COMPOSITIO MATHEMATICA

Effective limit distribution of the Frobenius numbers

Han Li

Compositio Math. **151** (2015), 898–916.

doi:10.1112/S0010437X14007866

FOUNDATION COMPOSITIO MATHEMATICA

LONDON MATHEMATICAL SOCIETY 150 YEARS

Compositio Math. 151 (2015) 898–916
doi:10.1112/S0010437X14007866

Effective limit distribution of the Frobenius numbers

Han Li

Abstract

The Frobenius number $F(\boldsymbol{a})$ of a lattice point $\boldsymbol{a}$ in $\mathbb{R}^d$ with positive coprime coordinates, is the largest integer which can *not* be expressed as a non-negative integer linear combination of the coordinates of $\boldsymbol{a}$. Marklof in [*The asymptotic distribution of Frobenius numbers*, Invent. Math. **181** (2010), 179–207] proved the existence of the limit distribution of the Frobenius numbers, when $\boldsymbol{a}$ is taken to be random in an enlarging domain in $\mathbb{R}^d$. We will show that if the domain has piecewise smooth boundary, the error term for the convergence of the distribution function is at most a polynomial in the enlarging factor.

1. Introduction

1.1 Some results before 1980

Let $\widehat{\mathbb{Z}}^d = \{\boldsymbol{a} = (a_1, \ldots a_d) \in \mathbb{Z}^d : \gcd(a_1, \ldots, a_d) = 1\}$ be the set of primitive lattice points, and $\widehat{\mathbb{Z}}^d_{\geqslant 2}$ be the subset of $\widehat{\mathbb{Z}}^d$ with coordinates $a_i \geqslant 2$. For any $\boldsymbol{a} \in \widehat{\mathbb{Z}}^d_{\geqslant 2}$, there exists a largest natural number $F(\boldsymbol{a})$ that is not representable as a linear non-negative integer combination of the coordinates of $\boldsymbol{a}$. The number $F(\boldsymbol{a})$ is called the Frobenius number of the vector $\boldsymbol{a}$, i.e.

$$F(\boldsymbol{a}) = \max \mathbb{N} \backslash \{\boldsymbol{k} \cdot \boldsymbol{a} : \boldsymbol{k} = (k_1, \ldots k_d) \in \mathbb{Z}^d, k_i \geqslant 0\}.$$

The Frobenius number problem is also known as ‘the Coin Exchange Problem’. It has been studied extensively in the past 50 years using various techniques, such as combinatorics, probabilistic methods, geometry of numbers, and more recently homogeneous dynamics. The book [Ram05] contains a good amount of information on the study of the Frobenius numbers. It is known that $F(\boldsymbol{a}) = a_1 a_2 - a_1 - a_2$ for $d = 2$, and no explicit formula is known when $d \geqslant 3$. Several upper bounds of the Frobenius numbers were obtained by the 1980s. With $\boldsymbol{a} \in \widehat{\mathbb{Z}}^d_{\geqslant 2}$ and $a_1 < a_2 < \cdots < a_d$, the estimates include the work by Erdős and Graham [EG72]

$$F(\boldsymbol{a}) \leqslant 2a_{d-1} \left[\frac{a_d}{d} \right] - a_d, \tag{1.1}$$

and the work by Selmer [Sel77]

$$F(\boldsymbol{a}) \leqslant 2a_d \left[\frac{a_1}{d} \right] - a_1. \tag{1.2}$$

Received 24 April 2013, accepted in final form 29 April 2014, published online 22 December 2014.
2010 Mathematics Subject Classification 37A25, 11K31 (primary), 11B57, 11D07, 11H06 (secondary).
Keywords: Frobenius numbers, Farey fractions, homogeneous dynamics, effective equidistribution.

1.2 The average behavior of the Frobenius numbers

There are also results on the limit distribution of Frobenius numbers from different perspectives. In dimension $d = 3$ using continued fractions, Bourgain and Sinai [BS07] showed that for ensembles $\Omega_N = \{a \in \widehat{\mathbb{Z}}^d_{\geqslant 2} : a_i \leqslant N\}$, the limit distribution of $F(\boldsymbol{a})/N^{3/2}$ exists. Marklof in [Mar10] generalized their result to higher dimensions. Before stating his results, let us first recall the notion of the covering radius. A lattice L in $\mathbb{R}^{d-1}$ is a discrete additive subgroup of $\mathbb{R}^{d-1}$ with finite covolume $\det(L)$, which equals the volume of the fundamental domain of the L action on $\mathbb{R}^{d-1}$. The covering radius Q_0 of the standard simplex $\Delta_{d-1} = \{\boldsymbol{x} \in \mathbb{R}^{d-1} : x_i \geqslant 0, \sum_{i=1}^{d-1} x_i \leqslant 1\}$ with respect to the lattice $L \subset \mathbb{R}^{d-1}$ is by definition

$$Q_0(L) = \inf\{t \in \mathbb{R}^+ : t\Delta_{d-1} + L = \mathbb{R}^{d-1}\}. \tag{1.3}$$

Lattices of covolume 1 are called unimodular. Let $G_0 = \mathrm{SL}_{d-1}(\mathbb{R})$, $\Gamma_0 = \mathrm{SL}_{d-1}(\mathbb{Z})$; then $\Omega_0 = G_0/\Gamma_0$ can be identified with the space of unimodular lattices in $\mathbb{R}^{d-1}$ ($g\Gamma_0 \leftrightarrow g\mathbb{Z}^{d-1}$). Let us fix a right invariant Riemannian metric on G_0, giving rise to a metric and a left G_0-invariant probability measure $\bar{\mu}_0$ on Ω_0. Let us recall the following theorem.

Theorem 1.1 [Mar10, Theorem 1]. *Let $d \geqslant 3$ and $E_R = \{L \in \Omega_0 : Q_0(L) \leqslant R\}$. Then:*

(i) *for any bounded set $\mathcal{D} \subset \mathbb{R}^d_{>0}$ with boundary of Lebesgue measure zero, and any $R \geqslant 0$,*

$$\lim_{T\to\infty} \frac{1}{T^d} \#\left\{\boldsymbol{a} \in \widehat{\mathbb{Z}}^d_{\geqslant 2} \cap T\mathcal{D} : \frac{F(\boldsymbol{a})}{(a_1 \cdots a_d)^{1/(d-1)}} \leqslant R\right\} = \frac{\mathrm{vol}(\mathcal{D})}{\zeta(d)} \bar{\mu}_0(E_R); \tag{1.4}$$

(ii) *Q_0 is a continuous function on Ω_0;*

(iii) *$\bar{\mu}_0(E_R)$ is continuous in R, i.e. $\bar{\mu}_0(\{L \in \Omega_0 : Q_0(L) = R\}) = 0$ for any $R > 0$.*

1.3 Some explanation of Marklof's work

We now briefly explain the existence of the limit distribution based on our private correspondence with Marklof. This is what we are going to follow in this paper which is more suitable for the purpose of 'effectivization', and is slightly different from [Mar10]. The method is based on homogeneous dynamics, which is also combined with the geometry of numbers. Here are the main ideas: Aliev and Gruber showed in [AG07] that for any $\boldsymbol{a} \in \widehat{\mathbb{Z}}^d_{\geqslant 2}$, one associates a $d-1$ dimensional unimodular lattice $L_{\boldsymbol{a}} \in \Omega_0$ (see Theorem 2.6) with

$$\frac{F(\boldsymbol{a}) + \sum_{i=1}^d a_i}{(a_1 \cdots a_d)^{1/(d-1)}} = Q_0(L_{\boldsymbol{a}}). \tag{1.5}$$

This is essentially due to a geometric interpretation of the Frobenius numbers found by Kannan [Kan92]. For this reason it will be more convenient for us to work with $F(\boldsymbol{a}) + \sum_{i=1}^d a_i$ instead of $F(\boldsymbol{a})$. It is easy to see that the result in (1.4) is not affected if $F(\boldsymbol{a})$ is replaced by $F(\boldsymbol{a}) + \sum_{i=1}^d a_i$. Notice that for every bounded connected non-empty open subset $\mathcal{D} \subseteq \{\boldsymbol{x} \in \mathbb{R}^d : 0 < x_d < 1;\ 0 < x_i < x_d\ (i = 1, \ldots, d-1)\}$ with boundary of Lebesgue measure zero,

$$|T\mathcal{D} \cap \widehat{\mathbb{Z}}^d| \sim \frac{T^d \mathrm{vol}(\mathcal{D})}{\zeta(d)}. \tag{1.6}$$

As we will see in § 4, the set of lattices $\{L_{\boldsymbol{a}} : \boldsymbol{a} \in T\mathcal{D} \cap \widehat{\mathbb{Z}}^d\}$ appearing in (1.5) becomes equidistributed in Ω_0 as $T \to \infty$, i.e. for any bounded continuous function ϕ on Ω_0, we have

$$\lim_{T\to\infty} \frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D} \cap \widehat{\mathbb{Z}}^d} \phi(L_a) = \frac{\mathrm{vol}(\mathcal{D})}{\zeta(d)} \int_{\Omega_0} \phi \, d\bar{\mu}_0. \tag{1.7}$$

https://doi.org/10.1112/S0010437X14007866 Published online by Cambridge University Press

As $E_R = \{L \in \Omega_0 : Q_0(L) \leqslant R\}$ has boundary of measure zero, with a standard approximation argument Theorem 1.1 can be deduced by applying χ_{E_R} to (1.7) in the place of ϕ.

Theorem 1.1 also implies that for large enough R, and large enough T (depending on the choice of R), the probability that a random lattice point $\boldsymbol{a} \in T\mathcal{D} \cap \widehat{\mathbb{Z}}^d$ satisfies $F(\boldsymbol{a}) < R(a_1 \cdots a_d)^{1/(d-1)}$ is greater than 99%. This gives a somewhat better estimate compared with (1.1) and (1.2), for most $\boldsymbol{a} \in T\mathcal{D} \cap \widehat{\mathbb{Z}}^d$.

1.4 Statement of the results

The aim of this paper is to estimate the decay of the function $\Psi(R) = 1 - \bar{\mu}_0(E_R)$ and the error term of (1.4).

Theorem 1.2. *There exists a constant $C > 0$ dependent only on d, such that for any $R > 0$ we have $\Psi(R) < CR^{-(d-1)}$.*

Theorem 1.2 improves the exponent compared with [AHH11, Theorem 1]. After this paper was completed, Marklof, in an unpublished work, proved that there exists a constant $c_d > 0$, so that $\Psi(R) > c_d R^{-(d-1)}$. Therefore our bound is actually sharp. An asymptotic formula for $\Psi(R)$ has recently been obtained by Strömbergsson in [Str12].

Definition 1.3. Let M be a connected smooth manifold. We say that a subset of M has thin boundary if the boundary is contained in a union of finitely many connected smooth submanifolds of M, and each of them has dimension $< \dim M$.

Theorem 1.4. *There exists $\kappa > 0$ dependent only on the dimension d satisfying the following property. For any $R > 0$, and any non-empty connected open subset $\mathcal{D} \subseteq \{\boldsymbol{x} \in \mathbb{R}^d : 0 < x_d < 1; 0 < x_i < x_d\ (i = 1, \ldots, d-1)\}$ which has thin boundary as a subset of the manifold $\mathbb{R}^d$, there exists a constant $C_{R,\mathcal{D}} > 0$ such that for every $T \geqslant 1$*

$$\left| \frac{1}{T^d} \# \left\{ \boldsymbol{a} \in T\mathcal{D} \cap \widehat{\mathbb{Z}}^d_{\geqslant 2} : \frac{F(\boldsymbol{a}) + \sum_{i=1}^d a_i}{(a_1 \cdots a_d)^{1/(d-1)}} \leqslant R \right\} - \frac{\mathrm{vol}(\mathcal{D})}{\zeta(d)} \bar{\mu}_0(E_R) \right| < \frac{C_{R,\mathcal{D}}}{T^\kappa}.$$

When $d = 3$ Ustinov [Ust10] has proved the existence of the limit distribution in Theorem 1.1 (with the average taken over only two of the three variables) and the polynomial convergence to the limit distribution. He has also obtained an explicit formula for the limit density function (see also [Str12, pp. 2–3] for a discussion).

