SINGULAR SYSTEMS OF LINEAR FORMS AND NON-ESCAPE OF MASS IN THE SPACE OF LATTICES

By

S. KADYROV[1], D. KLEINBOCK[2], E. LINDENSTRAUSS[3] AND G. A. MARGULIS[4]

Abstract. Singular systems of linear forms were introduced by Khintchine in the 1920s, and it was shown by Dani in the 1980s that they are in one-to-one correspondence with certain divergent orbits of one-parameter diagonal groups on the space of lattices. We give a (conjecturally sharp) upper bound on the Hausdorff dimension of the set of singular systems of linear forms (equivalently, the set of lattices with divergent trajectories) as well as the dimension of the set of lattices with trajectories "escaping on average" (a notion weaker than divergence). This extends work by Cheung, as well as by Chevallier and Cheung. Our method differs considerably from that of Cheung and Chevallier and is based on the technique of integral inequalities developed by Eskin, Margulis and Mozes.

1 Introduction

For any given $m, n \in \mathbb{N}$, we consider the space of unimodular $(m+n)$-lattices $X_{m+n} := G/\Gamma$, where $G = \mathrm{SL}(m+n, \mathbb{R})$ and $\Gamma = \mathrm{SL}(m+n, \mathbb{Z})$, and the one-parameter diagonal semigroup $\{g_t\}_{t\geq 0}$, where

$$g_t := \mathrm{diag}(e^{nt}, \ldots, e^{nt}, e^{-mt}, \ldots, e^{-mt}). \tag{1.1}$$

We consider the action $g_t : X_{m+n} \to X_{m+n}$ by left translations, $g_t \cdot x = g_t x$. The unstable horospherical subgroup U with respect to g_1 can be identified with the space $M_{m,n}$ of $m \times n$ real matrices:

$$U = \{u_s : s \in M_{m,n}\} \text{ where } u_s := \begin{pmatrix} I_m & s \\ 0 & I_n \end{pmatrix}. \tag{1.2}$$

[1]S. K. was supported by EPSRC through grant EP/H000091/1.

[2]D. K. was supported by the National Science Foundation through grant DMS-1101320. His stay at MSRI was supported in part by the NSF grant no. 0932078 000.

[3]E. L. was supported by the European Research Council through AdG Grant 267259 and by the ISF grant 983/09. His stay at MSRI was supported in part by the NSF grant no. 0932078 000.

[4]G. M. was supported by the National Science Foundation through grant DMS-1265695.

JOURNAL D'ANALYSE MATHÉMATIQUE, Vol. 133 (2017)
DOI 10.1007/s11854-017-0033-4

The one-parameter diagonal semigroup $\{g_t\}_{t\geq 0}$ and the corresponding horospherical subgroup $U < G$ are closely connected to the diophantine properties of $m \times n$ real matrices. One says that an $m \times n$ real matrix s is a **singular system of** m **linear forms in** n **variables** if for any $\varepsilon > 0$ there exists $T_0 \in \mathbb{N}$ such that for any $T > T_0$ there exist $\mathbf{q} \in \mathbb{Z}^n$ and $\mathbf{p} \in \mathbb{Z}^m$ such that[1]

$$\|s\mathbf{q} - \mathbf{p}\| < \frac{\varepsilon}{T^{n/m}} \text{ and } 0 < \|\mathbf{q}\| < T. \tag{1.3}$$

Equivalently, one can restrict T to be a power of a fixed natural number, e.g., consider only $T = 2^\ell$, $\ell \in \mathbb{N}$. This property was introduced by A. Khintchine [Kh26], who showed that the set of such matrices has Lebesgue measure zero, hence the name "singular". Later, it was shown by S. G. Dani [D85] that s is singular if and only if the trajectory $\{g_t u_s \Gamma : t \geq 0\}$ is divergent in X_{m+n}, that is, leaves every compact subset of X_{m+n}. Thus the zero measure of the set of singular systems follows from the ergodicity of the g_t-action on X_{m+n}.

When $m = n = 1$, it is easy to see that each divergent trajectory $\{g_t x\}$, where $x \in X_2$ is a unimodular lattice in $\mathbb{R}^2$, is, in Dani's terminology, **degenerate**, that is, there exists a subgroup of x contracted by the action. This is a manifestation of the fact that the group $\mathrm{SL}(2, \mathbb{R})$ has $\mathbb{Q}$-rank 1, cf. [D85, Theorem 6.1]. In particular, it follows that the set of singular real numbers (equivalently, 1×1 matrices) coincides with $\mathbb{Q}$ and thus has Hausdorff dimension zero[2]. However, when $\max(m, n) > 1$ (that is, when G has real rank bigger than 1), it is possible to construct trajectories which diverge for non-trivial reasons. This was first observed by Khintchine and then generalized by Dani [D85, Theorem 7.3]. Thus the set of points with divergent orbits has quite a complicated structure. In particular, computing its Hausdorff dimension is a difficult problem. It was shown by Y. Cheung [Che16] that when $m + n = 3$, the Hausdorff dimension of the set of singular pairs is $4/3$; hence the set of points in X_3 with divergent g_t-orbits has Hausdorff dimension $22/3$, that is, codimension $2/3$. Also, recently in [CC14], Cheung and N. Chevallier showed that the set of singular m-vectors has Hausdorff dimension $m^2/(m+1)$, which corresponds to codimension $m/(m+1)$ of the set of divergent trajectories in X_{m+1}.

Now let us say that a point $x \in X_{m+n}$ **escapes on average** (with respect to the semigroup g_t as in (1.1) which we fix for the duration of the paper) if

$$\lim_{N\to\infty} \frac{1}{N} \left| \left\{ \ell \in \{1, \dots, N\} : g_\ell x \in Q \right\} \right| = 0$$

[1]Note that this definition is independent of the choice of norms on $\mathbb{R}^m$ and $\mathbb{R}^n$; later it will be convenient to work with the euclidean norms.

[2]This can also be easily shown using continued fractions.

for any compact set Q in X_{m+n}. Observe that this notion is independent of the parametrization of the orbit; in other words, x escapes on average if and only if

$$\lim_{N\to\infty} \frac{1}{N}\left|\left\{\ell \in \{1,\ldots,N\} : g_{a\ell}x \in Q\right\}\right| = 0$$

for any $a > 0$ (or else one can replace summation by integration). In this paper, we prove

Theorem 1.1. *For any $x \in X_{m+n}$, the set*

$$\{u \in U : ux \text{ escapes on average}\} \tag{1.4}$$

has Hausdorff dimension at most $mn - \frac{mn}{m+n}$. Consequently,

$$\dim\left(\{x \in X_{m+n} : x \text{ escapes on average}\}\right) \leq \dim(X_{m+n}) - \frac{mn}{m+n}. \tag{1.5}$$

One can also consider a corresponding concept in Diophantine approximation, weakening the classical notion of singularity: say that an $m \times n$ real matrix s is **singular on average** if for any $\varepsilon > 0$ one has

$$\lim_{N\to\infty} \frac{1}{N}\left|\left\{\ell \in \{1,\ldots,N\} : \begin{array}{c} \exists\, \mathbf{q} \in \mathbb{Z}^n \text{ and } \mathbf{p} \in \mathbb{Z}^m \text{ such that} \\ (1.3) \text{ holds for } T = 2^\ell \end{array}\right\}\right| = 1.$$

Using Dani's correspondence, as an immediate corollary of Theorem 1.1 we obtain the following.

Corollary 1.2. *The Hausdorff dimension of the set of $s \in M_{m,n}$ which are singular on average (and hence of the set of singular $s \in M_{m,n}$) is at most $mn - \frac{mn}{m+n}$.*

Clearly, the set of points with divergent trajectories is contained in the set of points escaping on average. Thus the lower estimate from [CC14] implies that in the case

$$\min(m,n) = 1,\ m+n \geq 3, \tag{1.6}$$

the bound in (1.5) is sharp; the same has been established in [KP17] when $m = n = 1$. We conjecture that the equality in (1.5) holds in all dimensions m, n. Also, it is natural to conjecture that, unless $m = n = 1$, the set of points with divergent trajectories has the same Hausdorff dimension as that of those which escape on average; this follows from [CC14] and Theorem 1.1 in the case (1.6).

Our second result relates the entropy of an invariant measure on X_{m+n} to its mass in a compact set, in a way which is uniform over all invariant probability measures on X_{m+n}.

Theorem 1.3. *For every $\varepsilon > 0$, there exists a compact subset $Q = Q(\varepsilon)$ of X_{m+n} such that*

$$h_\mu(g_1) \le \big(m+n-1+\mu(Q)\big)mn + \varepsilon \tag{1.7}$$

for any g_1-invariant probability measure μ on X_{m+n}.

As a consequence of Theorem 1.3, we obtain the following.

Theorem 1.4. *For any $h > 0$ and any sequence $(\mu_k)_{k\ge 1}$ of g_1-invariant probability measures on X_{m+n} with entropies $h_{\mu_k}(g_1) \ge h$, any weak* limit μ of the sequence satisfies*

$$\mu(X_{m+n}) \ge \frac{h}{mn} - (m+n-1).$$

We remark that the maximal entropy is $(m+n)mn$, and for any value $h \in \big((m+n-1)mn, (m+n)mn\big]$ Theorem 1.4 produces a nontrivial result. A similar statement was first proved in [ELMV12] for the geodesic flow on the unit tangent bundle to the hyperbolic plane. Later, various generalizations were considered in [EK12, Kad12b, EKP15]. Theorem 1.3 can be thought as a generalization of analogous results from [ELMV12, EK12]. However the method used in the present paper is different from the previous work and crucially relies on the ideas from [EMM98]. There are also interesting parallels to these results in the context of moduli spaces of abelian and quadratic differentials, in particular [Ham11] by U. Hamenstädt; [Ath06] and [EM11] give applications of the [EMM98] techniques in the context of these moduli spaces that are also relevant.

