Continuity of subadditive pressure for matrix cocycles and the dimension of a self-affine set

Master's Thesis
Jussi Kormu
Department of Mathematical Sciences
University of Oulu
Spring 2016

Contents

Introduction

This master's thesis is based on the article *Non-conformal repellers and continuity of pressure for matrix cocycles* by De-Jun Feng and Pablo Shmerkin [12]. In the article Feng and Shmerkin prove that subadditive pressure

$$P(\mathbf{A}, s) = \lim_{n\to\infty} \frac{1}{n} \log \left(\sum_{(i_1,\ldots,i_n)\in\{1,\ldots,k\}^n} \varphi^s(A_{i_1}\cdots A_{i_n}) \right),$$

where φ^s is the singular value function, is a continuous function of the linear maps $\mathbf{A} = (A_1, \ldots, A_k)$. This, together with Kenneth Falconer's dimension formula 2.4 ([9, Theorem 5.3.]), implies that in some "typical" situation the dimension of a self-affine set varies continuously with the defining linear maps $\mathbf{A}$. As a matter of fact, Feng and Shmerkin prove a more general continuity result considering a pressure $P_g(A, s)$ for matrix cocycles. Continuity of $P(\cdot,\cdot)$ is then proved as a corollary.

In this thesis I prove the continuity results of Feng and Shmerkin following the arguments of the original article. The proofs are given in greater detail than the original ones. Also more background results and definitions are provided, so less preliminary knowledge is assumed on the reader's part.

The proofs of the continuity theorems make heavy use of ergodic theory and subadditive thermodynamic formalism. Although the necessary tools on these areas are covered, there is no attempt to give any comprehensive introduction to these subjects.

The necessary preliminaries considering ergodic theory, symbolic dynamics, singular values and exterior algebra are given in Section 1. Section 2 introduces the pressure functions and their connection to the dimension of a self-affine set. Also the statements of the continuity theorems that are to be proved, are given.

Matrix cocycles satisfying the cone condition play an important part in the proofs of the continuity theorems. The cone condition and related lemmas are covered in Section 3. In Section 4 a large subsystem satisfying the cone condition is constructed for a locally constant matrix cocycle by utilising the semi-invertible version of the multiplicative ergodic theorem.

Section 5 introduces the measure theoretic entropy and topological pressure for a subadditive potential. These two key concepts of the subadditive thermodynamic formalism are then linked together by the variational principle. Finally, in Section 6 the proofs of the continuity results are constructed with the help of the tools presented in the previous sections.

1 Preliminaries

In this section we cover some basic definitions of ergodic theory alongside with a few core results, the most important of them being Kingman's subadditive ergodic theorem, which we use repeatedly. The dynamical systems that we work on are symbolic systems, which are introduced in Subsection 1.2. In Subsection 1.3 we recall some linear algebraic facts mostly considering singular values. In the last subsection we introduce exterior algebras.

Firstly, we recall an important measure theoretic result, Egorov's theorem.

Theorem 1.1 (Egorov's theorem)**.** *Let $(X, \mathcal{B}, \mu)$ be a measure space and let Y be a separable metric space. Suppose that $f_n : A \to Y$ is a sequence of measurable functions on a subset $A \subset X$, with finite measure, such that f_n converges to f for μ-almost every $x \in A$. Then for every $\varepsilon > 0$ there exists a measurable set B such that $\mu(B) < \varepsilon$ and f_n converges uniformly to f on $A \setminus B$.*

Proof. See [6, Theorem 7.1.12, p. 72] which assumes the knowledge of the proof [5, Theorem 2.2.1, p. 110]. □

1.1 Ergodic theory

Ergodic theory studies measure preserving dynamical systems.

Definition 1.2. A quartet $(X, \mathcal{B}, \mu, T)$ is called a *measure preserving transformation* if $(X, \mathcal{B}, \mu)$ is a measure space and $T\colon X \to X$ is a measurable transformation such that $\mu(T^{-1}(E)) = \mu(E)$ for every $E \in \mathcal{B}$. Measure μ is referred as a *T-invariant measure.*

If $(X, \mathcal{B}, \mu)$ is a probability space, i.e. a measure space with $\mu(X) = 1$, then $(X, \mathcal{B}, \mu, T)$ is called a *probability preserving transformation* (ppt).

Definition 1.3. A probability preserving transformation $(X, \mathcal{B}, \mu, T)$ is called *ergodic* if for any measurable set $E \in \mathcal{B}$ such that $T^{-1}(E) = E$ either $\mu(E) = 0$ or $\mu(E) = 1$. The measure μ and the map T are called an *ergodic measure* and an *ergodic transformation*, respectively.

For every continuous transformation on a compact metric space, there exists an ergodic measure.

Proposition 1.4. *Let X be a compact metric space with the Borel sigma-algebra $\mathcal{B}$. Given any continuous transformation $T\colon X \to X$, there exists at least one ergodic T-invariant probability measure μ.*

Proof. [20, Proposition 9.5, p. 93] □

We use notation $T^k = T \circ \cdots \circ T$ to denote the k fold composition of T for an integer $k \geq 0$. For $E \subset X$, $T^{-k}(E)$ denotes the preimage $(T^k)^{-1}(E)$.

Proposition 1.5. *Let $(X, \mathcal{B}, \mu, T)$ be a ppt. Following statements are equivalent:*

(1) μ is ergodic

(2) Every measurable $L^1(X, \mu)$-function $f \colon X \to \mathbb{R}$ such that $f \circ T = f$ μ-almost everywhere, is almost everywhere a constant.

(3) There exists a family $\mathcal{A} \subset \mathcal{B}$ such that $\mathcal{A}$ generates $\mathcal{B}$, $A \cap B \in \mathcal{A}$ for every $A, B \in \mathcal{A}$ and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B) \tag{1.1}$$

for any $A, B \in \mathcal{A}$.

(4) Equality (1.1) *holds for every $A, B \in \mathcal{B}$.*

Proof. Claims (1) and (2) are equivalent by [4, Propositions 2.10 and 2.11, pp. 44–45]. Equivalence of claims (1), (3) and (4) is proved in [17, Lemma 2.2.2, p. 30]

□

Theorem 1.6 (Birkhoff's ergodic theorem)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a ppt. If $f \in L^1(X, \mu)$, then the limit*

$$\bar{f}(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x))$$

exists for μ-almost every $x \in X$, and in $L^1(X, \mu)$. For μ-almost everywhere, $\bar{f} \circ T = \bar{f}$. Moreover,

$$\int \bar{f} d\mu = \int f d\mu.$$

If in addition T is ergodic, then

$$\bar{f}(x) = \int f d\mu$$

for μ-almost every $x \in X$.

Proof. See [17, Theorem 2.1.5, p. 23] for the proof of all but the last claim. The last claim is then a direct corollary of Proposition 1.5. □

Theorem 1.7 (Kingman's subadditive ergodic theorem). *Let $(X, \mathcal{B}, \mu, T)$ be a ppt. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions such that $f_{n+m}(x) \leq f_n(x) + f_m(T^n(x))$ and f_1^+ is integrable, where $f^+(x) = \max\{f(x), 0\}$. Then f_n/n converges almost everywhere to a function $\bar{f}\colon X \to [-\infty, \infty)$ such that*

(1) $\bar{f}^+$ is integrable

(2) $\bar{f} \circ T = \bar{f}$ for μ-almost every $x \in X$

(3) $\int \bar{f} d\mu = \lim_{n\to\infty} \frac{1}{n} \int f_n d\mu = \inf_{n\in\mathbb{N}} \frac{1}{n} \int f_n d\mu$.

If in addition T is ergodic, then $\bar{f}$ is constant μ-almost everywhere and (3) can be rewritten as

$$\bar{f}(x) = \int \bar{f} d\mu = \lim_{n\to\infty} \frac{1}{n} \int f_n d\mu$$

for μ-almost every $x \in X$.

Proof. For the proof of (1), (2) and (3) see [3]. Note that even though the claim (2) is not included in the statement of the theorem in [3], it is proved during the proofs of other two claims. The ergodic case is then direct corollary of Proposition 1.5. □

Theorem 1.8 (Khintchine's recurrence theorem). *Let $(X, \mathcal{B}, \mu, T)$ be a ppt. For any $\varepsilon > 0$ and any measurable set $E \in \mathcal{B}$, the set*

$$F_\varepsilon = \{k \in \mathbb{Z}\colon \mu(T^{-k}(E) \cap E) \geq \mu(E)^2 - \varepsilon\}$$

has bounded gaps, i.e. there is $K \in \mathbb{N}$ such that $F_\varepsilon \cap \{j, j+1, \dots, j+K-1\} \neq \emptyset$, for every $j \in \mathbb{Z}$.

Proof. See [19, Theorem 3.3, p. 37] □

1.2 Symbolic dynamics

As a reference for the claims of this subsection see [7, Chapter 3] and [4, Chapter 3.5]

Definition 1.9. For $m > 1$, we refer to the set $\mathcal{A}_m = \{1, \ldots, m\}$ as an *alphabet* and its elements as *symbols.* Let $\Sigma_m^+ := \mathcal{A}_m^{\mathbb{N}}$ be the set of all one-sided infinite sequences of symbols of $\mathcal{A}_m$ and let $\Sigma_m := \mathcal{A}_m^{\mathbb{Z}}$ be the set of all two-sided infinite sequences. A finite sequence of symbols is called a *word* and an infinite sequence $x = (x_i)$, one or two-sided, is said to contain the word $\mathbf{i} = i_1, i_2, \ldots, i_k$ if for some j, $i_l = x_{j+l}$ for $l = 1, \ldots, k$. Length of a word $\mathbf{i}$ is denoted by $|\mathbf{i}|$. A word made of first n symbols of a given sequence x is denoted by $x|n := x_1, \ldots, x_n$.

We define the *shift map* $T\colon \Sigma_m^+ \to \Sigma_m^+$ by setting $T((x_i)_{i=1}^{\infty}) = (x_{i+1})_{i=1}^{\infty}$. In the case of two-sided sequences the shift map is defined as $T\colon \Sigma \to \Sigma$, $T((x_i)_{i=-\infty}^{\infty}) = (x_{i+1})_{i=-\infty}^{\infty}$. The pair (Σ_m^+, T) is called the *full one-sided shift* on alphabet $\mathcal{A}_m$ and the pair (Σ_m, T) is called the *full two-sided shift.*

Full shift is a compact metric space with respect to the metric

$$d(x, y) = d((x_i), (y_i)) = 2^{-\min\{|i|\colon x_i \neq y_i\}},$$

where $i \in \mathbb{N}$ in the case of one-sided shift or $i \in \mathbb{Z}$ in the case of two-sided shift. For a given word $\mathbf{i} = i_1, \ldots, i_n \in \mathcal{A}_m^n$ and $k \in \mathbb{N}$, we define a *cylinder set*

$$[\mathbf{i}]_k = \{(x_i)_{i=1}^{\infty} \in \Sigma_m^+\colon x_k, x_{k+1}, \ldots, x_{k+n-1} = \mathbf{i}\}.$$

The collection of all cylinder sets generates the Borel sigma-algebra of Σ_m^+. Actually the cylinder sets with $k = 1$ are enough to generate $\mathcal{B}$. In the case of two-sided shift we define a *cylinder set* as

$$[\mathbf{i}]_k = \{(x_i)_{i=-\infty}^{\infty} \in \Sigma_m\colon x_k, x_{k+1}, \ldots, x_{k+n-1} = \mathbf{i}\}$$

for $k \in \mathbb{Z}$. The collection of all cylinder sets generates the Borel sigma-algebra of Σ_m. We write $[\mathbf{i}] := [\mathbf{i}]_1$.

Definition 1.10. A subspace $(X, T|_X) \subset (\Sigma_m^+, T)$ is called *a one-sided subshift of finite type* if there exists an $m \times m$-matrix A with entries in $\{0, 1\}$, such that

$$X = \{x \in \Sigma_m^+\colon A_{x_i, x_{i+1}} = 1 \text{ for every } i \in \mathbb{N}\}.$$

The definition of *two-sided subshift of finite type* is obtained by replacing Σ_m^+ with Σ_m and $\mathbb{N}$ with $\mathbb{Z}$. The matrix A is called a *transition matrix.* A word w is *allowed* in X if there is a sequence $x \in X$ containing w. The set of all allowed finite words in X is denoted by X^* and the set of all allowed words of length n is denoted by X_n^*.

Remark 1.11. Subshifts of finite type are often defined more generally. Let L be a finite list of *forbidden* words. If $(X, T|_X) \subset (\Sigma_m^+, T)$ is such that X

is closed, $T(X) \subset X$ and no sequence in X contains a word from L, then $(X, T|_X)$ is called a *subshift of finite type*. Definition 1.10 is a special case of this definition, corresponding to the case where all forbidden words have length 2. Nevertheless, every subshift of finite type (X, T_X), defined in the general sense, is conjugate to some subshift of finite type (Y, T_Y), defined in the sense of Definition 1.10 (see [24, p. 4-5]). Being conjugate means that there exists a bijection $\phi\colon X \to Y$ such that $T_Y \circ \phi = \phi \circ T_X$. Thus, it is justified to consider only subshifts of finite type given by Definition 1.10.

If (X, T) is a subshift of finite type, then X is a closed subset of the full shift and hence a compact metric space. The shift map $T\colon X \to X$ is a continuous transformation. If $\mathcal{B}$ is the Borel sigma-algebra of X, then by Proposition 1.4 there exists an ergodic T-invariant probability measure μ.

1.3 Linear algebra

Let $E = \{e_1, e_2, \dots e_d\}$ be a set of linearly independent vectors on $\mathbb{R}^d$. We define an inner product $(\cdot|\cdot)$ with respect to this base by requiring

$$(e_i|e_j) = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}.$$

We call the basis E the *canonical basis* of the inner product space $(\mathbb{R}^d, (\cdot|\cdot))$. The inner product induces the norm $|v| = \sqrt{(v|v)}$ for every $v \in \mathbb{R}^d$.

In this thesis we work with linear maps from $\mathbb{R}^d$ to $\mathbb{R}^d$. For this reason, the following definitions are given in the context of real square matrices. On the set of matrices we use the operator norm $\|M\| = \sup_{v \in \mathbb{R}^d, |v|=1} |Mv|$, for every $M \in \mathbb{R}^{d \times d}$.

Definition 1.12. A symmetric matrix $M \in \mathbb{R}^{d \times d}$ with only non-negative eigenvalues is called *positive semi-definite*. If M has only positive eigenvalues, then M is called *positive definite*. An *nth root* of a positive semi-definite matrix M is a positive semi-definite matrix M_0 such that $M_0^n = M$. The nth root of M is uniquely defined and is denoted by $M^{1/n}$ ([16, Theorem 7.2.6, p. 405]).

Lemma 1.13. *For any matrix $M \in \mathbb{R}^{d \times d}$, the product MM^* is symmetric and positive semi-definite, where M^* denotes the transpose of M. If M is non-singular, then MM^* is positive definite.*

Proof. The matrix MM^* is symmetric since $(MM^*)^* = (M^*)^* M^* = MM^*$. Let λ be an eigenvalue of MM^* and let v be a unit eigenvector for λ. The eigenvalue λ is non-negative since

$$|Mv|^2 = (Mv|Mv) = (Mv)^*(Mv) = v^*M^*Mv = \lambda v^*v = \lambda.$$

Thus, all the eigenvalues of MM^* are non-negative. If M is non-singular, then $|Mv| > 0$, which implies that all the eigenvalues are positive. $\square$

Definition 1.14. Let $M\colon \mathbb{R}^d \to \mathbb{R}^d$ be a linear map. The eigenvalues $\alpha_1(M) \geq \ldots \geq \alpha_d(M) \geq 0$ of $(MM^*)^{1/2}$ listed with algebraic multiplicities are called the *singular values* of M. Equivalently singular values can be defined to be the lengths of semi-axes of the ellipsoid $M(B)$, where $B \subset \mathbb{R}^d$ is the unit ball. Throughout this thesis the singular values are indexed in decreasing order, usually without specifically mentioning it.

Definition 1.15. One can show that for every matrix $M \in \mathbb{R}^{d\times d}$ there exists a decomposition of the form $M = UDV$, where U, V are orthogonal matrices and D is a diagonal matrix with diagonal elements being the singular values of M (see [18, Theorem 2, p. 192]. Such decomposition is called a *singular value decomposition* (SVD) of M. We can assume that the constructed SVD is such that the diagonal matrix D lists the singular values of M in decreasing order, i.e. $D = \operatorname{diag}(\alpha_1(M), \ldots, \alpha_d(M))$.

Definition 1.16. A block matrix $A \in \mathbb{R}^{d\times d}$ of the form

$$A = \begin{bmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A_n \end{bmatrix},$$

where $A_i \in \mathbb{R}^{d_i\times d_i}$, $\sum_{i=1}^n d_i = d$ and the zeros are blocks of zeros, is called a *block-diagonal matrix*. We write

$$A = \bigoplus_{i=1}^n A_i.$$

The following lemma is a useful observation about how to construct a SVD of a block diagonal matrix from SVDs of the diagonal blocks.

Lemma 1.17. *Let A_i be a $d_i \times d_i$-square matrix and let $U_iD_iV_i$ be its singular value decomposition, for $i = 1, \ldots, n$. A singular value decomposition of $A = \bigoplus_{i=1}^n A_i$ is given by*

$$\left(\bigoplus_{i=1}^n U_i\right)\left(\bigoplus_{i=1}^n D_i\right)\left(\bigoplus_{i=1}^n V_i\right).$$

Note that the diagonal matrix $\bigoplus_{i=1}^{n} D_i$ does not necessarily list the singular values of A in decreasing order but this is the case if the singular values of A_i are greater than or equal to the singular values of A_{i+1} for $i = 1, \dots, n-1$.

The next proposition lists some facts about eigen- and singular values needed in the later sections.

Proposition 1.18. *Let $M \in \mathbb{R}^{d\times d}$ and let $\lambda_1(M), \dots, \lambda_d(M) \in \mathbb{C}$ be the eigenvalues of M listed with algebraic multiplicities and so that $|\lambda_i(M)| > |\lambda_{i+1}(M)|$ for every i. Also let $\alpha_1(M) \geq \dots \geq \alpha_d(M)$ be the singular values of M.*

(1) $|\det M| = \prod_{i=1}^{d} \alpha_i(M)$ *and* $\det M = \prod_{i=1}^{d} \lambda_i(M)$

(2) Maps $M \mapsto \det M$, $M \mapsto \alpha_i(M)$ and $M \mapsto \lambda_i(M)$ are continuous. The set of matrices is concidered with operator norm induced by euclidean norm.

(3) $\|M\| = \alpha_1(M)$

(4) For invertible M, $\|M^{-1}\| = 1/\alpha_d(M)$

(5) $\lambda_i((MM^*)^{1/2n}) = (\lambda_i(MM^*))^{1/2n}$ *for $i = 1, \dots, d$, $n \in N$.*

Proof. The first equality of (1) is obtained by taking SVD of M and calculating the determinant. Let $\lambda_1, \dots, \lambda_d$ be the eigenvalues of M. Recall that the eigenvalues of M are the roots of characteristic polynomial $p(t) = \det(M - tI)$. Factoring the characteristic polynomial as $p(t) = (-1)^d(t-\lambda_1)\dots(t-\lambda_d)$ and setting $t = 0$ results the second equality.

For (2), note that the coefficients of characteristic polynomial depend continuously on M and recall that the roots of polynomial depend continuously on the polynomial coefficients. Thus, the eigenvalues of M depend continuously on M. Now rest of (2) follows easily from (1) and the definition of singular values.

The truth of claim (3) is obvious if we adopt the alternative definition of singular values via lengths of semi-axes of transformed unit ball.

For (4), let α_d be the least singular value of M and let v be a unit vector such that $|Mv| = \alpha_d$. As $|M^{-1}Mv/\alpha_d| = 1/\alpha_d$, we have $\|M^{-1}\| \geq 1/\alpha_d$. On the other hand, it is impossible for $\|M^{-1}\|$ to be strictly greater than $1/\alpha_d$, since otherwise

$$\left| M\frac{M^{-1}w}{|M^{-1}w|} \right| = \frac{1}{|M^{-1}w|} < \alpha_d$$

for some unit vector $w \in \mathbb{R}^d$. This would contradict the fact that $\alpha_d = \min_{|x|=1} |Mx|$.

It is well known that for any square matrix $A \in \mathbb{R}^{d\times d}$ and integer $m \in \mathbb{N}$, the eigenvalues of A^m are the eigenvalues of A raised to the power of m. Using this and the fact that $(MM^*)^{1/2n}$ is positive semi-definite we have

$$\lambda_i(MM^*) = \lambda_i(((MM^*)^{1/2n})^{2n}) = \lambda_i((MM^*)^{1/2n})^{2n},$$

which implies that $\lambda_i((MM^*)^{1/2n}) = (\lambda_i(MM^*))^{1/2n}$ for $i = 1, \ldots, d$. $\square$

1.4 Multilinear algebra

In this thesis we need a few concepts from multilinear algebra, namely the m-dimensional exterior algebra of $\mathbb{R}^d$ and the exterior product of a linear map. Only few key properties, given in 1.22, of exterior algebra are needed later on. Therefore, only a non-constructive definition of exterior algebra via universal property is given.

Definition 1.19. Let V and W be real vector spaces. A map $f\colon V^k \to W$ is called a *k-linear map* if

$$f(v_1, \ldots, av_i + bv_i', \ldots, v_k) = af(v_1, \ldots, v_i, \ldots, v_k) + bf(v_1, \ldots, v_i', \ldots, v_k)$$

for every $a, b \in \mathbb{R}$ and $i = 1, \ldots, k$. A map $f\colon V^k \to W$ is called an *alternating map* if

$$f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -f(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)$$

for each $i \neq j$.

