



Negative Curvature, Matrix Products, and Ergodic Theory

by

Eduardo Oregón Reyes

Thesis submitted to the Faculty of Mathematics
of Pontificia Universidad Católica de Chile, as one of the requirements
to qualify for the academic Master's degree in Mathematics.

Supervisor : Dr. Jairo Bochi (PUC Chile)
Committee : Dr. Mario Ponce (PUC Chile)
Dr. Andrés Navas (USACH)

July, 2018

Santiago, Chile

Para Rosa

Este trabajo fue parcialmente financiado por la beca CONICYT Folio 22172003,
y por el proyecto Anillo ACT 172001

Contents

1	Properties of Sets of Isometries of Gromov Hyperbolic Spaces	11
1.1	Introduction	11
1.2	The case of the hyperbolic plane	14
1.2.1	Derivation of matrix inequalities	14
1.2.2	Finiteness conjecture on $Isom(\mathbb{H}^2)$	16
1.3	Proof of Theorems 1.1.1 and 1.1.2	17
1.4	Berger-Wang and further properties of the stable length and joint stable length	20
1.4.1	A Berger-Wang Theorem for sets of isometries	20
1.4.2	Dynamical interpretation and semigroups of isometries	20
1.4.3	Relation with the minimal length	21
1.5	Continuity	22
1.5.1	Continuity of the stable length	22
1.5.2	Vietoris topology and continuity of the joint stable length	23
2	A New Inequality about Matrix Products and a Berger-Wang Formula	26
2.1	Introduction	26
2.2	Proof of Theorem 2.1.3	28
2.2.1	Some computations in low dimension	30
2.3	A polynomial identity	31
2.4	Proof of Theorem 2.1.2	33
2.4.1	The case of the complex numbers	34
2.5	Ergodic-theoretical consequences	36
2.6	Ergodic characterization of the JSR	37
2.7	Geometric remarks	38

Appendix 41

A Vietoris topology over topological groups 41

Introduction

Since its discovery, hyperbolic geometry has played a prominent role in mathematics. For example, the Uniformization Theorem, one of the cornerstones of the theory of Riemann surfaces, asserts that most of the orientable surfaces possess a natural hyperbolic geometry (i.e. of constant negative curvature). Hyperbolic geometry is still an active research topic.

One of the most basic models of this geometry is the *upper-half plane* $\mathbb{H}^2 = \{z \in \mathbb{C} : \text{Im}(z) > 0\}$ endowed with the Riemannian metric $ds^2 = dz^2/\text{Im}(z)^2$. With this metric, the geodesics of \mathbb{H}^2 are the restrictions of vertical lines and circles with center on the real line. Even though this model does not satisfy Euclid's parallel postulate, it possesses a very rich and concrete geometry. For instance, it has good notions of angles, distances, and even trigonometric identities.

The group $\text{Isom}(\mathbb{H}^2)$ of isometries of \mathbb{H}^2 is also well understood, and has an inherent relation with linear algebra. Every orientation-preserving isometry of \mathbb{H}^2 is a Möbius transformation $z \mapsto \frac{az+b}{cz+d}$ for some $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(\mathbb{R})$. This relation is not only algebraic: quantities related to matrices in $\text{SL}_2(\mathbb{R})$ (such as norms or spectral radii) can be translated into quantities relating \mathbb{H}^2 and its isometries (distance between points in \mathbb{H}^2), and vice versa. This situation is reflected in the following proposition (see Section 1.2 for the proof):

Proposition 1. *Let $\|\cdot\|$ be the Euclidean operator norm on $\text{GL}_2(\mathbb{R})$. For every $A \in \text{SL}_2(\mathbb{R})$ with associated Möbius transformation \tilde{A} , the following holds:*

$$d_{\mathbb{H}^2}(\tilde{A}i, i) = 2 \log(\|A\|).$$

By using this translation, the spectral radius of a matrix A , which by Gelfand's formula equals $\rho(A) = \lim_{n \rightarrow \infty} \|A^n\|^{1/n}$, corresponds (after applying the function $x \rightarrow e^{x/2}$) to the number $\lim_{n \rightarrow \infty} \frac{d_{\mathbb{H}^2}(\tilde{A}^n i, i)}{n}$. In general, for a metric space (M, d) with isometry group $\text{Isom}(M)$, the *stable length* of an isometry $f \in \text{Isom}(M)$ is the number

$$d^\infty(f) = \lim_{n \rightarrow \infty} \frac{d(f^n x, x)}{n},$$

where x is an arbitrary point of M (as we will see, this quantity is well defined and the choice of x is irrelevant).

The stable length also appears in the next example of the translation between linear algebra and hyperbolic geometry, which we will present after the proof of the next proposition:

Proposition 2. *For all $A \in \mathrm{SL}_2(\mathbb{R})$, we have the identity:*

$$\|A^2\| - \|A^2\|^{-1} = (\|A\| - \|A\|^{-1}) \cdot |\mathrm{Tr} A|. \quad (1)$$

Proof. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$, the numbers $\|A\|$ and $\|A\|^{-1}$ are the singular values of A , and hence $\|A\|^2 + \|A\|^{-2}$ is the trace of A^*A (singular values and eigenvalues coincide for symmetric matrices), which equals $a^2 + b^2 + c^2 + d^2$. Applying the same reasoning to A^2 , which by Caley-Hamilton equals $\mathrm{Tr} A \cdot A - I$ (where I is the identity matrix), we obtain

$$\begin{aligned} (\|A^2\| - \|A^2\|^{-1})^2 &= \|A^2\|^2 + \|A^2\|^{-2} - 2 \\ &= (a \mathrm{Tr} A - 1)^2 + (b \mathrm{Tr} A)^2 + (c \mathrm{Tr} A)^2 + (d \mathrm{Tr} A - 1)^2 - 2 \\ &= (a^2 + b^2 + c^2 + d^2)(\mathrm{Tr} A)^2 - 2(a + d) \mathrm{Tr} A \\ &= (a^2 + b^2 + c^2 + d^2 - 2)(\mathrm{Tr} A)^2 \\ &= ((\|A\| - \|A\|^{-1}) \cdot \mathrm{Tr} A)^2. \quad \square \end{aligned}$$

Since every orientation-preserving isometry of \mathbb{H}^2 is of the form $f = \tilde{A}$ for some $A \in \mathrm{SL}_2(\mathbb{R})$, Proposition 1 converts the equation (1) into the following identity in hyperbolic geometry:

Corollary 3. *Let f be an orientation-preserving isometry of \mathbb{H}^2 , and let $x \in \mathbb{H}^2$.*

i) If $d^\infty(f) > 0$, then

$$\sinh\left(\frac{1}{2}d_{\mathbb{H}^2}(f^2x, x)\right) = 2 \sinh\left(\frac{1}{2}d_{\mathbb{H}^2}(fx, x)\right) \cosh\left(\frac{1}{2}d^\infty(f)\right).$$

ii) If $d^\infty(f) = 0$, then

$$\sinh\left(\frac{1}{2}d_{\mathbb{H}^2}(f^2x, x)\right) \leq 2 \sinh\left(\frac{1}{2}d_{\mathbb{H}^2}(fx, x)\right).$$

In particular, since for all $y \geq z \geq 0$ we have the inequality $2 \sinh(y) \cosh(z) \leq 2 \sinh(y+z) \leq \sinh(y+z + \log 2)$ (this is a consequence of the elementary trigonometric identities for hyperbolic functions), the previous corollary implies that

$$d_{\mathbb{H}^2}(f^2x, x) \leq d_{\mathbb{H}^2}(fx, x) + d^\infty(f) + 2 \log 2, \quad (2)$$

for all $x \in \mathbb{H}^2$ and f an orientation-preserving isometry of \mathbb{H}^2 .

These kinds of inequalities are in principle of hyperbolic nature, and do not apply in Euclidean spaces. In \mathbb{R}^2 , for a rotation f around the origin with small rotation angle, we have $d^\infty(f) = 0$ (since f has a fixed point). Then, for a point $x \in \mathbb{R}^2$ with $|x| \gg 1$, the circle containing x with center at the origin also contains fx and f^2x , and has geodesic curvature very close to 0. Since the rotation angle is small, the three points x, fx and f^2x are almost collinear, and we have $d(f^2x, x) \approx 2d(fx, x) \gg 1$. Therefore inequality (2) does not apply in this case. This complication does not occur in the hyperbolic plane, since the geodesic curvature of a hyperbolic circle is *always* greater than one, independently of the radius

(see e.g. [26, Sec. 2.3]). In fact, inequality *ii*) of Corollary 3 shows that for a rotation f in the hyperbolic plane, $d_{\mathbb{H}^2}(fx, x)$ and $d_{\mathbb{H}^2}(f^2x, x)$ have the *same* order of magnitude.

One of the goals of this thesis is to prove an analogue of inequality (2) valid for metric spaces presenting some sort of negative curvature similar to the space \mathbb{H}^2 . We will work with the notion of *Gromov hyperbolic metric spaces*, which includes geometric objects such as Riemannian manifolds with uniform negative curvature (the hyperbolic plane \mathbb{H}^2 satisfies this condition), but also combinatorial objects such as trees. This notion also applies to finitely generated groups, and in fact, in a rigorous sense *most* of these groups are Gromov hyperbolic [42].

We will prove the following inequality for a δ -hyperbolic metric space (M, d) (see Section 1.1 for a detailed definition):

Theorem 4. *If $x \in M$ and $f \in \text{Isom}(M)$, then:*

$$d(f^2x, x) \leq d(fx, x) + d^\infty(f) + 2\delta.$$

The situation becomes more interesting if two isometries are involved. In this case we obtain the following:

Theorem 5. *For every $x \in M$ and every $f, g \in \text{Isom}(M)$ we have:*

$$d(fgx, x) \leq \max\left(d(fx, x) + d^\infty(g), d(gx, x) + d^\infty(f), \frac{d(fx, x) + d(gx, x) + d^\infty(fg)}{2}\right) + 6\delta. \quad (3)$$

At first, inequality (3) may seem a bit technical. The key idea behind this result is the following: in general, we have $d^\infty(h) \leq d(hx, x)$, but it may occur that $d^\infty(h)$ is very small compared with $d(hx, x)$. Inequality (3) says that, if distance $d(fgx, x)$ is comparable to (that is, not much smaller than) the sum $d(fx, x) + d(gx, x)$, then $d^\infty(h)$ is not much smaller than $d(hx, x)$ for some $h \in \{f, g, fg\}$.

Since the almost-additive property $d(ghx, x) \approx d(fx, x) + d(gx, x)$ is in some sense frequent under ergodic-theoretical assumptions (see e.g. [21, Thm. 1.1]), Theorem 5 allows to produce isometries with $d(hx, x) \approx d^\infty(h)$, and hence becomes an effective tool in the study of cocycles of isometries (see Section 2.7). In particular, given a finite set $\Sigma \subset \text{Isom}(M)$ and $x \in M$, the asymptotic quantity

$$\mathfrak{D}(\Sigma) = \lim_{n \rightarrow \infty} \frac{1}{n} \sup_{f_i \in \Sigma} d(f_n \cdots f_1 x, x)$$

(which is well defined, see Section 1.1) may be approximated by using the stable length. More precisely, Theorem 5 implies the identity

$$\mathfrak{D}(\Sigma) = \limsup_{n \rightarrow \infty} \frac{1}{n} \sup_{f_i \in \Sigma} d^\infty(f_n \cdots f_1). \quad (4)$$

The number $\mathfrak{D}(\Sigma)$ that we call *joint stable length*, and the identity (4), are inspired in a well known identity for bounded sets of matrices: the *Berger-Wang Identity* (see Section 2.1 for a detailed explanation). Identity (4) has turned out to be very useful in the study of semigroups of isometries of non-positively curved spaces, as recently showed Breuillard and Fujiwara [10].

The dictionary between hyperbolic geometry and linear algebra also allows to use Theorem 5 to recover information regarding matrix products in $\mathrm{SL}_2(\mathbb{R})$, similar to Proposition 2. Then a natural question arises: what can we say about higher dimensional matrix products? Another goal of this thesis is to give an appropriate analogue of inequality (3) to matrix products in $M_d(\mathbb{R})$, the space of real matrices of order $d \times d$:

Theorem 6. *Let $d \in \mathbb{N}$, and $\|\cdot\|$ be a norm on $M_d(\mathbb{R})$. There exist constants $N = N(d) \in \mathbb{N}$, $0 < \delta = \delta(d) < 1$, and $C = C(d, \|\cdot\|) > 1$ satisfying the following inequality for all $A_1, \dots, A_N \in M_d(\mathbb{R})$:*

$$\|A_N \cdots A_1\| \leq C \left(\prod_{1 \leq i \leq N} \|A_i\| \right) \max_{1 \leq \alpha \leq \beta \leq N} \left(\frac{\rho(A_\beta \cdots A_\alpha)}{\prod_{\alpha \leq i \leq \beta} \|A_i\|} \right)^\delta, \quad (5)$$

where the right hand side expression is treated as zero if one of the A_i is the zero matrix.

The interpretation of this result is similar to the one given for Theorem 5. In general, for an operator norm $\|\cdot\|$ on $M_d(\mathbb{R})$, we have the inequalities $\rho(A_1) \leq \|A_1\|$ and $\|A_N \cdots A_1\| \leq \|A_N\| \cdots \|A_1\|$, for all $A_1, \dots, A_N \in M_d(\mathbb{R})$. Inequality (5) says that given a big enough N , if the norm of the product $A_N \cdots A_1$ is comparable to (that is, not much smaller than) the product of the norms of the A_j , then there exists a subproduct $A_\beta \cdots A_\alpha$ whose spectral radius is comparable to (that is, not much smaller than) $\prod_{\alpha \leq i \leq \beta} \|A_i\|$.

As we mentioned before, the property $\|A_N \cdots A_1\| \approx \|A_N\| \cdots \|A_1\|$ is natural in the context of ergodic theory. We will explore some of the consequences of Theorem 6, among which include a new proof of Berger-Wang identity for matrix products, and a simpler proof for a theorem of I. D. Morris that characterizes the upper Lyapunov exponent of a matrix cocycle in terms of spectral radii (see Theorem 2.1.4).

Organization of the thesis. This work is divided into two chapters, which are independent and can be read separately.

In the first chapter we prove Theorems 4 and 5 and some of their consequences. Section 1.1 introduces Gromov hyperbolic spaces and the important concepts that we will study. Section 1.2 is devoted to the relationship between Theorems 4 and 5 with matrix theory. We also give a counterexample to the finiteness conjecture in $\mathrm{Isom}(\mathbb{H}^2)$. Then, in Section 1.3 we prove Theorems 4 and 5. In Section 1.4 we prove the Berger-Wang theorem for sets of isometries and study the basic properties of the stable lengths, reviewing some known results, and their geometric or dynamical interpretations, specifically on the classification of semigroups of isometries in hyperbolic spaces. In Section 1.5 we study the continuity properties of the stable length and the joint stable length, with respect to very natural topologies on the space $\mathrm{Isom}(M)$ and on the space of compact subsets of $\mathrm{Isom}(M)$. The results of this chapter are contained in the paper [44]:

E. Oregón-Reyes: *Properties of sets of isometries of Gromov hyperbolic spaces.* To appear in *Groups, Geometry and Dynamics*.

In the second chapter we prove Theorem 6. In Section 2.1 we contextualize the problem and introduce relevant notation. The next three sections are devoted

to prove Theorem 6. In Section 2.2 we prove a theorem regarding products of nilpotent matrices over an arbitrary field. We use this theorem in Section 2.3, and via Nullstellensatz, we translate it into a polynomial identity, from which we deduce Theorem 6 in Section 2.4. We study the ergodic consequences of inequality (2.3) in Section 2.5. Finally, in Section 2.6 we show that Morris's theorem implies Berger-Wang theorem, and discuss some geometric analogies of these results for isometries of Gromov hyperbolic spaces in Section 2.7. Most of these results are contained in the preprint [43]:

E. Oregón-Reyes: *A new inequality of matrix products and a Berger-Wang formula.*

We leave Appendix A for the technical results that we used in Section 1.5, and we prove them for the Vietoris topology on an arbitrary topological group. Additional to the original preprint [43], we include Section 2.6 and the respective analogues of the added results for hyperbolic spaces in Section 2.7.

Chapter 1

Properties of Sets of Isometries of Gromov Hyperbolic Spaces

1.1 Introduction

Let (M, d) be a metric space with distance. We assume this space is δ -hyperbolic in the Gromov sense. This concept was introduced in 1987 [19] and has an important role in geometric group theory and negatively curved geometry [11, 19, 24]. There are several equivalent definitions [12], among which the following *four points condition* (f.p.c.): For all $x, y, s, t \in M$ the following holds:

$$d(x, y) + d(s, t) \leq \max(d(x, s) + d(y, t), d(x, t) + d(y, s)) + 2\delta. \quad (\text{f.p.c.})$$

This chapter deals with isometries of hyperbolic spaces. We do not assume M to be geodesic nor proper, since these conditions are irrelevant for many purposes [7, 13, 23, 49]. We also do not make use of the Gromov boundary, deriving our fundamental results directly from (f.p.c.).

Let us introduce some terminology and notation. Let $Isom(M)$ be the group of isometries of M . For $x \in M$ and $\Sigma \subset Isom(M)$ define

$$d(\Sigma, x) = \sup_{f \in \Sigma} d(fx, x).$$

We say that Σ is *bounded* if $d(\Sigma, x) < \infty$ for some (and hence any) $x \in M$.

For a single isometry f the *stable length* is defined by

$$d^\infty(f) = \lim_{n \rightarrow \infty} \frac{d(f^n x, x)}{n} = \inf_n \frac{d(f^n x, x)}{n}.$$

This quantity is well defined and finite by subadditivity and turns to be independent of $x \in M$.