1.5 Organization of the paper

In § 2 we will use the geometry of numbers to prove Theorem 1.2. We will also give an explicit description for the $L_{\boldsymbol{a}}$ appearing in formula (1.5). Sections 3 and 4 are devoted to proving effective equidistribution of the expanding horospheres, and a Farey sequence on a specified closed horosphere, respectively, under the translation of a one-sided diagonal flow. We will give an error term estimate of (1.7) for non-negative compactly supported C^1 test functions (Theorem 4.13). Using a standard approximation argument, we will prove Theorem 1.4 in § 5 by showing that $E_R = \{L \in \Omega_0 : Q_0(L) \leqslant R\}$ has thin boundary as a subset of Ω_0. We will borrow many ideas from [Mar10, § 4] to formulate a series of equidistribution results which lead to Theorem 4.13.

1.6 Notational convention

Throughout the paper we assume that the dimension $d \geqslant 3$, and always work with column vectors in Euclidean spaces. We use $A \ll B$ to represent inequalities $0 \leqslant A \leqslant cB$ in which the implicit constant c depends on the underlying Lie groups or Euclidean spaces. In a metric space X,

https://doi.org/10.1112/S0010437X14007866 Published online by Cambridge University Press

$B_X(x,r)$ stands for the open ball of radius r centered at x. On a Lie group G, $B_G(r) = B_G(e,r)$ with a specified metric on G; in $\mathbb{R}^n$, $B(r) = B_{\mathbb{R}^n}(0,r)$ with Euclidean norm. The exponents $\alpha_1, \alpha_2, \ldots$ in §§ 3 and 4 depend only on the dimension d.

2. Covering radius and the Frobenius numbers

2.1 Preliminaries

We call a subset K of $\mathbb{R}^{d-1}$ a *convex body* if K is a compact convex set with non-empty interior. A convex body is called centrally symmetric if it is symmetric with respect to the origin. For a centrally symmetric convex body K, its polar K^* is also a centrally symmetric convex body defined by $K^* = \{\boldsymbol{x} \in \mathbb{R}^{d-1} : \boldsymbol{x} \cdot \boldsymbol{y} \leqslant 1, \text{for any } \boldsymbol{y} \in K\}$.

We now recall the notion of *dual lattice*. Let $L = A\mathbb{Z}^{d-1} \in \Omega_0$ where $A \in G_0$, and let A^* be the inverse transpose of A. We call the lattice $L^* = A^*\mathbb{Z}^{d-1}$ the dual lattice of L. One can readily verify that the definition of L^* is independent of the choice A, and moreover the map $L \to L^*$ is a diffeomorphism of Ω_0 which preserves $\bar{\mu}_0$.

The covering radius $Q(K,L)$ of $K \subseteq \mathbb{R}^{d-1}$ with respect to a lattice L in $\mathbb{R}^{d-1}$ is defined by

$$Q(K,L) := \inf\{t \in \mathbb{R}^+ : tK + L = \mathbb{R}^{d-1}\}.$$

Clearly the function Q_0 defined in (1.3) satisfies $Q_0(L) = Q(\Delta_{d-1}, L)$ for any lattice L in $\mathbb{R}^{d-1}$. (We will abbreviate Δ_{d-1} as Δ in what follows.)

2.2 Minkowski's successive minima and the covering radius

The covering radius is related to *Minkowski's successive minima*. Let $K \subseteq \mathbb{R}^{d-1}$ be a centrally symmetric convex body, and L a lattice in $\mathbb{R}^{d-1}$. The ith minimum ($1 \leqslant i \leqslant d-1$) of K with respect to L is defined by

$$\lambda_i(K,L) := \min\{t \in \mathbb{R}^+ : \dim(\operatorname{span}(tK \cap L)) \geqslant i\}.$$

Clearly $0 < \lambda_1(K,L) \leqslant \lambda_2(K,L) \cdots \leqslant \lambda_{d-1}(K,L)$.

LEMMA 2.1 (Minkowski's second theorem). *Let $K \subseteq \mathbb{R}^{d-1}$ be a centrally symmetric convex body and L be a lattice in $\mathbb{R}^{d-1}$. Then*

$$\frac{2^{d-1}}{(d-1)!} \leqslant \frac{\operatorname{vol}(K)}{\det(L)} \prod_{i=1}^{d-1} \lambda_i(K,L) \leqslant 2^{d-1}.$$

LEMMA 2.2 [KL88, Lemma 2.4, Corollary 2.8]. *Let K be a convex body and L be a lattice in $\mathbb{R}^{d-1}$, and set $K - K = \{\boldsymbol{k}_1 - \boldsymbol{k}_2 : \boldsymbol{k}_1, \boldsymbol{k}_2 \in K\}$. Then:*

(i) $\lambda_{d-1}(K-K,L) \leqslant Q(K,L) \leqslant \sum_{i=1}^{d-1} \lambda_i(K-K,L)$;

(ii) *there exists a constant $C_d > 0$, so that $\lambda_{d-1}(K-K,L)\lambda_1((K-K)^*,L^*) < C_d$.*

LEMMA 2.3. *The function Q_0 defined in (1.3) is proper on Ω_0, i.e. $E_R = \{L \in \Omega_0 : Q_0(L) \leqslant R\}$ is a compact subset of Ω_0 for any $R \geqslant 0$.*

Proof. By Lemmas 2.1 and (i) of 2.2, $\lambda_1(\Delta - \Delta, L)$ is positively bounded below for $L \in E_R$. Mahler's Criterion implies that E_R is relatively compact in Ω_0. Since Q_0 is continuous (see (ii) of Theorem 1.1), E_R is compact. □

LEMMA 2.4 [AM09, Lemma 4.1]. *For any centrally symmetric convex body K in $\mathbb{R}^{d-1}$, there exists a constant $C_K > 0$ so that for any $r > 0$,*

$$\bar{\mu}_0(\{L \in \Omega_0 : \lambda_1(K,L) < r\}) < C_K r^{d-1}.$$

https://doi.org/10.1112/S0010437X14007866 Published online by Cambridge University Press

2.3 Proof of Theorem 1.2

Proof. By Lemma 2.2, for any $R > 0$,

$$\{L \in \Omega_0 : Q_0(L) > R\} \subseteq \left\{L \in \Omega_0 : \lambda_{d-1}(\Delta - \Delta, L) > \frac{R}{d-1}\right\}$$
$$\subseteq \left\{L \in \Omega_0 : \lambda_1((\Delta - \Delta)^*, L^*) < \frac{(d-1)C_d}{R}\right\}.$$

Since the map $L \mapsto L^*$ preserves $\bar{\mu}_0$, by Lemma 2.4

$$\Psi(R) = \bar{\mu}_0(\{L \in \Omega_0 : Q_0(L) > R\}) \ll_d R^{-(d-1)}. \qquad \square$$

2.4 An interpretation of a result of Kannan and Aliev–Gruber

For $T > 0$, $\boldsymbol{x} \in \mathbb{R}^{d-1}$ and $\boldsymbol{y} \in \mathbb{R}^{d-1}$ with each coordinate $y_i \neq 0$, we define

$$D(T) = \begin{pmatrix} T^{-1/(d-1)} 1_{d-1} & 0 \\ 0 & T \end{pmatrix} \quad n(\boldsymbol{x}) = \begin{pmatrix} 1_{d-1} & 0 \\ \boldsymbol{x}^t & 1 \end{pmatrix} \quad m(\boldsymbol{y}) = \begin{pmatrix} m'(\boldsymbol{y}) & 0 \\ 0 & 1 \end{pmatrix}$$

where

$$m'(\boldsymbol{y}) = (y_1 \cdots y_{d-1})^{-1/(d-1)} \operatorname{diag}(y_1, \ldots, y_{d-1}).$$

It is clear that $D(T)$, $n(\boldsymbol{x})$, and $m(\boldsymbol{y})$ all belong to $\mathrm{SL}_d(\mathbb{R})$.

Definition 2.5. For every $\boldsymbol{a} \in \mathbb{Z}^d$ with $a_d \neq 0$, we associate a vector $\widehat{\boldsymbol{a}} \in \mathbb{R}^{d-1}$ by

$$\widehat{\boldsymbol{a}} = \left(\frac{a_1}{a_d}, \ldots, \frac{a_{d-1}}{a_d}\right)^t.$$

For any $\boldsymbol{a} \in \widehat{\mathbb{Z}}^d_{\geqslant 2}$ let $M_{\boldsymbol{a}}$ be the lattice

$$M_{\boldsymbol{a}} := \left\{\boldsymbol{b} \in \mathbb{Z}^{d-1} : \sum_{i=1}^{d-1} a_i b_i \equiv 0 \ (\mathrm{mod}\ a_d)\right\}.$$

Since $\boldsymbol{a}$ is primitive, $M_{\boldsymbol{a}}$ has determinant a_d. Aliev and Gruber, based on the work of Kannan [Kan92], have shown in [AG07] that the Frobenius number $F(\boldsymbol{a})$ satisfies

$$\frac{F(\boldsymbol{a}) + \sum_{i=1}^d a_i}{(a_1 \cdots a_d)^{1/(d-1)}} = Q_0(m'(\widehat{\boldsymbol{a}})(a_d^{-1/(d-1)} M_{\boldsymbol{a}})). \tag{2.1}$$

This enables us to present an explicit description of the lattice $L_{\boldsymbol{a}}$ in formula (1.5).

Theorem 2.6. *For any $\boldsymbol{a} \in \widehat{\mathbb{Z}}^d_{\geqslant 2}$ the Frobenius number $F(\boldsymbol{a})$ satisfies*

$$\frac{F(\boldsymbol{a}) + \sum_{i=1}^d a_i}{(a_1 \cdots a_d)^{1/(d-1)}} = Q_0(m(\widehat{\boldsymbol{a}}) D(a_d) n(\widehat{\boldsymbol{a}}) \mathbb{Z}^d \cap \boldsymbol{e}_d^{\perp}), \tag{2.2}$$

where $\boldsymbol{e}_d = (0, \ldots, 0, 1)^t$, and we identify $\mathbb{R}^d \cap \boldsymbol{e}_d^{\perp}$ with $\mathbb{R}^{d-1}$ in the obvious way. In other words, $L_{\boldsymbol{a}} = m(\widehat{\boldsymbol{a}}) D(a_d) n(\widehat{\boldsymbol{a}}) \mathbb{Z}^d \cap \boldsymbol{e}_d^{\perp}$.

Proof. Note that for any $\boldsymbol{y} = (y_1, \ldots, y_d)^t \in \mathbb{R}^d$ with $\boldsymbol{y} \cdot \boldsymbol{a} = 0$, we have

$$n(\widehat{\boldsymbol{a}}) \boldsymbol{y} = (y_1, \ldots, y_{d-1}, 0)^t.$$

Therefore $M_{\boldsymbol{a}} = (n(\widehat{\boldsymbol{a}}) \mathbb{Z}^d) \cap \boldsymbol{e}_d^{\perp}$. The conclusion follows immediately from (2.1). $\square$

https://doi.org/10.1112/S0010437X14007866 Published online by Cambridge University Press

3. Translations of horospheres and effective equidistribution

3.1 Subgroups of $\mathrm{SL}_d(\mathbb{R})$ and their Haar measures

Let us fix the notation for the Lie groups which will be frequently used in what follows:

$$G = \mathrm{SL}_d(\mathbb{R}), \quad \Gamma = \mathrm{SL}_d(\mathbb{Z}), \quad G_0 = \mathrm{SL}_{d-1}(\mathbb{R}).$$

We will identify G_0 with the image of the embedding $A \to \left(\begin{smallmatrix} A & 0 \\ 0 & 1 \end{smallmatrix}\right)$. We denote by $F = \{D(s) : s > 0\}$ the subgroup of G, and set $F^+ = \{D(s) : s > 1\}$. For the subgroups of G

$$H = \{n(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{R}^{d-1}\}, \quad H^- = \{n(\boldsymbol{x})^t : \boldsymbol{x} \in \mathbb{R}^{d-1}\}, \quad H_0 = \left\{ \begin{pmatrix} A & 0 \\ 0 & c \end{pmatrix} : \det(A)c = 1 \right\},$$

their Lie algebras are invariant subspaces of the adjoint action of F on $\mathfrak{g} = \mathrm{Lie}(G)$.