We believe that Theorem 1.4 is sharp in the sense that for any constant $h \in \big[0, (m+n)mn\big]$ there should exists a sequence of probability invariant measures $(\mu_k)_{k\ge1}$ with $\lim_k h_{\mu_k}(g_1) = h$ such that the limit measure μ satisfies $\mu(X_{m+n}) = \max\{\frac{h}{mn} - (m+n-1), 0\}$. In [Kad12a], the claim was proved to be true when $\min(m,n) = 1$.

Both Theorem 1.1 and Theorem 1.3 are derived from Theorem 1.5, the main result of this paper. In what follows, we fix m, n and for the sake of brevity denote X_{m+n} by X. Take U as in (1.2) and let d_U be the distance induced by the euclidean norm $\|\cdot\|$ on $M_{m,n}$ via the map $s \mapsto u_s$. Also let

$$B_r^U := \{u_s : \|s\| < r\} = \{u \in U : d_U(u,e) < r\} \tag{1.8}$$

be the ball of radius r centered at the identity element. Then given a compact subset Q of X, $N \in \mathbb{N}$, $\delta \in (0,1)$, $t > 0$, and $x \in X$, define the set

$$Z_x(Q,N,t,\delta) := \left\{u \in B_1^U : \frac{1}{N}\big|\{\ell \in \{1,\dots,N\} : g_{t\ell}ux \notin Q\}\big| \ge \delta\right\}; \tag{1.9}$$

in other words, the set of $u \in B_1^U$ such that up to time N, the proportion of times ℓ for which the orbit point $g_{t\ell}ux$ is in the complement of Q is at least δ.

The following statement is a covering result which we need for the proof of previously mentioned theorems.

Theorem 1.5. *There exist $t_0 > 0$ and a function $C : X \to \mathbb{R}_+$ such that the following holds. For any $t > t_0$, there exists a compact set $Q := Q(t)$ of X such that for any $N \in \mathbb{N}$, $\delta \in (0, 1)$, and $x \in X$, the set $Z_x(Q, N, t, \delta)$ can be covered with $C(x)t^{3N}e^{(m+n-\delta)mntN}$ balls in U of radius $e^{-(m+n)tN}$.*

Remark 1.6. The function C can actually be made precise as follows. For any $i = 1, \dots, m+n$ and any $x \in X$, we let $F_i(x)$ denote the set of all i-dimensional subgroups of x (we recall that the latter is viewed as a unimodular lattice in $\mathbb{R}^{m+n}$). For any $L \in F_i(x)$, we let $\|L\|$ denote the volume of $L/(L \cap x)$ with respect to the standard euclidean structure on $\mathbb{R}^{m+n}$. (Equivalently, $\|L\| = \|v_1 \wedge \cdots \wedge v_i\|$ where $\{v_1, \dots, v_i\}$ is a basis for L.) Then, following [EMM98], define

$$\alpha_i(x) := \max\left\{\frac{1}{\|L\|} : L \in F_i(x)\right\} \tag{1.10}$$

and take

$$C(x) := \max\{\alpha_i^{\beta_i}(x) : i = 1, 2, \dots, m+n-1\}, \tag{1.11}$$

where for any $i \in \{1, \dots, m+n-1\}$ we let

$$\beta_i := \begin{cases} \frac{m}{i} & \text{if } i \leq m \\ \frac{n}{m+n-i} & \text{if } i > m. \end{cases} \tag{1.12}$$

In the next section, we show how to deduce Theorems 1.1 and 1.3 from Theorem 1.5. From §3, the rest of the paper is devoted to obtaining Theorem 1.5.

2 Proofs assuming Theorem 1.5

Notation. In what follows, by $x \ll y$ (resp., $x \asymp y$) we mean $x < Cy$ (resp., $cy < x < Cy$) for some absolute constants $c, C > 0$ depending only on m, n. Fix a right-invariant Riemannian metric on G, inducing a metric on X which will be denoted by d_X. We let B_r^G denote the open ball in G of radius r centered at the identity element.

Proof of Theorem 1.1. Let $x \in X$ be given. Fix $t_0 > 0$ as in Theorem 1.5, and for any $t > t_0$ choose the compact set Q as in Theorem 1.5. Then, using

Theorem 1.5, we get that for any $N \in \mathbb{N}$, each set $Z_x(Q, N, t, \delta)$ can be covered with $C(x)t^{3N}e^{(m+n-\delta)mntN}$ balls of radius $e^{-(m+n)tN}$.

Denote by Z_x the set of all $u \in B_1^U$ such that ux escapes on average. Note that

$$Z_x \subset \bigcup_{N_0 \geq 1} \bigcap_{N \geq N_0} Z_x(Q, N, t, \delta) \text{ for any } \delta \in (0, 1). \tag{2.1}$$

Using Theorem 1.5, we can write

$$\begin{aligned} \overline{\dim}_{\text{box}}\Big(\bigcap_{N \geq N_0} Z_x(Q, N, t, \delta)\Big) &\leq \overline{\lim}_{N \to \infty} \frac{\log\left(C(x)t^{3N}e^{(m+n-\delta)mntN}\right)}{-\log(e^{-(m+n)tN})} \\ &= \frac{3\log t + (m+n-\delta)mnt}{(m+n)t}. \end{aligned}$$

This is true for any $t > t_0$ and any $\delta \in (0, 1)$; thus, letting $t \to \infty$ and $\delta \to 1$, we get from (2.1) that $\dim(Z_x) \leq \frac{(m+n-1)mn}{m+n}$. Since the set (1.4) is contained in a countable union of sets of the form Z_x, the first part of Theorem 1.1 follows.

Now, let P^0 be the weak stable horospherical subgroup with respect to g_1, namely

$$P^0 := \left\{ \begin{pmatrix} s' & 0 \\ s & s'' \end{pmatrix} \,\middle|\, \begin{array}{c} s \in M_{n,m},\ s' \in M_{m,m},\ s'' \in M_{n,n} \\ \det(s')\det(s'') = 1 \end{array} \right\}. \tag{2.2}$$

Every element of a neighborhood of the identity in G can be written as gu, where u belongs to a neighborhood of the identity in U and $g \in P^0$. Note that for $g \in P^0$, the union $\bigcup_{t>0} g_t g g_{-t}$ is contained in a compact subset of G. Writing $g_t gu = g_t g g_{-t}(g_t u)$, one sees that ux escapes on average if and only if gux does also. Therefore, the "consequently" part follows from the slicing properties of the Hausdorff dimension. □

Remark 2.1. One can generalize the definition of escape on average, saying that a point $x \in X$ **δ-escapes on average**, where $0 < \delta \leq 1$, if

$$\lim_{N \to \infty} \frac{1}{N} \left| \left\{ \ell \in \{1, \ldots, N\} : g_\ell x \notin Q \right\} \right| \geq \delta$$

for any compact $Q \subset X$. The previous definition corresponds to $\delta = 1$. Our proof of Theorem 1.1 actually establishes that

$$\dim\left(\{x \in X : x\ \delta\text{-escapes on average}\}\right) \leq \dim(X) - \frac{\delta mn}{m+n}.$$

It seems plausible to conjecture that the above bound is sharp for all m, n and any $0 < \delta \leq 1$.

We now proceed with the proof of Theorem 1.3. Let us fix t_0 as in Theorem 1.5 and take $t > t_0$. We fix sufficiently small $\eta > 0$ such that the ball B_η^G is an injective image under the exponential map of a neighborhood of 0 in the Lie algebra of G. For any $N \in \mathbb{N}$, we define a **Bowen N-ball** to be any set of the form $\mathrm{Bow}(N)x$, where $x \in X$ and

$$\mathrm{Bow}(N) := \bigcap_{\ell=0}^{N-1} g_{-t\ell} B_\eta^G g_{t\ell}. \tag{2.3}$$

We need the following lemma, which relates the entropy and covers by Bowen balls. It essentially is due to Brin and Katok [BK83], though there are some modifications needed for the non-compact case.

Lemma 2.2. *Let μ be an ergodic g_t-invariant probability measure on X and let $A \subset X$ be a measurable subset with $\mu(A) > 0$. For any $N \geq 1$, let $BC(A, N)$ be the minimal number of Bowen N-balls needed to cover A. Then*

$$h_\mu(g_t) \leq \liminf_{N\to\infty} \frac{\log BC(A, N)}{N}.$$

For a proof for $G = \mathrm{SL}(2, \mathbb{R})$, see e.g., [ELMV12, Lemma B.2] and the remarks following it. The adaptation to $\mathrm{SL}(m + n, \mathbb{R})$ is straightforward.

Proof of Theorem 1.3. Note first that it is sufficient to consider ergodic measures. For if μ is not ergodic, we can write μ as an integral of its ergodic components $\mu = \int \mu_r d\nu(r)$ for some probability space (E, ν); see, for example, [EW11]. Therefore, for any compact subset Q of X, we have $\mu(Q) = \int_E \mu_r(Q)d\nu(r)$, but also $h_\mu(g_t) = \int_E h_{\mu_r}(g_t)d\nu(r)$; see for example, [Wal82, Theorem 8.4]. Hence the desired estimate follows from the ergodic case (note that we have used the fact that Q does not depend on the measure at hand). From now on, we assume μ to be ergodic.