Our first result gives a *lower* bound for the stable length:

Theorem 1.1.1. *If $x \in M$ and $f \in \text{Isom}(M)$ then:*

$$d(f^2x, x) \leq d(fx, x) + d^\infty(f) + 2\delta. \quad (1.1)$$

The main result of this chapter is a version of Theorem 1.1.1 for two isometries:

Theorem 1.1.2. *For every $x \in M$ and every $f, g \in \text{Isom}(M)$ we have:*

$$d(fgx, x) \leq \max \left(d(fx, x) + d^\infty(g), d(gx, x) + d^\infty(f), \frac{d(fx, x) + d(gx, x) + d^\infty(fg)}{2} \right) + 6\delta. \quad (1.2)$$

For the generalization of the stable length and Theorem 1.1.1 to bounded sets of isometries, some notation is required. If $\Sigma \subset \text{Isom}(M)$ we denote by Σ^n the set of all compositions of n isometries of Σ . Note that if Σ is bounded then each Σ^n is bounded. We define the *joint stable length* as the quantity

$$\mathfrak{D}(\Sigma) = \lim_{n \rightarrow \infty} \frac{d(\Sigma^n, x)}{n} = \inf_n \frac{d(\Sigma^n, x)}{n}.$$

Similarly as before, this function is well defined, finite and independent of x . Also, it is useful to define the *stable length* of Σ as

$$d^\infty(\Sigma) = \sup_{f \in \Sigma} d^\infty(f).$$

Taking supremum over $f, g \in \Sigma$ in both sides of (1.2) and noting that $d^\infty(\Sigma^2) \leq \mathfrak{D}(\Sigma^2) = 2\mathfrak{D}(\Sigma)$ we obtain a lower bound for the joint stable length similar to Theorem 1.1.1:

Corollary 1.1.3. *For every $x \in M$ and every bounded set $\Sigma \subset \text{Isom}(M)$ the following holds:*

$$d(\Sigma^2, x) \leq d(\Sigma, x) + \frac{d^\infty(\Sigma^2)}{2} + 6\delta \leq d(\Sigma, x) + \mathfrak{D}(\Sigma) + 6\delta. \quad (1.3)$$

Inequalities (1.1) and (1.3) are inspired by lower bounds for the spectral radius due to J. Bochi [5, Eq. 1 & Thm. A]. As we will see, the connection between the spectral radius and the stable length will allow us to deduce Bochi's inequalities from (1.1) and (1.3) (see Section 1.2 below), and actually improve them using (1.2).

We present some applications of Theorems 1.1.1 and 1.1.2:

Berger-Wang like theorem. The joint stable length is inspired by matrix theory. Let $M_d(\mathbb{R})$ be the set of real $d \times d$ matrices and let $\|\cdot\|$ be an operator norm on $M_d(\mathbb{R})$. We denote the spectral radius of a matrix A by $\rho(A)$. The *joint spectral radius* of a bounded set $\mathcal{M} \subset M_d(\mathbb{R})$ is defined by

$$\mathfrak{R}(\mathcal{M}) = \lim_{n \rightarrow \infty} \sup \left\{ \|A_1 \dots A_n\|^{1/n} : A_i \in \mathcal{M} \right\}.$$

Note the similarity with the definition of the joint stable length.

The joint spectral radius was introduced by Rota and Strang [46] and popularized by Daubechies and Lagarias [14]. This quantity has aroused research interest

in recent decades and it has appeared in several mathematical contexts (see e.g. [29, 32]). An important result related to the joint spectral radius is the Berger-Wang theorem [2] which says that for all bounded sets $\mathcal{M} \subset M_d(\mathbb{R})$ we have $\mathfrak{R}(\mathcal{M}) = \limsup_{n \rightarrow \infty} \sup \{\rho(A)^{1/n} : A \in \mathcal{M}^n\}$. From Corollary 1.1.3 we prove a similar result for the joint stable length in a δ -hyperbolic space¹:

Theorem 1.1.4. *Every bounded set $\Sigma \subset \text{Isom}(M)$ satisfies*

$$\mathfrak{D}(\Sigma) = \limsup_{n \rightarrow \infty} \frac{d^\infty(\Sigma^n)}{n} = \lim_{n \rightarrow \infty} \frac{d^\infty(\Sigma^{2n})}{2n}.$$

A question that arose from the Berger-Wang theorem is the *finiteness conjecture* proposed by Lagarias and Wang [34] which asserts that for every finite set $\mathcal{M} \subset M_d(\mathbb{R})$ there exists some $n \geq 1$ and $A_1, \dots, A_n \in \mathcal{M}$ such that $\mathfrak{R}(\mathcal{M}) = \rho(A_1 \cdots A_n)^{1/n}$. The failure of this conjecture was proved by Bousch and Mairesse [8].

In the context of sets of isometries, following an idea of I. D. Morris (personal communication) we refute the finiteness conjecture for $M = \mathbb{H}^2$.

Proposition 1.1.5. *There exists a finite set $\Sigma \subset \text{Isom}(\mathbb{H}^2)$ such that for all $n \geq 1$:*

$$\mathfrak{D}(\Sigma) > \frac{d^\infty(\Sigma^n)}{n}.$$

Let us interpret these facts in terms of Ergodic Theory. Given a compact set of matrices \mathcal{M} , the joint spectral radius equals the supremum of the Lyapunov exponents over all ergodic shift-invariant measures on the space $\mathcal{M}^{\mathbb{N}}$ (see [39] for details). Therefore, Berger-Wang says that instead of considering all shift-invariant measures, it is sufficient to consider those supported on periodic orbits. A far-reaching extension of this result was obtained by Kalinin [30].

Classification of semigroups of isometries. The stable length gives relevant information about isometries in hyperbolic spaces. Recall that for a δ -hyperbolic space M an isometry $f \in \text{Isom}(M)$ is either *elliptic*, *parabolic* or *hyperbolic*. This classification is directly related to the stable length [12, Chpt. 10, Prop. 6.3]:

Proposition 1.1.6. *An isometry f of M is hyperbolic if and only if $d^\infty(f) > 0$.*

There also exists a classification for semigroups of isometries in three disjoint families (also called *elliptic*, *parabolic* and *hyperbolic*) obtained by Das, Simmons and Urbański. An application of Theorem 1.1.4 is the following generalization of Proposition 1.1.6, which serves as a motivation to study the joint stable length $\mathfrak{D}(\Sigma)$:

Theorem 1.1.7. *The semigroup generated by a bounded set $\Sigma \subset \text{Isom}(M)$ is hyperbolic if and only if $\mathfrak{D}(\Sigma) > 0$.*

In addition, we give a sufficient condition for a product of two isometries to be hyperbolic, and a lower bound for the stable length of the product, improving [12, Chpt. 9, Lem. 2.2]:

¹Very recently, Breuillard and Fujiwara [10] gave a different proof of this result assuming that M is δ -hyperbolic and geodesic. They also proved the first formula in Theorem 1.1.4 when M is a symmetric space of non-compact type.

Proposition 1.1.8. *Let $K \geq 7\delta$ and $f, g \in \text{Isom}(M)$ be such that $d(fx, gx) > \max(d(fx, x) + d^\infty(g), d(gx, x) + d^\infty(f)) + K$ for some $x \in M$. Then fg is hyperbolic, and*

$$d^\infty(fg) > d^\infty(f) + d^\infty(g) + 2K - 14\delta.$$

Continuity results. The group $\text{Isom}(M)$ possesses a natural topology induced by the product topology on M^M , which is called the *point-open topology*. In this space it coincides with the compact-open topology [13, Prop. 5.1.2]. Using Theorem 1.1.1 we will prove that the stable length behaves well with respect to this topology:

Theorem 1.1.9. *The map $f \mapsto d^\infty(f)$ is continuous on $\text{Isom}(M)$ with the point-open topology.*

Remark 1.1.10. The stable length may be discontinuous if we do not assume that M is δ -hyperbolic. Take for example $M = \mathbb{C}$ with the Euclidean metric, and let $f_u : \mathbb{C} \rightarrow \mathbb{C}$ be given by $f_u(z) = uz + 1$, where u is a parameter in the unit circle. For $u \neq 1$ we have that f_u is a rotation, and hence $d^\infty(f_u) = 0$. But f_1 is a translation and $d^\infty(f_1) = 1$. However, the stable length is of course upper semi-continuous for *all* metric spaces.

Since in general the space $\text{Isom}(M)$ is not metrizable, we need a suitable generalization of the Hausdorff distance. We endow the set $\mathcal{C}(\text{Isom}(M))$ of non empty compact (with respect to the point-open topology) sets of isometries with the *Vietoris topology* [37]. This topology is natural in the sense that its separation, compactness and connectivity properties derive directly from the respective properties on $\text{Isom}(M)$ [37, §4]. In fact, when $\text{Isom}(M)$ is metrizable the Vietoris topology coincides with the one induced by the Hausdorff distance. The definition of the Vietoris topology is given in Subsection 1.5.2.

With these notions it is easy to check that every non empty compact set $\Sigma \subset \text{Isom}(M)$ is bounded and hence the joint stable length is well defined. As a consequence of Corollary 1.1.3 we have:

Theorem 1.1.11. *Endowing $\mathcal{C}(\text{Isom}(M))$ with the Vietoris topology, the joint stable length $\Sigma \mapsto \mathfrak{D}(\Sigma)$ and the stable length $\Sigma \mapsto d^\infty(\Sigma)$ are continuous.*

1.2 The case of the hyperbolic plane

1.2.1 Derivation of matrix inequalities

In this section we relate the stable lengths for sets of isometries and the spectral radii for 2×2 sets of matrices. To do this, we study the hyperbolic plane.

Let \mathbb{H}^2 be the upper-half plane $\{z \in \mathbb{C} : \text{Im}(z) > 0\}$ endowed with the Riemannian metric $ds^2 = dz^2/\text{Im}(z)^2$. This space is $\log 2$ -hyperbolic ($\log 2$ is the best possible constant [41, Cor. 5.4]). It is known that $\text{SL}_2^\pm(\mathbb{R}) = \{A \in M_2(\mathbb{R}) : \det A = \pm 1\}$ is isomorphic to $\text{Isom}(\mathbb{H}^2)$, with isomorphism given by

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \tilde{A}z = \begin{cases} \frac{az + b}{cz + d} & \text{if } \det A = 1, \\ \frac{a\bar{z} + b}{c\bar{z} + d} & \text{if } \det A = -1. \end{cases}$$

The relation between the distance $d_{\mathbb{H}^2}$ and the Euclidean operator norm $\|\cdot\|_2$ on $M_2(\mathbb{R})$ is established in the following proposition:

Proposition 1.2.1. *For every $A \in \mathrm{SL}_2^\pm(\mathbb{R})$ and every bounded set $\mathcal{M} \subset \mathrm{SL}_2^\pm(\mathbb{R})$ the following holds:*

- i) $d_{\mathbb{H}^2}(\tilde{A}i, i) = 2 \log(\|A\|_2)$.*
- ii) $d^\infty(\tilde{A}) = 2 \log(\rho(A))$.*
- iii) $\mathfrak{D}(\tilde{\mathcal{M}}) = 2 \log(\mathfrak{R}(\mathcal{M}))$, where $\tilde{\mathcal{M}} = \{\tilde{B} : B \in \mathcal{M}\} \subset \mathrm{Isom}(\mathbb{H}^2)$.*

Proof. By the definition of the joint stable length and Gelfand's formula $\rho(A) = \lim_{n \rightarrow \infty} (\|A^n\|_2)^{1/n}$ it is easy to see that *ii)* and *iii)* are consequences of *i)*.

The proof of *i)* is simple. In the case that \tilde{A} fixes i , that is, A is an orthogonal matrix, the equality is trivial. In the case that A is a diagonal matrix, the proof is a straightforward computation. The general case follows by considering the singular value decomposition. \square

Corollary 1.2.2. *For every $A \in \mathrm{SL}_2^\pm(\mathbb{R})$ and $z \in \mathbb{H}^2$:*

$$d_{\mathbb{H}^2}(\tilde{A}z, z) = 2 \log(\|SAS^{-1}\|_2)$$

where S is any element in $\mathrm{SL}_2^\pm(\mathbb{R})$ that satisfies $\tilde{S}z = i$.

Proof. By Proposition 1.2.1 *i)*, $d_{\mathbb{H}^2}(\tilde{A}z, z) = d_{\mathbb{H}^2}(\tilde{A}\tilde{S}^{-1}i, \tilde{S}^{-1}i) = d_{\mathbb{H}^2}(\tilde{S}\tilde{A}\tilde{S}^{-1}i, i) = 2 \log(\|SAS^{-1}\|_2)$, where we used that \tilde{S} is an isometry. \square

Now we present the lower bound for the spectral radius due to Bochi:

Proposition 1.2.3. *Let $d \geq 2$. For every $A \in M_d(\mathbb{R})$ and every operator norm $\|\cdot\|$ on $M_d(\mathbb{R})$:*

$$\|A^d\| \leq (2^d - 1)\rho(A)\|A\|^{d-1}. \quad (1.4)$$

The generalization of Proposition 1.2.3 to a lower bound for the joint spectral radius is as follows:

Theorem 1.2.4 (Bochi [5]). *There exists $C = C(d) > 1$ such that, for every bounded set $\mathcal{M} \subset M_d(\mathbb{R})$ and every operator norm $\|\cdot\|$ on $M_d(\mathbb{R})$:*

$$\sup_{A_i \in \mathcal{M}} \|A_1 \dots A_d\| \leq C \mathfrak{R}(\mathcal{M}) \sup_{A \in \mathcal{M}} \|A\|^{d-1}. \quad (1.5)$$

Dividing by 2, applying the exponential function in (1.1), and using Proposition 1.2.1 *i)* and Corollary 1.2.2 we obtain

$$\|SA^2S^{-1}\|_2 \leq 2\rho(A)\|SAS^{-1}\|_2. \quad (1.6)$$

To replace $\|\cdot\|_2$ by an arbitrary operator norm we use the following lemma [5, Lem. 3.2]:

Lemma 1.2.5. *There exists a constant $C_0 > 1$ such that for every operator norm $\|\cdot\|$ on $M_2(\mathbb{R})$ there exists some S in $\mathrm{SL}_2^\pm(\mathbb{R})$ such that for every $A \in M_2(\mathbb{R})$:*

$$C_0^{-1}\|A\| \leq \|SAS^{-1}\|_2 \leq C_0\|A\|.$$

With this lemma we can give another proof of Bochi's Proposition 1.2.3 for $d = 2$, replacing the constant $(2^2 - 1)$ by $2C_0^2$, where C_0 is the constant given by Lemma 1.2.5. This involves three steps:

Step 1. The result is valid for all operator norms and $A \in \mathrm{SL}_2^\pm(\mathbb{R})$.

Consider the operator norm $\|\cdot\|$ on $M_2(\mathbb{R})$ and the respective $S \in \mathrm{SL}_2^\pm(\mathbb{R})$ given by Lemma 1.2.5. Using this in (1.6) we obtain

$$\|A^2\| \leq C_0\|SA^2S^{-1}\|_2 \leq 2C_0\rho(A)\|SAS^{-1}\|_2 \leq 2C_0^2\rho(A)\|A\|.$$

Step 2. We extend the result to $A \in \mathrm{GL}_2(\mathbb{R})$.

This is easy since inequality (1.6) is homogeneous in A .

Step 3. We can consider A an arbitrary matrix in $M_2(\mathbb{R})$.

We use that $\mathrm{GL}_2(\mathbb{R})$ is dense in $M_2(\mathbb{R})$ considering the metric given by $\|\cdot\|_2$. In this case the matrix multiplication and the spectral radius are continuous functions. So the conclusion follows.

If we do the same process to recover Theorem 1.2.4 in dimension 2 from Theorem 1.1.2 we will obtain a stronger result:

Proposition 1.2.6. *For all pairs of matrices $A, B \in M_2(\mathbb{R})$ and all operator norms $\|\cdot\|$ on $M_2(\mathbb{R})$:*

$$\|AB\| \leq 8C_0^2 \max\left(\|A\|\rho(B), \|B\|\rho(A), \sqrt{\|A\|\|B\|\rho(AB)}\right). \quad (1.7)$$

Proof. The case with $\|\cdot\| = \|\cdot\|_2$ and $A, B \in \mathrm{SL}_2^\pm(\mathbb{R})$ is a consequence of applying Proposition 1.2.1 in (1.2). Steps 1 and 3 are exactly the same as we did before. Step 2 follows by noting that (1.7) is a bihomogeneous inequality, in the sense that when we fix A it is homogeneous in B and when we fix B it is homogeneous in A . \square

1.2.2 Finiteness conjecture on $\mathrm{Isom}(\mathbb{H}^2)$

We finish this section giving a negative answer to the finiteness conjecture when $M = \mathbb{H}^2$. This is equivalent to finding a counterexample to the finiteness conjecture for matrices in $\mathrm{SL}_2^\pm(\mathbb{R})$.

The following construction was communicated to us by I. D. Morris:

$$\text{Let } \mathcal{A}^{(t)} = \left(A_0^{(t)}, A_1^{(t)}\right) \in \mathrm{SL}_2^\pm(\mathbb{R}), \text{ where } A_0^{(t)} = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}, A_1^{(t)} = \begin{pmatrix} 2t^{-1} & 3t \\ t^{-1} & 2t \end{pmatrix}$$

and $t \in \mathbb{R}^+$. Our claim is the following:

Theorem 1.2.7. *The family $(\mathcal{A}^{(t)})_{t \geq 1}$ contains a counterexample to the finiteness conjecture.*

Proof. The argument is similar to the one used in [3]. First, note that for all $t \geq 1$ the set $\mathcal{A}^{(t)}$ satisfies the hypotheses of the Jenkinson-Pollicott's Theorem [28, Thm. 9] and therefore one of the following options holds: either $\mathcal{A}^{(t)}$ is a counterexample to the finiteness conjecture, or there exists a finite product $A_\sigma^{(t)} = A_{i_1}^{(t)} \cdots A_{i_n}^{(t)}$ with $\sigma = (i_1, \dots, i_n) \in \{0, 1\}^n$ not being a power such that $\rho(A_\sigma^{(t)})^{1/n} = \mathfrak{R}(\mathcal{A}^{(t)})$ and, more importantly, the word σ is unique modulo cyclic permutations.

