



# COMPOSITIO MATHEMATICA

## Equidistribution of primitive rational points on expanding horospheres

Manfred Einsiedler, Shahar Mozes, Nimish Shah and Uri Shapira

Compositio Math. **152** (2016), 667–692.

[doi:10.1112/S0010437X15007605](https://doi.org/10.1112/S0010437X15007605)



FOUNDATION  
COMPOSITIO  
MATHEMATICA



LONDON  
MATHEMATICAL  
SOCIETY  
EST. 1865



# Equidistribution of primitive rational points on expanding horospheres

Manfred Einsiedler, Shahar Mozes, Nimish Shah and Uri Shapira

## ABSTRACT

We confirm a conjecture of Marklof regarding the limiting distribution of certain sparse collections of points on expanding horospheres. These collections are obtained by intersecting the expanded horosphere with a certain manifold of complementary dimension and turns out to be of arithmetic nature. This result is then used along the lines suggested by Marklof to give an analogue of a result of Schmidt regarding the distribution of shapes of lattices orthogonal to integer vectors.

## 1. Main results

### 1.1 The main theorem

Given integers  $m \geq n \geq 1$ , let  $d \stackrel{\text{def}}{=} n + m$  and consider the space  $X_d \stackrel{\text{def}}{=} \text{SL}_d(\mathbb{R})/\text{SL}_d(\mathbb{Z})$  on which  $G \stackrel{\text{def}}{=} \text{SL}_d(\mathbb{R})$  and its subgroups act. Unless otherwise stated, when we write an element  $g \in G$  as a matrix  $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ , we shall mean that  $A, B, C, D$  represent matrices of dimensions  $m \times m$ ,  $m \times n$ ,  $n \times m$ , and  $n \times n$ , respectively. We refer to these as the *block components* of  $g$ . We shall denote the identity matrix in dimension  $k$  by  $I_k$  and often write just  $I = I_k$  if the dimension is clear from the context. Similarly,  $0$  will denote the zero matrix in various dimensions.

Consider the following subgroups of  $G$ :

$$\begin{aligned} U &\stackrel{\text{def}}{=} \left\{ \begin{pmatrix} I & 0 \\ \mathbf{u} & I \end{pmatrix} : \mathbf{u} \in \text{Mat}_{n \times m}(\mathbb{R}) \right\}, \\ V &\stackrel{\text{def}}{=} \left\{ \begin{pmatrix} I & \mathbf{v} \\ 0 & I \end{pmatrix} : \mathbf{v} \in \text{Mat}_{m \times n}(\mathbb{R}) \right\}, \\ H &\stackrel{\text{def}}{=} \left\{ \begin{pmatrix} h & \mathbf{v} \\ 0 & I \end{pmatrix} : h \in \text{SL}_m(\mathbb{R}), \mathbf{v} \in \text{Mat}_{m \times n}(\mathbb{R}) \right\}. \end{aligned}$$

Also define the diagonal matrices

$$a(y) \stackrel{\text{def}}{=} \begin{pmatrix} y^{-n/m} I_m & 0 \\ 0 & y I_n \end{pmatrix} \quad \text{for } y \in \mathbb{R}_{>0}.$$

We denote by  $x_0 \in X_d$  the identity coset  $\text{SL}_d(\mathbb{Z})$  and set, for  $\mathbf{u} \in \text{Mat}_{n \times m}(\mathbb{R})$ ,

$$x_{\mathbf{u}} \stackrel{\text{def}}{=} \begin{pmatrix} I & 0 \\ \mathbf{u} & I \end{pmatrix} x_0.$$

---

Received 18 June 2013, accepted in final form 11 May 2015, published online 9 November 2015.

*2010 Mathematics Subject Classification* 11-xx, 37-xx.

*Keywords:* equidistribution, homogeneous spaces, shapes of orthogonal lattices.

M. E. acknowledges support of SNF grant 200021-127145. S. M. acknowledges the support of ISF grant 1003/11, BSF grant 2010295, and the University of Zurich. N. S. acknowledges the support of NSF grant 1001654. Chaya Fellow U. S., acknowledges the partial support of the Advanced Research Grant 228304 from the European Research Council and ISF grant 357/13.

This journal is © Foundation Compositio Mathematica 2015.