Definition 1.20. Let V be a real vector space. The *j-th exterior power of V* is a pair $(V^{\wedge j}, \wedge^j)$, where $V^{\wedge j}$ is a vector space and $\wedge^j\colon V^j \to V^{\wedge j}$ is alternating j-linear map with the following property (called *the universal property*): If $\phi\colon V^j \to H$ is alternating j-linear map into any vector space H, then there exists a unique linear map $f : V^{\wedge j} \to H$ such that $\phi = f \circ \wedge^j$. The image $\wedge^j(v_1, \ldots, v_j)$ is denoted by $v_1 \wedge \ldots \wedge v_j$. The space $V^{\wedge j}$ is called the *j-th exterior algebra of V*. It can be shown that *j-th exterior power of V* exists and is unique up to isomorphism (see [15, Chapter 5]). Moreover, the span of the image $\wedge^j(V^k)$ is $V^{\wedge j}$.

We only work with the j-th exterior algebra of $\mathbb{R}^d$. The space $(\mathbb{R}^d)^{\wedge j}$ can be equipped with an inner product $(\cdot \mid \cdot)$ by defining

$$(v_1 \wedge \cdots \wedge v_j \mid w_1 \wedge \cdots \wedge w_j) = \det(v_a \cdot w_b)_{1\leq a,b\leq j},$$

where $v_a \cdot w_b$ stands for the usual inner product on $\mathbb{R}^d$ given by the canonical basis, and extending this to whole $(\mathbb{R}^d)^{\wedge j} \times (\mathbb{R}^d)^{\wedge j}$ by requiring bi-linearity of $(\cdot|\cdot)$.

Every linear map $A : \mathbb{R}^d \to \mathbb{R}^d$ induces a linear map acting on the j-th exterior algebra of $\mathbb{R}^d$.

Definition 1.21. Let $A : \mathbb{R}^d \to \mathbb{R}^d$ be a linear map. The *j-fold exterior product* of A is the linear map $A^{\wedge j} : (\mathbb{R}^d)^{\wedge j} \to (\mathbb{R}^d)^{\wedge j}$, satisfying

$$A^{\wedge j}(v_1 \wedge \cdots \wedge v_j) = Av_1 \wedge \cdots \wedge Av_j$$

for every $v_i \in \mathbb{R}^d$, $i = 1, \ldots, j$. We use the operator norm $\|A^{\wedge j}\|$ induced by the norm given by the inner product defined above.

The next proposition lists properties of $(\mathbb{R}^d)^{\wedge j}$ needed later on. Especially property (3) is the key reason why we consider exterior algebras in the first place.

Proposition 1.22. *Let $\{e_1, \ldots, e_d\}$ be an orthonormal basis of $\mathbb{R}^d$. Let $A, B \in \mathbb{R}^{d\times d}$. Then*

(1) The set
$$\{e_{i_1} \wedge \ldots \wedge e_{i_j} : 1 \le i_1 < \ldots < i_j \le d\}$$
is an orthonormal basis of $(\mathbb{R}^d)^{\wedge j}$.

(2) $(AB)^{\wedge j} = A^{\wedge j}B^{\wedge j}$ and as a consequence $\|(AB)^{\wedge j}\| \le \|A^{\wedge j}\|\|B^{\wedge j}\|$.

(3) $\|A^{\wedge j}\| = \alpha_1(A)\cdots\alpha_j(A)$ and therefore $\|A^{\wedge j}\| \le \|A\|^j$.

(4) $A^{\wedge d}(v_1 \wedge \ldots \wedge v_d) = \det(A) \cdot (v_1 \wedge \ldots \wedge v_d)$.

(5) $|v_1 \wedge \ldots \wedge v_j| \le |v_1 \wedge \ldots \wedge v_i||v_{i+1} \wedge \ldots \wedge v_j|$ for any $j = 2, \ldots, d$, $i = 1, \ldots, j-1$ and $\{v_k\}_{k=1}^{j} \subset \mathbb{R}^d$.

Proof. Proofs of (1) and (4) are standard material of books concerning multilinear algebra (for example [15]). Equality of (2) follows straight from definition 1.21 and the inequality is just a basic property of any operator norm. For proof of (3) see [21, Theorem 2.9, p. 50]. Property (5) can be proved using determinant inequality known as Fisher's inequality: *Let $M \in \mathbb{R}^{j\times j}$ be symmetric and positive definite. Let $M_i \in \mathbb{R}^{i\times i}$ be the submatrix of M formed by the first i rows and columns. Let $M_i^{\complement}$ be the submatrix of M formed by rows and columns with indices $i+1, \ldots, j$. Then*

$$\det M \le \det M_i \det M_i^{\complement}.$$

For the proof of this inequality see [14, p. 35]. It is a well-known fact that $v_1 \wedge \ldots \wedge v_j = 0$ if and only if the vectors $v_1, \ldots, v_j$ are linearly dependent. Thus, inequality (5) is clearly true when the vectors $v_1, \ldots, v_j$ are linearly dependent. Assume now that they are linearly independent. We may write

$$|v_1 \wedge \ldots \wedge v_j|^2 = \det(V^*V),$$

where V is a matrix whose columns are the vectors $v_1, \ldots, v_j$. By Lemma 1.13, V^*V is a symmetric positive-definite matrix and by Fisher's inequality we obtain that

$$|v_1 \wedge \ldots \wedge v_j|^2 \leq \det(V^*V)_i \det(V^*V)_i^{\complement} = |v_1 \wedge \ldots \wedge v_i|^2 |v_{i+1} \wedge \ldots \wedge v_j|^2.$$

□

Given $\varepsilon > 0$ and $A \in \mathbb{R}^{d\times d}$ we can find a neighbourhood $\mathcal{U}$ of A such that $\|A^{\wedge t} - B^{\wedge t}\| < \varepsilon$ for all $B \in \mathcal{U}$.

Proposition 1.23. *The map $A \mapsto A^{\wedge t}$ is continuous on $\mathbb{R}^{d\times d}$.*

Proof. Let $\{e_1, \ldots, e_d\}$ be an orthonormal basis of $\mathbb{R}^d$. Let

$$\mathcal{A} = \{(i_1, \ldots, i_t) \colon 1 \leq i_1 < \cdots < i_t \leq d\}.$$

For $\mathbf{i} = (i_1, \ldots, i_t) \in \mathcal{A}$, denote by $\mathcal{E}_{\mathbf{i}}$ the vector $e_{i_1} \wedge \ldots \wedge e_{i_t}$. By Proposition 1.22, the set $\{\mathcal{E}_{\mathbf{i}} \colon \mathbf{i} \in \mathcal{A}\}$ is an orthonormal basis of $(\mathbb{R}^d)^{\wedge t}$. Let $A, B \in \mathbb{R}^{d\times d}$ and a unit vector $v = \sum_{\mathbf{i}\in\mathcal{A}} a_{\mathbf{i}} \mathcal{E}_{\mathbf{i}} \in (\mathbb{R}^d)^{\wedge t}$, where $a_{\mathbf{i}} \in \mathbb{R}$. By the triangle and Cauchy-Schwarz inequalities,

$$\begin{aligned} |(A^{\wedge t} - B^{\wedge t})v|^2 &\leq \Big(\sum_{\mathbf{i}\in\mathcal{A}} |a_{\mathbf{i}}||A^{\wedge t}\mathcal{E}_{\mathbf{i}} - B^{\wedge t}\mathcal{E}_{\mathbf{i}}|\Big)^2 \\ &\leq \Big(\sum_{\mathbf{i}\in\mathcal{A}} a_{\mathbf{i}}^2\Big)\Big(\sum_{\mathbf{i}\in\mathcal{A}} |A^{\wedge t}\mathcal{E}_{\mathbf{i}} - B^{\wedge t}\mathcal{E}_{\mathbf{i}}|^2\Big) = \sum_{\mathbf{i}\in\mathcal{A}} |A^{\wedge t}\mathcal{E}_{\mathbf{i}} - B^{\wedge t}\mathcal{E}_{\mathbf{i}}|^2. \end{aligned}$$

Thus, it is sufficient to show that, for a fixed A, $|A^{\wedge t}\mathcal{E}_{\mathbf{i}} - B^{\wedge t}\mathcal{E}_{\mathbf{i}}|$ is small when B is close to A. Let $\mathcal{E}_{\mathbf{i}} = e_{i_1} \wedge \ldots \wedge e_{i_t}$. Now

$$\begin{aligned} |A^{\wedge t}\mathcal{E}_{\mathbf{i}} - B^{\wedge t}\mathcal{E}_{\mathbf{i}}|^2 = (A^{\wedge t}\mathcal{E}_{\mathbf{i}}|A^{\wedge t}\mathcal{E}_{\mathbf{i}}) &- (A^{\wedge t}\mathcal{E}_{\mathbf{i}}|B^{\wedge t}\mathcal{E}_{\mathbf{i}}) \\ &+ (B^{\wedge t}\mathcal{E}_{\mathbf{i}}|B^{\wedge t}\mathcal{E}_{\mathbf{i}}) - (B^{\wedge t}\mathcal{E}_{\mathbf{i}}|A^{\wedge t}\mathcal{E}_{\mathbf{i}}), \end{aligned}$$

so by symmetry it is sufficient to show that

$$(A^{\wedge t}\mathcal{E}_{\mathbf{i}}|A^{\wedge t}\mathcal{E}_{\mathbf{i}}) - (A^{\wedge t}\mathcal{E}_{\mathbf{i}}|B^{\wedge t}\mathcal{E}_{\mathbf{i}}) = \det(Ae_{i_a} \cdot Ae_{i_b})_{1\leq a,b\leq t} - \det(Ae_{i_a} \cdot Be_{i_b})_{1\leq a,b\leq t}$$

is small when $\|A - B\|$ is small. This follows by recalling that the determinant is continuous function of the matrix entries and by observing that

$$Ae_i \cdot Ae_j - Ae_i \cdot Be_j \leq \|A\|\|A - B\|.$$

□

2 Pressure functions and introduction of the continuity theorems

The topological pressure is an important tool in dimension theory. In the case of self affine sets the pressure for singular value function can be used to determine an upper bound for the dimension of the set. In some cases the dimension itself can be obtained. This motivates the study of the continuity properties of the pressure.

2.1 Pressure for singular value function and dimension of a self-affine set

For $d \geq 1$ and $k \geq 2$ let

$$\begin{aligned}
\mathcal{A}_{d,k} &= \{\mathbf{A} = (A_1, \ldots, A_k) \colon A_i \in \mathbb{R}^{d\times d}\}, \\
\mathcal{A}^C_{d,k} &= \{\mathbf{A} = (A_1, \ldots, A_k) \colon A_i \in \mathbb{R}^{d\times d}, \|A_i\| < 1\}, \\
\mathcal{G}_{d,k} &= \{\mathbf{A} = (A_1, \ldots, A_k) \colon A_i \in GL_d(\mathbb{R})\}, \\
\mathcal{G}^C_{d,k} &= \{\mathbf{A} = (A_1, \ldots, A_k) \colon A_i \in GL_d(\mathbb{R}), \|A_i\| < 1\},
\end{aligned}$$

where $\|\cdot\|$ is the operator norm induced by the euclidean norm and $GL_d(\mathbb{R})$ is the set of $d \times d$ invertible real matrices. Linear mappings with norm less than one are said to be *contractive*, hence the superscript "C". We endow $\mathcal{A}_{d,k}$ with the L^∞-topology, i.e. the topology given by the norm $\|\mathbf{A}\|_\infty = \max_{i=1,\ldots,k} \|A_i\|$.

For given $\mathbf{A} \in \mathcal{A}^C_{d,k}$ and translations $t = (t_1, \ldots, t_k) \in \mathbb{R}^{kd}$, there exists a unique non-empty compact set $F = F(\mathbf{A}, t)$ such that

$$F = \bigcup_{i=1}^{k} A_i(F) + t_i$$

(see, for example, [11, Theorem 9.1, p. 135]). The set F is called a *self-affine set*. The dimensional properties of the self-affine sets have been studied widely but no exact formula for the Hausdorff or box counting dimension of general self-affine set is known (for definitions see [11, Chapters 2 and 3]). However, in the case $\|A_i\| < 1/2$ for all i there is a formula relating a certain subadditive pressure to the dimension of a self-affine set. To state this result we need a few definitions and lemmas.

Lemma 2.1. *For every subadditive sequence $(a_k)_{k=1}^\infty$, i.e. $a_{m+n} \leq a_k + a_m$ for all $m, n \in \mathbb{N}$, the limit $\lim_{k\to\infty} a_k/k$ exists and equals $\inf_{k\in\mathbb{N}} a_k/k \in [-\infty, \infty)$.*

Proof. See [10, Proposition 1.1, p. 3]. □

Definition 2.2. For, $s \geq 0$ the function $\varphi^s \colon \mathbb{R}^{d\times d} \to [0,\infty)$ defined as

$$\varphi^s(A) = \begin{cases} \alpha_1(A)\cdots\alpha_m(A)\alpha_{m+1}(A)^{s-m} & \text{if } 0 \leq s < d \\ |\det(A)|^{s/d} & \text{if } s \geq d \end{cases},$$

where $m = \lfloor s \rfloor$ is the integer part of s, is called the *singular value function.* Here a convention $0^0 = 1$ is used. The singular value function is submultiplicative, i.e. $\varphi^s(AB) \leq \varphi^s(A)\varphi^s(B)$ [9, Proposition 2.1, p. 341]. For $\mathbf{i} = (i_1, \ldots, i_n)$ where $i_j \in \{1, \ldots, k\}$, we write $A_{\mathbf{i}} = A_{i_1}\cdots A_{i_n}$ and $|\mathbf{i}| = n$. We define a pressure

$$P(\mathbf{A}, s) = \lim_{n\to\infty} \frac{1}{n}\log\left(\sum_{|\mathbf{i}|=n} \varphi^s(A_{\mathbf{i}})\right) = \inf_{n\in\mathbb{N}} \frac{1}{n}\log\left(\sum_{|\mathbf{i}|=n} \varphi^s(A_{\mathbf{i}})\right) \in [-\infty, \infty).$$

This limit exists and equals the infimum by Lemma 2.1. For $\mathbf{A} \in \mathcal{A}^C_{d,k}$, the number

$$s(\mathbf{A}) = \inf\{s \geq 0 \colon P(\mathbf{A}, s) \leq 0\}$$

is called the *singularity dimension* of $F(\mathbf{A}, t)$.

Lemma 2.3. *The pressure $P(\mathbf{A}, s)$ is strictly decreasing in $s > 0$ for any $\mathbf{A} \in \mathcal{A}^C_{d,k}$. As a consequence $P(\mathbf{A}, t) < 0$ for any $t > s(\mathbf{A})$.*

Proof. There is $\overline{\alpha} < 1$ such that $\|A_i\| \leq \overline{\alpha}$ for $i = 1, 2, \ldots, k$. For any $s, \delta > 0$, we have

$$\varphi^{s+\delta}(A_{\mathbf{i}}) \leq \varphi^s(A_{\mathbf{i}})\overline{\alpha}^{\delta|\mathbf{i}|}.$$

We show this in the case where $s + \delta < d$ and $\lfloor s \rfloor < \lfloor s + \delta \rfloor$. Let $\ell = \lfloor s \rfloor$, $m = \lfloor s + \delta \rfloor$ and denote the singular values of $A_{\mathbf{i}}$ by $\alpha_1 \geq \ldots \geq \alpha_d$. Then

$$\begin{aligned} \varphi^{s+\delta}(A_{\mathbf{i}}) &= \alpha_1\cdots\alpha_m\alpha_{m+1}^{s+\delta-m} = \varphi^s(A_{\mathbf{i}})\alpha_{\ell+1}^{\ell-s+1}\alpha_{\ell+2}\cdots\alpha_m\alpha_{m+1}^{s+\delta-m} \\ &\leq \varphi^s(A_{\mathbf{i}})\alpha_1^{(\ell-s+1)+(m-\ell-1)+(s+\delta-m)} = \varphi^s(A_{\mathbf{i}})\|A_{\mathbf{i}}\|^\delta \\ &\leq \varphi^s(A_{\mathbf{i}})\overline{\alpha}^{\delta|\mathbf{i}|}. \end{aligned}$$

This is even easier to see in the case $\lfloor s \rfloor = \lfloor s + \delta \rfloor$. The case $s + \delta \geq d$ is handled similarly. Thus,

$$\begin{aligned} \frac{1}{n}\log\left(\sum_{|\mathbf{i}|=n} \varphi^{s+\delta}(A_{\mathbf{i}})\right) &\leq \frac{1}{n}\log\left(\overline{\alpha}^{\delta n}\sum_{|\mathbf{i}|=n} \varphi^s(A_{\mathbf{i}})\right) \\ &= \delta\log\overline{\alpha} + \frac{1}{n}\log\left(\sum_{|\mathbf{i}|=n} \varphi^s(A_{\mathbf{i}})\right). \end{aligned}$$

Taking limit as $n \to \infty$, we get $0 < -\delta\log\overline{\alpha} \leq P(\mathbf{A}, s) - P(\mathbf{A}, s+\delta)$. □

We denote the upper box counting dimension, the box counting dimension and the Hausdorff dimension of a set $F \subset \mathbb{R}^d$ by $\overline{\dim}_B(F)$, $\dim_B(F)$ and $\dim_H(F)$ respectively. The singularity dimension is related to the box counting and Hausdorff dimension of a self-affine set.

Theorem 2.4. *For any $\mathbf{A} \in \mathcal{A}^C_{d,k}$, the following holds:*

(1) $\overline{\dim}_B(F(\mathbf{A}, t)) \leq \min(s(\mathbf{A}), d)$ *for all $t \in \mathbb{R}^{kd}$.*

(2) If $\|A_i\| < 1/2$ for every i, then for Lebesgue almost all $t \in \mathbb{R}^{kd}$

$$\dim_H(F(\mathbf{A}, t)) = \dim_B(F(\mathbf{A}, t)) = \min(s(\mathbf{A}), d).$$

Proof. See [9, Theorem 5.3, p. 347] and [23, Proposition 3.1]. □

Remark 2.5. Falconer [9] proved this theorem for the norm bound 1/3 and later Solomyak [23] modified the proof so that this bound could be replaced with 1/2. In general this theorem won't hold in the case $\|A_i\| \geq 1/2$ [23, Proposition 3.1]. The original result was proved for the case $\mathbf{A} \in \mathcal{G}^C_{d,k}$ but Feng and Shmerkin state in [12], without pointing out the necessary modifications, that the proof goes through with no assumptions on the invertibility of mappings A_i.

In the light of Theorem 2.4, it is natural to ask whether $s(\mathbf{A})$ and $P(\mathbf{A}, s)$ depend continuously on $\mathbf{A}$. Feng and Shmerkin gave positive answer in [12] in the form of the following theorem.

Theorem 2.6.

(1) For any $s \geq 0$, the pressure $P(\mathbf{A}, s)$ depends continuously on $\mathbf{A}$ taking values on $\mathcal{A}_{d,k}$.

(2) The singularity dimension $s(\mathbf{A})$ depends continuously on $\mathbf{A}$ taking values on $\mathcal{A}^C_{d,k}$.

(3) The map $(\mathbf{A}, s) \mapsto P(\mathbf{A}, s)$ is continuous at a point $(\mathbf{A}, s) \in \mathcal{A}_{d,k} \times [0, \infty)$ if $\mathbf{A} \in \mathcal{G}_{d,k}$ or $s \notin \{0, 1, \ldots, d-1\}$.

Remark 2.7. There are discontinuity points $(\mathbf{A}, s)$ of $P(\cdot, \cdot)$ for $\mathbf{A} \notin \mathcal{G}_{d,k}$ and $s \in \{1, \ldots, d-1\}$. For example, if $s \in \{1, \ldots, d-1\}$, $\alpha_s(A) > 0$ and $\alpha_{s+1}(A) = 0$ for a non-invertible $A \in \mathbb{R}^{d \times d}$, then P is discontinuous at $(\mathbf{A}, s)$, where $\mathbf{A} = (A, A, \ldots, A)$. This is since $P(\mathbf{A}, s) > -\infty$ and $P(\mathbf{A}, t) = -\infty$ for any $t > s$.

As a direct consequence of Theorems 2.4 and 2.6 we have the following theorem.

Theorem 2.8. *Let* $\mathcal{A}_{d,k}^{1/2} = \{\mathbf{A} \in \mathcal{A}_{d,k} \colon \|A_i\| < 1/2 \text{ for } i = 1, \dots k\}$*. By Theorem 2.4, for* $\mathbf{A} \in \mathcal{A}_{d,k}^{1/2}$ *the Hausdorff and box counting dimension of* $F(\mathbf{A}, t)$ *share a common constant value for Lebesgue almost all* $t \in \mathbb{R}^{kd}$*. Denote this common dimension by* $D(\mathbf{A})$*. The map* $\mathbf{A} \mapsto D(\mathbf{A})$ *is continuous on* $\mathcal{A}_{d,k}^{1/2}$*.*

Feng and Shmerkin proved Theorem 2.6 as a consequence of a more general statement concerning the continuity of the pressure function for matrix cocycles (Theorem 2.12).