Suppose that no counterexample exists. As the map $t \mapsto \mathcal{A}^{(t)}$ is continuous, by the continuity of the spectral radius and joint spectral radius on $\mathrm{SL}_2^\pm(\mathbb{R})$ and $\mathrm{SL}_2^\pm(\mathbb{R})$ respectively, the maps $t \mapsto \mathfrak{R}(\mathcal{A}^{(t)})$ and $t \mapsto \rho(A_\sigma^{(t)})$ are continuous for all $\sigma \in \{0, 1\}^n$. So for all σ , the sets $P(\bar{\sigma}) = \left\{ t \in [1, \infty) : \rho(A_\sigma^{(t)})^{1/n} = \mathfrak{R}(\mathcal{A}^{(t)}) \right\}$ are closed in $[1, \infty)$, where $\bar{\sigma}$ denotes the class of σ modulo cyclic permutation.

If the cardinality of $\bar{\sigma}$ such that $P(\bar{\sigma}) \neq \emptyset$ was infinite countable, then the compact connected set $[1, \infty]$ would be partitioned in a countable family of non-empty closed sets, a contradiction (see [15, Thm. 6.1.27]). So, $P(\bar{\sigma})$ is empty for all but a finite number of $\bar{\sigma}$. But by connectedness, it happens that $[1, \infty) = P(\bar{\sigma})$ for a unique class $\bar{\sigma}$. Since $A_1^{(1)}$ is the transpose of $A_0^{(1)}$, the only option is the class of $\sigma = (0, 1) \in \{0, 1\}^2$, but for t large enough we have $\rho(A_\sigma^{(t)})^{1/n} < \mathfrak{R}(\mathcal{A}^{(t)})$, contradiction again. So, for some t_0 , $\mathcal{A}^{(t_0)}$ is a desired counterexample. \square

Remark 1.2.8. The continuity of the maps $t \mapsto \mathfrak{R}(\mathcal{A}^{(t)})$ and $t \mapsto \rho(A_\sigma^{(t)})$ also follows from the general results proved in Section 2.5.

Proof of Proposition 1.1.5. Just take $\Sigma = \tilde{\mathcal{A}}^{(t_0)}$ for t_0 found in Theorem 1.2.7. \square

1.3 Proof of Theorems 1.1.1 and 1.1.2

We begin with the proof of Theorem 1.1.1, which basically follows from (f.p.c.).

Proof of Theorem 1.1.1. We follow the arguments in [12, Chpt. 9, Lem. 2.2]. Fix x as base point and f isometry. Let $n \geq 2$ be an integer. Using (f.p.c.) on the points x, f^2x, fx and f^nx we obtain:

$$d(f^2x, x) + d(f^nx, fx) \leq \max(d(fx, x) + d(f^nx, f^2x), d(f^nx, x) + d(f^2x, fx)) + 2\delta.$$

As f is an isometry, if we define $a_n = d(f^nx, x)$, the inequality is equivalent to:

$$a_2 + a_{n-1} \leq \max(a_{n-2}, a_n) + a_1 + 2\delta. \quad (1.8)$$

Now, let $a = a_2 - a_1 - 2\delta$. We need to show that $a \leq d^\infty(f)$. If $a \leq 0$ there is nothing to prove. So, we assume that a is positive. We claim that $a + a_n \leq a_{n+1}$ for all $n \geq 1$, which is clear for $n = 1$. If we suppose it valid for some n , we know from (1.8):

$$a + a_{n+1} \leq \max(a_{n+2}, a_n).$$

If $a_{n+2} < a_n$, then

$$a_n < a + (a + a_n) \leq a + a_{n+1} \leq a_n$$

a contradiction. Therefore $a + a_{n+1} \leq a_{n+2}$, completing the proof of the claim.

So, by telescoping sum, $na \leq a_n$ for all n , and then

$$a \leq \lim_n \frac{a_n}{n} = d^\infty(f)$$

as we wanted to show. \square

Now we proceed with the proof of Theorem 1.1.2:

Proof of Theorem 1.1.2. First we suppose that $\delta > 0$. Let x be a base point and $f, g \in \text{Isom}(M)$. We use (f.p.c.) on the points x, fgx, fx and f^2x

$$d(fgx, x) + d(fx, x) \leq \max(d(fx, gx) + d(fx, x), d(f^2x, x) + d(gx, x)) + 2\delta, \quad (1.9)$$

and we separate into two cases:

$$\text{Case } i) \quad d(fx, gx) \leq \max(d(fx, x) + d^\infty(g), d(gx, x) + d^\infty(f)) + 4\delta:$$

Using this into (1.9) we obtain

$$\begin{aligned} d(fgx, x) &\leq \max(d(fx, gx), d(f^2x, x) + d(gx, x) - d(fx, x)) + 2\delta \quad (\text{by Thm. 1.1.1}) \\ &\leq \max(d(fx, gx), d^\infty(f) + d(gx, x) + 2\delta) + 2\delta \quad (\text{by Case } i) \\ &\leq \max(d^\infty(g) + d(fx, x), d^\infty(f) + d(gx, x)) + 6\delta \end{aligned}$$

completing the proof of the proposition in this case.

$$\text{Case } ii) \quad d(fx, gx) > \max(d(fx, x) + d^\infty(g), d(gx, x) + d^\infty(f)) + 4\delta:$$

Using this we get

$$d(f^2x, x) + d(gx, x) \leq d(fx, x) + d^\infty(f) + d(gx, x) + 2\delta < d(fx, x) + d(fx, gx) - 2\delta. \quad (1.10)$$

So, $\max(d(fx, x) + d(fx, gx), d(f^2x, x) + d(gx, x)) = d(fx, x) + d(fx, gx)$ and we obtain from (1.9) that

$$d(fgx, x) \leq d(fx, gx) + 2\delta. \quad (1.11)$$

Now, we use (f.p.c.) three times. First, on x, fx, fgx and f^2x :

$$d(fx, x) + d(fx, gx) \leq \max(d(fgx, x) + d(fx, x), d(f^2x, x) + d(gx, x)) + 2\delta.$$

But again by (1.10), it cannot happen that $d(fx, x) + d(fx, gx) \leq d(f^2x, x) + d(gx, x) + 2\delta$, so:

$$d(fx, gx) \leq d(fgx, x) + 2\delta$$

and combining with (1.11) we obtain:

$$|d(fgx, x) - d(fx, gx)| \leq 2\delta. \quad (1.12)$$

As our hypothesis is symmetric in f and g , an analogous reasoning allows us to conclude that

$$|d(gfx, x) - d(fx, gx)| \leq 2\delta. \quad (1.13)$$

Combining with (1.12) we obtain

$$|d(fgx, x) - d(gfx, x)| \leq 4\delta. \quad (1.14)$$

Next, we use (f.p.c.) on x, fgx, fx , and $fgfx$:

$$d(fgx, x) + d(gfx, x) \leq \max(2d(fx, x), d(fgfx, x) + d(gx, x)) + 2\delta. \quad (1.15)$$

But by (1.14) and assumption *ii*)

$$d(fgx, x) + d(gfx, x) \geq 2d(fx, gx) - 4\delta > 2(d(fx, x) + d^\infty(g) + 4\delta) - 4\delta > 2d(fx, x) + 2\delta.$$

So, using this with (1.14), in (1.15):

$$2d(fgx, x) \leq d(fgfx, x) + d(gx, x) + 6\delta. \quad (1.16)$$

Finally, by (f.p.c.) on $x, fgfx, fgx, (fg)^2x$ we obtain:

$$d(fgfx, x) + d(fgx, x) \leq \max(d(fgx, x) + d(gx, x), d((fg)^2x, x) + d(fx, x)) + 2\delta.$$

If the maximum in the right hand side were $d(fgx, x) + d(gx, x)$, we would have $d(fgfx, x) \leq d(gx, x) + 2\delta$. But then by (1.12) and (1.13):

$$\begin{aligned} 2d(fx, gx) - 4\delta &\leq (d(fgx, x) + d(gfx, x)) && \text{(by (1.15))} \\ &\leq \max(2d(fx, x), d(fgfx, x) + d(gx, x)) + 2\delta \\ &\leq 2 \max(d(fx, x), d(gx, x)) + 4\delta && \text{(by Case ii)} \\ &< 2d(fx, gx) - 4\delta. \end{aligned}$$

This contradiction and Theorem 1.1.1 applied to fg show us that

$$\begin{aligned} d(fgfx, x) &\leq d((fg)^2x, x) + d(fx, x) + 2\delta - d(fgx, x) \\ &\leq (d(fgx, x) + d^\infty(fg) + 2\delta) + d(fx, x) + 2\delta - d(fgx, x) \\ &\leq d(fx, x) + d^\infty(fg) + 4\delta. \end{aligned}$$

Using this with (1.16) we can finish:

$$\begin{aligned} d(fgx, x) &\leq (d(fgfx, x) + d(gx, x))/2 + 3\delta \\ &\leq (d^\infty(fg) + d(fx, x) + d(gx, x))/2 + 5\delta. \end{aligned}$$

In both cases our claim is true. To conclude the proof, note that a 0-hyperbolic space is δ -hyperbolic for all $\delta > 0$. \square

As a corollary of the proof of Theorem 1.1.2 we obtain a sufficient condition for a product of two isometries to be hyperbolic, and a lower bound for the stable length of the product, improving [12, Chpt. 9, Lem. 2.2]:

Proposition 1.3.1. *Let $K \geq 7\delta$ and $f, g \in \text{Isom}(M)$ be such that $d(fx, gx) > \max(d(fx, x) + d^\infty(g), d(gx, x) + d^\infty(f)) + K$ for some $x \in M$. Then fg is hyperbolic, and*

$$d^\infty(fg) > d^\infty(f) + d^\infty(g) + 2K - 14\delta.$$

Proof. Since $K \geq 4\delta$, we are in Case *ii*) of the previous proof. So we have $d(fgx, x) \leq (d^\infty(fg) + d(fx, x) + d(gx, x))/2 + 5\delta$. But by (1.12) and our assumption, we obtain

$$\begin{aligned} \frac{d(fx, x) + d(gx, x) + d^\infty(f) + d^\infty(g)}{2} &\leq \max(d(fx, x) + d^\infty(g), d(gx, x) + d^\infty(f)) \\ &< d(fx, gx) - K \\ &\leq d(fgx, x) + 2\delta - K \\ &\leq \frac{d(fx, x) + d(gx, x) + d^\infty(fg)}{2} + 7\delta - K \end{aligned}$$

The conclusion follows easily. \square

1.4 Berger-Wang and further properties of the stable length and joint stable length

1.4.1 A Berger-Wang Theorem for sets of isometries

Now we prove Theorem 1.1.4. We follow the arguments used in [5, Cor. 1]:

Proof of Theorem 1.1.4. It is clear that $\mathfrak{D}(\Sigma) \geq \limsup_{n \rightarrow \infty} d^\infty(\Sigma^n)/n$. Fixing a base point x and applying Corollary 1.1.3 to Σ^n we have

$$d(\Sigma^{2n}, x) \leq d(\Sigma^n, x) + d^\infty(\Sigma^{2n})/2 + 6\delta.$$

Dividing by n , taking \liminf when $n \rightarrow \infty$ and using that $\mathfrak{D}(\Sigma^2) = 2\mathfrak{D}(\Sigma)$, we obtain the result. \square

As a consequence we can describe the joint stable length of a bounded set of isometries in terms of the joint stable lengths of its finite non empty subsets.

Proposition 1.4.1. *If M is δ -hyperbolic then for every bounded set $\Sigma \subset \text{Isom}(M)$ we have:*

$$\mathfrak{D}(\Sigma) = \sup \{ \mathfrak{D}(F) : F \subset \Sigma \text{ and } F \text{ is finite and non empty} \}.$$

Proof. Let L be the supremum in the right hand side. Clearly $L \leq \mathfrak{D}(\Sigma)$. For the reverse inequality, let $\epsilon > 0$ and $n \geq 1$ be such that $|\mathfrak{D}(\Sigma) - d^\infty(\Sigma^n)/n| < \epsilon/2$. Also let $F = \{f_1, \dots, f_n\} \subset \Sigma$ be such that $d^\infty(\Sigma^n) \leq d^\infty(f_1 \cdots f_n) + \epsilon/2$. So we have

$$\begin{aligned} \mathfrak{D}(\Sigma) &\leq d^\infty(\Sigma^n)/n + \epsilon/2 \\ &< d^\infty(f_1 \cdots f_n)/n + \epsilon \\ &\leq \mathfrak{D}(F^n)/n + \epsilon \\ &= \mathfrak{D}(F) + \epsilon. \end{aligned}$$

Then it follows that $\mathfrak{D}(\Sigma) \leq L$. \square

1.4.2 Dynamical interpretation and semigroups of isometries

The stable length plays an important role in geometry and group theory (see e.g. [18] and the appendix in [17]). In this section we see its relation with isometries of Gromov hyperbolic spaces.

It is a well known fact that an isometry f of an hyperbolic metric space M belongs to exactly one of the following families:

- i) Elliptic:* if the orbit of some (and hence any) point by f is bounded.

- ii) *Parabolic*: if it is not elliptic and the orbit of some (and hence any) point by f has a unique accumulation point on the Gromov boundary ∂M .
- iii) *Hyperbolic*: if it is not elliptic and the orbit of some (and hence any) point by f has exactly two accumulation points on ∂M .

A proof of this classification for general hyperbolic spaces can be found in [13, Thm. 6.1.4], while for proper hyperbolic spaces this result is proved in [12, Chpt. 9, Thm. 2.1].

As we said in the introduction, an isometry of M is hyperbolic if and only if its stable length is positive. We want to extend this result for bounded sets of isometries. For our purpose we count with a classification for semigroups of isometries.

A semigroup $G \subset Isom(M)$ is:

- i) *Elliptic*: If Gx is a bounded subset of M for some (hence any) $x \in M$.
- ii) *Parabolic*: If it is not elliptic and there exists a unique point in ∂M fixed by all the elements of G .
- iii) *Hyperbolic*: If it contains some hyperbolic element.

An important result is that these are all the possibilities [13, Thm. 6.2.3]:

Theorem 1.4.2 (Das-Simmons-Urbański). *A semigroup $G \subset Isom(M)$ is either elliptic, parabolic or hyperbolic.*

So, as a corollary of Theorem 1.1.4 we obtain a criterion for hyperbolicity for a certain class of semigroups given by Theorem 1.1.7, that extends Proposition 1.1.6. Let Σ be a subset of $Isom(M)$ and denote by $\langle \Sigma \rangle$ the semigroup generated by Σ ; that is, $\langle \Sigma \rangle = \cup_{n \geq 1} \Sigma^n$.

Proof of Theorem 1.1.7. By Theorem 1.1.4, $\mathfrak{D}(\Sigma) > 0$ if and only if $d^\infty(\Sigma^n) > 0$ for some $n \geq 1$, which is equivalent to $d^\infty(f) > 0$ for some $f \in \cup_{n \geq 1} \Sigma^n = \langle \Sigma \rangle$. This is equivalent to the hyperbolicity of $\langle \Sigma \rangle$ by Proposition 1.1.6. \square

1.4.3 Relation with the minimal length

When we require M to be a *geodesic* space (i.e. every pair of points x, y can be joined by an arc isometric to an interval) we have another lower bound for the stable length. If $f \in Isom(M)$ define

$$d(f) = \inf_{x \in M} d(fx, x).$$

This number is called the *minimal length* of f . It is clear that $d^\infty(f) \leq d(f)$. On the other hand we have

Proposition 1.4.3. *If M is δ -hyperbolic and geodesic and $f \in Isom(M)$ then*

$$d(f) \leq d^\infty(f) + 16\delta.$$

For a proof of this proposition see [12, Chpt. 10, Prop. 6.4]. This gives us another lower bound for the joint stable length:

Proposition 1.4.4. *With the same assumptions of Proposition 1.4.3, for all bounded sets $\Sigma \subset \text{Isom}(M)$ we have:*

$$\sup_{f \in \Sigma} d(f) \leq \mathfrak{D}(\Sigma) + 16\delta.$$

Remark 1.4.5. A result similar to Proposition 1.4.3 is false if we do not assume M to be geodesic. Indeed, consider a δ -hyperbolic space M and $f \in \text{Isom}(M)$ with a fixed point and such that $\sup_{x \in M} d(fx, x) = \infty$. So, for all $R > 0$ the set $M_R = \{x \in M : d(fx, x) \geq R\}$ is a δ -hyperbolic space and f restricts to an isometry f_R of M_R . This is satisfied for example by every non-identity elliptic Möbius transformation in \mathbb{H}^2 . But $d^\infty(f_R) = d^\infty(f) = 0$ and $d(f_R) \geq R$.

This is one of the reasons, together with Proposition 1.1.6, we work with a generalization of the stable length instead of the minimal length (elliptic or parabolic isometries can satisfy $d(f) > 0$).

We finish this section showing that the generalizations of the minimum displacement and the stable distance in general may be different. This is the case of \mathbb{H}^2 :

Proposition 1.4.6. *There exists $\Sigma \subset \text{Isom}(\mathbb{H}^2)$ such that*

$$\mathfrak{D}(\Sigma) < \inf_{z \in \mathbb{H}^2} d(\Sigma, z).$$

Proof. Let $\mathcal{A} = \{F_0, F_1\}$ be a counterexample to the finiteness conjecture given by Theorem 1.2.7. We will prove that $\Sigma = \tilde{\mathcal{A}}$ satisfies our requirements.

