilibrium state which is a Gibbs-type measure for  $\Phi^s$ ; see [17, Proposition 1.2]. It is worthwhile to remark that irreducibility does not imply that  $\Phi^s$  is almost-additive. In the reducible case, there can be two distinct ergodic equilibrium states; see [17, Theorem 1.7]. Recall also that the set  $\{A \in GL_2(\mathbb{R})^N : A \text{ is irreducible}\}$  is open, dense, and of full Lebesgue measure in  $GL_2(\mathbb{R})^N$ . In fact, the complement of the set is a finite union of  $(4N - 1)$ -dimensional algebraic varieties; see [25, Propositions 3.4 and 3.6].

The following four results characterise different kind of properties equilibrium states for  $\Phi^s$  can have by means of the matrix tuple.

**Proposition 2.6.** *If  $A = (A_1, \dots, A_N) \in GL_2(\mathbb{R})^N$  and  $\mu$  is an ergodic equilibrium state for  $\Phi^s$ , then the following two statements are equivalent:*

- (1)  $\mu$  is a Gibbs-type measure for  $\Phi^s$ ,
- (2) at least one of the following three conditions hold:
  - (a)  $A$  is irreducible,
  - (b)  $A$  is strongly conformal,

- (c)  $A$  is reducible with a common invariant subspace  $V$  and there exists  $\varepsilon > 0$  such that either the closed  $\varepsilon$ -neighbourhood of  $V$  or the closure of its complement is an invariant unstable multicone.

Note that  $A$  can be both irreducible and strongly conformal and that neither condition imply each other.

**Proposition 2.7.** *If  $A = (A_1, \dots, A_N) \in GL_2(\mathbb{R})^N$  and  $\mu$  is an ergodic equilibrium state for  $\Phi^s$ , then the following two statements are equivalent:*

- (1)  $\mu$  is a Bernoulli measure,
- (2)  $A$  is reducible or  $A$  is strongly conformal.

In the previous two propositions, one has to assume that the equilibrium measure is ergodic; see [27, Example 6.2] for a counter-example. We remark that the Bernoulli property has been studied earlier in [30, Theorem 13]. Since the propositions give a complete characterisation of the properties in the reducible case, we can restrict our attention into irreducible matrix tuples.

**Theorem 2.8.** *If  $A = (A_1, \dots, A_N) \in GL_2(\mathbb{R})^N$  is irreducible and  $\mu$  is an equilibrium state for  $\Phi^s$ , then the following four statements are equivalent:*

- (1)  $\mu$  is a quasi-Bernoulli measure,
- (2)  $\mathcal{S}(A)$  is almost multiplicative,
- (3)  $A$  can be decomposed into two sets  $A_e$  and  $A_h$  such that  $A_e$  is strongly conformal and if  $A_h \neq \emptyset$ , then  $A_h$  has a strongly invariant multicone  $\mathcal{C}$  such that  $A\mathcal{C} = \mathcal{C}$  for all  $A \in A_e$ ,
- (4) there exist a constant  $C > 0$  and a  $\mu$ -almost everywhere continuous potential  $f \in L^1(\mu)$  such that

$$\left| \sum_{k=0}^{n-1} f(\sigma^k \mathbf{i}) - \log \|A_{\mathbf{i}|_n}\| \right| \leq C \quad (2.4)$$

for all  $\mathbf{i} \in \Sigma$  and  $n \in \mathbb{N}$ .

The previous theorem gives a positive answer to [2, Question 7.4] in the two dimensional case.

**Theorem 2.9.** *If  $A = (A_1, \dots, A_N) \in GL_2(\mathbb{R})^N$  is irreducible and  $\mu$  is an equilibrium state for  $\Phi^s$ , then the following three statements are equivalent:*

- (1)  $\mu$  is a Gibbs measure for some Hölder continuous potential,
- (2)  $A$  has a strongly invariant multicone or  $A$  is strongly conformal,
- (3) there exist a constant  $C > 0$  and a Hölder-continuous potential  $f$  such that

$$\left| \sum_{k=0}^{n-1} f(\sigma^k \mathbf{i}) - \log \|A_{\mathbf{i}|_n}\| \right| \leq C$$

for all  $\mathbf{i} \in \Sigma$  and  $n \in \mathbb{N}$ .

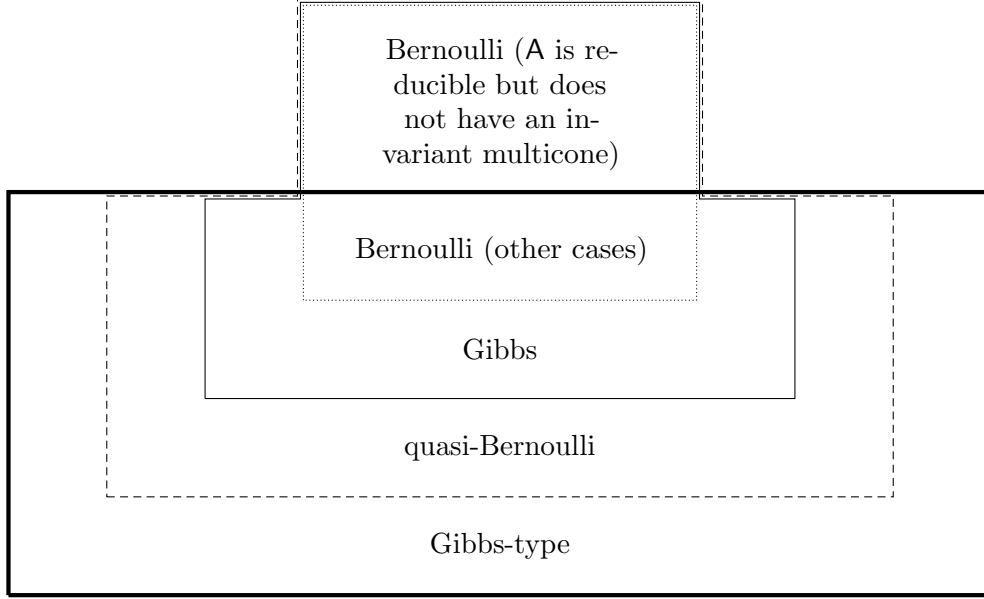
Figure 1 illustrates how different properties of equilibrium states for  $\Phi^s$  are related. The following example shows that the inclusions depicted in the figure are strict.

**Example 2.10.** (1) It can happen that an equilibrium state for  $\Phi^s$  is a Gibbs measure for some Hölder-continuous potential, but is not a Bernoulli measure: Choose two positive matrices

$$A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Then  $(A_1, A_2)$  is irreducible and has a strongly invariant multicone (i.e. the union of the first and third quadrants). The claim follows now from Theorem 2.9 and Proposition 2.7.

(2) It can happen that an equilibrium state for  $\Phi^s$  is a quasi-Bernoulli measure, but is not a Gibbs measure for any Hölder-continuous potential: Let  $A_1$  and  $A_2$  be as above. Then  $(A_1, A_2, I)$

FIGURE 1. Classification of equilibrium states for  $\Phi^s$ .

is irreducible and has an invariant multicone (i.e. the union of the first and third quadrants). The claim follows now from Theorems 2.8 and 2.9.

(3) It can happen that an equilibrium state for  $\Phi^s$  is a Gibbs-type measure for  $\Phi^s$ , but is not a quasi-Bernoulli measure: Choose two matrices

$$A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad A_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then  $(A_3, A_4)$  is irreducible, has no invariant multicone, and does not contain only conformal matrices. The claim follows now from Proposition 2.6 and Theorem 2.8. We remark that this phenomenon has been observed earlier in [18, §1.4]. Another way to see the claim is to consider two conformal irreducible matrices not sharing a conjugation matrix.

### 3. CHARACTERIZATION OF DOMINATION

In this section, we prove Theorem 2.1 and Propositions 2.2 and 2.3. Let  $A \subset GL_2(\mathbb{R})$  and recall that  $\mathcal{S}(A)$  is the sub-semigroup of  $GL_2(\mathbb{R})$  generated by  $A$ . Let  $\mathcal{S}(A) = \overline{\mathbb{R}\mathcal{S}(A)} \subset M_2(\mathbb{R})$  and note that  $\mathcal{S}(A)$  is a sub-semigroup of  $M_2(\mathbb{R})$ . Define

$$\mathcal{R}(A) = \{A \in \mathcal{S}(A) : \text{rank}(A) = 1\}.$$

**Lemma 3.1.** *If  $A \subset GL_2(\mathbb{R})$ , then  $\mathcal{R}(A) = \emptyset$  if and only if  $A$  is strongly conformal.*

*Proof.* If  $A$  is strongly conformal, then by definition there exists a conjugation matrix  $M \in GL_2(\mathbb{R})$  such that  $|\det(A)|^{-1/2} M A M^{-1} \in O(2)$  for all  $A \in A$ , which implies that  $|\det(A)|^{-1/2} M A M^{-1} \in O(2)$  for all nonzero  $A \in \mathcal{S}(A) = \overline{\mathbb{R}\mathcal{S}(A)}$ . In particular, all nonzero elements of  $\mathcal{S}(A)$  have rank 2 and therefore  $\mathcal{R}(A) = \emptyset$ .

Suppose conversely that  $\mathcal{R}(A) = \emptyset$ . We claim that set

$$\mathcal{S}'(A) = \{|\det(A)|^{-1/2} A : A \in \mathcal{S}(A) \setminus \{0\}\} = \mathcal{S}(A) \cap \{A \in M_2(\mathbb{R}) : |\det(A)| = 1\}$$

is compact. It is obviously closed, being the intersection of  $\mathcal{S}(A)$  with the closed set  $\{A \in M_2(\mathbb{R}) : |\det(A)| = 1\}$ . If it contains a sequence of elements  $(A_n)$  such that  $\|A_n\| \rightarrow \infty$  then this sequence can without loss of generality be taken to be a sequence of elements of  $\mathbb{R}\mathcal{S}(A)$ . The sequence of normalised matrices  $\|A_n\|^{-1} A_n$  has an accumulation point which necessarily has determinant zero and norm one and belongs to  $\mathcal{S}(A)$ ; this limit point is thus an element of  $\mathcal{R}(A)$ , which is a

contradiction, and we conclude that  $\{|\det(A)|^{-1/2}A : A \in \mathcal{S}(A) \setminus \{\mathbf{0}\}\}$  is bounded. It is therefore compact as claimed.

The set  $\mathcal{S}'(A)$  is thus a compact sub-semigroup of  $GL_2(\mathbb{R})$ . We claim that it is a group. To show this it is sufficient to show that the inverse of every  $A \in \mathcal{S}'(A)$  with  $|\det(A)| = 1$  belongs to  $\mathcal{S}'(A)$ . If  $A \in \mathcal{S}'(A)$  is arbitrary, take a convergent subsequence  $(A^{n_k})_{k=1}^\infty$  of the sequence  $(A^n)_{n=1}^\infty$  with limit  $B \in \mathcal{S}'(A) \subset GL_2(\mathbb{R})$ , say. The sequence  $(A^{-n_k})_{k=1}^\infty$  clearly converges to  $B^{-1}$  and therefore  $A^{n_{k+1}-n_k-1} \rightarrow A^{-1}$  as  $k \rightarrow \infty$ . Thus  $A^{-1}$  is the accumulation point of a sequence of elements of  $\mathcal{S}'(A)$ , hence an element of  $\mathcal{S}'(A)$ .

The set  $\mathcal{S}'(A)$  is therefore a compact subgroup of  $GL_2(\mathbb{R})$ . If  $m$  is Haar measure on  $\mathcal{S}'(A)$  and  $\langle \cdot, \cdot \rangle$  is the standard inner product on  $\mathbb{R}^2$  it is easy to see that  $\langle u, v \rangle' := \int \langle Au, Av \rangle dm(A)$  defines an inner product on  $\mathbb{R}^2$  which is invariant under every element of  $\mathcal{S}'(A)$ . Every inner product on  $\mathbb{R}^2$  is related to the standard one by a change of basis, so there exists  $X \in GL_2(\mathbb{R})$  such that  $\langle u, v \rangle' = \langle Xu, Xv \rangle$  for all  $u, v \in \mathbb{R}^2$ . In particular,  $\langle XAX^{-1}u, XAX^{-1}v \rangle = \langle AX^{-1}u, AX^{-1}v \rangle' = \langle X^{-1}u, X^{-1}v \rangle' = \langle u, v \rangle$  for all  $u, v \in \mathbb{R}^2$  and  $A \in \mathcal{S}'(A)$  which yields  $\mathcal{S}'(A) \subset XO(2)X^{-1}$ . Thus  $\mathcal{S}(A)$  is strongly conformal and therefore  $A$  is strongly conformal as required.  $\square$

We note that according to the previous lemma,  $\mathcal{R}(A) \neq \emptyset$  if and only if  $\mathcal{S}(A)$  contains at least one proximal or parabolic element. In the next lemma, we exclude parabolic elements.

**Lemma 3.2.** *Let  $A \in GL_2(\mathbb{R})$  with  $\mathcal{R}(A) \neq \emptyset$  be such that  $\mathcal{S}(A)$  is almost multiplicative. Then  $\mathcal{S}(A)$  does not contain parabolic elements and  $\mathcal{R}(A)$  does not contain nilpotent elements.*

*Proof.* Suppose that  $\mathcal{S}(A)$  contains a parabolic element. This means that, after a suitable change of basis, there exists  $A \in \mathcal{S}(A)$  such that

$$A = \begin{pmatrix} a & 0 \\ b & a \end{pmatrix},$$

where  $b \neq 0$ . It follows that there exists  $c > 0$  such that  $c^{-1}n|a^{n-1}b| \leq \|A^n\| \leq cn|a^{n-1}b|$  for all  $n \in \mathbb{N}$ . It follows directly that  $\lim_{n \rightarrow \infty} \|A^{2n}\|/\|A^n\|^2 = 0$  which contradicts the condition  $\|AB\| \geq \kappa\|A\|\|B\|$ .