Let $\varepsilon > 0$ be given. Fix $t > t_0$ so large that $(3 \log t)/t < \varepsilon$ with t_0 as in Theorem 1.5. For this t, let Q be as in Theorem 1.5. We establish the conclusion of Theorem 1.3 for this compact set Q.

Note that we can assume that $\mu(Q) < 1$, since otherwise (1.7) holds trivially due to the fact that $(m + n)mn$ is an upper bound for $h_\mu(g_1)$. Choose $c > 0$ such that

$$B_{c\eta}^U \subset B_{\eta/2}^G \tag{2.4}$$

(we recall that the metric on U does not coincide with the restriction of the metric on G, but locally near the identity element those metrics are close to each other).

Also choose an open neighborhood W of the identity in the group P^0 as in (2.2) such that

$$B^U_{c\eta} W \subset B^G_\eta, \tag{2.5}$$

which is possible in view of (2.4).

Let us fix any μ-generic point $x \in X$. Then

$$\varepsilon' := \frac{\mu(B^U_1 Wx)}{2} > 0.$$

The pointwise ergodic theorem implies

$$\frac{1}{N}\sum_{n=0}^{N-1} 1_{X\smallsetminus Q}\big(g_{nt}(y)\big) \to \mu(X \smallsetminus Q) \tag{2.6}$$

as $N \to \infty$ for μ-a.e. $y \in X$. In particular, for any $\varepsilon'' \in \big(0, \mu(X \smallsetminus Q)\big)$ there exists N_0 such that for $N > N_0$, the average in the left hand side of (2.6) is larger than $\mu(X \smallsetminus Q) - \varepsilon''$ for any $y \in Y$ for some $Y \subset X$ with measure $\mu(Y) > 1 - \varepsilon'$. Since $\mu(B^U_1 Wx) = 2\varepsilon'$, we see that $\mu(Z) > \varepsilon'$, where $Z := Y \cap B^U_1 Wx$. We now consider the covering of this set Z by Bowen N-balls. Note that $B^U_{c\eta e^{-(m+n)tN}} W \subset \mathrm{Bow}(N)$ in view of (2.5) and (2.3). Thus it suffices to consider a covering of

$$Z' := \{u \in B^U_1 : ux \in Z\} = \{u \in B^U_1 : ux \in Y\}$$

with balls of radius $c\eta e^{-(m+n)tN}$. Denote $\mu(X \smallsetminus Q) - \varepsilon''$ by δ; then clearly $Z' \subset \bigcap_{N>N_0} Z_x(Q, N, t, \delta)$. Applying Theorem 1.5, we get that for any $N > N_0$, the set Z' can be covered with $C(x)t^{3N} e^{(m+n-\delta)mntN}$ balls of radius $e^{-(m+n)tN}$. Observe that a ball of radius $e^{-(m+n)tN}$ in U can be covered by finitely many translates of a ball of radius $c\eta e^{-(m+n)tN}$. Thus for any $N > N_0$, the set Z can be covered with $\ll C(x)t^{3N} e^{(m+n-\delta)mntN}$ Bowen N-balls. Since $\mu(Z) > \varepsilon' > 0$, from Lemma 2.2 it follows that

$$h_\mu(g_t) \le \liminf_{N\to\infty} \frac{\log BC(Z, N)}{N} \le \big(m + n - \mu(X \smallsetminus Q) + \varepsilon''\big)mnt + 3\log t.$$

Since $\varepsilon'' \in \big(0, \mu(X \smallsetminus Q)\big)$ is arbitrary and $(3\log t)/t < \varepsilon$, we arrive at

$$h_\mu(g_1) = \frac{1}{t} h_\mu(g_t) \le \big(m + n - \mu\big(X \smallsetminus Q\big)\big)mn + \varepsilon. \tag{2.7}$$

This finishes the proof. □

We end this section by giving the proof of Theorem 1.4.

Proof of Theorem 1.4. Let us take $\varepsilon > 0$ and a compact subset $Q = Q(\varepsilon)$ as in Theorem 1.3. Since, $h_{\mu_k} \geq h$, using Theorem 1.3 we see that for any $k \in N$,

$$\mu_k(Q) \geq \frac{h}{mn} - (m+n) + 1 - \frac{\varepsilon}{mn}.$$

Pick a compactly supported continuous function $f : X \to [0, 1]$ such that $f(x) = 1$ on Q. Then $\int f d\mu_k \geq \mu_k(Q)$. Let μ be a weak* limit of $(\mu_k)_{\geq 1}$. Then, letting $k \to \infty$, we see that

$$\mu(X) \geq \int f d\mu \geq \frac{h}{mn} - (m+n) + 1 - \frac{\varepsilon}{mn}.$$

To finish the proof, we now let $\varepsilon \to 0$. □

For the rest of the paper, our goal is to prove Theorem 1.5.

3 Estimates for certain integral operators

We fix the standard euclidean structure on $\mathbb{R}^{m+n}$, let $\{e_1, \ldots, e_{m+n}\}$ be the standard basis of $\mathbb{R}^{m+n}$ and let $K = \mathrm{SO}(m+n)$ be the group of orientation-preserving linear isometries of $\mathbb{R}^{m+n}$ (maximal compact in G). The main goal of this section is to prove an estimate for averages of certain functions over K. We let dk stand for the normalized Haar measure on K.

Proposition 3.1. *For any $i \in \{1, \ldots, m+n-1\}$, let $\beta_i = \frac{m}{i}$ if $i \leq m$ and $\beta_i = \frac{n}{m+n-i}$ if $i > m$, as defined in* (1.12). *Then there exists $c > 0$ (dependent only on m, n) such that*

$$\int_K \|g_t k v\|^{-\beta_i}\, dk \leq c t e^{-mnt} \|v\|^{-\beta_i},$$

for any $t \geq 1$ and any decomposable $v \in \bigwedge^i \mathbb{R}^{m+n}$.

We note that except for $i = 1, m$ and $m+n-1$ the factor t in the right hand side is not necessary.

We need some lemmas.

Lemma 3.2. *Let $i \in \{1, \ldots, d\}$ and let $x_1, \ldots, x_i$ be independent K-invariant standard Gaussian random variables on $\mathbb{R}^d$. Then*

$$\mathbb{E}(\|x_1 \wedge \cdots \wedge x_i\|^{-\beta}) < \infty$$

for $\beta < d - i + 1$.

We will be applying the lemma both for $d = m$ and $d = m + n$.

Proof. Let $\pi_{\perp,j}$ denote the orthogonal projection to the space perpendicular to $x_1, \dots, x_j$. Then, for $\ell \in \mathbb{N}$,

$$\mathrm{Prob}(\|x_1 \wedge \cdots \wedge x_i\| \asymp e^{-\ell}) \asymp \sum_{0 \le \ell_1, \ell_2, \dots, \ell_i \le \ell} \prod_{j=1}^{i} \mathrm{Prob}(\|\pi_{\perp,j-1} x_j\| \asymp e^{-\ell_j}),$$

where $\ell_1, \dots, \ell_i$ run through integers with $\sum \ell_j = \ell$. Now,

$$\mathrm{Prob}\big(\|\pi_{\perp,j-1} x_j\| \asymp e^{-\ell}\big)$$

is the probability a standard Gaussian in $\mathbb{R}^{d-j+1}$ has size $\asymp e^{-\ell}$, which is $e^{-(d-j+1)\ell}$. It is now easy to conclude that

$$\mathrm{Prob}(\|x_1 \wedge \cdots \wedge x_i\| \asymp e^{-\ell}) \asymp e^{-(d-i+1)\ell}, \tag{3.1}$$

and the lemma follows. □

Corollary 3.3. *Let $i \in \{1, \dots, d\}$ and let A be some event depending on $x_1, \dots, x_i$, with $x_1, \dots, x_i$ standard Gaussians as before. Take $\beta < d - i + 1$. Then*

$$\mathbb{E}(\|x_1 \wedge \cdots \wedge x_i\|^{-\beta} \mid A) \ll \mathrm{Prob}(A)^{-\frac{\beta}{d-i+1}}.$$

Proof. Observe that from all events A with given $\mathrm{Prob}(A) = p$,

$$\mathbb{E}(\|x_1 \wedge \cdots \wedge x_i\|^{-\beta} \mid A)$$

is maximal for A of the form

$$A = \{(x_1, \dots, x_i) : \|x_1 \wedge \cdots \wedge x_i\| \le \sigma\},$$

where σ is chosen so that $\mathrm{Prob}(A) = p$. Write

$$A_\ell = \{x_1, \dots, x_i : \|x_1 \wedge \cdots \wedge x_i\| \asymp \sigma e^{-\ell}\}; \tag{3.2}$$

then $A = \bigcup_{\ell=0}^{\infty} A_\ell$ for an appropriate (uniform in ℓ) choice of implicit constants in (3.2). Thus, using (3.1), we get that

$$\begin{aligned}\mathbb{E}(\|x_1 \wedge \cdots \wedge x_i\|^{-\beta} \mid A) &\ll \sum_{\ell=0}^{\infty} \mathbb{E}(\|x_1 \wedge \cdots \wedge x_i\|^{-\beta} \mid A_\ell) \frac{\mathrm{Prob}(A_\ell)}{p} \\ &\ll p^{-1} \sum_{\ell=0}^{\infty} (\sigma e^{-\ell})^{-\beta} (\sigma e^{-\ell})^{d-i+1} \ll p^{-1} \sigma^{d-i+1-\beta}.\end{aligned}$$

Using (3.1) one more time, we see that $p = \mathrm{Prob}(A) \asymp \sigma^{d-i+1}$, so that

$$\mathbb{E}(\|x_1 \wedge \cdots \wedge x_i\|^{-\beta} \mid A) \le \mathrm{Prob}(A)^{-\frac{\beta}{d-i+1}}.$$ □

For later purposes we also need the following weaker estimate for $\beta = d - i + 1$.