Let $f_i = \tilde{F}_i$ for $i \in \{0, 1\}$. By the construction made in Subsection 1.2.2, it is a straightforward computation to see that f_0 and f_1 are hyperbolic isometries and that they have disjoint fixed point sets in $\partial\mathbb{H}^2 = \mathbb{R} \cup \{\infty\}$. Hence, by properties of hyperbolic geometry, given $K \geq 0$ the set $C_i(K) = \{z \in \mathbb{H}^2 : d(f_i z, z) \leq K\}$ is within bounded distance from the axis of f_i . We conclude that $C_0(K) \cap C_1(K)$ is compact and the map $z \rightarrow d(\Sigma, z) = \max(d(f_0 z, z), d(f_1 z, z))$ is proper.

Now suppose that $\mathfrak{D}(\Sigma) = \inf_{z \in \mathbb{H}^2} d(\Sigma, z)$ and let $(z_n)_n$ be a sequence in \mathbb{H}^2 such that $d(\Sigma, z_n) \rightarrow \mathfrak{D}(\Sigma)$. By the properness property the sequence $(z_n)_n$ must be bounded and by compactness we can suppose that it converges to $w \in \mathbb{H}^2$. So by continuity we have $\mathfrak{D}(\Sigma) = d(\Sigma, w)$. But then the set \mathcal{A} would have as extremal norm $\|A\| = \|SAS^{-1}\|_2$ where $S \in \text{SL}_2^\pm(\mathbb{R})$ satisfies $\tilde{S}w = i$, and by [34, Thm. 5.1], \mathcal{A} would satisfy the finiteness property, a contradiction. \square

1.5 Continuity

1.5.1 Continuity of the stable length

Now we study the continuity properties of the stable and joint stable lengths. Throughout the section we assume that $\text{Isom}(M)$ has the *finite-open topology*. It is generated by the subbasic open sets $\mathcal{G}(x, U) = \{f \in \text{Isom}(M) : f(x) \in U\}$

where $x \in M$ and U is open in M , and makes $Isom(M)$ a topological group [13, Prop. 5.1.3]. The finite-open topology is also called the *pointwise convergence topology* because of the following property [15, Prop. 2.6.5]:

Proposition 1.5.1. *A net $(f_\alpha)_{\alpha \in A} \subset Isom(M)$ converges to f if and only if $(f_\alpha x)_{\alpha \in A}$ converges to fx for all $x \in M$.*

Corollary 1.5.2. *For all $n \in \mathbb{Z}$ and $x \in M$ the function from $Isom(M)$ to \mathbb{R} that maps f to $d(f^n x, x)$ is continuous.*

Proof. As $Isom(M)$ is a topological group, by Proposition 1.5.1 the function $f \mapsto f^n x$ is continuous for all $x \in M$ and $n \in \mathbb{Z}$. The conclusion follows by noting that the map $f \mapsto d(f^n x, x)$ is a composition of continuous functions. \square

With Corollary 1.5.2 we can prove Theorem 1.1.9:

Proof of Theorem 1.1.9. We follow an idea of Morris (see [38]). By subadditivity, $d^\infty(f)$ is the infimum of continuous functions, hence is upper semi-continuous. For the lower semi-continuity, Theorem 1.1.1 implies that for any $x \in M$:

$$d^\infty(f) = \sup_{n \geq 1} \frac{d(f^{2n} x, x) - d(f^n x, x) - 2\delta}{n}.$$

So $d^\infty(f)$ is also the supremum of continuous functions. \square

1.5.2 Vietoris topology and continuity of the joint stable length

For the continuity of the joint stable length we need to work in the correct space. A natural candidate is $\mathcal{B}(Isom(M))$, the space of non empty bounded sets of $Isom(M)$. Also, let $\mathcal{BF}(Isom(M))$ be the set of closed and bounded subsets of $Isom(M)$. First of all, by the following lemma it is sufficient to consider closed (and bounded) sets of isometries:

Lemma 1.5.3. *If $\Sigma \in \mathcal{B}(Isom(M))$ then:*

- i) $\overline{\Sigma} \in \mathcal{BF}(Isom(M))$.
- ii) $d(\overline{(\Sigma^n)}, x) = d((\overline{\Sigma})^n, x) = d(\Sigma^n, x)$ for all $x \in M, n \geq 1$.
- iii) $\mathfrak{D}(\overline{\Sigma}) = \mathfrak{D}(\Sigma)$.

Proof. Assertion i) is trivial and iii) is immediate from ii). For the latter, let $f \in \overline{\Sigma}$ and f_α a net in Σ converging to f . As $d(f_\alpha x, x) \leq d(\Sigma, x)$ for all α , then $d(fx, x) \leq d(\Sigma, x)$. So $d(\overline{\Sigma}, x) \leq d(\Sigma, x) \leq d(\overline{\Sigma}, x)$ and

$$d(\overline{\Sigma}, x) = d(\Sigma, x). \quad (1.17)$$

Now, let $g = f^{(1)} f^{(2)} \dots f^{(n)} \in \overline{\Sigma}^n$ with $f^{(i)} \in \overline{\Sigma}$. There exist nets $(f_\alpha^{(i)})_{\alpha \in A_i}$ such that $f_\alpha^{(i)}$ tends to $f^{(i)}$ for all i . But since $Isom(M)$ is topological group,

$f_\alpha = f_{\alpha_1}^{(1)} f_{\alpha_2}^{(2)} \cdots f_{\alpha_n}^{(n)}$ (with $\alpha = (\alpha_1, \dots, \alpha_n) \in A_1 \times \cdots \times A_n$) defines a net in Σ^n that tends to g . We conclude that $(\overline{\Sigma})^n \subset \overline{(\Sigma^n)}$ and by (1.17) we obtain

$$d(\Sigma^n, x) \leq d((\overline{\Sigma})^n, x) \leq d(\overline{(\Sigma^n)}, x) = d(\Sigma^n, x).$$

The conclusion follows. \square

Our next step is to define a topology on $\mathcal{BF}(Isom(M))$. We follow the construction given by E. Michael [37]. Let $\mathcal{P}(Isom(M))$ be the set of non empty subsets of M . If U_1, \dots, U_n are non empty open sets in $Isom(M)$ let

$$\langle U_1, \dots, U_n \rangle := \left\{ E \in \mathcal{P}(Isom(M)) : E \subset \bigcup_i U_i \text{ and } E \cap U_i \neq \emptyset \text{ for all } i \right\}.$$

The *Vietoris topology* on $\mathcal{P}(Isom(M))$ is the one which has as base the collection of sets $\langle U_1, \dots, U_n \rangle$. We say that a subset of $\mathcal{P}(Isom(M))$ with the induced topology also has the Vietoris topology.

With this in mind the space $\mathcal{BF}(Isom(M))$ satisfies one of our requirements:

Proposition 1.5.4. *For all $x \in M$ the map $\Sigma \mapsto d(\Sigma, x)$ is continuous on $\mathcal{BF}(Isom(M))$.*

Proof. It follows from Theorem 1.1.9 and the fact that taking supremum preserves continuity on $\mathcal{BF}(Isom(M))$, see [37, Prop. 4.7]. \square

For the continuity of the composition map $(\Sigma, \Pi) \mapsto \Sigma\Pi$ we must impose further restrictions. So we work on $\mathcal{C}(Isom(M))$, the set of non empty compact subsets of $Isom(M)$. In this space all our claims are satisfied:

Proof of Theorem 1.1.11. The idea of the proof of the continuity of the joint stable length is the same one that we used in the proof of Theorem 1.1.9. We claim that in $\mathcal{C}(Isom(M))$ the maps $\Sigma \mapsto d(\Sigma, x)$ and $\Sigma \mapsto \Sigma^n$ are continuous for all $x \in M$ and $n \in \mathbb{Z}^+$. The first assertion is Proposition 1.5.4 and the second one comes from a general result in topological groups. We prove it in Appendix A (see Corollary A.3). Similarly the continuity of the stable length follows as in the proof of Proposition 1.5.4. \square

It follows from Theorem 1.1.11 that the joint stable length is continuous on the set of non empty finite subsets of $Isom(M)$. This affirmation together with Proposition 1.4.1 allows us to conclude a semi-continuity result on $\mathcal{BF}(Isom(M))$:

Theorem 1.5.5. *The map $\mathfrak{D}(\cdot) : \mathcal{BF}(Isom(M)) \rightarrow \mathbb{R}$ is lower semi-continuous.*

Proof. Let $\epsilon > 0$ and $\Sigma \in \mathcal{BF}(Isom(M))$. By Proposition 1.4.1 there is $F \subset \Sigma$ finite with $\mathfrak{D}(\Sigma) - \mathfrak{D}(F) < \epsilon/2$. As $F \in \mathcal{C}(Isom(M))$, by Theorem 1.1.11 there exist open sets $U_1, \dots, U_n \subset Isom(M)$ such that $V = \langle U_1, \dots, U_n \rangle$ is an open neighborhood of B and if G is finite and $G \in V$ then $|\mathfrak{D}(F) - \mathfrak{D}(G)| < \epsilon/2$.

Let $W = \langle Isom(M), U_1, \dots, U_n \rangle$. Clearly W is an open neighborhood of Σ , and if $A \in W$, then there exist f_1, \dots, f_n with $f_i \in A \cap U_i$ for all i . So $H = \{f_1, \dots, f_n\} \in V$ and then $|\mathfrak{D}(F) - \mathfrak{D}(H)| < \epsilon/2$. We have

$$\mathfrak{D}(\Sigma) < \mathfrak{D}(F) + \epsilon/2 < \mathfrak{D}(H) + \epsilon \leq \mathfrak{D}(A) + \epsilon$$

and the conclusion follows.

□

Chapter 2

A New Inequality about Matrix Products and a Berger-Wang Formula

2.1 Introduction

Let k be a field, and let $M_d(k)$ be the algebra of $d \times d$ matrices with coefficients in k . If $k = \mathbb{R}$ or \mathbb{C} , let $\|\cdot\|$ be any norm on k^d , with the corresponding operator norm on $M_d(k)$ also denoted by $\|\cdot\|$. The spectral radius of a matrix A will be denoted by $\rho(A)$. Given a bounded set $\mathcal{M} \subset M_d(k)$, the *joint spectral radius* of \mathcal{M} is defined by the formula

$$\mathfrak{R}(\mathcal{M}) = \lim_{n \rightarrow \infty} (\sup \{\|A_1 \cdots A_n\| : A_i \in \mathcal{M}\})^{1/n}. \quad (2.1)$$

By a submultiplicative argument, this quantity is well defined and finite, and the limit in the right hand side of (2.1) can be replaced by the infimum over n .

The joint spectral radius was introduced by Rota and Strang [46], and for a set $\mathcal{M} \subset M_d(k)$, represents the maximal exponential growth rate of the partial sequence of products $(A_1 \cdots A_n)_n$ of a sequence of matrices A_1, A_2, \dots with $A_i \in \mathcal{M}$. For this reason, this quantity has appeared in several mathematical contexts, making it an important object of study (see e.g. [22, 29, 39, 48]). In particular, the question of whether the joint spectral radius may be approximated by periodic sequences plays an important role. The Berger-Wang formula gives a positive answer to this question in the case of bounded sets of matrices [2]:

Theorem 2.1.1 (Berger-Wang formula). *If $\mathcal{M} \subset M_d(\mathbb{C})$ is bounded, then*

$$\mathfrak{R}(\mathcal{M}) = \limsup_{n \rightarrow \infty} (\sup \{\rho(A_1 \cdots A_n) : A_i \in \mathcal{M}\})^{1/n}. \quad (2.2)$$

This result has been generalized by Morris, to the context of linear cocycles (including infinite dimensional ones) [40], by using multiplicative ergodic theory. In the finite dimensional case, the problem of finding a formula similar to (2.2), when there is a Markov-type constraint on the allowed products was presented

by Kozyakin [33]. Although the result of Morris already applies to this kind of constraints, the novelty in Kozyakin's proof is that his arguments are purely linear algebraic, and are consequences of Theorem 2.1.1.

Another tool to obtain results related to joint spectral radius was found by J. Bochi in [5]. In that work, he proved some inequalities that may be seen as lower bounds for spectral radii of sets of matrices in terms of the norms of such matrices. Following that method, the purpose of this chapter is to present an inequality relating the norm of the product of matrices with the spectral radii of subproducts. We will give an upper bound for the norm of the product of matrices $A_N \cdots A_1$ in terms of the spectral radii of its subproducts $A_\beta A_{\beta-1} \cdots A_{\alpha+1} A_\alpha$. This inequality will allow us to obtain relations similar to (2.2). It holds in an arbitrary local field where the notions of absolute value, norm, and spectral radius are well defined (see Section 2.4 for a detailed explanation). Our main result is the following:

Theorem 2.1.2. *Let $d \in \mathbb{N}$, k be a local field, and $\|\cdot\|$ be a norm on $M_d(k)$. Then there exist constants $N = N(d) \in \mathbb{N}$, $0 < \delta = \delta(d, k) < 1$, and $C = C(d, k, \|\cdot\|) > 1$ satisfying the following inequality for all $A_1, \dots, A_N \in M_d(k)$:*

$$\|A_N \cdots A_1\| \leq C \left(\prod_{1 \leq i \leq N} \|A_i\| \right) \max_{1 \leq \alpha \leq \beta \leq N} \left(\frac{\rho(A_\beta \cdots A_\alpha)}{\prod_{\alpha \leq i \leq \beta} \|A_i\|} \right)^\delta, \quad (2.3)$$

where the right hand side is treated as zero if one of the A_i is the zero matrix.

So if the norm of the product $A_N \cdots A_1$ is comparable to (that is, not much smaller than) the product of the norms, then there exists a subproduct $A_\beta \cdots A_\alpha$ whose spectral radius is comparable to (that is, not much smaller than) $\prod_{\alpha \leq i \leq \beta} \|A_i\|$.

Note that inequality (2.3) is homogeneous in each variable A_i . We will later show that $N(d) \leq \prod_{i=1}^{d-1} \binom{d}{i} = \prod_{i=0}^d \binom{d}{i}$ for all $d \in \mathbb{N}$. We will also show that this upper bound is not sharp, because $N(3) \leq 5$ (see Proposition 2.2.4). In addition, when $k = \mathbb{C}$, the constant C in (2.3) may be chosen independent of the norm $\|\cdot\|$, provided that $\|\cdot\|$ is an operator norm (see Proposition 2.4.2).

The approach of using inequalities to prove results similar to (2.2) also has been applied by I. Morris to study matrix pressure functions [38] and by the author in the context of isometries in Gromov hyperbolic spaces [44]. The novelty of the inequality presented here is that it respects the order in which the matrices are multiplied. While previous works considered a sum or a maximum over *all* possible subproducts of length N with respect to a given alphabet of matrices, in Theorem 2.1.2 we consider just one product of length N together with its subproducts. This distinction allows inequality (2.3) to be used in cases where only some specific kinds of products are allowed.

The proof of this inequality is based in the non trivial case of equality, where the right hand side of (2.3) is zero. This occurs when $\rho(A_j \cdots A_i) = 0$ for all $1 \leq i < j \leq N$, that is, when $A_j \cdots A_i$ are all nilpotent. Denote by $\mathcal{N}_d(k)$ the set of nilpotent elements of $M_d(k)$. Then define, for $n \geq 1$, the set $\mathcal{N}_d^n(k)$ of n -tuples $(A_1, \dots, A_n) \in M_d(k)^n$ such that $A_j \cdots A_i \in \mathcal{N}_d(k)$ for all $1 \leq i < j \leq n$. The particular case of (2.3) that we highlighted can be restated as follows:

Theorem 2.1.3. *For all $d \geq 1$ there exists an integer $N = N(d) \geq 1$ such that, for every field k , if $(A_1, \dots, A_N) \in \mathcal{N}_d^N(k)$, then the product $A_N \cdots A_1$ is zero.*

The proof of Theorem 2.1.3 is purely linear algebraic, exploiting the properties of the n -exterior power functor. This result may be compared with Levitzki's Theorem (see [45, Thm. 2.1.7]), that asserts that for an algebraically closed field k , every semigroup $S \subset M_d(k)$ of nilpotent matrices is simultaneously triangularizable. That is, there is some $B \in \text{GL}_d(k)$ such that BAB^{-1} is upper triangular with zero diagonal for every $A \in S$. In particular, if $A_1, \dots, A_d \in S$, then the product $A_1 \cdots A_d$ is zero. As we show in Subsection 2.2.1, the optimal $N(d)$ in Theorem 2.1.3 is in general bigger than d . Therefore the matrices satisfying the hypothesis of the Theorem admit no normal form as simple as in Levitzki's Theorem.

Applications to Ergodic theory. Let (X, \mathcal{F}, μ) be a probability space, and let $T : X \rightarrow X$ be a measure preserving map. By a *linear cocycle* over X , we mean a measurable map $A : X \rightarrow M_d(k)$ together with the family of maps A^n defined by the formula

$$A^n(x) = A(T^{n-1}x) \cdots A(Tx)A(x), \quad \text{for } n \geq 1, x \in X.$$

These maps satisfy the multiplicative cocycle relation $A^{m+n}(x) = A^m(T^n x)A^n(x)$ for all $m, n \geq 1, x \in X$.

We usually denote a linear cocycle by $\mathcal{A} = (X, T, A)$, and say that \mathcal{A} is *integrable* if $\max(\log \|A\|, 0)$ is integrable. In this case, Kingman's theorem implies that, for μ -almost all $x \in X$, the limit $\lambda(x) = \lim_{n \rightarrow \infty} \frac{\log \|A^n(x)\|}{n} \in [-\infty, \infty)$ exists, and moreover, λ is T -invariant. This function is the upper *Lyapunov exponent* of A , and is one of the most important concepts in multiplicative ergodic theory.

As an application of our inequality, we reprove the following theorem due to I. Morris (first tested numerically in [20] and proved by Avila-Bochi for $\text{SL}_2(\mathbb{R})$ in [1, Thm. 15]).