We identity the Lie algebra $\mathfrak{g}$ with the space of $d \times d$ traceless matrices, and define an inner product on $\mathfrak{g}$ by setting

$$\langle X, Y \rangle = \mathrm{tr}(X^t Y), \quad X, Y \in \mathfrak{g}.$$

This gives rise to a right invariant Riemannian metric on G, and hence a metric d_S on any closed subgroup S of G. We have an orthogonal decomposition of $\mathfrak{g}$ into linear subspaces

$$\mathfrak{g} = \mathrm{Lie}(H) + \mathrm{Lie}(H^-) + \mathrm{Lie}(G_0) + \mathrm{Lie}(F).$$

We fix an orthonormal basis of $\mathfrak{g}$ coming from a basis of those subspaces

$$\mathcal{X} = \{X_i : i = 1, 2, \ldots, d^2 - 1\}. \tag{3.1}$$

We define for every $s > 0$ a map $\phi_s : G \to G$ by

$$\phi_s(g) = D(s) g D(s^{-1}). \tag{3.2}$$

The restriction of $\phi_s (s > 1)$ on $H' = H_0 H^-$ is thus contracting in the sense that, for any $r > 0$,

$$\phi_s(B_{H'}(r)) \subseteq B_{H'}(r).$$

The group H is called the expanding horospherical subgroup with respect to F^+ as

$$H = \{g \in G : D(s^{-1}) g D(s) \to 1_d, \text{as } s \to +\infty\}.$$

Let $\Omega = G/\Gamma$ with the metric d_Ω coming from G. Every H-orbit in Ω is called an expanding horosphere (with respect to F^+). We specify a closed horosphere

$$Y = \{h\Gamma : h \in H\} = \{n(\boldsymbol{x})\Gamma : \boldsymbol{x} \in \mathbb{T}^{d-1}\} \subseteq \Omega.$$

By μ and ν we denote the left Haar measures on G and H respectively, with the induced measures on Ω and Y satisfying $\bar{\mu}(\Omega) = 1$ and $\bar{\nu}(Y) = 1$ (which means ν and $\bar{\nu}$ correspond to the Lebesgue measures on $\mathbb{R}^{d-1}$ and $\mathbb{T}^{d-1}$ respectively). We choose a left Haar measure ν' on H' so that μ is locally the product of ν and ν'. This means, in view of [Kna02, Theorem 8.32], for any $f \in L^1(G)$:

$$\int_{H'H} f(g)\, d\mu = \int_{H' \times H} f(h'h)\, d\nu'(h')\, d\nu(h). \tag{3.3}$$

https://doi.org/10.1112/S0010437X14007866 Published online by Cambridge University Press

3.2 Decay of matrix coefficients and its consequences

A rich literature on the theory of 'the decay of the matrix coefficients' has evolved since the work of Harish-Chandra, including the works by Cowling [Cow79], Howe [How82], Moore [Moo87], Katok and Spatzier [KS94], and Oh [Oh02], to just mention a few. Let ρ be a (strongly continuous) unitary representation of G on a Hilbert space $\mathcal{H}$. We say that a vector $v \in \mathcal{H}$ is Lipschitz if (the metric d below refers to the fixed metric on G)

$$\mathrm{Lip}(v) := \sup\left\{ \frac{\|\rho(g)v - v\|}{d(e,g)} : g \neq e \right\} < \infty.$$

Based on previous works on the decay of the matrix coefficients, Kleinbock and Margulis proved a quantitative decay of matrix coefficients for Lipschitz vectors. For our purpose we only need the following theorem which follows from [KM96, Theorem A.4], combined with Kazhdan's property (T) for the groups $G = \mathrm{SL}_n(\mathbb{R})$ ($n \geqslant 3$).

Theorem 3.1. *There exists $\alpha_1 > 0$ so that for any unitary representation $(\rho, \mathcal{H})$ of G without G-invariant vectors, any Lipschitz vectors $v, w \in \mathcal{H}$ and every $s > 1$, we have that*

$$|\langle \rho(D(s))v, w\rangle| \ll s^{-\alpha_1}(\mathrm{Lip}(v) + \|v\|)(\mathrm{Lip}(w) + \|w\|).$$

Remark 3.2. Following [KM96, Theorem A.4] and the work of Oh [Oh02] one can give an explicit exponent α_1 in Theorem 3.1. It should in principle give an explicit exponent κ in Theorem 1.4 following the results in §§ 3 and 4. However, we shall not do this, to keep the computation simpler in this article.

Definition 3.3. We say that a real-valued function ψ on a metric space X (with metric 'dist') is Lipschitz if

$$\|\psi\|_{\mathrm{Lip}} := \sup\left\{ \frac{|\psi(x) - \psi(y)|}{\mathrm{dist}(x,y)} : x, y \in X, x \neq y \right\} < \infty.$$

The space of Lipschitz functions on X is denoted by $\mathrm{Lip}(X)$.

Remark 3.4. If X is a Riemannian manifold with distance coming from the Riemannian metric and ψ a real-valued smooth function on X, then $\|\psi\|_{\mathrm{Lip}} = \sup\{\|d\psi_x\| : x \in X\}$, where $d\psi_x$ is the tangent map of ψ at x, and its norm comes from the Riemannian metric on X.

We specify a metric on $H \times \Omega$ by setting $d((h_1, z_1), (h_2, z_2)) := d_H(h_1, h_2) + d_\Omega(z_1, z_2)$, where $h_1, h_2 \in H$ and $z_1, z_2 \in \Omega$. In what follows, the metrics on the product spaces are all defined in the same way. We now fix a subset

$$U = \{n(\boldsymbol{x}) : \boldsymbol{x} \in (-2, 2)^{d-1}\} \subset H.$$

Consider the action of G on $H \times \Omega$ by $g.(h, z) = (h, gz)$ and the associated unitary representation of G on $L^2(H \times \Omega)$. In this case any Lipschitz function on $H \times \Omega$ which is supported on $U \times \Omega$ is a Lipschitz vector in $L^2(H \times \Omega)$, and moreover $\mathrm{Lip}(\psi) \ll_U \|\psi\|_{\mathrm{Lip}}$. Notice that $\mathcal{H}_0 = \{\varphi : \varphi(h, z) = f(h),\ \text{for some } f \in L^2(H)\}$ is the linear subspace of the G-stable vectors in $L^2(H \times \Omega)$. Considering the representation of G on $\mathcal{H}_0^\perp$, we obtain the following corollary.

Corollary 3.5. *Let $\mathcal{P} : L^2(H \times \Omega) \to \mathcal{H}_0$ be the orthogonal projection. Then for any $s > 1$ and functions $\varphi, \psi \in \mathrm{Lip}(H \times \Omega) \cap L^2(H \times \Omega)$ which are supported on $U \times \Omega$, we have that*

$$|\langle D(s)\varphi, \psi\rangle - \langle \mathcal{P}\varphi, \mathcal{P}\psi\rangle| \ll_U (\|\varphi\|_{L^2} + \|\varphi\|_{\mathrm{Lip}})(\|\psi\|_{L^2} + \|\psi\|_{\mathrm{Lip}})s^{-\alpha_1}. \tag{3.4}$$

https://doi.org/10.1112/S0010437X14007866 Published online by Cambridge University Press

3.3 Effective equidistribution and F^+-translations

Definition 3.6. Let M be a smooth manifold on which G acts by diffeomorphisms. We define for every $X \in \mathfrak{g}$ a vector field on M by

$$\partial X(f)(m) := \lim_{t\to 0} \frac{f(\exp(tX)m) - f(m)}{t} \quad \text{for all } f \in C^\infty(M), m \in M.$$

The corresponding tangent vector $\partial_m X$ at $m \in M$ is by definition

$$\partial_m X(f) = \partial X(f)(m) \quad \text{for all } f \in C^\infty(M).$$

We are going to present a quantitative equidistribution result of the F^+-translations of the H-orbit $\{(h, hx) : h \in H\}$ (where $x \in \Omega$) in $H \times \Omega$. The method in our approach is by no means new. The proof here is a modification of the proofs for the equidistribution of the F^+-translations of $\{hx : h \in H\}$ in Ω (cf. [KM96, KM12]). The technique is sometimes known as 'equidistribution via mixing', which originated in Margulis' thesis. In contrast to [KM96, KM12], the additional variable on the horosperical subgroup H deserves special care when we do the 'thickening'. First we recall a well-known result.

Lemma 3.7. *For any $0 < r < 1$, there exists a non-negative function $\theta \in C^\infty(\mathbb{R}^n)$ supported in $B(r)$, such that $\theta(\mathbf{0}) = 1$, $\int_{\mathbb{R}^n} \theta = 1$, $\|\theta\|_{L^2} \ll r^{-n/2}$, $\|\theta\|_{L^\infty} \ll r^{-n}$ and $\|\theta\|_{\mathrm{Lip}} \ll r^{-n-1}$.*

We remark that the lemma will be used in what follows to thicken the functions, which are defined on submanifolds, in the transversal directions so as to get new functions defined on larger spaces which also satisfy certain norm bounds. The condition $\theta(\mathbf{0}) = 1$ in the lemma will be used specifically in the proof of Theorem 4.5.

Theorem 3.8. *Let $f \in C^\infty(H)$, $0 < r < 1$ be such that $B_H(r)\,\mathrm{supp}(f) \subset U$, and let $x \in \Omega$ be such that $\pi_x : G \to \Omega, \pi_x(g) = gx$ is injective on $B_G(r)\,\mathrm{supp}(f)$. Then for any $T > 1$ and $\varphi \in C^\infty(H \times \Omega)$ with $\mathrm{supp}(\varphi) \subseteq U \times \Omega$, we have that (the α_1 below is as in Theorem 3.1)*

$$\left| \int_H f(h)\varphi(h, D(T)hx)\, d\nu(h) - \int_{H\times\Omega} f(h)\varphi(h, z)\, d\nu(h)\, d\bar{\mu}(z) \right| \\ \ll \|\varphi\|_{\mathrm{Lip}} \cdot r \cdot \|f\|_{L^1} + r^{-d^2} \|f\|_{C^1} \|\varphi\|_{C^1} T^{-\alpha_1}. \tag{3.5}$$

The C^1 norms here for smooth functions on various manifolds are taken to be $\|f\|_\infty + \|f\|_{\mathrm{Lip}}$.

Remark 3.9. We assume here that $\|\varphi\|_{C^1} < \infty$, as there is nothing to prove otherwise. The same assumption applies to Theorems 3.11, 4.5, Proposition 4.6, Corollary 4.9 and Proposition 4.13.