Lemma 3.4. *For $i \in \{1, \dots, d\}$ and $\kappa > 0$,*

$$\mathbb{E}\left(\|x_1 \wedge \cdots \wedge x_i\|^{-(d-i+1)} \mathbb{1}(\|x_1 \wedge \cdots \wedge x_i\| > e^{-\kappa})\right) \ll \kappa,$$

where the implicit constant is independent of κ.

Proof. Let A_ℓ be as in (3.2). Again using the estimate (3.1), we see that

$$\begin{aligned}\mathbb{E}\big(\|x_1 \wedge \cdots \wedge x_i\|^{-(d-i+1)} \mathbb{1}(\|x_1 \wedge \cdots \wedge x_i\| > e^{-\kappa})\big) &\asymp \\ &\asymp \sum_{\ell=0}^{\kappa} \mathbb{E}(\|x_1 \wedge \cdots \wedge x_i\|^{-(d-i+1)} \mid A_\ell) \mathrm{Prob}(A_\ell) \\ &\asymp \sum_{\ell=0}^{\kappa} 1 = \kappa.\end{aligned}$$

□

Proof of Proposition 3.1. We assume that $i \leq m$, and deal with the other case by duality.

Consider $I_t(v) = \int_K \|g_t k v\|^{-\beta_i} \, dk$, where $v = x_1 \wedge \cdots \wedge x_i$ is decomposable. Since K acts transitively on the variety of the decomposable wedges up to homothety, we have $I_t(v) = C(t)\|v\|^{-\beta_i}$ for some function $C(t)$. It follows that if $v = x_1 \wedge \cdots \wedge x_i$ with x_i independent (K-invariant) standard Gaussians and $\beta_i < m + n - i + 1$ (which is certainly true in our choice of β_i), then

$$C(t) = \frac{1}{C}\mathbb{E} I_t(x_1 \wedge \cdots \wedge x_i) = \frac{1}{C}\mathbb{E}(\|g_t(x_1 \wedge \cdots \wedge x_i)\|^{-\beta_i}),$$

where the last equality follows from K-invariance and

$$C = \mathbb{E}(\|x_1 \wedge \cdots \wedge x_i\|^{-\beta_i})$$

is independent of t. Thus, to prove the proposition we need only show that if $x_1, \dots, x_i$ are standard Gaussian random variables, then

$$\mathbb{E}(\|g_t(x_1 \wedge \cdots \wedge x_i)\|^{-\beta_i}) \leq cte^{-mnt}, \tag{3.3}$$

where c is independent of t.

Let $V_u \subset \mathbb{R}^{m+n}$ denote the m-dimensional subspace spanned by $e_1, \dots, e_m$, and let V_s be the complementary subspace, so that

$$\|g_t v\| = e^{nt}\|v\| \text{ and } \|g_t w\| = e^{-mt}\|w\|$$

whenever $v \in V_u$ and $w \in V_s$. In particular, for any $v \in \bigwedge^i V_u$, we have $\|g_t v\| = e^{int}\|v\|$. Let $\pi_u^{(i)} : \bigwedge^i \mathbb{R}^{m+n} \to \bigwedge^i V_u$ be the natural (orthogonal) projection.

Clearly,

$$\pi_u^{(i)}(x_1 \wedge \cdots \wedge x_i) = \pi_u^{(1)}(x_1) \wedge \cdots \wedge \pi_u^{(1)}(x_i),$$

and each of $\pi_u^{(1)}(x_j)$ is a standard Gaussian random variable in m dimensions.

To show (3.3) we first assume that $1 < i < m$. Clearly, one has

$$\|g_t(x_1 \wedge \cdots \wedge x_i)\| \geq \|\pi_u^{(i)} g_t(x_1 \wedge \cdots \wedge x_i)\|.$$

Then, using Lemma 3.2 with $d = m$, we in fact get

$$\begin{aligned}\mathbb{E}(\|g_t(x_1 \wedge \cdots \wedge x_i)\|^{-\beta_i}) &\leq \mathbb{E}(\|\pi_u^{(i)} g_t(x_1 \wedge \cdots \wedge x_i)\|^{-\beta_i}) \\ &= e^{-in\beta_i t}\mathbb{E}(\|\pi_u^{(i)}(x_1 \wedge \cdots \wedge x_i)\|^{-\beta_i}) \ll e^{-mnt},\end{aligned}$$

as $\beta_i = \frac{m}{i} < m - i + 1$ for $1 < i < m$.

Now assume that $i = 1$ or $i = m$. In this case, we need to be a bit subtler, since $\beta_i = m - i + 1$. However, we note that

$$\|g_t(x_1 \wedge \cdots \wedge x_i)\| \geq \max\left(\|g_t \pi_u^{(i)}(x_1 \wedge \cdots \wedge x_i)\|, e^{-imt}\|x_1 \wedge \cdots \wedge x_i\|\right).$$

Hence, for any $\kappa > 0$,

$$\mathbb{E}(\|g_t(x_1 \wedge \cdots \wedge x_i)\|^{-\beta_i}) \leq E_1 + E_2,$$

where

$$E_1 = \mathbb{E}\left(\|g_t \pi_u^{(i)}(x_1 \wedge \cdots \wedge x_i)\|^{-\beta_i} \mathbb{1}(\|\pi_u^{(i)}(x_1 \wedge \cdots \wedge x_i)\| > e^{-\kappa t})\right)$$

and

$$E_2 = e^{im\beta_i t}\mathbb{E}\left(\|x_1 \wedge \cdots \wedge x_i\|^{-\beta_i} \mathbb{1}(\|\pi_u^{(i)}(x_1 \wedge \cdots \wedge x_i)\| < e^{-\kappa t})\right).$$

From Lemma 3.4 for $d = m$ we get

$$\begin{aligned}E_1 &= e^{-in\beta_i t}\mathbb{E}\left(\|\pi_u^{(i)}(x_1 \wedge \cdots \wedge x_i)\|^{-\beta_i} \mathbb{1}(\|\pi_u^{(i)}(x_1 \wedge \cdots \wedge x_i)\| > e^{-\kappa t})\right) \ll \kappa t e^{-in\beta_i t} \\ &= \kappa t e^{-mnt}.\end{aligned}$$

To estimate E_2, write

$$A = \{(x_1, \ldots, x_i) : \|\pi_u^{(i)}(x_1 \wedge \cdots \wedge x_i)\| < e^{-\kappa t}\}$$

and recall that $\mathrm{Prob}(A) \asymp e^{-\kappa t(m-i+1)}$. Hence, using Corollary 3.3 for $d = m + n$, we conclude

$$E_2 = e^{im\beta_i t}\mathrm{Prob}(A)\mathbb{E}(\|x_1 \wedge \cdots \wedge x_i\|^{-\beta_i} \mid A) \leq \begin{cases} e^{m^2 t} e^{-\kappa m t} e^{\frac{\kappa m^2 t}{m+n}} & \text{if } i = 1, \\ e^{m^2 t} e^{-\kappa t} e^{\frac{\kappa t}{n+1}} & \text{if } i = m. \end{cases}$$

Either way, if κ is a sufficiently large constant (depending on n and m), then $E_2 \ll e^{-mnt}$.

This concludes the proof of the proposition for $i \leq m$.

For $i > m$ we exploit duality. Define the linear map

$$* : \bigwedge^{j} \mathbb{R}^{m+n} \to \bigwedge^{m+n-j} \mathbb{R}^{m+n} \text{ by } *(\wedge_{i \in I} e_i) = \wedge_{i \notin I} e_i.$$

Then, $*(kv) = k(*v)$, $*(g_t v) = g_{-t}(*v)$, and $\| * v\| = \|v\|$. Therefore,

$$\int_U \|g_t k v\|^{-\beta}\, dk = \int_K \| * g_t k v\|^{-\beta}\, dk = \int_K \|g_{-t} k(*v)\|^{-\beta}\, dk,$$

and the desired estimate follows by applying the above for $j' = m + n - j$, $n' = m$ and $m' = n$. $\square$

We now show that Proposition 3.1 remains valid if integration over K is replaced with integration over a bounded subset of U (with the constant dependent on that subset).

Lemma 3.5. *There exists a neighborhood V of the identity in $M_{m,n}$ such that for any $s_0 \in M_{m,n}$, $t, \beta > 0$, $i \in \{1, \cdots, m + n - 1\}$, and decomposable $w \in \bigwedge^i \mathbb{R}^{m+n}$,*

$$\int_{V+s_0} \|g_t u_s w\|^{-\beta}\, ds \ll (1 + \|s_0\|)^{\beta} \int_K \|g_t k w\|^{-\beta}\, dk,$$

with the implied constant dependent only on m, n and β.

Proof.

We make use of the groups

$$N = \left\{ \begin{pmatrix} 1 & * & \cdots & * \\ 0 & 1 & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \right\} \quad \text{and} \quad N^0 = N \cap P^0,$$

where P^0 is as in (2.2). Note that $N = N^0 U$ and $U \lhd N$.

There is a local diffeomorphism $f : K \to N$ such that in a neighborhood $\mathcal{O}$ of the identity in K, we may write $k = p(k) f(k)$ for some $p(k) \in P^0$, with the Jacobian of f bounded from above and below in this neighborhood. Suppose $\mathcal{O}_U$, $\mathcal{O}_{N^0}$ are neighborhoods of the identity in U and N^0, respectively, so that $\mathcal{O}_{N^0}\mathcal{O}_U \subset f(\mathcal{O})$.