Observe that the relation  $\|AB\| \geq \kappa\|A\|\|B\|$  holds for all  $A, B \in \mathcal{S}(A)$  by continuity. So similarly, if there exists a nilpotent  $A \in \mathcal{R}(A)$ , then  $0 = \|A^n\| \geq \kappa^{n-1}\|A\|^n > 0$  for some  $n \in \mathbb{N}$ , which is again a contradiction.  $\square$

Assuming  $\mathcal{R}(A) \neq \emptyset$ , we define the set  $X_u$  of all *unstable directions* of proximal elements of  $\mathcal{S}(A)$  to be

$$X_u = X_u(A) = \{V \in \mathbb{RP}^1 : V = A\mathbb{R}^2 \text{ for some } A \in \mathcal{R}(A)\}$$

and the set  $X_s$  of all *stable directions* to be

$$X_s = X_s(A) = \{\ker(A) \in \mathbb{RP}^1 : A \in \mathcal{R}(A)\}.$$

**Lemma 3.3.** *Let  $A \in GL_2(\mathbb{R})$  with  $\mathcal{R}(A) \neq \emptyset$  be such that  $\mathcal{S}(A)$  is almost multiplicative. Then the sets  $X_u$  and  $X_s$  are nonempty, compact, and disjoint. Furthermore,  $AX_u \subset X_u$  for all  $A \in \mathcal{S}(A)$ .*

*Proof.* First, we note again that the relation  $\|AB\| \geq \kappa\|A\|\|B\|$  holds for all  $A, B \in \mathcal{S}(A)$ . To see that  $X_u$  and  $X_s$  are disjoint, note that if  $V \in X_u \cap X_s$  then there exist nonzero  $B_1, B_2 \in \mathcal{R}(A)$  such that  $B_2\mathbb{R}^2 = V$  and  $B_1V = \{0\}$ . Hence  $B_1B_2$  is the zero matrix but  $B_1$  and  $B_2$  are not, which contradicts  $\|B_1B_2\| \geq \kappa\|B_1\|\|B_2\| > 0$ . It follows that  $X_u \cap X_s$  is empty. The nonempty set

$$\mathcal{R}_1(A) = \{B \in \mathcal{R}(A) : \|B\| = 1\} = \{B \in \mathcal{S}(A) : \det(B) = 0 \text{ and } \|B\| = 1\}$$

is clearly a closed bounded subset of  $\mathcal{S}(A)$ , and in particular is compact. It follows that  $X_u$  and  $X_s$  are the images of continuous functions  $\mathcal{R}_1(A) \rightarrow \mathbb{RP}^1$  and hence are compact and nonempty.

To see the last claim, consider a subspace  $U$  such that  $U = AV$  for some  $V \in X_u$  and  $A \in \mathcal{S}(A)$ . We have  $V = B\mathbb{R}^2$  for some  $B \in \mathcal{R}(A)$ . Clearly  $AB$  has rank at most 1 and is nonzero since  $\|AB\| \geq \kappa\|A\|\|B\| > 0$ , so  $AB \in \mathcal{R}(A)$  and  $U \in X_u$ .  $\square$



The following lemma shows that the definitions of unstable and stable directions agree with the ones given in §2.1.

**Lemma 3.4.** *Let  $A \in GL_2(\mathbb{R})$  with  $\mathcal{R}(A) \neq \emptyset$  be such that  $\mathcal{S}(A)$  is almost multiplicative. Then*

$$X_u = \overline{\{u(A) : A \in \mathcal{S}(A) \text{ is proximal}\}},$$

$$X_s = \overline{\{s(A) : A \in \mathcal{S}(A) \text{ is proximal}\}}.$$

*Proof.* Let us first demonstrate the inclusions

$$X_u \subset \overline{\{u(A) : A \in \mathcal{S}(A) \text{ is proximal}\}}, \quad (3.1)$$

$$X_s \subset \overline{\{s(A) : A \in \mathcal{S}(A) \text{ is proximal}\}}. \quad (3.2)$$

Before doing so, we show that for every  $A \in \mathcal{R}(A)$  with  $\|A\| = 1$  and any sequence  $(B_n)_{n=1}^\infty$  of elements of  $\mathcal{S}(A)$  such that  $\|B_n\|^{-1}B_n \rightarrow A$  as  $n \rightarrow \infty$ , the sequence  $B_n$  contains only proximal elements for all sufficiently large  $n$ . By Lemma 3.2, no  $B_n$  may be a parabolic matrix. Let us contrarily assume that, after passing to a suitable subsequence, every  $B_n$  is conformal. Write  $B'_n := \|B_n\|^{-1}B_n$  for all  $n \in \mathbb{N}$ . Since  $A$  has rank one we have  $\det(A) = 0$  and therefore  $\det(B'_n) \rightarrow 0$ . Since every  $B'_n$  is conformal it satisfies  $(\text{tr } B'_n)^2 \leq 4|\det(B'_n)|$  and therefore  $\text{tr } B'_n \rightarrow 0$ . By the Cayley-Hamilton theorem, we have  $(B'_n)^2 - (\text{tr } B'_n)B'_n + (\det(B'_n))I = \mathbf{0}$  and since  $B'_n \rightarrow A$  we deduce that  $(B'_n)^2 \rightarrow 0$ . Since  $\|B'_n\| = 1$  for all  $n \in \mathbb{N}$  we get  $\|B_n^2\|/\|B_n\|^2 = \|(B'_n)^2\|/\|B'_n\|^2 = \|(B'_n)^2\| \rightarrow 0$ , but this contradicts  $\|B_n^2\| \geq \kappa\|B_n\|^2$ . We conclude that  $(B_n)_{n=1}^\infty$  is proximal for all sufficiently large  $n$  as claimed.

It is well known that the maps  $u(\cdot)$  and  $s(\cdot)$  are continuous on proximal matrices. Moreover, by Lemma 3.2, every  $A \in \mathcal{R}(A)$  is proximal. Hence, if  $V \in X_u$ , then there exists a proximal  $A \in \mathcal{R}(A)$  with  $\|A\| = 1$  such that  $V = A\mathbb{R}^2 = u(A)$ . Moreover, there exists a sequence of proximal matrices  $B_n \in \mathcal{S}(A)$  such that  $\|B_n\|^{-1}B_n \rightarrow A$  and thus, by the continuity of  $u$ ,  $u(B_n) = u(\|B_n\|^{-1}B_n) \rightarrow u(A) = V$ , which shows (3.1). Similarly, if  $V \in X_s$ , then there exists a proximal  $A \in \mathcal{R}(A)$  with  $\|A\| = 1$  such that  $V = \ker(A) = s(A)$ , and there exists a sequence of proximal matrices  $B_n \in \mathcal{S}(A)$  such that  $\|B_n\|^{-1}B_n \rightarrow A$ . Applying now the continuity of  $s$ , we get  $s(B_n) = s(\|B_n\|^{-1}B_n) \rightarrow s(A) = V$  showing (3.2).

To finish the characterization of  $X_u$  it is sufficient to show that

$$X_u \supset \{u(A) : A \in \mathcal{S}(A) \text{ is proximal}\}, \quad (3.3)$$

$$X_s \supset \{s(A) : A \in \mathcal{S}(A) \text{ is proximal}\}, \quad (3.4)$$

since we may then appeal to Lemma 3.3 and the fact that the sets  $X_u$  and  $X_s$  are closed. If  $V = u(A)$  for some proximal  $A \in \mathcal{S}(A)$ , then  $\|A^n\|^{-1}A^n \rightarrow B$  as  $n \rightarrow \infty$ , where  $B \in \mathcal{R}(A)$  is such that  $B\mathbb{R}^2 = u(B) = V$ . This shows (3.3). Similarly, if  $V = s(A)$  for some proximal  $A \in \mathcal{S}(A)$ , then  $\|A^n\|^{-1}A^n \rightarrow B$  as  $n \rightarrow \infty$ , where  $B \in \mathcal{R}(A)$  is such that  $\ker(B) = V$ . This shows (3.4) and completes the proof.  $\square$

Let  $d$  be the metric on  $\mathbb{RP}^1$  defined by taking  $d(U, V)$  to be the angle between the subspaces  $U$  and  $V$ . If  $A \in GL_2(\mathbb{R})$  is such that  $\mathcal{R}(A) \neq \emptyset$ , then we define

$$\mathcal{V}_n = \{U \in \mathbb{RP}^1 : d(U, V) < \frac{1}{n} \text{ for some } V \in X_u\}$$

and

$$\mathcal{U}_n = \bigcup_{A \in \mathcal{S}(A)} A\mathcal{V}_n$$

for all  $n \in \mathbb{N}$ .

**Lemma 3.5.** *Let  $A \in GL_2(\mathbb{R})$  with  $\mathcal{R}(A) \neq \emptyset$  be such that  $\mathcal{S}(A)$  is almost multiplicative. Then there is  $n_0 \in \mathbb{N}$  such that  $\overline{\mathcal{U}_n}$  as defined above is an invariant unstable multicone for all  $n \geq n_0$ .*

*Proof.* Note that for all  $n \in \mathbb{N}$  the invariance of  $\bar{\mathcal{U}}_n$  and the property (2) in the definition of the unstable multicone (see §2.1) follow immediately from the definition of the set  $\mathcal{U}_n$  and the continuity of each  $A \in \mathcal{S}(\mathbf{A})$  as an action on  $\mathbb{RP}^1$ . Let us prove the property (3) for all  $n \in \mathbb{N}$ . Obviously  $\mathcal{V}_n$  is open, and since each  $A \in \mathcal{S}(\mathbf{A})$  is invertible and therefore induces a homeomorphism of  $\mathbb{RP}^1$ , each  $\mathcal{U}_n$  is open too. It is clear from the definition that every connected component of  $\mathcal{V}_n$  intersects  $X_u$ . If  $U \in \mathcal{U}_n$ , then  $U = AU'$  for some  $A \in \mathcal{S}(\mathbf{A})$  and  $U' \in \mathcal{V}_n$ . Let  $\mathcal{I} \subset \mathcal{V}_n$  be an open connected set which contains  $U'$  and which also intersects  $X_u$ . The set  $A\mathcal{I}$  then contains  $U$ , is connected, and intersects  $AX_u$ . Since  $AX_u \subset X_u$  by Lemma 3.3, we conclude that each connected component of  $\mathcal{U}_n$  intersects  $X_u$ .

To show that the property (1) holds for all large enough  $n$ , let us suppose the contrary. In this case  $\bar{\mathcal{U}}_n \cap X_s$  must be nonempty for infinitely many  $n \in \mathbb{N}$ . This implies that in any prescribed neighbourhoods of  $X_u$  and  $X_s$  we may find a subspace  $U$  in the neighbourhood of  $X_u$  and a matrix  $A \in \mathcal{S}(\mathbf{A})$  such that  $AU$  belongs to the neighbourhood of  $X_s$ . It follows that we may choose a sequence of subspaces  $(U_n)$  converging to a limit  $U \in X_u$  and a sequence  $(A_n)$  of elements of  $\mathcal{S}(\mathbf{A})$  such that  $A_n U_n$  converges to a limit  $V \in X_s$ . Define  $B_n := \|A_n\|^{-1} A_n \in \mathcal{S}(\mathbf{A})$  for every  $n \in \mathbb{N}$ , and by passing to a subsequence if necessary we may suppose that  $(B_n)$  converges to a limit  $B \in \mathcal{S}(\mathbf{A})$  with norm 1.

We claim that  $BU = V$ . Let  $(u_n)$  be a sequence of unit vectors such that  $u_n \in U_n$  for every  $n \in \mathbb{N}$  and such that  $(u_n)$  converges to a unit vector  $u \in U$ . It is enough to show that  $(B_n u_n)$  converges to  $Bu$  and that  $Bu$  is nonzero, since we have then shown that  $V = \lim_{n \rightarrow \infty} B_n U_n = BU$ . To see that  $Bu$  is nonzero we note that  $u \in U \in X_u$  and  $B \in \mathcal{S}(\mathbf{A})$  with  $B \neq 0$ , so if  $Bu = 0$  then  $u \in \ker B \in X_s$  and we have  $U \in X_s \cap X_u$  contradicting Lemma 3.3. On the other hand since

$$0 \leq \|B_n u_n - Bu\| \leq \|B_n u_n - B_n u\| + \|B_n u - Bu\| \leq \|u_n - u\| + \|B_n - B\| \rightarrow 0$$

we have  $B_n u_n \rightarrow Bu$  as  $n \rightarrow \infty$  as required. But the equation  $BU = V$  is impossible since  $BU \in X_u$  by Lemma 3.3 and therefore  $V \in X_s \cap X_u$ , contradicting Lemma 3.3. We conclude that  $\bar{\mathcal{U}}_n \cap X_s$  must be empty for all large enough  $n$  and therefore property (1) holds for all  $n$  sufficiently large.

We are left to show that  $\bar{\mathcal{U}}_n$  is a multicone. To that end, it suffices to show that  $\partial\mathcal{U}_n$  contains only finitely many points. To see this suppose for a contradiction that  $U \in \mathbb{RP}^1$  is an accumulation point of a sequence  $(U_k)_{k=1}^\infty$  of distinct elements of  $\partial\mathcal{U}_n$ . We will find it convenient to identify a small open neighbourhood  $\mathcal{I}$  of  $U$  with a bounded open interval  $(a, b) \subset \mathbb{R}$ . By passing to a subsequence if necessary we may assume that  $(U_k)_{k=1}^\infty$  is monotone with respect to the natural order on  $\mathcal{I}$ , and without loss of generality we assume  $(U_k)_{k=1}^\infty$  to be strictly increasing.