2.2 Generalised pressure for matrix cocycles

Definition 2.9. Let (X, T) be a subshift of finite type. A function $A \colon X \times \mathbb{N} \to \mathbb{R}^{d \times d}$ is called a *matrix cocycle* if for all $x \in X$ and $n, m \in \mathbb{N}$ we have $A(x, 0) = I$ and $A(x, n+m) = A(T^m x, n)A(x, m)$, where I is the identity matrix. We call matrix cocycles simply *cocycles*. A cocycle A is uniquely determined by its generator $A(x) := A(x, 1)$ via

$$A(x, n) = A(T^{n-1}x) \cdots A(Tx)A(x).$$

We call a generator $A \colon X \to \mathbb{R}^{d \times d}$ also a *cocycle*. A cocycle A is called *locally constant* if $A(x)$ depends only on the first coordinate of x, i.e. $A(x) = A(x_1)$.

Definition 2.10. Let (X, T) be a one-sided subshift of finite type and denote by $C(X)$ the set of real-valued continuous functions on X. For a given $g \in C(X)$, we write

$$S_n g(x) = \sum_{i=0}^{n-1} g(T^i x).$$

For $g \in C(X)$, a cocycle $A \colon X \to \mathbb{R}^{d \times d}$ and $s \geq 0$, we define a subadditive pressure

$$P_g(A, s) = \lim_{n \to \infty} \frac{1}{n} \log \left(\sum_{\mathbf{i} \in X_n^*} \sup_{y \in [\mathbf{i}]} \exp(S_n g(y)) \varphi^s(A(y, n)) \right).$$

This limit exists and equals the infimum of the right hand side over $n \in \mathbb{N}$ due to the submultiplicativity of the expression between the parenthesis. More detailed reasoning is given in Section 5.2.

Remark 2.11. The pressure P_g includes the pressure P as a special case. Let X be the full shift over alphabet $\{1, \dots, k\}$. A sequence $\mathbf{A} = (A_1, \dots, A_k) \in \mathcal{A}_{d,k}$

of k matrices can be identified with a locally constant cocycle $A\colon X \to \mathbb{R}^{d\times d}$, $A(x) = A(x_1) = A_{x_1}$. By letting $g\colon X \to \mathbb{R}$ be the zero function, we have

$$\begin{aligned} P_g(A,s) &= \lim_{n\to\infty} \frac{1}{n} \log\left(\sum_{\mathbf{i}\in X_n} \sup_{y\in[\mathbf{i}]} \exp(S_n g(y))\varphi^s(A(y,n)) \right) \\ &= \lim_{n\to\infty} \frac{1}{n} \log\left(\sum_{|\mathbf{i}|=n} \varphi^s(A_{i_n}\cdots A_{i_1}) \right) = P(\mathbf{A}, s). \end{aligned}$$

Let $\mathcal{M}(X,d)$ be the collection of $d\times d$-matrix-valued cocycles on X. We endow $\mathcal{M}(X,d)$ with the $L^\infty(X,\mathbb{R}^{d\times d})$-topology, i.e. the topology given by the norm $\|A\|_\infty = \sup_{x\in X}\|Ax\|$. The main result of Feng and Shmerkin in [12] concerns the continuity of the subadditive pressure $P_g(A,s)$.

Theorem 2.12. *Let $g \in C(X)$ be a continuous real-valued function on a one-sided subshift of finite type X. Then the following statements hold:*

(1) For fixed $s \geq 0$, any locally constant cocycle A is a continuity point of the pressure map $P_g(\cdot, s)$ on $\mathcal{M}(X,d)$ with respect to $L^\infty(X,\mathbb{R}^{d\times d})$-topology.

(2) Let $A \in \mathcal{M}(X,d)$ be locally constant and $s \geq 0$. If the values of A are in $GL_d(\mathbb{R})$ or $s \notin \{0,1,\dots,d-1\}$, then $(A,s) \in \mathcal{M}(X,d)\times[0,\infty)$ is a continuity point of the pressure map $P_g(\cdot,\cdot)$, where $\mathcal{M}(X,d)$ is considered with the $L^\infty(X,\mathbb{R}^{d\times d})$-topology.

Quite a few tools must yet be presented until we are ready to prove this theorem in Section 6.

3 The cone condition

A key ingredient to the proof of Theorem 2.12 is finding a large subsystem of the original system which satisfies the cone condition. This section introduces the cone condition along with several lemmas.

3.1 Cones and the cone condition

Definition 3.1. A convex and closed subset $K \subset V$ of a finite-dimensional Banach space is called a *cone* if $tv \in K$ for every $t > 0$, $v \in K$ and $K \cap -K = \{0\}$. Let $L(V,V)$ denote the set of all linear maps from V to V. A linear map

$B \in L(V,V)$ is said to satisfy the *cone condition* with cones $K, K' \subset V$, if $K' \setminus \{0\} \subset \mathrm{int}(K)$ and

$$B(K) \subset K' \text{ or } B(K) \subset -K'.$$

Here $\mathrm{int}(K)$ denotes the interior of K.

Let X be a set. A collection $A\colon X \to L(V,V)$ of linear maps is said to satisfy the cone condition with cones $K, K' \subset V$ if $A(x)$ satisfies the cone condition with K, K' for every $x \in X$.

Lemma 3.2. *Let V be a d-dimensional Banach space. If $A \in L(V,V)$ satisfies the cone condition with cones K and K', then $\mathrm{int}(K) \cap \ker(A) = \emptyset$, where K is the outer cone and $\ker(A) = \{v \in V\colon Av = 0\}$ is the null space of A.*

Proof. Assume the contrary. There exists a non-zero vector $w \in \mathrm{int}(K) \cap \ker(A)$. The definition of the cone condition implies that $\dim(\ker(A)) < d = \dim(K)$. By this and the fact that $w \in \mathrm{int}(K) \cap \ker(A)$, there is a vector $v \in K \cap \ker(A)^{\complement}$ such that $w - v \in K$. Now $A(w-v) = -Av \in -A(K) \setminus \{0\}$, but this contradicts with the fact that $K' \cap -K' = \{0\}$. $\square$

The cone condition is robust, in a sense of the following proposition.

Proposition 3.3. *Let $A \in L(V,V)$ satisfy the cone condition with cones K, K', latter being the inner cone. Let $\widehat{K}$ and $\widehat{K}'$ be cones such that $K' \setminus \{0\} \subset \mathrm{int}(\widehat{K}')$, $\widehat{K}' \setminus \{0\} \subset \mathrm{int}(\widehat{K})$ and $\widehat{K} \setminus \{0\} \subset \mathrm{int}(K)$. There exists a neighbourhood $\mathcal{U}$ of A such that every $B \in \mathcal{U}$ satisfies the cone condition with cones $\widehat{K}, \widehat{K}'$.*

Proof. We assume $A(K) \subset K'$. The proof for the case $A(K) \subset -K'$ is similar. By the assumptions and Lemma 3.2, the minimum distances

$$D_1 := \min_{|v|=1, v \in K'} d(v, \widehat{K}'^{\complement}) > 0,$$
$$D_2 := \min_{|v|=1, v \in \widehat{K}} |Av| > 0$$

are positive. Let $\mathcal{U}$ be a A-centred ball with radius less than $D_1 D_2$. Let $B \in \mathcal{U}$. Then for any unit vector $v \in \widehat{K}$ and a vector $u \in \widehat{K}'^{\complement}$, $|Av - Bv| < D_1 D_2$ and $|Av - u| \geq D_1 |Av| \geq D_1 D_2$. Thus, $D_1 D_2 \leq |Av - Bv| + |Bv - u| < D_1 D_2 + |Bv - u|$ implying that $Bv \in \widehat{K}'$. This means that B satisfies the cone condition with cones $\widehat{K}'$ and $\widehat{K}$ for any $B \in \mathcal{U}$. $\square$

Lemma 3.4. *Let V be a normed vector space. Endow $L(V,V)$ with the operator norm $\|\cdot\|_{op}$ and let $A \in L(V,V)$. Let E be a subset of the set $\{(A,v) \in L(V,V) \times V \colon \|A\|_{op} = 1, |v| = 1\}$. Then the map $f \colon E \to \mathbb{R}$ defined by $f(A,v) = |Av|$ is continuous with respect to the topology given by the norm $\|(A,v)\| = \|A\|_{op} + |v|$.*

Proof. For every $(A,v),(B,w) \in E$, we have

$$\begin{aligned} |f(A,v) - f(B,w)| &= \big||Av| - |Bw|\big| \le |Av - Bw| \\ &= |Av - Bv + Bv - Bw| \\ &\le |Av - Bv| + |Bv - Bw| \\ &\le \|A - B\|_{\mathrm{op}}|v| + \|B\|_{\mathrm{op}}|v - w| \\ &= \|A - B\|_{\mathrm{op}} + |v - w| \\ &= \|(A,v) - (B,w)\|. \end{aligned}$$

This implies the continuity of f. □

For us, the significance of the cone condition is due the following lemma.

Lemma 3.5. *Let K, K' be cones in a finite-dimensional Banach space such that $K' \setminus \{0\} \subset \operatorname{int}(K)$. Then there exists a constant $c > 0$, depending on the cones K and K', such that for any two linear maps A_1, A_2 satisfying $A_i(K) \subset K'$ or $A_i(K) \subset -K'$, for $i = 1, 2$, the following holds:*

$$c\|A_1\|\|A_2\| \le \|A_1 A_2\| \le \|A_1\|\|A_2\|,$$

where $\|A\|$ is the operator norm of A.

Proof. The right-hand side inequality is clear. First we show that there exists $c_1 > 0$ such that

$$\frac{|Aw|}{|w|} \ge c_1\|A\|$$

for any linear map $A \colon V \to V$ such that $A(K) \subset K'$ or $A(K) \subset -K'$ and any $w \in K' \setminus \{0\}$. Assuming the contrary, we can find a sequence of linear maps A'_n such that $A'_n(K) \subset K'$, and a sequence of vectors $w'_n \in K'$ such that

$$\frac{|A'_n w'_n|}{|w'_n|} < \frac{1}{n}\|A'_n\|. \tag{3.1}$$

We can pick the mappings A'_n so that $A'_n(K) \subset K'$ since if A'_n would map the cone K inside $-K'$, then we could just pick the linear map $-A'_n$ instead of A'_n.

Write $A_n := A'_n/\|A'_n\|$ and $w_n := w'_n/|w'_n|$. For every $n \in \mathbb{N}$, the linear map A_n has norm 1 and $A_n(K) \subset K'$. Also, the vector $w_n \in K'$ is of unit norm and by (3.1) $|A_n w_n| < 1/n$ for every $n \in \mathbb{N}$. We show that the set

$$E = \{(A, w) \in L(V, V) \times V \,:\, \|A\| = |w| = 1, A(K) \subset K'\}$$

is compact in the product topology. Recall that the intersection of a compact set and a closed set is compact, product of compact sets is compact and the unit circles are compact. Using these facts we see that E is compact if the set $F = \{A \in L(V, V)\colon A(K) \subset K'\}$ is closed. Let B be a linear map from the complement of F. There exists a unit vector $x \in K$ such that the distance $D := d(Bx, K')$ between the point Bx and the cone K' is greater than 0. Let A be any linear map from B-centred and D-radius open ball. Then $Ax \notin K'$, since otherwise we would have $D \le \|Bx - Ax\| < D$. Thus, F is closed which in turn implies the compactness of E.

Recall that the product topology of $L(V, V) \times V$ is given by the norm $\|(A, v)\| = \|A\| + |v|$. Thus, by Lemma 3.4 the map $f\colon E \to \mathbb{R}$ given by $f(A, w) = |Aw|$ is continuous. By the extreme value theorem, f attains its infimum on E, which is zero since $|A_n w_n| < 1/n$ for every n. Thus, there exists a linear map A of norm one such that $A(K) \subset K'$ and a unit vector $w \in K'$ such that $Aw = 0$. This conflicts with the Proposition 3.2 which states that $\mathrm{int}(K) \cap \ker(A) = \emptyset$.

Now, rest of the proof follows easily. For a fixed unit vector $w \in K'$ we have

$$\|A_1 A_2\| \ge |A_1 A_2 w| \ge c_1 \|A_1\| |A_2 w| \ge c_1^2 \|A_1\| \|A_2\|.$$

□

3.2 Block-diagonal matrices

Definition 3.6. Let $\varepsilon > 0$ and $\lambda \in \mathbb{R}$. A matrix $A \in \mathbb{R}^{d\times d}$ such that

$$\exp(\lambda - \varepsilon)|x| \le |Ax| \le \exp(\lambda + \varepsilon)|x| \text{ for all } x \in \mathbb{R}^d,$$

is called (λ, ε)-*conformal*. This means that all singular values of A lie between $\exp(\lambda - \varepsilon)$ and $\exp(\lambda + \varepsilon)$.

Let $H_i \in \mathbb{R}^{d_i \times d_i}$ and $\tau_i \in \mathbb{R}$ for $i = 1, \dots, p$ and write $d = \sum_{i=1}^p d_i$. We say that the block-diagonal matrix $H = \bigoplus_{i=1}^p H_i \in \mathbb{R}^{d\times d}$ is of *hyperbolic class* $(d_1, \tau_1), \dots, (d_p, \tau_p)$ with tolerance ε if

(1) $H_i \in \mathbb{R}^{d_i \times d_i}$ is (τ_i, ε)-conformal for $i = 1, \dots, p - 1$,

(2) $|H_p x| \le \exp(\tau_p + \varepsilon)|x|$ for all $x \in \mathbb{R}^{d_p}$,

(3) $\tau_i - \varepsilon > \tau_{i+1} + \varepsilon$ for $i = 1, \ldots, p-1$.

This implies that the singular values of block H_i are larger than the singular values of H_{i+1} for $i = 1, \ldots, p-1$. The set of all matrices of hyperbolic class $(d_1, \tau_1), \ldots, (d_p, \tau_p)$ with tolerance ε is denoted by $\mathcal{H}^{\varepsilon,p}_{(d_i,\tau_i)}$.

Lemma 3.7. *Let $H \in \mathcal{H}^{\varepsilon,p}_{(d_i,\tau_i)}$ and set $d_0 = 0$. If $j \in \mathbb{N}$ is such that $\sum_{i=0}^{r-1} d_i < j \leq \sum_{i=0}^{r} d_i$ for some $1 \leq r \leq p-1$, then*

$$\exp(\tau_r - \varepsilon) \leq \alpha_j(H) \leq \exp(\tau_r + \varepsilon).$$

Proof. Let UDV be a singular value decomposition of H given by Lemma 1.17. Then

$$D = (\alpha_1(H), \alpha_2(H), \ldots, \alpha_d(H))$$

and thus $\alpha_j(H)$ is a singular value of H_r. The (τ_r, ε)-conformality of H_r implies the claim. $\square$

Lemma 3.8. *Let $\{e_1, \ldots, e_d\}$ be the canonical basis of $\mathbb{R}^d$. Let $H \in \mathcal{H}^{\varepsilon,p}_{(d_i,\tau_i)}$. Let $r \in \{1, \ldots, p-1\}$. Write $t = \sum_{i=1}^{r} d_i$ and $\Gamma = \sum_{i=1}^{r} d_i\tau_i$. Then $e_1 \wedge \cdots \wedge e_t$ is an eigenvector of $H^{\wedge t}$ with eigenvalue λ such that $|\lambda| \geq \exp(\Gamma - t\varepsilon)$. Also, the inequality*

$$|H^{\wedge t}(e_{i_1} \wedge \cdots \wedge e_{i_t})| \leq \exp(\Gamma + \tau_{r+1} - \tau_r + t\varepsilon)$$

holds whenever $1 \leq i_1 < \cdots < i_t \leq d$ and $(i_1, \ldots, i_t) \neq (1, \ldots, t)$.

Proof. We identify every H_i with the restriction of H to the corresponding d_i-dimensional subspace of $\mathbb{R}^d$. Then by Proposition 1.22

$$\begin{aligned} H^{\wedge t}(e_1 \wedge \cdots \wedge e_t) &= He_1 \wedge \cdots \wedge He_t \\ &= H_1e_1 \wedge \cdots \wedge H_1e_{d_1} \wedge H_2e_{d_1+1} \wedge \cdots \wedge H_2e_{d_1+d_2} \wedge \cdots \wedge H_re_t \\ &= \det(H_1) \cdots \det(H_r) \cdot (e_1 \wedge \cdots \wedge e_t). \end{aligned}$$

Hence, $e_1 \wedge \cdots \wedge e_t$ is an eigenvector of $H^{\wedge t}$ with the eigenvalue $\prod_{i=1}^{r} \det(H_i)$ which's absolute value equals the product of singular values of matrices H_i by Lemma 1.18. Thus, (τ_i, ε)-conformality of matrices H_i leads to the desired inequality

$$\begin{aligned} \left| \prod_{i=1}^{r} \det(H_i) \right| &\geq \exp(\tau_1 - \varepsilon)^{d_1} \cdots \exp(\tau_r - \varepsilon)^{d_r} = \exp\left(\sum_{i=1}^{r} (d_i\tau_i - d_i\varepsilon) \right) \\ &= \exp(\Gamma - t\varepsilon). \end{aligned}$$

Let $H = UDV$ be a singular value decomposition of H. The exterior product $U^{\wedge t}$ remains isometric since

$$|U^{\wedge t}v|^2 = \det(U(v_i) \cdot U(v_j))_{1\le a,b\le t} = \det(v_i \cdot v_j)_{1\le a,b\le t} = |v|^2$$

for any $v = v_1 \wedge \cdots \wedge v_t \in (\mathbb{R}^d)^{\wedge t}$, by orthogonality of U. Therefore,

$$|H^{\wedge t}(e_{i_1} \wedge \cdots \wedge e_{i_t})| = |D^{\wedge t}V^{\wedge t}(e_{i_1} \wedge \cdots \wedge e_{i_t})|.$$

Since V is an orthogonal matrix of the block form given by Lemma 1.17, we may write $V^{\wedge t}(e_{i_1} \wedge \cdots \wedge e_{i_t})$ in a form

$$V^{\wedge t}(e_{i_1} \wedge \cdots \wedge e_{i_t}) = \bigwedge_{j=1}^{p} w_1^j \wedge \cdots \wedge w_{t_j}^j,$$

where, for each j, $\{w_i^j\}_{i=1}^{t_j}$ is an orthonormal system in $\mathbb{R}^{d_j}$ embedded in $\mathbb{R}^d$. Note that the system $\{w_i^j\}_{i=1}^{t_j}$ is empty if all of the vectors $e_{i_1}, \ldots, e_{i_t}$ are zero when restricted to the subspace $\mathbb{R}^{d_j}$. The number of non-zero vectors amongst $\{e_{i_1}, \ldots, e_{i_t}\}$ when restricted to $\mathbb{R}^{d_j}$ is t_j. The number t_j is greater than zero for some $j > r$ since $(i_1, \ldots, i_t) \neq (1, \ldots, t)$. This in turn implies that for some $l \le r$ we have $t_l < d_l$. From this, together with the fact that $\tau_i > \tau_{i+1}$, we deduce that

$$\sum_{j=1}^{p} t_j\tau_j \le d_1\tau_1 + \cdots + d_{r-1}\tau_{r-1} + (d_r - 1)\tau_r + \tau_{r+1} = \Gamma + \tau_{r+1} - \tau_r.$$

Using (5), (3) of Proposition 1.22 and taking in to account that $w_1^j \wedge \cdots \wedge w_{t_j}^j$ is a unit vector in $(\mathbb{R}^{d_j})^{\wedge t_j}$ and $H \in \mathcal{H}_{d_i,\tau_i}^{\varepsilon,p}$, we have

$$\begin{aligned}
|H^{\wedge t}(e_{i_1} \wedge \cdots \wedge e_{i_t})| &= \left|D^{\wedge t} \bigwedge_{j=1}^{p} w_1^j \wedge \cdots \wedge w_{t_j}^j\right| \\
&= \left|\bigwedge_{j=1}^{p} D_j^{\wedge t_j}(w_1^j \wedge \cdots \wedge w_{t_j}^j)\right| \\
&\le \prod_{j=1}^{p} \left|D_j^{\wedge t_j}(w_1^j \wedge \ldots \wedge w_{t_j}^j)\right| \\
&\le \prod_{j=1}^{p} \left\|D_j^{\wedge t_j}\right\| \le \prod_{j=1}^{p} \|D_j\|^{t_j} \\
&\le \prod_{j=1}^{p} \exp(t_j(\tau_j + \varepsilon)) = \exp\left(\sum_{j=1}^{p} t_j\tau_j + t\varepsilon\right) \\
&\le \exp(\Gamma + \tau_{r+1} - \tau_r + t\varepsilon).
\end{aligned}$$

□

Definition 3.9. Let $(V, (\cdot \mid \cdot))$ be a finite dimensional inner product space and $v \in V$ a non-zero vector. For $0 < r < 1$ the r-cone around v is defined as a set

$$\mathcal{C}(v, r) = \{w \in V : (v|w) \geq (1-r)|v||w|\}.$$

This is the set of vectors whose angle with v is no more than $\cos^{-1}(1-r)$.

We use $v^{\perp}$ to denote the subspace of V orthogonal to the vector v.