Then

$$\begin{aligned}\int_K \|g_t k w\|^{-\beta}\, dk &\geq \int_{\mathcal{O}} \|g_t p(k) f(k) w\|^{-\beta}\, dk \\ &\geq \int_{\mathcal{O}} \left\|g_t p(k)^{-1} g_{-t}\right\|^{-\beta} \|g_t f(k) w\|^{-\beta}\, dk \\ &\gg \int_{\mathcal{O}_{N^0}} \int_{\mathcal{O}_U} \|g_t n_0 u w\|^{-\beta}\, dn_0 du \\ &= \int_{\mathcal{O}_{N^0}} \int_{\mathcal{O}_U} \|n_0 g_t u w\|^{-\beta}\, dn_0 du \\ &\gg \int_{\mathcal{O}_U} \|g_t u w\|^{-\beta}\, du\end{aligned}$$

We may assume that $\|w\| = 1$. For any $u \in U$, we may find $k \in K$ such that $kuw = \pm\|uw\| w$. Also, for any $s \in M_{m,n}$,

$$\|u_s w\| \geq \|u_s^{-1}\|^{-1} \geq (1 + \|s\|)^{-1}.$$

Setting $V = \{s \in M_{m,n} : u_s \in \mathcal{O}_U\}$, we have that for any $s_0 \in M_{m,n}$

$$\begin{aligned}\int_{V+s_0} \|g_t u_s w\|^{-\beta} ds &= \int_V \|g_t u_s u_{s_0} w\|^{-\beta} ds \ll \int_K \|g_t k(u_{s_0} w)\|^{-\beta}\, dk \\ &= \|u_{s_0} w\|^{-\beta} \int_K \|g_t k w\|^{-\beta}\, dk \leq (1 + \|s_0\|)^{\beta} \int_K \|g_t k w\|^{-\beta}\, dk. \square\end{aligned}$$

Recall that in Remark 1.6 we defined $\alpha_i(x)$ to be the maximum value of $1/\|L\|$, where $\|L\|$ is the volume of $L/(L \cap x)$ and L runs through the set $F_i(x)$ of all i-dimensional subgroups of x. Clearly, $\alpha_{m+n}(x) = 1$, and for convenience we also set $\alpha_0(x) = 1$ for all $x \in X$.

In the next corollary, we replace integration over a neighborhood of the identity in U with integration over all of U with respect to $d\rho$, which denotes the Gaussian probability measure on $M_{m,n}$, where each component is i.i.d. with mean 0 and variance 1. Using Proposition 3.1 and Lemma 3.5, we argue as in [EMM98, Lemma 5.7] to obtain the following corollary.

Corollary 3.6. *Let $\{\beta_i : i = 1, \ldots, m+n-1\}$ be as in Proposition 3.1. Then there exists $c_0 > 0$ with the following property: given any $t \geq 1$, one can choose $\omega > 0$ such that for any $x \in X$ and $i \in \{1, \ldots, m+n-1\}$,*

$$\int_U \alpha_i(g_t u_s x)^{\beta_i}\, d\rho(s) \leq c_0 t e^{-mnt} \alpha_i(x)^{\beta_i} + \omega^{2\beta_i} \max_{0 < j \leq \min\{m+n-i, i\}} \left(\sqrt{\alpha_{i+j}(x)\alpha_{i-j}(x)}\right)^{\beta_i}.$$

Proof. For a given $x \in X$, let $L_i \in F_i(x)$ be a subgroup of x such that

$$\alpha_i(x) = \frac{1}{\|L_i\|}. \tag{3.4}$$

For any $i \in \{1, \dots, m+n\}$, $L \in F_i(x)$, and $s \in M_{m,n}$, one has

$$(1 + \|s\|)^{-1}\|L\| \le \|u_s L\| \le (1 + \|s\|)\|L\|.$$

Let $\omega = \max_{0<j<m+n} \| \bigwedge^j g_t\|$; then

$$\omega^{-1} \le \frac{\|g_t v\|}{\|v\|} \le \omega \text{ for any } 0 < j < m+n,\ v \in \textstyle\bigwedge^j(\mathbb{R}^{m+n}) \setminus \{0\}. \tag{3.5}$$

Let us consider $\Psi_i := \{L : L \in F_i(x),\ \|L\| < \omega^2\|L_i\|\}$. For any $L \in F_i(x) \setminus \Psi_i$, we have

$$\|u_s L_i\| \le (1 + \|s\|)\|L_i\| \le \frac{1 + \|s\|}{\omega^2}\|L\| \le \frac{(1 + \|s\|)^2}{\omega^2}\|u_s L\| \text{ for any } s \in M_{m,n}.$$

Hence, for any $s \in M_{m,n}$, we see from (3.5) that

$$\|g_t u_s L_i\| \le (1 + \|s\|)^2\|g_t u_s L\|. \tag{3.6}$$

First, assume that $\Psi_i = \{L_i\}$. Then, (3.6) gives

$$\int_U \alpha_i(g_t u_s x)^{\beta_i}\, d\rho(s) \le \int_U (1 + \|s\|)^{2\beta_i}\|g_t u_s L_i\|^{-\beta_i}\, d\rho(s). \tag{3.7}$$

Clearly, for any $s_0 \in M_{m,n}$,

$$\begin{aligned}\int_{V+s_0} (1 + \|s\|)^{2\beta_i}\|g_t u_s L_i\|^{-\beta_i}\, d\rho(s) &\ll \left(\max_{s\in V+s_0}(1 + \|s\|)^{2\beta_i} e^{-\frac{\|s\|^2}{2}}\right)\int_{V+s_0} \|g_t u_s w\|^{-\beta_i} ds\\ &\le e^{-\frac{\|s_0\|^2}{2}+O(\|s_0\|)}\int_{V+s_0}\|g_t u_s w\|^{-\beta_i} ds,\end{aligned}$$

where the implied constants are independent of s_0. Summing over a lattice Λ in the vector space $M_{n,m}$, sufficiently fine so that $M_{n,m} = V + \Lambda$, and using Lemma 3.5, we see that

$$\begin{aligned}\int_U (1 + \|s\|)^{2\beta_i}\|g_t u_s L_i\|^{-\beta_i}\, d\rho(s) &\le \sum_{s'\in\Lambda}\int_{V+s'}(1 + \|s\|)^{2\beta_i}\|g_t u_s L_i\|^{-\beta_i}\, d\rho(s)\\ &\ll \sum_{s'\in\Lambda} e^{-\frac{\|s'\|^2}{2}+O(\|s'\|)}\int_{V+s'}\|g_t u_s L_i\|^{-\beta_i} ds\\ &\ll \sum_{s'\in\Lambda}(1 + \|s'\|)^{\beta_i} e^{-\frac{\|s'\|^2}{2}+O(\|s'\|)}\int_K \|g_t k L_i\|^{-\beta_i}\, dk\\ &\ll \int_K \|g_t k L_i\|^{-\beta_i}\, dk.\end{aligned}$$

Thus, Proposition 3.1 and (3.7) give

$$\int_U \alpha_i(g_t u_s x)^{\beta_i}\, d\rho(s) \le c_0 t e^{-mnt}\|L_i\|^{-\beta_i} = c_0 t e^{-mnt}\alpha_i(x)^{\beta_i}$$

when $\Psi_i = \{L_i\}$, where c_0 depends only on m, n. Otherwise, let $L' \in \Psi_i$, $L' \neq L_i$. Then $\dim(L_i + L') = i + j$ for some $j > 0$. From (3.4), (3.5) and [EMM98, Lemma 5.6] we get for all $s \in M_{m,n}$ that

$$\begin{aligned}\alpha_i(g_t u_s x) < (1 + \|s\|)^{\beta_i}\omega\alpha_i(x) = \frac{(1 + \|s\|)^{\beta_i}\omega}{\|L_i\|} &< \frac{(1 + \|s\|)^{\beta_i}\omega^2}{\sqrt{\|L_i\|\|L'\|}}\\ &\le \frac{(1 + \|s\|)^{\beta_i}\omega^2}{\sqrt{\|L_i \cap L'\|\|L_i + L'\|}} \le (1 + \|s\|)^{\beta_i}\omega^2\sqrt{\alpha_{i+j}(x)\alpha_{i-j}(x)}.\end{aligned}$$

Hence, if $\Psi_i \neq \{L_i\}$, then

$$\begin{aligned}&\int_U \alpha_i(g_t u_s x)^{\beta_i}\, d\rho(s)\\ &\qquad \le \omega^{2\beta_i} \max_{0<j\le\max\{m+n-i,i\}} \left(\sqrt{\alpha_{i+j}(x)\alpha_{i-j}(x)}\right)^{\beta_i} \int_U (1 + \|s\|)^{\beta_i}\, d\rho(s).\end{aligned}$$

Since $\int_U (1 + \|s\|)^{\beta_i} d\rho(s) \ll 1$, combining the above two cases leads to the desired result. $\square$

4 Choosing an appropriate height function

In this section, we study an abstract setting which allows us to choose certain functions on X to be used later for constructing the compact subset Q of Theorem 1.5.