Proof. Replacing φ by $\varphi(h, z) - \int_\Omega \varphi(h, z)\, d\bar{\mu}(z)$ if necessary, we may assume $\int_\Omega \varphi(h, z)\, d\bar{\mu}(z) = 0$ for every $h \in H$. We choose non-negative functions $\theta' \in C^\infty(H')$, $\theta_1 \in C^\infty(H)$ supported in $B_{H'}(r)$ and $B_H(r)$ respectively by Lemma 3.7. We define a function ψ on $H \times G$ by setting

$$\psi(h_1h_2, h'h_2) = \theta'(h')\theta_1(h_1)f(h_2) \quad \text{for all } h_1, h_2 \in H, h' \in H',$$

and setting $\psi(h, g) = 0$ outside the open subset $H \times H'H$ of $H \times G$. Since $\int_{H'} \theta' = \int_H \theta_1 = 1$, [Kna02, Theorem 8.32] and formula (3.3) implies

$$\int_H f(h)\varphi(h, D(T)hx)\, d\nu(h) = \int_{H\times H'\times H} \psi(h_1, h'h_2)\varphi(h_2, D(T)h_2x)\, d\nu(h_1)\, d\nu'(h')\, d\nu(h_2).$$

Let us define a function ψ_x on $H \times \Omega$ by setting $\psi_x(h, gx) = \psi(h, g)$ for $(h, g) \in H \times (B_G(r)\,\mathrm{supp}(f))$, and $\psi_x(h, z) = 0$ outside the open subset $H \times B_G(r)\,\mathrm{supp}(f)x \subseteq H \times \Omega$. The definition makes sense because of the injectivity assumption. It is easy to check that $\psi_x \in C^\infty(H \times \Omega)$, and $\mathrm{supp}(\psi_x) \subseteq U \times \Omega$. As the maps ϕ_s defined in (3.2) are contractions on H', one has that

$$\begin{aligned}
&\left| \int_H f(h)\varphi(h, D(T)hx)\,d\nu(h) - \langle D(T)\psi_x, \varphi\rangle \right| \\
&\quad = \left| \int_{H\times H'\times H} \psi(h_1, h'h_2)(\varphi(h_2, D(T)h_2x) - \varphi(h_1, D(T)h'h_2x))\,d\nu(h_1)\,d\nu'(h')\,d\nu(h_2) \right| \\
&\quad \ll \|\varphi\|_{\mathrm{Lip}} \cdot r \cdot \|\psi\|_{L^1} \\
&\quad = \|\varphi\|_{\mathrm{Lip}} \cdot r \cdot \|f\|_{L^1}.
\end{aligned} \tag{3.6}$$

On the other hand

$$\|\psi_x\|_{\mathrm{Lip}} \ll \|\theta'(h')\theta_1(h_1)f(h_2)\|_{C^1} \ll r^{-d^2}\|f\|_{C^1}, \quad \|\psi_x\|_{L^2} \ll r^{-d^2/2}\|f\|_{L^2}.$$

Let $\mathcal{P} : L^2(H \times \Omega) \to \mathcal{H}_0$ be the orthogonal projection as in Corollary 3.5. Then

$$\mathcal{P}(\varphi) = \int_\Omega \varphi(h, z)\,d\bar{\mu}(z) = 0$$

by our assumption. It follows from (3.4) that

$$|\langle D(T)\psi_x, \varphi\rangle| \ll r^{-d^2}\|f\|_{C^1}(\|\varphi\|_{\mathrm{Lip}} + \|\varphi\|_{L^2})T^{-\alpha_1} \ll r^{-d^2}\|f\|_{C^1}\|\varphi\|_{C^1}T^{-\alpha_1}.$$

The theorem follows immediately. □

Remark 3.10. It is the estimate (3.6) that makes us use the Lipschitz norm in (3.5), instead of the 'Sobolev norm' which is very common in the recent literature.

The following theorem concerns the equidistribution of F^+-translations of the Lebesgue measure on $\{(\boldsymbol{x}, n(\boldsymbol{x})\Gamma) : \boldsymbol{x} \in I^{d-1}\}$ where $I = (0, 1)$. The result without the additional variable on I^{d-1} is the classical equidistribution of large closed horospheres. The reason that we also consider a variable on I^{d-1} here is that the matrix $m(\boldsymbol{x})$, which is related to Frobenius numbers via Theorem 2.6, is defined for $\boldsymbol{x} \in I^{d-1}$. This will be relevant in Proposition 4.6.

Theorem 3.11. *There exists a constant $\alpha_2 > 0$ so that for any $\phi \in C^\infty(I^{d-1} \times \Omega)$ and $T > 1$,*

$$\left| \int_{I^{d-1}} \phi(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma)\,d\boldsymbol{x} - \int_{I^{d-1}\times\Omega} \phi(\boldsymbol{x}, z)\,d\boldsymbol{x}\,d\bar{\mu}(z) \right| \ll \|\phi\|_{C^1} T^{-\alpha_2}. \tag{3.7}$$

Here $d\boldsymbol{x}$ is the Lebesgue measure, and

$$\|\phi\|_{C^1} := \|\phi\|_{L^\infty} + \max\left\{ \left\| \frac{\partial}{\partial x_i}\phi \right\|_{L^\infty}, \|\partial X(\phi)\|_{L^\infty} \right\},$$

where $\partial/\partial x_i$ are the standard Euclidean vector fields for the I^{d-1} factor, and the $X \in \mathcal{X}$ are vector fields for the Ω factor. (See (3.1) and Definition 3.6.)

Proof. To outline the idea of the proof, we will approximate $\chi_{I^{d-1}}$, ϕ by smooth functions on $\mathbb{R}^{d-1}$, $\mathbb{R}^{d-1}\times\Omega$ respectively. This enables us to apply Theorem 3.8 and get an error term estimate.

Let's fix a partition $\{E_i : 1 \leqslant i \leqslant N\}$ of I^{d-1} with the interior of each E_i being an open cube, and choose $r_0 > 0$ so that for each i we have that $B_H(r_0)\{n(\boldsymbol{x}) : \boldsymbol{x} \in E_i\} \subseteq U$, and the restriction of $\pi : G \to \Omega,\ \pi(g) = g\Gamma$ to $\{gn(\boldsymbol{x}) : g \in B_G(r_0), \boldsymbol{x} \in E_i\}$ is injective. For every $0 < r < r_0$ and $1 \leqslant i \leqslant N$, we fix a function $p_i \in C^\infty(\mathbb{R}^{d-1})$ supported in E_i,

$$0 \leqslant p_i \leqslant \chi_{E_i}, \quad \mathrm{vol}(\{\boldsymbol{x} \in E_i : p_i(\boldsymbol{x}) \neq 1\}) \ll r, \quad \|p_i\|_{C^1} \ll r^{-1}.$$

We fix also a function $p \in C^\infty(\mathbb{R}^{d-1})$ supported in I^{d-1} with

$$0 \leqslant p \leqslant \chi_{I^{d-1}}, \quad \mathrm{vol}(\{\boldsymbol{x} \in I^{d-1} : p(\boldsymbol{x}) \neq 1\}) \ll r^{1/2}, \quad \|p\|_{C^1} \ll r^{-1/2}.$$

Letting $\tilde{\phi}(\boldsymbol{x}, z) = p(\boldsymbol{x})\phi(\boldsymbol{x}, z) \in C^\infty(\mathbb{R}^{d-1}\times\Omega)$, we then have

$$\left|\int_{\mathbb{R}^{d-1}} \chi_{E_i}(\boldsymbol{x})\phi(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma)\, d\boldsymbol{x} - \int_{\mathbb{R}^{d-1}} p_i(\boldsymbol{x})\tilde{\phi}(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma)\, d\boldsymbol{x}\right| \ll \|\phi\|_{L^\infty} r^{1/2},$$

$$\left|\int_{\mathbb{R}^{d-1}\times\Omega} (\chi_{E_i}(\boldsymbol{x})\phi(\boldsymbol{x}, z) - p_i(\boldsymbol{x})\tilde{\phi}(\boldsymbol{x}, z))\, d\boldsymbol{x}\, d\bar{\mu}(z)\right| \ll \|\phi\|_{L^\infty} r^{1/2}.$$

As $p_i(\boldsymbol{x})$ and $\tilde{\phi}(\boldsymbol{x}, z)$ satisfy the assumptions of Theorem 3.8, we have that

$$\begin{aligned}&\left|\int_{\mathbb{R}^{d-1}} p_i(\boldsymbol{x})\tilde{\phi}(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma)\, d\boldsymbol{x} - \int_{\mathbb{R}^{d-1}\times\Omega} p_i(\boldsymbol{x})\tilde{\phi}(\boldsymbol{x}, z)\, d\boldsymbol{x}\, d\bar{\mu}(z)\right| \\ &\quad \ll \|\tilde{\phi}\|_{\mathrm{Lip}} \cdot r + r^{-d^2-1}\|\tilde{\phi}\|_{C^1} T^{-\alpha_1} \\ &\quad \ll \|\phi\|_{\mathrm{Lip}} \cdot r + \|\phi\|_{L^\infty} r^{1/2} + r^{-d^2-3/2}\|\phi\|_{C^1} T^{-\alpha_1}.\end{aligned}$$

Setting $r = r_0 T^{-2\alpha_2}$ for some appropriate $\alpha_2 > 0$, we get (r_0 depends only on the dimension)

$$\left|\int_{I^{d-1}} \phi(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma)\, d\boldsymbol{x} - \int_{I^{d-1}\times\Omega} \phi(\boldsymbol{x}, z)\, d\boldsymbol{x}\, d\bar{\mu}(z)\right| \ll \|\phi\|_{C^1} T^{-\alpha_2}. \qquad \square$$

4. Translations of a Farey sequence and effective equidistribution

4.1 Description of the F^+-translations of a Farey sequence

The Farey fractions on the torus $\mathbb{T}^{d-1}$ are those points whose coordinates are rational numbers. We already know that the expanding horosphere $Y = \{h\Gamma : h \in H\}$ becomes equidistributed under F^+-translations. We are going to study the equidistribution property of Farey fractions on Y in this section. We denote by K the subgroup

$$K = \left\{A \ltimes \boldsymbol{b} := \begin{pmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{pmatrix} : A \in G_0, \boldsymbol{b} \in \mathbb{R}^{d-1}\right\} \subseteq G.$$

Let $\Lambda = \{D(s)k\Gamma : s > 1, k \in K\}$. This is an embedded submanifold of Ω by [Mar10, Lemma 2]. For any element $\lambda \in \Lambda$, there exist unique $s > 1$ and $z \in K\Gamma/\Gamma$ such that $\lambda = D(s)z$. Let

$$\mathcal{D}_0 = \{\boldsymbol{x} \in \mathbb{R}^d : 0 < x_d < 1;\ \forall\, i < d,\ 0 < x_i < x_d\}.$$

Marklof in [Mar10] proved that under F^+-translation, the Farey fractions $\{n(\widehat{\boldsymbol{a}})\Gamma : \boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d\}$ on the closed horosphere $\{n(\boldsymbol{x})\Gamma : \boldsymbol{x} \in \mathbb{T}^{d-1}\}$ become equidistributed on Λ. We will prove an effective version of this result in Theorem 4.5. The following lemma, which describes the behavior of F^+-translations of the Farey fractions, is also hinted at in [Mar10].

Lemma 4.1. *For any $T > 1$, the lattice points in $T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d$ are in one-to-one correspondence with the intersection of $\{D(T)n(\boldsymbol{x})\Gamma : \boldsymbol{x} \in I^{d-1}\}$ with the submanifold Λ. More precisely,*

$$\left\{\widehat{\boldsymbol{a}} = \left(\frac{a_1}{a_d}, \ldots, \frac{a_{d-1}}{a_d}\right)^t : \boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d\right\} = \{\boldsymbol{x} \in I^{d-1} : D(T)n(\boldsymbol{x})\Gamma \in \Lambda\}.$$

Proof. ('$\subseteq$') In view of Theorem 2.6, for every $\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d$, one has that

$$D(a_d)n(\widehat{\boldsymbol{a}})\Gamma \in K\Gamma.$$

It follows that $D(T)n(\widehat{\boldsymbol{a}})\Gamma \in \Lambda$.

