

# ESCAPE OF MASS AND ENTROPY FOR DIAGONAL FLOWS IN REAL RANK ONE SITUATIONS

BY

M. EINSIEDLER\*

*Departement Mathematik, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland*  
*e-mail: manfred.einsiedler@math.ethz.ch*

AND

S. KADYROV\*\*

*School of Mathematics, University of Bristol, Bristol, UK*  
*e-mail: shirali.kadyrov@bristol.ac.uk*

AND

A. POHL†.

*Mathematisches Institut, Georg-August-Universität Göttingen*  
*Bunsenstr. 3-5, 37073 Göttingen, Germany*  
*e-mail: pohl@uni-math.gwdg.de*

ABSTRACT

Let  $G$  be a connected semisimple Lie group of real rank 1 with finite center, let  $\Gamma$  be a non-uniform lattice in  $G$  and  $a$  any diagonalizable element in  $G$ . We investigate the relation between the metric entropy of  $a$  acting on the homogeneous space  $\Gamma \backslash G$  and escape of mass. Moreover, we provide bounds on the escaping mass and, as an application, we show that the Hausdorff dimension of the set of orbits (under iteration of  $a$ ) which miss a fixed open set is not full.

---

\* M. E. acknowledges the support by the SNF (Grant 200021-127145).

\*\* S. K. acknowledges the support by the EPSRC.

† A. P. acknowledges the support by the SNF (Grant 200021-127145) and the Volkswagen Foundation

Received September 25, 2011 and in revised form September 7, 2014

CONTENTS

1. Introduction . . . . . 246  
 2. Fundamental domains in the cusps . . . . . 248  
 3. The height function . . . . . 250  
 4. Coordinate system for  $G$  . . . . . 252  
 5. Variation of height . . . . . 254  
 6. Common cusp excursions of nearby points . . . . . 258  
 7. Estimate of metric entropy and proof of Theorem A . . . . . 264  
 8. Hausdorff dimension of orbits missing a fixed open subset . . . . . 278  
 9. Modification of the partition from [EL10] . . . . . 292  
 References . . . . . 294

**1. Introduction**

Let  $G$  be a connected semisimple (real) Lie group of  $\mathbb{R}$ -rank 1 with finite center and  $\Gamma$  a lattice in  $G$ . Suppose that

$$\mathcal{X} := \Gamma \backslash G$$

denotes the associated homogeneous space. Let  $A$  be a one-parameter subgroup consisting of  $\mathbb{R}$ -diagonalizable elements. Pick an element  $\tilde{a} \in A \setminus \{\text{id}\}$  and consider the right action

$$T: \begin{cases} \mathcal{X} \rightarrow \mathcal{X} \\ x \mapsto x\tilde{a} \end{cases}$$

of  $\tilde{a}$  on  $\mathcal{X}$ . Further let  $(\mu_n)_{n \in \mathbb{N}}$  be a sequence of  $T$ -invariant probability measures on  $\mathcal{X}$  which converges in the weak\* topology to the measure  $\nu$ .

If  $\nu$  is itself a probability measure (which is always the case if  $\Gamma$  is cocompact), then upper semi-continuity of metric entropy is well-known, that is

$$\limsup_{n \rightarrow \infty} h_{\mu_n}(T) \leq h_{\nu}(T).$$

In this article we investigate the case that  $\Gamma$  is non-cocompact and  $\nu$  is not a probability measure. We show that if upper semi-continuity does not hold, the amount by which it fails is controlled by the escaping mass. More precisely, the main result can be stated as follows.

**THEOREM A:** *Let  $h_m(T)$  denote the maximal metric entropy of  $T$  and suppose that  $\nu(\mathcal{X}) > 0$ . Then*

$$\nu(\mathcal{X})h_{\frac{\nu}{\nu(\mathcal{X})}}(T) + \frac{1}{2}h_m(T) \cdot (1 - \nu(\mathcal{X})) \geq \limsup_{n \rightarrow \infty} h_{\mu_n}(T).$$

In [KP] it is shown that the factor  $\frac{1}{2}$  is sharp. A consequence of this theorem is the following result about escape of mass, which is of interest on its own.

**COROLLARY:** *Suppose that  $\limsup h_{\mu_n}(T) \geq c$ . Then*

$$\nu(\mathcal{X}) \geq \frac{2c}{h_m(T)} - 1.$$

Thus, if the entropy on the sequence  $(\mu_n)$  is high, meaning at least  $\frac{1}{2}h_m(T) + \varepsilon$ , then not all of the mass can escape and the remaining mass can be bounded quantitatively.

For  $\mathcal{X} = \text{SL}_2(\mathbb{Z}) \backslash \text{SL}_2(\mathbb{R})$  and  $T$  being the time-one map, this control on escape of mass is already shown in [ELMV12]. For recent results of this kind in different settings and their applications we refer to [EK12, Kad12, KKLm].

In case of equality in the corollary above, Theorem A yields the following consequence for the remaining normalized measure.

**COROLLARY:** *If  $\limsup h_{\mu_n}(T) \geq c$  and*

$$\nu(\mathcal{X}) = \frac{2c}{h_m(T)} - 1 > 0,$$

*then  $h_{\frac{\nu}{\nu(\mathcal{X})}}(T) = h_m(T)$  and  $\frac{\nu}{\nu(\mathcal{X})}$  is the Haar measure on  $\mathcal{X}$ .*

As an application of these results and the methods for their proofs we show in Section 8 the following observation, thereby answering a question of Barak Weiss. Its positive solution is already used in [HW13].

**THEOREM B:** *Let  $\mathcal{O}$  be an open nonempty subset of  $\mathcal{X}$ , and let  $\mathcal{E}$  be the set of points in  $\mathcal{X}$  whose forward trajectories (forward  $A$ -orbits) do not intersect  $\mathcal{O}$ . Then the Hausdorff dimension of  $\mathcal{E}$  is strictly smaller than the (Hausdorff) dimension of  $\mathcal{X}$ .*

We outline the strategy of proof for Theorem A. The key tool for its proof is the existence of a finite partition  $\eta$  of  $\mathcal{X}$  such that for each  $T$ -invariant probability measure  $\mu$  on  $\mathcal{X}$  the entropy of  $\mu$ , the entropy of the partition  $\eta$  and the mass “high” in the cusps of  $\mathcal{X}$  are seen to be related as in Theorem A. More

precisely, if  $\mathcal{X}_{>s}$  denotes the part of  $\mathcal{X}$  above height  $s$  (the notion of height is defined in Section 3 below), then

$$h_\mu(T) \leq h_\mu(T, \eta) + c_s + \frac{1}{2}h_m(T)\mu(\mathcal{X}_{>s})$$

with a global constant  $c_s$  such that  $c_s \rightarrow 0$  as  $s \rightarrow \infty$ . We remark that  $\eta$  is independent of  $\mu$ . To achieve this we use a partition of  $\mathcal{X}$  into a fixed compact part, the part  $\mathcal{X}_{>s}$  above height  $s$ , and the strip between the compact part and  $\mathcal{X}_{>s}$ . The compact part is refined into very small sets, depending on the width of the strip, such that this part and the strip do not contribute to entropy.

The entropy of  $\mu$  is estimated from above using the Brin–Katok Lemma, which reduces this task to counting Bowen balls needed to cover some set of fixed positive measure. In Lemma 7.4 below we provide a non-trivial bound for this number. In order to be able to establish this result, we translate the situation to Siegel sets in  $G$  (which is possible thanks to a result of Garland and Raghunathan [GR70] on fundamental domains), and conduct a detailed study how nearby trajectories behave high up in the cusp.

These investigations do not use the classification of  $\mathbb{R}$ -rank 1 simple Lie groups. Rather we take advantage of the uniform and easy to manipulate construction of rank 1 symmetric spaces of noncompact type provided by [CDKR91] and [CDKR98] and the coordinate system of the associated Lie groups adapted to their geometry.

ACKNOWLEDGMENT. We thank the anonymous referees for many valuable comments that helped to improve the presentation of the paper.

## 2. Fundamental domains in the cusps

Let  $A$  be a one-parameter  $\mathbb{R}$ -diagonalizable subgroup in  $G$  containing the diagonalizable element  $\tilde{a}$  defining the transformation  $T$  on  $\mathcal{X}$  via  $x \mapsto x\tilde{a}$ . Let  $C = C_A(G)$  denote the centralizer of  $A$  in  $G$  and let  $\mathfrak{c}$  be its Lie algebra. Let  $\mathfrak{g}$  denote the Lie algebra of  $G$ . Since  $G$  is of  $\mathbb{R}$ -rank 1, there exists a group homomorphism  $\alpha: A \rightarrow (\mathbb{R}_{>0}, \cdot)$  such that with

$$\mathfrak{g}_j := \{X \in \mathfrak{g} \mid \forall a \in A: \text{Ad}_a X = \alpha(a)^{\frac{j}{2}} X\}, \quad j \in \{\pm 1, \pm 2\},$$

we have the direct sum decomposition

$$(1) \quad \mathfrak{g} = \mathfrak{g}_{-2} \oplus \mathfrak{g}_{-1} \oplus \mathfrak{c} \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2.$$

We choose the homomorphism  $\alpha$  such that  $\alpha(\tilde{a}) > 1$ . The Lie algebra  $\mathfrak{g}$  is the direct product of a simple Lie algebra and a compact one. Unless this simple Lie algebra is isomorphic to  $\mathfrak{so}(1, n)$ , the homomorphism  $\alpha$  is then unique and (1) is the restricted root space decomposition of  $\mathfrak{g}$ . If the simple factor of  $\mathfrak{g}$  is isomorphic to  $\mathfrak{so}(1, n)$  for some  $n \in \mathbb{N}$ ,  $n \geq 2$ , then there are two choices for  $\alpha$ . Depending on the choice, either  $\mathfrak{g}_2$  or  $\mathfrak{g}_1$  is trivial. In this case (1) simplifies to

$$\mathfrak{g} = \mathfrak{g}_{-1} \oplus \mathfrak{c} \oplus \mathfrak{g}_1 \quad \text{resp.} \quad \mathfrak{g} = \mathfrak{g}_{-2} \oplus \mathfrak{c} \oplus \mathfrak{g}_2,$$

each of which is the restricted root space decomposition of  $\mathfrak{g}$ . The first one corresponds to the Cayley–Klein models of real hyperbolic spaces, the second one to the Poincaré models. Define  $\mathfrak{n} := \mathfrak{g}_2 \oplus \mathfrak{g}_1$  and let  $N$  be the connected, simply connected Lie subgroup of  $G$  with Lie algebra  $\mathfrak{n}$ . By the theorem concerning Iwasawa decompositions of  $G$ , there exists a maximal compact subgroup  $K$  of  $G$  such that

$$N \times A \times K \rightarrow G, \quad (n, a, k) \mapsto nak$$

is a diffeomorphism. Let

$$M := K \cap C.$$

For any  $s > 0$  we set

$$A_s := \{a \in A \mid \alpha(a) > s\}.$$

Moreover, for any  $s > 0$  and any compact subset  $\eta$  of  $N$  we define the Siegel set

$$\Omega(s, \eta) := \eta A_s K.$$

Garland and Raghunathan provide the following result on fundamental domains for the non-cocompact lattice  $\Gamma$  in  $G$ .

**PROPOSITION 2.1** (Theorem 0.6 and 0.7 in [GR70]): *There exists  $s_0 > 0$ , a compact subset  $\eta_0$  of  $N$  and a finite subset  $\Xi$  of  $G$  such that*

- (i)  $G = \Gamma \Xi \Omega(s_0, \eta_0)$ ,
- (ii) for all  $\xi \in \Xi$ , the group  $\Gamma \cap \xi N \xi^{-1}$  is a cocompact lattice in  $\xi N \xi^{-1}$ ,
- (iii) for all compact subsets  $\eta$  of  $N$  the set

$$\{\gamma \in \Gamma \mid \gamma \Xi \Omega(s_0, \eta) \cap \Omega(s_0, \eta) \neq \emptyset\}$$

is finite,

- (iv) for each compact subset  $\eta$  of  $N$  containing  $\eta_0$ , there exists  $s_1 > s_0$  such that for all  $\xi_1, \xi_2 \in \Xi$  and all  $\gamma \in \Gamma$  with  $\gamma \xi_1 \Omega(s_0, \eta) \cap \xi_2 \Omega(s_1, \eta) \neq \emptyset$  we have  $\xi_1 = \xi_2$  and  $\gamma \in \xi_1 N M \xi_1^{-1}$ .

For the remainder of this article we fix  $s_1 > s_0 > 0$ , a compact subset  $\eta_0$  of  $N$  and a finite subset  $\Xi$  of  $G$  which satisfy (i)–(iv) of Proposition 2.1 with  $\eta := \eta_0$ .

The elements of  $\Xi$  are a minimal set of representatives for the cusps of

$$\mathcal{X} := \Gamma \backslash G,$$

and for each  $\xi \in \Xi$ , the Siegel set  $\xi\Omega(s_1, \eta)$  modulo  $\Gamma \cap \xi NM\xi^{-1}$  is a neighborhood of the corresponding cusp of  $\mathcal{X}$ . In the following we will often identify this cusp with its neighborhood  $(\Gamma \cap \xi NM\xi^{-1}) \backslash \xi\Omega(s_1, \eta) \subseteq \mathcal{X}$ , and also refer to the latter one as the cusp represented by  $\xi$ .

### 3. The height function

For each  $\xi \in \Xi$ , we introduce a height function which measures how far a point  $x \in \mathcal{X}$  is “in the cusp represented by  $\xi$ ”. More precisely, the  $\xi$ -height of  $x$  is the maximal value  $\alpha(a)$  for an  $x$ -representative  $\xi nak$  in  $G = \xi NAK$ . The maximum over all  $\xi$ -heights gives the total height of  $x \in \mathcal{X}$ . For a coordinate-free definition of the height functions, we introduce a representation derived from the adjoint representation. This representation was also used in [Dan84].

For each  $\xi \in \Xi$  we set

$$L_\xi := \xi NM\xi^{-1}$$

and denote its Lie algebra by  $\mathfrak{l}_\xi$ . Set  $\ell := \dim \mathfrak{l}_\xi$  (which in fact is independent of  $\xi$ ) and let  $V$  be the  $\ell$ -th exterior power of  $\mathfrak{g}$ ,

$$V := \bigwedge^\ell \mathfrak{g}.$$

Let  $\varrho$  be the right<sup>1</sup>  $G$ -action on  $V$  given by the  $\ell$ -th exterior power of

$$\text{Ad} \circ (\cdot)^{-1}: G \rightarrow \text{End}(\mathfrak{g}), \quad g \mapsto \text{Ad}_{g^{-1}},$$

hence

$$\varrho := \bigwedge^\ell (\text{Ad} \circ (\cdot)^{-1}): G \rightarrow \text{End}(V).$$

We fix a non-zero element  $v_\xi$  in the one-dimensional space

$$W_\xi := \bigwedge^\ell \mathfrak{l}_\xi$$

and let

$$\theta_\xi: \xi NMA\xi^{-1} \rightarrow \mathbb{R}_{>0}$$

---

<sup>1</sup> When applying  $\varrho(g)$  for  $g \in G$  to  $v \in V$ , we will write  $v\varrho(g)$  instead of  $\varrho(g)v$  to stress that it is a right action.

be the unique group homomorphism into the multiplicative group  $(\mathbb{R}_{>0}, \cdot)$  such that for all  $g \in \xi NMA\xi^{-1}$  we have

$$v_\xi \varrho(g) = \theta_\xi(g)v_\xi.$$

One easily shows that  $\theta_\xi(g) = 1$  for  $g$  in the connected component of  $L_\xi$ , and

$$\theta_\xi(\xi a \xi^{-1}) = \alpha(a)^{-\left(\frac{1}{2} \dim \mathfrak{g}_1 + \dim \mathfrak{g}_2\right)}$$

for  $a \in A$ . Let

$$q := \frac{1}{2} \dim \mathfrak{g}_1 + \dim \mathfrak{g}_2.$$

We choose a  $\varrho(K)$ -invariant inner product  $\langle \cdot, \cdot \rangle$  on  $V$  (e.g., induced by the Killing form) and denote its associated norm by  $\| \cdot \|$ .

For  $\xi \in \Xi$ , the  $\xi$ -**height** of  $x \in \mathcal{X}$  is defined as

$$(2) \quad \text{ht}_\xi(x) := \sup \left\{ \left( \frac{\|v_\xi \varrho(g)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}} \mid g \in G, x = \Gamma g \right\}.$$

If  $g \in G$  is represented as  $g = \xi n a k$  with  $n \in N$ ,  $a \in A$  and  $k \in K$ , then by definition

$$\left( \frac{\|v_\xi \varrho(g)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}} = \alpha(a).$$

Hence this value only depends on the  $A$ -components of  $g$  when represented in  $\xi NAK (= G)$ , of which we may think as an Iwasawa decomposition of  $G$  relative to  $\xi$ .

The **height** of  $x \in \mathcal{X}$  is

$$\text{ht}(x) := \max\{\text{ht}_\xi(x) \mid \xi \in \Xi\}.$$

For  $s > 0$  and  $\xi \in \Xi$  we set

$$\mathcal{X}(\xi, s) := \{x \in \mathcal{X} \mid \text{ht}_\xi(x) > s\}$$

and

$$(3) \quad \mathcal{X}_{>s} := \{x \in \mathcal{X} \mid \text{ht}(x) > s\} = \bigcup_{\xi \in \Xi} \mathcal{X}(\xi, s).$$

In the following we will see that the points in  $\mathcal{X}(\xi, s)$  correspond to the elements in the Siegel set  $\xi\Omega(s, \eta)$ . To that end let  $B_\delta$  denote the open  $\| \cdot \|$ -ball in  $V$  with radius  $\delta > 0$ , centered at 0. We define

$$\delta_\xi(s) := s^{-q} \|v_\xi \varrho(\xi)\|.$$

PROPOSITION 3.1 (Corollary 2.3 in [Dan84]): *Let  $\xi \in \Xi$ ,  $s > 0$ , and  $g \in G$ . Then  $\Gamma g \in \Gamma \backslash \Gamma \xi \Omega(s, \eta)$  if and only if  $v_\xi \varrho(\gamma g) \in B_{\delta_\xi(s)}$  for some  $\gamma \in \Gamma$ . Further, if  $s \geq s_1$  and  $\gamma_1, \gamma_2 \in \Gamma$  satisfy  $v_\xi \varrho(\gamma_j g) \in B_{\delta_\xi(s)}$  for  $j = 1, 2$ , then*

$$v_\xi \varrho(\gamma_1 g) \in \{\pm v_\xi \varrho(\gamma_2 g)\}.$$

Thus

$$\mathcal{X}(\xi, s) = \Gamma \backslash \Gamma \xi \Omega(s, \eta)$$

for all  $\xi \in \Xi$  and  $s > 0$ . If  $s \geq s_1$ , the supremum in (2) is attained. Moreover, by Proposition 2.1(iv),

$$\mathcal{X}(\xi, s) \cap \mathcal{X}(\xi', s) = \emptyset$$

if  $\xi \neq \xi' \in \Xi$ . Hence the sets  $\mathcal{X}(\xi, s)$  are then disjoint neighborhoods of the cusps of  $\mathcal{X}$ , and the union in (3) is disjoint.