Proposition 4.1. *Let $d \in \mathbb{N}$ be given, and let $\beta : \{0, \ldots, d\} \to \mathbb{R}_+$ be a concave[3] function such that $\beta(0) = \beta(d) = 0$. Let H be a set and let A be a linear operator in the space of real functions on H with $A(1) = 1$. Suppose we are given functions $f_i : H \to \mathbb{R}_+$, where $i = 0, \ldots, d$, such that $f_0 = f_d = 1$ and the following inequalities hold:*

$$A(f_i^{\beta_i}) \le a f_i^{\beta_i} + \omega^{2\beta_i} \max_{0<j\le\min\{d-i,i\}} \left(\sqrt{f_{i+j} f_{i-j}}\right)^{\beta_i}, \ i = 1, \ldots, d-1, \tag{4.1}$$

where $\beta_i = 1/\beta(i)$ and a, ω are some positive constants. Then for any $a' > a$, there exist constants $\omega_0, \ldots, \omega_d > 0$ and $C_0 > 1$ such that the linear combination

$$f := \sum_{i=0}^{d} (\omega_i f_i)^{\beta_i} \tag{4.2}$$

satisfies

$$(Af)(h) \le a' f(h) + C_0 \quad \textit{for all } h \in H.$$

[3] More precisely, if the piecewise linear interpolation of $i \mapsto \beta(i)$ is a concave function $[0, d] \to \mathbb{R}_+$.

We note that the special case of the proposition when

$$\beta(1) = \beta(2) = \cdots = \beta(d-1)$$

appears in [EMM98]. In our context, $H = X$, $f_i = \alpha_i$ as in (1.10), $d = m+n$, A is defined by

$$(Af)(x) := \int_U f(g_t u_s x)\, d\rho(s)$$

for a fixed $t \geq 1$, and the function β is given by $\beta(i) = 1/\beta_i$, where β_i's are as in (1.12). Clearly, β satisfies the convexity assumption. Take c_0 as in Corollary 3.6, and then, for arbitrary $t \geq 1$, choose ω as in Corollary 3.6 and let $a = c_0 te^{-mnt}$ and $a' = 2c_0 te^{-mnt}$. After that take $\omega_0, \ldots, \omega_{m+n}$ and C_0 as in Proposition 4.1. Now consider the linear combination

$$\tilde{\alpha} := \sum_{i=0}^{m+n} (\omega_i \alpha_i)^{\beta_i}; \tag{4.3}$$

Proposition 4.1 then implies that for any $x \in X$, one has

$$\int_U \tilde{\alpha}(g_t u_s x)\, d\rho(s) \leq 2c_0 te^{-mnt}\tilde{\alpha}(x) + C_0.$$

Furthermore, the right hand side of the above inequality is not greater than $3c_0 te^{-mnt}\tilde{\alpha}(x)$ if $\tilde{\alpha}(x) > C_0 e^{mnt}/c_0 t$. This way, one arrives at the following result.

Corollary 4.2. *There exists a constant $c > 0$ with the following property: for any $t \geq 1$, one can choose constants $\omega_0, \ldots, \omega_{m+n}$ and T such that for any $x \in X$ with $\tilde{\alpha}(x) > T$, where $\tilde{\alpha}$ is as in* (4.3)*, one has*

$$\int_U \tilde{\alpha}(g_t u_s x)\, d\rho(s) \leq cte^{-mnt}\tilde{\alpha}(x).$$

Proof of Proposition 4.1. Let

$$\Psi := \left\{(i,j) \in \mathbb{N}^2 : 0 < i < d, \quad 0 < j \leq \min(i, d-i)\right\}$$

and

$$\Phi = \left\{(i,j) \in \Psi : \beta(i) = \frac{1}{2}\big(\beta(i-j) + \beta(i+j)\big)\right\},$$

and define

$$b := \max_{(i,j)\in\Psi\smallsetminus\Phi} \left(\frac{\beta(i-j)+\beta(i+j)}{2\beta(i)}\right)$$

(note that $b < 1$ in view of the definition of Φ). Let $\varepsilon \in (0,1)$ be a small parameter to be set later. Let I be the set of indices $i \in \{0, \ldots, d\}$ at which $\beta(\cdot)$ is strictly concave, i.e., such that $(i, j) \notin \Phi$ for all j. By definition, $0, d \in I$. Define for every $i \in \{0, \ldots, d\}$

$$d_-(i) := \min_{i' \leq i, i' \in I} (i - i'), \quad d_+(i) := \min_{i' \geq i, i' \in I} (i' - i);$$

note in particular that $d_-(0) = d^+(d) = 0$. Now, for every $i \in \{0, \ldots, d\}$, define

$$\omega_i := \varepsilon^{d_-(i) d_+(i)}.$$

Clearly, for any $i \in \{0, \ldots, d\}$ we have $f_i \leq \omega_i^{-1} f^{\beta(i)}$, where f is defined as in (4.2). Applying (4.1), we can write

$$\begin{aligned} Af &= \sum_{i=0}^{d} \omega_i^{\beta_i} A(f_i^{\beta_i}) = 2 + \sum_{i=1}^{d-1} \omega_i^{\beta_i} A(f_i^{\beta_i}) \\ &\leq 2 + \sum_{i=1}^{d-1} \omega_i^{\beta_i} \left(a f_i^{\beta_i} + \omega^{2\beta_i} \max_{0 < j \leq \min\{d-i, i\}} \left(\sqrt{f_{i+j} f_{i-j}} \right)^{\beta_i} \right) \\ &\leq 2(1-a) + af + \sum_{(i,j) \in \Psi} \omega_i^{\beta_i} \omega^{2\beta_i} \left(\sqrt{f_{i+j} f_{i-j}} \right)^{\beta_i}. \end{aligned}$$

Now, for $(i, j) \in \Psi \smallsetminus \Phi$, it follows that

$$\left(f_{i-j}(h) f_{i+j}(h) \right)^{1/2\beta(i)} \leq C_{i,j,\varepsilon} f(h)^{\frac{\beta(i-j)\beta(i+j)}{2\beta(i)}} \leq C_{i,j,\varepsilon} \left(1 + f(h)^b \right)$$

for some constants $C_{i,j,\varepsilon}$ and all $h \in H$. On the other hand, $(i, j) \in \Phi$ implies

$$d_-(i-j) = d_-(i) - j = d_-(i+j) - 2j$$

and

$$d_+(i-j) = d_+(i) + j = d_+(i+j) + 2j.$$

Hence $\sqrt{\omega_{i+j} \omega_{i-j}} = \omega_i \varepsilon^{j^2}$, and thus

$$\begin{aligned} \omega_i^{\beta_i} \left(f_{i-j}(h) f_{i+j}(h) \right)^{\beta_i/2} &\leq \left(\frac{\omega_i}{\sqrt{\omega_{i-j} \omega_{i+j}}} \right)^{\beta_i} \left(f(h)^{\beta(i-j)} f(h)^{\beta(i+j)} \right)^{\beta_i/2} \\ &= \varepsilon^{\beta_i j^2} f(h). \end{aligned}$$

We now estimate separately this sum for $(i, j) \in \Psi \smallsetminus \Phi$ and for $(i, j) \in \Phi$:

$$\sum_{(i,j) \in \Psi \smallsetminus \Phi} \omega^{2\beta_i} \left(\sqrt{f_{i+j} f_{i-j}} \right)^{\beta_i} \leq C_{\varepsilon,\omega}(1 + f^b),$$

for some constant $C_{\varepsilon,\omega}$, and

$$\sum_{(i,j)\in\Phi} \omega^{2\beta_i} \left(\sqrt{f_{i+j} f_{i-j}}\right)^{\beta_i} \leq \left(\sum_{(i,j)\in\Psi} \omega^{2\beta_i} \varepsilon^{\beta_i j^2}\right) f.$$

Recall that we are given an arbitrary $a' > a$. If we take ε so small that

$$\sum_{(i,j)\in\Psi} \omega^{2\beta_i} \varepsilon^{\beta_i j^2} < \frac{a'-a}{2},$$

then, using the fact that $\omega_i^{\beta_i} \leq 1$, we can conclude that

$$\begin{aligned} Af &\leq 2(1-a) + af + C_{\varepsilon,\omega}(1+f^b) + \frac{a'-a}{2} f \\ &< 2 + a'f - \frac{a'-a}{2} f + C_{\varepsilon,\omega}(1+f^b) \leq a'f + C'_{\varepsilon,\omega} \end{aligned}$$

for an appropriate constant $C'_{\varepsilon,\omega}$. □

5 Averages of $\tilde{\alpha}$ under g_{Nt} and coverings by small balls

For the next two statements, we take c as in Corollary 4.2 and, for a given $t \geq 1$, fix $\tilde{\alpha}$ as in (4.3) and T as in Corollary 4.2. Also, for any $x \in X$, $M > 0$ and $N \in \mathbb{N}$, let us define

$$Z_x(M, N, t) := \left\{u \in B_1^U : \tilde{\alpha}(g_{t\ell} ux) > M \text{ for all } \ell \in \{1, \dots, N\}\right\}.$$

Observe that the function $\tilde{\alpha}$ is proper due to Mahler's Compactness Criterion; hence, for any $M > 0$, the set $X_{\leq M} := \{x \in X : \tilde{\alpha}(x) \leq M\}$ is compact. Note also that clearly $Z_x(M, N, t) = Z_x(X_{\leq M}, N, t, 1)$, where the latter is defined as in (1.9).