('$\supseteq$') For every $\boldsymbol{x} \in I^{d-1}$ with $D(T)n(\boldsymbol{x})\Gamma \in \Lambda$, there exists $T' > 1$ such that

$$D(T/T')n(\boldsymbol{x})\Gamma \in K\Gamma/\Gamma.$$

For any lattice in $K\Gamma/\Gamma$, the last coordinates of its lattice points form the set $\mathbb{Z}$. This means that

$$\frac{T}{T'}(x_1, \ldots, x_{d-1}, 1)^t \in \widehat{\mathbb{Z}}^d.$$

Hence $\boldsymbol{x} = \widehat{\boldsymbol{a}}$ for some $\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d$. □

4.2 Transversal injectivitity radius of $\Lambda \subseteq \Omega$

To study effective equidistribution of the Farey sequence, we need to introduce the following.

Definition 4.2. Let $\pi : G \to \Omega$ be the natural projection given by $\pi(g) = g\Gamma$. For $g \in G$ and $x \in \Omega$, we set

$$|g|_\infty := \max\{|a_{ij}| : g = (a_{ij})\}, \quad |x|_\infty := \inf\{|g|_\infty : \pi(g) = x\}.$$

Let $\mathcal{C} \subseteq \Omega$ be a Borel subset; we define

$$|\mathcal{C}| := \max(1, \sup\{|x|_\infty : x \in \mathcal{C}\}).$$

Remark 4.3. It follows from the definition that for every $g \in G, A \in G_0, x \in \Omega$ and $\boldsymbol{b} \in \mathbb{R}^{d-1}$:

(i) $|gx|_\infty \ll |g|_\infty |x|_\infty$;
(ii) $|(A \ltimes \boldsymbol{b})\Gamma|_\infty \ll |A\Gamma_0|_\infty$; (as $(A \ltimes \boldsymbol{b})\Gamma = (A \ltimes \boldsymbol{b}')\Gamma$ for some $\|\boldsymbol{b}'\| \ll |A|_\infty$);
(iii) $|\mathcal{C}| < \infty$ for every relatively compact subset $\mathcal{C} \subset \Omega$.

Lemma 4.4. *Let $\tilde{\pi}$ be the map $\mathbb{R}^{d-1} \times F^+ \times K\Gamma/\Gamma \to \Omega$ defined by $\tilde{\pi}(\boldsymbol{x}, D(s), z) = n(\boldsymbol{x})D(s)z$, and $\mathcal{C}$ be a relatively compact subset of $K\Gamma/\Gamma$. Then the restriction of $\tilde{\pi}$ on $B(1/(4d|\mathcal{C}|)) \times F^+ \times \mathcal{C}$ is injective.*

Proof. Let $r = 1/(2d|\mathcal{C}|)$. It is enough to show that if $n(\boldsymbol{x})D(T)k_1\mathbb{Z}^d = k_2\mathbb{Z}^d$ for some $\boldsymbol{x} \in B(r)$, $T \geqslant 1$, $k_1\Gamma, k_2\Gamma \in \mathcal{C}$, then $\boldsymbol{x} = 0, T = 1$. To prove this, we choose k_1, k_2 so that $|k_1|_\infty, |k_2|_\infty < 2|\mathcal{C}|$, and let $k_1 = A \ltimes \boldsymbol{b}$. The last coordinates of the lattice points in $k_2\mathbb{Z}^d$ form the set $\mathbb{Z}$, so does the $\mathbb{Z}$-span of the entries in the last row of $n(\boldsymbol{x})D(T)k_1$, i.e.

$$T^{-1/(d-1)}(\mathbb{Z}\boldsymbol{x} \cdot \boldsymbol{a}_1 + \cdots + \mathbb{Z}\boldsymbol{x} \cdot \boldsymbol{a}_{d-1}) + \mathbb{Z}(T^{-1/(d-1)}\boldsymbol{x} \cdot \boldsymbol{b} + T) = \mathbb{Z},$$

where the $\boldsymbol{a}_i$ are the columns of A. By the choice of r we have $|\boldsymbol{x} \cdot \boldsymbol{a}_i| < 1$, so $\boldsymbol{x} = 0, T = 1$. □

https://doi.org/10.1112/S0010437X14007866 Published online by Cambridge University Press

4.3 The main equidistribution result

Let dk be the left Haar measure on K such that $dk = d\mu_0 d\boldsymbol{b}$, where $d\boldsymbol{b}$ is the Lebesgue measure on $\mathbb{R}^{d-1}$, and let $d\bar{k}$ be the induced probability measure on $K\Gamma/\Gamma$. According to Siegel's volume formula (cf. [Sie45, Mar10]) and [Kna02, Theorem 8.32], for any $f \in L^1(G)$

$$\int_{HFK} f(g)\, d\mu(g) = \frac{1}{\zeta(d)} \int_{H\times\mathbb{R}^+\times K} f(hD(s)k)\, d\nu(h) \frac{ds}{s^{d+1}}\, dk. \tag{4.1}$$

This naturally defines a Borel measure on Λ : $d\lambda = s^{-(d+1)}\, ds\, d\bar{k}$. We also consider for every smooth function ϕ on $I^{d-1} \times \Lambda$ the C^1-norm given by

$$\|\phi\|_{C^1} := \|\phi\|_{L^\infty} + \sum_{i=1}^{d-1} \left\| \frac{\partial}{\partial x_i} \phi \right\|_{L^\infty} + \sum_X \|\partial X(\phi)\|_{L^\infty}, \quad X \in \mathcal{X} \cap (\mathrm{Lie}\,(F) + \mathrm{Lie}\,(G_0) + \mathrm{Lie}\,(H^-)).$$

THEOREM 4.5. *There exists a constant $\alpha_3 > 0$ satisfying the following property. Let $\mathcal{C}$ be any relatively compact, open subset of $K\Gamma/\Gamma$, and $\mathcal{C}'$ be a compact subset of $\mathcal{C}$. For every $\varphi \in C^\infty(I^{d-1} \times \Lambda)$ with $\mathrm{supp}(\varphi) \subseteq I^{d-1} \times F^+\mathcal{C}'$, and $T > 1$ we have*

$$\left| \frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d} \varphi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma) - \frac{1}{\zeta(d)} \int_{I^{d-1}\times\Lambda} \varphi(\boldsymbol{x}, \lambda)\, d\boldsymbol{x}\, d\lambda \right| \ll |\mathcal{C}|^d \|\varphi\|_{C^1} T^{-\alpha_3}.$$

Proof. Step (i). Thicken an approximation of φ to a function $\psi \in C^\infty(I^{d-1} \times \Omega)$.

Let $0 < r < r_0 = 1/(4d|\mathcal{C}|)$. In the proof of this Theorem, we temporarily set $B_H(r) := \{n(\boldsymbol{x}) : \boldsymbol{x} \in B(r)\}$. We choose $\theta \in C^\infty(\mathbb{R}^{d-1})$ supported in $B(r)$ according to Lemma 3.7; and $\beta \in C^\infty(F)$ so that $0 \leqslant \beta \leqslant 1, \mathrm{supp}(\beta) \subseteq \{D(s) : s > e^{r/2}\}$, $\beta = 1$ on $\{D(s) : s \geqslant e^r\}$, and $\|\beta\|_{C^1} \ll r^{-1}$. We define a function ψ on the open submanifold $I^{d-1} \times B_H(r)F^+\mathcal{C}$ of $I^{d-1} \times \Omega$:

$$\psi(\boldsymbol{x}, n(\boldsymbol{y})D(s)z) = \beta(D(s))\theta(\boldsymbol{y})\varphi(\boldsymbol{x}, D(s)z) \quad \text{for all } \boldsymbol{x} \in I^{d-1}, \boldsymbol{y} \in B(r), s > 1, z \in \mathcal{C}.$$

The function ψ is well-defined by the injectivity result proved in Lemma 4.4 and the fact that $r < r_0$. By [Mar10, Lemma 2], $\{D(s) : s \geqslant s_0\}K\Gamma/\Gamma$ is a closed embedded submanifold of Ω for any $s_0 > 0$. It follows that $\overline{B_H(r)} \cdot \{D(s) : s \geqslant s_0\} \cdot \mathcal{C}'$ is a closed subset of Ω; hence the support of ψ in $I^{d-1} \times B_H(r)F^+\mathcal{C}$ is a closed subset of $I^{d-1} \times \Omega$. Therefore if we extend ψ to a function on $I^{d-1} \times \Omega$ by setting $\psi = 0$ outside the open subest $I^{d-1} \times B_H(r)F^+\mathcal{C} \subseteq I^{d-1} \times \Omega$, we get a smooth function which we, with abuse of notation, also denote by ψ. Moreover

$$\|\psi\|_{C^1} \ll (\|\beta\|_{L^\infty}\|\theta\|_{C^1} + \|\beta\|_{C^1}\|\theta\|_{L^\infty})\|\varphi\|_{C^1} \ll r^{-d}\|\varphi\|_{C^1}. \tag{4.2}$$

Notice that we have

$$\int_{I^{d-1}\times\Lambda} |\varphi(\boldsymbol{x}, \lambda) - \psi(\boldsymbol{x}, \lambda)|\, d\boldsymbol{x}\, d\lambda \ll r\|\varphi\|_{L^\infty}. \tag{4.3}$$

If $T > e^r a_d$, then

$$\varphi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma) = \psi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma).$$

On the other hand,

$$\#\{\boldsymbol{a} \in T\mathcal{D}_0 : T < e^r a_d\} \ll rT^d.$$

It follows that

$$\sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d} |\varphi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma) - \psi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma)| \ll T^d r\|\varphi\|_{L^\infty}. \tag{4.4}$$

Step (ii). Compare the average of ψ over the Farey sequences and horospheres.

Let $T > 1, \boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d$, and set $r' = r/T^{d/(d-1)}$ where $r < r_0$ as before. Let

$$\mathcal{E}_r = \{x \in I^{d-1} : \mathrm{dist}(x, \partial I^{d-1}) > r\}, \quad \mathcal{M}_{T,r} = \{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d : \widehat{\boldsymbol{a}} \in \mathcal{E}_{r'}, D(T)n(\widehat{\boldsymbol{a}})\Gamma \in F^+\mathcal{C}\}.$$

By our construction $\psi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma) \neq 0$ only if $D(T)n(\widehat{\boldsymbol{a}})\Gamma \in F^+\mathcal{C}$; hence

$$\sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d, \widehat{\boldsymbol{a}} \in \mathcal{E}_{r'}} \psi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma) = \sum_{\boldsymbol{a} \in \mathcal{M}_{T,r}} \int_{B(r)} \theta(\boldsymbol{y})\psi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma)\, d\boldsymbol{y}.$$

Let us consider the subset $\mathcal{Z}_{T,r} := \bigcup_{\boldsymbol{a} \in \mathcal{M}_{T,r}} \{\widehat{\boldsymbol{a}} + \boldsymbol{y} : \boldsymbol{y} \in B(r')\}$ of I^{d-1}. Our injectivity assumption implies that the union in $\mathcal{Z}_{T,r}$ is disjoint. Hence we have that

$$\begin{aligned}
&\left| \frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d, \widehat{\boldsymbol{a}} \in \mathcal{E}_{r'}} \psi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma) - \int_{\mathcal{Z}_{T,r}} \psi(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma)\, d\boldsymbol{x} \right| \\
&\quad = \left| \sum_{\boldsymbol{a} \in \mathcal{M}_{T,r}} \int_{B(r')} \psi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}} + \boldsymbol{y})\Gamma)\, d\boldsymbol{y} - \sum_{\boldsymbol{a} \in \mathcal{M}_{T,r}} \int_{B(r')} \psi(\widehat{\boldsymbol{a}} + \boldsymbol{y}, D(T)n(\widehat{\boldsymbol{a}} + \boldsymbol{y})\Gamma)\, d\boldsymbol{y} \right| \\
&\quad \ll T^d \|\psi\|_{\mathrm{Lip}} \int_{B(r')} \|\boldsymbol{y}\|\, d\boldsymbol{y} \ll \|\varphi\|_{C^1} T^{-d/(d-1)}.
\end{aligned} \tag{4.5}$$

On the other hand, we have that $\{\boldsymbol{x} \in \mathcal{E}_{2r'} : \psi(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma) \neq 0\} \subseteq \mathcal{Z}_{T,r}$. To see this, notice that for any such $\boldsymbol{x}$, we have $n(\boldsymbol{x}_1)D(T)n(\boldsymbol{x})\Gamma \in F^+\mathcal{C}$ for some $\boldsymbol{x}_1 \in B(r)$, because our function ψ is supported in $I^{d-1} \times B_H(r)F^+\mathcal{C}$. By Lemma 4.1, $n(\boldsymbol{x}_1)D(T)n(\boldsymbol{x})\Gamma = D(T)n(\widehat{\boldsymbol{a}})\Gamma$ for some $\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d$. As $\boldsymbol{x} \in \mathcal{E}_{2r'}$, we have that $\widehat{\boldsymbol{a}} \in \mathcal{E}_{r'}$. The above discussion implies that

$$\left| \int_{I^{d-1} - \mathcal{Z}_{T,r}} \psi(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma)\, d\boldsymbol{x} \right| \ll r' \|\psi\|_{L^\infty}. \tag{4.6}$$

Because $\#\{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d : \widehat{\boldsymbol{a}} \notin \mathcal{E}_{r'}\} \ll T^d r'$, and $r < r_0 < 1$ (as $|\mathcal{C}| \geqslant 1$) we get that

$$\begin{aligned}
&\left| \int_{I^{d-1}} \psi(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma)\, d\boldsymbol{x} - \frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d} \psi(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma) \right| \\
&\quad \ll r' \|\psi\|_{L^\infty} + (4.5) + (4.6) \ll r^{(1-d)} \|\varphi\|_{C^1} T^{-d/(d-1)}.
\end{aligned} \tag{4.7}$$

Step (iii). Apply the equidistribution result of expanding horospheres.