#### 4. Coordinate system for $G$

Recall that the Lie algebra  $\mathfrak{g}$  is the direct sum of a simple Lie algebra of rank 1 and a compact one. Since the height function is right- $\varrho(K)$ -invariant and all further considerations are right- $\varrho(K)$ -invariant, we can restrict to  $\mathfrak{g}$  being simple. [CDKR91] and [CDKR98] provide a classification-free construction of all Riemannian symmetric spaces of noncompact type and rank one. Their results rely on the choice of a certain coordinate system for real simple Lie groups  $G$  of real rank 1, which allows us to treat all these groups without referring to their classification. In the following we recall this coordinate system, the one for the associated symmetric spaces and some essential formulas.

The semidirect product  $NA$  is parametrized by

$$\mathbb{R}_{>0} \times \mathfrak{g}_2 \times \mathfrak{g}_1 \rightarrow NA, \quad (s, Z, X) \mapsto \exp(Z + X) \cdot a_s,$$

where we may assume  $s := \alpha(a_s)$ . The (left) action of  $a_s = (s, 0, 0) \in A$  on  $n = (1, Z, X) \in N$  is then given by

$$a_s n = (s, sZ, s^{1/2}X).$$

We define an inner product on  $\mathfrak{n} = \mathfrak{g}_2 \oplus \mathfrak{g}_1$  as follows. Let  $\mathfrak{k}$  be the Lie algebra of  $K$ . Let  $\theta$  be a Cartan involution of  $\mathfrak{g}$  such that  $\mathfrak{k}$  is its 1-eigenspace. For  $X, Y \in \mathfrak{n}$  we define

$$\langle X, Y \rangle := -\frac{1}{\dim \mathfrak{g}_1 + 4 \dim \mathfrak{g}_2} B(X, \theta Y)$$

where  $B$  is the Killing form of  $\mathfrak{g}$ . It is well-known that  $\langle \cdot, \cdot \rangle$  is an inner product on  $\mathfrak{n}$ . As in [CDKR91, CDKR98], we identify  $G/K \cong NA \cong \mathbb{R}_{>0} \times \mathfrak{g}_2 \times \mathfrak{g}_1$  with

$$D := \{(t, Z, X)_D \in \mathbb{R} \times \mathfrak{g}_2 \times \mathfrak{g}_1 \mid t > \frac{1}{4}|X|^2\}$$

via

$$\mathbb{R}_{>0} \times \mathfrak{g}_2 \times \mathfrak{g}_1 \rightarrow D, \quad (t, Z, X) \mapsto (t + \frac{1}{4}|X|^2, Z, X)_D.$$

We will include the subscript  $D$  when denoting elements  $(\cdot, \cdot, \cdot)_D$  of the symmetric space  $D$  to avoid confusion with elements of the group  $NA$ . The (left) action of an element  $s = (t_s, Z_s, X_s) \in NA$  on a point  $p = (t_p, Z_p, X_p)_D \in D$  becomes

$$s.p = (t_s t_p + \frac{1}{4}|X_s|^2 + \frac{1}{2}t_s^{1/2}\langle X_s, X_p \rangle, Z_s + t_s Z_p + \frac{1}{2}t_s^{1/2}[X_s, X_p], X_s + t_s^{1/2}X_p)_D.$$

These coordinates of  $G/K$  enable us to use [CDKR91, CDKR98], and they simplify some of the expressions below, in particular the one for the geodesic inversion. To state the geodesic inversion, we define the linear map

$$J: \mathfrak{g}_2 \rightarrow \text{End}(\mathfrak{g}_1), \quad Z \mapsto J_Z,$$

via

$$\langle J_Z X, Y \rangle = \langle Z, [X, Y] \rangle \quad \text{for all } X, Y \in \mathfrak{g}_1.$$

Then the geodesic inversion  $\sigma$  of  $D$  at  $o := (1, 0, 0)_D$  is given by (see [CDKR98])

$$\sigma(t, Z, X)_D = \frac{1}{t^2 + |Z|^2}(t, -Z, (-t + J_Z)X)_D.$$

We identify  $\sigma$  with an element in  $K$  which acts as geodesic inversion on  $D = G/K$  at  $o$ . Then  $G$  has the Bruhat decomposition ([CDKR98, Theorem 6.4])

$$G = NAM \cup NAM\sigma N.$$

Multiplying this with  $\xi \in \Xi$  from the left and  $\sigma$  from the right, we get

$$G = \xi NAM\sigma \cup \xi NAMU$$

with  $U := \sigma N\sigma$ . This decomposition provides a coordinate system on  $G$  adapted to the cusp represented by  $\xi$ . The set  $\xi NAM\sigma$  we call the **small  $\xi$ -Bruhat cell** and  $\xi NAMU$  the **big  $\xi$ -Bruhat cell**.

The group  $M$  is parametrized by the pairs  $(\varphi, \psi)$  consisting of the orthogonal endomorphisms  $\varphi$  on  $\mathfrak{g}_2$  resp.  $\psi$  on  $\mathfrak{g}_1$  such that  $\psi(J_Z X) = J_{\varphi(Z)}\psi(X)$  for all  $(Z, X) \in \mathfrak{g}_2 \times \mathfrak{g}_1$ . The action of  $(\varphi, \psi) \in M$  on  $p = (t, Z, X)_D \in D$  is given by

$$(\varphi, \psi).p = (t, \varphi(Z), \psi(X))_D.$$

By [CDKR98, Proposition 7.1],  $|J_Z X| = |Z||X|$  for all  $Z \in \mathfrak{g}_2, X \in \mathfrak{g}_1$ .

**5. Variation of height**

Suppose that the point  $x \in X$  is of big height and its trajectory stays far out for some time. In this section, we provide non-trivial bounds on the unstable components of a group element  $g \in G$  representing  $x$ . In Proposition 6.3 below, this bound implies constraints on the perturbation allowed for  $x$  without destroying the qualitative behavior of its trajectory during this time.

LEMMA 5.1: *Let  $a_t, a_r \in A, m \in M$  and  $n \in N$  with  $n = (1, Z, X)$  such that  $\sigma m n a_t \in N a_r K$ . Then*

$$r = \frac{t}{(t + \frac{1}{4}|X|^2)^2 + |Z|^2}.$$

*Proof.* By Iwasawa decomposition we know that  $\sigma m n a_t = n' a_r k$  for suitable  $n' \in N, k \in K$  and  $r \in \mathbb{R}_{>0}$ . Suppose that  $m = (\varphi, \psi)$ . Applying both  $\sigma m n a_t$  and  $n' a_r k$  to the base point  $o = (1, 0, 0)_D$  in  $D$ , we find

$$\sigma m n a_t \cdot o = n' a_r k \cdot o = n' a_r \cdot o.$$

In the coordinates of  $D$  one easily calculates that

$$\sigma m n a_t \cdot o = \frac{1}{(t + \frac{1}{4}|X|^2)^2 + |Z|^2} (t + \frac{1}{4}|X|^2, -\varphi(Z), (-t - \frac{1}{4}|X|^2 + J_{\varphi(Z)})\psi(X))_D.$$

Suppose that  $n' = (1, Z', X')$ . Then

$$n' a_r \cdot o = (r + \frac{1}{4}|X'|^2, Z', X')_D.$$

Thus

$$\begin{aligned} X' &= \frac{1}{(t + \frac{1}{4}|X|^2)^2 + |Z|^2} (-t - \frac{1}{4}|X|^2 + J_{\varphi(Z)})\psi(X), \\ |X'|^2 &= \frac{1}{((t + \frac{1}{4}|X|^2)^2 + |Z|^2)^2} \\ &\quad \times ((t + \frac{1}{4}|X|^2)^2 |X|^2 + |J_{\varphi(Z)}\psi(X)|^2 - 2(t + \frac{1}{4}|X|^2)\langle\psi(X), J_{\varphi(Z)}\psi(X)\rangle) \\ &= \frac{|X|^2}{(t + \frac{1}{4}|X|^2)^2 + |Z|^2}, \end{aligned}$$

and

$$r = \frac{t + \frac{1}{4}|X|^2}{(t + \frac{1}{4}|X|^2)^2 + |Z|^2} - \frac{1}{4}|X'|^2 = \frac{t}{(t + \frac{1}{4}|X|^2)^2 + |Z|^2}. \quad \blacksquare$$

LEMMA 5.2: *Let  $\xi \in \Xi$  and  $g \in G$ . If  $g = \xi na_s m \sigma$  with  $n \in N$  and  $m \in M$ , then*

$$\left( \frac{\|v_\xi \varrho(ga_t)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}} = \frac{s}{t}.$$

*If  $g = \xi na_s m \sigma(1, Z, X) \sigma$  with  $n \in N$  and  $m \in M$ , then*

$$\left( \frac{\|v_\xi \varrho(ga_t)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}} = s \cdot \frac{\frac{1}{t}}{\left(\frac{1}{t} + \frac{1}{4}|X|^2\right)^2 + |Z|^2}.$$

Recall the identification of the cusp represented by  $\xi \in \Xi$  with the cusp neighborhood  $(\Gamma \cap \xi N M \xi^{-1}) \backslash \xi \Omega(s_1, \eta)$  from Section 2. Let us note that the first case corresponds to a trajectory pointing straight out of the cusp represented by  $\xi$ . In the second case, the element  $u = \sigma(1, Z, X) \sigma$  determines the perturbation to the trajectory pointing straight into the cusp. If  $(Z, X) = (0, 0)$ , the second case corresponds to a trajectory pointing straight into the cusp, and the formula simplifies to

$$\left( \frac{\|v_\xi \varrho(ga_t)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}} = st.$$

*Proof of Lemma 5.2.* At first we suppose that  $g = \xi na_s m \sigma$ . Then

$$ga_t = \xi na_{s/t} m \sigma = \xi na_{s/t} \xi^{-1} \xi m \sigma \in (\xi N A \xi^{-1})(\xi K).$$

Hence

$$\|v_\xi \varrho(ga_t)\| = \theta_\xi(\xi a_{s/t} \xi^{-1}) \|v_\xi \varrho(\xi)\| = \left(\frac{s}{t}\right)^{-q} \|v_\xi \varrho(\xi)\|.$$

Suppose now that  $g = \xi n m a_s u$  with  $u = \sigma n' \sigma$  and  $n' = (1, Z, X)$ . Then (for some  $m' \in M$ )

$$\begin{aligned} \|v_\xi \varrho(ga_t)\| &= s^{-q} \|v_\xi \varrho(\xi \sigma m' n' \sigma a_t)\| = s^{-q} \|v_\xi \varrho(\xi \sigma m' n' a_{1/t} \sigma)\| \\ &= s^{-q} \|v_\xi \varrho(\xi \sigma m' n' a_{1/t})\|. \end{aligned}$$

Lemma 5.1 yields

$$\sigma m' n' a_{1/t} = n'' a_r k$$

for some  $n'' \in N$ ,  $k \in K$  and

$$r = \frac{\frac{1}{t}}{\left(\frac{1}{t} + \frac{1}{4}|X|^2\right)^2 + |Z|^2}.$$

Thus,

$$\|v_\xi \varrho(ga_t)\| = \left( \frac{\frac{1}{t}}{\left(\frac{1}{t} + \frac{1}{4}|X|^2\right)^2 + |Z|^2} \right)^{-q} s^{-q} \|v_\xi \varrho(\xi)\|. \quad \blacksquare$$

The following proposition describes the amount of time a trajectory spends in a neighborhood of the cusp represented by  $\xi$ .

**PROPOSITION 5.3:** *Let  $\xi \in \Xi$  and  $g \in G$ . Write  $\delta := \|v_\xi \varrho(g)\|$ . If  $g \in \xi NAM\sigma$ , then  $v_\xi \varrho(ga_t) \in B_\delta$  if and only if  $t < 1$ . If  $g = \xi na_s mu \in \xi NAMU$  with  $u = \sigma(1, Z, X)\sigma$ , then  $v_\xi \varrho(ga_t) \in B_\delta$  if and only if*

$$t \in \left( \frac{1}{\frac{1}{16}|X|^4 + |Z|^2}, 1 \right) \cup \left( 1, \frac{1}{\frac{1}{16}|X|^4 + |Z|^2} \right).$$

If  $u = \text{id}$ , then  $(\frac{1}{16}|X|^4 + |Z|^2)^{-1}$  is to be understood as  $\infty$ .

*Proof.* The first part of the statement follows immediately from Lemma 5.2. Suppose now that  $g = \xi nma_s u$  with  $u = \sigma n' \sigma$  and  $n' = (1, Z, X)$ . By Lemma 5.2,

$$(4) \quad \|v_\xi \varrho(ga_r)\| = \left( \frac{\frac{1}{r}}{(\frac{1}{r} + \frac{1}{4}|X|^2)^2 + |Z|^2} \right)^{-q} s^{-q} \|v_\xi \varrho(\xi)\|.$$

Applying (4) for  $r = 1$  and  $r = t$ , we see that

$$\|v_\xi \varrho(ga_t)\| < \|v_\xi \varrho(g)\|$$

if and only if

$$\frac{1}{(1 + \frac{1}{4}|X|^2)^2 + |Z|^2} < \frac{\frac{1}{t}}{(\frac{1}{t} + \frac{1}{4}|X|^2)^2 + |Z|^2},$$

which is equivalent to

$$\left(1 - \frac{1}{t}\right) \left(-\frac{1}{t} + \frac{1}{16}|X|^4 + |Z|^2\right) < 0.$$

This is the case if and only if

$$|Z|^2 + \frac{1}{16}|X|^4 < \frac{1}{t} < 1 \quad \text{or} \quad |Z|^2 + \frac{1}{16}|X|^4 > \frac{1}{t} > 1. \quad \blacksquare$$

Suppose that  $\|v_\xi \varrho(ga_t)\| = \|v_\xi \varrho(\gamma ga_t)\|$  for some  $g \in G$ ,  $\gamma \in \Gamma$  and all  $t$  in a non-trivial interval (i.e., an interval which contains at least two points). Then Lemma 5.2 yields that  $g$  and  $\gamma g$  have the same  $A$ -component in  $\xi NAK$  and they are in the same  $\xi$ -Bruhat cell. If, moreover,  $g$  and  $\gamma g$  are in the big  $\xi$ -Bruhat cell, then also the norms of their  $U$ -components are equal. The following lemma shows that far out in the cusp much more is true.

LEMMA 5.4: *Let  $\xi \in \Xi$  and suppose that  $g \in G$  and  $\gamma \in \Gamma$  are such that*

$$\|v_\xi \varrho(g)\| = \|v_\xi \varrho(\gamma g)\| < \delta_\xi(s_1).$$

*Then  $\gamma \in \xi NM\xi^{-1}$ . In particular, if  $g = \xi nam\sigma$  resp.  $g = \xi namu$  with  $n \in N$ ,  $a \in A$ ,  $m \in M$  and  $u \in U$ , then  $\gamma g = \xi n'am'\sigma$  resp.  $\gamma g = \xi n'am'u$  for some  $n' \in N$ ,  $m' \in M$ .*

*Proof.* By [Dan84, Lemma 2.2] (see also Proposition 3.1), for each  $s > 0$  we have

$$L_\xi \xi A_s K = \{g \in G \mid v_\xi \varrho(g) \in B_{\delta_\xi(s)}\}.$$

Hence  $g, \gamma g \in L_\xi \xi A_{s_1} K$ . By [Dan84, Remark 1.3] (with  $\eta$  as in Proposition 2.1),

$$L_\xi \xi A_{s_1} K = (\Gamma \cap L_\xi) \xi \eta A_{s_1} K.$$

Hence there exist  $\gamma_1, \gamma_2 \in \Gamma \cap L_\xi$ ,  $h_1, h_2 \in \eta A_{s_1} K$  such that

$$g = \gamma_1 \xi h_1, \quad \gamma g = \gamma_2 \xi h_2.$$

Therefore

$$g \in \gamma_1 \xi \Omega(s_1, \eta) \cap \gamma^{-1} \gamma_2 \xi \Omega(s_1, \eta).$$

Proposition 2.1(iv) yields  $\gamma_1^{-1} \gamma^{-1} \gamma_2 \in \xi NM\xi^{-1}$ . Thus,  $\gamma \in \xi NM\xi^{-1}$ . ■

For the proof of the following proposition we recall that the supremum in the definition of  $\xi$ -height (2) is realized if  $\text{ht}_\xi(x) \geq s_1$ .

PROPOSITION 5.5: *Let  $s > s_1$  and  $x \in \mathcal{X}$ . Suppose that there exists an interval  $I$  in  $\mathbb{R}$  such that  $\text{ht}(xa_t) > s$  for all  $t \in I$ . Then there exists a unique cusp representative  $\xi \in \Xi$  and a (non-unique) element  $g \in G$  with  $x = \Gamma g$  such that*

$$\text{ht}(xa_t) = \text{ht}_\xi(xa_t) = \left( \frac{\|v_\xi \varrho(ga_t)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}}$$

*for all  $t \in I$ . Moreover, if  $1 \in I$  and if there exists  $t \in I$  with  $t > 1$  and  $\text{ht}(xa_t) > \text{ht}(x)$ , then  $g = \xi na_r mu$  for some  $r > 0$ ,  $n \in N$ ,  $m \in M$  and  $u \in U$ . The elements  $a_r$  and  $u$  do not depend on the choice of  $g$ . Finally, if  $u = \sigma(1, Z, X)\sigma$ , then*

$$|X| < 2t^{-1/4} \quad \text{and} \quad |Z| < t^{-1/2}.$$

*Proof.* If  $y \in \mathcal{X}$  and  $\xi \in \Xi$  such that  $\text{ht}_\xi(y) > s_1$ , then there exists  $h \in G$  such that  $y = \Gamma h$  and

$$\text{ht}_\xi(y) = \left( \frac{\|v_\xi \varrho(h)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}}.$$

Since the function

$$\begin{cases} \mathbb{R}_{>0} \rightarrow \mathbb{R} \\ r \mapsto \|v_\xi \varrho(ga_r)\| \end{cases}$$

is continuous, there exists an open neighborhood  $J$  of 1 in  $\mathbb{R}_{>0}$  such that  $\text{ht}_\xi(ya_r) > s_1$  for all  $r \in J$ . For  $\xi \in \Xi$  let

$$J_\xi := \{t \in I \mid \text{ht}_\xi(xa_t) > s\}.$$

These sets are pairwise disjoint, open in  $I$  and cover  $I$ . Since  $I$  is connected, there exists a unique  $\xi \in \Xi$  with  $I = J_\xi$ . Thus

$$\text{ht}(xa_t) = \text{ht}_\xi(xa_t)$$

for all  $t \in I$ . For each  $t \in I$  pick an element  $g_t \in G$  such that  $x = \Gamma g_t$  and

$$\text{ht}_\xi(xa_t) = \left( \frac{\|v_\xi \varrho(g_t a_t)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}}.$$

Let  $J_t$  be the set of  $p \in I$  such that

$$\text{ht}_\xi(xa_p) = \left( \frac{\|v_\xi \varrho(g_t a_p)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}}.$$

Then  $I$  is covered by the sets  $J_t$ , and these are open in  $I$  by Proposition 3.1. If  $J_t$  and  $J_r$  overlap for some  $t, r \in I, t \neq r$ , then Lemmas 5.4 and 5.2 imply that  $J_t = J_r$ . In turn,  $J_t = I$  for each  $t \in I$ .