We assert that every interval  $(U_k, U_{k+2})$  contains a point of  $X_u$ . Since  $U_{k+1}$  is in the closure of  $\mathcal{U}_n$ , there exists a point of  $\mathcal{U}_n$  in the interval  $(U_k, U_{k+2})$ . Since neither  $U_k$  nor  $U_{k+2}$  can belong to  $\mathcal{U}_n$ , it follows that some connected component of  $\mathcal{U}_n$  is contained wholly within the interval  $(U_k, U_{k+2})$ . By (3), this implies that a point of  $X_u$  must lie in the interval  $(U_k, U_{k+2})$ . Since this is true for every  $k \in \mathbb{N}$ , it follows that  $U$  is an accumulation point of  $X_u$  and hence, by Lemma 3.3,  $U$  belongs to  $X_u$ . But  $X_u$  is a subset of  $\mathcal{U}_n$  and therefore  $U \in \mathcal{U}_n$ , which implies that  $U_k \in \mathcal{U}_n$  for all sufficiently large  $k$ . This is clearly impossible since no element of  $\partial\mathcal{U}_n$  can be an element of  $\mathcal{U}_n$ . This contradiction proves that  $\partial\mathcal{U}_n$  must be finite.  $\square$

The above lemmas prove Theorem 2.1:

*Proof of Theorem 2.1.* If  $\mathcal{R}(\mathbf{A}) = \emptyset$ , then, by Lemma 3.1, the set  $\mathbf{A}$  is strongly conformal. If  $\mathcal{R}(\mathbf{A}) \neq \emptyset$ , then the claim follows from Lemmas 3.2 and 3.5.  $\square$

Let us next turn to the proof of the propositions.

**Lemma 3.6.** *Let  $A \in GL_2(\mathbb{R})$  and let  $\mathcal{C}$  be a multicone such that  $AC \subset \mathcal{C}$ . If  $A$  is conformal, then  $AC = \mathcal{C}$ .*

*Proof.* By a suitable change of basis, we may assume that  $A \in O(2)$ . In this case  $A$  preserves Lebesgue measure on  $\mathbb{RP}^1$ . If  $AC \subsetneq \mathcal{C}$ , then, since  $\mathcal{C}$  is a finite union of closed projective intervals and  $A$  is a homeomorphism,  $AC$  must have smaller Lebesgue measure than  $\mathcal{C}$ , which is a contradiction.  $\square$

We remark that the converse statement is false: if  $A$  is proximal and  $\mathcal{C}$  is a closed projective interval with one endpoint equal to  $u(A)$  and the other endpoint equal to  $s(A)$ , then  $A\mathcal{C} = \mathcal{C}$  but  $A$  is not conformal.

If  $A \in GL_2(\mathbb{R})$  and  $A_e$  is the collection of all conformal elements of  $A$ , then we write

$$\mathcal{F}(A) := \mathcal{S}(\{|\det(A)|^{-1/2}A : A \in A_e\}).$$

**Lemma 3.7.** *Let  $A \in GL_2(\mathbb{R})$  be such that  $\mathcal{R}(A) \neq \emptyset$  and  $A_e$  be the collection of all conformal elements of  $A$ . If  $\mathcal{C}$  is an invariant unstable multicone of  $A$ , then  $A_e = \{A \in A : A\mathcal{C} = \mathcal{C}\}$  is strongly conformal and  $\mathcal{F}(A)$  is finite.*

*Proof.* Since  $A$  has an invariant multicone  $\mathcal{C}$ , it follows from Lemma 3.6 that  $A\mathcal{C} = \mathcal{C}$  for all  $A \in A_e$ . Hence  $A_e \subset \{A \in A : A\mathcal{C} = \mathcal{C}\}$ .

Write  $A'_e = \{|\det(A)|^{-1/2}A : A \in A \text{ and } A\mathcal{C} = \mathcal{C}\}$ . Let us first assume that  $\#\partial\mathcal{C} > 2$ . Let  $B_1, B_2 \in \mathcal{S}(A'_e)$  and suppose that  $B_1$  and  $B_2$  induce the same permutation of  $\partial\mathcal{C}$ . Then  $B_1^{-1}B_2$  fixes every point of  $\partial\mathcal{C}$  and therefore has more than 2 invariant subspaces and is necessarily equal to  $\pm I$ . It follows that in this case  $\mathcal{S}(A'_e)$  has at most  $2(\#\partial\mathcal{C})!$  distinct elements. Let us now assume that  $\#\partial\mathcal{C} = 2$ . Write  $\partial\mathcal{C} = \{U_1, U_2\}$ , and let  $u_1 \in U_1$  and  $u_2 \in U_2$  be so that  $\{u_1, u_2\}$  is a basis for  $\mathbb{R}^2$ . Every element of  $\mathcal{S}(A'_e)$  preserves  $\partial\mathcal{C}$  and hence is either diagonal or antidiagonal in this basis (where by an antidiagonal matrix we mean a  $2 \times 2$  matrix with both main diagonal entries equal to zero and both other entries nonzero). Let  $D$  be the matrix which  $Du_1 = u_1$  and  $Du_2 = -u_2$ . A diagonal element of  $\mathcal{S}(A'_e)$  cannot be proximal since then either  $U_1$  or  $U_2$  would be the stable space of that matrix contradicting the property  $X_s \cap \mathcal{C} = \emptyset$  of the unstable multicone  $\mathcal{C}$ . It follows that every diagonal element of  $\mathcal{S}(A'_e)$  must belong to  $\{\pm I, \pm D\}$ . Let  $A_1, \dots, A_\ell$  be the anti-diagonal elements of  $A'_e$  and define  $S = \{\pm I, \pm D\} \cup \{\pm A_1, \dots, \pm A_\ell\} \cup \{\pm DA_1, \dots, \pm DA_\ell\}$ . The set  $S$  is a semigroup since  $A_i D = -DA_i$  and since each  $A_i A_j$  is diagonal and hence equal to  $\pm I$  or  $\pm D$ . In particular,  $\mathcal{S}(A'_e)$  is contained in a finite semigroup. Thus,  $A'_e$  is strongly conformal, which implies that  $\{A \in A : A\mathcal{C} = \mathcal{C}\} \subset A_e$ .  $\square$

**Lemma 3.8.** *Let  $A \in GL_2(\mathbb{R})$  be such that  $A$  has an invariant unstable multicone  $\mathcal{C}$  and  $\mathcal{S}(A)$  does not contain parabolic elements. Let  $A_e$  be the collection of all conformal elements of  $A$ . Then*

$$A_1 F_1 \cdots A_n F_n \mathcal{C} \subset \mathcal{C}^\circ$$

for all  $n \geq (\#\partial\mathcal{C})^2 + 1$ ,  $A_1, \dots, A_n \in A \setminus A_e$ , and  $F_1, \dots, F_n \in \mathcal{F}(A)$ .

*Proof.* It is sufficient to show that every point of  $\partial\mathcal{C}$  is mapped into  $\mathcal{C}^\circ$  by  $A_1 F_1 \cdots A_n F_n$ . Clearly, if there exists  $\ell \in \{1, \dots, n\}$  such that  $A_\ell F_\ell \cdots A_n F_n \mathcal{C} \subset \mathcal{C}^\circ$ , then our claim follows.

Suppose for a contradiction that there exist  $n \geq (\#\partial\mathcal{C})^2 + 1$ ,  $A_1, \dots, A_n \in A \setminus A_e$ , and  $F_1, \dots, F_n \in \mathcal{F}(A)$  such that for every  $\ell \in \{1, \dots, n\}$  there exist  $V_\ell, W_\ell \in \partial\mathcal{C}$  for which

$$A_\ell F_\ell \cdots A_n F_n V_\ell = W_\ell.$$

Since  $n \geq (\#\partial\mathcal{C})^2 + 1$ , there exist  $\ell_1 < \ell_2$  such that  $V_{\ell_1} = V_{\ell_2}$  and  $W_{\ell_1} = W_{\ell_2}$ . Hence,

$$A_{\ell_1} F_{\ell_1} \cdots A_{\ell_2-1} F_{\ell_2-1} W_{\ell_2} = W_{\ell_2}.$$

Thus, if  $A_{\ell_1} F_{\ell_1} \cdots A_{\ell_2-1} F_{\ell_2-1}$  is proximal, then  $W_{\ell_2} \in X_u \cup X_s$ . This is impossible, since  $\partial\mathcal{C} \cap (X_s \cup X_u) = \emptyset$ . If  $A_{\ell_1} F_{\ell_1} \cdots A_{\ell_2-1} F_{\ell_2-1}$  is conformal, then  $A_{\ell_1} F_{\ell_1} \cdots A_{\ell_2-1} F_{\ell_2-1} \mathcal{C} = \mathcal{C}$  by Lemma 3.6. This is also impossible, since  $\mathcal{C} \supsetneq A_{\ell_1} \mathcal{C} \supset A_{\ell_1} F_{\ell_1} \cdots A_{\ell_2-1} F_{\ell_2-1} \mathcal{C}$  by Lemma 3.7.  $\square$

**Lemma 3.9.** *Let  $A \in GL_2(\mathbb{R})$  be such that  $A$  has an invariant unstable multicone  $\mathcal{C}$  and  $\mathcal{S}(A)$  does not contain parabolic elements. Let  $A_e$  be the collection of all conformal elements of  $A$ . If  $A \setminus A_e$  is compact, then*

$$B = \{A_1 A_2 : A_1 \in A \setminus A_e \text{ and } A_2 \in \mathcal{F}(A)\}$$

has a strongly invariant multicone.

*Proof.* Write  $m = (\#\partial\mathcal{C})^2 + 1$  and note that, by Lemma 3.8,  $\mathbf{B}^m$  has a strongly invariant multicone. Since  $\mathbf{A} \setminus \mathbf{A}_e$  is compact by the assumption and  $\mathcal{F}(\mathbf{A})$  is finite by Lemma 3.7,  $\mathbf{B}^m$  is compact. Hence, by [7, Theorem B],  $\mathbf{B}^m$  is dominated, i.e. there exist constants  $C > 0$  and  $\tau > 1$  such that

$$\frac{\|B_1 \cdots B_n\|}{\|(B_1 \cdots B_n)^{-1}\|^{-1}} \geq C\tau^n.$$

for all  $B_1, \dots, B_n \in \mathbf{B}^m$  and all  $n \in \mathbb{N}$ . Choose  $k \in \mathbb{N}$  and let  $A_i F_i \in \mathbf{B}$  for all  $i \in \{1, \dots, k\}$ . Write  $k = qm + p$ , where  $q \in \mathbb{N} \cup \{0\}$  and  $p \in \{0, \dots, m-1\}$ . Then

$$\begin{aligned} \frac{\|A_1 F_1 \cdots A_k F_k\|}{\|(A_1 F_1 \cdots A_k F_k)^{-1}\|^{-1}} &= \frac{\|B_1 \cdots B_q \cdot A_{k-p} F_{k-p} \cdots A_k F_k\|}{\|(B_1 \cdots B_q \cdot A_{k-p} F_{k-p} \cdots A_k F_k)^{-1}\|^{-1}} \\ &\geq \frac{\|B_1 \cdots B_q\| \|(A_{k-p} F_{k-p} \cdots A_k F_k)^{-1}\|^{-1}}{\|(B_1 \cdots B_q)^{-1}\|^{-1} \|A_{k-p} F_{k-p} \cdots A_k F_k\|} \\ &\geq C\tau^q \frac{\|(A_{k-p} F_{k-p} \cdots A_k F_k)^{-1}\|^{-1}}{\|A_{k-p} F_{k-p} \cdots A_k F_k\|}. \end{aligned}$$

By choosing  $C' = C\tau^{-1} \min_{\ell \in \{1, \dots, m-1\}} \|(A_1 F_1 \cdots A_\ell F_\ell)^{-1}\|^{-1} / \|A_1 F_1 \cdots A_\ell F_\ell\|$  and  $\tau' = \tau^{1/m}$ , it follows again from [7, Theorem B] that  $\mathbf{B}$  has a strongly invariant multicone.  $\square$

The following lemma is [8, Lemma 2.2].

**Lemma 3.10.** *Let  $\mathcal{C}_0, \mathcal{C} \subset \mathbb{RP}^1$  be multicones such that  $\mathcal{C}_0 \subset \mathcal{C}^o$ . Then there exists a constant  $\kappa_0 > 0$  such that  $\|A|V\| \geq \kappa_0 \|A\|$  for all  $V \in \mathcal{C}_0$  and for every matrix  $A \in GL_2(\mathbb{R})$  with  $AC \subset \mathcal{C}_0$ .*

We are now ready to prove the propositions:

*Proof of Proposition 2.2.* The assertion (2) follows immediately from Lemma 3.7. Let us verify (1). If  $\mathbf{A}_e = \mathbf{A}$ , then  $\mathcal{S}(\mathbf{A})$  is strongly conformal since  $\mathbf{A}$  is. This means that  $\mathcal{S}(\mathbf{A})$  does not contain an proximal matrix and thus,  $\mathbf{A}$  cannot have an unstable multicone by definition. Therefore, (1) holds.

To prove the final claim, it is sufficient to show that, by assuming  $\mathbf{A} \setminus \mathbf{A}_e$  to be compact, there exists an invariant multicone  $\mathcal{C}$  such that  $AC \subset \mathcal{C}^o$  for all  $A \in \mathbf{A} \setminus \mathbf{A}_e$  and  $AC = \mathcal{C}$  for all  $A \in \mathbf{A}_e$ . By Lemma 3.7, the set  $\mathcal{F}(\mathbf{A})$  is finite. Therefore, the set  $\mathbf{B} = \{A_1 A_2 : A_1 \in \mathbf{A} \setminus \mathbf{A}_e \text{ and } A_2 \in \mathcal{F}(\mathbf{A})\}$  is compact and, by Lemma 3.9, it has a strongly invariant multicone  $\mathcal{C}_0$ . Defining

$$\mathcal{C} = \bigcup_{F \in \mathcal{F}(\mathbf{A})} F\mathcal{C}_0,$$

we have

$$AC = \bigcup_{F \in \mathcal{F}(\mathbf{A})} AFC_0 \subset \mathcal{C}_0^o \subset \mathcal{C}^o.$$

for all  $A \in \mathbf{A} \setminus \mathbf{A}_e$ . We have finished the proof since for any  $A \in \mathbf{A}_e$ ,  $AC = \mathcal{C}$  holds trivially.  $\square$

*Proof of Proposition 2.3.* Let  $\varepsilon > 0$  and define

$$\mathcal{C}_0 = \bigcup_{F \in \mathcal{F}(\mathbf{A})} F \left( \left\{ U \in \mathbb{RP}^1 : d(U, V) \leq \varepsilon \text{ for some } V \in \bigcup_{A \in \mathbf{A}_h} AC \right\} \right)$$

Recall that  $\mathcal{F}(\mathbf{A})$  is finite by Lemma 3.7. By compactness of  $\mathbf{A}_h$ , we may choose  $\varepsilon > 0$  small enough so that  $\mathcal{C}_0 \subset \mathcal{C}^o$ ,  $AC \subset \mathcal{C}_0$  for all  $A \in \mathbf{A}_h$ , and  $AC_0 = \mathcal{C}_0$  for all  $A \in \mathbf{A}_e$ . Observe that every element  $A \in \mathcal{S}(\mathbf{A}_h \cup \mathbf{A}_e)$  can be written in the form  $(c_0 c_1 \cdots c_k) F_0 \prod_{i=1}^k A_i F_i$ , where  $c_i \in \mathbb{R} \setminus \{0\}$ ,  $k \in \mathbb{N} \cup \{0\}$ ,  $A_i \in \mathbf{A}_h$ , and  $F_i \in \mathcal{F}(\mathbf{A})$ . Therefore,  $AC \subset \mathcal{C}_0$  for all  $A \in \mathcal{S}(\mathbf{A}_h \cup \mathbf{A}_e) \setminus \mathcal{S}(\mathbf{A}_e)$ .