By (4.2) and Theorem 3.11

$$\left| \int_{I^{d-1}} \psi(\boldsymbol{x}, D(T)n(\boldsymbol{x})\Gamma)\, d\boldsymbol{x} - \int_{I^{d-1} \times \Omega} \psi(\boldsymbol{x}, z)\, d\boldsymbol{x}\, d\bar{\mu}(z) \right| \ll r^{-d} \|\varphi\|_{C^1} T^{-\alpha_2}. \tag{4.8}$$

Let $\mathcal{K}$ be a Borel subset of K which is mapped bijectively onto $\mathcal{C}$ by π. By equation (4.1)

$$\begin{aligned}
\int_{I^{d-1} \times \Omega} \psi(\boldsymbol{x}, z)\, d\boldsymbol{x}\, d\bar{\mu}(z) &= \int_{I^{d-1} \times B_H(r) \cdot F^+ \cdot \mathcal{K}} \psi(\boldsymbol{x}, g\Gamma)\, d\boldsymbol{x}\, d\mu(g) \\
&= \frac{1}{\zeta(d)} \int_{I^{d-1} \times \mathbb{R}^{d-1} \times \mathbb{R}_{>1} \times \mathcal{K}} \theta(\boldsymbol{y})\psi(\boldsymbol{x}, D(s)k\Gamma)\, d\boldsymbol{x}\, d\boldsymbol{y} \frac{ds}{s^{d+1}}\, dk \\
&= \frac{1}{\zeta(d)} \int_{I^{d-1} \times \Lambda} \psi(\boldsymbol{x}, \lambda)\, d\boldsymbol{x}\, d\lambda.
\end{aligned} \tag{4.9}$$

https://doi.org/10.1112/S0010437X14007866 Published online by Cambridge University Press

Setting $r = \frac{1}{2} r_0 T^{-\alpha_3}$ for some suitable constant $\alpha_3 > 0$ and combining (4.3)–(4.9), we conclude that for every $T > 1$

$$\left| \frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d} \varphi(\widehat{\boldsymbol{a}}, D(T) n(\widehat{\boldsymbol{a}})\Gamma) - \frac{1}{\zeta(d)} \int_{I^{d-1} \times \Lambda} \varphi(\boldsymbol{x}, \lambda)\, d\boldsymbol{x}\, d\lambda \right| \ll |\mathcal{C}|^d \|\varphi\|_{C^1} T^{-\alpha_3}. \qquad \square$$

4.4 Consequences of Theorem 4.5

Recall from Theorem 2.6 that for any primitive lattice point $\boldsymbol{a} \in \widehat{\mathbb{Z}}^d_{\geqslant 2}$, the lattice $L_{\boldsymbol{a}}$ appearing in (1.5) which produces the Frobenius number $F(\boldsymbol{a})$ is given by $L_{\boldsymbol{a}} = m(\widehat{\boldsymbol{a}}) D(a_d) n(\widehat{\boldsymbol{a}}) \mathbb{Z}^d \cap e_d^{\perp}$. Let $L'_{\boldsymbol{a}} = m(\widehat{\boldsymbol{a}}) D(a_d) n(\widehat{\boldsymbol{a}}) \Gamma \in K\Gamma/\Gamma$. Translating $L'_{\boldsymbol{a}}$ by $D(T/a_d)$, we get

$$D(T/a_d) L'_{\boldsymbol{a}} = m(\widehat{\boldsymbol{a}}) D(T) n(\widehat{\boldsymbol{a}}) \Gamma.$$

Moreover, we have $D(T/a_d) L'_{\boldsymbol{a}} \in \Lambda$ if and only if $\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d$. The following theorem shows that under this translation, the lattices $\{L'_{\boldsymbol{a}} : \boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d\}$ become equidistributed in Λ.

Proposition 4.6. *There exists a constant $\alpha_4 > 0$ with the following property. With $\mathcal{C}$, $\mathcal{C}'$ as in Theorem 4.5, for any $\varphi \in C^\infty(I^{d-1} \times \Lambda)$ with $\mathrm{supp}(\varphi) \subseteq I^{d-1} \times F^+ \mathcal{C}'$, and any $T > 1$, we have*

$$\left| \frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d} \varphi(\widehat{\boldsymbol{a}}, m(\widehat{\boldsymbol{a}}) D(T) n(\widehat{\boldsymbol{a}})\Gamma) - \frac{1}{\zeta(d)} \int_{I^{d-1} \times \Lambda} \varphi(\boldsymbol{x}, \lambda)\, d\boldsymbol{x}\, d\lambda \right| \ll |\mathcal{C}|^d \|\varphi\|_{C^1} T^{-\alpha_4}.$$

Remark 4.7. The non-effective result can be derived from Theorem 4.5 via the following simple fact. Let μ_n, μ be Borel measures on $I^{d-1} \times \Lambda$ so that μ_n converges to μ in the weak* topology. Let $\mathcal{T} : I^{d-1} \times \Lambda \to I^{d-1} \times \Lambda$ be a smooth map given by $\mathcal{T}(\boldsymbol{x}, \lambda) = (\boldsymbol{x}, m(\boldsymbol{x})\lambda)$. Then the push-forward Borel measures $\mathcal{T}^*(\mu_n)$ also weak* converge to $\mathcal{T}^*(\mu)$.

Proof. Let $\mathcal{T}$ be as in Remark 4.7. Since $\|\varphi \circ \mathcal{T}\|_{C^1}$ is not necessarily finite, we need to approximate $\varphi \circ \mathcal{T}$ by compactly supported functions to get an error estimate. Recall that $\mathcal{E}_r = \{x \in I^{d-1} : \mathrm{dist}(x, \partial I^{d-1}) > r\}$. Let $\theta \in C^1(I^{d-1})$ be a function so that $\chi_{\mathcal{E}_r} \leqslant \theta \leqslant \chi_{\mathcal{E}_{r/2}}$ and $\|\theta\|_{C^1} \ll r^{-1}$. The function $\tilde{\varphi}(\boldsymbol{x}, \lambda) = \theta(\boldsymbol{x}) \varphi(\boldsymbol{x}, m(\boldsymbol{x})\lambda)$ satisfies $\mathrm{supp}(\tilde{\varphi}) \subseteq I^{d-1} \times F^+ \mathcal{C}'_1$ where $\mathcal{C}'_1$ is a compact subset of $\mathcal{C}_1 = \bigcup_{\boldsymbol{x} \in \mathcal{E}_{r/2}} m(\boldsymbol{x})^{-1} \mathcal{C}$. Since the entries of each $m(\boldsymbol{x})^{-1}$ ($\boldsymbol{x} \in \mathcal{E}_{r/2}$) are bounded by $2/r$ in absolute value, it follows from Remark 4.3 that $|\mathcal{C}_1| \ll |\mathcal{C}| r^{-1}$.

Claim. *There exists $n = n(d) > 0$, such that $\|\tilde{\varphi}\|_{C^1} \ll r^{-n} \|\theta\|_{C^1} \|\varphi\|_{C^1}$.*

Proof of the claim. We take $(\boldsymbol{x}, \lambda) \in \mathcal{E}_{r/2} \times \Lambda$ and consider the differential $d\mathcal{T}$ of $\mathcal{T}$ at this point. We use $\partial/\partial x_i$ as the usual Euclidean tangent vector at $\boldsymbol{x} \in I^{d-1}$, and let $\partial_\lambda X$ denote a tangent vector at $\lambda \in \Lambda$ (see Definition 3.6 and Theorem 4.5). It is easy to check that

$$\begin{aligned} d\mathcal{T}\left(\frac{\partial}{\partial x_i}\right) &= \frac{\partial}{\partial x_i} + \frac{1}{(d-1)x_i} \partial_{m(\boldsymbol{x})\lambda} E_i, \quad 1 \leqslant i \leqslant d-1 \\ d\mathcal{T}(\partial_\lambda X) &= \partial_{m(\boldsymbol{x})\lambda} Ad(m(\boldsymbol{x}))(X), \end{aligned}$$

where $E_i = \mathrm{diag}(-1, \ldots, d-2, \ldots, -1, 0)$ with $d-2$ in the (i,i)th entry. Hence the norm of $d\mathcal{T}$ satisfies $\|d\mathcal{T}\| \ll r^{-n}$ for some $n = n(d)$ at every $(x, \lambda) \in \mathcal{E}_{r/2} \times \Lambda$. $\square$

https://doi.org/10.1112/S0010437X14007866 Published online by Cambridge University Press

Note that $\varphi(\widehat{\boldsymbol{a}}, m(\widehat{\boldsymbol{a}})D(T)n(\widehat{\boldsymbol{a}})\Gamma) = \tilde{\varphi}(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma)$ if $\widehat{\boldsymbol{a}} \in \mathcal{E}_r$. We thus have

$$\begin{aligned}
&\left|\frac{1}{T^d}\sum_{\boldsymbol{a}\in T\mathcal{D}_0\cap\widehat{\mathbb{Z}}^d}\varphi(\widehat{\boldsymbol{a}}, m(\widehat{\boldsymbol{a}})D(T)n(\widehat{\boldsymbol{a}})\Gamma) - \frac{1}{\zeta(d)}\int_{I^{d-1}\times\Lambda}\varphi(\boldsymbol{x},\lambda)\,d\boldsymbol{x}\,d\lambda\right| \\
&\quad\leqslant \left|\frac{1}{T^d}\sum_{\boldsymbol{a}\in T\mathcal{D}_0\cap\widehat{\mathbb{Z}}^d}\tilde{\varphi}(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma) - \frac{1}{\zeta(d)}\int_{I^{d-1}\times\Lambda}\tilde{\varphi}(\boldsymbol{x},\lambda)\,d\boldsymbol{x}\,d\lambda\right| \\
&\qquad+\left|\frac{1}{T^d}\sum_{\boldsymbol{a}\in T\mathcal{D}_0\cap\widehat{\mathbb{Z}}^d,\widehat{\boldsymbol{a}}\in I^{d-1}\setminus\mathcal{E}_r}(\varphi(\widehat{\boldsymbol{a}}, m(\widehat{\boldsymbol{a}})D(T)n(\widehat{\boldsymbol{a}})\Gamma) - \tilde{\varphi}(\widehat{\boldsymbol{a}}, D(T)n(\widehat{\boldsymbol{a}})\Gamma))\right| \\
&\qquad+\left|\frac{1}{\zeta(d)}\int_{I^{d-1}\times\Lambda}\varphi(\boldsymbol{x},\lambda)\,d\boldsymbol{x}\,d\lambda - \frac{1}{\zeta(d)}\int_{I^{d-1}\times\Lambda}\tilde{\varphi}(\boldsymbol{x},\lambda)\,d\boldsymbol{x}\,d\lambda\right|. \qquad (4.10)
\end{aligned}$$