Proposition 5.1. *There exists $M_0 > T$ such that for any $x \notin X_{\leq T}$, any $N \in \mathbb{N}$ and any $M > M_0$,*

$$\int_{Z_x(M,N-1,t)} \tilde{\alpha}(g_{Nt} u_s x) ds \ll c^N t^N e^{-mntN} \tilde{\alpha}(x).$$

Proof. For any $\sigma > 0$, let ρ_{σ^2} denote the Gaussian probability measure on $\mathbb{R}^{mn}$ where each component is i.i.d. with mean 0 and variance σ^2. In particular, $\rho = \rho_1$. We use the following fact: for any continuous function f on U and $\varepsilon > 0$,

$$\int_U \int_U f(\varepsilon x + y)\, d\rho_1(x)\, d\rho_1(y) = \int_U f(z)\, d\rho_{1+\varepsilon^2}(z). \tag{5.1}$$

Let $t \geq 2c$ be given and let $\tilde{\alpha}$, T be as above. Corollary 4.2 gives

$$\int_U \tilde{\alpha}(g_t u_s x)\, d\rho_1(s) \leq cte^{-mnt}\tilde{\alpha}(x) \text{ for any } x \in X \text{ with } \tilde{\alpha}(x) > T. \tag{5.2}$$

It follows from the definition of $\tilde{\alpha}$ that there is a constant $C_{\tilde{\alpha}}$, dependent only on m, n, such that

$$C_{\tilde{\alpha}}^{-1} \leq \frac{\tilde{\alpha}(ux)}{\tilde{\alpha}(x)} \leq C_{\tilde{\alpha}} \quad \text{for any } u \in B_2^U \text{ and } x \in X. \tag{5.3}$$

Pick $M \geq C_{\tilde{\alpha}} T$. Let $N \in \mathbb{N}$ be given. Using (5.2) repeatedly, we get

$$\int \cdots \int_Z \tilde{\alpha}(g_t u_{s_N} \cdots g_t u_{s_1} x)\, d\rho_1(s_N) \cdots d\rho_1(s_1) \leq c^N t^N e^{-mntN} \tilde{\alpha}(x), \tag{5.4}$$

where

$$Z = \{(s_1, \ldots, s_N) \in U^N : \tilde{\alpha}(g_t u_k \cdots g_t u_1 x) > T \text{ for all } k = 1, \ldots, N-1\}.$$

Write $g_t u_{s_N} \ldots g_t u_{s_1} = g_{Nt} u_{\phi(s_1,\ldots,s_N)}$, where

$$\phi(s_1, \ldots, s_N) := \sum_{k=1}^{N} e^{-(k-1)(m+n)t} s_k.$$

Thus, using (5.1), we see that (5.4) takes the form

$$\int_U \mathbb{1}_{\phi(Z)}(s)\tilde{\alpha}(g_{Nt} u_s x)\, d\rho_{\sigma^2}(s) \leq c^N t^N e^{-mntN} \tilde{\alpha}(x), \tag{5.5}$$

where $\sigma^2 = \sum_{k=1}^{N} e^{-2(k-1)(m+n)t}$. Although σ^2 depends on N, it is in $[1, 2]$ whenever $t \geq 1$. This implies that ds is absolutely continuous with respect to $d\rho_{\sigma^2}$ on B_1^U, with a uniform bound on the Radon-Nikodym derivative. Hence, (5.5) gives

$$\int_{B_1^U} \mathbb{1}_{\phi(Z)}(s)\tilde{\alpha}(g_{Nt} u_s x) ds \ll c^N t^N e^{-mntN} \tilde{\alpha}(x). \tag{5.6}$$

We claim that $\phi(s_1, \ldots, s_N) \in Z_x(M, N-1, t)$ implies that $(s_1, \ldots, s_N) \in Z$. Assuming the claim, we get that

$$\int_{Z_x(M,N-1,t)} \tilde{\alpha}(g_{Nt} u_s x) ds \ll c^N t^N e^{-mntN} \tilde{\alpha}(x).$$

So it suffices to show the claim. Suppose that

$$s = \phi(s_1, \ldots, s_N) = \sum_{k=1}^{N} e^{-(k-1)(m+n)t} s_k \in Z_x(M, N-1, t).$$

This means that for any $\ell = 1, \ldots, N-1$, we have

$$\tilde{\alpha}(g_{\ell t} u_s x) = \tilde{\alpha}\big((g_{\ell t} u_{s'} g_{-\ell t}) g_{\ell t} u_{s''}\big) > M, \tag{5.7}$$

where $s' = \sum_{k=\ell+1}^{N} e^{-(k-1)(m+n)t} s_k$ and $s'' = \sum_{k=1}^{\ell} e^{-(k-1)(m+n)t} s_k$. Clearly, $g_{\ell t} u_{s'} g_{-\ell t} \in B_2^U$; therefore, (5.7) together with our choice of M give

$$\tilde{\alpha}(g_{\ell t} u_{s''} x) = \tilde{\alpha}(g_t u_\ell \cdots g_t u_1 x) > T.$$

Thus $(s_1, \ldots, s_N) \in Z$, which finishes the proof of the claim. $\square$

Corollary 5.2. *For all sufficiently large $M > 0$, any $N \in \mathbb{N}$ and $x \in X$, the set $Z_x(M, N, t)$ can be covered with $\frac{\tilde{\alpha}(x)}{M} C_1^N t^N e^{(m+n-1)mntN}$ balls in U of radius $e^{-(m+n)tN}$ for some $C_1 > 0$ independent of t, N and x.*

Proof. We partition B_1^U into $p \leq Ce^{(m+n-1)mntN}$ disjoint subsets $D_1, \ldots, D_p$, each containing a ball of radius $r/10$ and contained in a ball of radius r for $r = e^{-(m+n)tN}$. This can be done, e.g., by choosing a maximal $r/2$-separated subset of B_1^U, say $\{u_1, \ldots, u_p\}$, and letting

$$D_i := \left(B_r^U u_i \setminus \Big(\bigcup_{j=1}^{i-1} D_j \cup \bigcup_{j=i+1}^{p} B_{r/2}^U u_i \Big) \right) \cap B_1^U.$$

Note that for u_i near the boundary of B_1^U, the set D_i may fail to contain the ball $B_{r/2}^U(x_i)$, but $B_{r/2}^U u_i \cap B_1^U \subset D_i$ certainly contains some ball of radius $r/10$.

Let ν denote Lebesgue measure on U, normalized so that $\nu(B_1^U) = 1$. By Proposition 5.1, we know that for M sufficiently large (depending on t)

$$\int_{Z_x(C_{\tilde{\alpha}}^{-1} M, N-1, t)} \tilde{\alpha}(g_{Nt} u_s x) ds \ll c^N t^N e^{-mntN} \tilde{\alpha}(x)$$

with $C_{\tilde{\alpha}}$ as in the proof of that proposition; hence

$$\nu\big(Z_x(C_{\tilde{\alpha}}^{-1} M, N, t)\big) \ll \frac{\tilde{\alpha}(x)}{M} c^N t^N e^{-mntN}.$$

Since $\nu(D_i) \gg e^{-mn(m+n)tN}$ for all i, it follows that the number p_1 of the D_i's that are contained in $Z_x(C_{\tilde{\alpha}}^{-1} M, N, t)$ satisfies

$$p_1 \leq \frac{\tilde{\alpha}(x)}{M} C_1^N t^N e^{mn(m+n-1)tN}$$

(for an appropriate constant C_1 independent of t, N, x). Reordering the D_i's, if necessary, we can assume that these are exactly $D_1, \ldots, D_{p_1}$.

Take now $i > p_1$. Then D_i contains at least one element u outside the set $Z_x(C_{\tilde{\alpha}}^{-1}M, N, t)$; therefore, for some $1 \le \ell \le N$, one has $\tilde{\alpha}(g_{t\ell}ux) \le M/C_{\tilde{\alpha}}$. But then, as $D_i \subset B_r^U u_i$,

$$g_{t\ell}D_i \subset g_{t\ell}B_{2r}^U u = B^U_{e^{(m+n)t\ell}2r} g_{t\ell}u.$$

Hence, since $e^{(m+n)t\ell}2r \le 2$, by definition of $C_{\tilde{\alpha}}$ one has $g_{t\ell}D_i x \subset X_{\le M}$, so D_i is disjoint from $Z_x(M, N, t)$. Thus $Z_x(M, N, t) \subset \bigcup_{i=1}^{p_1} D_i$, and the proposition follows. $\square$

Proof of Theorem 1.5. Let $t > t_0$ be given, with t_0 a large real number, to be determined later, depending only on m, n. Let $\tilde{\alpha}$ and T be as in Corollary 4.2. We will find large enough $M > 0$ such that the compact set $Q = X_{\le M}$ satisfies the conclusion of the theorem.

For a given $N \in \mathbb{N}$ and $x \in X$, we consider a subset J_x of $\{1, \dots, N\}$ given by

$$J_x := \big\{\ell \in \{1, \dots, N\} : g_{\ell t}x \notin Q\big\}.$$

Then one can write the set $Z_x(Q, N, t, \delta)$ as

$$Z_x(Q, N, t, \delta) = \{u \in B_1^U : |J_{ux}| \ge \delta N\}.$$

For any subset J of $\{1, \dots, N\}$, we set $Z(J) := \{u \in B_1^U : J_{ux} = J\}$ We note that $Z_x(Q, N, t, \delta) = \bigcup_J Z(J)$, where the union runs over all subsets J of $\{1, \dots, N\}$ with cardinality at least δN. Clearly, the number of such subsets of $\{1, \dots, N\}$ is at most $2^N \le t^N$. Thus, it suffices to show that for a given subset $J \subset \{1, \dots, N\}$, the set $Z(J)$ can be covered with $C(x)t^{2N}e^{mnt[(m+n)N-|J|]}$ balls of radius $e^{-(m+n)tN}$, for

$$C(x) = \max\{\alpha_i^{\beta_i}(x) : i = 1, 2, \dots, m+n-1\}; \tag{5.8}$$

cf. (1.11).