By Lemma 3.10, there exists a constant  $\kappa_0 = \kappa_0(\mathcal{C}_0, \mathcal{C})$  such that  $\|A|V\| \geq \kappa_0 \|A\|$  for all  $V \in \mathcal{C}_0$  and for every matrix  $A \in GL_2(\mathbb{R})$  with  $AC \subset \mathcal{C}_0$ . Hence,

$$\|AB\| \geq \|AB|V\| = \|A|BV\| \|B|V\| \geq \kappa_0^2 \|A\| \|B\|.$$

for all  $A, B \in \mathcal{S}(\mathbf{A}_h \cup \mathbf{A}_e) \setminus \mathcal{S}(\mathbf{A}_e)$ . If  $A \in \mathcal{S}(\mathbf{A}_e)$  or  $B \in \mathcal{S}(\mathbf{A}_e)$ , then  $\|AB\| \geq \kappa' \|A\| \|B\|$  holds trivially for some  $\kappa' > 0$  by the finiteness of  $\mathcal{F}(\mathbf{A})$ .  $\square$

## 4. CLASSIFICATION OF EQUILIBRIUM STATES

This section is devoted to the proofs of Propositions 2.6 and 2.7, and Theorems 2.8 and 2.9. In order to keep the proof of Theorem 2.9 as readable as possible, we have postponed the proof of a key technical lemma, Lemma 4.6, into §5. Before we start with the proof of the propositions, we state a couple of auxiliary lemmas.

We recall that  $\lambda_u(A)$  is the eigenvalue of  $A$  with the largest absolute value, and similarly,  $\lambda_s(A)$  is the eigenvalue of  $A$  with the smallest absolute value. Note that  $|\lambda_u(A)| = \|A|u(A)\|$  and  $|\lambda_s(A)| = \|A|s(A)\|$ , where  $u(A)$  is the eigenspace corresponding to  $\lambda_u(A)$  and  $s(A)$  the eigenspace corresponding to  $\lambda_s(A)$ .

The following two lemmas are special cases of the result of Protasov and Voynov; see [32, Theorem 2]. In order to keep the paper as self-contained as possible, we give here alternative proofs.

**Lemma 4.1.** *Let  $\mathbf{A} = (A_1, \dots, A_N) \in GL_2(\mathbb{R})^N$  be such that all the elements of  $\mathbf{A}$  are proximal. Then the following two statements are equivalent:*

- (1)  $\lambda_u(A_i A_j) = \lambda_u(A_i) \lambda_u(A_j)$  for all  $i, j$ ,
- (2)  $u(A_i) = u(A_j)$  for all  $i, j$  or  $s(A_i) = s(A_j)$  for all  $i, j$ .

*Proof.* It is easy to see that (2) implies (1). Let us show that (1) implies (2). By the assumption and the multiplicativity of the determinant, we have  $\lambda_s(A_i A_j) = \lambda_s(A_i) \lambda_s(A_j)$  for all  $i, j$ . First note that  $s(A_i) \neq u(A_j)$ , for any  $i \neq j$ . Indeed,  $s(A_i) = u(A_j)$  would imply that the matrix  $A_i A_j$  has eigenvalue  $\lambda_s(A_i) \lambda_u(A_j)$ . Thus, either  $\lambda_s(A_i) \lambda_u(A_j) = \lambda_u(A_i) \lambda_u(A_j)$  or  $\lambda_s(A_i) \lambda_u(A_j) = \lambda_s(A_i) \lambda_s(A_j)$ , which implies that either  $\lambda_s(A_i) = \lambda_u(A_i)$  or  $\lambda_u(A_j) = \lambda_s(A_j)$ , which contradicts to the proximality.

We prove the statement by induction. Since  $s(A_1) \neq u(A_2)$ , after a suitable change of basis, the matrices  $A_1$  and  $A_2$  have the form

$$A_1 = \begin{pmatrix} \lambda_u(A_1) & 0 \\ a & \lambda_s(A_1) \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} \lambda_u(A_2) & b \\ 0 & \lambda_s(A_2) \end{pmatrix}.$$

Hence,  $\text{tr}(A_1 A_2) = \lambda_u(A_1 A_2) + \lambda_s(A_1 A_2) = \lambda_u(A_1) \lambda_u(A_2) + \lambda_s(A_1) \lambda_s(A_2) + ab$ . So  $ab = 0$ , which implies that if  $b = 0$  then  $s(A_1) = s(A_2)$  or if  $a = 0$  then  $u(A_1) = u(A_2)$ .

Let us then assume that the first  $N - 1$  matrices have the property that either  $u(A_i) = u(A_j)$  for all  $i, j \in \{1, \dots, N - 1\}$  or  $s(A_i) = s(A_j)$  for all  $i, j \in \{1, \dots, N - 1\}$ . We may assume without loss of generality that  $u(A_i) = u(A_j)$  for all  $i, j \in \{1, \dots, N - 1\}$ . For a fixed  $i \in \{1, \dots, N - 1\}$  the equation  $\lambda_u(A_i) \lambda_u(A_N) = \lambda_u(A_i A_N)$  holds only if  $u(A_i) = u(A_N)$  or  $s(A_i) = s(A_N)$ . If  $u(A_i) = u(A_N)$  for some  $i \in \{1, \dots, N - 1\}$ , then the proof is complete; otherwise  $s(A_i) = s(A_N)$  must hold for all  $i \in \{1, \dots, N - 1\}$ , which again implies the claimed property.  $\square$

**Lemma 4.2.** *Let  $\mathbf{A} = (A_1, \dots, A_N) \in GL_2(\mathbb{R})^N$  be such that all the elements of  $\mathbf{A}$  are proximal. The following two statements are equivalent:*

- (1)  $|\lambda_u(AB)| = |\lambda_u(A) \lambda_u(B)|$  for all  $A, B \in \mathcal{S}(\mathbf{A})$ ,
- (2)  $u(A_i) = u(A_j)$  for all  $i, j$  or  $s(A_i) = s(A_j)$  for all  $i, j$ .

*Proof.* It is again easy to see that (2) implies (1). Therefore, we assume that (1) holds. Let us first show that  $\lambda_u(A_i A_j) = \lambda_u(A_i) \lambda_u(A_j)$  or  $\lambda_u(A_i A_j^2) = \lambda_u(A_i) \lambda_u(A_j)^2$  for every  $i \neq j$ . Suppose for a contradiction that there exist  $i \neq j$  such that

$$\lambda_u(A_i A_j) = -\lambda_u(A_i) \lambda_u(A_j) \quad \text{and} \quad \lambda_u(A_i A_j^2) = -\lambda_u(A_i) \lambda_u(A_j)^2.$$

Hence,  $\lambda_u(A_i A_j) \lambda_u(A_j) = -\lambda_u(A_i) \lambda_u(A_j)^2 = \lambda_u(A_i A_j^2)$  and, by Lemma 4.1 applied to the matrix pair  $(A_i A_j, A_j)$ , we have  $u(A_i A_j) = u(A_j)$  or  $s(A_i A_j) = s(A_j)$ . Assuming  $u(A_i A_j) = u(A_j)$ , we have  $-\lambda_u(A_i) \lambda_u(A_j)^2 v(A_j) = A_j A_j^2 v(A_j) = \lambda_u(A_j)^2 A_i v(A_j)$ , where  $v(A_j) \in u(A_j)$  is a unit vector. But this is a contradiction since this would imply that  $\lambda_u(A_i) = 0$  or  $\lambda_s(A_i) = -\lambda_u(A_i)$ . The case  $s(A_i A_j) = s(A_j)$  is similar.

If  $\lambda_u(A_i A_j) = \lambda_u(A_i) \lambda_u(A_j)$ , then (2) follows from Lemma 4.1. Similarly, if  $\lambda_u(A_i A_j^2) = \lambda_u(A_i) \lambda_u(A_j)^2$ , then again by Lemma 4.1,  $u(A_j) = u(A_j^2) = u(A_i)$  or  $s(A_j) = s(A_j^2) = s(A_i)$ . The proof can be finished by induction similarly to the proof of Lemma 4.1.  $\square$

The following lemma is a simple application of [17, Theorem 1.7(ii)–(iii)].

**Lemma 4.3.** *Let  $\mathbf{A} = (A_1, \dots, A_N) \in GL_2(\mathbb{R})^N$  be such that*

$$A_i = \begin{pmatrix} a_i & b_i \\ 0 & c_i \end{pmatrix}$$

*for all  $i \in \{1, \dots, N\}$ , where  $a_i, b_i, c_i \in \mathbb{R}$ , and let  $\mu_a$  and  $\mu_c$  be the Bernoulli measures obtained from the probability vectors  $(\sum_{i=1}^N |a_i|^s)^{-1}(|a_1|^s, \dots, |a_N|^s)$  and  $(\sum_{i=1}^N |c_i|^s)^{-1}(|c_1|^s, \dots, |c_N|^s)$ , respectively. If  $\mu$  is an ergodic equilibrium state for  $\Phi^s$ , then*

$$\mu \in \begin{cases} \{\mu_a\}, & \text{if } \sum_{i=1}^N |a_i|^s > \sum_{i=1}^N |c_i|^s, \\ \{\mu_c\}, & \text{if } \sum_{i=1}^N |a_i|^s < \sum_{i=1}^N |c_i|^s, \\ \{\mu_a, \mu_c\}, & \text{if } \sum_{i=1}^N |a_i|^s = \sum_{i=1}^N |c_i|^s. \end{cases}$$

The following lemma is [17, Proposition 1.2].

**Lemma 4.4.** *If  $\mathbf{A} \in GL_2(\mathbb{R})^N$  is irreducible, then there is unique equilibrium state which is a Gibbs-type measure for  $\Phi^s$ .*

We are now ready to prove the propositions.

*Proof of Proposition 2.6.* Let us first show that (2) implies (1). Lemma 4.4 shows that if  $\mathbf{A}$  is irreducible then the equilibrium state is a Gibbs-type measure for  $\Phi^s$ . Also, if  $\mathbf{A}$  is strongly conformal, the conclusion is straightforward. We may thus assume that  $\mathbf{A}$  is reducible with a common invariant subspace  $V$  and that there exists  $\varepsilon > 0$  such that either the closed  $\varepsilon$ -neighbourhood of  $V$  or the closure of its complement is an invariant unstable multicone. Note that  $\mathcal{S}(\mathbf{A})$  cannot contain any parabolic elements, since in this case the neighbourhood (or its complement) cannot be invariant.

We may, by Proposition 2.2, assume that for some  $M \in \mathbb{N}$  the tuple  $\mathbf{A}_h = (A_1, \dots, A_M)$  has a strongly invariant multicone  $\mathcal{C}$  and  $\mathbf{A}_e = (A_{M+1}, \dots, A_N)$  is such that  $A_i \mathcal{C} = \mathcal{C}$  for all  $i \in \{M+1, \dots, N\}$ . Thus, either  $V \in \mathcal{C}^o$  or  $V \notin \mathcal{C}$ . If  $V \in \mathcal{C}^o$ , then  $u(A_i) = V$  for all  $i \in \{1, \dots, M\}$  and if  $V \notin \mathcal{C}$ , then  $s(A_i) = V$  for all  $i \in \{1, \dots, M\}$ . By the invariance of  $V$  and since  $\mathcal{S}(\mathbf{A})$  does not contain parabolic element, every  $A \in \mathcal{S}(\mathbf{A})$  is diagonalisable. So in the first case, for any  $A_{i_1}, \dots, A_{i_n} \in \mathbf{A}$ ,

$$|\lambda_u(A_{i_1} \cdots A_{i_n})| = \|A_{i_1} \cdots A_{i_n} V\| = \prod_{\ell=1}^n \|A_{i_\ell} V\| = \prod_{\ell=1}^n |\lambda_u(A_{i_\ell})|.$$

In the second case similarly,  $|\lambda_s(A_{i_1} \cdots A_{i_n})| = \prod_{\ell=1}^n |\lambda_s(A_{i_\ell})|$ , but by the multiplicity of the determinant  $|\lambda_u(A_{i_1} \cdots A_{i_n})| = \prod_{\ell=1}^n |\lambda_u(A_{i_\ell})|$ . Moreover, by Lemma 3.10, there exists a constant  $C > 0$  such that for every  $A \in \mathcal{S}(\mathbf{A}) \setminus \mathcal{S}(\mathbf{A}_e)$

$$|\lambda_u(A)| \leq \|A\| \leq C |\lambda_u(A)|,$$

and  $|\lambda_u(A)| = \|A\|$  for  $A \in \mathcal{S}(\mathbf{A}_e)$  trivially. Hence, the Bernoulli measure  $\lambda$  obtained from the probability vector

$$\left( \frac{|\lambda_u(A_1)|^s}{\sum_{i=1}^N |\lambda_u(A_i)|^s}, \dots, \frac{|\lambda_u(A_N)|^s}{\sum_{i=1}^N |\lambda_u(A_i)|^s} \right)$$

is a  $\sigma$ -invariant Gibbs-type measure for  $\Phi^s$ . Therefore,  $\mu = \lambda$ .