Notice that the Haar measure $d\bar{k}$ on $K\Gamma/\Gamma$ is left-invariant, so

$$\int_{I^{d-1}\times\Lambda}\tilde{\varphi}(\boldsymbol{x},\lambda)\,d\boldsymbol{x}\,d\lambda = \int_{I^{d-1}\times\Lambda}\theta(\boldsymbol{x})\varphi(\boldsymbol{x},\lambda)\,d\boldsymbol{x}\,d\lambda.$$

Applying Theorem 4.5 to the function $\tilde{\varphi}$ and using the claim above, we conclude that

$$(4.10) \ll |\mathcal{C}|^d r^{-d}\|\tilde{\varphi}\|_{C^1}T^{-\alpha_3} + r\|\varphi\|_{L^\infty} + r\|\varphi\|_{L^\infty} \ll |\mathcal{C}|^d r^{-n-d-1}\|\varphi\|_{C^1}T^{-\alpha_3} + r\|\varphi\|_{L^\infty}.$$

We complete the proof by setting $r = T^{-\alpha_4}$ for suitable $\alpha_4 > 0$. □

Recall that the space Ω_0 is naturally embedded into $K\Gamma/\Gamma$. The map from $K\Gamma/\Gamma$ to Ω_0 given by $\iota : k\mathbb{Z}^d \mapsto k\mathbb{Z}^d \cap \mathbf{e}_d^\perp$ is smooth. (As before we identify $\mathbb{R}^d \cap \mathbf{e}_d^\perp$ with $\mathbb{R}^{d-1}$ in the obvious way.) We will use for every $z \in K\Gamma/\Gamma$ the notations $\iota(z)$ and $z \cap \mathbf{e}_d^\perp$ interchangeably. Consider the product Riemannian manifold of $\mathcal{D}_0$ (Euclidean metric) and Ω_0

$$\mathcal{M}_{\mathcal{D}_0} = \{(\boldsymbol{x}, y, z) : (\boldsymbol{x}, y) \in \mathcal{D}_0, z \in \Omega_0\}. \qquad (4.11)$$

The product Borel measure on $\mathcal{M}_{\mathcal{D}_0}$ is written as $d_{\mathcal{D}_0}(\boldsymbol{x}, y, z) = d\boldsymbol{x}\,dy\,d\bar{\mu}_0(z)$.

Notational Convention 4.8. To simplify the notation, whenever $d_{\mathcal{D}_0}$ is used to abbreviate $d_{\mathcal{D}_0}(\boldsymbol{x}, y, z)$ we always assume implicitly that $\mathcal{M}_{\mathcal{D}_0}$ is parametrized as in (4.11).

We set for every smooth function f on $\mathcal{M}_{\mathcal{D}_0}$

$$\|f\|_{C^1} := \|f\|_{L^\infty} + \sum_{i=1}^{d-1}\left\|\frac{\partial}{\partial x_i}f\right\|_{L^\infty} + \left\|\frac{\partial}{\partial y}f\right\|_{L^\infty} + \sum_X \|\partial X(f)\|_{L^\infty}, \quad X \in \mathcal{X} \cap \mathrm{Lie}(G_0).$$

Corollary 4.9. *Let $\mathcal{C}$ be a relatively compact, open subset of Ω_0, and $\mathcal{C}'$ be a compact subset of $\mathcal{C}$. Then for any $\psi \in C^\infty(\mathcal{M}_{\mathcal{D}_0})$ with $\mathrm{supp}(\psi) \subseteq \{(\boldsymbol{x}, y, z) \in \mathcal{M}_{\mathcal{D}_0} : z \in \mathcal{C}'\}$, and for every $T > 1$ we have (the α_4 below is as in Proposition 4.6)*

$$\left|\frac{1}{T^d}\sum_{\boldsymbol{a}\in T\mathcal{D}_0\cap\widehat{\mathbb{Z}}^d}\psi\left(\frac{\boldsymbol{a}}{T}, m(\widehat{\boldsymbol{a}})D(a_d)n(\widehat{\boldsymbol{a}})\Gamma \cap \mathbf{e}_d^\perp\right) - \frac{1}{\zeta(d)}\int_{\mathcal{M}_{\mathcal{D}_0}}\psi(\boldsymbol{x}, y, z)d_{\mathcal{D}_0}\right| \ll |\mathcal{C}|^d\|\psi\|_{C^1}T^{-\alpha_4}.$$

Proof. Let $\mathcal{Q}$ be the smooth map from $I^{d-1} \times \Lambda$ to $\mathcal{M}_{\mathcal{D}_0}$ defined by

$$\mathcal{Q}(\boldsymbol{x}, D(y)^{-1}z') = (y\boldsymbol{x}, y, z' \cap \mathbf{e}_d^\perp), \quad (\boldsymbol{x}, y) \in I^d, z' \in K\Gamma/\Gamma.$$

Let $\tilde{\psi} = \psi \circ \mathcal{Q}$ be a smooth function on $I^{d-1} \times \Lambda$. We are going to show that $\|\tilde{\psi}\|_{C^1} < \infty$. Note that for every $A \in G_0, \boldsymbol{b} \in \mathbb{R}^{d-1}$

$$(A \ltimes \boldsymbol{b})D(s) = D(s)(A \ltimes (s^{d/(d-1)}\boldsymbol{b})).$$

Thus at every $w = (\boldsymbol{x}, D(y)^{-1}z') \in I^{d-1} \times \Lambda$, the directional derivatives (see Definition 3.6) satisfy that $\partial_w Z(\tilde{\psi}) = \partial_{\mathcal{Q}(w)} Z(\psi)$ for every $Z \in \mathrm{Lie}(G_0)$, and $\partial_w Y(\tilde{\psi}) = 0$ for every $Y \in \mathrm{Lie}(H^-)$. Let $X = \mathrm{diag}(1/(d-1), \ldots, 1/(d-1), -1) \in \mathrm{Lie}(F)$. We have

$$d\mathcal{Q}(\partial_w X) = \sum_{i=1}^{d-1} x_i y \frac{\partial}{\partial x_i} + y \frac{\partial}{\partial y}.$$

It follows easily that $\|\tilde{\psi}\|_{C^1} \ll \|\psi\|_{C^1}$. Moreover,

$$\begin{aligned}\int_{I^{d-1}\times\Lambda} \tilde{\psi}(\boldsymbol{x}, \lambda)\, d\boldsymbol{x}\, d\lambda &= \int_{I^{d-1}\times I \times K\Gamma/\Gamma} \psi(y\boldsymbol{x}, y, z' \cap \mathbf{e}_d^{\perp}) y^{d-1}\, d\boldsymbol{x}\, dy\, d\bar{k}(z') \\ &= \int_{\mathcal{M}_{\mathcal{D}_0}} \psi(\boldsymbol{x}, y, z)\, d_{\mathcal{D}_0}.\end{aligned}$$

On the other hand, Mahler's criterion implies that the set $\mathcal{C}_1 = \{z \in K\Gamma/\Gamma : z \cap \mathbf{e}_d^{\perp} \in \mathcal{C}\}$ is a relatively compact, open subset of $K\Gamma/\Gamma$, and $\mathcal{C}_1' = \{z \in K\Gamma/\Gamma : z \cap \mathbf{e}_d^{\perp} \in \mathcal{C}'\}$ is compact. Moreover $|\mathcal{C}_1| \ll |\mathcal{C}|$ (Remark 4.3). Since $\mathrm{supp}(\tilde{\psi}) \subseteq I^{d-1} \times F^+\mathcal{C}_1'$, by Proposition 4.6

$$\begin{aligned}&\left|\frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d} \psi\left(\frac{\boldsymbol{a}}{T}, m(\widehat{\boldsymbol{a}})D(a_d)n(\widehat{\boldsymbol{a}})\Gamma \cap \mathbf{e}_d^{\perp}\right) - \frac{1}{\zeta(d)} \int_{\mathcal{M}_{\mathcal{D}_0}} \psi(\boldsymbol{x}, y, z)\, d_{\mathcal{D}_0}\right| \\ &= \left|\frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d} \tilde{\psi}(\widehat{\boldsymbol{a}}, m(\widehat{\boldsymbol{a}})D(T)n(\widehat{\boldsymbol{a}})\Gamma) - \frac{1}{\zeta(d)} \int_{I^{d-1}\times\Lambda} \tilde{\psi}(\boldsymbol{x}, \lambda)\, d\boldsymbol{x}\, d\lambda\right| \\ &\ll |\mathcal{C}|^d \|\psi\|_{C^1} T^{-\alpha_4}.\end{aligned}$$

□

Remark 4.10. The equidistribution result in Corollary 4.9 enables us to derive (1.7). Indeed, for any $\phi \in C(\Omega_0)$ let ϕ_0 be the function on $\mathcal{M}_{\mathcal{D}_0}$ defined by $\phi_0(\boldsymbol{x}, y, z) := \chi_{\mathcal{D}}(\boldsymbol{x}, y)\phi(z)$. (Recall that $L_{\boldsymbol{a}} = m(\widehat{\boldsymbol{a}})D(a_d)n(\widehat{\boldsymbol{a}})\Gamma \cap \mathbf{e}_d^{\perp}$.) Then

$$\frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D} \cap \widehat{\mathbb{Z}}^d} \phi(L_{\boldsymbol{a}}) - \frac{\mathrm{vol}(\mathcal{D})}{\zeta(d)} \int_{\Omega_0} \phi\, d\bar{\mu}_0 = \frac{1}{T^d} \sum_{\boldsymbol{a} \in T\mathcal{D}_0 \cap \widehat{\mathbb{Z}}^d} \phi_0\left(\frac{\boldsymbol{a}}{T}, L_{\boldsymbol{a}}\right) - \frac{1}{\zeta(d)} \int_{\mathcal{M}_{\mathcal{D}_0}} \phi_0\, d_{\mathcal{D}_0}. \quad (4.12)$$

Suppose $\mathcal{D} \subseteq \mathcal{D}_0$ has boundary of Lebesgue measure zero. We can apply Corollary 4.9 and a weak* convergence argument to show that the above expression tends to zero as $T \to \infty$. This completes the proof of (1.7).

4.5 Approximation of indicator functions by smooth functions

Our discussion suggests that to study the error of (4.12) we have to deal with the error term in the equidistribution result of Corollary 4.9 when indicator functions are involved. Technically we consider sets with thin boundary. Let us recall a well-known result.

Lemma 4.11. *Let $\mathcal{D}$ be a bounded open subset of $\mathbb{R}^d$ with thin boundary, and m be the Lebesgue measure. Then for every $0 < r < 1$ there exist smooth functions f_1, f_2 so that $0 \leqslant f_1 \leqslant \chi_{\mathcal{D}} \leqslant f_2$, $\|f_i - \chi_{\mathcal{D}}\|_{L^1(m)} \ll_{\mathcal{D}} r$, and $\|f_i\|_{C^1} \ll r^{-1}$ $(i = 1, 2)$.*

Remark 4.12. The key fact which guarantees the approximation in Lemma 4.11 is that, for every $0 < r < 1$, we have $m(\{x \in \mathbb{R}^d : d(x, \partial\mathcal{D}) < r\}) \ll_{\mathcal{D}} r$, where d is the Euclidean distance. In view of [SV05, Lemma 1], the statement of Lemma 4.11 remains valid when $(\mathbb{R}^d, m)$ is replaced by $(\Omega, \bar{\mu})$.