Theorem 2.1.4. [40, Thm. 1.6] *Let T be a measure-preserving transformation of a probability space (X, \mathcal{F}, μ) and let $A : X \rightarrow M_d(k)$ be an integrable linear cocycle. If λ is as before, then for μ -almost all $x \in X$ we have*

$$\limsup_{n \rightarrow \infty} \frac{\log(\rho(A^n(x)))}{n} = \lambda(x). \quad (2.4)$$

While Morris's proof of this result relies on Oseledets Theorem, we will mainly use Theorem 2.1.2 and a quantitative version of Poincaré's recurrence Theorem.

2.2 Proof of Theorem 2.1.3

We begin the proof of Theorem 2.1.3 with some useful results. For a given vector space V (over an arbitrary field), let $\text{End}(V)$ be the algebra of linear endomorphisms of V . The dimension of the image of a linear transformation $T \in \text{End}(V)$ will be denoted as $\text{rank}(T)$. Also, let $\mathcal{N}^n(V)$ be the set of n -tuples $(T_1, \dots, T_n) \in \text{End}(V)^n$ such that $T_j \cdots T_i$ is nilpotent for all $1 \leq i \leq j \leq n$. With our previous notation, we have $\mathcal{N}^n(k^d) = \mathcal{N}_d^n(k)$.

Proposition 2.2.1. *Let $1 \leq n \leq \dim V$ and $(T_1, \dots, T_n) \in \mathcal{N}^n(V)$ be such that $\text{rank}(T_j) \leq 1$ for all $1 \leq j \leq n-1$. If $v \in V$ and $T_n \cdots T_1 v \neq 0$, then $v, T_1 v, T_2 T_1 v, \dots, T_n \cdots T_1 v$ are all distinct and form a linearly independent set.*

Proof. We will use induction on n . The case $n = 1$ comes from the nilpotence of T_1 . So, assume that the result holds for tuples in $\mathcal{N}^{n-1}(V)$ and let $(T_1, \dots, T_n) \in \mathcal{N}^n(V)$ and $v \in V$ be as in the hypothesis. Take a linear combination of $v, T_1v, \dots, T_n \cdots T_1v$ of the form

$$\lambda_0 v + \lambda_1 T_1 v + \cdots + \lambda_{n-1} T_{n-1} \cdots T_1 v + \lambda_n T_n \cdots T_1 v = 0, \quad (2.5)$$

and suppose that this linear combination is non trivial. As $(T_1, \dots, T_{n-1}) \in \mathcal{N}^{n-1}(V)$ also satisfies the hypothesis with respect to v , by our inductive assumption we have $\lambda_n \neq 0$. Now, apply $T_n \cdots T_1$ in (2.5). The rank condition over the maps T_j and the fact that $(T_1, \dots, T_n) \in \mathcal{N}^n(V)$ imply that $(T_j \cdots T_1)^2 = 0$, for all $1 \leq j \leq n$. Hence, the left hand side of (2.5) becomes $\lambda_0 T_n \cdots T_1 v$, forcing $\lambda_0 = 0$. But in that case, equation (2.5) would be a non trivial linear combination of $\{w, T_2 w, T_3 T_2 w, \dots, T_n \cdots T_2 w\}$, with $w = T_1 v$. This is impossible by our inductive assumption, since $(T_2, \dots, T_n) \in \mathcal{N}^{n-1}(V)$ satisfies the hypothesis of the proposition with respect to w . We conclude that all linear combinations of $v, T_1 v, T_2 T_1 v, \dots, T_n \cdots T_1 v$ of the form (2.5) are trivial, and hence this set is linearly independent with exactly $n + 1$ elements. \square

Corollary 2.2.2. *If $(T_1, \dots, T_d) \in \mathcal{N}^d(V)$ and $\text{rank}(T_j) \leq 1$ for all $1 \leq j \leq d - 1$, then $T_d \cdots T_1 = 0$.*

Proof. Assume the contrary and let $v \in V$ be such that $T_d \cdots T_1 v \neq 0$. Then by Proposition 2.2.1, the set $\{v, T_1 v, T_2 T_1 v, \dots, T_d \cdots T_1 v\}$ would be a linearly independent set of cardinality greater than $\dim V$. A contradiction. \square

For the next steps in our proof we need some fact about exterior powers. Recall that if V is a vector space of dimension d , the r -fold exterior power $\Lambda^r V$ is the vector space of alternating r -linear forms on the dual space V^* (see e.g. [35, XIX.1]). Given a basis $\{v_1, \dots, v_d\}$ of V , the set $\{v_{i_1} \wedge \cdots \wedge v_{i_r} : 1 \leq i_1 < \cdots < i_r \leq d\}$ is a basis of $\Lambda^r V$. Hence $\dim \Lambda^r V = \binom{d}{r}$.

The exterior power also induces a map $\Lambda^r : \text{End}(V) \rightarrow \text{End}(\Lambda^r V)$ given by the linear extension of $(\Lambda^r T)(w_1 \wedge \cdots \wedge w_r) = (T w_1 \wedge \cdots \wedge T w_r)$. This map is functorial: The relation $\Lambda^r(ST) = \Lambda^r(S) \Lambda^r(T)$ holds for all $S, T \in \text{End}(V)$. This induces a map $\Lambda^r : \mathcal{N}(V) \rightarrow \mathcal{N}(\Lambda^r V)$ that extends to $\mathcal{N}^n(V) \rightarrow \mathcal{N}^n(\Lambda^r V)$ for all $n \geq 1$.

Another important fact is that, when $T \in \mathcal{N}(V)$ and $\text{rank}(T) = r > 0$, then $\text{rank}(\Lambda^r T) = 1$. This is because the image of $\Lambda^r T$ is generated by any r -form associated to the r -dimensional subspace $T(V)$. This remark is crucial in the end of our proof.

Lemma 2.2.3. *Let $1 \leq r \leq d$ and $m = \binom{d}{r}$. Given $(T_1, \dots, T_m) \in \mathcal{N}^m(V)$, with $\text{rank}(T_j) \leq r$ for all $1 \leq j \leq m - 1$, we have $\text{rank}(T_m \cdots T_1) < r$.*

Proof. If that is the case then we will have $\text{rank}(T_j T_{j-1} \cdots T_i) \leq r$ for all $1 \leq i \leq j \leq m - 1$. Then the tuple $(\Lambda^r T_1, \dots, \Lambda^r T_m) \in \mathcal{N}^m(\Lambda^r V)$ will satisfy the hypothesis of Corollary 2.2.2, and hence $\Lambda^r(T_m \cdots T_1) = 0$, which implies that $\text{rank}(T_m \cdots T_1) < r$. \square

Proof of Theorem 2.1.3. Let $1 \leq l < d$ and $r(l) = \binom{d}{1} \cdots \binom{d}{l}$. We claim that for all $(T_1, \dots, T_{r(l)}) \in \mathcal{N}^{r(l)}(V)$ we have $\text{rank}(T_{r(l)} \cdots T_1) < d - l$. If so, the result follows with $N = r(d - 1) = \binom{d}{1} \cdots \binom{d}{d-1}$.

We will argue by induction. The case $l = 1$ is Lemma 2.2.3 with $r = d - 1$. Now, assume the result for some $l < d$, and for $1 \leq j \leq \binom{d}{l+1}$, define $\hat{T}_j = T_{r(l)j} \cdots T_{r(l)(j-1)+1}$. Then $(\hat{T}_1, \dots, \hat{T}_{\binom{d}{l+1}}) \in \mathcal{N}^{\binom{d}{l+1}}(V)$, and by our inductive hypothesis, we obtain $\text{rank}(\hat{T}_j) \leq d - l - 1$. So, we are in the assumption of 2.2.3 with $r = d - l - 1$ and we conclude that $\text{rank}(T_{r(l+1)} \cdots T_1) = \text{rank}(\hat{T}_{\binom{d}{l+1}} \cdots \hat{T}_1) < d - l - 1$. This proves the claim and concludes the proof of the theorem. \square

2.2.1 Some computations in low dimension

Let $N(d)$ be the least value of N for which Theorem 2.1.3 (and therefore also Theorem 2.1.2) holds true. From the proof of Theorem 2.1.3, we can obtain the bound $N(d) \leq \binom{d}{1} \binom{d}{2} \cdots \binom{d}{d-1}$ for all d . Also, since for all d we can construct a matrix $A \in \mathcal{N}_d(k)$ of rank $d - 1$, the tuple $(A, \dots, A) \in \mathcal{N}^{d-1}(k^d)$ satisfies $A^{d-1} \neq 0$ and hence we have the lower bound $N(d) \geq d$. In particular, we conclude that $N(2) = 2$, and for higher dimensions we get the bounds $3 \leq N(3) \leq 9$ and $4 \leq N(4) \leq 96$. We end this section by finding a better bound for $N(3)$.

Proposition 2.2.4. *For any field k , we have $N(3) \leq 5$. In addition, if $\text{char } k \neq 2$, then $N(3) = 5$.*

To prove this, we need a lemma:

Lemma 2.2.5. *Let $(C, B, A) \in \mathcal{N}^3(k^3)$. If $\text{rank } B = 1$, then $AB = \lambda B$ or $BC = \lambda B$ for some $\lambda \in k$.*

Proof. Assume that $B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ and $C = \begin{pmatrix} p & q & r \\ s & t & u \\ v & w & x \end{pmatrix}$.

Then $AB = \begin{pmatrix} 0 & 0 & b \\ 0 & 0 & e \\ 0 & 0 & h \end{pmatrix}$ and $BC = \begin{pmatrix} 0 & 0 & 0 \\ v & w & x \\ 0 & 0 & 0 \end{pmatrix}$. The nilpotence of AB and BC

implies $h = \text{Tr } AB = w = \text{Tr } BC = 0$. Then $ABC = \begin{pmatrix} bv & 0 & bx \\ ev & 0 & ex \\ 0 & 0 & 0 \end{pmatrix}$, and by the nilpotence of ABC , $bv = \text{Tr } ABC = 0$. The case $b = 0$ is $AB = eB$ and the case $v = 0$ is $BC = xB$. \square

Corollary 2.2.6. *If $(C, B, A) \in \mathcal{N}^3(k^3)$ and $\text{rank}(B) \leq 1$, then $ABC = 0$.*

Proof. Assume that $\text{rank}(B) = 1$, and $A, C \neq 0$. By Lemma 2.2.5 and after rescaling A or C , we may suppose that $BC = B$ or $AB = B$. In the first case we will have $(C^2, B, A) \in \mathcal{N}^3(k^3)$, and by Corollary 2.2.2, $ABC = A(BC)C = ABC^2 = 0$. For the case $AB = B$, applying a similar argument to the tuple (A^t, B^t, C^t) of the transposes of A, B, C , we will obtain $(ABC)^t = C^t B^t (A^t)^2 = 0$, and hence $ABC = 0$. \square

Proof of Proposition 2.2.4. Let $(E, D, C, B, A) \in \mathcal{N}^5(k^3)$. Then (E, BCD, A) belongs to $\mathcal{N}^3(k^3)$, and by Lemma 2.2.3 with $d = 3, r = 2$, $\text{rank}(BCD) \leq 1$. Then,

by Corollary 2.2.6, $ABCDE = 0$ and $N(3) \leq 5$. Moreover, when $\text{char } k \neq 2$, it is a straightforward computation to show that $(D, C, B, A) \in \mathcal{N}^4(k^3)$, with

$$A = \begin{pmatrix} -2 & -6 & 1 \\ 3 & 9 & 16 \\ -1 & -3 & -7 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, C = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 4 \\ -1 & -1 & -2 \end{pmatrix}, D = \begin{pmatrix} -1 & 3 & 16 \\ 1 & -3 & -16 \\ 1 & 2 & 4 \end{pmatrix}$$

$$\text{and } ABCD = \begin{pmatrix} 4 & 8 & 16 \\ -6 & -12 & -24 \\ 2 & 4 & 8 \end{pmatrix} \neq 0. \quad \square$$

Remark 2.2.7. This last proposition shows that, in general we cannot expect $N(d) = d$. For that reason, the hypothesis of Theorem 2.1.3 does not imply any kind of simultaneous triangularization. In fact, it is not hard to prove that the matrices in the last example we gave in $\mathcal{N}^4(k^3)$ are not simultaneously triangularizable, since A and B do not have a common invariant subspace of dimension 1.

2.3 A polynomial identity

For the proof of Theorem 2.1.2 we need some notation. Let k be a field with algebraic closure \bar{k} . For $d, N \in \mathbb{N}$, consider Nd^2 variables $x_{i,j}$ with $1 \leq i \leq N$, $1 \leq j \leq d^2$ and let $R_{d,N}$ be the polynomial ring $k[x_{i,j}]$. If $A_1, \dots, A_N \in M_d(\bar{k})$ and $f \in R_{d,N}$, by $f(A_1, \dots, A_N)$ we mean the element $f((a_{i,j})_{i,j})$ where $(a_{i,j})_j$ are the coefficients of A_i in some fixed order.

Recall that a polynomial $f \in k[y_1, \dots, y_m]$ is *homogeneous* of degree $\lambda \geq 0$ if it is of the form $\sum_{i_1 + \dots + i_m = \lambda} c_{i_1 \dots i_m} y_1^{i_1} \dots y_m^{i_m}$ for some $c_{i_1 \dots i_m} \in k$, $i_1, \dots, i_m \geq 0$. We say that monomial $f \in R_{d,N}$ is *multihomogeneous* of degree $\text{deg } f = (\lambda_1, \dots, \lambda_N) \in (\mathbb{N} \cup \{0\})^N$ if it is of the form $f((x_{i,j})_{i,j}) = c \prod_{i,j} x_{i,j}^{u_{i,j}}$, where $c \in k$, $u_{i,j} \geq 0$ and $\sum_j u_{i,j} = \lambda_i$ for all $1 \leq i \leq N$, and that a polynomial $p \in R_{d,N}$ is said to be multihomogeneous of degree $\text{deg } p$ if it is a finite sum of multihomogeneous monomials of degree $\text{deg } p$. This is equivalent to say that, for each $1 \leq i \leq N$, p is homogeneous of degree λ_i in the variables $x_{i,1} \dots x_{i,d^2}$.

For $1 \leq j \leq d^2$ denote by f_j the polynomial in $R_{d,N}$ representing the map that sends the N -tuple $(A_1, \dots, A_N) \in \bar{k}^{Nd^2}$ to the j -th entry of $A_N \dots A_1$. Also, for $1 \leq \ell \leq d$ and $1 \leq \alpha \leq \beta \leq N$, let $T_{\alpha,\beta}^\ell \in R_{d,N}$ be the polynomial that represents the map $(A_1, \dots, A_N) \mapsto \text{Tr } \Lambda^\ell(A_\beta \dots A_\alpha)$.

It is not hard to see that f_j are multihomogeneous of degree $(1, 1, \dots, 1, 1)$ and that $T_{\alpha,\beta}^\ell$ are multihomogeneous of degree $(0, \dots, 0, \ell, \dots, \ell, 0, \dots, 0)$, with the ℓ 's in positions $\alpha, \alpha + 1, \dots, \beta$.

Our purpose is to prove the following:

Theorem 2.3.1. *If $N = N(d)$ is given by Theorem 2.1.3, there is some $r \in \mathbb{N}$ such that for all $1 \leq j \leq d^2$ there exist multihomogeneous polynomials $p_{j,\ell}^{\alpha,\beta} \in R_{d,N}$ of degree $r \text{deg } f_j - \text{deg } T_{\alpha,\beta}^\ell \in (\mathbb{N} \cup \{0\})^N$ such that*

$$(f_j)^r = \sum_{\alpha,\beta,\ell} p_{j,\ell}^{\alpha,\beta} T_{\alpha,\beta}^\ell. \quad (2.6)$$

The natural tool to prove this result is Hilbert's Nullstellensatz. If $I \subset k[y_1, \dots, y_m]$ is a homogeneous ideal (i.e. generated by homogeneous polynomials), let $Z(I)$ be its zero locus in $\mathbb{P}^{m-1} = \mathbb{P}^{m-1}(\bar{k})$. Also, for $Z \subset \mathbb{P}^{m-1}(\bar{k})$, let $I(Z) \in k[y_1, \dots, y_m]$ be the homogeneous ideal of polynomials f that vanish on Z . The statement of Projective Nullstellensatz is the following (see e.g. [16, Sec. 4.2] and [27, Thm. 30.6]).

Theorem 2.3.2 (Projective Nullstellensatz). *If $I \subset k[y_1, \dots, y_m]$ is a homogeneous ideal, then $I(Z(I))$ is equal to \sqrt{I} , the radical ideal of I , provided that $Z(I) \neq \emptyset$.*

Proof of Theorem 2.3.1. Consider $\bar{k}^{(d^2)^N}$ with coordinates $(z_{i_1 \dots i_N})_{1 \leq i_1, \dots, i_N \leq d^2}$. Let $\varphi : (\bar{k}^{d \times d})^N \rightarrow \bar{k}^{(d^2)^N}$ be the Segre map such that the $(i_1 \dots i_N)$ -coordinate of $\varphi(A_1, \dots, A_N)$ is $a_{1, i_1} \dots a_{N, i_N}$, where a_{j, i_j} is the i_j -th coordinate of A_j . Let $\hat{\varphi} : (\mathbb{P}^{d^2-1})^N \rightarrow \mathbb{P}^{(d^2)^N-1}$ be the induced projective Segre embedding. As $\text{Im } \hat{\varphi} \subset \mathbb{P}^{(d^2)^N-1}$ is an algebraic set there is a homogeneous ideal $J \subset k[z_{i_1 \dots i_N}]$ such that $\text{Im } \hat{\varphi} = Z(J)$ (for references see e.g. [25, Ex. 2.11]). Let $I \subset R_{d, N}$ be the ideal generated by the polynomials $T_{\alpha, \beta}^\ell$ and let W be the zero locus of I in $(\mathbb{P}^{d^2-1})^N$ (which is well defined since $T_{\alpha, \beta}^\ell$ are multihomogeneous).