Let us then show that (1) implies (2). We may assume without loss of generality that  $\mathbf{A}$  is reducible with common subspace  $V$ . Moreover, let us assume that neither any  $\varepsilon$ -neighbourhood of  $V$  nor the closures of the complements are invariant unstable multicone. Our goal is to show that the only remaining possibility,  $\mathbf{A}$  is strongly conformal, holds.

By reducibility, after a change of basis, every  $A_i \in \mathcal{S}(\mathbf{A})$  has the form

$$A_i = \begin{pmatrix} a_i & b_i \\ 0 & c_i \end{pmatrix},$$

where  $a_i = \prod_{k=1}^{|\mathbf{i}|} a_{i_k}$  and  $c_i = \prod_{k=1}^{|\mathbf{i}|} c_{i_k}$  with some  $a_i, b_i, c_i \in \mathbb{R}$  for  $i \in \{1, \dots, N\}$ . Then, by Lemma 4.3,  $\mu = \mu_a$  or  $\mu = \mu_c$ , where  $\mu_a$  and  $\mu_c$  are defined in the formulation of Lemma 4.3. If one of the matrices, say  $A_i \in \mathcal{S}(\mathbf{A})$ , is parabolic, then  $a_i = c_i$  and  $b_i \neq 0$ . It follows that there exists  $c > 0$  such that  $c^{-1}na_i^{n-1}b_i \leq \|A_i^n\| \leq cna_i^{n-1}b_i$  for all  $n \in \mathbb{N}$ . By [17, Theorem 1.7(ii)], we may assume that  $P(\Phi^s) = \log \sum_{i=1}^N |a_i|^s$  and that  $\mu = \mu_a$ . The definition of  $\mu_a$  thus implies that

$$C^{-1} \frac{|a_i|^{ns}}{n^s |a_i|^{s(n-1)} |b_i|^s} \leq \frac{\mu([i^n])}{\|A_i^n\|^s \exp(-nP(\Phi^s))} \leq C \frac{|a_i|^{ns}}{n^s |a_i|^{s(n-1)} |b_i|^s}$$

for all  $n \in \mathbb{N}$ . This is a contradiction since  $\mu$  was assumed to be a Gibbs-type measure for  $\Phi^s$ . Thus,  $\mathcal{S}(\mathbf{A})$  does not contain any parabolic element.

The common subspace  $V$  and the fact that  $\mathcal{S}(\mathbf{A})$  does not contain parabolic element implies that all the matrices in  $\mathbf{A}$  are diagonalisable. Since neither any  $\varepsilon$ -neighbourhood of  $V$  nor the closures of the complements are invariant unstable multicones, then either  $|a_k| = |c_k|$  and  $b_k = 0$  for every  $k \in \{1, \dots, N\}$  (which implies that  $\mathbf{A}$  is strongly conformal) or there exist  $i \neq j$  such that  $|a_i| < |c_i|$  and  $|a_j| > |c_j|$ . If  $\mu = \mu_a$ , then

$$C^{-1} < \frac{\mu([i^n])}{\|A_i^n\|^s \exp(-nP(\Phi^s))} \leq C' \frac{|a_i|^{sn}}{|c_i|^{sn}}$$

for all  $n \in \mathbb{N}$ , and similarly, if  $\mu = \mu_c$ , then

$$C^{-1} < \frac{\mu([j^n])}{\|A_j^n\|^s \exp(-nP(\Phi^s))} \leq C' \frac{|c_j|^{sn}}{|a_j|^{sn}}$$

for all  $n \in \mathbb{N}$ . Since both inequalities lead to a contradiction, it follows that  $\mathbf{A}$  must be strongly conformal.  $\square$

*Proof of Proposition 2.7.* Let us first show that (2) implies (1). If  $\mathbf{A}$  is reducible, then the statement follows directly from Lemma 4.3. If  $\mathbf{A}$  is strongly conformal, then the statement is straightforward.

Let us then show that (1) implies (2). Let us contrarily assume that  $\mu$  is a Bernoulli measure,  $\mathbf{A}$  is irreducible, and not strongly conformal. By Lemma 4.4,  $\mu$  is a Gibbs-type measure for  $\Phi^s$ , that is, there exists a constant  $C > 0$  such that

$$C^{-1} \leq \frac{\mu([i])}{\|A_i\|^s \exp(-nP(\Phi^s))} \leq C \quad (4.1)$$

for all  $i \in \Sigma_n$  and  $n \in \mathbb{N}$ . Since  $\mu$  is a Bernoulli measure and  $\mathbf{A}$  is not strongly conformal, Theorem 2.1 implies that  $\mathbf{A}$  has an invariant unstable multicone  $\mathcal{C}$  and  $\mathcal{S}(\mathbf{A})$  does not contain any parabolic element. We may, by Proposition 2.2, assume that for some  $M \in \mathbb{N}$  the tuple  $\mathbf{A}_h = (A_1, \dots, A_M)$  has a strongly invariant multicone  $\mathcal{C}$  and  $\mathbf{A}_e = (A_{M+1}, \dots, A_N)$  is strongly conformal with  $A_i \mathcal{C} = \mathcal{C}$  for all  $i \in \{M+1, \dots, N\}$ .

By (4.1) and the Bernoulli property of  $\mu$ ,

$$C^{-1/n} \leq \frac{\mu([i])}{\|A_i^n\|^{s/n} \exp(-|i|P(\Phi^s))} \leq C^{1/n}$$

for all  $i \in \Sigma_*$  and  $n \in \mathbb{N}$ . Thus, by letting  $n \rightarrow \infty$ , we see that

$$|\lambda_u(A_i)| = \mu([i])^{1/s} \exp(|i|P/s).$$

for all  $i \in \Sigma_* \setminus \bigcup_{k \in \mathbb{N}} \{M+1, \dots, N\}^k$ . Since  $\mu$  is a Bernoulli measure, we see that  $|\lambda_u(A_{ij})| = |\lambda_u(A_i)\lambda_u(A_j)|$  for any two  $i, j \in \Sigma_* \setminus \bigcup_{k \in \mathbb{N}} \{M+1, \dots, N\}^k$ . Thus Lemma 4.2 implies that there exists a subspace  $V$  such that  $u(A_i) = V$  for all  $i \in \Sigma_* \setminus \bigcup_{k \in \mathbb{N}} \{M+1, \dots, N\}^k$  or  $s(A_i) = V$  for

all  $\mathbf{i} \in \Sigma_* \setminus \bigcup_{k \in \mathbb{N}} \{M+1, \dots, N\}^k$ . Without loss of generality, we may assume that we are in the first case.

Since  $|\lambda_u(A_{\mathbf{i}}^2 A_j)| = |\lambda_u(A_{\mathbf{i}}) \lambda_u(A_{\mathbf{i}} A_j)|$  and  $|\lambda_u(A_{\mathbf{i}}^3 A_j)| = |\lambda_u(A_{\mathbf{i}})^2 \lambda_u(A_{\mathbf{i}} A_j)|$ , for every  $j \in \{M+1, \dots, N\}$ , we have by Lemma 4.1 that  $u(A_{\mathbf{i}}^k A_j) = u(A_{\mathbf{i}})$ , where  $k = 1$  or  $k = 2$ . Therefore  $A_{\mathbf{i}}^k A_j u(A_{\mathbf{i}}^k A_j) = A_{\mathbf{i}}^k u(A_{\mathbf{i}})$ , which implies that  $A_j V = V$ . Thus,  $V$  is an invariant subspace for  $\mathbf{A}$ . This contradicts the irreducibility assumption.  $\square$

Let us next prove the theorems. For the existence of the function in the statement (4) of Theorem 2.8 we need the following lemma.

**Lemma 4.5.** *Let  $\mathbf{A} \subset GL_2(\mathbb{R})$  be a finite set such that  $\mathbf{A} = \mathbf{A}_h \cup \mathbf{A}_e$ , where  $\mathbf{A}_e$  is strongly conformal and  $\mathbf{A}_h \neq \emptyset$  has a strongly invariant multicone  $\mathcal{C}$  such that  $\mathbf{A}\mathcal{C} = \mathcal{C}$  for all  $\mathbf{A} \in \mathbf{A}_e$ . Let  $m$  be the Haar measure generated by  $\mathbf{A}_e$  normalised on  $\mathcal{C}$ . Then for every  $\mathbf{i} \in \Sigma$  there exists a probability measure  $\nu_{\mathbf{i}}$  on  $\mathcal{C}$  such that*

$$\nu_{\mathbf{i}} = \lim_{n \rightarrow \infty} (A_{\mathbf{i}|_n})_* m.$$

In particular,  $(A_j)_* \nu_{\mathbf{i}} = \nu_{j\mathbf{i}}$ .

*Proof.* Write  $\mathbf{A}_h = \{A_1, \dots, A_M\}$  and  $\mathbf{A}_e = \{A_{M+1}, \dots, A_N\}$ . Let us divide  $\Sigma$  into two disjoint sets

$$\hat{\Sigma} = \{i_1 i_2 \dots \in \Sigma : i_n \in \{1, \dots, M\} \text{ for infinitely many } n \in \mathbb{N}\}, \quad (4.2)$$

$$\Upsilon = \{i_1 i_2 \dots \in \Sigma : \text{there is } n_0 \in \mathbb{N} \text{ such that } i_n \in \{M+1, \dots, N\} \text{ for all } n > n_0\}. \quad (4.3)$$

Fix  $\mathbf{i} \in \hat{\Sigma}$ . By the definition,  $\mathbf{i} \in \hat{\Sigma}$  can be written as  $\mathbf{i} = \mathbf{i}_1 j_1 \mathbf{i}_2 j_2 \dots$ , where

$$\mathbf{i}_k \in \bigcup_{n \in \mathbb{N}} \{M+1, \dots, N\}^n \cup \{\emptyset\}$$

and  $i_k \in \{1, \dots, M\}$  for all  $k \in \mathbb{N}$ . Thus,  $A_{\mathbf{i}_k j_k} \mathcal{C} \subset \mathcal{C}^o$  for every  $k \in \mathbb{N}$ , and there exists a unique  $V = V(\mathbf{i}) \in \mathbb{RP}^1$  such that  $V = \bigcap_{n=0}^{\infty} A_{\mathbf{i}_1 j_1} \dots A_{\mathbf{i}_n j_n}(\mathcal{C})$ .

Let  $g: \mathbb{RP}^1 \rightarrow \mathbb{R}$  be a continuous function. Since  $\mathbb{RP}^1$  is compact, for every  $\varepsilon > 0$  there exists  $r > 0$  such that for every  $V, W \in \mathbb{RP}^1$  with  $d(V, W) < r$ ,  $|g(V) - g(W)| < \varepsilon$ . Thus, by choosing  $n$  sufficiently large so that  $\text{diam}(A_{\mathbf{i}_1 j_1} \dots A_{\mathbf{i}_n j_n}(\mathcal{C})) < r$ , we have

$$\left| \int g(V) d(A_{\mathbf{i}|_n})_* m(V) - g(V(\mathbf{i})) \right| \leq \varepsilon.$$

Hence,  $\lim_{n \rightarrow \infty} (A_{\mathbf{i}|_n})_* m$  exists and equals to  $\delta_{V(\mathbf{i})}$ .

On the other hand, if  $\mathbf{i} \in \Upsilon$ , then clearly  $\lim_{n \rightarrow \infty} (A_{\mathbf{i}|_n})_* m = (A_{\mathbf{i}|_k})_* m$ , where  $k$  is the smallest  $n_0$  satisfying the condition in (4.3).  $\square$

*Proof of Theorem 2.8.* The equivalence of (2) and (3) follows directly from Corollary 2.5. By Lemma 4.4, the equilibrium state  $\mu$  is unique and a Gibbs-type measure for  $\Phi^s$ . Thus, also (1) and (2) can be immediately seen to be equivalent.

Let us show that (4) implies (1). Plugging (4) into (2.3), we see that

$$C^{-1} \exp\left(-nP(\Phi) + s \sum_{k=0}^{n-1} f(\sigma^k \mathbf{i})\right) \leq \mu([\mathbf{i}|_n]) \leq C \exp\left(-nP(\Phi) + s \sum_{k=0}^{n-1} f(\sigma^k \mathbf{i})\right)$$

holds for every  $\mathbf{i} \in \Sigma$ , from which the quasi-Bernoulli property clearly follows.