Proposition 4.13. *There exists a constant $\alpha_5 > 0$ with the following property. With $\mathcal{C}, \mathcal{C}'$ as in Corollary 4.9, for any non-empty open subset $\mathcal{D} \subseteq \mathcal{D}_0$ which has thin boundary as a subset of $\mathbb{R}^d$, any non-negative function $\phi \in C^\infty(\Omega_0)$ with $\mathrm{supp}(\phi) \subseteq \mathcal{C}'$ and any $T > 1$, we have that*

$$\left|\frac{1}{T^d}\sum_{\boldsymbol{a}\in T\mathcal{D}\cap\widehat{\mathbb{Z}}^d}\phi(L_{\boldsymbol{a}}) - \frac{\mathrm{vol}(\mathcal{D})}{\zeta(d)}\int_{\Omega_0}\phi\,d\bar{\mu}_0\right| \ll_{\mathcal{D}} |\mathcal{C}|^d\|\phi\|_{C^1}T^{-\alpha_5}. \tag{4.13}$$

Here $\|\phi\|_{C^1} := \|\phi\|_{L^\infty} + \sum_{X\in(\mathcal{X}\cap\mathrm{Lie}(G_0))}\|\partial X(\phi)\|_{L^\infty}$.

Proof. For every function f on $\mathcal{D}_0$ and function ψ on Ω_0, we denote by $f \otimes \psi$ the function on $\mathcal{M}_{\mathcal{D}_0}$ defined by $(f\otimes\psi)(\boldsymbol{x}, y, z) = f(\boldsymbol{x}, y)\psi(z)$. Since $\mathcal{D}$ has thin boundary in $\mathbb{R}^d$, for every $0 < r < 1$ we take f_1, f_2 as in Lemma 4.11 and consider their restrictions to $\mathcal{D}_0$. By Corollary 4.9 and Lemma 4.11, for $i = 1, 2$

$$\left|\frac{1}{T^d}\sum_{\boldsymbol{a}\in T\mathcal{D}_0\cap\widehat{\mathbb{Z}}^d}(f_i\otimes\phi)\left(\frac{\boldsymbol{a}}{T}, L_{\boldsymbol{a}}\right) - \frac{1}{\zeta(d)}\int_{\mathcal{M}_{\mathcal{D}_0}}(f_i\otimes\phi)\,d_{\mathcal{D}_0}\right| \ll |\mathcal{C}|^d r^{-1}\|\phi\|_{C^1}T^{-\alpha_4}. \tag{4.14}$$

Again by Lemma 4.11, we have that $\|(f_1\otimes\phi) - (f_2\otimes\phi)\|_{L^1(\mathcal{M}_{\mathcal{D}_0})} \ll_{\mathcal{D}} r\|\phi\|_{L^\infty}$. Notice that $f_1\otimes\phi \leqslant \chi_{\mathcal{D}}\otimes\phi \leqslant f_2\otimes\phi$ as ϕ is non-negative. It follows easily that estimate (4.13) holds. □

5. The Proof of Theorem 1.4

Proof. Let Q_0 be the covering radius function as before. For any fixed $R > 0$ we show that $\{L \in \Omega_0 : Q_0(L) < R\}$ has thin boundary as a subset of Ω_0. That is, the set $E_R = \{L \in \Omega_0 : Q_0(L) = R\}$ is contained in a union of finitely many bounded connected submanifolds of Ω_0 of dimensions $< \dim(\Omega_0)$. In fact the statement is 'almost' established in [Mar10, Lemma 7], and we only need to provide some further explanation. In view of Remark 4.12, Theorem 1.4 can be deduced with essentially the same argument as in Proposition 4.13.

Let $\Sigma_1, \ldots, \Sigma_d$ be the faces of the standard simplex Δ_{d-1}. We fix a Borel Γ_0-fundamental domain $\mathcal{F}_0$ in G_0 so that every compact subset in Ω_0 corresponds to a relatively compact subset in $\mathcal{F}_0$, and set $\mathcal{L}_R = \{A \in \mathcal{F}_0 : Q_0(A\mathbb{Z}^{d-1}) = R\}$. By [Mar10, Lemma 7],

$$\mathcal{L}_R \subseteq \bigcup_{\mathbf{n}_1,\ldots,\mathbf{n}_d\in\mathbb{Z}^{d-1}}\{A\in\mathcal{F}_0 : \text{there exists } \zeta\in\mathbb{R}^{d-1} \text{ so that } A\mathbf{n}_i\cap(R\Sigma_i^\circ+\zeta)\neq\emptyset\ (i=1,\ldots,d)\}.$$

As $\mathcal{L}_R$ is a relatively compact subset in G_0 (Lemma 2.3), there exists $c > 0$ so that $B_{\mathbb{R}^{d-1}}(c)$ contains a fundamental domain of $A\mathbb{Z}^{d-1}$ for every $A \in \mathcal{L}_R$. Hence $\mathcal{L}_R$ is a subset of

$$\bigcup_{\mathbf{n}_1,\ldots,\mathbf{n}_d\in\mathbb{Z}^{d-1}}\{A\in\mathcal{F}_0 : \text{there exists } \|\zeta\|<c \text{ so that } A\mathbf{n}_i\cap(R\Sigma_i^\circ+\zeta)\neq\emptyset\ (i=1,\ldots,d)\}. \tag{5.1}$$

Since $\mathcal{L}_R$ is relatively compact, we have that $\|A\mathbf{n}_i\| \gg_R \|\mathbf{n}_i\|$ whenever $A \in \mathcal{L}_R$. As the set $R\Sigma_i^\circ + B_{\mathbb{R}^{d-1}}(c)$ is bounded, it follows that in (5.1) $\mathcal{L}_R$ is contained in a finite union. To complete the proof of the theorem, it suffices to show that for every fixed d integral vectors $\mathbf{n}_1, \ldots, \mathbf{n}_d \in \mathbb{Z}^{d-1}$, the set

$$\{A \in \mathcal{F}_0 : \text{there exists } \|\zeta\| < c \text{ so that } A\mathbf{n}_i \cap (R\Sigma_i^\circ + \zeta) \neq \emptyset \ (i = 1, \ldots, d)\} \tag{5.2}$$

is contained in a union of finitely many bounded connected submanifolds of G_0, and that each of them has dimension $< \dim G_0$. Indeed, as the map $\pi_0 : G_0 \to \Omega_0$ given by $\pi_0(g) = g\Gamma_0$ ($g \in G_0$) is a local diffeomorphism, we can further conclude that $\{L \in \Omega_0 : Q_0(L) < R\}$ has thin boundary as a subset of Ω_0.

It was shown in the proof of [Mar10, Lemma 7] that

$$(5.2) \subseteq \{A = (a_{ij}) \in G_0 : \operatorname{tr}(LA) = R\}, \tag{5.3}$$

where L is the $(n-1) \times (n-1)$ matrix whose ith column is $\mathbf{n}_i - \mathbf{n}_d$. Because $\mathcal{L}_R$ is relatively compact, there is a constant $C_R > 0$, so that (5.3) can be refined as

$$(5.2) \cap \mathcal{L}_R \subseteq \{A = (a_{ij}) \in G_0 : |a_{ij}| < C_R, \operatorname{tr}(LA) = R\}. \tag{5.4}$$

The set $\{A = (a_{ij}) \in G_0 : |a_{ij}| < C_R, \operatorname{tr}(LA) = R\}$ is a semi-algebraic set, and standard results in real algebraic geometry (see for example [BCR98, (2.9)]) imply that it can be written as a union of finitely many bounded connected submanifold of G_0 of dimensions $< \dim G_0$. □

Acknowledgements

I would like to thank J. Marklof and J. Athreya for their discussions on the concepts and ideas related to this project, and for their comments on the early draft of this paper. I am grateful to Marklof for pointing out a problem in the original formulation of Theorem 1.4. I am deeply indebted to my advisor Professor G. A. Margulis for suggesting this project as part of my PhD thesis, and for his invaluable discussions and constant encouragement. I thank H. Oh and A. Strömbergsson for their insightful discussions. I also want to thank the anonymous referee for many helpful suggestions and corrections.

References

AG07 I. M. Aliev and P. M. Gruber, *An optimal lower bound for the Frobenius problem*, J. Number Theory **123** (2007), 71–79.

AHH11 I. M. Aliev, M. Henk and A. Hinrichs, *Expected Frobenius numbers*, J. Combin. Theory Ser. A **118** (2011), 525–531.

AM09 J. S. Athreya and G. A. Margulis, *Logarithm laws for unipotent flows, I*, J. Mod. Dyn. **3** (2009), 359–378.

BCR98 J. Bochnak, M. Coste and M.-F. Roy, *Real algebraic geometry* (Springer, 1998).

BS07 J. Bourgain and Ya. G. Sinai, *Limit behavior of large Frobenius numbers*, Russian Math. Surveys **62** (2007), 713–725.

Cow79 M. Cowling, *Sur le coefficients des représentations unitaries des groupes de Lie simple*, Lecture Notes in Mathematics, vol. 739 (Springer, New York, 1979), 132–178.

EG72 P. Erdős and R. Graham, *On a linear Diophantine problem of Frobenius*, Acta Arith. **21** (1972), 399–408.

How82 R. Howe, *On a notion of rank for unitary representations of the classical groups*, in *Harmonic analysis and group representations* (Liguori, Naples, 1982), 223–331.

Kan92 R. Kannan, *Lattice translates of a polytope and the Frobenius problem*, Combinatorica **12** (1992), 161–177.

KL88 R. Kannan and L. Lovász, *Covering minima and lattice-point-free convex bodies*, Ann. of Math. (2) **128** (1988), 577–602.

KS94 A. Katok and R. Spatzier, *First cohomology of Anosov actions of higher rank Abelian groups and applications to rigidity*, Publ. Math. Inst. Hautes Études Sci. **79** (1994), 131–156.

KM96 D. Kleinbock and G. A. Margulis, *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, Amer. Math. Soc. Transl. **171** (1996), 141–172.

KM12 D. Kleinbock and G. A. Margulis, *On effective equidistribution of expanding translates of certain orbits in the space of lattices*, in *Number theory, analysis and geometry* (Springer, New York, 2012), 385–396.

Kna02 A. Knapp, *Lie groups beyond an introduction* (Birkhäuser, Boston, 2002).

Mar10 J. Marklof, *The asymptotic distribution of Frobenius numbers*, Invent. Math. **181** (2010), 179–207.

Moo87 C. C. Moore, *Exponential decay of correlation coefficients for geodesic flows*, in *Group representations, ergodic theory, operator algebras, and mathematical physics (Berkeley, California, 1984)*, Mathematical Sciences Research Institute Publications (Springer, New York, 1987), 163–181.

Oh02 H. Oh, *Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants*, Duke Math. J. **113** (2002), 133–192.

Ram05 J. Ramirez Alfonsin, *The diophantine Frobenius problem* (Oxford University Press, 2005).

Sel77 E. Selmer, *On the linear Diophantine problem of Frobenius*, J. Reine Angew. Math. **293/294** (1977), 1–17.

Sie45 C. L. Siegel, *A mean value theorem in geometry of numbers*, Ann. of Math. (2) **46** (1945), 340–347.

Str12 A. Strömbergsson, *On the limit distribution of Frobenius numbers*, Acta Arith. **152** (2012), 81–107.

SV05 A. Strömbergsson and A. Venkatesh, *Small solutions to linear congruences and Hecke equidistribution*, Acta Arith. **118** (2005), 41–78.

Ust10 A. Ustinov, *On the distribution of Frobenius numbers with three arguments*, Izv. RAN. Ser. Mat. **74** (2010), 145–170.

Han Li lihan85math@gmail.com
Department of Mathematics, Yale University, New Haven, CT 06520, USA

Current address: Department of Mathematics, The University of Texas at Austin,
Austin, TX 78750, USA.