The remaining statements follow immediately from Proposition 5.3 and Lemma 5.4. ■

### 6. Common cusp excursions of nearby points

For  $s > 0$  we define

$$\mathcal{X}_{\leq s} := \mathcal{X} \setminus \mathcal{X}_{> s}.$$

Further, we let

$$r_0 := \alpha(\tilde{a})$$

and recall that  $r_0 > 1$  by our choice of  $\alpha$ .

Each connected component of  $\mathcal{X}$  of height above  $s_1$  can essentially be identified with a Siegel set (cf. Proposition 2.1). For the proof of the main theorem, trajectories of points  $x \in \mathcal{X}$  are only considered time-discretized by the map  $T$ . In the following lemma we construct a height level  $s_2$  above which we can identify pieces of these discretized trajectories with trajectory segments in a

Siegel set. More specifically, as soon as we know that two consecutive points of the discretized trajectory stay above height  $s \geq s_2$ , then the (continuous) trajectory segment of the corresponding geodesic also stays above height  $s$  and, in particular, does not visit the compact set  $\mathcal{X}_{\leq s_1}$ . Then we construct a second height level  $s_3 > s_2$  such that any discretized trajectory entering  $\mathcal{X}_{> s_3}$  can locally be identified with a continuous trajectory segment in the Siegel set. In Section 8 below this will be crucial to effectively determine the behavior of nearby starting trajectories. Of special importance for Section 7 below is the item (v) of the following lemma, which states that if we start to descend somewhere high in a cusp, then we actually descend up to below height  $s_3$ .

LEMMA 6.1: *There exist  $s_3 > s_2 > s_1$  such that we have the following properties:*

- (i) *If  $x \in \mathcal{X}_{> s_2}$ , then  $\text{ht}(xa_t) > 2s_1$  for all  $t \in [r_0^{-1}, r_0]$ .*
- (ii) *If  $s \geq s_2$  and  $x, Tx \in \mathcal{X}_{> s}$ , then  $\text{ht}(xa_t) > s$  for all  $t \in [1, r_0]$ .*
- (iii) *Let  $s > s_3$ . If  $x \in \mathcal{X}_{\leq s_3}$  and  $T^j x \in \mathcal{X}_{> s}$  for some  $j \in \mathbb{N}$ , then there exists  $n \in \{0, \dots, j - 1\}$  such that  $\text{ht}(T^n x) \leq s_3$  and  $\text{ht}(xa_t) > s_2$  for all  $t \in [r_0^n, r_0^j]$ .*
- (iv) *Let  $s > s_3$ . If  $x \in \mathcal{X}_{> s}$  and  $T^j x \in \mathcal{X}_{\leq s_3}$  for some  $j \in \mathbb{N}$ , then there exists  $n \in \{1, \dots, j\}$  such that  $\text{ht}(T^n x) \leq s_3$  and  $\text{ht}(xa_t) > s_2$  for all  $t \in [1, r_0^n]$ .*
- (v) *Let  $s > s_3$ . If  $x \in \mathcal{X}_{> s}$  and  $Tx \in \mathcal{X}_{\leq s}$ , then there exists  $n \in \mathbb{N}$  such that  $T^n x \in \mathcal{X}_{\leq s_3}$  and  $T^k x \in \mathcal{X}_{\leq s}$  for all  $k = 1, \dots, n$ .*

*Proof.* We will choose  $s_2 > s_1$  below. Let  $x \in \mathcal{X}_{> s_2}$ . We wish to prove that  $xa_t \in \mathcal{X}_{> 2s_1}$  for all  $t \in [r_0^{-1}, r_0]$ . Since  $s_2 > s_1$ , there exist by Proposition 3.1 a unique  $\xi \in \Xi$  and an element  $g \in G$  such that  $x = \Gamma g$  and

$$\text{ht}(x) = \text{ht}_\xi(x) = \left( \frac{\|v_\xi \varrho(g)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}}.$$

Further, for all  $t \in [r_0^{-1}, r_0]$ , we have

$$\text{ht}(xa_t) \geq \text{ht}_\xi(xa_t) \geq \left( \frac{\|v_\xi \varrho(ga_t)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}}.$$

However, now it is clear that if  $s_2$  is sufficiently big<sup>2</sup> or equivalently  $\|v_\xi \varrho(g)\|$  is sufficiently small, this will force  $\|v_\xi \varrho(ga_t)\|$  for  $t \in [r_0^{-1}, r_0]$  sufficiently small to get the claim in (i).

---

<sup>2</sup> A more careful analysis using Lemma 5.2 reveals that  $s_2 > 2r_0^2 s_1$  suffices.

For the proof of the remaining properties we will use the (quite natural) monotonicity properties of the functions appearing in Lemma 5.2. So assume that

$$\text{ht}(x) = \text{ht}_\xi(x) = \left( \frac{\|v_\xi \varrho(g)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}} > s_1$$

for  $x = \Gamma g$  (with the cusp representative  $\xi$  and  $\pm v_\xi \varrho(g)$  uniquely determined by Proposition 3.1). If  $g = \xi n a_s m \sigma$  is as in the first part of Lemma 5.2, then the trajectory comes straight out of the cusp. Hence  $\text{ht}(x a_t) = \frac{s}{t}$  is monotonically decreasing until it reaches the value  $s_1$  (at which point Proposition 3.1 will not apply any longer). In the more general case, if  $g = \xi n a_s m \sigma(1, Z, X) \sigma$  is as in the second part of Lemma 5.2, then the height of  $x a_t$  is given by the formula

$$\text{ht}(x a_t) = s \cdot \frac{\frac{1}{t}}{\left(\frac{1}{t} + \frac{1}{4}|X|^2\right)^2 + |Z|^2} = \frac{s}{\frac{1}{t} + \frac{1}{4}|X|^2 + \left(|Z|^2 + \frac{|X|^4}{16}\right)t},$$

at least for all  $t$  for which the right-hand side is  $\geq s_1$ . If  $X = 0$  and  $Z = 0$ , the right-hand side equals  $st$  and the orbit points straight into the cusp. However, in general the right-hand side has a unique maximum, is monotonically increasing left to the maximum and monotonically decreasing to the right of the maximum.

Property (i) and these monotonicity properties imply (ii).

We choose  $s_3$  in the same way as  $s_2$  but with  $s_2$  replacing  $s_1$  in (i). Assume now  $s > s_3$ ,  $x \in \mathcal{X}_{\leq s_3}$  and  $T^j x \in \mathcal{X}_{> s}$  for some  $j \in \mathbb{N}$ . We choose the maximal integer  $n < j$  with  $\text{ht}(T^n x) \leq s_3$ . By our choice of  $s_3$  we have  $\text{ht}(x a_t) > 2s_2$  for  $t \in [r_0^{n-1}, r_0^{n+1}]$ . Using the above monotonicity properties now implies (iii). Property (iv) follows in the same way using the first  $n \leq j$  with  $\text{ht}(T^n x) \leq s_3$ .

Property (v) follows directly from the monotonicity properties. ■

Given a point  $x \in \mathcal{X}$  whose orbit stays near the cusp represented by  $\xi$  for the next  $S$  steps, Proposition 6.3 below provides non-trivial constraints on small perturbations of  $x$  which do not destroy the qualitative behavior of the orbit for these next  $S$  steps. The following lemma is needed for its proof.

LEMMA 6.2: *Let  $D^U$  be a bounded subset of  $U$ . Let  $\xi \in \Xi$  and  $g = \xi n a_r m u \in G$  with  $n \in N$ ,  $a_r \in A$ ,  $m \in M$  and  $u = \sigma(1, Z, X) \sigma \in D^U$ . Suppose that*

$$\left( \frac{\|v_\xi \varrho(g a_t)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}} > \lambda \left( \frac{\|v_\xi \varrho(g)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}}$$

for some  $t > 1$  and  $\lambda > 0$ . Then there exist  $c_1, c_2 > 0$ , only depending on  $D^U$  and  $\lambda$ , such that

$$|X| < c_1 t^{-1/4} \quad \text{and} \quad |Z| < c_2 t^{-1/2}.$$

*Proof.* For  $\lambda \geq 1$ , the statement is already proven in Proposition 5.5. So suppose  $1 > \lambda > 0$ . Invoking Lemma 5.2 we find

$$(5) \quad t \left[ \left( \frac{1}{t} + \frac{1}{4} |X|^2 \right)^2 + |Z|^2 \right] < \frac{1}{\lambda} \left[ \left( 1 + \frac{1}{4} |X|^2 \right)^2 + |Z|^2 \right].$$

Thus,

$$t \left( \frac{1}{t} + \frac{1}{4} |X|^2 \right)^2 < \frac{1}{\lambda} \left( 1 + \frac{1}{4} |X|^2 \right)^2 + (\lambda^{-1} - t) |Z|^2.$$

For  $t > \lambda^{-1}$ , it follows that

$$t \left( \frac{1}{t} + \frac{1}{4} |X|^2 \right)^2 < \lambda^{-1} \left( 1 + \frac{1}{4} |X|^2 \right)^2.$$

Therefore,

$$|X|^2 < 4 \frac{(t\lambda^{-1})^{\frac{1}{2}} - 1}{t^{\frac{1}{2}} - \lambda^{-\frac{1}{2}}} t^{-\frac{1}{2}}.$$

Hence, for  $t > \lambda^{-1} + 1$ , we have

$$|X| < c_1 t^{-\frac{1}{4}}$$

for some constant  $c_1 > 0$ . Since  $|X|$  is bounded, by possibly choosing a larger  $c_1$ , this estimate holds for all  $t > 1$ . To deduce the bound for  $|Z|$  we note that (5) yields

$$\begin{aligned} (t - \lambda^{-1}) |Z|^2 &< \lambda^{-1} \left( 1 + \frac{1}{4} |X|^2 \right)^2 - t \left( \frac{1}{t} + \frac{1}{4} |X|^2 \right)^2 \\ &< \lambda^{-1} \left( 1 + \frac{1}{4} c_1^2 \right)^2 = c_3. \end{aligned}$$

Suppose that  $t > \lambda^{-1} + 1$ . Then

$$|Z|^2 < \frac{c_3}{t - \lambda^{-1}} = \frac{c_3}{1 - (t\lambda)^{-1}} t^{-1}.$$

The factor in front of  $t^{-1}$  is bounded. Thus,

$$|Z| < c_2 t^{-\frac{1}{2}}$$

for some constant  $c_2 > 0$ . As before, since  $|Z|$  is bounded, this estimate holds for all  $t > 1$  after possibly choosing a larger  $c_2$ . This completes the proof. ■

Let  $d$  be the left- $G$ -invariant metric on  $G$  induced from a left-invariant Riemannian metric that is induced by an inner product on  $\mathfrak{g}$ . For  $r > 0$  let  $B_r^G$  denote the open  $d$ -ball in  $G$  centered at the identity of  $G$  with radius  $r$ . For  $\kappa > 0$  let  $D_\kappa^U$  denote the subset of  $U$  consisting of the elements  $u = \sigma(1, Z, X)\sigma$  with  $|Z| < \kappa$  and  $|X| < \kappa$ , and let  $D_\kappa^{NAM} := B_\kappa^G \cap NAM$ . Further, let

$$(6) \quad D_\kappa := D_\kappa^U D_\kappa^{NAM}.$$

Then  $D_\kappa$  is open. We choose  $\kappa > 0$  such that for all  $h \in D_\kappa$  we have

$$(7) \quad \|\varrho(h)\|, \|\varrho(h^{-1})\| \leq \left(\frac{s_1}{s_2}\right)^{-q}.$$

We consider  $\kappa$  to be fixed throughout and will shrink it if necessary (e.g., in the paragraph before Lemma 7.3).

**PROPOSITION 6.3:** *There exist  $c_3, c_4 > 0$  such that the following holds: Let  $x \in \mathcal{X}$ ,  $S \in \mathbb{N}$ ,  $h \in D_\kappa$  be such that  $\text{ht}(T^j x) > s_2$  and  $\text{ht}(T^j(xh)) > s_2$  for  $j = 0, \dots, S$ ,  $\text{ht}(T^S x) > \text{ht}(x)$  and  $\text{ht}(T^S(xh)) > \text{ht}(xh)$ . Suppose that  $h = \sigma(1, Z, X)\sigma na_r m$ . Then*

$$|X| \leq c_3 r_0^{-S/4} \quad \text{and} \quad |Z| \leq c_4 r_0^{-S/2}.$$

*Proof.* By Lemma 6.1 we have  $\text{ht}(xa_t) > s_2$  and  $\text{ht}(xha_t) > s_2$  for all  $t \in [1, r_0^S]$ . Since  $s_2 > s_1$ , Proposition 5.5 shows that there exist a unique cusp representative  $\xi \in \Xi$  and an element  $g \in G$  such that  $x = \Gamma g$  and

$$\text{ht}(xa_t) = \left(\frac{\|v_\xi \varrho(ga_t)\|}{\|v_\xi \varrho(\xi)\|}\right)^{-\frac{1}{q}}$$

for all  $t \in [1, r_0^S]$ . Moreover, there exist a unique cusp representative  $\xi_1 \in \Xi$  and an element  $g_1 \in G$  such that  $xh = \Gamma g_1 h$  and

$$\text{ht}(xha_t) = \left(\frac{\|v_{\xi_1} \varrho(g_1 ha_t)\|}{\|v_{\xi_1} \varrho(\xi_1)\|}\right)^{-\frac{1}{q}}$$

for all  $t \in [1, r_0^S]$ . In the following we show that  $\xi = \xi_1$  and that we can choose  $g_1 = g$ . We have

$$\|v_\xi \varrho(gha_t)\| = \|v_\xi \varrho(ga_t a_{t-1} ha_t)\| \leq \|v_\xi \varrho(ga_t)\| \cdot \|\varrho(a_{t-1} ha_t)\|.$$

Now,  $a_{t-1} ha_t \in D_\kappa$  for  $t$  near 1, say in the non-trivial interval  $I$ . By (7), for  $t \in I$  this yields

$$\|\varrho(a_{t-1} ha_t)\| \leq \left(\frac{s_1}{s_2}\right)^{-q}.$$

Thus, for  $t \in I$ ,

$$\begin{aligned} \|v_\xi \varrho(gha_t)\| &\leq \|v_\xi \varrho(ga_t)\| \left(\frac{s_1}{s_2}\right)^{-q} < \left(\frac{s_1}{s_2}\right)^{-q} s_2^{-q} \|v_\xi \varrho(\xi)\| \\ &= s_1^{-q} \|v_\xi \varrho(\xi)\|. \end{aligned}$$

Hence, for  $t \in I$ ,

$$\left(\frac{\|v_\xi \varrho(gha_t)\|}{\|v_\xi \varrho(\xi)\|}\right)^{-\frac{1}{q}} > s_1.$$

The uniqueness of  $\xi_1$  yields  $\xi_1 = \xi$ . Moreover, we can choose  $g_1 = g$  for  $t \in I$ . As in the proof of Proposition 5.5, we see that we can choose  $g_1 = g$  for all  $t \in [1, r_0^S]$ .

Proposition 5.5 shows that  $g \in \xi NAMU$ , say  $g = \xi n_4 a_{r_1} m_1 u_1$  with  $u_1 = \sigma(1, Z_1, X_1)\sigma$ , and that

$$(8) \quad |X_1| < 2r_0^{-S/4} \quad \text{and} \quad |Z_1| < r_0^{-S/2}.$$

Suppose that  $h = u_2 n_3 a_{r_2} m_2$  and set  $h_2 := n_3 a_{r_2} m_2$ . Then

$$\|v_\xi \varrho(gh)\| = \|v_\xi \varrho(gu_2 h_2)\| \leq \|v_\xi \varrho(gu_2)\| \|\varrho(h_2)\| \leq \|v_\xi \varrho(gu_2)\| \left(\frac{s_1}{s_2}\right)^{-q}$$

and

$$\begin{aligned} \|v_\xi \varrho(gu_2 a^S)\| &= \|v_\xi \varrho(gha^S a^{-S} h_2^{-1} a^S)\| \\ &\leq \|v_\xi \varrho(gha^S)\| \|\varrho(a^{-S} h_2^{-1} a^S)\| \leq \|v_\xi \varrho(gha^S)\| \left(\frac{s_1}{s_2}\right)^{-q}. \end{aligned}$$

This yields

$$\begin{aligned} \left(\frac{\|v_\xi \varrho(gu_2 a^S)\|}{\|v_\xi \varrho(\xi)\|}\right)^{-\frac{1}{q}} &\geq \frac{s_1}{s_2} \left(\frac{\|v_\xi \varrho(gha^S)\|}{\|v_\xi \varrho(\xi)\|}\right)^{-\frac{1}{q}} = \frac{s_1}{s_2} \text{ht}(xha^S) \\ &> \frac{s_1}{s_2} \text{ht}(xh) = \frac{s_1}{s_2} \left(\frac{\|v_\xi \varrho(gh)\|}{\|v_\xi \varrho(\xi)\|}\right)^{-\frac{1}{q}} \\ &\geq \left(\frac{s_1}{s_2}\right)^2 \left(\frac{\|v_\xi \varrho(gu_2)\|}{\|v_\xi \varrho(\xi)\|}\right)^{-\frac{1}{q}}. \end{aligned}$$

Let  $u_2 = \sigma(1, Z_2, X_2)\sigma$ . Then

$$u_1 u_2 = \sigma(1, Z_1 + Z_2 + \frac{1}{2}[X_1, X_2], X_1 + X_2)\sigma.$$

From (8) and  $u_2 \in D_\kappa^U$  it follows that

$$|X_1 + X_2| \leq |X_1| + |X_2| < 2 + \kappa.$$

Moreover, using triangle inequality and [Poh10, Lemma 2.12, Proposition 3.3] we find

$$|Z_1 + Z_2 + \frac{1}{2}[X_1, X_2]| \leq |Z_1| + |Z_2| + \frac{1}{2}|X_1||X_2| < 1 + 2\kappa.$$

Thus,  $u_1u_2$  is contained in the bounded set  $D_{2+2\kappa}^U$ . Note that this set only depends on  $\kappa$ . Then Lemma 6.2 gives

$$|X_1 + X_2| < c_1r_0^{-S/4} \quad \text{and} \quad |Z_1 + Z_2 + \frac{1}{2}[X_1, X_2]| < c_2r_0^{-S/2},$$

where the constants  $c_1, c_2$  only depend on  $s_1, s_2$  and  $\kappa$ . It follows that

$$|X_2| < c_1r_0^{-S/4} + |X_1| < (c_1 + 2)r_0^{-S/4}$$

and

$$\begin{aligned} |Z_2| &\leq |Z_1 + Z_2 + \frac{1}{2}[X_1, X_2]| + |Z_1| + \frac{1}{2}|X_1||X_2| \\ &\leq c_2r_0^{-S/2} + r_0^{-S/2} + (c_1 + 2)r_0^{-S/2}. \end{aligned}$$

This completes the proof. ■

### 7. Estimate of metric entropy and proof of Theorem A

This section, in which we prove Theorem A, can be understood independently from the previous ones if one is willing to accept the following facts previously shown: The height level  $s_3$  is chosen such that the connected parts of  $\mathcal{X}_{>s_3}$  (thus, cuspidal ends of uniform “length”) can be identified with  $(\Gamma \cap P) \backslash C$ , where  $C$  is the cylindrical set  $C = \xi A_{s_3} NK$  at the cusp represented by  $\xi$  of the considered end and  $P$  is the corresponding minimal parabolic subgroup in  $G$ . In particular, this means that connected parts of geodesic trajectories in  $\mathcal{X}_{>s_3}$  can be identified with any representing geodesic trajectories in  $C$ . As a consequence we know (see Lemma 6.1) that (discretized) geodesic trajectories in  $\mathcal{X}_{>s_3}$  which start to move out of the cusp actually descend to below height level  $s_3$ , and geodesics in  $\mathcal{X}$  which move from one of these cuspidal ends to another one necessarily have to pass through the compact part  $\mathcal{X}_{\leq s_3}$ . Moreover, if the trajectories of two nearby points  $x, xh$  in  $\mathcal{X}$  ( $h \in G$ ) stay together near a cusp (meaning in the same connected component of  $\mathcal{X}_{>s_3}$ ) for “time”  $t$ , then the unstable component of  $h$  is restricted (up to a multiplicative constant) by  $t^{-1/2}$  in the direction of the long root and by  $t^{-1/4}$  in the direction of the short root (see Proposition 6.3).