DEFINITION 1.1. Given a subgroup  $L < G$  and a point  $x \in X_d$  we say that the orbit  $Lx$  is *periodic* if it supports an  $L$ -invariant probability measure. In this case we denote by  $\mathbf{m}_{Lx}$  the unique  $L$ -invariant probability measure supported on  $Lx$  and refer to it as the *Haar* measure on the orbit.

The orbits  $Ux_0, Vx_0$  and  $Hx_0$  are periodic and their respective Haar measures will play a prominent role in our discussion. In our discussion the cases  $m > n$  and  $m = n$  will exhibit different phenomena and in order to be able to have unified statements we use the following notation throughout:

$$\boldsymbol{\theta} \stackrel{\text{def}}{=} \begin{cases} \mathbf{m}_{Vx_0} & \text{if } m = n, \\ \mathbf{m}_{Hx_0} & \text{if } m > n. \end{cases}$$

DEFINITION 1.2. We say that a matrix  $\mathbf{u} \in \text{Mat}_{n \times m}(\mathbb{Z})$  is *k-primitive*, where  $k$  is a positive integer, if the reduction modulo  $k$  of its columns span  $(\mathbb{Z}/k\mathbb{Z})^n$ .

Consider the finite<sup>1</sup> set

$$\tilde{\mathcal{P}}_k \stackrel{\text{def}}{=} \{(x_{k^{-1}\mathbf{u}}, x_{k^{-1}\mathbf{u}}) : \mathbf{u} \text{ is } k\text{-primitive}\} \subset Ux_0 \times Ux_0, \tag{1.1}$$

and let  $\tilde{\boldsymbol{\mu}}_k$  denote the normalized counting measure on  $\tilde{\mathcal{P}}_k$ ; i.e.

$$\tilde{\boldsymbol{\mu}}_k \stackrel{\text{def}}{=} \frac{1}{|\tilde{\mathcal{P}}_k|} \sum_{(x,x) \in \tilde{\mathcal{P}}_k} \delta_{(x,x)}. \tag{1.2}$$

Finally, let us denote  $\tilde{a}(k) = (e, a(k)) \in G \times G$ . With this notation we can state our main result.

THEOREM 1.3. As  $k \rightarrow \infty$ ,  $\tilde{a}(k)_* \tilde{\boldsymbol{\mu}}_k \xrightarrow{w^*} \mathbf{m}_{Ux_0} \times \boldsymbol{\theta}$ .

We remark that our main tool is [Sha98], which extends the fundamental work of Ratner [Rat91a]. We also note that some of our arguments are similar to [MS95].

### 1.2 An application

Theorem 1.3 is a generalization of a result by Marklof [Mar10, Theorem 6]. Marklof’s result has various applications, most notably to the distribution of Frobenius numbers, circulant graphs, and the shapes of co-dimension 1 primitive subgroups of  $\mathbb{Z}^d$  (see [Mar10, MS13]). Naturally, in each of these discussions an application of Theorem 1.3 gives new results. We give one such application which is an analogue of a certain equidistribution result of Schmidt [Sch98]. We follow closely the viewpoint of [Mar10].

Assume  $n = 1, m \geq 2$  so that  $d = m + 1 \geq 3$ . The quotient  $Z_m \stackrel{\text{def}}{=} \text{SO}_m(\mathbb{R}) \backslash X_m$  will be referred to as the *space of shapes of m-dimensional lattices*. We equip  $Z_m$  with the probability measure  $\mathbf{m}_{Z_m}$  which is by definition the image of  $\mathbf{m}_{X_m}$  under the natural projection.

For any integer vector  $v \in \mathbb{Z}^d$ , let us denote  $\Lambda_v \stackrel{\text{def}}{=} \mathbb{Z}^d \cap \{v\}^\perp$ ; that is,  $\Lambda_v$  is the lattice of integer points in the  $m$ -dimensional orthocomplement of  $v$  in  $\mathbb{R}^d$ . We may choose a matrix  $k_v \in \text{SO}_d(\mathbb{R})$  so that  $k_v \Lambda_v \subset \mathbb{R}^m \times \{0\} \subset \mathbb{R}^d$ , and so that  $k_v v$  lies on the positive half of the  $d$ th coordinate axis. After normalizing the covolume of  $k_v \Lambda_v$  to be 1, we obtain a point in  $X_m$ . As the choice of  $k_v$  is only well defined up to the action of  $\text{SO}_m(\mathbb{R})$ , we obtain a well-defined point in  $Z_m$  which we denote henceforth by  $[\Lambda_v]$ . There is a certain redundancy in considering  $\Lambda_v$  if  $v \in \mathbb{Z}^d$  is non-primitive (that is, if it is an integer multiple of another integer vector). We therefore