It remains to show that (3) implies (4). By Lemma 4.5,  $\nu_{\mathbf{i}} = \lim_{n \rightarrow \infty} (A_{\mathbf{i}|_n})_* m$  exists for every  $\mathbf{i} \in \Sigma$ . Define  $f: \Sigma \rightarrow \mathbb{R}$  by setting

$$f(\mathbf{i}) = \int \log \|A_{\mathbf{i}|_1} V\| d\nu_{\sigma \mathbf{i}}(V)$$

for all  $\mathbf{i} \in \Sigma$ . Clearly,

$$\int |f(\mathbf{i})| d\mu(\mathbf{i}) \leq \iint |\log \|A_{\mathbf{i}|_1} V\|| d\nu_{\sigma \mathbf{i}}(V) d\mu(\mathbf{i}) \leq \iint C d\nu_{\sigma \mathbf{i}}(V) d\mu(\mathbf{i}) = C,$$



where  $C = \log \max\{\max_i\{\|A_i\|\}, \max_i\{\|A_i^{-1}\|\}\}$ . Let  $\hat{\Sigma}$  and  $\Upsilon$  be as in (4.2) and (4.3), respectively. Since  $\mu$  is fully supported,  $\mu(\Upsilon) = 0$  and  $\hat{\Sigma}$  has full  $\mu$  measure. Furthermore, every  $\mathbf{i} \in \hat{\Sigma}$  satisfies

$$\text{diam}(A_{\mathbf{i}|_n}(\mathcal{C})) \leq C\tau^{\#\mathbf{i}|_n + \dots + \#\mathbf{M}|\mathbf{i}|_n} \rightarrow 0$$

as  $n \rightarrow \infty$ . Therefore, for  $\mu$ -almost every  $\mathbf{i}$  and for any sequence  $(\mathbf{j}_n)_{n \in \mathbb{N}}$  converging to  $\mathbf{i}$  and sufficiently large  $n$ ,

$$\begin{aligned} |f(\mathbf{i}) - f(\mathbf{j}_n)| &= \left| \int \log \|A_{\mathbf{i}|_1}|V\| d\nu_{\sigma\mathbf{i}}(V) - \int \log \|A_{\mathbf{j}_n|_1}|V\| d\nu_{\sigma\mathbf{j}_n}(V) \right| \\ &= \left| \log \|A_{\mathbf{i}|_1}|V(\sigma\mathbf{i})\| - \int \log \|A_{\mathbf{i}|_1}|V\| d\nu_{\sigma\mathbf{j}_n}(V) \right| \\ &\leq C \text{dist}(\delta_{V(\sigma\mathbf{i})}, \nu_{\sigma\mathbf{j}_n}) \leq C' \text{diam}(A_{\sigma\mathbf{i} \wedge \sigma\mathbf{j}_n}(\mathcal{C})), \end{aligned}$$

which converges to 0 as  $n \rightarrow \infty$ . Note that

$$\sum_{k=0}^{n-1} f(\sigma^k \mathbf{i}) = \int \log \|A_{\mathbf{i}|_n}|V\| d\nu_{\sigma^n \mathbf{i}}(V)$$

for every  $n \in \mathbb{N}$  and  $\mathbf{i} \in \Sigma$ . By Lemma 3.10, there exists  $\kappa > 0$  such that  $\|A_{\mathbf{i}|_n}\| \geq \|A_{\mathbf{i}|_n}|V\| \geq \kappa \|A_{\mathbf{i}|_n}\|$  for all  $V \in \mathcal{C}$ . Therefore, (4) follows.  $\square$

The following lemma, which we refer to as the three matrices lemma, is the key observation in the proof of Theorem 2.9.

**Lemma 4.6.** *If  $\mathbf{A} = (A_1, A_2, A_3) \in GL_2(\mathbb{R})^3$  is such that  $A_3 = cI$  for some  $c \in \mathbb{R} \setminus \{0\}$  and  $(A_1, A_2)$  is irreducible and dominated, then for every Hölder continuous potential  $f: \{1, 2, 3\}^{\mathbb{N}} \rightarrow \mathbb{R}$  and every  $C > 0$  there exists  $\mathbf{i} \in \{1, 2, 3\}^{\mathbb{N}}$  and  $n \in \mathbb{N}$  such that*

$$\left| \sum_{k=0}^{n-1} f(\sigma^k \mathbf{i}) - \log \|A_{\mathbf{i}|_n}\| \right| > C.$$

The proof of the lemma takes several pages. Trying not to disrupt the flow of the proofs in this section, we have postponed it into §5.

*Proof of Theorem 2.9.* By Lemma 4.4, the equilibrium state  $\mu$  is unique and a Gibbs-type measure for  $\Phi^s$ . Taking the potential  $f$  in (3), it is clear that  $\mu$  is Gibbs for the potential  $sf$ . On the other hand, if  $\mu$  is Gibbs for the potential  $g$  then  $\frac{1}{s}g$  clearly satisfies (3).

Let us show that (2) implies (3). If  $\mathbf{A}$  has a strongly invariant multicone  $\mathcal{C}$ , then, by e.g. [3, Lemma 2.4], there exist Hölder-continuous functions  $V: \Sigma \rightarrow \mathbb{R}^1$  and  $f: \Sigma \rightarrow \mathbb{R}$  such that

$$V(\mathbf{i}) = \bigcap_{n=1}^{\infty} A_{\mathbf{i}|_n}(\mathcal{C}) \quad \text{and} \quad f(\mathbf{i}) = \log \|A_{\mathbf{i}|_1}|V(\sigma\mathbf{i})\| \quad (4.4)$$

for all  $\mathbf{i} \in \Sigma$ . Moreover, by Lemma 3.10, there is a constant  $C > 0$  such that

$$\left| \sum_{k=0}^{m-1} f(\sigma^k \mathbf{i}) - \log \|A_{\mathbf{i}|_m}\| \right| \leq C.$$

for all  $\mathbf{i} \in \Sigma$  and  $m \in \mathbb{N}$ . On the other hand, if  $\mathbf{A}$  is strongly conformal then, by choosing  $f(\mathbf{i}) = \frac{1}{2} \log |\det(A_{\mathbf{i}|_1})|$ , the claimed properties follow.

It remains to show that (3) implies (2). Let us assume contrarily that there exist a constant  $C > 0$  and a Hölder-continuous function  $f$  such that

$$\left| \sum_{k=0}^{n-1} f(\sigma^k \mathbf{i}) - \log \|A_{\mathbf{i}|_n}\| \right| \leq C, \quad (4.5)$$

for all  $\mathbf{i} \in \Sigma$ ,  $\mathbf{A}$  does not have strongly invariant multicone, and  $\mathbf{A}$  is not strongly conformal. Thus, by Theorem 2.8,  $\mathbf{A}$  can be decomposed into  $\mathbf{A}_h \neq \emptyset$  and strongly conformal set  $\mathbf{A}_e \neq \emptyset$  such that  $\mathbf{A}_h$

has strongly invariant multicone  $\mathcal{C}$  and  $A\mathcal{C} = \mathcal{C}$  for every  $A \in \mathbf{A}_e$ . As usual, let  $\mathbf{A}_h = \{A_1, \dots, A_M\}$  and  $\mathbf{A}_e = \{A_{M+1}, \dots, A_N\}$ . The equilibrium state  $\mu$  is a quasi-Bernoulli measure. Recall that, by Proposition 2.2,  $\{|\det(A)|^{-1/2}A : A \in \mathcal{S}(\mathbf{A}_e)\}$  is finite. Hence, there exists  $A_j \in \mathcal{S}(\mathbf{A}_e)$  such that  $A_j = cI$ .

Since  $\mathbf{A}_h$  is non-empty and  $\mathbf{A}$  is irreducible,  $X_u(\mathbf{A})$  and  $X_s(\mathbf{A})$  contain at least two points each. Then there exist four proximal matrices  $A_{i_1}, A_{i_2}, A_{i_3}, A_{i_4} \in \mathcal{S}(\mathbf{A})$  such that  $u(\mathbf{A}_{i_1}) \neq u(\mathbf{A}_{i_2})$  and  $s(\mathbf{A}_{i_3}) \neq s(\mathbf{A}_{i_4})$ . Taking  $q > 0$  sufficiently large we have that  $A_{i_1}^q \mathcal{C} \cap A_{i_2}^q \mathcal{C} = \emptyset$  and  $A_{i_3}^{-q}(\mathcal{C}^o)^c \cap A_{i_4}^{-q}(\mathcal{C}^o)^c = \emptyset$ . Clearly,  $u(A_{i_1}^q A_{i_3}^q) \in A_{i_1}^q \mathcal{C}$  and  $u(A_{i_2}^q A_{i_4}^q) \in A_{i_2}^q \mathcal{C}$  and so  $u(A_{i_1}^q A_{i_3}^q) \neq u(A_{i_2}^q A_{i_4}^q)$ . Similarly,  $s(A_{i_1}^q A_{i_3}^q) \in A_{i_3}^{-q}(\mathcal{C}^o)^c$  and  $s(A_{i_2}^q A_{i_4}^q) \in A_{i_4}^{-q}(\mathcal{C}^o)^c$  and so  $s(A_{i_1}^q A_{i_3}^q) \neq s(A_{i_2}^q A_{i_4}^q)$ . Thus,  $(A_{i_1}^q A_{i_3}^q, A_{i_2}^q A_{i_4}^q)$  is dominated and irreducible.

There exist  $n_1, n_2, n_3 \geq 1$  such that  $\ell := n_3|j| = n_1q(|i_1| + |i_3|) = n_2q(|i_2| + |i_4|)$ . Let us define  $\Gamma = \{(i_1^q i_3^q)^{n_1}, (i_2^q i_4^q)^{n_2}, j^{n_3}\}^{\mathbb{N}}$ . By (4.5), the Hölder continuous potential  $h = \sum_{j=0}^{\ell-1} f \circ \sigma^j$  satisfies

$$\left| \sum_{k=0}^{m-1} h(\sigma^k \mathbf{i}) - \log \|A_{\mathbf{i}|_m}\| \right| \leq C$$

for all  $m \in \mathbb{N}$  and  $\mathbf{i} \in \Gamma$ , where  $\sigma$  denotes the left-shift operator on  $\Gamma$ . Since this contradicts Lemma 4.6, we have finished the proof.  $\square$

## 5. THE THREE MATRICES LEMMA

In this section, we prove Lemma 4.6. Throughout the section, we assume that  $\mathbf{A} = (A_1, A_2, A_3) \in GL_2(\mathbb{R})^3$  is such that  $A_3 = cI$  for some  $c \in \mathbb{R} \setminus \{0\}$ , and  $(A_1, A_2)$  is irreducible and has a strongly invariant multicone  $\mathcal{C}$ . Note that there exists a multicone  $\mathcal{C}_0 \subset \mathcal{C}^o$  such that  $A_i \mathcal{C} \subset \mathcal{C}_0$  for  $i = 1, 2$ . Without loss of generality, by multiplying the matrix triple  $\mathbf{A}$  by  $c^{-1}$ , we may assume that  $c = 1$ . This does not affect on the existence of a Hölder continuous potential.

For simplicity, let us denote  $\Sigma = \{1, 2, 3\}^{\mathbb{N}}$  and  $\Gamma = \{1, 2\}^{\mathbb{N}}$ . Let the Borel  $\sigma$ -algebras of  $\Sigma$  and  $\Gamma$  be  $\mathcal{B}_\Sigma$  and  $\mathcal{B}_\Gamma$ , respectively. As in (4.3), let  $\Upsilon = \bigcup_{n=0}^{\infty} \bigcup_{\mathbf{i} \in \Sigma_n} \{\mathbf{i}3^\infty\} \subset \Sigma$  be the countable set of infinite words whose tail consists only 3's, and define  $\hat{\Sigma} = \Sigma \setminus \Upsilon$ . Notice that each  $\mathbf{i} \in \hat{\Sigma}$  can be written in the form  $\mathbf{i} = 3^{k_1} i_1 3^{k_2} i_2 \dots$ , where  $k_i \in \mathbb{N} \cup \{0\}$  and  $i_k \in \{1, 2\}$  for all  $k \in \mathbb{N}$ . Relying on this representation, let us define a function  $\kappa: \hat{\Sigma} \rightarrow \Gamma$  by setting

$$\kappa(3^{k_1} i_1 3^{k_2} i_2 \dots) = i_1 i_2 \dots$$

for all  $\mathbf{i} \in \hat{\Sigma}$ . The definition of  $\kappa$  can be naturally extended to  $\Sigma_*$  by  $\kappa(3^{k_1} i_1 3^{k_2} \dots, \dots, i_n 3^{k_{n+1}}) = (i_1, \dots, i_n)$  and  $\kappa(3^k) = \emptyset$ , where  $k_i \in \mathbb{N} \cup \{0\}$ .

Observe that  $\kappa^{-1}(C)$  is a countable union of cylinder sets in  $\Sigma$  for every cylinder set  $C$  in  $\Gamma$ . Thus  $\kappa: (\Sigma, \mathcal{B}_\Sigma) \rightarrow (\Gamma, \mathcal{B}_\Gamma)$  is measurable. With a slight abuse of notation, we denote both left-shift operators on  $\Sigma$  and  $\Gamma$  by  $\sigma$ . Finally, let us observe that

$$\kappa(\sigma \mathbf{i}) = \begin{cases} \kappa(\mathbf{i}), & \text{if } \mathbf{i}|_1 = 3 \\ \sigma \kappa(\mathbf{i}), & \text{if } \mathbf{i}|_1 \neq 3. \end{cases} \quad (5.1)$$

Let  $\mu_h$  be the unique ergodic Gibbs measure on  $\Gamma$  for the Hölder continuous potential  $h: \Gamma \rightarrow \mathbb{R}$  defined by

$$h(\mathbf{i}) = \log \|A_{\mathbf{i}|_1}\| V(\sigma \mathbf{i}),$$

where  $V(\mathbf{i}) = \bigcap_{n=1}^{\infty} A_{\mathbf{i}|_n}(\mathcal{C})$ . Since

$$\sum_{k=0}^{m-1} h(\sigma^k \mathbf{i}) = \log \|A_{\mathbf{i}|_m}\| V(\sigma^m \mathbf{i}),$$

Lemma 3.10 implies

$$\left| \sum_{k=0}^{m-1} h(\sigma^k \mathbf{i}) - \log \|A_{\mathbf{i}|_m}\| \right| \leq C.$$

for all  $\mathbf{i} \in \Gamma$  and  $m \in \mathbb{N}$ .

Let us assume contrarily that the statement of Lemma 4.6 fails. This means that there is a Hölder continuous potential  $f: \Sigma \rightarrow \mathbb{R}$  and a constant  $C > 0$  such that

$$\left| \sum_{k=0}^{n-1} f(\sigma^k \mathbf{i}) - \log \|A_{\mathbf{i}|_n}\| \right| \leq C \quad (5.2)$$

for all  $n \in \mathbb{N}$  and  $\mathbf{i} \in \Sigma$ . Our goal is to show that in this case the Gibbs measure  $\mu_h$  is a Bernoulli measure. By Proposition 2.7, as the tuple  $(A_1, A_2)$  is irreducible and contains only proximal matrices, this is a contradiction. We will show this after some auxiliary lemmas.

The proof of the following lemma follows easily from the definition of  $\kappa$  and the domination of the tuple  $(A_1, A_2)$ , and we leave it to the reader.

**Lemma 5.1.** *There exists  $C > 0$  such that*

$$\left| \log \|A_{\mathbf{i}|_n}\| - \sum_{k=0}^{n-1-\#\mathbf{3}\mathbf{i}|_n} h(\sigma^k \kappa(\mathbf{i})) \right| \leq C.$$

for all  $\mathbf{i} \in \hat{\Sigma}$  and  $n \in \mathbb{N}$ .