Let  $\mathcal{M}_1(\mathcal{X})^T$  denote the set of  $T$ -invariant probability measures on  $\mathcal{X}$ . Let  $\mu \in \mathcal{M}_1(\mathcal{X})^T$  and suppose that  $\mathcal{P}$  is a partition of  $\mathcal{X}$  (consisting of measurable sets). We denote the static entropy of  $\mathcal{P}$  with respect to  $\mu$  by

$$(9) \quad H_\mu(\mathcal{P}) = - \sum_{P \in \mathcal{P}} \mu(P) \log \mu(P).$$

For  $n \in \mathbb{N}_0$  let

$$\mathcal{P}_0^n := \bigvee_{j=0}^n T^{-j}\mathcal{P} = \{P_{j_0} \cap T^{-1}P_{j_1} \cap \dots \cap T^{-n}P_{j_n} \mid P_{j_i} \in \mathcal{P}\}.$$

Then

$$h_\mu(T, \mathcal{P}) = \inf_{n \in \mathbb{N}} \frac{1}{n} H_\mu(\mathcal{P}_0^{n-1})$$

is the dynamical entropy of  $(T, \mathcal{P})$  with respect to  $\mu$ . Finally,

$$\begin{aligned} h_\mu(T) &= \sup\{h_\mu(T, \mathcal{P}) \mid \mathcal{P} \text{ partition of } \mathcal{X}, H_\mu(\mathcal{P}) < \infty\} \\ &= \sup\{h_\mu(T, \mathcal{P}) \mid \mathcal{P} \text{ finite partition of } \mathcal{X}\} \end{aligned}$$

is the (metric) entropy of  $T$  with respect to  $\mu$ .

In our set-up there exists a unique maximal entropy measure for  $T$ . We provide a reference for this statement and recall how to calculate its value in the following proposition. Set  $p_1 := \dim \mathfrak{g}_1, p_2 := \dim \mathfrak{g}_2$  and recall that  $\tilde{a} = a_{r_0}$ .

**PROPOSITION 7.1:** *The maximal entropy of  $T$  is achieved by the Haar measure  $m$  on  $\mathcal{X}$  and is given by*

$$h_m(T) = \max\{h_\mu(T) \mid \mu \in \mathcal{M}_1(\mathcal{X})^T\} = \left(\frac{p_1}{2} + p_2\right) \log r_0.$$

*Moreover, the Haar measure is the only  $T$ -invariant probability measure that achieves this maximal entropy.*

*Proof.* The statement follows from a combination of the proposition in Section 9.3 in [MT94] and Lemma 9.5 and Proposition 9.6 in [MT94]. If  $G$  is algebraic, a more accessible reference is [EL10, Theorem 7.6]. Note that

$$-\log \det(\text{Ad}_a \mid_{\mathfrak{g}_{-1} \oplus \mathfrak{g}_{-2}}) = \left(\frac{p_1}{2} + p_2\right) \log r_0. \quad \blacksquare$$

For  $r > 0$  we call

$$(10) \quad B_L := B_L(r) := \bigcap_{j=0}^{L-1} \tilde{a}^j B_r^G \tilde{a}^{-j}$$

a (**forward**) **Bowen  $L$ -ball** in  $G$  with (**radius**) **parameter**  $r$ . Further, any subset of  $\mathcal{X}$  of the form

$$(11) \quad xB_L = xB_L(r)$$

with  $x \in \mathcal{X}$  is called a **Bowen  $L$ -ball** in  $\mathcal{X}$  with **center**  $x$  and (**radius**) **parameter**  $r$ .

Through the work of Brin–Katok [BK83] it is well known that entropy is strongly related to the decay rate of the measure of Bowen  $L$ -balls. For the Haar measure this can be established quite directly and in the following strong form (which will be used in many covering arguments below).

LEMMA 7.2: *Let  $r > 0$  be sufficiently small (depending only on  $G$ ) and  $L \in \mathbb{N}$ . Then*

$$r^{\dim G} e^{-h_m(T)L} \ll m(B_L(r)) \ll r^{\dim G} e^{-h_m(T)L},$$

where the implied constants only depend on  $G$  and  $\tilde{a}$ .

*Proof.* Recall from (6) the definition of  $D_r$ . We find  $r_1, r_2 > 0$  (uniform for small  $r$ ) such that

$$D_{r_1 r} \subseteq B_r^G \subseteq D_{r_2 r}.$$

Then

$$D^{(L)}(r_1 r) := \bigcap_{j=0}^{L-1} \tilde{a}^j D_{r_1 r} \tilde{a}^{-j} \subseteq B_L(r) = \bigcap_{j=0}^{L-1} \tilde{a}^j B_r^G \tilde{a}^{-j} \subseteq \bigcap_{j=0}^{L-1} \tilde{a}^j D_{r_2 r} \tilde{a}^{-j}.$$

One easily checks that

$$D^{(L)}(r) = \tilde{a}^{L-1} D_r^U \tilde{a}^{-(L-1)} D_r^{NAM}$$

and

$$\tilde{a}^{L-1} D_r^U \tilde{a}^{-(L-1)} = \{\sigma(1, Z, X)\sigma \mid |Z| < rr_0^{-(L-1)/2}, |X| < rr_0^{-(L-1)/4}\}.$$

Let  $du, dn, da$  and  $dm$  be Haar measures on  $U, N, A$  and  $M$ , respectively, and let  $dg$  denote the Haar measure on  $G$ . With appropriate normalizations we have ([Hel00, Chapter I, Proposition 5.21], and [Hel00, Chapter I, Corollary 5.2] for the change of order of integration)

$$\int_G f(g) dg = \int_{U \times N \times A \times M} f(unam) dudndadm$$

for all  $f \in C_c(G)$ . Further, we recall from [Hel00, Chapter I, Theorem 1.14] that if the support of  $f \in C_c(G)$  is contained in the canonical coordinate neighborhood of  $G$ , then

$$(12) \quad \int_G f(g)dg = \int_{\mathfrak{g}} f(\exp W) \det \left( \frac{1 - e^{-\text{ad } W}}{\text{ad } W} \right) dW,$$

where  $dW$  is the Euclidean measure on  $\mathfrak{g}$  which coincides with  $(dg)_{\text{id}}$ .

We now use the coordinates  $(Z, X) \in \mathfrak{g}_2 \times \mathfrak{g}_1$  for the Lie algebra  $\mathfrak{u}$  of  $U$ . Since  $\mathfrak{u}$  is two-step nilpotent, the Jacobian determinant in (12) (applied for  $G = U$ ) equals 1 for all  $W \in \mathfrak{u}$ . With an appropriate global constant  $c_U$ , the Haar measure  $du$  is then

$$m_U(f) := \int_U f(u)du = c_U \int_{\mathfrak{u}} f(\sigma(1, Z, X)\sigma) dZ dX.$$

Thus,

$$m_U(\tilde{a}^{L-1} D_r^U \tilde{a}^{-(L-1)}) = c_U r^{\dim U} e^{-h_m(T)(L-1)}.$$

Hence

$$m(D^{(L)}(r_1 r)) = c_U r_1^{\dim U} r^{\dim U} e^{-h_m(T)(L-1)} m_{NAM}(D_{r_1 r}^{NAM}),$$

where  $m_{NAM} := dn \otimes da \otimes dm$ . We may assume that  $D_r^{NAM} = B_r^{NAM}$ . For sufficiently small  $r > 0$ , the parameter space in  $\mathfrak{n} \times \mathfrak{a} \times \mathfrak{m}$  for the set  $D_r^{NAM}$  is the spherical normal neighborhood  $V_r = \{W \in \mathfrak{n} \times \mathfrak{a} \times \mathfrak{m} \mid \|W\| < r\}$  (see [Hel01, Chapter I, Proposition 9.4]). On this neighborhood, the Jacobian determinant (12) (applied to  $G = NAM$ ) is bounded from above and from below by some positive constants. Hence, (12) yields

$$r^{\dim NAM} \ll m_{NAM}(D_r^{NAM}) \ll r^{\dim NAM}.$$

This completes the proof. ■

We pick  $\lambda > 0$  such that  $r_0 \lambda$  is an injectivity radius of  $\mathcal{X}_{\leq s_3}$  and use it throughout as radius parameter for Bowen balls. Recall the set  $D_\kappa$  and the choice of  $\kappa$  from (6)–(7). We may choose  $\lambda$  so small such that  $B_\lambda^G \subseteq D_\kappa$ .

In Lemma 7.4 below we will estimate how many Bowen  $L$ -balls are needed to cover  $P \in \eta_0^{L-1}$ ,  $P \subseteq \mathcal{X}_{\leq s_3}$ , for certain partitions  $\eta$  of  $\mathcal{X}$ .

**LEMMA 7.3:** *Let  $s > s_3$ . Then there exists  $k_{\max} \in \mathbb{N}$  such that whenever  $x \in \mathcal{X}_{> s}$  satisfies  $Tx, \dots, T^k x \in \mathcal{X}_{\leq s} \cap \mathcal{X}_{> s_3}$ , then  $k \leq k_{\max}$ .*

*Proof.* Let  $x \in \mathcal{X}_{>s}$  be as in the statement of the lemma. Lemma 6.1 and Proposition 5.5 show that there is a unique  $\xi \in \Xi$  and some  $g \in G$  such that  $\Gamma g = x$  and

$$\text{ht}(xa_t) = \left( \frac{\|v_\xi \varrho(ga_t)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{4}}$$

for  $t \in [1, r_0^k]$ . We suppose first that  $g = \xi na_r m \sigma$  for some  $n \in N, r > 0$  and  $m \in M$ . Then  $\text{ht}(xa_t) = \frac{r}{t}$  for  $t \in [1, r_0^k]$ . Therefore  $s \geq \text{ht}(Tx) = \frac{r}{r_0}$ . This and  $\text{ht}(T^k x) = \frac{r}{r_0^k} > s_3$  yield

$$k < \frac{\log \frac{r}{s_3}}{\log r_0} \leq \frac{\log \frac{sr_0}{s_3}}{\log r_0}.$$

Now we suppose that  $g = \xi na_r m \sigma(1, Z, X) \sigma$  for some  $n, (1, Z, X) \in N, r > 0$  and  $m \in M$ . For  $t \in [1, r_0^k]$  we have

$$\text{ht}(xa_t) = r \cdot \frac{t^{-1}}{(t^{-1} + \frac{1}{4}|X|^2) + |Z|^2}.$$

Then  $\text{ht}(xa_t) > s_3$  is equivalent to

$$(13) \quad 0 > (t^{-1} - \lambda_-)(t^{-1} - \lambda_+)$$

where

$$\lambda_\pm = -\frac{1}{2} \left( \frac{1}{2}|X|^2 - \frac{r}{s_3} \right) \pm \sqrt{\frac{1}{4} \left( \frac{1}{2}|X|^2 - \frac{r}{s_3} \right)^2 - \left( \frac{1}{16}|X|^4 + |Z|^2 \right)}.$$

Since  $\text{ht}(x) > s_3$ , (13) is satisfied at least for  $t = 1$ . Therefore, the roots  $\lambda_\pm$  are real and

$$\lambda_+ > 1 > \lambda_-.$$

From  $\lambda_+ > 1$  it follows that

$$\frac{1}{2} \left( \frac{r}{s_3} - \frac{1}{2}|X|^2 \right) > 0.$$

In turn,  $\lambda_- > 0$ . Now  $\text{ht}(T^k x) > s_3$  implies

$$r_0^k < \lambda_-^{-1} = \frac{\frac{1}{2} \left( \frac{r}{s_3} - \frac{1}{2}|X|^2 \right) + \sqrt{\frac{1}{4} \left( \frac{r}{s_3} - \frac{1}{2}|X|^2 \right)^2 - \left( \frac{1}{16}|X|^4 + |Z|^2 \right)}}{\frac{1}{16}|X|^4 + |Z|^2}.$$

From  $s \geq \text{ht}(Tx)$  it follows that

$$r \leq r_0 s \left[ \left( r_0^{-1} + \frac{1}{4}|X|^2 \right)^2 + |Z|^2 \right].$$

Therefore

$$\begin{aligned} \lambda_{s'}^{-1} &\leq \frac{r_0 s}{s_3} \cdot \frac{(r_0^{-1} + \frac{1}{4}|X|^2)^2 + |Z|^2}{\frac{1}{16}|X|^4 + |Z|^2} \\ &= \frac{s}{s_3} \cdot \frac{r_0^{-1} + \frac{1}{2}|X|^2}{\frac{1}{16}|X|^4 + |Z|^2} + \frac{r_0 s}{s_3}. \end{aligned}$$

From  $\text{ht}(x) > \text{ht}(Tx)$ , a straightforward deduction yields

$$\frac{1}{16}|X|^4 + |Z|^2 > r_0^{-1}.$$

Hence,

$$\frac{r_0^{-1} + \frac{1}{2}|X|^2}{\frac{1}{16}|X|^4 + |Z|^2}$$

is bounded from above (independent of  $x$ ), and so is  $\lambda_{s'}^{-1}$ . This completes the proof. ■

In the following, for  $s' > s > s_3$  we define various numbers which vary with  $s$  and  $s'$ .

Let  $\ell$  denote the maximal number of  $T$ -steps between  $\mathcal{X}_{>s}$  and  $\mathcal{X}_{\leq s_3}$ , that is,

$$(14) \quad \ell := \max\{k \in \mathbb{N} \mid \exists x \in \mathcal{X}_{>s} : Tx, \dots, T^k x \in \mathcal{X}_{>s_3} \cap \mathcal{X}_{\leq s}, T^{k+1} x \in \mathcal{X}_{\leq s_3}\}.$$

We note that this maximum exists by Lemma 7.3. It equals the maximal number of  $T$ -steps between  $\mathcal{X}_{\leq s_3}$  and  $\mathcal{X}_{>s}$  in the sense that

$$\ell := \max\{k \in \mathbb{N} \mid \exists x \in \mathcal{X}_{\leq s_3} : Tx, \dots, T^k x \in \mathcal{X}_{>s_3} \cap \mathcal{X}_{\leq s}, T^{k+1} x \in \mathcal{X}_{>s}\},$$

which follows since all sets of the form  $\mathcal{X}_{\leq t}$  or  $\mathcal{X}_{>t}$  are invariant under  $\sigma$  and since  $\sigma \tilde{a} \sigma = \tilde{a}^{-1}$ . We note that the maximal amount of time a trajectory can spend continuously within  $\mathcal{X}_{>s_3} \cap \mathcal{X}_{\leq s}$  is then bounded by  $2\ell + 5$  (corresponding to a trajectory that reaches about height  $s$  and then returns to  $\mathcal{X}_{\leq s_3}$ ), i.e., that

$$\max\{k \in \mathbb{N} \mid \exists x : x, Tx, \dots, T^k x \in \mathcal{X}_{>s_3} \cap \mathcal{X}_{\leq s}\} \leq 2\ell + 5.$$

We define  $\ell'$  in the same way using  $s'$  in place of  $s$ .

Let  $s > s_3$  and  $L \in \mathbb{N}$ . Let  $\eta$  be a finite partition of  $\mathcal{X}$  of the form

$$\eta = \{\mathcal{X}_{>s}, \mathcal{X}_{>s_3} \cap \mathcal{X}_{\leq s}, P_1, \dots, P_r\}$$

with  $P_i \subseteq \mathcal{X}_{\leq s_3}$  for  $i = 1, \dots, r$ . For any  $P \in \eta_0^{L-1}$  we define

$$(15) \quad V_P := \{j \in \{0, \dots, L-1\} \mid T^j P \subseteq \mathcal{X}_{>s}\}.$$

For brevity we use the notation

$$[m, n) := \{m, m + 1, \dots, n - 1\}$$

for an interval of integer points with endpoints  $m \leq n \in \mathbb{N}$ .

An interval  $k + [0, K) \subseteq [0, L)$  of a trajectory of a set  $P \in \eta_0^{L-1}$  is said to be an **excursion into**  $\mathcal{X}_{>s}$  (of length  $K$ ) if

$$T^{k-1}P \subseteq \mathcal{X}_{\leq s}, \quad T^k P, \dots, T^{k+K-1}P \subseteq \mathcal{X}_{>s},$$

and either  $T^{k+K}P \subseteq \mathcal{X}_{\leq s}$  or  $k + K = L$ .

Clearly,  $V_P$  is a disjoint union of intervals which are excursions into  $\mathcal{X}_{>s}$ .

For the statement of the following lemma we remark that  $\mathcal{X}_{\leq s}$  is compact by [Dan84, p. 27]. Further, we recall that  $\lambda$  is the parameter used in the definition of Bowen balls, and that  $B_\lambda^G \subseteq D_\kappa$ .

LEMMA 7.4: *Let  $s' > s > s_3$  and define  $\ell, \ell'$  as above. Let  $\lambda' \in (0, \lambda]$  be such that  $r_0\lambda'$  is an injectivity radius of  $\mathcal{X}_{\leq s'}$ . Suppose that*

$$\eta = \{\mathcal{X}_{>s}, \mathcal{X}_{>s_3} \cap \mathcal{X}_{\leq s}, P_1, \dots, P_r\}$$

*is a finite partition of  $\mathcal{X}$  such that  $\text{diam } T^j(P_i) \leq \lambda'$  for each  $j = 0, \dots, 2\ell' + 5$  and  $i = 1, \dots, r$ . Then for each  $L \in \mathbb{N}$  and  $P \in \eta_0^{L-1}$  with  $P \subseteq \mathcal{X}_{\leq s_3}$  the set  $P$  can be covered by*

$$c^m e^{h_m(T)\ell m} e^{\frac{1}{2}h_m(T)|V_P|}$$

*Bowen  $L$ -balls. Here the constant  $c$  only depends on  $G, r_0, s_1, s_2$  and  $\lambda$ . The constant  $m$  (not to be confused with the Haar measure  $m$ ) is the number of excursions of  $P$  into  $\mathcal{X}_{>s'}$ .*

We note that while the partition  $\eta$  is (in a strong way) adapted to the heights  $s, s'$ , our definition of Bowen  $L$ -ball does not depend on  $s, s'$ .