Given α, β, ℓ and $\gamma = (j_1, \dots, j_{\alpha-1}, j_{\beta+1}, \dots, j_N) \in \{1, \dots, d^2\}^{N-\beta+\alpha-1}$, define $u_\gamma \in R_{d, N}$ as $u_\gamma(A_1, \dots, A_N) = (a_{1, j_1} \dots a_{(\alpha-1), j_{\alpha-1}} a_{(\beta+1), j_{\beta+1}} \dots a_{N, j_N})^\ell$, with the convention $a_{1, j_1} \dots a_{0, j_0} = a_{(N+1), j_{N+1}} \dots a_{N, j_N} = 1$. Also, let $S_{\alpha, \beta, \gamma}^\ell \in k[z_{i_1 \dots i_N}]$ be the homogeneous polynomial of degree ℓ such that

$$S_{\alpha, \beta, \gamma}^\ell(\varphi(A_1, \dots, A_N)) = T_{\alpha, \beta}^\ell(A_1, \dots, A_N)(u_\gamma(A_1, \dots, A_N))^\ell \quad (2.7)$$

for all $A_1, \dots, A_N \in \bar{k}^{d^2}$. In a similar way, define $g_j \in k[z_{i_1 \dots i_N}]$ as the homogeneous polynomial of degree 1 such that $g_j \circ \varphi = f_j$. It is clear that, for $P \in (\mathbb{P}^{d^2-1})^N$, $T_{\alpha, \beta}^\ell(P) = 0$ if and only if $S_{\alpha, \beta, \gamma}^\ell(\hat{\varphi}(P)) = 0$ for all γ . We deduce from this that $\hat{\varphi}(W) = Z(I') \cap \text{Im } \hat{\varphi} = Z(I' + J)$, where $I' \subset k[z_{i_1 \dots i_N}]$ is the homogeneous ideal generated by the polynomials $S_{\alpha, \beta, \gamma}^\ell$.

Now, note the following: for a matrix A of order $d \times d$, the non leading coefficients of its characteristic polynomial are precisely $(-1)^\ell \text{Tr } \Lambda^\ell(A)$, with $1 \leq \ell \leq d$. By this observation, the set W is precisely the set of N -tuples $(A_1, \dots, A_N) \in (\mathbb{P}^{d^2-1})^N$ such that $(B_1, \dots, B_N) \in \mathcal{N}_d^N(\bar{k})$ for all B_i in the class of A_i . Clearly this set is non-empty, since it contains the N -tuples (M, \dots, M) with M being the class of a nilpotent non zero matrix. This implies that $Z(I' + J) = \hat{\varphi}(W) \neq \emptyset$. Hence, by our choice of N , Theorem 2.1.3 guarantees us that $f_j(P) = 0$ for all $P \in W$ and $g_j(Q) = 0$ for all $Q \in \hat{\varphi}(W)$. Then Nullstellensatz applies and $g_j \in I(\hat{\varphi}(W)) = \sqrt{I' + J}$.

Let $r \in \mathbb{N}$ be big enough such that $(g_j)^r \in I' + J$ for all j . There are polynomials $q_{j, \ell}^{\alpha, \beta, \gamma} \in k[z_{i_1, \dots, i_N}]$ and $h_j \in J$ such that

$$(g_j)^r = h_j + \sum_{\alpha, \beta, \ell, \gamma} q_{j, \ell}^{\alpha, \beta, \gamma} S_{\alpha, \beta, \gamma}^\ell \quad (2.8)$$

Since J is a homogeneous ideal, and comparing homogeneous degrees, we may assume that the h_j and the $q_{j, \ell}^{\alpha, \beta, \gamma}$ are homogeneous of degree $r \deg g_j = r$ and $r \deg g_j - \deg S_{\alpha, \beta, \gamma}^\ell = r - \ell$ respectively. Composing (2.8) with φ , and using (2.7)

and the fact that $h_j \circ \varphi = 0$ for all j (by the definition of J), we obtain (2.6) and our desired result with $p_{j,\ell}^{\alpha,\beta} = \sum_{\gamma} (g_{j,\ell}^{\alpha,\beta,\gamma} \circ \varphi) \cdot (u_{\gamma})^{\ell}$. \square

Remark 2.3.3. Although in the hypothesis and in the proof of Theorem 2.3.1 the constant r depends on the field k , this can be avoided by means of the effective Nullstellensatz [47]. Roughly, it says that the constant r in equation (2.8) can be chosen lower than a constant depending only on the degree of g_j , and on the number and degrees of generators of the ideal $I' + J$. The degrees of the polynomials g_j and $S_{\alpha,\beta}^{\ell}$ have a bound depending only on d and N , and the same can be said for a suitable set generators of J . A description this ideal for the case $N = 2$ is given in [25, Ex. 2.11].

2.4 Proof of Theorem 2.1.2

Theorem 2.3.1 is the fundamental relation that we will need to prove inequality (2.3).

For the next we will assume that k is a *local field*. That is, a field together with an absolute value $|\cdot| : k \rightarrow \mathbb{R}^+$ that inherits a non-discrete locally compact topology on k via the induced metric. Examples of these include \mathbb{R}, \mathbb{C} with the standard absolute values and fields of p -adic numbers \mathbb{Q}_p for a prime p . For more information about local fields, see [36].

We will work on the finite dimensional vector space k^d , where k is a local field with absolute value $|\cdot|$. In this situation, we consider the norm on $M_d(k)$ given by $\|A\|_0 = \max_{1 \leq j \leq d} |a_j|$, where a_j are the entries of A . Since the absolute value on k extends in a unique way to an absolute value on \bar{k} (see Lang's Algebra [35, XII.2, Prop. 2.5]), the spectral radius of a matrix $A \in M_d(k)$ is then defined in the usual way.

We begin with a lemma.

Lemma 2.4.1. *If $f \in k[x_{i,j}]$ is a multihomogeneous polynomial of degree $\deg f = (\lambda_1, \dots, \lambda_N)$, then there exists some $C > 0$ such that*

$$|f(A_1, \dots, A_N)| \leq C \|A_1\|_0^{\lambda_1} \cdots \|A_N\|_0^{\lambda_N}$$

for all $A_1, \dots, A_N \in M_d(k)$.

Proof. Since f is a finite sum of multihomogeneous monomials of degree $\deg f$, it is enough to prove the result when f is a monomial. In that case, $f(X_1, \dots, X_N) = c \prod_{i=1}^N \prod_{j=1}^{\lambda_i} X_{i,\ell_{i,j}}$, for some $1 \leq \ell_{i,j} \leq d^2$ and $c \in k$. So, given $A_1, \dots, A_N \in M_d(k)$,

$$|f(A_1, \dots, A_N)| = |c| \prod_{i=1}^N \prod_{j=1}^{\lambda_i} |A_{i,\ell_{i,j}}| \leq |c| \prod_{i=1}^N \|A_i\|_0^{\lambda_i}.$$

The lemma is then proved. \square

Proof of Theorem 2.1.2. Let $N = N(d)$ and $r > 1$ be given by Theorems 2.1.3 and 2.3.1 respectively. Since in a finite dimensional local field all norms are equivalent [35, XII.2, Prop. 2.2], we only have to check the result for the norm

$\|\cdot\|_0$. Let $A_1, \dots, A_N \in M_d(k)$. First, note that for $1 \leq \alpha \leq \beta \leq N$ and $1 \leq \ell \leq d$, $T_{\alpha, \beta}^\ell(A_1, \dots, A_N)$ is the ℓ -th symmetric polynomial evaluated at the eigenvalues of $A_\beta \cdots A_\alpha$. Hence we have $|T_{\alpha, \beta}^\ell(A_1, \dots, A_N)| \leq \binom{d}{\ell} \rho(A_\beta \cdots A_\alpha)^\ell$. Also, as the polynomials $p_{j, \ell}^{\alpha, \beta}$ in the statement of Theorem 2.3.1 have degree $(r, \dots, r, r-\ell, \dots, r-\ell, r, \dots, r)$, by Lemma 2.4.1 there is a constant C_0 independent of A_1, \dots, A_N such that $|p_{j, \ell}^{\alpha, \beta}(A_1, \dots, A_N)| \leq C_0 (\prod_{s=1}^N \|A_s\|_0)^r (\prod_{t=\alpha}^\beta \|A_t\|_0)^{-\ell}$ for all j, α, β, ℓ . Thus, from (2.6) we obtain the following:

$$\begin{aligned} \|A_N \cdots A_1\|_0^r &= \max_j |f_j(A_1, \dots, A_N)|^r \\ &\leq \max_j \sum_{\alpha, \beta, \ell} |p_{j, \ell}^{\alpha, \beta}(A_1, \dots, A_N)| |T_{\alpha, \beta}^\ell(A_1, \dots, A_N)| \\ &\leq C_0 \sum_{\alpha, \beta, \ell} \left(\prod_{s=1}^N \|A_s\|_0 \right)^r \left(\prod_{t=\alpha}^\beta \|A_t\|_0 \right)^{-\ell} \binom{d}{\ell} \rho(A_\beta \cdots A_\alpha)^\ell \\ &\leq C_1^r \left(\prod_{i=1}^N \|A_i\|_0 \right)^r \max_{\alpha, \beta, \ell} \left(\frac{\rho(A_\beta \cdots A_\alpha)}{\prod_{t=\alpha}^\beta \|A_t\|_0} \right)^\ell, \end{aligned}$$

for some $C_1 > 0$.

Now, let $\Lambda = \max_{\alpha, \beta} \left(\frac{\rho(A_\beta \cdots A_\alpha)}{\prod_{t=\alpha}^\beta \|A_t\|_0} \right)$. An easy computation shows that $\|AB\|_0 \leq d\|A\|_0\|B\|_0$ for all $A, B \in M_d(k)$. Moreover, comparing the norm $\|\cdot\|_0$ with some operator norm on $M_d(k)$, we can find some $D \geq 1$ such that $\rho(A) \leq D\|A\|_0$ for all $A \in M_d(k)$. This facts together imply that $\Lambda \leq Dd^N$, and hence $\Lambda^d \leq (Dd^N)^{d-1}\Lambda$. Also, depending on whether Λ is greater to 1 or not, we have $\Lambda^\ell \leq \max(\Lambda, \Lambda^d)$ for all $1 \leq \ell \leq d$. Thus we conclude

$$\|A_N \cdots A_1\|_0^r \leq C_1^r \left(\prod_{i=1}^N \|A_i\|_0 \right)^r \max(\Lambda, \Lambda^d) \leq C^r \left(\prod_{i=1}^N \|A_i\|_0 \right)^r \cdot \Lambda$$

for some $C > 0$. Applying r -th root to the last inequality, we obtain (2.3) with $\delta = 1/r$. \square

2.4.1 The case of the complex numbers

When the base field is $k = \mathbb{C}$ we can say a little more. Recall that for a norm $\|\cdot\|$ on \mathbb{C}^d , the operator norm on $M_d(\mathbb{C})$ (also denoted by $\|\cdot\|$) is defined by $\|A\| = \sup_{v \in \mathbb{C}^d \setminus \{0\}} \frac{\|Av\|}{\|v\|}$.

Proposition 2.4.2. *For $d \in \mathbb{N}$ and $N(d)$ given by Theorem 2.1.2, there are constants $C > 1$ and $0 < \delta < 1$ such that inequality (2.3) holds for all operator norms $\|\cdot\|$ on $M_d(\mathbb{C})$ and $A_1, \dots, A_N \in M_d(\mathbb{C})$.*

We will need the following lemma [5, Lem. 3.2]:

Lemma 2.4.3. *For all $d \in \mathbb{N}$ there exists a constant $C_0 > 1$ such that for every two operator norms $\|\cdot\|$ and $\|\cdot\|_1$ on $M_d(\mathbb{C})$ there exists some $S \in \text{GL}_d(\mathbb{C})$ such that for every $A \in M_d(\mathbb{C})$:*

$$C_0^{-1}\|A\| \leq \|SAS^{-1}\|_1 \leq C_0\|A\|. \quad (2.9)$$

Proof of Proposition 2.4.2: Fix an operator norm $\|\cdot\|_1$ on $M_d(\mathbb{C})$ with respective constant C given by Theorem 2.1.2. Let C_0 be given by Lemma 2.4.3, and let $\|\cdot\|$ be an arbitrary norm on $M_d(\mathbb{C})$. Let $S \in \text{GL}_d(\mathbb{C})$ be relating $\|\cdot\|$ and $\|\cdot\|_1$ as in (2.9).

Given $A_1, \dots, A_N \in M_d(\mathbb{C})$ let $B_i = SA_iS^{-1}$ for all i . We have

$$\begin{aligned} \|A_N \cdots A_1\| &\leq C_0 \|B_N \cdots B_1\|_1 \\ &\leq CC_0 \left(\prod_{1 \leq i \leq N} \|B_i\|_1 \right) \max_{1 \leq \alpha \leq \beta \leq N} \left(\frac{\rho(B_\beta \cdots B_\alpha)}{\prod_{\alpha \leq i \leq \beta} \|B_i\|_1} \right)^\delta \\ &\leq C(C_0)^{N+1} \left(\prod_{1 \leq i \leq N} \|A_i\| \right) \max_{1 \leq \alpha \leq \beta \leq N} \left(\frac{(C_0)^{\beta-\alpha+1} \rho(A_\beta \cdots A_\alpha)}{\prod_{\alpha \leq i \leq \beta} \|A_i\|} \right)^\delta \\ &\leq C(C_0)^{2N+1} \left(\prod_{1 \leq i \leq N} \|A_i\| \right) \max_{1 \leq \alpha \leq \beta \leq N} \left(\frac{(\rho(A_\beta \cdots A_\alpha))^\delta}{\prod_{\alpha \leq i \leq \beta} \|A_i\|} \right)^\delta. \end{aligned}$$

It is clear that $C(C_0)^{2N+1}$ does not depend on $\|\cdot\|$. \square

Proposition 2.4.2 allows to conclude the following inequality:

Theorem 2.4.4. *Given $d \in \mathbb{N}$ there exists a constant $C = C(d) > 1$ such that following inequality is valid for all bounded sets $\mathcal{M} \subset M_d(\mathbb{C})$:*

$$\mathfrak{R}(\mathcal{M}) \leq C \sup_{1 \leq j \leq N(d)} (\sup \{\rho(A_1 \cdots A_j) : A_i \in \mathcal{M}\})^{1/j}.$$

This inequality was first proved by Bochi in [5], and it has Theorem 2.1.1 as an immediate consequence. In [9], Breuillard gave another proof of this inequality, and used it to study semigroups of invertible matrices.

Proof of Theorem 2.4.4: For $1 \leq j \leq N(d)$, define $\rho_j = \sup \{\rho(A_1 \cdots A_j) : A_i \in \mathcal{M}\}$. For an arbitrary norm $\|\cdot\|$ on $M_d(\mathbb{C})$, take supremum for $A_i \in \mathcal{M}$ in both sides of (2.3). We obtain

$$\begin{aligned} \mathfrak{R}(\mathcal{M})^N &\leq \sup_{A_i \in \mathcal{M}} \|A_N \cdots A_1\| \\ &\leq C \max_{1 \leq j \leq N} \left(\sup_{A \in \mathcal{M}} \|A\|^{N-j\delta} \cdot (\rho_j)^\delta \right). \end{aligned} \quad (2.10)$$

Now, recall that $\mathfrak{R}(\mathcal{M}) = \inf_{\|\cdot\|} \sup_{A \in \mathcal{M}} \|A\|$, where the infimum is taken over all operator norms on $M_d(\mathbb{C})$ (for a proof, see [46]), and let $\|\cdot\|_n$ be a sequence of operator norms on $M_d(\mathbb{C})$ such that $\sup_{A \in \mathcal{M}} \|A\|_n \rightarrow \mathfrak{R}(\mathcal{M})$. Taking a subsequence, we may assume that for all $\|\cdot\|_n$, the maximum in the right hand side of (2.10) is achieved by the same index $j \in \{1, \dots, N\}$. Then, taking limit as n tends to infinity in (2.10) we will have

$$\mathfrak{R}(\mathcal{M})^N \leq C(\rho_j)^\delta \cdot \mathfrak{R}(\mathcal{M})^{N-j\delta} \quad (2.11)$$

(here is where we use Proposition 2.4.2 since C does not depend on n). If $\mathfrak{R}(\mathcal{M}) = 0$ the conclusion is obvious. Otherwise, dividing by $\mathfrak{R}(\mathcal{M})^{N-j\delta}$ and taking $j\delta$ -th root in (2.11) we obtain the desired inequality. \square

2.5 Ergodic-theoretical consequences

For the proof of Theorem 2.1.4, we will need the following result which may be seen as a quantitative version of Poincaré's recurrence Theorem for measure preserving transformations. It is a consequence of Birkhoff Ergodic Theorem, and the fact that for a measurable set U of positive measure, for almost all points x in U , the frequency of points of the sequence x, Tx, T^2x, \dots that belong to U is positive. For a detailed proof, see [4, Lem. 3.12].

Lemma 2.5.1. *Let $T : X \rightarrow X$ be a measure preserving map over the probability space (X, \mathcal{F}, μ) , and let $U \in \mathcal{F}$ have positive measure. Given $\gamma > 0$, there exists a measurable map $N_0 : U \rightarrow \mathbb{N}$ such that, for μ -a.e. $x \in U$ and $n \geq N_0(x)$ and $t \in [0, 1]$ there is some $\ell \in \{1, \dots, n\}$ with $T^\ell(x) \in U$ and $|(\ell/n) - t| < \gamma$.*

Proof of Theorem 2.1.4. Fix an operator norm $\|\cdot\|$ on $M_d(k)$, and let $Y = \{x \in X : \lambda(x) \in \mathbb{R}\}$. This is a measurable T -invariant set, and since $\rho(A) \leq \|A\|$ for all $A \in M_d(k)$, we have that both sides of (2.16) equal $-\infty$ for μ -almost all $x \in X \setminus Y$. So we only have to check the result μ -a.e. in Y .