Lemma 3.10. *Let $A\colon V \to V$ be a linear map on a finite dimensional inner product space. Let v be an eigenvector of A with eigenvalue $\lambda > 0$ such that $\|A|_{v^{\perp}}\| \leq \lambda/18$. Then $A(\mathcal{C}(v, 1/2)) \subset \mathcal{C}(v, 1/5)$.*

Proof. If v satisfies the conditions of the statement, then does also the vector $v/|v|$ and they define the same r-cone. Thus, we may assume that $|v| = 1$. Let $w \in \mathcal{C}(v, 1/2)$ be a non-zero vector. We can write the vector w as $w = av + bz$ for some $a, b \in \mathbb{R}$ and $z \in v^{\perp}$, $|z| = 1$. Since $w \in \mathcal{C}(v, 1/2)$,

$$\begin{aligned}(v \mid w) &= (v \mid av + bz) = a(v \mid v) + b(v \mid z) = a \\ &\geq (1 - 1/2)|v||w| = \frac{1}{2}|w| = \frac{1}{2}\sqrt{a^2 + b^2},\end{aligned}$$

and thus $\sqrt{3}a \geq |b|$. Also note that $a > 0$. Using these and the Cauchy-Schwarz inequality, we may calculate

$$\begin{aligned}(v \mid Aw) &= (v \mid a\lambda v + bAz) = a\lambda + (v \mid bAz) \geq a\lambda - |b||v||Az| \\ &\geq a\lambda - |b|\lambda/18 \geq a\lambda(1 - \sqrt{3}/18) > 0\end{aligned}$$

and

$$|Aw| \leq |a\lambda v| + |bAz| \leq a\lambda + |b||Az| \leq a\lambda + \sqrt{3}a\lambda/18 = a\lambda(1 + \sqrt{3}/18).$$

Therefore,

$$\frac{(v \mid Aw)}{|v||Aw|} \geq \frac{1 - \sqrt{3}/18}{1 + \sqrt{3}/18} \geq \frac{1 - 2/18}{1 + 2/18} = 4/5 = 1 - 1/5,$$

so $Aw \in \mathcal{C}(v, 1/5)$.

□

Corollary 3.11. *Let $H \in \mathcal{H}^{\varepsilon,p}_{(d_i,\tau_i)}$ be such that*

$$\sqrt{\binom{d}{\lfloor d/2 \rfloor}} \max_{1\le r\le p-1} \exp(-\tau_r + \tau_{r+1} + 2d\varepsilon) < 1/18.$$

Denote $e^{\wedge t} = e_1 \wedge \cdots \wedge e_t$ and $t_r = \sum_{i=1}^r d_i$. Then $H^{\wedge t_r}$ satisfies the cone condition with cones $\mathcal{C}(e^{\wedge t_r}, 1/2)$ and $\mathcal{C}(e^{\wedge t_r}, 1/5)$ for $r = 1, 2, \ldots, p-1$.

Proof. Let $1 \le r \le p-1$ and denote $t = t_r$. Let $\mathcal{A}$ be the collection of indices $(i_1, \ldots, i_t)$ such that $1 \le i_1 < \cdots < i_t \le d$. The number of elements in $\mathcal{A}$ is $|\mathcal{A}| = \binom{d}{t} \le \binom{d}{\lfloor d/2 \rfloor}$. Let $v = e^{\wedge t}$. By Lemma 3.8 v is an eigenvector of $H^{\wedge t}$ and for the corresponding eigenvalue λ, $|\lambda| \ge \exp(\Gamma - t\varepsilon)$, where $\Gamma = \sum_{i=1}^r d_i\tau_i$. Also, by Lemma 3.8,

$$|H^{\wedge t}(e_{i_1} \wedge \cdots \wedge e_{i_t})| \le \exp(\Gamma + \tau_{r+1} - \tau_r + t\varepsilon)$$

for every $(i_1, \ldots, i_t) \in \widetilde{\mathcal{A}}$, where $\widetilde{\mathcal{A}} = \mathcal{A} \setminus \{(1, \ldots, t)\}$.

We show that $\|H^{\wedge t}|_{v^\perp}\| \le |\lambda|/18$ and use Lemma 3.10 to establish the conclusion. For $\mathbf{i} = (i_1, \ldots, i_t) \in \mathcal{A}$, denote by $\mathcal{E}_\mathbf{i}$ the vector $(e_{i_1} \wedge \ldots \wedge e_{i_t})$. Recall from Proposition 1.22 that the set $\{\mathcal{E}_\mathbf{i} \colon \mathbf{i} \in \mathcal{A}\}$ is an orthonormal basis of $(\mathbb{R}^d)^{\wedge t}$. Let w be a unit vector in $v^\perp$. The vector w can be written as

$$w = \sum_{\mathbf{i}\in\widetilde{\mathcal{A}}} a_\mathbf{i}\mathcal{E}_\mathbf{i}$$

for some $a_\mathbf{i} \in \mathbb{R}$. Remark that $\sum_{\mathbf{i}\in\widetilde{\mathcal{A}}} a_\mathbf{i}^2 = 1$ since $|w| = 1$ and the basis we are working on is orthonormal. By the Cauchy-Schwarz inequality

$$\begin{aligned}|H^{\wedge t}w|^2 &= \Big|\sum_{\mathbf{i}\in\widetilde{\mathcal{A}}} a_\mathbf{i} H^{\wedge t}\mathcal{E}_\mathbf{i}\Big|^2 \le \Big(\sum_{\mathbf{i}\in\widetilde{\mathcal{A}}} |a_\mathbf{i}||H^{\wedge t}\mathcal{E}_\mathbf{i}|\Big)^2 \le \Big(\sum_{\mathbf{i}\in\widetilde{\mathcal{A}}} a_\mathbf{i}^2\Big)\Big(\sum_{\mathbf{i}\in\widetilde{\mathcal{A}}} |H^{\wedge t}\mathcal{E}_\mathbf{i}|^2\Big)\\ &\le |\widetilde{\mathcal{A}}| \max_{\mathbf{i}\in\widetilde{\mathcal{A}}} |H^{\wedge t}\mathcal{E}_\mathbf{i}|^2 \le |\mathcal{A}| \max_{\mathbf{i}\in\widetilde{\mathcal{A}}} |H^{\wedge t}\mathcal{E}_\mathbf{i}|^2 \le \binom{d}{\lfloor d/2 \rfloor} \max_{\mathbf{i}\in\widetilde{\mathcal{A}}} |H^{\wedge t}\mathcal{E}_\mathbf{i}|^2.\end{aligned}$$

Using the inequalities given by Lemma 3.8, we have

$$\begin{aligned}|H^{\wedge t}w| &\le \sqrt{\binom{d}{\lfloor d/2 \rfloor}} \max_{\mathbf{i}\in\widetilde{\mathcal{A}}} |H^{\wedge t}\mathcal{E}_\mathbf{i}| \le \sqrt{\binom{d}{\lfloor d/2 \rfloor}} \exp(\Gamma + \tau_{r+1} - \tau_r + t\varepsilon)\\ &\le \sqrt{\binom{d}{\lfloor d/2 \rfloor}} \exp(\Gamma - t\varepsilon) \exp(-\tau_r + \tau_{r+1} + 2d\varepsilon) < \frac{\exp(\Gamma - t\varepsilon)}{18}\\ &\le \frac{|\lambda|}{18}\end{aligned}$$

and thus $\|H|_{v^\perp}\| \leq |\lambda|/18$. If $\lambda > 0$, then the corollary follows directly from Lemma 3.10. If $\lambda < 0$, then $-\lambda$ is an eigenvalue of $-H^{\wedge t}$ with eigenvector $e^{\wedge t}$ and $\|-H^{\wedge t}\| \leq -\lambda/18$. By Lemma 3.10 $-H^{\wedge t}(\mathcal{C}(e^{\wedge t}, 1/2)) \subset \mathcal{C}(e^{\wedge t}, 1/5)$, so $H^{\wedge t}(\mathcal{C}(e^{\wedge t}, 1/2)) \subset -\mathcal{C}(e^{\wedge t}, 1/5)$. □

3.3 Space of splittings

Definition 3.12. The space of k-dimensional subspaces of $\mathbb{R}^d$ is called the *Grassmannian* $G(d,k)$. We endow it with the metric $d(V,W) = \inf \|O - I\|$ where the infimum is over all orthogonal maps $O \in \mathbb{R}^{d\times d}$ such that $OV = W$. A direct sum presentation $\mathbf{V} = \bigoplus_{i=1}^p V_i$ of $\mathbb{R}^d$, with subspaces V_i, is called a *splitting* of $\mathbb{R}^d$. Let $d_1, \dots, d_p \in \mathbb{N}$ be such that $d = \sum_{i=1}^p d_i$. The collection of all splittings $\mathbf{V} = \bigoplus_{i=1}^p V_i$, where $V_i \in G(d, d_i)$, is denoted by $\mathcal{S} = \mathcal{S}_{d_1,\dots,d_p}$.

Denote the product $\prod_{i=1}^p G(d, d_i)$ of Grassmannians by $\mathcal{G} = \mathcal{G}_{d_1,\dots,d_p}$. We may identify $\mathcal{S}$ with the subset

$$\left\{(V_1, \dots, V_p) \in \mathcal{G} \colon \mathbb{R}^d = \bigoplus_{i=1}^p V_i\right\} \subset \mathcal{G}$$

of product space of Grassmannians. We endow $\mathcal{G}$ with the l^∞-metric so the distance between splittings $\mathbf{V}, \mathbf{W} \in \mathcal{S}$ is given by

$$d_\infty(\mathbf{V}, \mathbf{W}) = \max_{i=1}^p d(V_i, W_i).$$

With respect to this metric, $\mathcal{S}$ is an open subset of $\mathcal{G}$, i.e. for any splitting $\mathbf{V} \in \mathcal{S}$ there exits $r > 0$ such that if $d_\infty(\mathbf{V}, \mathbf{W}) < r$ for $\mathbf{W} \in \mathcal{G}$, then $\mathbf{W} \in \mathcal{S}$.

Let $\{e_j\}_{j=1}^d$ be the canonical basis of $\mathbb{R}^d$ and let V_i^* be the subspace of $\mathbb{R}^d$ generated by the vectors $\{e_j\}_{j=d_0+\dots+d_{i-1}+1}^{d_0+\dots+d_i}$, where $d_0 = 0$. We fix the *canonical splitting* $\mathbf{V}^* = \bigoplus_{i=1}^p V_i^*$.

Lemma 3.13. *There exits $\delta = \delta(d) > 0$ such that, if $A \in \mathbb{R}^{d\times d}$ is a linear map satisfying*

(1) there are splittings $\mathbf{V}, \mathbf{W} \in \mathcal{S}_{d_1,\dots,d_p}$ with $AV_p \subset W_p$, $AV_i = W_i$ for $i = 1, \dots, p-1$ and

$$d(\mathbf{V}, \mathbf{V}^*), d(\mathbf{W}, \mathbf{V}^*) < \delta,$$

(2) there are real numbers $\{\tau_i\}_{i=1}^p$ and $\varepsilon > 0$ such that

$$\sqrt{\binom{d}{\lfloor d/2 \rfloor}} \max_{1\leq r\leq p-1} \exp(-\tau_r + \tau_{r+1} + 2d\varepsilon) < 1/18,$$

and

$$\exp(\tau_i - \varepsilon)|v| \leq |Av| \leq \exp(\tau_i + \varepsilon)|v|$$

for every $v \in V_i \setminus \{0\}$, $i = 1, \ldots, p-1$, *and*

$$|Av| \leq \exp(\tau_p + \varepsilon)|v|$$

for every $v \in V_p \setminus \{0\}$,

then

$$A^{\wedge t_r}\big(\mathcal{C}(e^{\wedge t_r}, 1/3)\big) \subset \pm\mathcal{C}(e^{\wedge t_r}, 1/4) \text{ for } r = 1, \ldots, p-1,$$

where t_r *and* $e^{\wedge t_r}$ *are as in Corollary 3.11 and the notation* $E \subset \pm F$ *means that either* $E \subset F$ *or* $E \subset -F$.

Proof. Let $\delta > 0$ be small enough to guarantee that if $\|U - I\| < (d+1)\delta$ for $U \in \mathbb{R}^{d\times d}$, then

$$U^{\wedge t}\big(\mathcal{C}(e^{\wedge t}, 1/3)\big) \subset \mathcal{C}(e^{\wedge t}, 1/2) \quad \text{and} \quad U^{\wedge t}\big(\mathcal{C}(e^{\wedge t}, 1/5)\big) \subset \mathcal{C}(e^{\wedge t}, 1/4) \quad (3.2)$$

for $t = 1, \ldots, d$. Let $A\colon \mathbb{R}^d \to \mathbb{R}^d$ be a linear map such that conditions (1) and (2) are satisfied. Let $\mathbf{V}$ and $\mathbf{W}$ be the splittings of condition (1). Condition (1) guarantees that there is an orthogonal map $O_i \in \mathbb{R}^{d\times d}$ such that $O_i V_i^* = V_i$ and $\|O_i - I\| \leq \delta$ for $i = 1, \ldots, p$. Let O be the linear map that equals O_i on V_i^* for all i, i.e. $Ov = \sum_{i=1}^p O_i v_i$ for every $v = v_1 + \ldots + v_p \in \mathbb{R}^d = \bigoplus_{i=1}^p V_i^*$. Let $v \in \mathbb{R}^d$ be a unit vector. Then

$$|(O-I)v| = \Big|\sum_{i=1}^p (O_i - I)v_i\Big| \leq \sum_{i=1}^p \Big|(O_i - I)v_i\Big| \leq \sum_{i=1}^p \Big\|O_i - I\Big\| \leq p\delta$$

implying that $\|O - I\| \leq p\delta < (d+1)\delta$. Similarly, we define a linear map O' such that $O'|_{V_i^*}$ is an orthogonal map onto W_i and $\|O' - I\| < (d+1)\delta$.

Let $H = (O')^{-1}AO$. We show $H \in \mathcal{H}^{\varepsilon,p}_{(d_i,\tau_i)}$. It is easy to check that the first inequality of condition (2) implies that $\tau_i - \varepsilon < \tau_{i+1} + \varepsilon$. By condition (1) and the way O and O' are constructed, we have $HV_i^* = V_i^*$ for $i = 1, \ldots, p-1$ and $HV_p^* \subset V_p^*$. Thus, H is of the block form $H = \bigoplus_{i=1}^p H_i$, where $H_i \in \mathbb{R}^{d_i \times d_i}$. By condition (2) the blocks H_i are (τ_i, ε)-conformal for $i = 1, \ldots, p-1$ and $|H_p v| \leq \exp(\tau_p + \varepsilon)$. Thus, H is of hyperbolic class $\mathcal{H}^{\varepsilon,p}_{(d_i,\tau_i)}$. By Corollary 3.11 $H^{\wedge t_r}$ satisfies the cone condition with cones $\mathcal{C}(e^{\wedge t_r}, 1/2)$ and $\mathcal{C}(e^{\wedge t_r}, 1/5)$. Thus, by (3.2) and by noting that $(B^{\wedge t})^{-1} = (B^{-1})^{\wedge t}$ for any $B \in \mathrm{GL}_d(\mathbb{R})$,

$$A^{\wedge t_r} O^{\wedge t_r}\big(\mathcal{C}(e^{\wedge t_r}, 1/2)\big) \subset \pm(O')^{\wedge t_r}\big(\mathcal{C}(e^{\wedge t_r}, 1/5)\big) \subset \pm\mathcal{C}(e^{\wedge t_r}, 1/4). \quad (3.3)$$

The norm of an orthogonal matrix is 1, thus

$$\|O - I\| = \|O(I - O^{-1})\| \leq \|I - O^{-1}\| = \|O^{-1}(O - I)\| \leq \|O - I\|,$$

which implies that $\|O^{-1} - I\| = \|O - I\| \le (d+1)\delta$. Therefore, by (3.2),

$$\mathcal{C}(e^{\wedge t_r}, 1/3) \subset O^{\wedge t_r}\left(\mathcal{C}(e^{\wedge t_r}, 1/2)\right). \tag{3.4}$$

By (3.3) and (3.4),

$$A^{\wedge t_r}\left(\mathcal{C}(e^{\wedge t_r}, 1/3)\right) \subset \pm\mathcal{C}(e^{\wedge t_r}, 1/4)$$

for $r = 1, \dots, p-1$. $\square$

4 Subcocycles satisfying the cone condition

The proof of Theorem 2.12 is based on finding a nicely behaving subsystem (after iteration) that captures almost all of the topological pressure of the original system. One of these nice properties is that the subcocycle and its exterior powers satisfy the cone condition.

In this section (X, T) is a subshift of finite type with the Borel sigma-algebra $\mathcal{B}$. We denote the family of T-invariant ergodic measures on X by $\mathcal{E}$.

To find the subsystem we use the semi-invertible version of Oseledets multiplicative ergodic theorem formulated for symbolic dynamics.

Theorem 4.1. *Let (X, T) be a two sided subshift and $\mu \in \mathcal{E}$. Let $A\colon X \to \mathbb{R}^{d\times d}$ be a measurable map such that*

$$\int \log^+ \|A(x)\| d\mu(x) < \infty,$$

where $\log^+(a) = \max\{0, a\}$ for $a \in \mathbb{R}$. Then, there exists $\lambda_1 > \cdots > \lambda_p \ge -\infty$, dimensions $d_1, \dots, d_p$ with $\sum_{i=1}^p d_i = d$ and a measurable family of splittings $\mathbf{E}(x) = \bigoplus_{i=1}^p E_i(x) \in \mathcal{S}_{d_1,\dots,d_p}$ such that for μ-almost every x the following holds:

(1) $A(x)E_i(x) \subset E_i(Tx)$, with equality if $\lambda_i > -\infty$

(2) for all $v \in E_i(x) \setminus \{0\}$,

$$\lim_{n\to\infty} \frac{\log |A(x,n)v|}{n} = \lambda_i, \tag{4.1}$$

and the convergence is uniform on any compact subset of $E_i(x) \setminus \{0\}$.

Proof. All but the uniform convergence is proved in [13, Theorem 4.1]. To prove the uniform convergence, we note that the proofs of Oseledets theorems provide the existence of the limit

$$\lim_{n\to\infty} (A(x,n)^* A(x,n))^{1/2n} = B(x) \tag{4.2}$$

for almost every x, and the different eigenvalues of B are $\exp(\lambda_1) > \cdots > \exp(\lambda_p)$ (see, for example, [2, Theorem 3.4.1, p. 134]). Fix $x \in X$ such that the limits (4.1) and (4.2) exist. Let $U_1, \ldots, U_p$ be the eigenspaces of $B(x)$ and write for brevity $B(x) = B$ and $E_i(x) = E_i$ for $i = 1, \ldots, p$. The matrix B is symmetric as a limit of symmetric matrices and hence diagonalisable, so $\mathbb{R}^d = U_1 \oplus \cdots \oplus U_p$. First, assume that $\lambda_p > -\infty$. Again, by the standard proofs of Oseledets Theorem, for any $\varepsilon > 0$ there exists a constant C (depending on ε and x) such that for all n

$$C|v| \exp(n(\lambda_i - \varepsilon)) \le |A(x,n)v| \le C|v| \exp(n(\lambda_i + \varepsilon)) \tag{4.3}$$

for all $v \in U_i$ (see [2, Theorem 3.4.11]). Thus, $(1/n)\log|A(x,n)v|$ converges uniformly to λ_i on any compact $K_i \subset U_i \setminus \{0\}$ since $\max_{v\in K_i} |v|$ and $\min_{v\in K_i} |v|$ are finite and non-zero. This implies that for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\exp(n(\lambda_i - \varepsilon)) \le |A(x,n)v| \le \exp(n(\lambda_i + \varepsilon)) \tag{4.4}$$

for every $n > N$ and $v \in K_i$.

Let $0 < \varepsilon < \min_{i=1,\ldots,p-1}(\lambda_i - \lambda_{i+1})/2$. Let $F \subset \mathbb{R}^d \setminus \{0\}$ be a compact set such that $F \subset U_\ell \oplus U_{\ell+1} \oplus \cdots \oplus U_p$ and $F \cap (U_{\ell+1} \oplus \cdots \oplus U_p) = \{0\}$. Every $v \in F$ can be written as $v = \sum_{j=\ell}^p u_j$, with $u_j \in U_j$, $j = \ell, \ell+1, \ldots, p$, and $u_\ell \neq 0$. Since $\lambda_\ell > \lambda_j$ for any $j > \ell$, there exists $N_0 \in \mathbb{N}$ such that

$$|A(x,n)v| \le \sum_{j=\ell}^{p} |A(x,n)u_j| \le \sum_{j=\ell}^{p} \exp(n(\lambda_j + \varepsilon)) \le 2\exp(n(\lambda_\ell + \varepsilon)) \tag{4.5}$$

for every $n > N_0$ and $v \in F$. On the other hand, since $\lambda_\ell - \varepsilon > \lambda_j + \varepsilon$ for any $j > \ell$, N_0 can be chosen large enough that

$$|A(x,n)u_\ell| - \sum_{j=\ell+1}^{p} |A(x,n)u_j| \ge \exp(n(\lambda_\ell - \varepsilon)) - \sum_{j=\ell+1}^{p} \exp(n(\lambda_j + \varepsilon))$$
$$\ge \frac{1}{2}\exp(n(\lambda_\ell - \varepsilon)) > 0$$

for every $n > N_0$ and $v \in F$. Thus, we may use the reverse triangle inequality to obtain

$$|A(x,n)v| \geq |A(x,n)u_\ell| - \sum_{j=\ell+1}^{p} |A(x,n)u_j| \geq \frac{1}{2}\exp(n(\lambda_\ell - \varepsilon)). \tag{4.6}$$

By (4.6) and (4.5), $\lim_{n\to\infty}(1/n)\log|A(x,n)v| = \lambda_\ell$ uniformly on F.

This with the claim (2) of the theorem implies that $E_i \subset U_i \oplus U_{i+1} \oplus \cdots \oplus U_p$ and $E_i \cap (U_{i+1} \oplus \cdots \oplus U_p) = \{0\}$ for $i = 1,\dots,p$. Assuming the contrary, there exists $v \in E_i$ such that $\{v\} \subset U_\ell \oplus U_{\ell+1} \oplus \cdots \oplus U_p$ and $\{v\} \cap (U_{\ell+1} \oplus \cdots \oplus U_p) = \{0\}$ for some $\ell \neq i$. But this would imply that $\lim_{n\to\infty}(1/n)\log|A(x,n)v| = \lambda_\ell$ contradicting (4.1). Thus, the convergence in (4.2) is uniform on any compact $K_i \subset E_i \setminus \{0\}$.