*Proof.* Let  $P \in \eta_0^{L-1}$  with  $P \subseteq \mathcal{X}_{\leq s_3}$ . We decompose  $V_P$  into a disjoint union of excursions into  $\mathcal{X}_{>s}$ . We denote those intervals that contain excursions into  $\mathcal{X}_{>s'}$  by  $V_j = [k_j, k_j + K_j)$  and their union by

$$(16) \quad V = \bigcup_{j=1}^m V_j = [k_1, k_1 + K_1) \cup \dots \cup [k_m, k_m + K_m) \subset V_P.$$

We may suppose that  $k_1 < k_2 < \dots < k_m$ . We note that each of those excursion  $V_j$  is contained in an excursion  $\tilde{V}_j = [n_j, n_j + h_j) \subseteq [k_j - \ell, k_j + K_j + \ell)$

into  $\mathcal{X}_{>s_3}$ . We define  $\tilde{V} = \bigcup_{j=1}^m \tilde{V}_j$  (which is a disjoint union). Analogously, we decompose  $W := [0, L) \setminus \tilde{V}$  into a disjoint union

$$W = \bigcup_{j=1}^{m+1} W_j$$

where each  $W_j$  is a maximal subset of  $W$  of the form  $[l_j, l_j + L_j)$  with  $0 = l_1 < l_2 < \dots < l_{m+1}$ . The set  $W_{m+1}$  might be empty. It follows that  $[0, L)$  is the disjoint union of  $W_1, \tilde{V}_1, \dots, \tilde{V}_m, W_{m+1}$  in that order.

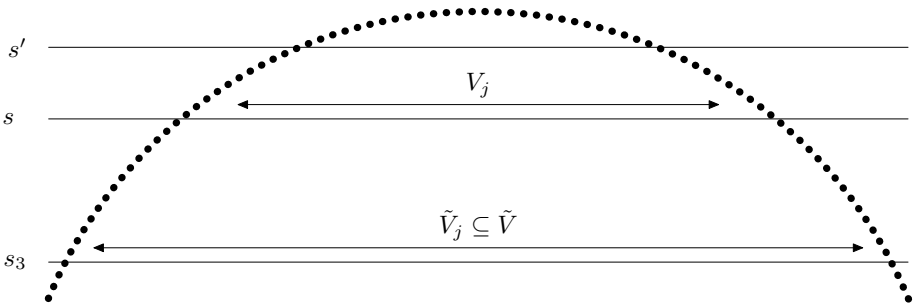


Figure 1. An excursion  $V_j$  into  $\mathcal{X}_{>s}$  containing an excursion into  $\mathcal{X}_{>s'}$  and contained in the excursion  $\tilde{V}_j$  into  $\mathcal{X}_{>s_3}$ .

For  $R$  running iteratively through  $|W_1|, |W_1 \cup \tilde{V}_1|, \dots, |W_1 \cup \tilde{V}_1 \cup \dots \cup W_m \cup \tilde{V}_m|, |W_1 \cup \tilde{V}_1 \cup \dots \cup W_m \cup \tilde{V}_m \cup W_{m+1}|$ , we now show that the number of  $R$ -boxes with center in  $P$  needed to cover  $P$  is bounded by

$$c^{j-1} e^{\frac{1}{2} h_m(T) |\tilde{V}_1 \cup \dots \cup \tilde{V}_{j-1}|}$$

for  $R \in \{|W_1 \cup \tilde{V}_1 \cup \dots \cup W_j|, |W_1 \cup \tilde{V}_1 \cup \dots \cup W_{j-1} \cup \tilde{V}_{j-1}|\}$ ,  $j = 1, \dots, m + 1$ , where  $c$  is a constant depending only on  $G, \tilde{a}, s_1, s_2$  and  $\lambda$ , but not on  $s$  or  $s'$ . We determine  $c$  in (21) below.

THE STEP CORRESPONDING TO ADDING  $W_j$ : Note first that even though an interval  $W_j = [l_j, l_j + L_j)$  may contain one or more excursions into  $\mathcal{X}_{>s}$ , it does not contain an excursion into  $\mathcal{X}_{>s'}$ . Each of the excursions into  $\mathcal{X}_{>s}$  is of length at most  $2\ell' + 5$ .

Suppose now that we have already found (at most)

$$(17) \quad c^{j-1} e^{\frac{1}{2} h_m(T) |\tilde{V}_1 \cup \dots \cup \tilde{V}_{j-1}|}$$

Bowen  $l_j$ -balls with center in  $P$  whose union contains the given element  $P \in \eta_0^{L'-1}$ . If  $j = 1$ , and hence  $l_j = 0$ , we may define a Bowen 0-ball as a Bowen 1-ball and the claim on the covering number is trivial. We claim that  $T^{l_j}(P), \dots, T^{l_j+L_j-1}(P)$  being contained in  $\mathcal{X}_{\leq s'}$  implies that  $T^{l_j}P$  is so small such that if  $P$  is covered by the Bowen  $l_j$ -balls with centers  $x_1, \dots, x_r \in P$ , then  $P$  is already covered by the smaller Bowen  $l_j + L_j$ -balls with these centers.

The choice of  $\lambda', \ell'$  and  $\eta$  implies that any element  $Q \in \eta$  with  $Q \subseteq \mathcal{X}_{\leq s_3}$  is contained in

$$(18) \quad w \bigcap_{i=0}^k \tilde{a}^i B_{\lambda'}^G \tilde{a}^{-i}$$

for  $k \leq 2\ell' + 5$  and for any  $w \in Q$  as long as  $Q, T(Q), \dots, T^k(Q) \subseteq \mathcal{X}_{\leq s'}$ . A simple induction shows that  $Q \in \eta_0^{L'-1}$ ,  $Q \subseteq \mathcal{X}_{\leq s_3}$ , belongs to the ‘‘small’’ Bowen  $L'$ -ball (18) (defined with the radius  $\lambda'$ ) with  $k = L' - 1$  if it is known that  $Q, T(Q), \dots, T^{L'-1}(Q) \subseteq \mathcal{X}_{\leq s'}$ . This applies to the partition element  $Q$  containing  $T^{l_j}(P)$  and  $L' = L_j$ , and shows that  $T^{l_j}(P)$  is contained in

$$(19) \quad w \bigcap_{i=0}^{L_j-1} \tilde{a}^i B_{\lambda'}^G \tilde{a}^{-i}$$

for any  $w \in T^{l_j}(P)$ . Let  $x B_{l_j}$  be one of the Bowen balls used in the cover associated to (17). Set  $w := x \tilde{a}^{l_j}$ . From  $x \tilde{a}^{l_j} \in \mathcal{X}_{\leq s_3}$  and the choice of  $\lambda' \leq \lambda$  it follows that

$$\begin{aligned} P \cap x B_{l_j} &\subseteq \left( (x \tilde{a}^{l_j}) \bigcap_{i=0}^{L_j-1} \tilde{a}^i B_{\lambda'}^G \tilde{a}^{-i} \cap (x \tilde{a}^{l_j}) \tilde{a}^{-l_j} B_{l_j} \tilde{a}^{l_j} \right) \tilde{a}^{-l_j} \\ &\subseteq (x \tilde{a}^{l_j}) \left( \bigcap_{i=0}^{L_j-1} \tilde{a}^i B_{\lambda'}^G \tilde{a}^{-i} \cap \tilde{a}^{-l_j} B_{l_j} \tilde{a}^{l_j} \right) \tilde{a}^{-l_j} \subseteq x B_{l_j+L_j}. \end{aligned}$$

Note that  $n_j = l_j + L_j$ , which is the starting point of  $\tilde{V}_j$  for  $j \leq m$ . Thus,  $P$  can be covered with the same number of Bowen  $n_j$ -balls with center in  $P$  as with Bowen  $l_j$ -balls, so that (17) is still an upper bound for the necessary number.

THE STEP CORRESPONDING TO ADDING  $\tilde{V}_j$ : Suppose now that we have already found (17)-many Bowen  $n_j$ -balls with center in  $P$  whose union covers  $P$ . Let  $x B_{n_j}$  be any Bowen  $n_j$ -ball used in this covering. Define

$$E := \{y \in x B_{n_j} \mid T^{n_j-1}y \in \mathcal{X}_{\leq s_3}, T^{n_j}y, \dots, T^{n_j+h_j-1}y \in \mathcal{X}_{> s_3}\}.$$

Then  $P \cap xB_{n_j} \subseteq E$ . If  $y \in E$ , then there exists  $g \in B_{n_j}$  with  $y = xg$ . This implies  $h := \tilde{a}^{-n_j+1}g\tilde{a}^{n_j-1} = \sigma(1, Z, X)\sigma g' \in B_\lambda^G$  for some bounded (with the bound only depending on  $G$  and  $\lambda$ )  $g' \in NAM$  and  $\sigma(1, Z, X)\sigma \in U$ . Set  $x' := T^{n_j-1}x$  and  $y' := T^{n_j-1}y$  so that  $y' = x'h$ . Moreover, we have

$$\begin{aligned} Tx', \dots, T^{h_j}x' &\in \mathcal{X}_{>s_3}, & x' &\in \mathcal{X}_{\leq s_3} \cap \mathcal{X}_{>s_2}, \\ Ty', \dots, T^{h_j}y' &\in \mathcal{X}_{>s_3}, & y' &\in \mathcal{X}_{\leq s_3} \cap \mathcal{X}_{>s_2}. \end{aligned}$$

Applying Proposition 6.3 to  $x', y'$  shows that

$$(20) \quad |Z| \leq c_4 r_0^{-h_j/2} \quad \text{and} \quad |X| \leq c_3 r_0^{-h_j/4}.$$

We claim that this bound implies that we can cover  $P \cap xB_{n_j}$  with not more than  $c r_0^{(\frac{p_1}{4} + \frac{p_2}{2})h_j}$  Bowen  $l_{j+1}$ -boxes with centers in  $P$  (recall that  $l_{j+1} = n_j + h_j$ ). Here  $c$  is some constant that does not depend on  $s, s'$ .

To see that, notice first that (20) implies that  $\tilde{a}^{-\lceil h_j/2 \rceil} h \tilde{a}^{\lceil h_j/2 \rceil}$  is still bounded uniformly in  $y$  and  $j$ . More precisely, we may accommodate the constants  $c_3, c_4$  by finding some absolute integer  $b > 0$  (depending on  $r_0, c_3, c_4$  and  $\lambda$ ) such that (20) implies that  $\tilde{a}^{-\lceil h_j/2 \rceil + b} \sigma(1, Z, X) \sigma \tilde{a}^{\lceil h_j/2 \rceil - b} \in B_\lambda^G$ . Note that we also have

$$g', \tilde{a}^{-\lceil h_j/2 \rceil + b} g' \tilde{a}^{\lceil h_j/2 \rceil - b} \in B_{c'\lambda}^G$$

for some constant  $c' > 0$  that only depends on the choice of the Riemannian metric on  $G$  and  $b$ . Together we see that  $g \in B_{n_j}$  actually belongs to the Bowen  $(n_j + \lceil h_j/2 \rceil - b)$ -ball  $B'$  defined by the radius  $(1 + c')\lambda$ . Define  $B''$  to be the Bowen  $l_{j+1}$ -ball defined by the radius  $\lambda/2$ . Let now  $y_1, \dots, y_q \in P \cap xB'$  be a maximal collection of points for which the sets  $y_1 B'', \dots, y_q B''$  are pairwise disjoint. By definition we have  $y_1 B'' \cup \dots \cup y_q B'' \subseteq xB' B''$ . Also note that  $B' B'' \subseteq B^{(3)}$ , where  $B^{(3)}$  is the Bowen  $(n_j + \lceil h_j/2 \rceil - b)$ -ball defined by the radius  $(3/2 + c')\lambda$ . With the lower and upper bounds for the Haar measure of a Bowen  $n$ -ball of the form  $ce^{-h_m(T)n}$  from Lemma 7.2 for two values of  $c$  (which depend on the radius used) we obtain

$$(21) \quad q \leq \frac{m(B^{(3)})}{m(B'')} \leq c'' e^{-h_m(T)(n_j + \lceil h_j/2 \rceil - b) + h_m(T)l_{j+1}}.$$

Finally note that by maximality of the collection  $y_1, \dots, y_q$  it follows that  $P \cap xB$  is covered by  $y_1 B_{l_{j+1}}, \dots, y_q B_{l_{j+1}}$ , which gives the claim (by recalling that  $l_{j+1} = n_j + h_j$ ).

Finally note that the bound (17) for  $j = m$  (if  $W_{m+1} = \emptyset$ ) or for  $j = m + 1$  gives the conclusion of the proposition since  $|\tilde{V}_j| \leq |V_j| + 2\ell$ . ■

PROPOSITION 7.5: *For all  $s > s_3$  there exists a finite partition*

$$\eta = \{\mathcal{X}_{>s}, \mathcal{X}_{>s_3} \cap \mathcal{X}_{\leq s}, P_1, \dots, P_r\}$$

of  $\mathcal{X}$  such that for each  $T$ -invariant probability measure  $\mu$  on  $\mathcal{X}$  we have

$$h_\mu(T) \leq h_\mu(T, \eta) + \frac{1}{s} + \frac{1}{2}h_m(T)(1 - \mu(\mathcal{X}_{\leq s})).$$

*Proof.* Using the ergodic decomposition of  $T$ -invariant measures we may restrict to ergodic  $T$ -invariant measures. Also note that every  $T$ -orbit visits  $\mathcal{X}_{\leq s_3}$  (Lemma 6.1(v), which also holds for  $T^{-1}$  in place of  $T$ ), so that we have

$$\delta_1 := \mu(\mathcal{X}_{\leq s_3}) > 0.$$

Let  $s' > s$ . Define  $\ell$  as in (14) and let

$$\ell'' := \min\{k \in \mathbb{N} \mid \exists x \in \mathcal{X}_{\leq s_3} : T^{k+1}x \in \mathcal{X}_{>s'}\}.$$

Let  $\eta = \{\mathcal{X}_{>s}, \mathcal{X}_{>s_3} \cap \mathcal{X}_{\leq s}, P_1, \dots, P_r\}$  be as in Lemma 7.4.

We will show the proposition using [BK83], more precisely in the form of Lemma B.2 in [ELMV12]: There it is shown that for any small  $\delta > 0$ , the entropy of  $\mu$  is the limit as  $\lambda \rightarrow 0$  of

$$\liminf_{L \rightarrow \infty} \frac{\log N_\lambda(\delta, L)}{L},$$

where  $N_\lambda(\delta, L)$  is the minimal number of Bowen  $L$ -balls, with  $\lambda$  being the parameter used in the definition, that are needed to cover some set of  $\mu$ -measure  $\delta$ . However, there is one important difference between our definition of Bowen  $L$ -balls and that of [BK83]. In the latter, one takes the intersections of pre-images of  $\lambda$ -balls within  $\mathcal{X}$ . In our definition of Bowen  $L$ -ball we took the intersection in the group, which results in general in a smaller set (namely in those cases where the orbit ventures near the cusp). As we are seeking an upper bound of entropy, we may use [BK83] also together with our definition of Bowen  $L$ -balls. This has the advantage that we do not have to take the limit as  $\lambda \rightarrow 0$ . Within the group a bounded number of translates<sup>3</sup> of Bowen  $L$ -balls defined by  $\lambda > 0$

---

<sup>3</sup> Let  $\lambda < \lambda'$ . If  $g_1B_L(\lambda/2), \dots, g_bB_L(\lambda/2)$  is a maximal collection of disjoint left translates for some  $g_1, \dots, g_b \in B_L(\lambda')$ , then  $b \leq m(B_L(\lambda' + \lambda/2))/m(B_L(\lambda/2))$  is bounded independently (see Lemma 7.2) of  $L$  and  $B_L(\lambda') \subseteq g_1B_L(\lambda) \cup \dots \cup g_bB_L(\lambda)$ .

can be used to cover a Bowen  $L$ -ball defined by  $\lambda' > 0$ , and as  $L \rightarrow \infty$  this difference becomes unimportant.

By ergodicity  $\bigcup_{j=0}^{\infty} T^{-j}\mathcal{X}_{\leq s_3}$  has full measure. Hence there exists  $M$  with

$$\mu\left(\bigcup_{j=0}^{M-1} T^{-j}\mathcal{X}_{\leq s_3}\right) > 1 - \frac{\delta_1}{2}.$$

The intersection of the preimage of this set under  $T^{L'}$  with  $\mathcal{X}_{\leq s_3}$  has measure at least  $\delta_1/2$ . It follows that there are infinitely many  $L$  (of the form  $L' + j$  for some  $j \in [0, M)$ ) for which

$$\mu(\mathcal{X}_{\leq s_3} \cap T^{-L}\mathcal{X}_{\leq s_3}) > \frac{\delta_1}{2M}.$$

We now proceed making  $Y_L := \mathcal{X}_{\leq s_3} \cap T^{-L}\mathcal{X}_{\leq s_3}$  smaller, taking care that the resulting sets have measures that do not approach zero, and obtaining more information on the smaller sets.

Since  $\mu$  is ergodic, the value  $h_\mu(T, \eta)$  has the following interpretation: for every  $\varepsilon > 0$  and every sufficiently large  $L$  there exists a set  $Z_{L,\varepsilon}$  such that its measure is bigger than  $1 - \varepsilon$  and  $Z_{L,\varepsilon}$  can be covered with  $e^{(h_\mu(T,\eta)+\varepsilon)L}$  elements of  $\eta_0^{L-1}$ . We choose  $\varepsilon = \min(\frac{1}{3s}, \frac{\delta_1}{4M})$ , and take the intersection  $Y'_L = Y_L \cap Z_{L,\varepsilon}$ . We now know that  $\mu(Y'_L) > \frac{\delta_1}{4M}$  and that  $Y'_L$  can be covered with  $e^{(h_\mu(T,\eta)+\frac{1}{3s})L}$  elements of  $\eta_0^{L-1}$ .