Assume the contrary. That is, assume the existence of some $\epsilon > 0, K \in \mathbb{N}$ and a measurable set $U \subset Y$ of positive measure such that, for all $x \in U$, if $n \geq K$, then $\log \rho(A^n(x))/n + \epsilon \leq \lambda(x)$. By Egorov's theorem, and restricting to a smaller subset if necessary, we may assume that on U , $\log \|A^n(x)\|/n$ converges uniformly to $\lambda(x)$.

Let N, δ and C be as in the statement of Theorem 2.1.2 and let $\epsilon' = \epsilon/(2 + 6N\delta^{-1})$. By the uniform convergence assumption, there is some $M \geq 1$ such that, $n \geq M$ implies

$$|\log \|A^n(x)\| - n\lambda(x)| < n\epsilon' \text{ for all } x \in U. \quad (2.12)$$

Take $x \in U$ and $N_0(x) \in \mathbb{N}$ such that Lemma 2.5.1 holds with $\gamma = 1/3N$, and let $n \geq \max(3NM, 3NK, 3N \log C/\delta\epsilon', N_0(x))$. Let $m_0 = 0$, and given $1 \leq i \leq N$ let $1 \leq m_i \leq n$ be such that

$$\left| \frac{m_i}{n} - \frac{i}{N} \right| < \frac{1}{3N} \quad (2.13)$$

and $T^{m_i}x \in U$. We have that $m_i - m_{i-1} > (in/N - n/3N) - ((i-1)/N + n/3N) = n/3N \geq \max(M, K, \log C/\delta\epsilon')$ for all $1 \leq i \leq N$.

Now apply Theorem 2.1.2 to $A_i = A^{m_i - m_{i-1}}(T^{m_{i-1}}x)$. By the cocycle relation, we obtain $A_N \cdots A_1 = A^{m_N}(x)$, and hence

$$\begin{aligned} \log \|A^{m_N}(x)\| &\leq \log C + \sum_{i=1}^N \log \|A^{m_i - m_{i-1}}(T^{m_{i-1}}x)\| \\ &\quad + \delta \left(\log \rho(A^{m_\beta - m_{\alpha-1}}(T^{m_{\alpha-1}}x)) - \sum_{i=\alpha}^{\beta} \log \|A^{m_i - m_{i-1}}(T^{m_{i-1}}x)\| \right) \end{aligned} \quad (2.14)$$

for some $1 \leq \alpha \leq \beta \leq N$. But, by definition, $T^{m_i}x \in U$ for all i , and as $m_i - m_{i-1} \geq$

M , (2.12) applies. Combining it with (2.14) we have

$$\begin{aligned} \log \rho(A^{m_\beta - m_{\alpha-1}}(T^{m_{\alpha-1}}x)) &\geq \sum_{i=\alpha}^{\beta} \log \|A^{m_i - m_{i-1}}(T^{m_{i-1}}x)\| \\ &\quad + \delta^{-1} \left(\log \|A^{m_N}(x)\| - \sum_{i=1}^N \log \|A^{m_i - m_{i-1}}(T^{m_{i-1}}x)\| - \log C \right) \\ &> (m_\beta - m_{\alpha-1})(\lambda(x) - \epsilon') - \delta^{-1}(\log C + 2\epsilon' m_N) \\ &= (m_\beta - m_{\alpha-1})\lambda(x) - (\epsilon'((m_\beta - m_{\alpha-1}) + 2\delta^{-1}m_N)) + \delta^{-1} \log C. \end{aligned}$$

On the other hand, by (2.13) we have

$$\frac{m_N}{m_\beta - m_{\alpha-1}} < \frac{n}{\frac{n(\beta-\alpha+1)}{N} - \frac{2n}{3N}} \leq 3N.$$

But, since $T^{m_{\alpha-1}}x \in U$, and $(m_\beta - m_{\alpha-1}) \geq K$ we conclude

$$\begin{aligned} \frac{\epsilon'((m_\beta - m_{\alpha-1}) + 2\delta^{-1}(m_N)) + \delta^{-1} \log C}{m_\beta - m_{\alpha-1}} &= \epsilon' + \frac{\epsilon' 2(m_N)}{\delta(m_\beta - m_{\alpha-1})} + \frac{\log C}{\delta(m_\beta - m_{\alpha-1})} \\ &\leq \epsilon' + \epsilon' 6N\delta^{-1} + \frac{\log C}{\delta(m_\alpha - m_{\alpha-1})} \\ &\leq (2 + 6N\delta^{-1})\epsilon' = \epsilon. \end{aligned}$$

This is the desired contradiction and the proof is complete. \square

2.6 Ergodic characterization of the JSR

In this section we interpret the joint spectral radius in terms of Ergodic Theory. Let $\mathcal{M} \subset M_d(k)$ be a compact set of matrices, and consider the respective one-step cocycle $A : X = \mathcal{M}^{\mathbb{N}} \rightarrow M_d(k)$. The logarithm of the joint spectral radius of \mathcal{M} equals

$$\log(\mathfrak{R}(\mathcal{M})) = \lim_{n \rightarrow \infty} \sup_{A_i \in \mathcal{M}} \left(\frac{\log \|A_n \cdots A_1\|}{n} \right) = \lim_{n \rightarrow \infty} \sup_{x \in X} \left(\frac{\log \|A^n(x)\|}{n} \right).$$

The right hand side of the previous identity equals the supremum of the upper Lyapunov exponents over all ergodic shift-invariant measures on the space $\mathcal{M}^{\mathbb{N}}$ (see [39] for details). Specifically, we have

$$\lim_{n \rightarrow \infty} \sup_{x \in X} \left(\frac{\log \|A^n(x)\|}{n} \right) = \sup_{\mu \in \mathcal{E}} \inf_{n \geq 1} \left(\frac{1}{n} \int_X \log \|A^n(x)\| d\mu \right) = \sup_{\mu \in \mathcal{E}} \lambda_+(\mu), \quad (2.15)$$

where \mathcal{E} denotes the set of ergodic shift-invariant measures on $\mathcal{M}^{\mathbb{N}}$. Therefore, Berger-Wang says that instead of considering all shift-invariant ergodic measures, it is sufficient to consider those supported on periodic orbits. A far-reaching extension of this result was obtained by Kalinin [30].

We now show how Theorem 2.1.4 implies Berger-Wang. First, we present a Poincaré's recurrence-like version of 2.1.4:

Corollary 2.6.1. *With the same assumptions of Theorem 2.1.4, let $U \subset X$ be a measurable set of positive measure. Then for μ -almost all $x \in U$ there is a sequence $1 \leq n_1 < n_2 < \cdots$ such that $T^{n_j}x \in U$ for all j , and satisfying:*

$$\lambda_+(x) = \lim_{j \rightarrow \infty} \frac{\log(\rho(A^{n_j}(x)))}{n_j}.$$

Proof. Given $x \in U$, let $\mathcal{N}_x = \{n \geq 1: T^n x \in U\} = \{n_1(x) < n_2(x) < \dots\}$. By Poincaré's recurrence, \mathcal{N}_x is an infinite set for μ -almost all $x \in U$. Let $\hat{T}: U \rightarrow U$ be defined by $\hat{T}x = T^{n_1(x)}x$, which is well defined μ -a.e. on U , and consider the induced cocycle $\hat{A}: U \rightarrow M_d(k)$ given by $\hat{A}(x) = A^{n_1(x)}(x)$. It is a well known fact that \hat{T} preserves the probability measure $\hat{\mu} = \mu(\cdot)/\mu(U)$ on U , and that the cocycle (\hat{T}, \hat{A}) is $\hat{\mu}$ -integrable [6, Sec. 3.2]. Then Theorem 2.1.4 applies and we obtain

$$\lim_{k \rightarrow \infty} \frac{\log \|\hat{A}^k(x)\|}{k} = \lim_{k \rightarrow \infty} \frac{\log \|A^{(n_1 + \dots + n_k)}(x)\|}{k} = \limsup_{k \rightarrow \infty} \frac{\log(\rho(A^{(n_1 + \dots + n_k)}(x)))}{k}$$

for μ -almost all $x \in U$. The conclusion follows by noting that the limit $\lim_{k \rightarrow \infty} \frac{n_1(x) + \dots + n_k(x)}{k}$ exists and is a finite positive number for μ -almost all $x \in U$ (this is a consequence of Birkhoff Ergodic Theorem). \square

We use Corollary 2.6.1 to prove Berger-Wang Theorem over a subshift of finite type.

Proposition 2.6.2. *Let X be a subshift of finite type with shift map $T: X \rightarrow X$ and let μ be a T -invariant probability measure on X . Let $A: X \rightarrow M_d(k)$ be a one-step cocycle. Then for μ -almost all $x \in X$ there is a sequence $(p_k)_k \subset X$ of periodic points for T such that*

$$\lambda_+(x) = \lim_{k \rightarrow \infty} \lambda_+(p_k).$$

In particular we obtain the identity

$$\lim_{n \rightarrow \infty} \left(\sup_{x \in X} \|A^n(x)\| \right)^{1/n} = \sup_{n \geq 1} \left(\sup_{x \in X; T^n x = x} \rho(A^n(x)) \right)^{1/n}. \quad (2.16)$$

Proof. Suppose that the alphabet of X is $\{1, \dots, N\}$ and let $1 \leq \alpha \leq N$ be such that $\mu([\alpha]) > 0$, with $[\alpha] = \{(x_1, x_2, \dots) \in X: x_1 = \alpha\}$. By Corollary 2.6.1, for μ -almost all $x \in [\alpha]$ there is a sequence $1 \leq n_1 < n_2 < \dots$ such that $T^{n_k}x \in [\alpha]$ for all k , and such that $\lambda_+(x) = \lim_{k \rightarrow \infty} \frac{1}{n_k} \log(\rho(A^{n_k}(x)))$. If $x = (x_1, x_2, \dots)$ the condition $T^{n_k}x \in [\alpha]$ is equivalent to $x_{n_k+1} = \alpha$, which implies that the periodic sequence p_k with period (x_1, \dots, x_{n_k}) also belongs to $[\alpha]$ (and hence to X) for all k . Since α was arbitrary and $X = \cup_{1 \leq \alpha \leq N} [\alpha]$, the conclusion follows.

Moreover, if we apply this result to an ergodic measure μ that maximizes the right hand side of (2.15), we obtain (2.16) after applying the exponential function. \square

Remark 2.6.3. Applying Proposition 2.6.2 to the full shift we obtain the classical Berger-Wang identity (2.2) for all finite sets $\mathcal{F} \subset M_d(k)$.

2.7 Geometric remarks

We can observe that the main ingredients of the proof of Theorem 2.1.4 were Theorem 2.1.2 and Poincaré's recurrence Theorem. Therefore, if we had in another situation where an analogue of inequality (2.3) holds, then we should obtain a result similar to Theorem 2.1.4. This is the case of cocycles of isometries of Gromov

hyperbolic spaces. For definition and further properties of Gromov hyperbolicity see [11, 12, 13].

As it was proved in Chapter 1, if M is a Gromov hyperbolic space with distance d , then there is a constant $C > 0$ such that, for all $w \in M$ and f, g isometries of M we have

$$d(fgw, w) \leq C + \max \left(d(fw, w) + d^\infty(g), d^\infty(f) + d(gw, w), \frac{d(fw, w) + d(gw, w) + d^\infty(fg)}{2} \right),$$

where $d^\infty(h) = \lim_{n \rightarrow \infty} \frac{d(h^n w, w)}{n}$ is the stable length.

In this context, given a probability space (X, \mathcal{F}, μ) and a measure preserving map $T : X \rightarrow X$, a *cocycle of isometries* of M is a measurable map $A : X \rightarrow \text{Isom}(M)$, where $\text{Isom}(M)$ is the group of isometries of M , endowed with the Borel σ -algebra induced by the compact-open topology. We say that the cocycle A is *integrable* if the map $x \mapsto d(A(x)w, w)$ is integrable for some (and hence all) $w \in M$. In the same way as for linear cocycles, we define the family of maps $A^n : X \rightarrow \text{Isom}(M)$. For references about cocycles of isometries, see e.g. [21, 31].

Following the same steps of the proof of Theorem 2.1.4, we can obtain the following:

Proposition 2.7.1. *Let M be a Gromov hyperbolic space, $w \in M$, and let T be a measure-preserving transformation of a probability space (X, \mathcal{F}, μ) . Also, let $A : X \rightarrow \text{Isom}(M)$ be an integrable cocycle of isometries of M . Then for μ -almost all $x \in X$ and we have the following limits exist in \mathbb{R}_0^+ and are equal:*

$$\limsup_{n \rightarrow \infty} \frac{d^\infty(A^n(x))}{n} = \lim_{n \rightarrow \infty} \frac{d(A^n(x)w, w)}{n}.$$

A result similar to Proposition 2.7.1 is far from being true if we do not assume a negative curvature condition on M .

Example 2.7.2. Let $X = \mathbb{S}^1$ and μ be the Lebesgue measure on X . If $T(z) = z^2$ is the doubling map in X , which preserves μ , and $R_a(p) = p + a$ is the translation by $a \neq 0$ in \mathbb{R}^2 , define the cocycle $A : \mathbb{S}^1 \rightarrow \text{Isom}(\mathbb{R}^2)$ as $A(z)p = T(z)R_a(z^{-1}p)$ for all $p \in \mathbb{R}^2$. Note that $A^n(z)p = T^n(z)R_a^n(z^{-1}p)$ and hence the limit $\lim_{n \rightarrow \infty} \frac{d(A^n(z)p, p)}{n}$ exists and equals $|a| > 0$ for all $z \in \mathbb{S}^1$ and $p \in \mathbb{R}^2$. On the other hand, if z is not a periodic point for T , then $A^n(z)$ is not a translation and hence has a fixed point. Thus we have that $d^\infty(A^n(z)) = 0$ for all $n \in \mathbb{N}$ and all z in the set of non periodic point of T , which is a full measure set with respect to μ .

We also have the corresponding analogues of Corollary 2.6.1 and Proposition 2.6.2:

Corollary 2.7.3. *With the same assumptions of Proposition 2.7.1, let $U \subset X$ be a measurable set of positive measure. Then for μ -almost all $x \in U$ there is a sequence $1 \leq n_1 < n_2 < \dots$ such that $T^{n_j}x \in U$ for all j , and satisfying:*

$$\lim_{n \rightarrow \infty} \frac{d(A^n(x)w, w)}{n} = \lim_{j \rightarrow \infty} \frac{d^\infty(A^{n_j}(x))}{n_j}.$$

Proposition 2.7.4. *Let X be a subshift of finite type with shift map $T : X \rightarrow X$ and let μ be a T -invariant probability measure on X . Let $A : X \rightarrow \text{Isom}(M)$ be*

a one-step cocycle of isometries of M . Then for μ -almost all $x \in X$ there is a sequence $(p_k)_k \subset X$ of periodic points for T such that

$$\lim_{n \rightarrow \infty} \frac{d(A^n(x)w, w)}{n} = \lim_{k \rightarrow \infty} \left(\lim_{n \rightarrow \infty} \frac{d(A^n(p_k)w, w)}{n} \right).$$

In particular we obtain the identity

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left(\sup_{x \in X} d(A^n(x)w, w) \right) = \sup_{n \geq 1} \frac{1}{n} \left(\sup_{x \in X; T^n x = x} d^\infty(A^n(x)) \right).$$

Appendix

A Vietoris topology over topological groups

This appendix is dedicated to the topological results that we used in Section 1.5. Assume that X is a Hausdorff topological space and let $\mathcal{P}(X)$ be the set of non empty subsets of X endowed with the Vietoris topology defined in Section 1.5. Also let $\mathcal{C}(X)$ be the set of non empty compact subsets of X .

The following theorem is a criterion for convergence of nets in $\mathcal{P}(X)$ when the limit is compact. We need some notation: If A, B are directed sets, the notation $B \prec_h A$ means that $h : B \rightarrow A$ is a function satisfying the following condition: for all $\alpha \in A$ there is some $\beta \in B$ such that $\gamma \geq \beta$ implies $h(\gamma) \geq \alpha$. We say that a net $(x_{h(\beta)})_{\beta \in B}$ is a *subnet* of the net $(x_\alpha)_{\alpha \in A}$ if $B \prec_h A$. For our purposes the criterion is as follows:

Theorem A.1. *A net $(\Sigma_\alpha)_{\alpha \in A} \subset \mathcal{P}(X)$ converges to $\Sigma \in \mathcal{C}(X)$ if and only if both conditions below hold:*

- i) For every $f \in \Sigma$ and every open set U containing f there exists $\alpha \in A$ such that $\beta \geq \alpha$ implies $\Sigma_\beta \cap U \neq \emptyset$.*
- ii) Every net $(f_{h(\beta)})_{\beta \in B}$ with $B \prec_h A$ and $f_{h(\beta)} \in \Sigma_{h(\beta)}$ has a convergent subnet $(f_{h(k(\gamma))})_{\gamma \in C}$ with $C \prec_k B$ and with limit in Σ .*

This result is perhaps known, but in the lack of an exact reference we provide a proof (compare with [11, Chpt. I.5, Lem. 5.32]).

Proof. We first prove the “if” part:

Let $\langle U_1, \dots, U_n \rangle$ a basic open containing Σ . We must show that for some $\alpha \in A$, if $\beta \geq \alpha$ then $\Sigma_\beta \subset \bigcup_{1 \leq i \leq n} U_i$ and $\Sigma_\beta \cap U_i \neq \emptyset$ for all i .

Suppose that our first claim is false. Then for all $\alpha \in A$ there exists $h(\alpha) \geq \alpha$ such that $\Sigma_{h(\alpha)} \not\subset \bigcup_{1 \leq i \leq n} U_i$. That is, for all α there exists $f_{h(\alpha)} \in \Sigma_{h(\alpha)}$ such that $f_{h(\alpha)} \notin U_i$ for all i . Hence $(f_{h(\alpha)})_{\alpha \in A}$ is a net with $A \prec_h A$ and since we are assuming *ii*), it has a convergent subnet $(f_{h(k(\gamma))})_{\gamma \in C}$ with limit $f \in \Sigma$ and $C \prec_k A$. But $f \in U_j$ for some j and there is $\gamma \in C$ with $f_{h(k(\gamma))} \in U_j$, contradicting the definition of $h(k(\gamma))$. So there exists α_0 such that $\beta \geq \alpha_0$ implies $\Sigma_\beta \subset \bigcup_{1 \leq i \leq n} U_i$.