Now, assume that $\lambda_p = -\infty$. In this case (4.4) holds for $i = 1,\dots,p-1$. The reasons that lead to the right hand side inequality of (4.3) in [2] also give that for any $M < 0$ there exists a constant C such that

$$|A(x,n)v| \leq C|v|\exp(n(M+\varepsilon))$$

for every $v \in U_p$ and $n \in \mathbb{N}$. This implies that $(1/n)\log|A(x,n)v|$ converges uniformly to $-\infty$ on any compact set $K \subset U_p$. Rest of the proof is similar to the case $\lambda_p > -\infty$.

□

The data $(\lambda_i, d_i)_{i=1}^p$ is called the *Lyapunov spectrum* of (A,μ). If (X,T) is a one-sided subshift and $A\colon X \to \mathbb{R}^{d\times d}$ satisfies the assumptions of Theorem 4.1, then the limiting matrix $B(x)$ of (4.2) exists for μ-almost every $x \in X$ and has eigenvalues $\exp(\lambda_1) > \cdots > \exp(\lambda_p)$ which do not depend on x, neither does their algebraic multiplicities $d_1,\dots,d_p$. Again the data $(\lambda_i, d_i)_{i=1}^p$ is called the *Lyapunov spectrum* of (A,μ).

Remark 4.2. Let E_1 be the canonical basis of $\mathbb{R}^d$ (for the definition see the beginning of Section 1.3) and denote the corresponding norm by $|\cdot|_1$. Let $A\colon X \to \mathbb{R}^{d\times d}$ be a map satisfying the assumptions of Theorem 4.1, $(\lambda_i, d_i)_{i=1}^p$ the Lyapunov spectrum and $\mathbf{E}(x)$ a splitting given by Theorem 4.1. Consider $\mathbb{R}^d$ with another canonical basis E_2 and denote the induced norm by $|\cdot|_2$. Since $\mathbb{R}^d$ is a finite dimensional vector space, the norms $|\cdot|_1$ and $|\cdot|_2$ are equivalent, i.e. there exists constants $c_1, c_2 > 0$ such that $c_1|v|_1 \leq |v|_2 \leq c_2|v|_1$ for every $v \in \mathbb{R}^d$. Thus, the splitting $\mathbf{E}(x)$ satisfies the claims of Theorem 4.1 in the new canonical basis E_2 since, for typical $x \in X$,

$$\lim_{n\to\infty} \frac{\log c_j|A(x,n)v|_1}{n} = \lim_{n\to\infty}\left(\frac{\log|A(x,n)v|_1}{n} + \frac{\log c_j}{n}\right) = \lambda_i$$

for every $v \in E_i(x) \setminus \{0\}$ and $j = 1,2$.

Lemma 4.3. *Let $A : X \to \mathbb{R}^{d\times d}$ be a measurable cocycle and $\mu \in \mathcal{E}$ be such that $\int \log^+ \|A(x)\| d\mu(x) < \infty$. Let $\{(\lambda_i, d_i)\}_{i=1}^p$ be the Lyapunov spectrum. Write $d_0 = 0$, $t_r = \sum_{i=0}^r d_i$ and $\Gamma_r = \sum_{i=0}^r d_i\lambda_i$. If s is such that $t_r < s \leq t_{r+1}$ for some $0 \leq r < p$, then*

$$\lim_{n\to\infty} \frac{1}{n} \int \log \varphi^s(A(x,n))d\mu(x) = \Gamma_r + (s - t_r)\lambda_{r+1}.$$

Proof. Fix $r \in \{0, 1, \dots, p-1\}$ and $t_r < s \leq t_{r+1}$. Denote by m the integer part of s, i.e. $m = \lfloor s \rfloor$. Since $\|B^{\wedge k}\| = \alpha_1(B)\cdots\alpha_k(B)$ for any $B \in \mathbb{R}^{d\times d}$ and $k \in \{1, \dots, d\}$, we have

$$\begin{aligned}\varphi^s(B) &= \alpha_1(B)\cdots\alpha_m(B)\alpha_{m+1}(B)^{s-m}\\ &= (\alpha_1(B)\cdots\alpha_m(B))^{m-s+1}(\alpha_1(B)\cdots\alpha_{m+1}(B))^{s-m}\\ &= \|B^{\wedge m}\|^{m-s+1}\|B^{\wedge m+1}\|^{s-m}. \end{aligned} \tag{4.7}$$

This includes also the case $m = s = t_{r+1} = d$ if we ignore the fact that in this case $\alpha_{m+1}(B)$ is not defined and set $\alpha_{m+1}(B)^0 = \|B^{\wedge m+1}\|^0 = 1$. Next, denote by $\gamma_k(B)$ the product of k largest eigenvalues of symmetric $B \in \mathbb{R}^{d\times d}$, with algebraic multiplicities, for $k = 1, \dots, d$, and set $\gamma_0(B) = 1$. By Lemma 1.18, the map $B \mapsto \gamma_k(B)$ is continuous. To use the Kingman's subaddtive theorem, we note that the sequence $\{\log \|A(x,n)^{\wedge k}\|\}_{n\in\mathbb{N}}$ is subadditive for every $x \in X$ since

$$\|A(x,n+m)^{\wedge k}\| = \|A(T^m x,n)^{\wedge k}A(x,m)^{\wedge k}\| \leq \|A(T^m x,n)^{\wedge k}\|\|A(x,m)^{\wedge k}\|.$$

For

$$k = \begin{cases} m \text{ or } m+1 & \text{if } s \neq t_{r+1} \\ m & \text{if } s = t_{r+1} \end{cases}$$

we have, by Kingman's subadditive ergodic theorem 1.7 , Propositions 1.22

and 1.18 and (4.2),

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\int \log\|A(x,n)^{\wedge k}\|d\mu(x) &= \lim_{n\to\infty}\frac{1}{n}\log\|A(x,n)^{\wedge k}\| \\
&= \lim_{n\to\infty}\frac{1}{n}\log\big[\alpha_1(A(x,n))\cdots\alpha_k(A(x,n))\big] \\
&= \lim_{n\to\infty}\frac{1}{n}\log\gamma_k\Big(\big(A(x,n)^*A(x,n)\big)^{1/2}\Big) \\
&= \lim_{n\to\infty}\log\gamma_k\Big(\big(A(x,n)^*A(x,n)\big)^{1/2n}\Big) \\
&= \log\gamma_k\Big(\lim_{n\to\infty}\big(A(x,n)^*A(x,n)\big)^{1/2n}\Big) \\
&= \log\Big[\exp(\lambda_{r+1})^{k-t_r}\prod_{i=1}^{r}\exp(\lambda_i)^{d_i}\Big] \\
&= \sum_{i=0}^{r} d_i\lambda_i + (k-t_r)\lambda_{r+1} = \Gamma_r + (k-t_r)\lambda_{r+1}. \qquad (4.8)
\end{aligned}$$

First assume that $s \neq t_{r+1}$. Then equations (4.7) and (4.8) give the conclusion

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}\int \log\varphi^s(A(x,n))d\mu(x) \\
&= \lim_{n\to\infty}\frac{1}{n}\int \log\left(\|A(x,n)^{\wedge m}\|^{m+1-s}\|A(x,n)^{\wedge(m+1)}\|^{s-m}\right)d\mu(x) \\
&=(m+1-s)\lim_{n\to\infty}\frac{1}{n}\int \log\|A(x,n)^{\wedge m}\|d\mu(x) \\
&\quad + (s-m)\lim_{n\to\infty}\frac{1}{n}\int \log\|A(x,n)^{\wedge(m+1)}\|d\mu(x) \\
&=(m+1-s)[\Gamma_r + (m-t_r)\lambda_{r+1}] + (s-m)[\Gamma_r + (m+1-t_r)\lambda_{r+1}] \\
&=\Gamma_r + (s-t_r)\lambda_{r+1}.
\end{aligned}$$

If $s = t_{r+1}$, then $m = t_{r+1}$, so

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\int \log\varphi^s(A(x,n))d\mu(x) &= \lim_{n\to\infty}\frac{1}{n}\int \log\|A(x,n)^{\wedge t_{r+1}}\|d\mu(x) \\
&= \Gamma_r + (s-t_r)\lambda_{r+1}.
\end{aligned}$$

□

Before constructing a subcocycle satisfying the cone condition, we recall a simple lemma considering measure of an intersection.

Lemma 4.4. *Let (X,μ) be a probability space. If A and B are measurable subsets of X then $\mu(A\cap B) \geq \mu(A) + \mu(B) - 1$.*

Proof. We can write $A \cup B = (A \cap B^{\complement}) \cup (A \cap B) \cup (B \cap A^{\complement})$, where the right hand side is a union of disjoint sets. Thus,

$$\begin{aligned}\mu(A \cup B) + \mu(A \cap B) &= \mu(A \cap B^{\complement}) + \mu(A \cap B) + \mu(B \cap A^{\complement}) + \mu(B \cap A) \\ &= \mu(A) + \mu(B).\end{aligned}$$

This implies the claim

$$\mu(A \cap B) = \mu(A) + \mu(B) - \mu(A \cup B) \geq \mu(A) + \mu(B) - 1.$$

□

Theorem 4.5. *Let $A : X \to \mathbb{R}^{d\times d}$ be a locally constant cocycle over two-sided subshift of finite type X, and $\mu \in \mathcal{E}$. Let $\{(\lambda_i, d_i)\}_{i=1}^{p}$ be the Lyapunov spectrum. As before, let $t_r = \sum_{i=1}^{r} d_i$. Then there exists $\eta > 0$, a set $S \subset \mathbb{N}$ with bounded gaps, and a collections $\{\Sigma_n \subset X_n^*\colon n \in S\}$ of allowed finite words such that:*

(1) Concatenations of arbitrary words in Σ_n are in X^* for all $n \in S$

(2) $\sum_{\mathbf{i}\in\Sigma_n} \mu[\mathbf{i}] \geq \eta$ for all $n \in S$

(3) Moreover, if there are at least two distinct Lyapunov exponents, i.e. $p > 1$, then there exist cones $K_r, K_r' \subset (\mathbb{R}^d)^{\wedge t_r}$, for every $1 \leq r \leq p-1$, with $K_r' \backslash \{0\} \subset \operatorname{int}(K_r)$, such that $A(x, n)^{\wedge t_r}$ satisfies the cone condition with cones K_r and K_r' whenever $n \in S$ and $x|n \in \Sigma_n$.

Proof. Let $\mathbf{E}(x)$ be the collection of splittings given by Theorem 4.1. The pushforward measure $\mathbf{E}_*\mu$ is defined via $\mathbf{E}_*\mu(\mathcal{V}) = \mu(\mathbf{E}^{-1}(\mathcal{V}))$, for every measurable collection of splittings $\mathcal{V} \subset \mathcal{S}_{d_1,\dots d_p}$. We can assume that the canonical splitting $\mathbf{V}^*$ is in the support of $\mathbf{E}_*\mu$. If not, then let $\mathbf{V} = \bigoplus_{i=1}^{p} V_i$ be a splitting from the support of $\mathbf{E}_*\mu$ and let $E = \{f_1, \dots, f_d\}$ be a basis of $\mathbb{R}^d$ such that the vectors $f_{d_i+1}, \dots, f_{d_{i+1}}$ span the subspace V_{i+1} for $i = 0, \dots, p-1$. Making E the canonical basis of $\mathbb{R}^d$ guarantees that the new canonical splitting $\mathbf{V}^*$ is in the support of $\mathbf{E}_*\mu$. By Remark 4.2, $\mathbf{E}(x)$ satisfies the claims of Oseledets theorem 4.1 also in this new basis. Also the change of canonical basis does not affect the statement.

Let δ be the constant from Lemma 3.13. Then $\mu(\mathbf{E}^{-1}(B(\mathbf{V}^*, \delta))) > 0$. Furthermore, for some symbol a in the alphabet of X we have that $\mu(Y) > 0$, where $Y = \mathbf{E}^{-1}(B(\mathbf{V}^*, \delta)) \cap [a]$.

The proof differs slightly for the cases $\lambda_p = -\infty$ and $\lambda_p \neq -\infty$. We cover these cases in parallel. Let $\varepsilon = \frac{1}{11+2d} \min_{r=1}^{p-1}(\lambda_r - \lambda_{r+1})$. Let S^{d-1} be the set of unit vectors in $\mathbb{R}^d$. Let $f_n^i\colon X \to \mathbb{R} \cup \{-\infty\}$ be the functions defined as

$$f_n^i(x) = \sup_{v\in E_i(x)\cap S^{d-1}} \left| \frac{\log |A(x, n)v|}{n} - \lambda_i \right|$$

for every i with $\lambda_i \neq -\infty$. By part (2) of the semi-invertible Oseledets theorem 4.1, $\lim_{n\to\infty} f_n^i(x) = 0$ for μ-almost every $x \in X$. If $\lambda_p = -\infty$, define

$$f_n^p(x) = \sup_{v\in E_i(x)\cap S^{d-1}} \left(\frac{\log |A(x,n)v|}{n} \right).$$

In this case $\lim_{n\to\infty} f_n^p(x) = -\infty$ for μ-almost every $x \in X$. Measurability of f_n^i can be verified by observing that, by compactness, the supremum can be replaced with maximum and then using the Measurable maximum theorem [1, Theorem 18.19, p. 605]. Let $\varepsilon' = \min\{\varepsilon, \mu(Y)^2/3\}$. By Egorov's theorem 1.1 there is a set $B \subset X$ such that $\mu(B) < \varepsilon'$ and the convergence of f_n^i is uniform on $Z := X \setminus B$. We may assume that B contains all of the non-typical points of Theorem 4.2. If $\lambda_p = -\infty$, fix a constant $\lambda < 0$ such that $-\lambda_{p-1} + \lambda \leq \max_{1\leq r\leq p-2}(-\lambda_r + \lambda_{r+1})$. There exists $N_0 \in \mathbb{N}$ such that if $n > N_0$, then

$$\left| \frac{\log\left|A(x,n)\frac{v}{|v|}\right|}{n} - \lambda_i \right| < \varepsilon \tag{4.9}$$

for every $x \in Z$ and every $v \in E_i(x) \setminus \{0\}$, for such i that $\lambda_i > -\infty$ and

$$\frac{\log\left|A(x,n)\frac{v}{|v|}\right|}{n} - \lambda < \varepsilon \tag{4.10}$$

for every $x \in Z$ and every $v \in E_p(x) \setminus \{0\}$ if $\lambda_p = -\infty$. Inequalities (4.9) and (4.10) imply that for $x \in Z$, $n > N_0$ we have

$$|A(x,n)v| \in [\exp(n(\lambda_i - \varepsilon))|v|, \exp(n(\lambda_i + \varepsilon))|v|] \text{ if } v \in E_i(x) \setminus \{0\} \tag{4.11}$$
$$|A(x,n)v| \in [0, \exp(n(\lambda + \varepsilon))|v|] \text{ if } v \in E_p(x) \text{ and } \lambda_p = -\infty. \tag{4.12}$$

Also note that $\mu(Z) = \mu(X) - \mu(B) \geq 1 - \mu(Y)^2/3$. We may assume N_0 is large enough, so that $\exp(N_0\varepsilon) \geq \sqrt{\binom{d}{\lfloor d/2 \rfloor}}$.

Khintchine's recurrence theorem 1.8 assures that the set

$$S := [N_0, \infty) \cap \{n \in \mathbb{Z} \colon \mu(Y \cap T^{-n}(Y)) > \mu(Y)^2/2\}$$

is infinite (with bounded gaps). For every $n \in S$ we define a set

$$\Sigma_n = \{x|n \colon x \in Y \cap T^{-n}(Y) \cap Z\}.$$

Every sequence in Y starts with a. If $x \in Y \cap T^{-n}(Y)$, then $x_1 = a$ and x_n is followed by a. Thus, concatenations of words in Σ_n are allowed. This proves claim (1).

Using Lemma 4.4 we have

$$\begin{aligned}\Sigma_{\mathbf{i}\in\Sigma_n}\mu[\mathbf{i}] &\geq \mu\Bigg(\bigcup_{\mathbf{i}\in\Sigma_n}[\mathbf{i}]\Bigg)\\ &\geq \mu(Y\cap T^{-n}(Y)\cap Z) \geq \mu(Y\cap T^{-n}(Y)) + \mu(Z) - 1\\ &\geq \frac{\mu(Y)^2}{2} + 1 - \frac{\mu(Y)^2}{3} - 1 = \frac{\mu(Y)^2}{6} =: \eta,\end{aligned}$$

which proves claim (2).

For the last claim, assume that $p > 1$. Let $n > N_0$ and $x \in X$ be such that $x|n \in \Sigma_n$. There exists $y \in Y \cap T^{-n}(Y) \cap Z$ such that $y|n = x|n$. Since A is locally constant, $A(x,n) = A(y,n)$. We proceed to show that $A(y,n)$ satisfies the assumptions of Lemma 3.13 with $\mathbf{V} = \mathbf{E}(y)$, $\mathbf{W} = \mathbf{E}(T^n y)$, $\tau_i = n\lambda_i$ and epsilon replaced with $n\varepsilon$. If $\lambda_p = -\infty$, then we must pick $\tau_p = n\lambda$. First observe that $d(\mathbf{V},\mathbf{V}^*)$, $d(\mathbf{W},\mathbf{V}^*) < \delta$ since $y \in Y$ and $T^n y \in Y$. Also by induction

$$\begin{aligned}A(y,n)V_i &= A(y,n)E_i(y) = A(T^{n-1}y)\cdots A(y)E_i(y)\\ &= A(T^{n-1}y)\cdots A(Ty)E_i(Ty) = \cdots = E_i(T^n y) = W_i\end{aligned}$$

for $i = 1,\ldots,p-1$. This holds also for $i = p$ if $\lambda_p \neq -\infty$. If $\lambda_p = -\infty$, then $A(y,n)V_p \subset W_p$. Let

$$K = \begin{cases} p-2 & \text{if } \lambda_p = -\infty \\ p-1 & \text{if } \lambda_p \neq -\infty \end{cases}.$$

Assumption (2) of Lemma 3.13 holds by (4.11), (4.12) and the following calculation:

$$\begin{aligned}&\sqrt{\binom{d}{\lfloor d/2\rfloor}}\max_{1\leq r\leq p-1}\exp(-\tau_r + \tau_{r+1} + 2dn\varepsilon)\\ =&\sqrt{\binom{d}{\lfloor d/2\rfloor}}\max_{1\leq r\leq K}\exp(-n\lambda_r + n\lambda_{r+1} + 2dn\varepsilon)\\ =&\sqrt{\binom{d}{\lfloor d/2\rfloor}}\exp(-n\min_{1\leq r\leq K}(\lambda_r - \lambda_{r+1}) + 2dn\varepsilon)\\ =&\sqrt{\binom{d}{\lfloor d/2\rfloor}}\exp(-n\varepsilon(11+2d) + 2dn\varepsilon) = \sqrt{\binom{d}{\lfloor d/2\rfloor}}\exp(n\varepsilon)^{-11}\\ \leq&\sqrt{\binom{d}{\lfloor d/2\rfloor}}\sqrt{\binom{d}{\lfloor d/2\rfloor}}^{-11} = \sqrt{\binom{d}{\lfloor d/2\rfloor}}^{-10} \leq \sqrt{\binom{2}{\lfloor 2/2\rfloor}}^{-10} = \frac{1}{32} < \frac{1}{18}.\end{aligned}$$

Thus, by Lemma 3.13 $A(y,n)^{\wedge t_r} = A(x,n)^{\wedge t_r}$ satisfies the cone condition, for $1 \le r \le p-1$, with cones $K_r = \mathcal{C}(e^{\wedge t_r}, 1/3)$ and $K'_r = \mathcal{C}(e^{\wedge t_r}, 1/4)$. $\square$

Theorem 4.5 remains true if we assume X to be one-sided subshift. To see this, we construct an ergodic two-sided extension of X and utilise Theorem 4.5.

Theorem 4.6. *Theorem 4.5 holds also for the one-sided subshifts of finite type with Borel sigma-algebra.*

Proof. In the following we deal with cylinder sets of both the one-sided and two-sided subshifts of finite type. There is no notational distinction between these sets, but the context is always clear from the associated measure.

Let

$$\mathcal{C} = \left\{ \bigcup_{j=1}^{n} [\mathbf{i}_j]_k \colon \mathbf{i}_j \in \widetilde{X}_m^*, k \in \mathbb{Z}, n, m \in \mathbb{N} \right\}$$

be the set of all (finite) unions of cylinder sets of same length and position. It is quite straightforward to check that for any words $\mathbf{i} \in X_{m_1}^*$, $\mathbf{j} \in X_{m_2}^*$, we have $[\mathbf{i}]_{k_1} \cup [\mathbf{j}]_{k_2} \in \mathcal{C}$, $[\mathbf{i}]_{k_1} \setminus [\mathbf{j}]_{k_2} \in \mathcal{C}$ and $[\mathbf{i}]_{k_1} \cap [\mathbf{j}]_{k_2} \in \mathcal{C}$ for any $k_1, k_2 \in \mathbb{Z}$. This clearly implies that $\mathcal{C}$ is closed under union and intersection. Further, it implies that $\mathcal{C}$ is closed under relative complement since for $E_1 = \bigcup_{j=1}^{n} [\mathbf{i}_j]_k \in \mathcal{C}$ and $E_2 = \bigcup_{\ell=1}^{m} [\mathbf{j}_\ell]_h \in \mathcal{C}$,

$$E_1 \setminus E_2 = \bigcup_{j=1}^{n} \left([\mathbf{i}_j]_k \setminus \bigcup_{\ell=1}^{m} [\mathbf{j}_\ell]_h \right) = \bigcup_{j=1}^{n} \bigcap_{\ell=1}^{m} ([\mathbf{i}_j]_k \setminus [\mathbf{j}_\ell]_h).$$

This makes $\mathcal{C}$ a ring of sets that generates $\widetilde{\mathcal{B}}$. We define a set function $\widetilde{\mu} \colon \mathcal{C} \to \mathbb{R}$ by

$$\widetilde{\mu}\left(\bigcup_{j=1}^{n} [\mathbf{i}_j]_k \right) = \mu\left(\bigcup_{j=1}^{n} [\mathbf{i}_j] \right).$$

By Caratheodory's extension theorem, $\widetilde{\mu}$ extends to a Borel measure. From now on, let $\widetilde{\mu}$ be this extension.