Finally, we may make  $Y'_L$  again a bit smaller to ensure that the ergodic averages for the characteristic function  $\chi_{\mathcal{X}_{>s}}$  are correct up to an error of  $(3sh_m(T))^{-1}$  and for sufficiently large  $L$ . More precisely, there exists a subset  $Y''_L \subset Y'_L$  (obtained by intersecting  $Y'_L$  with a set of near full measure) with  $\mu(Y''_L) > \delta = \frac{\delta_1}{5M}$  and some  $L_0$  such that for all  $L > L_0$  and all  $x \in Y''_L$  we have

$$\left| \frac{1}{L} \sum_{i=0}^{L-1} \chi_{\mathcal{X}_{>s}}(T^i x) - \mu(\mathcal{X}_{>s}) \right| < \frac{1}{3sh_m(T)}.$$

In the following we assume  $L > L_0$ . Now apply Lemma 7.4 to each of the partition elements of  $\eta_0^{L-1}$  obtained earlier. Notice that for each of the partition elements we have  $m \leq \frac{L}{v^n}$ . Finally, notice that the above ergodic sum is constant on each  $P \in \eta_0^{L-1}$ . For those  $P$  that intersect  $Y''_L$  we then have

$$|V_P| < (\mu(\mathcal{X}_{>s}) + \frac{1}{3sh_m(T)})L.$$

Thus, Lemma 7.4 together with the above covering estimate on  $Y'_L$  implies that  $Y''_L$  can be covered with  $N_L$  Bowen  $L$ -balls, where

$$N_L \leq e^{(h_\mu(T,\eta) + \frac{1}{3s})L} (ce^{h_m(T)\ell})^{\frac{L}{\ell''}} e^{\frac{1}{2}h_m(T)\mu(\mathcal{X}_{>s})L + \frac{1}{6s}L}.$$

Choose  $s'$  so big such that  $(\log c + h_m(T)\ell)/\ell'' < \frac{1}{6s}$ . This implies the proposition. ■

We now restate and prove Theorem A from the introduction.

**THEOREM 7.6:** *Let  $(\mu_j)$  be a sequence of  $T$ -invariant probability measures on  $\mathcal{X}$  which converges to the measure  $\nu$ . Then*

$$\nu(\mathcal{X})h_{\frac{\nu}{\nu(\mathcal{X})}}(T) + \frac{1}{2}h_m(T)(1 - \nu(\mathcal{X})) \geq \limsup_{j \rightarrow \infty} h_{\mu_j}(T),$$

where it does not matter how we interpret  $h_{\frac{\nu}{\nu(\mathcal{X})}}(T)$  if  $\nu(\mathcal{X}) = 0$ .

*Proof.* Pick  $s > s_3$  such that  $\nu(\partial\mathcal{X}_{\leq s}) = 0$  (this holds for all but countably many  $s$ ). Let  $\eta = \{\mathcal{X}_{>s}, \mathcal{X}_{>s_3} \cap \mathcal{X}_{\leq s}, P_1, \dots, P_r\}$  be a partition of  $\mathcal{X}$  as in Proposition 7.5 such that  $\nu(\partial P_j) = 0$  for  $j = 1, \dots, r$ . Let  $\varepsilon > 0$ . Suppose now that  $\nu(\mathcal{X}) > 0$ . By definition of entropy we may fix  $m \in \mathbb{N}$  such that

$$h_{\frac{\nu}{\nu(\mathcal{X})}}(T) + \varepsilon > \frac{1}{m}H_{\frac{\nu}{\nu(\mathcal{X})}}(\eta_0^{m-1})$$

and

$$\frac{2e^{-1}}{m} < \frac{\varepsilon}{2} \quad \text{and} \quad -\frac{1}{m} \log \nu(\mathcal{X}) < \varepsilon.$$

Then using (9) we get

$$\nu(\mathcal{X})h_{\frac{\nu}{\nu(\mathcal{X})}}(T) + 2\varepsilon > -\frac{1}{m} \sum_{P \in \eta_0^{m-1}} \nu(P) \log \nu(P).$$

Note that this holds trivially if  $\nu(\mathcal{X}) = 0$ .

Let

$$Q := \bigcap_{k=0}^{m-1} T^{-k}\mathcal{X}_{>s}.$$

Since

$$\sum_{P \in \eta_0^{m-1} \setminus \{Q\}} \mu_j(P) \log \mu_j(P) \xrightarrow{j \rightarrow \infty} \sum_{P \in \eta_0^{m-1} \setminus \{Q\}} \nu(P) \log \nu(P),$$

we find  $j_0 \in \mathbb{N}$  such that for all  $j \geq j_0$  we have

$$\begin{aligned} & \left| -\frac{1}{m} \sum_{P \in \eta_0^{m-1}} \nu(P) \log \nu(P) - \frac{1}{m} H_{\mu_j}(\eta_0^{m-1}) \right| \\ & \leq \frac{1}{m} \left| \sum_{P \in \eta_0^{m-1} \setminus \{Q\}} (\mu_j(P) \log \mu_j(P) - \nu(P) \log \nu(P)) \right| \\ & \quad + \frac{1}{m} |\mu_j(Q) \log \mu_j(Q) - \nu(Q) \log \nu(Q)| \\ & \leq \frac{\varepsilon}{2} + \frac{2e^{-1}}{m} < \varepsilon. \end{aligned}$$

This and Proposition 7.5 yield

$$\begin{aligned} \nu(\mathcal{X}) h_{\frac{\nu}{\nu(\mathcal{X})}}(T) + 3\varepsilon &> \frac{1}{m} H_{\mu_j}(\eta_0^{m-1}) \geq h_{\mu_j}(T, \eta) \\ &> h_{\mu_j}(T) - \frac{1}{s} - \frac{1}{2} h_m(T) \cdot (1 - \mu_j(\mathcal{X}_{\leq s})). \end{aligned}$$

Hence

$$\nu(\mathcal{X}) h_{\frac{\nu}{\nu(\mathcal{X})}}(T) + \frac{1}{2} h_m(T) \cdot (1 - \nu(\mathcal{X}_{\leq s})) + 3\varepsilon + \frac{1}{s} \geq \limsup_{j \rightarrow \infty} h_{\mu_j}(T).$$

Letting  $\varepsilon$  tend to 0 and  $s$  tend to infinity, it follows that

$$\nu(\mathcal{X}) h_{\frac{\nu}{\nu(\mathcal{X})}}(T) + \frac{1}{2} h_m(T) \cdot (1 - \nu(\mathcal{X})) \geq \limsup_{j \rightarrow \infty} h_{\mu_j}(T). \quad \blacksquare$$

As an immediate consequence of Proposition 7.1 and Theorem 7.6 we obtain the corollaries stated in the introduction.

**COROLLARY 7.7:** *Let  $(\mu_j)_{j \in \mathbb{N}}$  be a sequence of  $T$ -invariant probability measures on  $\mathcal{X}$  such that  $\liminf_{j \rightarrow \infty} h_{\mu_j}(T) \geq c$ . Let  $\nu$  be any weak\* limit point of  $(\mu_j)$ . Then*

$$\nu(\mathcal{X}) \geq \frac{2c}{h_m(T)} - 1.$$

Moreover, if

$$\nu(\mathcal{X}) = \frac{2c}{h_m(T)} - 1 > 0,$$

then  $h_{\frac{\nu}{\nu(\mathcal{X})}}(T) = h_m(T)$  and  $\frac{\nu}{\nu(\mathcal{X})}$  is the Haar measure on  $\mathcal{X}$ .

**8. Hausdorff dimension of orbits missing a fixed open subset**

In this section we prove Theorem B from the introduction, which is an application of Theorem 7.6 and the methods for its proof and answers a question by Barak Weiss about the Hausdorff dimension of the set of all orbits which miss a fixed open subset of  $\mathcal{X}$ . We note that for a compact quotient this is a simple corollary of semi-continuity of entropy and uniqueness of the measure of maximal entropy. In the presence of cusps, the methods of this paper become relevant. We also note that related results have been obtained by Shi [Shi12] but to our knowledge these do not provide the following results as corollaries.

Let  $\mathcal{O} \subseteq \mathcal{X}$  be a non-empty open subset. Let  $\mathcal{E}$  denote the set of points in  $\mathcal{X}$  whose forward- $A$ -orbits do not intersect  $\mathcal{O}$ , that is

$$\mathcal{E} := \{x \in \mathcal{X} \mid \forall t \geq 0: xa_t \notin \mathcal{O}\}.$$

In the following we will show that  $\mathcal{E}$  cannot have full Hausdorff dimension as claimed in Theorem B of the introduction. Instead of Theorem B we will prove a (stronger) discretized version. To that end we now define  $T$  using  $\tilde{a} = a_e$  ( $e = \exp(1)$ ), so that

$$T: \mathcal{X} \rightarrow \mathcal{X}, \quad x \mapsto xa_e,$$

denotes the time-one (discrete) geodesic flow. Note that then the maximal entropy of  $T$  is

$$h_m(T) = \frac{p_1}{2} + p_2.$$

We consider the set

$$\mathcal{E}' := \{x \in \mathcal{X} \mid \forall n \in \mathbb{N}_0: T^n x \notin \mathcal{O}\}.$$

Then Theorem B is implied by the following

**THEOREM 8.1:** *The Hausdorff dimension of  $\mathcal{E}'$  satisfies*

$$\dim_H \mathcal{E}' < \dim \mathcal{X} = \dim G.$$

For convenience we recall the following definitions, adapted to our current set-up. For  $r > 0$ , the open ball in  $G$  centered at the identity of  $G$  with radius  $r$  is denoted by  $B_r^G$ . For  $L \in \mathbb{N}$ , the Bowen  $L$ -ball in  $G$  with radius parameter  $r$  is

$$B_L = B_L(r) = \bigcap_{j=0}^{L-1} a_e^j B_r^G a_e^{-j}.$$

Finally, for each  $x \in \mathcal{X}$ , the Bowen  $L$ -ball in  $\mathcal{X}$  with center  $x$  is

$$xB_L = xB_L(r).$$

*Strategy for the proof of Theorem 8.1.* We cover  $\mathcal{E}'$  by countably many (small) bounded open sets, say by  $A(n), n \in \mathbb{N}$ , and estimate the Hausdorff dimension of each of the sets

$$\mathcal{W}_n := \mathcal{E}' \cap A(n).$$

By countable stability of Hausdorff dimension we have

$$\dim_H \mathcal{E}' = \sup\{\dim_H \mathcal{W}_n \mid n \in \mathbb{N}\}.$$

Thus, we have to show that

$$\dim_H \mathcal{W}_n < \dim \mathcal{X} - \varepsilon_0 = \dim G - \varepsilon_0$$

for some  $\varepsilon_0 > 0$  not depending on  $n \in \mathbb{N}$ . To seek a contradiction we assume that for (some small, to be determined below)  $\varepsilon_0 > 0$  we find a set  $\mathcal{W} = \mathcal{W}(\varepsilon_0)$  among the sets  $\mathcal{W}_n$  such that

$$(22) \quad d := \dim_H \mathcal{W} \geq \dim \mathcal{X} - \varepsilon_0.$$

Frostman’s Lemma assures the existence of a probability measure  $\mu$  on  $\mathcal{W}$  such that

$$(23) \quad \mu(xB_r^G) \ll r^{d-\varepsilon_0} \leq r^{\dim G - 2\varepsilon_0}$$

for any  $x \in \mathcal{X}$  and any  $r \in (0, 1]$ , with an implied constant only depending on  $\mu$ . Then we will give an upper bound for the number of Bowen  $L$ -balls needed to cover  $\mathcal{W}$  as well as the  $\mu$ -mass of an Bowen  $L$ -ball. Bounding the  $\mu$ -mass of  $\mathcal{W}$  (which is 1) via these Bowen balls will result in a contradiction.

We start by choosing a good value for  $\varepsilon_0$ . Recall that  $\mathcal{M}_1(\mathcal{X})^T$  denotes the space of  $T$ -invariant probability measures on  $\mathcal{X}$  and that  $h_m(T)$  is the maximal entropy of  $T$  (see Proposition 7.1).

LEMMA 8.2: *There exists  $\delta_0 > 0$  such that  $h_\nu(T) \leq h_m(T) - \delta_0$  for all  $\nu \in \mathcal{M}_1(\mathcal{X})^T$  with  $\text{supp } \nu \subseteq \mathcal{E}'$ .*

*Proof.* To seek a contradiction assume that there exists a sequence  $(\nu_n)_{n \in \mathbb{N}}$  in  $\mathcal{M}_1(\mathcal{X})^T$  with  $\text{supp } \nu_n \subseteq \mathcal{E}'$  for each  $n \in \mathbb{N}$  such that

$$h_{\nu_n}(T) \geq h_m(T) - \frac{1}{n}.$$

Let  $\nu$  be any weak\* limit point of  $(\nu_n)$ . Then  $\text{supp } \nu \subseteq \mathcal{E}'$  as  $\mathcal{E}'$  is closed. Since  $\lim h_{\nu_n}(T) = h_m(T)$ , Corollary 7.7 yields  $\nu(\mathcal{X}) = 1$  and  $h_\nu(T) = h_m(T)$ . Thus,  $\nu$  is the Haar measure on  $\mathcal{X}$  and hence  $\text{supp } \nu = \mathcal{X}$ . This is a contradiction. ■

We fix  $\delta_0 \in (0, h_m(T)/4]$  with the properties as in Lemma 8.2 (the upper bound will be needed for Proposition 8.6 below), set

$$\varepsilon_0 := \frac{\delta_0}{20},$$

and assume the existence of  $\mathcal{W} \subset \mathcal{E}'$  satisfying (22). We also fix a probability measure  $\mu$  on  $\mathcal{W}$  satisfying (23).

We pick a weak\* limit point  $\nu$  of

$$\frac{1}{L} \sum_{j=0}^{L-1} T_*^j \mu \quad \text{as } L \rightarrow \infty.$$

Then  $\nu$  is  $T$ -invariant and  $\nu(\emptyset) = 0$ . We will see in Proposition 8.6 below that  $\nu(\mathcal{X}) > 0$ . Then Lemma 8.2 shows

$$h_{\frac{\nu}{\nu(\mathcal{X})}}(T) \leq h_m(T) - \delta_0,$$

which will allow us to establish nontrivial bounds on the number of Bowen  $L$ -balls needed to cover  $\mathcal{W}$ .

We start by deriving a lower bound for  $\nu(\mathcal{X})$ , where the following will be needed.

LEMMA 8.3: *For any sufficiently (depending only on  $G$ ) small  $r > 0$ , any  $x \in \mathcal{X}$  and any  $L \in \mathbb{N}$  we have*

$$\mu(\overline{x B_L(r)}) \leq c_\mu r^{\dim G - 2\varepsilon_0} e^{(-h_m(T) + 2\varepsilon_0)L},$$

where the constant  $c_\mu$  depends on the implied constant in (23) and on  $G$ .

*Proof.* Choose a maximal collection of elements  $g_1, \dots, g_q \in \overline{B_L(r)}$  for which the balls  $g_1 B_{r e^{-L}}, \dots, g_q B_{r e^{-L}}$  are pairwise disjoint. Note that  $B_{r e^{-L}} \subseteq B_L(r)$  and so  $g_j B_{r e^{-L}} \subseteq B_L(2r)$ . By Lemma 7.2 the Haar measure of  $B_L(r)$  is between bounded multiples (depending on  $G$ ) of  $r^{\dim G} e^{-h_m(T)L}$  and the Haar measure of  $B_{r e^{-L}}$  is between bounded multiples of  $r^{\dim G} e^{-L \dim G}$ . Hence it follows that  $q \leq c' e^{L(\dim G - h_m(T))}$  for some constant  $c'$  that only depends on  $G$ .

The maximality of  $q$  implies  $\overline{x B_L(r)} \subseteq x g_1 B_{2r e^{-L}} \cup \dots \cup x g_q B_{2r e^{-L}}$ . We now apply (23) and obtain

$$\begin{aligned} \mu(\overline{x B_L(r)}) &\leq q c'_\mu 2^{\dim G - 2\varepsilon_0} r^{\dim G - 2\varepsilon_0} e^{-L(\dim G - 2\varepsilon_0)} \\ &\leq c_\mu r^{\dim G - 2\varepsilon_0} e^{(\dim G - h_m(T) - \dim G + 2\varepsilon_0)L}, \end{aligned}$$

where  $c'_\mu$  is the implied constant in (23). ■

For a subset  $V \subseteq [0, L - 1]$  we set

$$Q_{s,V} := \{x \in \mathcal{X} \mid \forall j \in [0, L - 1]: (T^j x \in \mathcal{X}_{>s} \Leftrightarrow j \in V)\}.$$

LEMMA 8.4: *Let  $s > s_3$  and*

$$L \geq 2 \log\left(\frac{s}{s_3}\right) + 1.$$

*Then there are at most  $e^{f(s)L}$  subsets  $V \subseteq [0, L - 1]$  for which  $Q_{s,V}$  is nonempty, where*

$$f(s) := \frac{4 \log(2 \log(\frac{s}{s_3}) + 2)}{\log(\frac{s}{s_3})}.$$

*Proof.* Let  $V \subseteq [0, L - 1]$  and decompose  $V$  as in (16) (using  $s' := s$ ). We will show that  $Q_{s,V}$  being nonempty implies that there is a uniform nontrivial minimal distance between  $k_n + K_n$  and  $k_{n+1}$  since a trajectory going down from  $\mathcal{X}_{>s}$  cannot go back up before entering  $\mathcal{X}_{\leq s_3}$  (see Lemma 6.1(v)). The existence of this distance yields restrictions on those  $V$  for which  $Q_{s,V} \neq \emptyset$ .