Now, fix $j \in \{1, \dots, n\}$ and suppose that for all α , $\Sigma_{\beta(\alpha)} \cap U_j = \emptyset$ for some $\beta(\alpha) \geq \alpha$. Noting that $\langle U_j, X \rangle$ also contains Σ , there exists $f \in \Sigma \cap U_j$, and by *i*)

there is some α such that $\Sigma_\beta \cap U_j \neq \emptyset$ for $\beta \geq \alpha$, contradicting the existence of $\Sigma_{\beta(\alpha)}$. So for all j , there is some α_j such that $\Sigma_\beta \cap U_j \neq \emptyset$ for $\beta \geq \alpha_j$ and hence any $\alpha \geq \alpha_j$ for $0 \leq j \leq n$ satisfies our requirements.

For the converse, suppose that Σ_α tends to Σ . Let $f \in \Sigma$ and U be an open neighborhood of f . The set $\langle U, X \rangle$ is open and contains Σ . So there exists some α such that for all $\beta \geq \alpha$, $\Sigma_\beta \in \langle U, X \rangle$ and hence $\Sigma_\beta \cap U \neq \emptyset$.

Finally, let $(f_{h(\beta)})_{\beta \in B}$ be a net with $B \prec_h A$ and $f_{h(\beta)} \in \Sigma_{h(\beta)}$. We claim that this net has a subnet converging to an element of Σ . For $\beta \in B$ consider the set $E(\beta) = \{f_{h(\gamma)} : \gamma \in B \text{ and } \gamma \geq \beta\} \subset X$ and let $F(\beta) = \overline{E(\beta)}$.

If $\bigcap_{\beta \in B} F(\beta) \cap \Sigma = \emptyset$, the collection $\{X \setminus F(\beta)\}_{\beta \in B}$ is an open cover of Σ and by compactness it has a minimal finite subcover $\{X \setminus F(\beta_i)\}_{1 \leq i \leq m}$. This implies that $\Sigma \in \langle X \setminus F(\beta_i) \rangle_{1 \leq i \leq m}$ and by our convergence assumption, for some $\alpha_0 \in A$ it happens that $\Sigma_\alpha \subset \bigcup_{1 \leq i \leq m} X \setminus F(\beta_i)$ when $\alpha \geq \alpha_0$. But $B \prec_h A$, so if we take $\beta_0 \in B$ such that $h(\beta_0) \geq \alpha_0$ and β' greater than β_i for all $0 \leq i \leq m$, then $f_{h(\beta')} \notin F(\beta_i)$ for some i . This contradicts that $f_{h(\beta')} \in E(\beta_i) \subset F(\beta_i)$. So there exists some $f \in \bigcap_{\beta \in B} F(\beta) \cap \Sigma$.

Then for every open neighborhood U of f and every $\beta \in B$, there exists some $k(U, \beta) \geq \beta$ such that $f_{h(k(U, \beta))} \in U \cap E(\beta)$. Let \mathcal{N} be the set of open neighborhoods of f partially ordered by reverse inclusion. In this way $\mathcal{N} \times B$ with the product order becomes a directed set. Now consider the map $k : \mathcal{N} \times B \rightarrow B$ and let $\beta \in B$. For some $U_0 \in \mathcal{N}$, every $(V, \gamma) \in \mathcal{N} \times B$ with $(V, \gamma) \geq (U_0, \beta)$ satisfies $k(V, \gamma) \geq \gamma \geq \beta$. So $\mathcal{N} \times B \prec_k B$ and $(f_{h \circ k(\lambda)})_{\lambda \in \mathcal{N} \times B}$ is a subnet of $(f_{h(\beta)})_{\beta \in B}$. To finish the proof we must verify that f is the limit to this subnet. So, let $U \in \mathcal{N}$. For $(U, \beta) \in \mathcal{N} \times B$ we have that $(V, \gamma) \geq (U, \beta)$ implies $f_{h \circ k(V, \gamma)} \in V \cap E(\gamma) \subset U$. So f is our desired limit and our claim is proved. \square

As application to Theorem A.1 let G be a Hausdorff topological group with identity e . If $o : G \times G \rightarrow G$ is the composition map and $\Sigma, \Pi \in \mathcal{C}(G)$ then $\Sigma\Pi = o(\Sigma \times \Pi) \in \mathcal{C}(G)$, so it induces a composition map $\pi : \mathcal{C}(G) \times \mathcal{C}(G) \rightarrow \mathcal{C}(G)$. We establish that this map is continuous.

Theorem A.2. *The composition map $\pi : \mathcal{C}(G) \times \mathcal{C}(G) \rightarrow \mathcal{C}(G)$ given by $\pi(\Sigma, \Pi) = \Sigma\Pi$ is continuous.*

Proof. Let $(\Sigma_\alpha, \Pi_\alpha)_{\alpha \in A}$ be a net that converges to (Σ, Π) . We claim that $(\Sigma_\alpha \Pi_\alpha)_{\alpha \in A}$ tends to $\Sigma\Pi$. For that, we use the equivalence given by Theorem A.1. Let $f \in \Sigma$, $g \in \Pi$ and U be an open neighborhood of fg . So $f^{-1}Ug^{-1}$ is an open neighborhood of e and hence there exists V open with $e \in V \subset V^2 \subset f^{-1}Ug^{-1}$. Then we have $f \in fV$ and $g \in Vg$.

So there exists α_1 and α_2 such that $\beta \geq \alpha_1$ implies $\Sigma_\beta \cap fV \neq \emptyset$ and $\beta \geq \alpha_2$ implies $\Sigma_\beta \cap Vg \neq \emptyset$. If we take α_0 greater than α_1 and α_2 , for $\beta \geq \alpha_0$ there exists $f_\beta \in \Sigma_\beta$ and $g_\beta \in \Sigma_\beta$ such that $f_\beta \in fV$ and $g_\beta \in Vg$. We conclude that for all $\beta \geq \alpha_0$, $f_\beta g_\beta \in fV^2g \subset U$, hence $\Sigma_\beta \Pi_\beta \cap U \neq \emptyset$ for all $\beta \geq \alpha_0$.

Now, let $B \prec_h A$ be such that $(f_{h(\beta)}g_{h(\beta)})_{\beta \in B}$ is a net with $f_{h(\beta)} \in \Sigma_{h(\beta)}$ and $g_{h(\beta)} \in \Pi_{h(\beta)}$. We must exhibit a subnet converging to an element in $\Sigma\Pi$. But it is easy. Since $\Sigma_{h(\beta)} \rightarrow \Sigma$, there exists $C \prec_k B$ such that $(f_{h \circ k(\gamma)})_{\gamma \in C}$ is a net that tends to $f \in \Sigma$. Also, as $C \prec_{h \circ k} A$ there exists $D \prec_l C$ with $(g_{h \circ k \circ l(\lambda)})_{\lambda \in D}$ a

net that converges to $g \in \Pi$. Then $(f_{h \circ kol(\lambda)} g_{h \circ kol(\lambda)})_{\lambda \in D}$ tends to $fg \in \Sigma\Pi$. Our proof is complete. \square

Corollary A.3. *The map $\Sigma \mapsto \Sigma^n$ is continuous in $\mathcal{C}(G)$ for all $n \in \mathbb{Z}^+$.*

Bibliography

- [1] A. Avila and J. Bochi, A formula with some applications to the theory of Lyapunov exponents. *Isr. J. Math.* **131** (2002), 125-137. [Zbl 1022.37019](#) [MR 1942304](#)
- [2] M. A. Berger and Y. Wang, Bounded semigroups of matrices. *Linear Algebra Appl.* **166** (1992), 21-27. [Zbl 0818.15006](#) [MR 1152485](#)
- [3] V. D. Blondel, J. Theys and A. A. Vladimirov, An elementary counterexample to the finiteness conjecture. *SIAM J. Matrix Anal. Appl.* **24** (2003), no. 4, 963-970. [Zbl 1043.15007](#) [MR 2003315](#)
- [4] J. Bochi, Genericity of zero Lyapunov exponents. *Ergodic Theory Dynam. Systems* **22** (2002), no. 6, 1667-1696. [Zbl 1023.37006](#) [MR 1944399](#)
- [5] J. Bochi, Inequalities for numerical invariants of sets of matrices. *Linear Algebra Appl.* **368** (2003), 71-81. [Zbl 1031.15023](#) [MR 1983195](#)
- [6] J. Bochi, The basic ergodic theorems, yet again. Preprint.
- [7] M. Bonk and O. Schramm, Embeddings of Gromov hyperbolic spaces. *Geom. Funct. Anal.* **10** (2000), no. 2, 266-306. [Zbl 0972.53021](#) [MR 1771428](#)
- [8] T. Bousch and J. Mairesse, Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture. *J. Amer. Math. Soc.* **15** (2002), no. 1, 77-111. [Zbl 1057.49007](#) [MR 1862798](#)
- [9] E. Breuillard, A height gap theorem for finite subsets of $GL_d(\overline{\mathbb{Q}})$ and nonamenable subgroups. *Ann. of Math.*(2) **174** (2011), no. 2, 1057-1110. [Zbl 1243.11071](#) [MR 2831113](#)
- [10] E. Breuillard and K. Fujiwara, On the joint spectral radius for isometries of non-positively curved spaces and uniform growth. Preprint 2018. [arXiv:1804.00748](#) [math.GR]
- [11] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*. Grundlehren Math. Wiss. 319, Springer-Verlag, Berlin, 1999. [Zbl 0988.53001](#) [MR 1744486](#)
- [12] M. Coornaert, T. Delzant and A. Papadopoulos, *Géométrie et théorie des groupes. Les groupes hyperboliques de Gromov*. Lecture Notes in Math. 1441, Springer-Verlag, Berlin, 1990. [Zbl 0727.20018](#) [MR 1075994](#)
- [13] T. Das, D. Simmons and M. Urbański, *Geometry and dynamics in Gromov hyperbolic metric spaces. With an emphasis on non-proper settings*. Math. Surveys Monogr. 218, Amer. Math. Soc., Providence, RI, 2017. [Zbl 06729361](#) [MR 3558533](#)
- [14] I. Daubechies and J. C. Lagarias, Sets of matrices all infinite products of which converge. *Linear Algebra Appl.* **161** (1992), 227-263. [Zbl 0746.15015](#) [MR 1142737](#)
- [15] R. Engelking, *General topology*. Second edition. Sigma Series in Pure Math. 6, Heldermann Verlag, Berlin, 1989. [Zbl 0684.54001](#) [MR 1039321](#)
- [16] W. Fulton, *Algebraic curves: an Introduction to algebraic geometry*. 2008. <http://www.math.lsa.umich.edu/~wfulton/CurveBook.pdf>

- [17] S. M. Gersten and H. B. Short, Rational subgroups of biautomatic groups. *Ann. of Math.* (2) **134** (1991), no. 1, 125-158. [Zbl 0744.20035](#) [MR 1114609](#)
- [18] M. Gromov, Hyperbolic manifolds, groups and actions. In *Riemann surfaces and related topics: Proc. 1978 Stony Brook Conf.* Ann. of Math. Stud. 97, Princeton Univ. Press, Princeton, N.J., 1981, 183–213. [Zbl 0467.53035](#) [MR 0624814](#)
- [19] M. Gromov, Hyperbolic groups. In *Essays in group theory*, Math. Sci. Res. Inst. Publ. 8, Springer, New York, 1987, 75-263. [Zbl 0634.20015](#) [MR 0919829](#)
- [20] I. Goldhirsch, P.-L. Sulem, and S. A. Orszag, Stability and Lyapunov stability of dynamical systems: a differential approach and a numerical method. *Physica. D.* **27** (1987), 311-337. [Zbl 0656.34045](#)
- [21] S. Gouëzel and A. Karlsson, Subadditive and Multiplicative Ergodic Theorems. Preprint 2015. [arXiv:1509.07733](#) [math.DS]
- [22] L. Gurvits, Stability of discrete linear inclusion. *Linear Algebra Appl.* **231** (1995), 47-85. [Zbl 0845.68067](#) [MR 1361100](#)
- [23] M. Hamann, Group actions on metric spaces: fixed points and free subgroups. *Abh. Math. Semin. Univ. Hambg.* **87** (2017), no. 2, 245–263. [Zbl 1377.05199](#) [MR 3696149](#)
- [24] P. de la Harpe, *Topics in geometric group theory*. Chicago Lectures in Math. Univ. Chicago Press, Chicago, IL, 2000. [Zbl 0965.20025](#) [MR 1786869](#)
- [25] J. Harris, *Algebraic Geometry: a First Course*. Graduate Texts in Math. 133, Springer-Verlag, New York, 1992. [Zbl 0779.14001](#) [MR 1416564](#)
- [26] J. Hubbard, *Teichmüller theory and applications to geometry, topology, and dynamics*. Vol. 1, Matrix Editions, Ithaca, NY, 2006. [Zbl 1102.30001](#) [MR 2245223](#)
- [27] I. M. Isaacs, *Algebra: a graduate course*. Brooks/Cole, Pacific Grove, 1994. [Zbl 0805.00001](#) [MR 1276273](#)
- [28] O. Jenkinson and M. Pollicott, Joint spectral radius, Sturmian measures, and the finiteness conjecture. *Ergodic Theory Dynam. Systems*, to appear.
- [29] R. Jungers, *The joint spectral radius. Theory and applications*. Lect. Notes Control Inf. Sci. 385, Springer-Verlag, Berlin, 2009. [MR 2507938](#)
- [30] B. Kalinin, Livšic theorem for matrix cocycles. *Ann. of Math.* (2) **173** (2011), no. 2, 1025–1042. [Zbl 1238.37008](#) [MR 2776369](#)
- [31] A. Karlsson and G. Margulis, A multiplicative ergodic theorem and nonpositively curved spaces. *Comm. Math. Phys.* **208** (1999), 107-123. [Zbl 0979.37006](#) [MR 1729880](#)
- [32] V. Kozyakin, An annotated bibliography on convergence of matrix products and the theory of joint/generalized spectral radius. Preprint.
- [33] V. Kozyakin, The Berger-Wang formula for the Markovian joint spectral radius. *Linear Algebra Appl.* **448** (2014), 315-328. [Zbl 1321.15021](#) [MR 3182989](#)
- [34] J. C. Lagarias and Y. Wang, The finiteness conjecture for the generalized spectral radius of a set of matrices. *Linear Algebra Appl.* **214** (1995), 17–42. [Zbl 0818.15007](#) [MR 1311628](#)
- [35] S. Lang, *Algebra*. 3rd revised ed. Graduate Texts in Math. 211, Springer-Verlag, New York, 2002. [Zbl 0984.00001](#) [MR 1878556](#)
- [36] F. Lorenz, *Algebra. Volume II: Fields with structure, algebras, and advanced topics*. Universitext, Springer, New York, 2008. [Zbl 1130.12001](#)
- [37] E. Michael, Topologies on spaces of subsets. *Trans. Amer. Math. Soc.* **71** (1951), 152–182. [Zbl 0043.37902](#) [MR 0042109](#)
- [38] I. D. Morris, An inequality for the matrix pressure function and applications. *Adv. Math.* **302** (2016), 280-308. [Zbl 1350.15005](#) [MR 3545931](#)
- [39] I. D. Morris, Mather sets for sequences of matrices and applications to the study of joint spectral radii. *Proc. Lond. Math. Soc.* (3) **107** (2013), no. 1, 121-150. [Zbl 1277.15009](#) [MR 3083190](#)

- [40] I. D. Morris, The generalised Berger-Wang formula and the spectral radius of linear cocycles. *J. Funct. Anal.* **262** (2012), no. 3, 811–824. [Zbl 1254.47006](#) [MR 2863849](#)
- [41] B. Nica and J. Špakula, Strong Hyperbolicity. *Groups Geom. Dyn.* **10** (2016), no. 3, 951–964. [Zbl 1368.20057](#) [MR 3551185](#)
- [42] A. Y. Ol'shanskii, Almost every group is hyperbolic. *Internat. J. Algebra Comput.* **2** (1992), no. 1, 1-17. [Zbl 0779.20016](#) [MR 1167524](#)
- [43] E. Oregón-Reyes, A new inequality about matrix products and a Berger-Wang formula. Preprint 2017. [arXiv:1710.00639](#) [math.DS]
- [44] E. Oregón-Reyes, Properties of sets of isometries of Gromov hyperbolic spaces. *Groups Geom. Dyn.*, to appear.
- [45] H. Radjavi and P. Rosenthal, *Simultaneous Triangularization*. Springer-Verlag, New York, 2000. [Zbl 0981.15007](#) [MR 1736065](#)
- [46] G. C. Rota and G. Strang, A note on the joint spectral radius. *Indag. Math. (N.S.)* **22** (1960), 379-381. [Zbl 0095.09701](#) [MR 0147922](#)
- [47] M. Sombra, A sparse effective Nullstellensatz. *Adv. Appl. Math.* **22** (1999), no. 2, 271-295. [Zbl 0933.14001](#) [MR 1659402](#)
- [48] J. N. Tsitsiklis and V. D. Blondel, The Lyapunov exponent and joint spectral radius of pairs of matrices are hard when not impossible to compute and to approximate. *Math. Control Signals Syst.* **10** (1997), no. 1, 31-40. [Zbl 0888.65044](#) [MR 1462278](#)
- [49] J. Väisälä, Gromov hyperbolic spaces. *Expo. Math.* **23** (2005), no. 3, 187-231. [Zbl 1087.53039](#) [MR 2164775](#)