For $E = \bigcup_{j=1}^{n} [\mathbf{i}_j]_k \in \mathcal{C}$

$$\widetilde{\mu}(\widetilde{T}^{-1}(E)) = \widetilde{\mu}\left(\bigcup_{j=1}^{n} [\mathbf{i}_j]_{k+1} \right) = \mu\left(\bigcup_{j=1}^{n} [\mathbf{i}_j] \right) = \widetilde{\mu}(E),$$

so $\widetilde{\mu}$ is $\widetilde{T}$-invariant.

Next we show that $\widetilde{\mu}$ is ergodic. Let $E_1 = \bigcup_{j=1}^{n_1}[\mathbf{i}_j]_k \in \mathcal{C}$ and $E_2 = \bigcup_{\ell=1}^{n_2}[\mathbf{j}_\ell]_h \in \mathcal{C}$. Let $s = -\min\{k-1, h-1, 0\}$, so that $k+s, h+s \geq 1$. By ergodicity of μ and Proposition 1.5, we have

$$\begin{aligned}
&\lim_{n\to\infty}\sum_{m=0}^{n-1}\widetilde{\mu}(\widetilde{T}^{-m}(E_1)\cap E_2) = \lim_{n\to\infty}\sum_{m=0}^{n-1}\widetilde{\mu}(\widetilde{T}^{-m-s}(E_1)\cap \widetilde{T}^{-s}(E_2))\\
&= \lim_{n\to\infty}\sum_{m=0}^{n-1}\mu\left(T^{-m}\left(\bigcup_{j=1}^{n_1}[\mathbf{i}_j]_{k+s}\right)\cap\left(\bigcup_{\ell=1}^{n_2}[\mathbf{j}_\ell]_{h+s}\right)\right)\\
&= \mu\left(\bigcup_{j=1}^{n_1}[\mathbf{i}_j]_{k+s}\right)\mu\left(\bigcup_{\ell=1}^{n_2}[\mathbf{j}_\ell]_{h+s}\right) = \widetilde{\mu}(E_1)\widetilde{\mu}(E_2).
\end{aligned}$$

Thus, by Proposition 1.5, $\widetilde{\mu}$ is an ergodic measure.

Next, define the map $\pi\colon \widetilde{X} \to X$,

$$\pi(\dots x_{-1}, x_0, x_1, x_2, \dots) = (x_1, x_2, x_3, \dots).$$

Consider the locally constant cocycle $\widetilde{A}\colon \widetilde{X} \to \mathbb{R}^{d\times d}$, $\widetilde{A}(x) := A(\pi(x))$. By (4.2), (A,μ) and $(\widetilde{A},\widetilde{\mu})$ have the same Lyapunov spectrum since $\widetilde{A}(x,n) = A(\pi(x),n)$. Let $\eta > 0$, $S \subset \mathbb{N}$ and $\{\Sigma_n \subset \widetilde{X}_n^*\colon n \in S\}$ be as in Theorem 4.5 when applied to $(\widetilde{X},\widetilde{\mathcal{B}},\widetilde{\mu},\widetilde{T},\widetilde{A})$. The collection $\{\Sigma_n \subset \widetilde{X}_n^*\colon n \in S\}$ satisfies the claims (1)-(3) of 4.5 also for the original system $(X,\mathcal{B},\mu,T,A)$. Claim (1) follows from the fact that $\widetilde{X}$ and X share the same set of allowed finite words. Second claim is true since $\widetilde{\mu}([\mathbf{i}]) = \mu([\mathbf{i}])$. □

5 Subadditive thermodynamic formalism

Proof of Theorem 2.12 depends on the subadditive version of a theorem called the variational principle. The variational principle is an important result in the area of thermodynamic formalism linking together two of its key concepts, measure theoretic entropy and topological pressure.

5.1 Entropy

Let $(X,\mathcal{B},\mu)$ be a probability space.

Definition 5.1. A finite family $\xi \subset \mathcal{B}$ such that $\mu(\cup_{C\in\xi}C) = 1$ and $\mu(C\cap D) = 0$ for any distinct $C, D \in \xi$ is called a *pre-partition* of $(X,\mathcal{B},\mu)$. We define an equivalence relation $\sim$ on the set of pre-partitions by setting $\eta \sim \xi$ if the elements of η and ξ differ by sets of zero measure, i.e.

(1) For every $D \in \eta$, there exists $C \in \xi$ such that $\mu(D \setminus C) = 0$.

(2) For every $C \in \xi$, there exists $D \in \eta$ such that $\mu(C \setminus D) = 0$.

The equivalence classes of this relation are called *measurable partitions* of X. We define a new partition called *common refinement* of measurable partitions ξ and η by

$$\xi \vee \eta = \{C \cap D \colon C \in \xi, D \in \eta\}.$$

Definition 5.2. Let ξ be a measurable partition of X. The number

$$H_\mu(\xi) = -\sum_{C \in \xi} \mu(C) \log \mu(C)$$

is called the *entropy* of the partition ξ.

Definition 5.3. Let $(X, \mathcal{B}, \mu, T)$ be a ppt. Let ξ be a measurable partition of X and denote by ξ_n the common refinement

$$\xi_n = \bigvee_{k=0}^{n-1} T^{-k}(\xi).$$

Then the limit

$$h_\mu(T, \xi) = \lim_{n \to \infty} \frac{1}{n} H_\mu(\xi_n)$$

exists [4, Proposition 4.2, p. 111]. The number $h_\mu(T, \xi)$ is called the *measure theoretic entropy* of T with respect to μ and the partition ξ. The *measure theoretic entropy* of T with respect to μ is defined as

$$h_\mu(T) = \sup\{h_\mu(T, \xi) \colon \xi \text{ is a measurable partition of } X\}. \tag{5.1}$$

We write h_μ instead of $h_\mu(T)$ when the transfromation is clear from the context.

In the case of symbolic dynamics there is a simple partition giving the supremum in (5.1).

Theorem 5.4 (Kolmogorov-Sinai)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a ppt. If ξ is a partition of X such that the partitions $\{\xi_n\}_{n \in \mathbb{N}}$ generate the sigma algebra $\mathcal{B}$, i.e. $\mathcal{B}$ is the smallest sigma algebra containing every ξ_n, then $h_\mu(T) = h_\mu(T, \xi)$.*

Proof. [4, Theorem 4.3, p. 121] □

Lemma 5.5. *Let $(X, \mathcal{B}, \mu, T)$ be a probability preserving one-sided subshift of finite type with Borel σ-algebra $\mathcal{B}$. For a partition*

$$\xi = \{[i] \subset X \colon i \textit{ is a symbol from the alphabet of } X\},$$

$h_\mu(T) = h_\mu(T, \xi)$.

Proof. It is easy to see that $\xi_n = \{[\mathbf{i}] \subset X \colon |\mathbf{i}| = n\}$. These cylinder sets generate $\mathcal{B}$, so by Theorem 5.4 $h_\mu(T) = h_\mu(T, \xi)$. □

The following theorem, known as the *Shannon-McMillan-Breiman theorem*, is an important result in the areas of information theory and ergodic theory.

Theorem 5.6 (Shannon-McMillan-Breiman)**.** *Let ξ be a measurable partition of an ergodic ppt $(X, \mathcal{B}, \mu, T)$. For μ-almost every $x \in X$ and all $n \in \mathbb{N}$, there is a unique element $\xi_n(x)$ of the partition ξ_n such that $x \in \xi_n(x)$. For a.e. $x \in X$ and in $L^1(X, \mathcal{B}, \mu)$*

$$\lim_{n\to\infty} -\frac{1}{n} \log \mu(\xi_n(x)) = h_\mu(T, \xi).$$

Proof. [19, Theorem 2.3, p. 261] □

Corollary 5.7. *Let ξ be a measurable partition of an ergodic probability preserving transformation $(X, \mathcal{B}, \mu, T)$. Then for a given $\varepsilon > 0$ there is $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$ there exists a subset S_n of ξ_n with the following properties:*

(1) $\mu\big(\bigcup_{C\in S_n} C\big) > 1 - \varepsilon$

(2) for every $C \in S_n$ we have

$$-n(h_\mu(T, \xi) + \varepsilon) < \log \mu(C) < -n(h_\mu(T, \xi) - \varepsilon).$$

Proof. Let $\varepsilon > 0$. By Egorov's and the Shannon-McMillan-Breiman theorems 1.1 and 5.6 there is a set $E \subset X$ of measure greater than $1 - \varepsilon$ such that $-\frac{1}{n} \log \mu(\xi_n(x))$ converges uniformly to $h_\mu(T, \xi)$ on E. Therefore, for large enough n, the measure of a set

$$A_{\varepsilon,n} = \left\{x \in X \colon \left|-\frac{1}{n} \log \mu(\xi_n(x)) - h_\mu(T, \xi)\right| < \varepsilon\right\}$$

is greater than $1 - \varepsilon$. Further, a collection

$$S_n = \{C \in \xi_n \colon C \cap A_{\varepsilon,n} \neq \emptyset\}$$

satisfies property (1) for large n since $\mu(A_{\varepsilon,n}) > 1-\varepsilon$. Every $C \in S_n$ contains a point from $A_{\varepsilon,n}$ and thus $| - \frac{1}{n} \log \mu(C) - h_\mu(T, \xi)| < \varepsilon$. This implies the second property. □

We need corollary 5.7 in the context of symbolic dynamics.

Corollary 5.8. *Let (X, T, μ) be an ergodic probability preserving subshift of finite type. For every $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for every $n > n_0$, there is $\Delta_n \subset X_n^*$ such that*

(1) $\sum_{\mathbf{i} \in \Delta_n} \mu[\mathbf{i}] > 1 - \varepsilon$.

(2) for every $\mathbf{i} \in \Delta_n$

$$-n(h_\mu + \varepsilon) < \log \mu[\mathbf{i}] < -n(h_\mu - \varepsilon).$$

Proof. Let $\xi = \{[i] \subset X : i \text{ is a symbol from the alphabet of } X\}$. The claim follows directly from Corollary 5.7 and Lemma 5.5 $\square$

5.2 Topological pressure and variational principle for subadditive potentials

Definition 5.9. Let (X, T) be a subshift of finite type. A sequence $\mathcal{F} = \{\log f_n\}_{n \in \mathbb{N}}$ of real-valued functions on X such that

$$0 \leq f_{n+m}(x) \leq f_n(x) f_m(T^n x)$$

for all $x \in X$ and $n, m \in \mathbb{N}$ is called a *subadditive potential*. The limit

$$P(T, \mathcal{F}) = \lim_{n \to \infty} \frac{1}{n} \log \left(\sum_{\mathbf{i} \in X_n^*} \sup_{y \in [\mathbf{i}]} f_n(y) \right)$$

is called *the topological pressure of a subadditive potential* $\mathcal{F}$ with respect to T.

The next proposition shows that the topological pressure of a subadditive potential is well-defined.

Proposition 5.10. *The limit of Definition 5.9 exists (possibly as $-\infty$) and the limit may be replaced by infimum.*

Proof. If $\mathbf{i} = (i_1, \ldots, i_{k+m}) \in X_{k+m}^*$, then $(i_1, \ldots, i_k) \in X_k^*$ and $(i_{k+1}, \ldots, i_m) \in X_m^*$. Also

$$\sup_{y \in [\mathbf{i}]} f_k(y) f_m(T^k y) \leq \left(\sup_{y \in [(i_1, \ldots, i_k)]} f_k(y) \right) \left(\sup_{x \in [(i_{k+1}, \ldots, i_m)]} f_m(x) \right).$$

Let $h_n = \sum_{\mathbf{i}\in X_n^*} \sup_{y\in[\mathbf{i}]} f_n(y)$ for $n \in \mathbb{N}$. The sequence h_n is submultiplicative since

$$\begin{aligned}
h_{k+m} &= \sum_{\mathbf{i}\in X_{k+m}^*} \sup_{y\in[\mathbf{i}]} f_{k+m}(y) \le \sum_{\mathbf{i}\in X_{k+m}^*} \sup_{y\in[\mathbf{i}]} f_k(y) f_m(T^k y) \\
&\le \sum_{\mathbf{i}\in X_m^*} \sum_{\mathbf{j}\in X_k^*} \sup_{y\in[\mathbf{j}]} f_k(y) \sup_{x\in[\mathbf{i}]} f_m(x) = \sum_{\mathbf{j}\in X_k^*} \sup_{y\in[\mathbf{j}]} f_k(y) \sum_{\mathbf{i}\in X_m^*} \sup_{x\in[\mathbf{i}]} f_m(x) \\
&= h_k h_m.
\end{aligned}$$

This means that the sequence $\{\log h_n\}_{n\in\mathbb{N}}$ is subadditive. Thus, by Lemma 2.1, the limit

$$P(T,\mathcal{F}) = \lim \frac{1}{n} \log h_n = \inf \frac{1}{n} \log h_n$$

exists. □

Example 5.11. For $g \in C(X)$, $s \ge 0$ and a matrix cocycle $A\colon X \to \mathbb{R}^{d\times d}$, let $f_n(x) = \exp(S_n g(x))\varphi^s(A(x,n))$ for $n \in \mathbb{N}$. Then $\mathcal{F} = \{\log f_n\}_{n\in\mathbb{N}}$ is a subadditive potential by submultiplicativity of φ^s. Moreover, $P(T,\mathcal{F}) = P_g(A,s)$.

Theorem 5.12 (The variational principle)**.** *Let $\mathcal{F} = \{\log f_n\}$ be a subadditive potential on a subshift of finite type $(X,\mathcal{B},T)$ with Borel σ-algebra $\mathcal{B}$. Suppose that the functions f_n are continuous on X. Then*

$$P(T,\mathcal{F}) = \sup_{\mu\in\mathcal{E}} \left(h_\mu(T) + \lim_{n\to\infty} \frac{1}{n} \int \log f_n(x) d\mu(x) \right).$$

Proof. This theorem is proved for arbitrary continuous dynamical system on compact space in [8, Theorem 1.1]. The case of symbolic dynamics is discussed in Section 4 of [8]. □

6 Proofs of the continuity results

Before we prove the Theorems 2.12 and 2.6, the last few preparations must be done.

In the proof of Theorem 2.12 we consider separately the lower and upper semi-continuity of the pressure.

Definition 6.1. Let X be a metric space and let $f\colon X \to \overline{\mathbb{R}}$ be an extended real-valued function. A function f is said to be *upper semi-continuous at* $x_0 \in X$ if

$$\limsup_{x\to x_0} f(x) \le f(x_0).$$

A function f is said to be *lower semi-continuous at* $x_0 \in X$ if

$$\liminf_{x \to x_0} f(x) \geq f(x_0).$$

A function f is called upper semi-continuous (lower semi-continuous) if f is upper semi-continuous (lower semi-continuous) at every $x \in X$.

Recall that the standard topology of the extended real line is defined so that collections $\{[-\infty, a) \colon a \in \mathbb{R}\}$ and $\{(a, \infty] \colon a \in \mathbb{R}\}$ form a local basis of $-\infty$ and ∞, respectively. The continuity of a function can be shown by verifying both upper and lower semi-continuity.

Lemma 6.2. *A function $f \colon X \to \overline{\mathbb{R}}$ is continuous if and only if f is upper and lower semi-continuous.*

The difficult part of Theorems 2.12 and 2.6 is to prove the lower semi-continuity of the pressure. Upper semi-continuity is proved with the help of the following lemma.

Lemma 6.3. *Let $\mathcal{F} = \{\log f_n\}$, $\mathcal{G}^{(k)} = \{\log g_n^{(k)}\}$, $k \in \mathbb{N}$, be subadditive potentials on a subshift of finite type X. If for every $n \in \mathbb{N}$ there exists a decreasing sequence of positive numbers $\delta_k(n) \to 0$ as $k \to \infty$ such that $g_n^{(k)}(x) - f_n(x) \leq \delta_k(n)$ for all $x \in X$ and $k \in \mathbb{N}$, then $\limsup_{k \to \infty} P(T, \mathcal{G}^{(k)}) \leq P(T, \mathcal{F})$.*

Proof. Let $n \in \mathbb{N}$ and write $\delta_k := \delta_k(n)$. Since $g_n^{(k)}(x) \leq \delta_k + f_n(x)$, we have

$$\sum_{\mathbf{i} \in X_n^*} \sup_{y \in [\mathbf{i}]} g_n^{(k)}(y) \leq |X_n^*| \delta_k + \sum_{\mathbf{i} \in X_n^*} \sup_{y \in [\mathbf{i}]} f_n(y).$$

Denote $\varepsilon_k = |X_n^*| \delta_k$. Clearly $\varepsilon_k \searrow 0$ as $k \to \infty$. By Proposition 5.10 the limit in the definition of pressure may be replaced by infimum. Thus, for every $n \in \mathbb{N}$ we have

$$P(T, \mathcal{G}^{(k)}) = \inf_{n' \in \mathbb{N}} \frac{1}{n'} \log \left(\sum_{\mathbf{i} \in X_{n'}^*} \sup_{y \in [\mathbf{i}]} g_{n'}^{(k)}(y) \right) \leq \frac{1}{n} \log \left(\varepsilon_k + \sum_{\mathbf{i} \in X_n^*} \sup_{y \in [\mathbf{i}]} f_n(y) \right).$$

Hence,

$$\limsup_{k \to \infty} P(T, \mathcal{G}^{(k)}) \leq \frac{1}{n} \log \left(\sum_{\mathbf{i} \in X_n^*} \sup_{y \in [\mathbf{i}]} f_n(y) \right)$$

and the claim follows by taking limit as n tends to infinity. $\square$

Lemma 6.4. *Let X be a subshift of finite type with a Borel probability measure μ. Let $B \subset X$ be a measurable set and let*

$$\Delta_n = \{\mathbf{i} \in X_n^* \colon [\mathbf{i}] \cap B = \emptyset\}$$

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

$$\lim_{n\to\infty} \mu\Bigg(\bigcup_{\mathbf{i}\in\Delta_n} [\mathbf{i}]\Bigg) = 1 - \mu(B).$$

Proof. The claim is clear if $B = X$, so we assume that $B \neq X$. Let $A_n := \bigcup_{\mathbf{i}\in\Delta_n} [\mathbf{i}]$ for $n \in \mathbb{N}$. If $x \in A_n$, then $[x|n] \cap B = \emptyset$ which implies that $[x|(n+1)] \cap B = \emptyset$ and $x \in A_{n+1}$. Since $A_i \subset A_{i+1}$ for every i, $A_n = \bigcup_{i=1}^n A_i$. Let $C_1 = A_1$ and $C_i = A_i \setminus A_{i-1}$ for $i \geq 2$. The sets C_i are pairwise disjoint and $\bigcup_{i=1}^n C_i = A_n$. Thus,

$$\lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \mu\Bigg(\bigcup_{i=1}^n C_i\Bigg) = \sum_{i=1}^\infty \mu(C_i) = \mu\Bigg(\bigcup_{i=1}^\infty C_i\Bigg) = \mu\Bigg(\bigcup_{i=1}^\infty A_i\Bigg). \tag{6.1}$$

Let $x \in X \setminus B$. Then $[x|m] \cap B = \emptyset$ for some $m \in \mathbb{N}$ since otherwise $x \in \bigcap_{n=1}^\infty ([x|n] \cap B) \subset B$. Thus, $x|m \in \Delta_m$ implying that $x \in \bigcup_{i=1}^\infty A_i$. Therefore, $X \setminus B \subset \bigcup_{i=1}^\infty A_i$. The inclusion $\bigcup_{i=1}^\infty A_i \subset X \setminus B$ is clear. By (6.1),

$$\lim_{n\to\infty} \mu\Bigg(\bigcup_{\mathbf{i}\in\Delta_n} [\mathbf{i}]\Bigg) = \mu\Bigg(\bigcup_{i=1}^\infty A_i\Bigg) = \mu(X \setminus B) = 1 - \mu(B).$$

□

Lemma 6.5. *Let $B \in \mathbb{R}^{d\times d}$. If m and n are integers with $0 \leq m < n \leq d$, then for any $s \in [m,n]$, we have*

$$\varphi^s(B) \geq (\varphi^m(B))^{\frac{n-s}{n-m}} (\varphi^n(B))^{\frac{s-m}{n-m}}.$$

Especially, $\varphi^s(B) \geq (\varphi^n(B))^{\frac{s}{n}}$.