At first, suppose that we have  $x \in \mathcal{X}$  with  $\text{ht}(x) \leq s_3$  and  $\text{ht}(T^j x) > s$  for some  $j \in \mathbb{N}$ . By Lemma 6.1(iii) we may suppose  $\text{ht}(x a_t) > 2s_1$  for all  $t \in [1, e^j]$ . We aim to prove a nontrivial lower bound on  $j$ . By Proposition 5.5 there exists a unique cusp representative  $\xi \in \Xi$  and an element  $g \in G$  with  $x = \Gamma g$  such that

$$\text{ht}(x a_t) = \text{ht}_\xi(x a_t) = \left( \frac{\|v_\xi \varrho(g a_t)\|}{\|v_\xi \varrho(\xi)\|} \right)^{-\frac{1}{q}}$$

for all  $t \in [1, e^j]$ . Proposition 5.3 implies  $g \in \xi N A M U$ , say  $g = \xi n a_r m u$  with  $u = \sigma(1, Z, X)\sigma$ . Lemma 5.2 shows

$$\text{ht}_\xi(x a_t) = r \cdot \frac{\frac{1}{t}}{\left(\frac{1}{t} + \frac{1}{4}|X|^2\right)^2 + |Z|^2}$$

for all  $t \in [1, e^j]$ . In particular,

$$s_3 > \text{ht}_\xi(x) = \frac{r}{\left(1 + \frac{1}{4}|X|^2\right)^2 + |Z|^2}$$

and

$$\text{ht}(xa_{e^j}) = r \cdot \frac{e^{-j}}{(e^{-j} + \frac{1}{4}|X|^2)^2 + |Z|^2} > s.$$

Therefore

$$\frac{(1 + \frac{1}{4}|X|^2)^2 + |Z|^2}{(e^{-j} + \frac{1}{4}|X|^2)^2 + |Z|^2} \cdot e^{-j} > \frac{s}{s_3}.$$

Together with the elementary estimate

$$e^{2j} \geq \frac{(1 + \frac{1}{4}|X|^2)^2 + |Z|^2}{(e^{-j} + \frac{1}{4}|X|^2)^2 + |Z|^2}$$

it follows that

$$j > \log\left(\frac{s}{s_3}\right).$$

Suppose now that we have  $x \in \mathcal{X}$  with  $\text{ht}(x) > s$  and  $\text{ht}(T^j x) \leq s_3$  for some  $j \in \mathbb{N}$ . Invoking Lemma 6.1(iv), we can deduce as before that

$$j > \log\left(\frac{s}{s_3}\right).$$

We set

$$j_0 := \left\lceil \log\left(\frac{s}{s_3}\right) \right\rceil.$$

Lemma 6.1(v) implies that

$$k_n + K_n + 2j_0 \leq k_{n+1}$$

for  $n = 1, \dots, m_1 - 1$ . Let

$$Q_0^L(s) := \bigvee_{j=0}^{L-1} T^{-j} \{ \mathcal{X}_{\leq s}, \mathcal{X}_{> s} \}.$$

If  $L = 2j_0 - 1$ , the definition of  $j_0$  yields that the cardinality of  $Q_0^L(s)$  is (note  $j_0 \geq 1$ )

$$\leq 1 + \binom{2j_0}{2} = \frac{(2j_0)^2}{2} + 1 - j_0 \leq (2j_0)^2.$$

For an arbitrary  $L$  the set  $[0, L - 1]$  is covered by the disjoint union

$$[0, L - 1] \subseteq \bigcup_{h=0}^{k_L} h \cdot (2j_0 - 1) + [0, 2j_0 - 2]$$

with

$$k_L := \left\lceil \frac{L}{2j_0 - 1} \right\rceil.$$

For each  $h \in \{0, \dots, k_L\}$ , the cardinality of

$$\{V \cap (h \cdot (2j_0 - 1) + [0, 2j_0 - 2]) \mid V \subseteq [0, L - 1], Q_{s,V} \neq \emptyset\}$$

is at most  $(2j_0)^2$ . Therefore, there are at most

$$(2j_0)^{2k_L}$$

subsets  $V \subseteq [0, L - 1]$  with  $Q_{s,V} \neq \emptyset$ . Hence the cardinality of  $Q_0^L(s)$  is bounded from above by

$$\exp(2k_L \cdot \log(2j_0)).$$

Using

$$k_L \leq \frac{L}{2j_0 - 1} + 1$$

and

$$\frac{\log(2 \lceil \log(\frac{s}{s_3}) \rceil)}{2 \lceil \log(\frac{s}{s_3}) \rceil - 1} \leq \frac{1}{\log(\frac{s}{s_3})} \log \left( 2 \log \left( \frac{s}{s_3} \right) + 2 \right),$$

the statement of the proposition follows easily. ■

For  $L \in \mathbb{N}$ , a subset  $V \subseteq [0, L - 1]$  and  $s > 0$  we set

$$Z_{s,L}(V) := \{x \in \mathcal{W} \cap \mathcal{X}_{\leq s} \mid \forall j \in [0, L - 1]: (T^j x \in \mathcal{X}_{> s} \Leftrightarrow j \in V)\}.$$

Lemma 8.4 immediately provides an upper bound on the number of nonempty sets  $Z_{s,L}(V)$ . By increasing  $s_3$  we may assume from now on that  $\mathcal{W} \subseteq \mathcal{X}_{\leq s_3}$ .

LEMMA 8.5: *Suppose that the parameter  $r$  in the definition of Bowen balls is an injectivity radius of  $\mathcal{X}_{\leq s_3}$ . Let  $s > s_3$ ,  $L$  as in Lemma 8.4 and  $V \subseteq [0, L - 1]$ . Then the set  $Z_{s,L}(V)$  can be covered with*

$$c_{r,\mathcal{W}} e^{c(s)L + h_m(T)(L - \frac{1}{2}|V|)}$$

Bowen  $L$ -balls, where  $c(s) \rightarrow 0$  as  $s \rightarrow \infty$  and the constant  $c_{r,\mathcal{W}}$  does not depend on  $s, L$  and  $V$ .

*Proof.* The proof is similar to that of Lemma 7.4 with  $s = s'$ . We can cover  $Z_{s,L}(V)$  with finitely many balls  $x B_r$  with  $x \in Z_{s,L}(V)$ , say

$$(24) \quad Z_{s,L}(V) \subseteq \bigcup_{j=1}^{c_{r,\mathcal{W}}} x_j B_r.$$

The necessary number  $c_{r,W}$  of such balls is bounded by a constant independent of  $s, L$  and  $V$ . Suppose now that  $x_0 B_r$  is one of the sets used in the covering (24) and consider

$$\mathcal{Z} := Z_{s,L}(V) \cap x_0 B_r.$$

If  $\mathcal{Z}$  is covered with say  $d_1$  Bowen  $\ell$ -balls with center in  $\mathcal{Z}$ , then, as in Lemma 7.4 “step corresponding to adding  $\tilde{V}_j$ ”, an excursion into  $\mathcal{X}_{>s}$  of length  $\ell_1$  starting at  $\ell + 1$  has the effect that  $\mathcal{Z}$  is covered by

$$cd_1 e^{\frac{1}{2}h_m(T)\ell_1}$$

Bowen  $(\ell + \ell_1)$ -balls. In contrast (as in Lemma 7.4 “step corresponding to adding  $W_j$ ”), a stay in  $\mathcal{X}_{\leq s}$  of length  $\ell_2$  only leads to a trivial estimate, that is  $\mathcal{Z}$  is covered by

$$cd_1 e^{h_m(T)\ell_2}$$

Bowen  $(\ell + \ell_2)$ -balls.

By iteratively applying these two steps, we see that  $\mathcal{Z}$  can be covered with

$$c^{2m} e^{h_m(T)(L - \frac{1}{2}|V|)}$$

Bowen  $L$ -balls, where  $c$  is a constant independent of  $s, L$  and  $V$ , and  $m$  is the number of excursions into  $\mathcal{X}_{>s}$ . The proof of Lemma 8.4 now shows that

$$m \leq \frac{2L}{\log(\frac{s}{s_3})} + 1.$$

This completes the proof. ■

PROPOSITION 8.6: *We have*

$$\nu(\mathcal{X}) \geq 1 - \frac{4\varepsilon_0}{h_m(T)} \geq 1 - \frac{1}{20}.$$

*Proof.* For  $L \in \mathbb{N}$  set

$$(25) \quad \mu_L := \frac{1}{L} \sum_{j=0}^{L-1} T_*^j \mu.$$

Then  $\mu_L$  converges along some subsequence to  $\nu$  in the weak\* topology. For any  $s > s_3$  we have

$$\mu_L(\mathcal{X}_{>s}) = \frac{1}{L} \sum_{j=0}^{L-1} \mu(T^{-j}\mathcal{X}_{>s}) = \frac{1}{L} \sum_{j=0}^{L-1} \mu(\mathcal{X}_{\leq s} \cap T^{-j}\mathcal{X}_{>s}),$$

since  $\mathcal{W} \subset \mathcal{X}_{\leq s_3}$ . For  $x \in \mathcal{X}$  we set

$$V_x := \{j \in [0, L - 1] \mid T^j x \in \mathcal{X}_{> s}\}.$$

In the following let  $\varrho > 0$  be a constant depending on  $s$ , to be fixed below. Then, it follows that

$$\begin{aligned} \frac{1}{L} \sum_{j=0}^{L-1} \mu(\mathcal{X}_{\leq s} \cap T^{-j} \mathcal{X}_{> s}) &= \frac{1}{L} \sum_{n=1}^L n \mu(\{x \in \mathcal{X}_{\leq s} \mid |V_x| = n\}) \\ &= \frac{1}{L} \sum_{n=1}^{\lceil \varrho L \rceil - 1} n \mu(\{x \in \mathcal{X}_{\leq s} \mid |V_x| = n\}) \\ &\quad + \frac{1}{L} \sum_{n=\lceil \varrho L \rceil}^L n \mu(\{x \in \mathcal{X}_{\leq s} \mid |V_x| = n\}) \\ &\leq \frac{1}{L} (\lceil \varrho L \rceil - 1) \mu(\mathcal{X}_{\leq s}) + \frac{1}{L} \cdot L \cdot \mu(\{x \in \mathcal{X}_{\leq s} \mid |V_x| \geq \varrho L\}) \\ &\leq \varrho + \mu(\{x \in \mathcal{X}_{\leq s} \mid |V_x| \geq \varrho L\}). \end{aligned}$$

Using Lemmas 8.4 and 8.5 for sufficiently large  $L$  allows us to estimate the latter term. Combining this with the estimate in Lemma 8.3 we get

$$\begin{aligned} \mu(\{x \in \mathcal{X}_{\leq s} \mid |V_x| \geq \varrho L\}) &\leq \mu\left(\bigcup_{|V| \geq \varrho L} Z_{s,L}(V)\right) \\ &\leq c_{r,\mathcal{W}} e^{f(s)L} e^{c(s)L + h_m(T)(L - \frac{1}{2}\varrho L)} \cdot c_\mu r^{\dim G - 2\varepsilon_0} e^{(-h_m(T) + 2\varepsilon_0)L} \\ &\leq c' e^{(\tilde{c}(s) + 2\varepsilon_0 - \frac{1}{2}h_m(T)\varrho)L}, \end{aligned}$$

where  $c'$  is a constant depending on  $r$  but not on  $s$  or  $L$ , and  $\tilde{c}(s) := f(s) + c(s)$  tends to 0 as  $s \rightarrow \infty$ . If we choose  $\varrho$  so that the exponent is negative, i.e. with

$$\varrho = \varrho(s) > \frac{2\tilde{c}(s) + 4\varepsilon_0}{h_m(T)},$$

then

$$\varepsilon(L) = \mu(\{x \in \mathcal{X}_{\leq s} \mid |V_x| \geq \varrho(s)L\}) \rightarrow 0 \quad \text{as } L \rightarrow \infty.$$

Therefore, for sufficiently large  $s$  and  $L$ , we have

$$\mu_L(\mathcal{X}_{> s}) \leq \varrho(s) + \varepsilon(L).$$

In turn,

$$\mu_L(\mathcal{X}_{\leq s}) \geq 1 - (\varrho(s) + \varepsilon(L)).$$

Letting  $L \rightarrow \infty$  along the subsequence that gives  $\nu$  as the limit we obtain

$$\nu(\mathcal{X}_{\leq s}) \geq 1 - \varrho(s),$$

and

$$\nu(\mathcal{X}) \geq 1 - \varrho \quad \text{for all } \varrho > \frac{4\varepsilon_0}{h_m(T)}.$$

This proves the claim if one recalls that  $\delta_0 \leq \frac{h_m(T)}{4}$  and  $\varepsilon_0 = \frac{\delta_0}{20}$ . ■

Our next goal is to use  $h_{\frac{\nu}{\nu(\mathcal{X})}}(T) \leq h_m(T) - \delta_0$  to give an upper bound on the number of Bowen  $L$ -balls needed to cover  $\mathcal{W}$ . Let  $\mathcal{E}_1(\mathcal{X})^T$  denote the space of  $T$ -invariant ergodic probability measures on  $\mathcal{X}$ .

**PROPOSITION 8.7:** *Let  $\varepsilon > 0$ . There exist  $L_1 \in \mathbb{N}$  and  $Y \subseteq \mathcal{X}$  with  $\nu(Y) > \nu(\mathcal{X}) - \varepsilon$  such that for all  $L \geq L_1$ , the set  $Y$  can be covered with*

$$e^{(h_m(T) - \delta_0 + \varepsilon)L}$$

*Bowen  $L$ -balls (with centers in  $Y$ ). Here we may use a radius parameter  $r = r(\varepsilon)$  in the definition of the Bowen balls such that  $10er$  is an injectivity radius on the compact set  $\bar{Y}$ .*

*Proof.* We normalize the measure  $\nu$  to

$$\sigma := \frac{\nu}{\nu(\mathcal{X})}.$$

By the ergodic decomposition of  $\sigma$  and since  $\sigma(\mathcal{O}) = 0$  we find a subset  $\mathcal{X}' \subseteq \mathcal{E}'$  with  $\sigma(\mathcal{X}') = 1$  and a measurable map  $\mathcal{X}' \rightarrow \mathcal{E}_1(\mathcal{X})^T$ ,  $x \mapsto \sigma_x$ , such that

$$\sigma = \int_{\mathcal{X}'} \sigma_x d\sigma(x)$$

and  $\sigma_x(\mathcal{E}') = 1$  for all  $x \in \mathcal{X}'$ . Let  $\mathcal{P}$  be any countable partition of  $\mathcal{X}$  with finite partition entropy  $H_\sigma(\mathcal{P}) < \infty$ . For any  $n \in \mathbb{N}$  and any  $x \in \mathcal{X}$  let  $[x]_{\mathcal{P}_0^{n-1}}$  denote the partition element in  $\mathcal{P}_0^{n-1}$  which contains  $x$ .

We want to find a lower estimate of  $\sigma([x]_{\mathcal{P}_0^{n-1}})$ . For this, let

$$I_\sigma(\mathcal{P}_0^{n-1})(x) := -\log \sigma([x]_{\mathcal{P}_0^{n-1}})$$

denote the information function of  $\mathcal{P}_0^{n-1}$ . By the Shannon–McMillan–Breiman Theorem (see, e.g., [ELW, Theorem 3.2]) there exists a subset  $\mathcal{X}'' \subseteq \mathcal{X}'$  with  $\sigma(\mathcal{X}'') = 1$  such that

$$\frac{1}{n} I_\sigma(\mathcal{P}_0^{n-1})(x) \rightarrow h_{\sigma_x}(T, \mathcal{P}) \quad \text{as } n \rightarrow \infty$$

for all  $x \in \mathcal{X}''$  and in  $L^1$ .

Fix some  $\varepsilon > 0$ . By the above there exists a subset  $Y_1 \subseteq \mathcal{X}''$  with  $\sigma(Y_1) > 1 - \varepsilon$  and  $L_1 \in \mathbb{N}$  such that for all  $L \geq L_1$  and all  $x \in Y_1$  we have

$$\frac{1}{L} I_\sigma(\mathcal{P}_0^{L-1})(x) < h_{\sigma_x}(T, \mathcal{P}) + \varepsilon.$$

From Lemma 8.2 (note that  $\text{supp } \sigma_x \subseteq \mathcal{E}'$  for  $x \in \mathcal{X}'$ ) and the definition of the dynamical entropy it follows that

$$h_{\sigma_x}(T, \mathcal{P}) \leq h_{\sigma_x}(T) \leq h_m(T) - \delta_0$$

for each  $x \in \mathcal{X}'$ . Thus, for any  $L \geq L_1$  and  $x \in Y_1$  we have

$$-\frac{1}{L} \log \sigma([x]_{\mathcal{P}_0^{L-1}}) < h_m(T) - \delta_0 + \varepsilon.$$

Hence

$$(26) \quad \sigma([x]_{\mathcal{P}_0^{L-1}}) \geq e^{-L(h_m(T) - \delta_0 + \varepsilon)}.$$

We now pick a more specific partition as follows. A slight modification of the proof of [EL10, 7.53] (we provide more details in Section 9) allows us to choose a relatively compact subset  $Q \subseteq \mathcal{X}$  with  $\sigma(Q) > 1 - \varepsilon$  such that there exists a countable partition  $\mathcal{P}$  of  $\mathcal{X}$  with  $H_\sigma(\mathcal{P}) < \infty$  such that<sup>4</sup>

$$(27) \quad [x]_{\mathcal{P}_0^{L-1}} \subseteq xB_L$$

for all  $x \in Q$  and all  $L \in \mathbb{N}$ .

Set  $Y := Y_1 \cap Q$ . Then  $\sigma(Y) > 1 - 2\varepsilon$ . Let  $L \geq L_1$ . For any  $x \in Y$  the formula (26) gives a lower bound on the  $\sigma$ -measure of the partition element that contains  $x$ . It follows that we get the upper bound  $e^{L(h_m(T) - \delta_0 + \varepsilon)}$  for the number of distinct (and hence disjoint) elements that intersect  $Y$  nontrivially. For each of those we pick an element in  $Y$  to use as the center for a Bowen  $L$ -ball, which then covers the associated partition element by (27). These Bowen  $L$ -balls now cover  $Y$  as required. ■

LEMMA 8.8: *There exist  $Y \subseteq \mathcal{X}$  and  $L_1 \in \mathbb{N}$  with the following properties. Let  $L_0 \geq L_1$  and choose a cover by Bowen  $L_0$ -balls as in Proposition 8.7. Denote*

---

<sup>4</sup> The construction of the partition  $\mathcal{P}$  takes into account any previously fixed injectivity radius  $r > 0$  on  $Q$ , and the Bowen balls in (27) will use this as the radius parameter.

the closure of the union of the cover by  $Y_0$ . Then there exist infinitely many  $K$  with the property that

$$(28) \quad \frac{1}{K} \sum_{k=0}^{K-1} T_*^{kL_0} \mu(Y_0 a_e^{-n}) > 1 - \frac{5\varepsilon_0}{h_m(T)}$$

for some  $n \in \{0, \dots, L_0 - 1\}$ .

*Proof.* Let  $\varepsilon \in (0, 1)$  (determined below). Let  $(Y, L_1)$  be as in Proposition 8.7. Let  $L_0 \geq L_1$  and define  $Y_0$  as in the statement of the lemma. As  $Y_0$  is closed and by the definition of  $\nu$  we have

$$\liminf \frac{1}{L} \sum_{j=0}^{L-1} T_*^j \mu(Y_0) \geq \nu(Y_0),$$

where  $\liminf$  is taken over a subsequence of  $L$ 's such that  $(\mu_L)$  converges to  $\nu$  (cf. the definition of  $\nu$  and (25)). Since  $Y_0 \supseteq Y$  we have  $\nu(Y_0) \geq \nu(Y) > \nu(\mathcal{X}) - \varepsilon$ . Hence there exist infinitely many  $L$  with

$$\frac{1}{L} \sum_{j=0}^{L-1} T_*^j \mu(Y_0) > \nu(\mathcal{X}) - 2\varepsilon.$$

We divide  $L$  by  $L_0$  with remainder, write  $K = \lceil \frac{L}{L_0} \rceil$  and obtain

$$\begin{aligned} \frac{1}{L} \sum_{j=0}^{L-1} T_*^j \mu(Y_0) &= \frac{1}{L} \left( \sum_{n=0}^{L_0-1} \sum_{k=0}^{K-1} T_*^{kL_0+n} \mu(Y_0) + \sum_{j=L_0K}^{L-1} T_*^j \mu(Y_0) \right) \\ &= \frac{K}{L} \sum_{n=0}^{L_0-1} \frac{1}{K} \sum_{k=0}^{K-1} T_*^{kL_0+n} \mu(Y_0) + \frac{1}{L} \sum_{j=L_0K}^{L-1} T_*^j \mu(Y_0). \end{aligned}$$

The last sum has at most  $L_0$  summands and so converges to 0 as  $L \rightarrow \infty$ . Moreover,  $\frac{K}{L}$  converges to  $\frac{1}{L_0}$ . Therefore, for large enough  $L$  from our sequence we must have

$$\nu(\mathcal{X}) - 3\varepsilon < \frac{1}{L_0} \sum_{n=0}^{L_0-1} \frac{1}{K} \sum_{k=0}^{K-1} T_*^{kL_0+n} \mu(Y_0).$$

Hence we find  $n \in \{0, \dots, L_0 - 1\}$  such that

$$\frac{1}{K} \sum_{k=0}^{K-1} T_*^{kL_0+n} \mu(Y_0) > \nu(\mathcal{X}) - 3\varepsilon.$$

Proposition 8.6 now gives

$$\frac{1}{K} \sum_{k=0}^{K-1} T^{kL_0} \mu(Y_0 a_e^{-n}) > 1 - \frac{4\varepsilon_0}{h_m(T)} - 3\varepsilon.$$

Choosing  $\varepsilon > 0$  sufficiently small, the lemma follows. ■

PROPOSITION 8.9: *Assuming that  $L_0$  in Lemma 8.8 is sufficiently big, there exist infinitely many  $K$  such that a  $\mu$ -proportion of at least  $1/2$  of the set  $\mathcal{W}$  can be covered with*

$$e^{(h_m(T) - \frac{7}{8}\delta_0 + 2\varepsilon_0)KL_0}$$

Bowen  $L_0K$ -balls.