Let J be as above. We decompose J into ordered subintervals $J_1, \dots, J_p$ of maximal possible sizes such that $J = \bigsqcup_{i=1}^p J_i$. Let $I_1, \dots, I_{p'}$ be the ordered maximal subintervals of $\{1, \dots, N\} \setminus J$ such that

$$\{1, \dots, N\} = \bigsqcup_{i=1}^{p} J_i \sqcup \bigsqcup_{j=1}^{p'} I_j.$$

We now inductively prove the following claim: for any integer $L \le N$, if

$$\{1, \dots, L\} = \bigsqcup_{i=1}^{\ell} J_i \sqcup \bigsqcup_{j=1}^{\ell'} I_j, \tag{5.9}$$

then the set $Z(J)$ can be covered with

$$S \le \max\left(1, \frac{\tilde{\alpha}(x)}{M}\right) t^{2L} e^{mnt[(m+n)L-|J_1|-\cdots-|J_\ell|]} \tag{5.10}$$

sets of the form $Du_1, \dots, Du_S$ where $D = g_t^{-L} B_\eta^U g_t^L$, i.e. a ball of radius $e^{-(m+n)tL}$. Comparing (4.3) and (5.8), we note that for sufficiently large M, we have

$$\max\left(1, \frac{\tilde{\alpha}(x)}{M}\right) \le C(x). \tag{5.11}$$

Thus, by letting $L = N$, we establish the claim for $|J| \ge \delta N$. If in the first step we have $\{1, \dots, L\} = J_1$, then (5.10) follows from Corollary 5.2 once $t \ge C_1$ and M is large enough to satisfy the conclusion of the corollary. If $\{1, \dots, L\} = I_1$, then as $Z(J) \subset B_1^U$, it is obvious that the set $Z(J)$ can be covered with $\le C_2 e^{mnt(m+n)L}$ balls of radius $e^{-(m+n)L}$, with C_2 depending only on nm.

Assume now the set $Z(J)$ can be covered with S balls of radius $e^{-(m+n)tL}$ for some L satisfying (5.9). In the inductive step, for the next $L' > L$ satisfying an equation similar to (5.9), we have two cases: either

$$\{1, \dots, L'\} = \{1, \dots, L\} \sqcup I_{\ell'+1} \tag{5.12}$$

or

$$\{1, \dots, L'\} = \{1, \dots, L\} \sqcup J_{\ell+1}. \tag{5.13}$$

Consider first the case (5.12). Obviously, each box Du_i in U of size $\eta e^{-(m+n)tL}$ can be covered by $C_2 e^{(m+n)t|I_{\ell'+1}|mn}$ balls of radius $e^{-(m+n)t(L+|I_{\ell'+1}|)}$. Thus, noting that $L + |I_{\ell'+1}| = L'$, from (5.10) and (5.11) it follows that if we assume, as we may, that $t > C_2$, the set $Z(J)$ can be covered by $C(x)t^{2L'} e^{mnt[(m+n)L'-|J_1|-\cdots-|J_\ell|]}$ balls of radius $\eta e^{-(m+n)tL'}$, as claimed.

Now assume (5.13) and consider one of the balls Du_i of radius $e^{-(m+n)tL}$. We are interested in bounding the number of balls of radius $e^{-(m+n)t(L+|J_{\ell+1}|)}$ needed to cover $Z(J) \cap Du_i$. If $Z(J) \cap Du_i = \varnothing$, there is nothing to cover.

So let $u \in Z(J) \cap Du_i$. By definition of $Z(J)$, this implies $\tilde{\alpha}(g_t^L ux) \le M$; on the other hand, $\tilde{\alpha}(g_t^j ux) > M$ for all $j \in J_{\ell+1}$. Since g_t expands every vector in $\bigwedge \mathbb{R}^{m+n}$ by at most $e^{C_3 t}$, with C_3 depending only on m, n, it follows that

$$\tilde{\alpha}(g_t^L u_i x) \ge e^{-C_3 t} M.$$

Hence, using (5.3), one gets

$$C_{\tilde{\alpha}} M \ge \tilde{\alpha}(g_t^L x) \ge C_{\tilde{\alpha}}^{-1} e^{-C_3 t} M.$$

Assuming M is large enough (depending on t) for Corollary 5.2 to be applicable to $x' = g_t^L u_i x$, we see that $Z_{x'}(M, |J_{\ell+1}|, t)$ can be covered by

$$C_{\tilde{\alpha}} C_1^{|J_{\ell+1}|} t^{|J_{\ell+1}|} e^{(m+n-1)mnt|J_{\ell+1}|} \leq t^{2|J_{\ell+1}|} e^{(m+n-1)mnt|J_{\ell+1}|}$$

balls in U of radius $e^{-(m+n)|J_{\ell+1}|}$ (assuming $t_0 > C_{\tilde{\alpha}} C_1$).

Note that by definition of $Z(J)$ and $Z_x(\cdot, \cdot, \cdot)$, one has

$$Z(J) \cap Du_i \subset g_t^{-L} Z_{x'}(M, |J_{\ell+1}|, t) g_t^L.$$

Our bound on the number of $e^{-(m+n)|J_{\ell+1}|}$-balls needed to cover $Z_{x'}(M, |J_{\ell+1}|, t)$ implies that $g_t^{-L} Z_{x'}(M, |J_{\ell+1}|, t) g_t^L$ can be covered by at most

$$t^{2|J_{\ell+1}|} e^{(m+n-1)mnt|J_{\ell+1}|}$$

balls of radius $e^{-(m+n)(L+|J_{\ell+1}|)} = e^{-(m+n)(L')}$. Hence $Z(J)$ can be covered by

$$t^{2L'} e^{mnt[(m+n)L' - |J_1| - \cdots - |J_{\ell+1}|]}$$

balls of radius $e^{-(m+n)(L')}$, establishing the inductive hypothesis. □

Acknowledgements. We thank J. Athreya, Y. Cheung, M. Einsiedler, A. Eskin and M. Mirzakhani for helpful discussions, and the referee for useful comments. We are also grateful to MSRI, where this paper was finalized, for its hospitality during Spring 2015. E. L. also thanks the Israel IAS for ideal working conditions during the fall of 2013.

References

[Ath06] J. Athreya. *Quantitative recurrence and large deviations for Teichmuller geodesic flow*, Geom. Dedicata **119** (2006), 121–140.

[BK83] M. Brin and A. Katok, *On local entropy*, in *Geometric Dynamics (Rio de Janeiro, 1981)*, Springer, Berlin, 1983, pp. 30–38.

[Che16] Y. Cheung, *Hausdorff dimension of the set of singular pairs*, Ann. of Math. (2) **173** (2011), 127–167.

[CC14] Y. Cheung and N. Chevallier, *Hausdorff dimension of singular vectors*, Duke Math. J. **165** (2016), 2273–2329.

[D85] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*, J. Reine Angew. Math. **359** (1985), 55–89.

[EK12] M. Einsiedler and S. Kadyrov, *Entropy and escape of mass for* $\mathrm{SL}(3,\mathbb{Z})\backslash \mathrm{SL}(3,\mathbb{R})$, Israel J. Math. **190** (2012), 253–288.

[EKP15] M. Einsiedler, S. Kadyrov, and A. D. Pohl, *Escape of mass and entropy for diagonal flows in real rank one situations*, Israel J. Math. **210** (2015), 245–295.

[ELMV12] M. Einsiedler, E. Lindenstrauss, P. Michel, and A. Venkatesh, *The distribution of closed geodesics on the modular surface, and Duke's theorem*, Enseign. Math. (2) **58** (2012), 249–313.

[EMM98] A. Eskin, G. A. Margulis , and S. Mozes, *Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture*, Ann. of Math. (2) **147** (1998), 93–141.

[EM11] A. Eskin and M. Mirzakhani, *Counting closed geodesics in moduli space*, J. Mod. Dyn. **5** (2011), 71–105.

[EW11] M. Einsiedler and T. Ward, *Ergodic Theory with a View Towards Number Theory*, Springer-Verlag London, Ltd. London, 2011.

[Ham11] U. Hamenstädt, *Symbolic dynamics for the Teichmueller flow*, arXiv:1112.6107[math.DS].

[Kad12a] S. Kadyrov, Positive entropy invariant measures on the space of lattices with escape of mass, Ergodic Theory Dynam. Systems **32** (2012), 141–157.

[Kad12b] S. Kadyrov, *Entropy and escape of mass for Hilbert modular spaces*, J. Lie Theory **22** (2012), 701–722.

[Kh26] A. Khintchine, *Über eine Klasse linearer diophantischer Approximationen*, Rend. Circ. Mat. Palermo **50** (1926), 170–195.

[KP17] S. Kadyrov and A. D. Pohl, *Amount of failure of upper-semicontinuity of entropy in non-compact rank one situations, and Hausdorff dimension*, Ergodic Theory Dynam. Systems **37** (2017), 539–563.

[Wal82] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, New York-Berlin, 1982.

S. Kadyrov
MATHEMATICS DEPARTMENT
NAZARBAYEV UNIVERSITY
ASTANA, KAZAKHSTAN
email: shirali.kadyrov@nu.edu.kz

D. Kleinbock
DEPARTMENT OF MATHEMATICS
BRANDEIS UNIVERSITY
WALTHAM, MA, USA
email: kleinboc@brandeis.edu

E. Lindenstrauss
EINSTEIN INSTITUTE OF MATHEMATICS
THE HEBREW UNIVERSITY OF JERUSALEM
JERUSALEM, ISRAEL
email: elon@math.huji.ac.il

G. A. Margulis
DEPARTMENT OF MATHEMATICS
YALE UNIVERSITY
NEW HAVEN, CT, USA
email: grigorii.margulis@yale.edu

(Received July 19, 2014 and in revised form June 12, 2015)