Proof. Let $p = \lfloor s \rfloor$, and let $\alpha_1, \ldots, \alpha_d$ be the singular values of B in decreasing order. We have an inequality $\varphi^n(B) = \alpha_1 \cdots \alpha_n \leq \alpha_1 \cdots \alpha_p (\alpha_{p+1})^{n-p}$. Applying this, we have

$$\begin{aligned}
&\frac{(\varphi^s(B))^{n-m}}{(\varphi^m(B))^{n-s}(\varphi^n(B))^{s-m}} \geq \frac{(\alpha_1 \cdots \alpha_p (\alpha_{p+1})^{s-p})^{n-m}}{(\alpha_1 \cdots \alpha_m)^{n-s} (\alpha_1 \cdots \alpha_p (\alpha_{p+1})^{n-p})^{s-m}} \\
=&\frac{(\alpha_1 \cdots \alpha_m)^{n-m} (\alpha_{m+1} \cdots \alpha_p)^{n-m}}{(\alpha_1 \cdots \alpha_m)^{(n-s)+(s-m)} (\alpha_{m+1} \cdots \alpha_p)^{s-m} (\alpha_{p+1})^{(n-p)(s-m)-(s-p)(n-m)}} \\
=&\frac{(\alpha_{m+1} \cdots \alpha_p)^{n-s}}{(\alpha_{p+1})^{(n-s)(p-m)}} \geq 1.
\end{aligned}$$

Thus, $\varphi^s(B) \geq (\varphi^m(B))^{\frac{n-s}{n-m}}(\varphi^n(B))^{\frac{s-m}{n-m}}$. The latter claim is obtained by letting $m = 0$. □

In the proof of Theorem 2.12 we use the continuity of additive pressure. The *additive pressure* for continuous *potential* $\phi\colon X \to \mathbb{R}$ is

$$P(\phi) = \lim_{n\to\infty} \frac{1}{n} \log\left(\sum_{\mathbf{i}\in X_n^*} \sup_{y\in[\mathbf{i}]} \exp(S_n\phi(y))\right). \tag{6.2}$$

We extend this definition to cover all potentials $\phi\colon X \to \mathbb{R}$ such that the limit (6.2) exists, not just the continuous potentials.

Lemma 6.6. *The additive pressure $P(\phi)$ is a continuous function of a potential ϕ on L^∞-topology.*

Proof. Let ϕ be such a potential that $P(\phi)$ exists. Let $\varepsilon > 0$. Let $\mathcal{U}$ be a ϕ-centred ball of radius ε. Let $\psi \in \mathcal{U}$. Since $\|\phi-\psi\|_\infty = \sup_{x\in X}|\phi(x)-\psi(x)| < \varepsilon$,

$$S_n\phi(y) - n\varepsilon < S_n\psi(y) < S_n\phi(y) + n\varepsilon$$

for every $y \in X$. Thus,

$$P(\phi) - \varepsilon \leq P(\psi) \leq P(\phi) + \varepsilon,$$

which implies the lower and upper semi-continuity at ϕ. □

We are finally ready to prove the continuity theorems 2.12 and 2.6.

Proof of Theorem 2.12, claim (1). Let $A \in \mathcal{M}(X,d)$ be a locally constant cocycle. The claim is trivially true for the case $s = 0$ since $P_g(\cdot, 0)$ is a constant. Assume first that $s \geq d$. Let $P(\phi_A)$ be the additive pressure of a potential $\phi_A(x) = g(x) + \log|\det(Ax)|^{s/d}$. Then $P(\phi_A) = P_g(A,s)$. Indeed,

$$S_n\phi_A(x) = S_ng(x) + \sum_{i=0}^{n-1} \log|\det A(T^ix)|^{s/d} = S_ng(x) + \log|\det A(x,n)|^{s/d},$$

and thus

$$\begin{aligned} P(\phi_A) &= \lim_{n\to\infty} \frac{1}{n} \log\left(\sum_{\mathbf{i}\in X_n^*} \sup_{y\in[\mathbf{i}]} \exp(S_ng(y) + \log|\det A(y,n)|^{s/d})\right) \\ &= \lim_{n\to\infty} \frac{1}{n} \log\left(\sum_{\mathbf{i}\in X_n^*} \sup_{y\in[\mathbf{i}]} \exp(S_ng(y))|\det A(y,n)|^{s/d}\right) = P_g(A,s). \end{aligned}$$

Note that ϕ_A is a continuous potential since A is locally constant, but the above discussion shows that $P(\phi_B)$ exists for any cocycle B since $P_g(B,s)$ exists. The map $A \mapsto \phi_A$ is continuous in the L^∞- topology. By Lemma 6.6, the map $\phi \mapsto P(\phi)$ is continuous. Putting this together gives that the map $A \mapsto P(\phi_A) = P_g(A,s)$ is continuous for $s \geq d$.

Assume now that $0 < s < d$. We first show that $P_g(\cdot,s)$ is upper semi-continuous at A. Let B_k be a sequence of matrix cocycles, from some bounded neighbourhood of A, converging to A. Consider the potentials $\mathcal{F} = \{\log f_n\}$, $\mathcal{G}^{(k)} = \{\log g_n^{(k)}\}$ with functions

$$\begin{aligned} f_n(x) &= \exp(S_n g(x))\varphi^s(A(x,n)) \\ g_n^{(k)}(x) &= \exp(S_n g(x))\varphi^s(B_k(x,n)). \end{aligned}$$

For every $n \in \mathbb{N}$, $\exp(S_n g(\cdot))$ is a bounded function by the compactness of X and the continuity of g. Since $B_k \to A$, $\varphi^s(B_k(x,n)) \to \varphi^s(A(x,n))$ uniformly as $k \to \infty$ for every $n \in \mathbb{N}$. Therefore, for every $n \in \mathbb{N}$, there exists numbers $\delta_k(n) > 0$, $k \in \mathbb{N}$ such that $\delta_k(n) \searrow 0$ as $k \to \infty$, and $g_n^{(k)}(x) - f_n(x) \leq \delta_k(n)$ for every $x \in X$. Thus, by Example 5.11 and Lemma 6.3

$$\limsup_{k\to\infty} P_g(B_k,s) = \limsup_{k\to\infty} P(T,\mathcal{G}^{(k)}) \leq P(T,\mathcal{F}) = P_g(A,s).$$

This proves the upper semi-continuity.

For the proof of the lower semi-continuity fix $\varepsilon > 0$. If $P_g(A,s) = -\infty$, the lower semi-continuity is clear, so assume that $P_g(A,s) > -\infty$. By the variational principle 5.12, there exists an invariant ergodic measure $\mu \in \mathcal{E}$ such that

$$\begin{aligned} P_g(A,s) - \varepsilon &= P(T,\mathcal{F}) - \varepsilon \\ &\leq h_\mu + \lim_{n\to\infty} \frac{1}{n} \int \log(\exp(S_n g(x))\varphi^s(A(x,n)))d\mu(x) \\ &= h_\mu + \lim_{n\to\infty} \frac{1}{n} \int \left[S_n g(x) + \log \varphi^s(A(x,n))\right] d\mu(x) \\ &= h_\mu + \int g(x)d\mu(x) + \lim_{n\to\infty} \frac{1}{n} \int \log \varphi^s(A(x,n))d\mu(x), \end{aligned}$$

where the last equality is obtained by Birkhoff's ergodic theorem 1.6. Note that g is integrable since X is compact and $g \in C(X)$. We write $\mu(g) := \int g(x)d\mu(x)$, $E_\mu(A,s) := \lim_{n\to\infty} \frac{1}{n} \int \log \varphi^s(A(x,n))d\mu(x)$, and rewrite the inequality above in a form

$$P_g(A,s) - \varepsilon \leq h_\mu + \mu(g) + E_\mu(A,s). \tag{6.3}$$

As in Theorem 4.5, let $(\lambda_i, d_i)_{i=1}^p$ be the Lyapunov spectrum of (A, μ) with λ_i in decreasing order. Let $t_m = \sum_{i=1}^m d_i$ and $\Gamma_m = \sum_{i=1}^m d_i\lambda_i$ for $m = 1, \dots, p$. Set $t_0 = \Gamma_0 = 0$. Let $r \in \{0, 1, \dots, p-1\}$ be such that $t_r < s \le t_{r+1}$. Since A is locally constant, we have $\int \log^+ \|A(x)\| d\mu(x) < \infty$. By Lemma 4.3,

$$E_\mu(A, s) = \Gamma_r + (s - t_r)\lambda_{r+1}. \tag{6.4}$$

This implies that $\lambda_{r+1} > -\infty$, for otherwise we would have $P_g(A, s) = -\infty$.

Let η, S and $\{\Sigma_n : n \in S\}$ be as in Theorem 4.5. We show that there exists arbitrarily large $n \in S$ and a subset $\Delta_n \subset \Sigma_n \subset X_n^*$ such that

$$\sum_{\mathbf{i} \in \Delta_n} \mu[\mathbf{i}] > \frac{\eta}{2} \tag{6.5}$$

and if $x|n \in \Delta_n$, then

$$S_n g(x) > n(\mu(g) - \varepsilon), \tag{6.6}$$

$$\mu[x|n] < \exp(n(\varepsilon - h_\mu)), \tag{6.7}$$

$$\exp(n(\Gamma_m - \varepsilon)) < \varphi^{t_m}(A(x, n)) < \exp(n(\Gamma_m + \varepsilon)),\ m = 1, \dots, p-1, \tag{6.8}$$

$$\varphi^{t_p}(A(x, n)) > \exp(n(\Gamma_p - \varepsilon)) \ \text{ if } \lambda_p \ne -\infty. \tag{6.9}$$

Let $\varepsilon' = \min\{\eta/7, \varepsilon\}$. Remember that $\sum_{\mathbf{i} \in \Sigma_n} \mu[\mathbf{i}] = \mu\big(\bigcup_{\mathbf{i} \in \Sigma_n} [\mathbf{i}]\big) > \eta$. By Corollary 5.8, there exists $n_1 \in \mathbb{N}$ such that for every $n \ge n_1$ there exists $\Delta_n^{(0)} \subset X_n^*$ such that $\sum_{\mathbf{i} \in \Delta_n^{(0)}} \mu[\mathbf{i}] > 1 - \varepsilon'$, and for every $\mathbf{i} \in \Delta_n^{(0)}$

$$-n(h_\mu + \varepsilon) \le -n(h_\mu + \varepsilon') < \log \mu[\mathbf{i}] < -n(h_\mu - \varepsilon') \le -n(h_\mu - \varepsilon). \tag{6.10}$$

Let $\Delta_n^{(1)} = \Sigma_n \cap \Delta_n^{(0)}$, for $n \in S$ with $n \ge n_1$. By Lemma 4.4, and by noting that for distinct $\mathbf{i}, \mathbf{j} \in X_n^*$ the intersection $[\mathbf{i}] \cap [\mathbf{j}]$ is empty, we have

$$\begin{aligned}\sum_{\mathbf{i} \in \Delta_n^{(1)}} \mu[\mathbf{i}] &= \mu\left(\left(\bigcup_{\mathbf{i} \in \Delta_n^{(0)}} [\mathbf{i}]\right) \cap \left(\bigcup_{\mathbf{j} \in \Sigma_n} [\mathbf{j}]\right)\right) \\ &\ge \mu\left(\bigcup_{\mathbf{i} \in \Delta_n^{(0)}} [\mathbf{i}]\right) + \mu\left(\bigcup_{\mathbf{j} \in \Sigma_n} [\mathbf{j}]\right) - 1 \\ &\ge 1 - \varepsilon' + \eta - 1 = \eta - \varepsilon'.\end{aligned}$$

By Birkhoff's ergodic theorem 1.6, $S_n g(x)/n$ converges to $\mu(g)$ as n tends to infinity for almost every $x \in X$. By Egorov's theorem 1.1 there exists a set $B \subset X$, with $\mu(B) < \varepsilon'$, such that $S_n g(\cdot)/n$ converges uniformly to $\mu(g)$ on $X \setminus B$. Thus, there exists $n_2 \in \mathbb{N}$ such that for any $y \in X \setminus B$ we have

$$\mu(g) - \varepsilon < S_n g(y)/n, \tag{6.11}$$

for all $n \geq n_2$. For every $n \geq n_2$ define a set $\Delta_n^{(2)} = \{\mathbf{i} \in X_n^* \colon [\mathbf{i}] \cap B = \emptyset\}$. By Lemma 6.4, $\lim_{n\to\infty} \mu(\bigcup_{\mathbf{i}\in\Delta_n^{(2)}}[\mathbf{i}]) > 1 - \varepsilon'$. We may assume that n_2 is picked large enough such that $\mu(\bigcup_{\mathbf{i}\in\Delta_n^{(2)}}[\mathbf{i}]) > 1 - \varepsilon'$ for every $n \geq n_2$. Let

$$\Delta_n^{(3)} = \Delta_n^{(1)} \cap \Delta_n^{(2)}$$

for $n \in S$ with $n \geq \max\{n_1, n_2\}$. As above, we estimate by Lemma 4.4

$$\sum_{\mathbf{i}\in\Delta_n^{(3)}} \mu[\mathbf{i}] \geq \mu\bigg(\bigcup_{\mathbf{i}\in\Delta_n^{(1)}} [\mathbf{i}]\bigg) + \mu\bigg(\bigcup_{\mathbf{j}\in\Delta_n^{(2)}} [\mathbf{j}]\bigg) - 1 \geq \eta - \varepsilon' + 1 - \varepsilon' - 1 = \eta - 2\varepsilon'.$$

By Lemma 4.3, for all $m = 1, \dots, p$,

$$\lim_{n\to\infty} \frac{1}{n} \int \log \varphi^{t_m}(A(x,n)) d\mu(x) = \Gamma_{m-1} + (t_m - t_{m-1})\lambda_m = \Gamma_m. \tag{6.12}$$

By Kingman's subadditive ergodic theorem $\log \varphi^{t_m}(A(x,n))/n$ converges to a constant value for μ-almost every $x \in X$ and this constant is given by the limit of (6.12). Thus, $\lim_{n\to\infty} \frac{1}{n} \log \varphi^{t_m} A(x,n) = \Gamma_m$ for μ-almost every $x \in X$. Again, by Egorov's theorem the convergence is uniform, for all $m = 1, \dots, p$, on some set $X \setminus C$, where C is a set of measure less than ε'. Thus, for some $n_3 \in \mathbb{N}$,

$$\left| \lim_{n\to\infty} \frac{1}{n} \log(\varphi^{t_m}(A(x,n))) - \Gamma_m \right| < \varepsilon \tag{6.13}$$

for every $x \in X \setminus C$, $m = 1, \dots, p-1$, $n \geq n_3$ and also for $m = p$ if $\lambda_p \neq -\infty$. For $n \geq n_3$, define a set $\Delta_n^{(4)} = \{x|n \colon x \in X \setminus C\}$. Finally, let $n \in S$ be such that $n \geq \max\{n_1, n_2, n_3\}$ and define the set

$$\Delta_n = \Delta_n^{(3)} \cap \Delta_n^{(4)}.$$

Let $x \in X$ be such that $x|n \in \Delta_n$. Then $x \in X \setminus B$ since $x|n \in \Delta_n^{(2)}$. Thus, (6.6) holds as stated in (6.11). Inequality (6.10) implies (6.7). There exists $y \in X \setminus C$ such that $x|n = y|n$. Since $A(x,n) = A(y,n)$, inequalities (6.8) and (6.9) follow from (6.13). Also,

$$\sum_{\mathbf{i}\in\Delta_n} [\mathbf{i}] \geq \mu\bigg(\bigcup_{\mathbf{i}\in\Delta_n^{(3)}} [\mathbf{i}]\bigg) + \mu\bigg(\bigcup_{\mathbf{j}\in\Delta_n^{(4)}} [\mathbf{j}]\bigg) - 1 \geq \eta - 3\varepsilon' \geq 4\eta/7 > \eta/2,$$

which proves (6.5).

It follows from (6.5) and (6.7) that $\eta/2 < \sum_{\mathbf{i}\in\Delta_n}\mu[\mathbf{i}] < |\Delta_n|\exp(n(\varepsilon - h_\mu))$. Thus,

$$|\Delta_n| > \frac{\eta}{2}\exp(n(h_\mu - \varepsilon)). \tag{6.14}$$

Denote by $\Delta_n^\ell \subset X_{\ell n}^*$ the family of juxtapositions of ℓ words in Δ_n, i.e.

$$\Delta_n^\ell = \{(x_1,\dots,x_{\ell n}) \in X_{\ell n}^* \colon (x_{in+1},\dots,x_{(i+1)n}) \in \Delta_n, i = 0,\dots,\ell-1\}.$$

Note that $|\Delta_n^\ell| = |\Delta_n|^\ell$. Combining this with (6.14) gives

$$|\Delta_n^\ell| \geq \left(\frac{\eta}{2}\right)^\ell \exp(\ell n(h_\mu - \varepsilon)). \tag{6.15}$$

Consider the case when $p = 1$. This means that the cocycle A has only one Lyapunov exponent $\lambda = \lambda_1$ with respect to the measure μ. Since we assumed $P_g(A,s) \neq -\infty$, we have $\lambda \neq -\infty$. For any $M \in \mathbb{R}^{d\times d}$, we have $\varphi^{t_1}(M) = |\det(M)|$ since $t_1 = d$. The determinant of a matrix depends continuously on the entries of the matrix. Thus, if $B \in \mathcal{M}(X,d)$ is close to A, that is

$$\|A - B\|_\infty = \sup_{x\in X}\|A(x) - B(x)\|$$

is small, then the entries of the matrices $A(x)$ and $B(x)$ are close to each other for all $x \in X$ and further, the determinants are close. Hence, we can find an L^∞-neighbourhood $\mathcal{U}$ of A such that $\varphi^d(B(x)) \geq e^{-\varepsilon}\varphi^d(A(x))$ for all $B \in \mathcal{U}$ and $x \in X$. Thus, by multiplicativity of determinant and (6.9),

$$\varphi^d(B(x,n)) \geq e^{-n\varepsilon}\varphi^d(A(x,n)) > e^{-n\varepsilon}e^{n(d\lambda-\varepsilon)} = e^{nd(\lambda-2\varepsilon/d)} \geq e^{nd(\lambda-2\varepsilon)}$$

if $x|n \in \Delta_n$. Therefore, if $x|\ell n \in \Delta_n^\ell$, then

$$\varphi^d(B(x,\ell n)) = \prod_{k=0}^{\ell-1}\varphi^d(B(T^{kn}x,n)) \geq e^{\ell nd(\lambda-2\varepsilon)}.$$

Using Lemma 6.5, we have

$$\varphi^s(B(x,\ell n)) \geq (\varphi^d(B(x,\ell n)))^{s/d} \geq e^{\ell nd(\lambda-2\varepsilon)\cdot s/d} = e^{\ell ns(\lambda-2\varepsilon)} \tag{6.16}$$

for $x|\ell n \in \Delta_n^\ell$.

Since $p = 1$, it follows from equation (6.4) that $E_\mu(A,s) = s\lambda$. Thus, by (6.3),

$$P_g(A,s) \leq h_\mu + \mu(g) + \lambda s + \varepsilon. \tag{6.17}$$

For a given $x \in X^*$, let $\bar{x} \in X$ be any infinite word starting with x. If $x \in \Delta_n^\ell$, then by (6.6) we have

$$S_{\ell n} g(\bar{x}) = \sum_{i=0}^{\ell-1} S_n g(T^{in}\bar{x}) > \ell n(\mu(g) - \varepsilon). \tag{6.18}$$

Using this, (6.16), (6.15) and (6.17), we have

$$\begin{aligned}
P_g(B,s) =& \lim_{m\to\infty} \frac{1}{m} \log\left(\sum_{\mathbf{i}\in X_m^*} \sup_{y\in[\mathbf{i}]} \exp(S_m g(y))\varphi^s(B(y,m))\right)\\
\geq& \limsup_{\ell\to\infty} \frac{1}{\ell n} \log\left(\sum_{x\in\Delta_n^\ell} \exp(S_{\ell n} g(\bar{x}))\varphi^s(B(\bar{x},\ell n))\right)\\
\geq& \limsup_{\ell\to\infty} \frac{1}{\ell n} \log\left(\sum_{x\in\Delta_n^\ell} \exp(\ell n(\mu(g)-\varepsilon))\exp(\ell n s(\lambda - 2\varepsilon))\right)\\
\geq& \limsup_{\ell\to\infty} \frac{1}{\ell n} \log\left(\left(\frac{\eta}{2}\right)^\ell \exp(\ell n(h_\mu - \varepsilon))\exp(\ell n(\mu(g)-\varepsilon))\exp(\ell n s(\lambda - 2\varepsilon))\right)\\
=& h_\mu - \varepsilon + \mu(g) - \varepsilon + s(\lambda - 2\varepsilon) + \frac{1}{n}\log\eta/2\\
=& h_\mu + \mu(g) + \lambda s + \varepsilon - [(3+2s)\varepsilon - \frac{1}{n}\log\eta/2]\\
\geq& P_g(A,s) - [(3+2s)\varepsilon + C/n],
\end{aligned} \tag{6.19}$$

where $C = -\log\eta/2 > 0$ is independent of B, $\mathcal{U}$ and n. This proves the lower semi-continuity for the case p=1 since $(3+2s)\varepsilon + C/n$ can be made arbitrarily small.