Let  $f$  be the Hölder continuous potential in (5.2) and let  $\mu_f$  be the unique ergodic Gibbs measure for the potential  $f$  on  $\Sigma$ . By the definition of the pressure and (5.2), we have

$$P\left(\left(\sum_{k=0}^{n-1} f \circ \sigma^k\right)_n\right) = P((\mathbf{i} \mapsto \log \|A_{\mathbf{i}|_n}\|)_n) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{\mathbf{i} \in \Sigma_n} \|A_{\mathbf{i}}\|.$$

Let us denote the common quantity by  $Q$ . Then by the definition of Gibbs measures (2.2), there exists a constant  $C > 0$  such that

$$C^{-1} \exp\left(\sum_{k=0}^{n-1} f(\sigma^k \mathbf{i}) - nQ\right) \leq \mu_f([\mathbf{i}|_n]) \leq C \exp\left(\sum_{k=0}^{n-1} f(\sigma^k \mathbf{i}) - nQ\right),$$

for every  $\mathbf{i} \in \Sigma$ . Let us write

$$R = \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{\mathbf{i} \in \Gamma_n} \|A_{\mathbf{i}}\|.$$

By a simple calculation, recalling that  $A_3 = I$ , we see that

$$Q = \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{\mathbf{i} \in \Sigma_n} \|A_{\mathbf{i}}\| = \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{\ell=0}^n \binom{n}{\ell} \sum_{\mathbf{i} \in \Gamma_\ell} \|A_{\mathbf{i}}\|.$$

Since for every  $\varepsilon > 0$  there exists a constant  $K > 0$  such that

$$K^{-1} e^{(R-\varepsilon)\ell} \leq \sum_{\mathbf{i} \in \Gamma_\ell} \|A_{\mathbf{i}}\| \leq K e^{(R+\varepsilon)\ell}$$

for every  $\ell \in \mathbb{N}$ , we see that  $\log(1 + e^{R-\varepsilon}) \leq Q \leq \log(1 + e^{R+\varepsilon})$ . Since  $\varepsilon > 0$  was arbitrary, we get

$$Q = \log(1 + e^R). \quad (5.3)$$

Let us define the Perron-Frobenius operators  $\mathcal{L}_f$  and  $\mathcal{L}_h$  on  $\Sigma$  and on  $\Gamma$ , respectively, for the Hölder-continuous potentials  $f$  and  $h$  as

$$(\mathcal{L}_f(\psi))(\mathbf{i}) = \sum_{i=1}^3 e^{f(i\mathbf{i})} \psi(i\mathbf{i}) \quad \text{and} \quad (\mathcal{L}_h(\phi))(\mathbf{i}) = \sum_{i=1}^2 e^{h(i\mathbf{i})} \phi(i\mathbf{i}).$$

By [12, Theorem 1.7 and the proof of Theorem 1.16], there exist unique functions  $\psi_f: \Sigma \rightarrow \mathbb{R}$  and  $\phi_h: \Gamma \rightarrow \mathbb{R}$  (i.e. eigenfunctions) and unique probability measures  $\nu_f$  on  $\Sigma$  and  $\nu_h$  on  $\Gamma$  (i.e. eigenmeasures) such that

$$\mathcal{L}_f(\psi_f) = e^Q \psi_f, \quad \mathcal{L}_h(\phi_h) = e^R \phi_h, \quad \mathcal{L}_f^* \nu_f = e^Q \nu_f, \quad \mathcal{L}_h^* \nu_h = e^R \nu_h,$$

and  $\int \psi_f d\nu_f = 1 = \int \phi_h d\nu_h$ . Moreover, by [12, Lemmas 1.8 and 1.10], the potentials  $\log \psi_f$  and  $\log \phi_h$  are Hölder continuous. By induction, it is easy to see that for any function  $\varphi$

$$\begin{aligned} (\mathcal{L}_f^n(\varphi))(\mathbf{i}) &= \sum_{j_1} e^{f(j_1 \mathbf{i})} (\mathcal{L}_f^{n-1}(\varphi))(j_1 \mathbf{i}) = \sum_{j_1, j_2} e^{f(j_1 \mathbf{i})} e^{f(j_2 j_1 \mathbf{i})} (\mathcal{L}_f^{n-2}(\varphi))(j_2 j_1 \mathbf{i}) \\ &= \sum_{j_1, \dots, j_n} e^{f(j_1 \mathbf{i}) + \dots + f(j_n \dots j_1 \mathbf{i})} \varphi(j_n \dots j_1 \mathbf{i}) = \sum_{\mathbf{k} \in \Sigma_n} e^{\sum_{k=0}^{n-1} f(\sigma^k \mathbf{k} \mathbf{i})} \varphi(\mathbf{k} \mathbf{i}). \end{aligned}$$

By [12, Proposition 1.14] and the uniqueness of the ergodic Gibbs measure,

$$\mu_f(B) = \int_B \psi_f d\nu_f \quad \text{and} \quad \mu_h(B') = \int_{B'} \phi_h d\nu_h \quad (5.4)$$

for every  $B \in \mathcal{B}_\Sigma$  and  $B' \in \mathcal{B}_\Gamma$ . Thus, for any  $\mathbf{j} \in \Sigma_n$  and every  $B \in \mathcal{B}_\Sigma$

$$\begin{aligned} \mu_f([\mathbf{j}] \cap \sigma^{-|\mathbf{j}|}(B)) &= \int \psi_f(\mathbf{i}) \mathbf{1}_{[\mathbf{j}] \cap \sigma^{-|\mathbf{j}|}(B)}(\mathbf{i}) d\nu_f(\mathbf{i}) \\ &= \int \psi_f(\mathbf{i}) \mathbf{1}_{[\mathbf{j}] \cap \sigma^{-|\mathbf{j}|}(B)}(\mathbf{i}) e^{-nQ} d(\mathcal{L}_f^*)^n(\nu_f)(\mathbf{i}) \\ &= \int \mathcal{L}_f^n(\psi_f \mathbf{1}_{[\mathbf{j}] \cap \sigma^{-|\mathbf{j}|}(B)})(\mathbf{i}) e^{-nQ} d\nu_f(\mathbf{i}) \\ &= \int \sum_{\mathbf{k} \in \Sigma_n} \exp\left(\sum_{k=0}^{n-1} f(\sigma^k \mathbf{k} \mathbf{i}) - nQ\right) \psi_f(\mathbf{k} \mathbf{i}) \mathbf{1}_{[\mathbf{j}] \cap \sigma^{-|\mathbf{j}|}(B)}(\mathbf{k} \mathbf{i}) d\nu_f(\mathbf{i}) \quad (5.5) \\ &= \int_B \exp\left(\sum_{k=0}^{n-1} f(\sigma^k \mathbf{j} \mathbf{i}) - nQ\right) \psi_f(\mathbf{j} \mathbf{i}) d\nu_f(\mathbf{i}) \\ &= \int_B \exp\left(\sum_{k=0}^{n-1} f(\sigma^k \mathbf{j} \mathbf{i}) - nQ\right) \frac{\psi_f(\mathbf{j} \mathbf{i})}{\psi_f(\mathbf{i})} d\mu_f(\mathbf{i}). \end{aligned}$$

Let  $\hat{f}(\mathbf{i}) = f(\mathbf{i}) + \log \psi_f(\mathbf{i}) - \log \psi_f(\sigma \mathbf{i})$ . Since  $\log \psi_f$  is Hölder continuous and thus, uniformly bounded over  $\Sigma$ , there exists  $C > 0$  such that

$$\left| \sum_{k=0}^{n-1} \hat{f}(\sigma^k \mathbf{i}) - \log \|A_{\mathbf{i}|_n}\| \right| \leq C$$

for all  $n \in \mathbb{N}$  and  $\mathbf{i} \in \Sigma_n$ . By (5.5),

$$\mu_f([\mathbf{i}] \cap \sigma^{-|\mathbf{i}|}(B)) = \int_B \exp\left(\sum_{k=0}^{|\mathbf{i}|-1} \hat{f}(\sigma^k \mathbf{j} \mathbf{i}) - |\mathbf{i}|Q\right) d\mu_f(\mathbf{i}). \quad (5.6)$$

for all  $\mathbf{i} \in \Sigma_*$  and  $B \in \mathcal{B}_\Sigma$ .

Let us denote the ratio  $(1 + e^R)^{-1}$  by  $q$ . Define

$$\eta([\mathbf{i}]) = q^{\#_3 \mathbf{i}} (1 - q)^{|\mathbf{i}| - \#_3 \mathbf{i}} \mu_h([\kappa(\mathbf{i})])$$

for all  $\mathbf{i} \in \Sigma_*$  and notice that

$$\begin{aligned} \sum_{i=1}^3 \eta([\mathbf{i}i]) &= \sum_{i=1}^2 q^{\#_3 \mathbf{i}} (1 - q)^{|\mathbf{i}|+1 - \#_3 \mathbf{i}} \mu_h([\kappa(\mathbf{i}i)]) + q^{\#_3 \mathbf{i}+1} (1 - q)^{|\mathbf{i}| - \#_3 \mathbf{i}} \mu_h([\kappa(\mathbf{i})]) \\ &= q^{\#_3 \mathbf{i}} (1 - q)^{|\mathbf{i}|+1 - \#_3 \mathbf{i}} \sum_{i=1}^2 \mu_h([\kappa(\mathbf{i})i]) + q^{\#_3 \mathbf{i}+1} (1 - q)^{|\mathbf{i}| - \#_3 \mathbf{i}} \mu_h([\kappa(\mathbf{i})]) \quad (5.7) \\ &= q^{\#_3 \mathbf{i}} (1 - q)^{|\mathbf{i}| - \#_3 \mathbf{i}} \mu_h([\kappa(\mathbf{i})]) (1 - q + q) = \eta([\mathbf{i}]). \end{aligned}$$

for all  $\mathbf{i} \in \Sigma_*$ . Thus, by Kolmogorov's extension theorem,  $\eta$  can be extended to a probability measure on  $(\Sigma, \mathcal{B}_\Sigma)$ . We shall denote the extension by  $\eta$  too. The following lemma shows that  $\eta$  is ergodic.

**Lemma 5.2.** *The measure  $\eta$  is  $\sigma$ -invariant and mixing on  $\Sigma$ .*

*Proof.* Since  $\mu_h$  is  $\sigma$ -invariant, the proof of  $\sigma$ -invariance of  $\eta$  is similar to (5.7), and therefore, we omit it. To prove that  $\eta$  is mixing, it is sufficient to show that

$$\lim_{n \rightarrow \infty} \eta([\mathbf{i}] \cap \sigma^{-n}[\mathbf{j}]) = \eta([\mathbf{i}])\eta([\mathbf{j}]).$$

for all  $\mathbf{i}, \mathbf{j} \in \Sigma_*$ . Let  $n > |\mathbf{i}|$  and observe that

$$\begin{aligned} \eta([\mathbf{i}] \cap \sigma^{-n}[\mathbf{j}]) &= \sum_{\mathbf{h} \in \Sigma_{n-|\mathbf{i}|}} \eta([\mathbf{i}\mathbf{h}\mathbf{j}]) = \sum_{\mathbf{h} \in \Sigma_{n-|\mathbf{i}|}} q^{\#_3\mathbf{i} + \#_3\mathbf{j} + \#_3\mathbf{h}} (1-q)^{|\mathbf{i}| - \#_3\mathbf{i} + |\mathbf{j}| - \#_3\mathbf{j} + |\mathbf{h}| - \#_3\mathbf{h}} \mu_h([\kappa(\mathbf{i}\mathbf{h}\mathbf{j})]) \\ &= q^{\#_3\mathbf{i} + \#_3\mathbf{j}} (1-q)^{|\mathbf{i}| - \#_3\mathbf{i} + |\mathbf{j}| - \#_3\mathbf{j}} \sum_{\mathbf{h} \in \Sigma_{n-|\mathbf{i}|}} q^{\#_3\mathbf{h}} (1-q)^{|\mathbf{h}| - \#_3\mathbf{h}} \mu_h([\kappa(\mathbf{i})\kappa(\mathbf{h})\kappa(\mathbf{j})]) \\ &= q^{\#_3\mathbf{i} + \#_3\mathbf{j}} (1-q)^{|\mathbf{i}| - \#_3\mathbf{i} + |\mathbf{j}| - \#_3\mathbf{j}} \sum_{\ell=0}^{n-|\mathbf{i}|} \binom{n-|\mathbf{i}|}{\ell} q^{n-|\mathbf{i}|-\ell} (1-q)^\ell \sum_{\mathbf{k} \in \Gamma_\ell} \mu_h([\kappa(\mathbf{i})\mathbf{k}\kappa(\mathbf{j})]). \end{aligned}$$

Hence,

$$\frac{\eta([\mathbf{i}] \cap \sigma^{-n}[\mathbf{j}])}{\eta([\mathbf{i}])\eta([\mathbf{j}])} = \sum_{\ell=0}^{n-|\mathbf{i}|} \binom{n-|\mathbf{i}|}{\ell} q^{n-|\mathbf{i}|-\ell} (1-q)^\ell \frac{\mu_h([\kappa(\mathbf{i}) \cap \sigma^{-\ell-|\kappa(\mathbf{i})|}\kappa(\mathbf{j})])}{\mu_h([\kappa(\mathbf{i})])\mu_h([\kappa(\mathbf{j})])}. \quad (5.8)$$

By [12, Proposition 1.14], the measure  $\mu_h$  is mixing. Thus, for every  $\varepsilon > 0$  there exists  $N$  such that if  $\ell \geq N$ , then

$$e^{-\varepsilon} \leq \frac{\mu_h([\kappa(\mathbf{i})] \cap \sigma^{-\ell-|\kappa(\mathbf{i})|}\kappa(\mathbf{j}))}{\mu_h([\kappa(\mathbf{i})])\mu_h([\kappa(\mathbf{j})])} \leq e^\varepsilon.$$