*Proof.* Let  $\varepsilon = \varepsilon_0 > 0$  and apply Proposition 8.7 to obtain  $L_1$ . Let  $L_0 \geq L_1$  (below we will give another constraint for  $L_0$ ) and let  $Y_0$  be the closure of the union of the Bowen  $L_0$ -balls covering  $Y$  as in Lemma 8.8. Fixing  $L_0$  there are infinitely many  $K$  for which the conclusion (28) of Lemma 8.8 holds. Fix one such  $K$  and the corresponding  $n$ . For a given  $y \in \mathcal{W}$  we define

$$V(y) := \{k \in [0, K) \mid T^{n+kL_0}(y) \in Y_0\}.$$

We fix one subset  $V$  of  $[0, K)$  and define

$$\mathcal{Z}(V) := \{y \in \mathcal{W} \mid V(y) = V_0\}.$$

We claim that we can cover  $\mathcal{Z}(V)$  with

$$(29) \quad c_{r, \mathcal{W}} b_1 e^{h_m(T)n} b_1^K e^{|V|(h_m(T) - \delta_0 + \varepsilon_0)L_0} e^{(K-|V|)h_m(T)L_0}$$

Bowen  $KL_0$ -balls, where  $b_1$  is a constant independent of  $K, L, V$  and  $s$ .

We may phrase the conclusion of Lemma 8.8 (in the case where  $n = 0$ ) by saying that most points in  $\mathcal{W}$  spend a lot of time in relatively few Bowen  $L_0$ -balls under the orbit w.r.t.  $T^{L_0}$ . We will use this together with a concatenation procedure to bound how many Bowen  $L_0K$ -balls are needed to cover a portion of  $\mathcal{W}$ . To ensure that we always work with Bowen balls with the same radius, we have to multiply the number of possibilities with an extra factor of  $b_1$  at each concatenation step resulting in the  $b_1^K$ -factor (and need the  $b_1 e^{h_m(T)L_0}$ -factor to handle the case where  $n \neq 0$ ).

More precisely, let  $b_1$  be an upper bound for the number of Bowen  $L$ -balls  $gB_L$  (with parameter  $r$ ) that are needed to cover  $B_L(2r)$  independent of  $L$ . The existence of  $b_1$  follows using a maximal collection of pairwise disjoint Bowen  $L$ -balls  $gB_L(r/2)$  with  $g \in B_L(2r)$  and Lemma 7.2. The simple concatenation step

is now as follows. If  $y_1 B_{L_1}$  and  $y_2 B_{L_2}$  are two Bowen balls such that  $y_2 \in Y_0$ , then  $y_1 B_{L_1} \cap (y_2 \overline{B_{L_2} a_e^{-L_1}})$  can be covered by  $b_1$  Bowen  $L_1 + L_2$ -balls. Indeed take some  $y \in y_1 B_{L_1} \cap (y_2 \overline{B_{L_2} a_e^{-L_1}})$  and we get

$$(y_1 B_{L_1}(r)) \cap (y_2 \overline{B_{L_2}(r) a_e^{-L_1}}) \subseteq y B_{L_1+L_2}(2r)$$

and the claim follows from the definition of  $b_1$ . For the inclusion in the last step we need to know that  $2er$  is an injectivity radius at  $y_2$ , which we know as  $y_2 \in Y_0$  is assumed.

We may also assume that  $b_1 e^{h_m(T)L_2}$  is an upper bound on the number of  $(L_1 + L_2)$ -Bowen balls that are needed to cover  $B_{L_1}$  (choose a pairwise disjoint collection of Bowen  $(L_1 + L_2)$ -balls  $g B_{L_0}(r/2)$  with  $g \in B_{L_1}$  and apply Lemma 7.2), and note that in the case where  $L_1 = 0$  we may define  $B_0$  to be the  $r$ -ball around the identity. The second possible step (which we will call an extension step) then concerns the case where the first Bowen ball  $y_1 B_{L_1}$  is given, but we have no additional information about  $y_1 a_e^{L_1}$ . In this case we cover  $y_1 B_{L_1}$  by  $b_1 e^{h_m(T)L_2}$  Bowen  $(L_1 + L_2)$ -balls.

Using the above concatenation and extension steps iteratively we see that  $\mathcal{Z}(V_0)$  can be covered by (29)-many Bowen  $KL_0$ -balls. Indeed we first cover  $\mathcal{W}$  by  $c_{r,\mathcal{W}}$  balls of radius  $r$  (or equivalently by Bowen 0-balls) and use the extension step with  $L_1 = 0$  and  $L_2 = n$ . Starting with  $k = 0$  and if  $k \in V$ , we use the concatenation for  $L_1 = n + kL_0$  and  $L_2 = L_0$  using all possibilities for the Bowen  $L_0$ -balls appearing in the definition of  $Y_0$  in Lemma 8.8, or use the extension step if  $k \notin V$ . For  $k = K$  we obtain Bowen  $n + KL_0$ -balls and the claim follows.

Note that the above estimate is monotonically decreasing as  $|V|$  increases. Fix some  $\alpha \in (0, 1)$  and consider all subset  $V \subseteq [0, K)$  with  $|V| \geq K\alpha$ . The above now gives that

$$(30) \quad \bigcup_{|V| \geq K\alpha} \mathcal{Z}(V)$$

is covered by

$$c_{r,\mathcal{W}} b_1^{1+K} e^{h_m(T)L_0} e^{\alpha(h_m(T) - \delta_0 + \varepsilon_0)KL_0} e^{(1-\alpha)h_m(T)KL_0} \sum_{k=\lfloor \alpha K \rfloor}^K \binom{K}{k}$$

Bowen  $KL_0$ -balls. It is well known that Sterling’s formula can be used to estimate the last sum. In fact, using Sterling’s formula, one obtains

$$\binom{K}{k} = e^{KH(\frac{k}{K})+o(K)} \quad \text{as } K \rightarrow \infty$$

independent of  $k \in [0, K]$  and where

$$H(x) = -x \log x - (1 - x) \log (1 - x).$$

Note that there are at most  $K = e^{o(K)}$  summands. Bounding  $H$  by  $\log 2$  and taking the sum we obtain the upper bound

$$(31) \quad c_r \mathcal{W} b_1^{1+K} e^{h_m(T)L_0} e^{\alpha(h_m(T)-\delta_0+\varepsilon_0)KL_0} e^{(1-\alpha)h_m(T)KL_0} e^{K(\log 2+1)}$$

if  $K$  is sufficiently large. Note that the product of the constant and the factors containing  $K$  or  $L_0$ , but not both, can be bounded by  $e^{\varepsilon_0 L_0 K}$  if  $L_0$  and  $K$  are sufficiently big. We now choose  $\alpha = 1 - \frac{10\varepsilon_0}{h_m(T)}$  and recall that  $\delta_0 \leq \frac{h_m(T)}{4}$  and  $\varepsilon_0 = \frac{\delta_0}{20}$ , which gives  $\alpha \geq \frac{7}{8}$ . Putting these estimates into (31) we get

$$(32) \quad e^{\frac{7}{8}(h_m(T)-\delta_0+\varepsilon_0)KL_0} e^{\frac{1}{8}h_m(T)KL_0} e^{\varepsilon_0 KL_0} \leq e^{(h_m(T)-\frac{7}{8}\delta_0+2\varepsilon_0)KL_0}$$

for the upper bound on the number of Bowen  $KL_0$ -balls.

We finally apply Lemma 8.8 to bound the set that is not covered by the Bowen balls in (32). In fact, using the above notation we may rewrite (28) to get

$$\sum_{k=0}^K \frac{k}{K} \mu(\{y \in \mathcal{W} \mid |V(y)| = k\}) > 1 - \frac{5\varepsilon_0}{h_m(T)}$$

or

$$\sum_{k=0}^K \left(1 - \frac{k}{K}\right) \mu(\{y \in \mathcal{W} \mid |V(y)| = k\}) < \frac{5\varepsilon_0}{h_m(T)}.$$

The latter implies

$$(1 - \alpha) \mu(\{y \in \mathcal{W} \mid |V(y)| < \alpha K\}) < \frac{5\varepsilon_0}{h_m(T)}$$

or

$$\mu(\{y \in \mathcal{W} \mid |V(y)| < \alpha K\}) < \frac{1}{2}.$$

From this, and since (32) gives an upper bound for the number of Bowen  $KL_0$ -balls needed to cover (30), we obtain the proposition. ■

*Proof of Theorem 8.1.* Let  $L_0$  be sufficiently big for Proposition 8.9 to hold. Together with Lemma 8.3 we obtain

$$\begin{aligned} \frac{1}{2} &= \frac{1}{2} \mu(\mathcal{W}) \leq c_\mu r^{\dim G - 2\varepsilon_0} e^{(-h_m(T) + 2\varepsilon_0)KL_0} e^{(h_m(T) - \frac{7}{8}\delta_0 + 2\varepsilon_0)KL_0} \\ &= \tilde{c} e^{(4\varepsilon_0 - \frac{7}{8}\delta_0)KL_0} \end{aligned}$$

for some constant  $\tilde{c}$  independent of  $K$ , and for infinitely many  $K$ . Note that  $4\varepsilon_0 - \frac{7}{8}\delta_0 < 0$ , which leads to a contradiction. This completes the proof. ■

### 9. Modification of the partition from [EL10]

The partition  $\mathcal{P}$  in the proof of Proposition 8.7 is essentially identical with the partition in [EL10, 7.51]. However, our situation is slightly different to the one in [EL10] for which reason we outline the necessary steps of the proof. The differences are as follows:

- The measure  $\sigma$  is not necessarily ergodic, and we cannot reduce to an ergodic situation as in [EL10].
- We want to find a big set  $Q$  on which the inclusion relation

$$[x]_{\mathcal{P}_0^{N-1}} \subseteq xB_N$$

holds for all  $N \in \mathbb{N}$  (in [EL10], the mass of  $Q$  does not matter as long as it is positive, and it is asked for the (weaker) relation

$$[x]_{\mathcal{P}_0^\infty} \subseteq xB_N$$

for  $N \in \mathbb{N}$  such that  $xa^N \in Q$ ).

- We do not need the lower bounds from [EL10] on the atoms, which allows us to simplify  $\mathcal{P}$  a bit.

The construction of  $\mathcal{P}$  and the proofs of its properties proceeds in a number of steps.

- (1) Pick a subset  $Q \subseteq \mathcal{X}$  which is open, relatively compact and which has mass  $\sigma(Q) > 1 - \varepsilon$ . Pick an injectivity radius  $r$  of  $Q$  and decompose  $Q$  (up to measure zero) into finitely many subsets  $Q_1, \dots, Q_R$  with positive measure such that each of these subsets is contained in a Bowen ball with parameter  $r/16$ . Set

$$\mathcal{Q} := \{Q_1, \dots, Q_R, \mathcal{X} \setminus Q\}.$$

- (2) For each  $i = 1, \dots, R$  and each  $j \in \mathbb{N}$  let  $Q_{ij}$  be the set of points  $x \in Q_i$  which return to  $Q$  with the  $j$ -th step but not earlier, that is

$$Q_{ij} = \{x \in Q_i \mid xa_e^j \in Q, xa_e^\ell \notin Q \text{ for } \ell = 1, \dots, j - 1\}.$$

Set

$$\tilde{Q} := \{\mathcal{X} \setminus Q, Q_{ij} \mid i = 1, \dots, R, j \in \mathbb{N}\}.$$

- (3) We now decompose the partition elements  $Q_{ij}$  into smaller subsets. For this we remark that each ball  $xB_{r/16}^G$  can be covered with

$$ce^{\kappa j}$$

balls with parameter  $e^{-j}r/8$  (here  $c = c(G)$  is a constant and  $\kappa = \dim G$  would work). Let  $i \in \{1, \dots, R\}$  and  $j \in \mathbb{N}$ . Note that  $Q_{ij} \subseteq x_i B_{r/16}^G$  for some  $x_i \in Q$ . We fix a cover  $B_1, \dots, B_{N(j)}$  with  $N(j) \leq ce^{\kappa j}$  by balls with parameter  $e^{-j}r/8$ . We define

$$\begin{aligned} Q_{ij1} &:= Q_{ij} \cap B_1, \\ Q_{ij2} &:= Q_{ij} \cap (B_2 \setminus B_1), \\ Q_{ij3} &:= Q_{ij} \cap (B_3 \setminus (B_1 \cup B_2)), \end{aligned}$$

and so on. Set

$$\mathcal{P} := \{Q_{ijk}, \mathcal{X} \setminus Q \mid i = 1, \dots, R, j \in \mathbb{N}, k = 1, \dots, N(j)\}.$$

LEMMA 9.1: For  $\sigma$ -almost every  $x \in Q$  we have

$$[x]_{\mathcal{P}_0^{N-1}} \subseteq xB_N$$

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

*Proof.* By the Poincaré Recurrence Theorem,  $\sigma$ -almost every point in  $Q$  returns infinitely often to  $Q$ . We restrict to these points and pick such an  $x$ . Let  $N \in \mathbb{N}$  and let

$$y \in [x]_{\mathcal{P}_0^{N-1}}.$$

There exist  $i \in \{1, \dots, R\}$ ,  $j \in \mathbb{N}$  and  $k \in \{1, \dots, N(j)\}$  such that  $x, y \in Q_{ijk}$ . Thus, there exists  $g \in B_{e^{-j}r/4}^G$  such that  $y = xg$ . Then for any  $\ell = 0, \dots, j$  we have

$$g \in a_e^\ell B_{e^{-j}r/4}^G a_e^{-\ell} \subseteq a_e^\ell B_r^G a_e^{-\ell}.$$

In the case  $j \geq N - 1$  it follows immediately that

$$y \in xB_N.$$

Suppose that  $j < N - 1$ . Then we find  $i_2, j_2, k_2$  such that  $xa_e^j, ya_e^j \in Q_{i_2 j_2 k_2}$ . Since

$$ya_e^j = xa_e^j(a_e^{-j}ga_e^j),$$

and  $a_e^{-j}ga_e^j \in B_r^G$  and  $r$  is an injectivity radius, it follows that actually

$$a_e^{-j}ga_e^j \in B_{e^{-j_2}r/4}^G.$$

Now reasoning inductively as above finally shows  $y \in xB_N$ .  $\blacksquare$

LEMMA 9.2: *The partition  $\mathcal{P}$  has finite partition entropy  $H_\sigma(\mathcal{P})$ .*

*Proof.* We note that<sup>5</sup>

$$\sum_{i,j} j\sigma(Q_{ij}) \leq 1.$$

Using this, the proof is parallel to that in [EL10].  $\blacksquare$

## References

- [BK83] M. Brin and A. Katok, *On local entropy*, in *Geometric Dynamics (Rio de Janeiro, 1981)*, Lecture Notes in Mathematics, Vol. 1007, Springer, Berlin, 1983, pp. 30–38.
- [CDKR91] M. Cowling, A. Dooley, A. Korányi and F. Ricci, *H-type groups and Iwasawa decompositions*, *Advances in Mathematics* **87** (1991), 1–41.
- [CDKR98] M. Cowling, A. Dooley, A. Korányi and F. Ricci, *An approach to symmetric spaces of rank one via groups of Heisenberg type*, *Journal of Geometric Analysis* **8** (1998), 199–237.
- [Dan84] S. G. Dani, *On orbits of unipotent flows on homogeneous spaces*, *Ergodic Theory and Dynamical Systems* **4** (1984), 25–34.
- [EK12] M. Einsiedler and S. Kadyrov, *Entropy and escape of mass for  $SL_3(\mathbb{Z}) \backslash SL_3(\mathbb{R})$* , *Israel Journal of Mathematics* **190** (2012), 253–288.
- [EL10] M. Einsiedler and E. Lindenstrauss, *Diagonal actions on locally homogeneous spaces*, in *Homogeneous Flows, Moduli Spaces and Arithmetic*, Clay Mathematics Proceedings, Vol. 10, American Mathematical Society, Providence, RI, 2010, pp. 155–241.
- [ELMV12] M. Einsiedler, E. Lindenstrauss, Ph. Michel and A. Venkatesh, *The distribution of closed geodesics on the modular surface, and Duke’s theorem*, *L’Enseignement Mathématique* **58** (2012), 249–313.
- [ELW] M. Einsiedler, E. Lindenstrauss and T. Ward, *Entropy in ergodic theory and homogeneous dynamics*, book project, <http://maths.dur.ac.uk/~tpcc68/entropy/>.
- [GR70] H. Garland and M. S. Raghunathan, *Fundamental domains for lattices in  $\mathbb{R}$ -rank 1 semisimple Lie groups*, *Annals of Mathematics* **92** (1970), 279–326.

---

<sup>5</sup> Actually there is equality if  $Q \supseteq \mathcal{X}_{\leq s_3}$ , but this is not needed for the proof.

- [Hel00] S. Helgason, *Groups and Geometric Analysis. Integral Geometry, Invariant Differential Operators, and Spherical Functions.*, Mathematical Surveys and Monographs, Vol. 83, American Mathematical Society, Providence, RI, 2000.
- [Hel01] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Graduate Studies in Mathematics, Vol. 34, American Mathematical Society, Providence, RI, 2001.
- [HW13] P. Hubert and B. Weiss, *Ergodicity for infinite periodic translation surfaces*, *Compositio Mathematica* **149** (2013), 1364–1380.
- [Kad12] S. Kadyrov, *Entropy and escape of mass for Hilbert modular spaces*, *Journal of Lie Theory* **22** (2012), no. 3, 701–722.
- [KKLM] S. Kadyrov, D. Y. Kleinbock, E. Lindenstrauss and G. A. Margulis *Singular systems of linear forms and nonescape of mass in the space of lattices*, *Journal d'Analyse Mathématique*, to appear.
- [KP] S. Kadyrov and A. Pohl, *Amount of failure of upper-semicontinuity of entropy in noncompact rank one situations, and Hausdorff dimension*, arXiv.org:1211.3019.
- [MT94] G. A. Margulis and G. M. Tomanov, *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, *Inventiones Mathematicae* **116** (1994), 347–392.
- [Poh10] A. Pohl, *Ford fundamental domains in symmetric spaces of rank one*, *Geometriae Dedicata* **147** (2010), 219–276.
- [Shi12] R. Shi, *Convergence of measures under diagonal actions on homogeneous spaces*, *Advances in Mathematics* **229** (2012), 1417–1434.