Now consider the case $p > 1$. By Theorem 4.5, the families

$$\{A(y,n)^{\wedge t_m} : y|n \in \Delta_n, y \in X\}$$

satisfy the cone condition with some cones $\widehat{K}_m, \widehat{K}'_m$ for $m = 1, \ldots, p-1$. Fix new cones K_m and K'_m such that $\widehat{K}'_m \setminus \{0\} \subset \mathrm{int}(K'_m)$, $K'_m \setminus \{0\} \subset \mathrm{int}(K_m)$ and $K_m \setminus \{0\} \subset \mathrm{int}(\widehat{K}_m)$ for $m = 1, \ldots, p-1$. Since the cone condition is robust (see Proposition 3.3) and A is locally constant, there is a L^∞-neighbourhood $\mathcal{U}$ of A (depending on n) such that for every $B \in \mathcal{U}$, $\{B(y,n)^{\wedge t_m} : y|n \in \Delta_n, y \in X\}$ satisfies the cone condition with the cones K_m, K'_m. Note that the cones $\widehat{K}_m, \widehat{K}'_m$ do not depend on n and hence do not K_m and K'_m. The singular values of $A(x,n)$ and $B(x,n)$ are close to each other for all $x \in X$ if A and B are sufficiently close to each other in

L^∞-topology, where being sufficiently close depends on n. Thus, by taking $\mathcal{U}$ small enough, inequalities (6.8) and (6.9) hold for every $B \in \mathcal{U}$ in place of A:

$$\exp(n(\Gamma_m - \varepsilon)) < \varphi^{t_m}(B(x,n)) < \exp(n(\Gamma_m + \varepsilon)), \ 1 \le m \le p-1 \quad (6.20)$$
$$\varphi^{t_p}(B(x,n)) > \exp(n(\Gamma_p - \varepsilon)) \ \text{ if } \lambda_p \neq -\infty. \quad (6.21)$$

If $x|\ell n \in \Delta_n^\ell$ and $B \in \mathcal{U}$, then by writing

$$\varphi^{t_m}(B(x,\ell n)) = \|B(x,\ell n)^{\wedge t_m}\| = \left\| \prod_{k=0}^{\ell-1} B(T^{(\ell-1-k)n}x, n)^{\wedge t_m} \right\|$$

and applying Lemma 3.5 and inequality (6.20), we have

$$c^\ell \exp(\ell n(\Gamma_m - \varepsilon)) \le \varphi^{t_m}(B(x,\ell n)) \le \exp(\ell n(\Gamma_m + \varepsilon)) \quad (6.22)$$

for $m = 1, \dots, p-1$, where $0 < c \le 1$ does not depend on n, $\mathcal{U}$ or B. In the case $m = p$, we have $t_m = t_p = d$ and $\varphi^{t_p}(B(x,\ell n)) = |\det(B(x,\ell n))| = \prod_{k=0}^{\ell-1} \varphi^{t_p}(B(T^{kn}x,n))$. By (6.21), we have

$$\varphi^{t_p}(B(x,\ell n)) \ge \exp(\ell n(\Gamma_p - \varepsilon)) \quad (6.23)$$

for $B \in \mathcal{U}$ and $x|\ell n \in \Delta_n^\ell$. Recall, that $t_m = \sum_{i=1}^m d_i$ and $r \in \{0,1,\dots,p-1\}$ is such that $t_r < s \le t_{r+1}$. By Lemma 6.5, (6.22) and (6.23), we have

$$\begin{aligned}
\varphi^s(B(x,\ell n)) \ge & \big(\varphi^{t_r}(B(x,\ell n))\big)^{\frac{t_{r+1}-s}{t_{r+1}-t_r}} \big(\varphi^{t_{r+1}}(B(x,\ell n))\big)^{\frac{s-t_r}{t_{r+1}-t_r}} \\
\ge & c^\ell \exp\left(\frac{t_{r+1}-s}{d_{r+1}} \ell n(\Gamma_r - \varepsilon) + \frac{s-t_r}{d_{r+1}} \ell n(\Gamma_{r+1} - \varepsilon) \right) \\
= & c^\ell \exp\left(\frac{\ell n}{d_{r+1}} \Big[t_{r+1}\Gamma_r - s\Gamma_r - t_r\Gamma_{r+1} + s\Gamma_{r+1} - \varepsilon d_{r+1} \Big] \right) \\
= & c^\ell \exp\left(\frac{\ell n}{d_{r+1}} \Big[(t_r + d_{r+1})\Gamma_r - s\Gamma_r + (s-t_r)(\Gamma_r + d_{r+1}\lambda_{r+1}) - \varepsilon d_{r+1} \Big] \right) \\
= & c^\ell \exp\left(\frac{\ell n}{d_{r+1}} \Big[d_{r+1}\Gamma_r + d_{r+1}(s-t_r)\lambda_{r+1} - \varepsilon d_{r+1} \Big] \right) \\
= & c^\ell \exp(\ell n(\Gamma_r + (s-t_r)\lambda_{r+1} - \varepsilon)). \qquad (6.24)
\end{aligned}$$

By combining (6.18), (6.24), (6.3) and (6.4), we have

$$\begin{aligned}
& \exp(S_{\ell n} g(x)) \varphi^s(B(x,\ell n)) \\
& \quad \ge \exp(\ell n(\mu(g) - \varepsilon)) \cdot c^\ell \exp(\ell n[\Gamma_r + (s-t_r)\lambda_{r+1} - \varepsilon]) \\
& \quad \ge c^\ell \exp\big(\ell n[P_g(A,s) - E_\mu(A,s) - h_\mu - 2\varepsilon]\big) \exp\big(\ell n[E_\mu(A,s) - \varepsilon]\big) \\
& \quad = c^\ell \exp\big(\ell n[P_g(A,s) - h_\mu - 3\varepsilon]\big) \qquad (6.25)
\end{aligned}$$

for $x|\ell n \in \Delta_n^\ell$.

Again, for given $x \in X^*$, let $\bar{x} \in X$ be any infinite word starting with x. By (6.15) and (6.25), for every $B \in \mathcal{U}$ we have

$$\begin{aligned}
P_g(B,s) \geq &\limsup_{\ell\to\infty} \frac{1}{\ell n} \log\left(\sum_{x\in\Delta_n^\ell} \exp(S_{\ell n} g(\bar{x}))\varphi^s(B(\bar{x},\ell n)) \right) \\
\geq &\limsup_{\ell\to\infty} \frac{1}{\ell n} \log\left(\left(\frac{\eta}{2}\right)^l \exp\big(\ell n(h_\mu - \varepsilon)\big) c^\ell \exp\big(\ell n[P_g(A,s) - h_\mu - 3\varepsilon]\big) \right) \\
= &h_\mu - \varepsilon + P_g(A,s) - h_\mu - 3\varepsilon + \frac{1}{n}\big(\log(\eta/2) + \log c\big) \\
= &P_g(A,s) - (4\varepsilon + C/n),
\end{aligned} \tag{6.26}$$

where $C = -\log(\eta/2) - \log c > 0$ is independent of $B, \mathcal{U}$ and n. This implies the lower semi-continuity for the case $p > 1$ since $(4\varepsilon + C/n)$ can be made arbitrarily small. $\square$

Proof of Theorem 2.12, claim (2). We first consider the case $s \notin \{0,\ldots,d-1\}$. Let $A \in \mathcal{M}(X,d)$ be locally constant. If $s > d$, then $P_g(A,s) = P(\phi_A^s)$ for the additive potential

$$\phi_A^s(x) = g(x) + \log|\det A(x,n)|^{s/d}.$$

Again, if (B,s') is close enough to (A,s), then $P_g(B,s') = P(\phi_B^{s'})$ and $\phi_B^{s'}$ is close to ϕ_A^s. Thus, the continuity of additive pressure implies the continuity of $P_g(\cdot,\cdot)$ as in part (1). The same reasoning shows that P_g is jointly right continuous at (A,d), where the right continuity is, naturally, with respect to the latter coordinate.

Let $s \leq d$. To see the upper semi-continuity of P_g at (A,s), let (B_k, s'_k) be a sequence converging to (A,s) and consider the subadditive potentials

$$\begin{aligned}
f_n &= \exp(S_n g(x))\varphi^s(A(x,n)) \\
g_n^{(k)} &= \exp(S_n g(x))\varphi^{s'_k}(B_k(x,n)).
\end{aligned}$$

The upper semi-continuity, in the case $s \notin \{0,\ldots,d-1\}$, is then obtained by observing that $\varphi^{s'_k}(B_k(x,n)) \to \varphi^s(A(x,n))$ as $k \to \infty$ for every $n \in \mathbb{N}$ and using Lemma 6.3 in a similar manner as in part (1). Note that even though $\varphi^{s'_k}(B_k(x,n))$ does not necessary converge to $\varphi^s(A(x,n))$ as $k \to \infty$ for $s \in \{0,\ldots,d-1\}$, it does for $s = d$.

We next prove the lower semi-continuity. Let $s < d$. There is $0 \leq r \leq p-1$ such that $t_r < s < t_{r+1}$. Let I be a neighbourhood of s such that $t_r < s' < t_{r+1}$ for every $s' \in I$. We need to consider separately the cases $p = 1$ and $p > 1$.

Assume first that $p = 1$. Following the reasoning of part (1), we find a $L^\infty-$neighbourhood $\mathcal{U}$ of A such that

$$\varphi^{s'}(B(x,\ell n)) \geq e^{\ell n s'(\lambda-2\varepsilon)}$$

for all $x|\ell n \in \Delta_n^\ell$ and $B \in \mathcal{U}$ and $s' \in I$, cf. (6.16). By calculating as in (6.19), we obtain

$$\begin{aligned} P_g(B,s') &\geq h_\mu + \mu(g) - 2\varepsilon + s'(\lambda - 2\varepsilon) + \frac{1}{n}\log(\eta/2) \\ &\geq P_g(A,s) - \lambda s - 3\varepsilon + s'(\lambda - 2\varepsilon) + \frac{1}{n}\log(\eta/2) \\ &\geq P_g(A,s) + (s'-s)\lambda - [(3+2s')\varepsilon - \frac{1}{n}\log(\eta/2)]. \end{aligned}$$

Here $\log(\eta/2)$ is again a constant independent of $B, s', I, \mathcal{U}$ and n. This proves the lower semi-continuity for the case $p = 1$.

Assume now, that $p > 1$. Again, proceeding as in part (1), we can find a neighbourhood $\mathcal{U} \subset \mathcal{M}(X,d)$ of A such that

$$\varphi^{s'}(B(x,\ell n)) \geq c^\ell \exp(\ell n(\Gamma_r + (s'-t_r)\lambda_{r+1}) - \varepsilon) \tag{6.27}$$

for every $B \in \mathcal{U}$, $x|\ell n \in \Delta_n^\ell$ and $s' \in I$. This is obtained in a similar manner as (6.24). Noting that $\Gamma_r + (s'-t_r)\lambda_{r+1} = E_\mu(A,s) + (s'-s)\lambda_{r+1}$ by (6.4), and calculating as in (6.25), we obtain

$$\begin{aligned} &\exp(S_{\ell n} g(x))\varphi^{s'}(B(x,\ell n)) \\ \geq &c^\ell \exp\big(\ell n[P_g(A,s) + (s'-s)\lambda_{r+1} - h_\mu - 3\varepsilon]\big). \end{aligned} \tag{6.28}$$

Using (6.28) and following the steps of (6.26), we have

$$P_g(B,s') \geq P_g(A,s) + (s'-s)\lambda_{r+1} - (4\varepsilon + C/n),$$

where $C = -(\log(\eta/2) + \log c) > 0$ is independent of B, s', I, $\mathcal{U}$ and n. Thus, the pressure map is lower semi-continuous, and therefore continuous, at (A,s). The same reasoning shows that the pressure P_g is jointly left continuous at (A,d). This, together with the right continuity, implies the continuity at (A,d).

Next, assume that $0 \leq s < d$ and A is a locally constant cocycle taking values in $GL_d(\mathbb{R})$. Since A is locally constant, the set $\{A(x)\colon x \in X\}$ is finite. Let $C = 2\max_{x\in X}\|A(x)\| > 0$ and $D = \frac{1}{2}\min_{x\in X} 1/\|A(x)^{-1}\| > 0$. By taking neighbourhood $\mathcal{U}$ of A small enough, for any $x \in X$ and $B \in \mathcal{U}$ we have

$$D < \frac{1}{\|B(x)^{-1}\|} \leq \|B(x)\| < C,$$

where the inequality in the middle follows from the fact that $1 = \|I\| = \|MM^{-1}\| \leq \|M\|\|M^{-1}\|$ for every $M \in GL_d(\mathbb{R})$. Thus, for $n \in \mathbb{N}$,

$$D^n < \frac{1}{\|B(x,n)^{-1}\|} \leq \|B(x,n)\| < C^n.$$

This and Proposition 1.18 implies that all the singular values of $B(x,n)$ are between D^n and C^n. Assume that $s < s' < d$, and let $h = \lfloor s' \rfloor$ and $k = \lfloor s \rfloor$. Denote the singular values of $B(x,n)$ by $\alpha_1 \geq \cdots \geq \alpha_d$. Then

$$\begin{aligned}\frac{\varphi^{s'}(B(x,n))}{\varphi^s(B(x,n))} &= \frac{\alpha_1 \cdots \alpha_k \cdots \alpha_h \alpha_{h+1}^{s'-h}}{\alpha_1 \cdots \alpha_k \alpha_{k+1}^{s-k}} = \alpha_{k+1}^{1-s+k} \alpha_{k+2} \cdots \alpha_h \alpha_{h+1}^{s'-h} \\ &\leq \alpha_1^{1-s+k+(h-(k+2)+1)+s'-h} = \alpha_1^{(s'-s)} < C^{(s'-s)n}.\end{aligned}$$

Also

$$\alpha_{k+1}^{1-s+k} \alpha_{k+2} \cdots \alpha_h \alpha_{h+1}^{s'-h} \geq \alpha_d^{1-s+k+(h-(k+2)+1)+s'-h} > D^{(s'-s)n}.$$

Thus,

$$D^{(s'-s)n} \leq \frac{\varphi^{s'}(B(x,n))}{\varphi^s(B(x,n))} \leq C^{(s'-s)n},$$

when $s < s' < d$. Similar calculations show that if $s \geq s'$, then

$$C^{(s'-s)n} \leq \frac{\varphi^{s'}(B(x,n))}{\varphi^s(B(x,n))} \leq D^{(s'-s)n}.$$

Write

$$\mathcal{P} = P_g(B,s') - P_g(B,s) = \lim_{n\to\infty} \frac{1}{n} \log \frac{\sum_{\mathbf{i}\in X_n^*} \sup_{y\in[\mathbf{i}]} \exp(S_n g(y)) \varphi^{s'}(B(y,n))}{\sum_{\mathbf{j}\in X_n^*} \sup_{z\in[\mathbf{j}]} \exp(S_n g(z)) \varphi^{s}(B(z,n))}.$$

If $s < s' < d$, then $D^{(s'-s)n}\varphi^s(B(x,n)) \leq \varphi^{s'}(B(x,n)) \leq C^{(s'-s)n}\varphi^s(B(x,n))$. Therefore,

$$\begin{aligned}\mathcal{P} &\leq \lim_{n\to\infty} \frac{1}{n} \log \left(C^{(s'-s)n} \frac{\sum_{\mathbf{i}\in X_n^*} \sup_{y\in[\mathbf{i}]} \exp(S_n g(y)) \varphi^{s}(B(y,n))}{\sum_{\mathbf{j}\in X_n^*} \sup_{z\in[\mathbf{j}]} \exp(S_n g(z)) \varphi^{s}(B(z,n))} \right) \\ &= (s'-s) \log C\end{aligned}$$

and $\mathcal{P} \geq (s'-s)\log D$. So

$$(s'-s)\log D \leq \mathcal{P} \leq (s'-s)\log C,$$

if $s < s' < d$. Similarly, if $s \geq s'$,

$$(s' - s)\log C \leq \mathcal{P} \leq (s' - s)\log D.$$

Thus,

$$|P_g(B, s') - P_g(B, s)| \leq |s - s'| \cdot \max\{|\log C|, |\log D|\}.$$

This implies that $|P_g(B, s') - P_g(B, s)|$ can be made arbitrarily small by choosing smaller and smaller neighbourhoods I of s. On the other hand by part (1), by choosing suitably small $\mathcal{U}$, $|P_g(B, s) - P_g(A, s)|$ can be made small. Thus,

$$\begin{aligned} |P_g(B, s') - P_g(A, s)| &= |P_g(B, s') - P_g(B, s) + P_g(B, s) - P_g(A, s)| \\ &\leq |P_g(B, s') - P_g(B, s)| + |P_g(B, s) - P_g(A, s)| \end{aligned}$$

can be made arbitrarily small, establishing the continuity at (A, s). □

Now, Theorem 2.6 is just a simple corollary of Theorem 2.12.

Proof of Theorem 2.6. By remark 2.11, the first and third claim are just special cases of Theorem 2.12.

For the second claim, we first assume that $s(A) > 0$. Fix $\mathbf{A} \in \mathcal{A}^C_{d,k}$ and $0 < \varepsilon < s(\mathbf{A})$. Then $P(\mathbf{A}, s(\mathbf{A}) + \varepsilon) < 0$ since by Lemma 2.3 $P(\mathbf{A}, \cdot)$ is strictly decreasing for positive values. By (1) there exists a neighbourhood $\mathcal{U}_1$ of $\mathbf{A}$ such that $P(\mathbf{B}, s(\mathbf{A}) + \varepsilon) < 0$ for every $\mathbf{B} \in \mathcal{U}_1$. Thus, $s(\mathbf{B}) \leq s(\mathbf{A}) + \varepsilon$. On the other hand, $P(\mathbf{A}, s(\mathbf{A}) - \varepsilon) > 0$. By (1) there exists a neighbourhood $\mathcal{U}_2$ of $\mathbf{A}$ such that $P(\mathbf{B}, s(\mathbf{A}) - \varepsilon) > 0$ for any $\mathbf{B} \in \mathcal{U}_2$. Thus, $s(\mathbf{B}) \geq s(\mathbf{A}) - \varepsilon$. Hence, $|s(\mathbf{B}) - s(\mathbf{A})| < \varepsilon$ for every $\mathbf{B} \in \mathcal{U}_1 \cap \mathcal{U}_2$. Suppose now $s(\mathbf{A}) = 0$. We deduce as above that $s(\mathbf{B}) \leq s(\mathbf{A}) + \varepsilon$ for every $\mathbf{B} \in \mathcal{U}_1$. This shows that $s(\mathbf{A})$ is upper semi-continuous at $\mathbf{A}$. The lower semi-continuity is clear since $s(\mathbf{A}) \leq s(\mathbf{B})$ for every $\mathbf{B}$. □

References

[1] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, 2007.

[2] L. Arnold. *Random Dynamical Systems*. Monographs in Mathematics. Springer, 1998.

[3] A. Avila and J. Bochi. On the subadditive ergodic theorem. `https://www.researchgate.net/publication/253322374_ON_THE_SUBADDITIVE_ERGODIC_THEOREM`, 2009. [Accessed 24 February 2016].

[4] L. Barreira. *Ergodic Theory, Hyperbolic Dynamics and Dimension Theory.* Universitext. Springer, 2012.

[5] V. I. Bogachev. *Measure Theory*, volume 1. Springer Berlin Heidelberg, 2007.

[6] V. I. Bogachev. *Measure Theory*, volume 2. Springer Berlin Heidelberg, 2007.

[7] M. Brin and G. Stuck. *Introduction to Dynamical Systems.* Cambridge University Press, 2002.

[8] Y. L. Cao, D. J. Feng, and W. Huang. The thermodynamic formalism for sub-additive potentials. *Discrete and Continuous Dynamical Systems*, 20(3):639–657, 2008.

[9] K. J. Falconer. The hausdorff dimension of self-affine fractals. *Mathematical Proceedings of the Cambridge Philosophical Society*, 103(2):339–350, 1988.

[10] K. J. Falconer. *Techniques in Fractal Geometry.* Wiley, 1997.

[11] K. J. Falconer. *Fractal Geometry: Mathematical Foundations and Applications.* Wiley, 2014.

[12] D. J. Feng and P. Shmerkin. Non-conformal repellers and the continuity of pressure for matrix cocycles. *Geometric and Functional Analysis*, 24(4):1101–1128, 2014.

[13] G. Froyland, S. Lloyd, and A. Quas. Coherent structures and isolated spectrum for perron–frobenius cocycles. *Ergodic Theory and Dynamical Systems*, 30(3):729–756, 2010.

[14] E. Gover and N. Krikorian. Determinants and the volumes of parallelotopes and zonotopes. *Linear Algebra and its Applications*, 433(1):28–40, 2010.

[15] W. Greub. *Multilinear Algebra.* Universitext. Springer New York, 2012.

[16] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, 1994.

[17] G. Keller. *Equilibrium States in Ergodic Theory.* London Mathematical Society Student Texts. Cambridge University Press, 1998.

[18] P. Lancaster and M. Tismenetsky. *The Theory of Matrices: With Applications*. Computer Science and Scientific Computing Series. Academic Press, 1985.

[19] K. Petersen. *Ergodic Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1989.

[20] M. Pollicott and M. Yuri. *Dynamical Systems and Ergodic Theory*. London Mathematical Society Student Texts. Cambridge University Press, 1998.

[21] O. Sarig. Lecture notes on ergodic theory. `http://www.wisdom.weizmann.ac.il/~sarigo/506/ErgodicNotes.pdf`, 2009. [Accessed 24 February 2016].

[22] P. Shmerkin. Self-affine sets and the continuity of subadditive pressure. In De-Jun Feng and Ka-Sing Lau, editors, *Geometry and Analysis of Fractals*, volume 88 of *Springer Proceedings in Mathematics and Statistics*, chapter 12, pages 325–342. Springer, 2014.

[23] B. Solomyak. Measure and dimension for some fractal families. *Mathematical Proceedings of the Cambridge Philosophical Society*, 124(3):531–546, 1998.

[24] S. G. Williams. *Symbolic Dynamics and Its Applications: American Mathematical Society, Short Course, January 4-5, 2002, San Diego, California*. AMS Short Course Lecture Notes. American Mathematical Society, 2004.