Hence, for  $n > N + |\mathbf{i}|$  we get

$$\begin{aligned} &\sum_{\ell=0}^{n-|\mathbf{i}|} \binom{n-|\mathbf{i}|}{\ell} q^{n-|\mathbf{i}|-\ell} (1-q)^\ell \frac{\mu_h([\kappa(\mathbf{i}) \cap \sigma^{-\ell-|\kappa(\mathbf{i})|}\kappa(\mathbf{j})])}{\mu_h([\kappa(\mathbf{i})])\mu_h([\kappa(\mathbf{j})])} \\ &\leq e^\varepsilon \sum_{\ell=N}^{n-|\mathbf{i}|} \binom{n-|\mathbf{i}|}{\ell} q^{n-|\mathbf{i}|-\ell} (1-q)^\ell + \mu_h([\kappa(\mathbf{j})])^{-1} \sum_{\ell=0}^{N-1} \binom{n-|\mathbf{i}|}{\ell} q^{n-|\mathbf{i}|-\ell} (1-q)^\ell \\ &\leq e^\varepsilon + \mu_h([\kappa(\mathbf{j})])^{-1} N(n-|\mathbf{i}|)^N (1-q)^{n-|\mathbf{i}|-N}, \end{aligned}$$

where in the last inequality we used  $\binom{n}{k} \leq n^k$ . By a similar argument,

$$\begin{aligned} &\sum_{\ell=0}^{n-|\mathbf{i}|} \binom{n-|\mathbf{i}|}{\ell} q^{n-|\mathbf{i}|-\ell} (1-q)^\ell \frac{\mu_h([\kappa(\mathbf{i}) \cap \sigma^{-\ell-|\kappa(\mathbf{i})|}\kappa(\mathbf{j})])}{\mu_h([\kappa(\mathbf{i})])\mu_h([\kappa(\mathbf{j})])} \\ &\geq e^{-\varepsilon} \sum_{\ell=N}^{n-|\mathbf{i}|} \binom{n-|\mathbf{i}|}{\ell} q^{n-|\mathbf{i}|-\ell} (1-q)^\ell \geq e^{-\varepsilon} - N(n-|\mathbf{i}|)^N (1-q)^{n-|\mathbf{i}|-N}. \end{aligned}$$

By (5.8) and letting  $n \rightarrow \infty$ , we see that for every  $\varepsilon > 0$

$$e^{-\varepsilon} \leq \lim_{n \rightarrow \infty} \frac{\eta([\mathbf{i}] \cap \sigma^{-n}[\mathbf{j}])}{\eta([\mathbf{i}])\eta([\mathbf{j}])} \leq e^\varepsilon.$$

Since  $\varepsilon > 0$  was arbitrary, the definition of  $\eta$  finishes the proof.  $\square$

**Proposition 5.3.** *If  $\eta$  and  $\mu_f$  are as above, then  $\eta = \mu_f$ .*

*Proof.* Since both  $\eta$  and  $\mu_f$  are ergodic measures, it suffices to show that they are equivalent. By Lemma 5.1 and our assumption (5.2) on  $f$ ,

$$\begin{aligned} \eta([i]) &= q^{\#_3 i} (1-q)^{|i|-\#_3 i} \mu_h([\kappa(i)]) \\ &\leq C q^{\#_3 i} (1-q)^{|i|-\#_3 i} \exp\left(\sum_{k=0}^{|\kappa(i)|-1} h(\sigma^k \kappa(i)j) - |\kappa(i)|R\right) \\ &\leq C' (1-q)^{|i|-\#_3 i} \|A_i\| \exp(-(|i| - \#_3 i)R - \#_3 i \log(1+e^R)) \\ &= C' \|A_i\| \exp(-|i| \log(1+e^R)) \\ &\leq C'' \exp\left(\sum_{k=0}^{|i|-1} f(\sigma^k i j') - |i| \log(1+e^R)\right) \leq C''' \mu_f([i]) \end{aligned}$$

for all  $i \in \Sigma_*$ . The other inequality follows by a similar argument. Since for every cylinder set  $[i]$ , the ratio  $\eta([i])/\mu_f([i])$  is bounded away from 0 and  $\infty$  uniformly, the statement follows.  $\square$

By (5.6) and (5.3),

$$\mu_f([i] \cap \sigma^{-1}(B)) = \int_B \exp(\hat{f}(i i) - Q) d\mu_f(i) = \frac{1}{1+e^R} \int_B \exp(\hat{f}(i i)) d\mu_f(i)$$

for all  $i \in \{1, 2, 3\}$  and  $B \in \mathcal{B}_\Sigma$ . By Proposition 5.3 and recalling the definition of  $\eta$ , we have

$$\mu_f([3] \cap \sigma^{-1}(B)) = \eta([3] \cap \sigma^{-1}(B)) = \frac{1}{1+e^R} \eta(B) = \frac{1}{1+e^R} \mu_f(B).$$

Since  $\hat{f}$  is Hölder continuous and the above two equations hold for every  $B \in \mathcal{B}_\Sigma$ , we conclude that

$$\hat{f}(i) = 0 \tag{5.9}$$

for all  $i \in [3]$ . By (5.6), we have

$$\begin{aligned} \mu_f([i_1 3^{k_1} \dots i_n 3^{k_n}]) &= \int \exp\left(\sum_{\ell=0}^{k_1+\dots+k_n+n-1} \hat{f}(\sigma^\ell i_1 3^{k_1} \dots i_n 3^{k_n} j) - (k_1+\dots+k_n+n)Q\right) d\mu_f(j) \\ &= q^{k_1+\dots+k_n} \int \exp(\hat{f}(i_1 3^{k_1} \dots j) + \dots + \hat{f}(i_n 3^{k_n} j) - nQ) d\mu_f(i) \end{aligned}$$

for every  $k_1, \dots, k_n \in \mathbb{N}$ ,  $i_1, \dots, i_n \in \{1, 2\}$ , and  $n \in \mathbb{N}$ . Since  $\hat{f}$  is Hölder continuous, we have

$$\lim_{k_1, \dots, k_n \rightarrow \infty} \hat{f}(i_\ell 3^{k_\ell} \dots i_n 3^{k_n} j) = \hat{f}(i_\ell 3^\infty)$$

uniformly for all  $\ell \in \{1, \dots, n\}$  and  $j \in \Sigma$ . Hence by the dominated convergence theorem

$$\lim_{k_1, \dots, k_n \rightarrow \infty} \frac{\mu_f([i_1 3^{k_1} \dots i_n 3^{k_n}])}{q^{k_1+\dots+k_n}} = \prod_{\ell=1}^n e^{\hat{f}(i_\ell 3^\infty) - Q}.$$

On the other hand, by the definition of  $\eta$  and Proposition 5.3,

$$\frac{\mu_f([i_1 3^{k_1} \dots i_n 3^{k_n}])}{q^{k_1+\dots+k_n}} = (1-q)^n \mu_h([i_1 \dots i_n]).$$

It follows that

$$\mu_h([i_1 \dots i_n]) = \prod_{\ell=1}^n e^{\hat{f}(i_\ell 3^\infty) - R}$$

and hence,  $\mu_h$  is a Bernoulli measure. This contradicts Proposition 2.7 and finishes the proof of Lemma 4.6.

*Acknowledgement.* Balázs Bárány acknowledges support from the grants NKFI PD123970, OTKA K123782, and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. Antti Käenmäki was supported by the Finnish Center of Excellence in Analysis and Dynamics Research. Ian Morris was supported by the Leverhulme Trust (Research Project Grant number RPG-2016-194). All the authors were partially supported by the ERC grant 306494. The research was started in the Hebrew University of Jerusalem. The authors thank the HUJI, and especially Professor Michael Hochman, for warm hospitality. Bárány and Käenmäki also thank the Institut Mittag-Leffler, where the paper was finished. Finally, the authors thank the anonymous referee for the careful reading and valuable comments which improved the paper.

## REFERENCES

- [1] A. Avila, J. Bochi, and J.-C. Yoccoz. Uniformly hyperbolic finite-valued  $SL(2, \mathbb{R})$ -cocycles. *Comment. Math. Helv.*, 85(4):813–884, 2010.
- [2] B. Bárány, A. Käenmäki, and H. Koivusalo. Dimension of self-affine sets for fixed translation vectors. *J. Lond. Math. Soc. (2)*, 98(1):223–252, 2018.
- [3] B. Bárány and M. Rams. Dimension maximizing measures for self-affine systems. *Trans. Amer. Math. Soc.*, 370(1):553–576, 2018.
- [4] L. Barreira. Nonadditive thermodynamic formalism: equilibrium and Gibbs measures. *Discrete Contin. Dyn. Syst.*, 16(2):279–305, 2006.
- [5] L. Barreira. Almost additive thermodynamic formalism: some recent developments. *Rev. Math. Phys.*, 22(10):1147–1179, 2010.
- [6] L. Barreira and P. Doutor. Almost additive multifractal analysis. *J. Math. Pures Appl. (9)*, 92(1):1–17, 2009.
- [7] J. Bochi and N. Gourmelon. Some characterizations of domination. *Math. Z.*, 263(1):221–231, 2009.
- [8] J. Bochi and I. D. Morris. Continuity properties of the lower spectral radius. *Proc. Lond. Math. Soc. (3)*, 110(2):477–509, 2015.
- [9] J. Bochi and M. Viana. The Lyapunov exponents of generic volume-preserving and symplectic maps. *Ann. of Math. (2)*, 161(3):1423–1485, 2005.
- [10] T. Bomfim and P. Varandas. Multifractal analysis of the irregular set for almost-additive sequences via large deviations. *Nonlinearity*, 28(10):3563–3585, 2015.
- [11] C. Bonatti, L. J. Díaz, and E. R. Pujals. A  $C^1$ -generic dichotomy for diffeomorphisms: weak forms of hyperbolicity or infinitely many sinks or sources. *Ann. of Math. (2)*, 158(2):355–418, 2003.
- [12] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, revised edition, 2008. With a preface by David Ruelle, Edited by Jean-René Chazottes.
- [13] Y. Cao. Dimension spectrum of asymptotically additive potentials for  $C^1$  average conformal repellers. *Nonlinearity*, 26(9):2441–2468, 2013.
- [14] Y. Cao, D.-J. Feng, and W. Huang. The thermodynamic formalism for sub-additive potentials. *Discrete Contin. Dyn. Syst.*, 20(3):639–657, 2008.
- [15] D.-J. Feng. Lyapunov exponents for products of matrices and multifractal analysis. II. General matrices. *Israel J. Math.*, 170:355–394, 2009.
- [16] D.-J. Feng and W. Huang. Lyapunov spectrum of asymptotically sub-additive potentials. *Comm. Math. Phys.*, 297(1):1–43, 2010.
- [17] D.-J. Feng and A. Käenmäki. Equilibrium states of the pressure function for products of matrices. *Discrete Contin. Dyn. Syst.*, 30(3):699–708, 2011.
- [18] J. Fraser, T. Jordan, and N. Jurga. Dimensions of equilibrium measures on a class of planar self-affine sets. *J. Fractal Geom.* To appear, available at arXiv:1706.06833, 2017.
- [19] H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960.
- [20] I. Hueter and S. P. Lalley. Falconer’s formula for the Hausdorff dimension of a self-affine set in  $\mathbb{R}^2$ . *Ergodic Theory Dynam. Systems*, 15(1):77–97, 1995.
- [21] G. Iommi and Y. Yayama. Almost-additive thermodynamic formalism for countable Markov shifts. *Nonlinearity*, 25(1):165–191, 2012.
- [22] G. Iommi and Y. Yayama. Weak Gibbs measures as Gibbs measures for asymptotically additive sequences. *Proc. Amer. Math. Soc.*, 145(4):1599–1614, 2017.
- [23] R. M. Jungers. On asymptotic properties of matrix semigroups with an invariant cone. *Linear Algebra Appl.*, 437(5):1205–1214, 2012.
- [24] A. Käenmäki. On natural invariant measures on generalised iterated function systems. *Ann. Acad. Sci. Fenn. Math.*, 29(2):419–458, 2004.
- [25] A. Käenmäki and B. Li. Genericity of dimension drop on self-affine sets. *Statist. Probab. Lett.*, 126:18–25, 2017.

- [26] A. Käenmäki and H. W. J. Reeve. Multifractal analysis of Birkhoff averages for typical infinitely generated self-affine sets. *J. Fractal Geom.*, 1(1):83–152, 2014.
- [27] A. Käenmäki and M. Vilppolainen. Dimension and measures on sub-self-affine sets. *Monatsh. Math.*, 161(3):271–293, 2010.
- [28] R. Mañé. Contributions to the stability conjecture. *Topology*, 17(4):383–396, 1978.
- [29] R. Mañé. Oseledec’s theorem from the generic viewpoint. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983)*, pages 1269–1276. PWN, Warsaw, 1984.
- [30] I. D. Morris. Ergodic properties of matrix equilibrium states. *Ergodic Theory Dynam. Systems*, 38(6):2295–2320, 2018.
- [31] M. Pollicott. Maximal Lyapunov exponents for random matrix products. *Invent. Math.*, 181(1):209–226, 2010.
- [32] V. Yu. Protasov and A. S. Voynov., Matrix semigroups with constant spectral radius. *Linear Algebra Appl.*, 513:376–408, 2017.
- [33] Y. Yayama. On factors of Gibbs measures for almost additive potentials. *Ergodic Theory Dynam. Systems*, 36(1):276–309, 2016.

(Balázs Bárány) BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS, MTA-BME STOCHASTICS RESEARCH GROUP, DEPARTMENT OF STOCHASTICS, P.O. BOX 91, 1521 BUDAPEST, HUNGARY  
*E-mail address:* `balubsheep@gmail.com`

(Antti Käenmäki) DEPARTMENT OF PHYSICS AND MATHEMATICS, UNIVERSITY OF EASTERN FINLAND, P.O. BOX 111, FI-80101 JOENSUU, FINLAND  
*E-mail address:* `antti.kaenmaki@uef.fi`

(Ian D. Morris) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SURREY, GUILDFORD GU2 7XH, UNITED KINGDOM  
*E-mail address:* `i.morris@surrey.ac.uk`