---

<sup>1</sup> Note that for two integer matrices, if  $\mathbf{u}_1 = \mathbf{u}_2 \pmod k$  then  $x_{k^{-1}\mathbf{u}_1} = x_{k^{-1}\mathbf{u}_2}$ .

denote by  $\widehat{\mathbb{Z}}^d$  the subset of *primitive* integer vectors. Note that a vector  $(u_1, \dots, u_m, k) \in \mathbb{Z}^d$  is primitive if and only if the vector  $(u_1, \dots, u_m) \in \mathbb{Z}^m$  is  $k$ -primitive as in Definition 1.2. Finally, let  $B_\infty \stackrel{\text{def}}{=} \{x \in \mathbb{R}^d : \|x\|_\infty \leq 1\}$  and  $\partial B_\infty$  denote its boundary. As an application of Theorem 1.3 we prove the following result in § 8.

THEOREM 1.4. *Let  $\mathcal{F} \subset \partial B_\infty$  be a measurable set of positive measure having a boundary in  $\partial B_\infty$  of measure 0 with respect to the  $m$ -dimensional Lebesgue measure on  $\partial B_\infty$ . For any positive integer  $k$ , let*

$$\eta_k \stackrel{\text{def}}{=} \frac{1}{|\widehat{\mathbb{Z}}^d \cap k\mathcal{F}|} \sum_{v \in \widehat{\mathbb{Z}}^d \cap k\mathcal{F}} \delta_{[\Lambda_v]}.$$

Then  $\eta_k \xrightarrow{w^*} \mathbf{m}_{Z_m}$  as  $k \rightarrow \infty$ .

Note that the choice of the  $\ell_2$ -norm yields a much more elegant statement as we have the relation

$$\{\Lambda_v : v \in \widehat{\mathbb{Z}}^d, \|v\|_2 = k\} = \left\{ \begin{array}{l} \text{Primitive } m\text{-dimensional subgroups} \\ \text{of } \mathbb{Z}^d \text{ of covolume } = k \end{array} \right\}.$$

We suggest the following conjecture.<sup>2</sup>

CONJECTURE 1.5. Let  $d \geq 3$  and let  $B_2(r)$  denote the Euclidean ball of radius  $r > 0$  in  $\mathbb{R}^d$ . Then if  $r_n \rightarrow \infty$  is a sequence of radii such that  $|\partial B_2(r_n) \cap \widehat{\mathbb{Z}}^d| \rightarrow \infty$ , then the collections  $\{[\Lambda_v] : v \in \partial B_2(r_n) \cap \widehat{\mathbb{Z}}^d\}$  equidistribute in  $Z_m$ .

## 2. Outline and initial steps

### 2.1 A motivating low-dimensional example

We now prove Theorem 1.3 in the case  $n = m = 1$  to which the techniques of the later parts of this paper do not apply. We learned the argument we give here from Marklof, and to the best of our understanding this result was his reason for anticipating the validity of Theorem 1.3.

We identify both of the orbits  $Ux_0 = \left\{ \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix} x_0 : s \in \mathbb{R} \right\}$  and  $Vx_0 = \left\{ \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} x_0 : t \in \mathbb{R} \right\}$  with  $\mathbb{R}/\mathbb{Z}$  and use the parameters  $s, t$  respectively to describe them. We wish to analyze for which values of  $y$

$$a(y)Ux_0 \cap Vx_0 \neq \emptyset, \tag{2.1}$$

and if this happens, we wish to understand the asymptotics of the joint distribution of the set

$$\left\{ (s, t) \in (\mathbb{R}/\mathbb{Z})^2 : a(y) \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix} x_0 = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} x_0 \right\}. \tag{2.2}$$

Following the definitions, one sees that  $(s, t)$  is in the set in (2.2) if and only if there exists  $\begin{pmatrix} k & \ell \\ m & n \end{pmatrix} \in \text{SL}_2(\mathbb{Z})$  which solves the equation

$$\begin{pmatrix} y^{-1} & 0 \\ ys & y \end{pmatrix} \begin{pmatrix} k & \ell \\ m & n \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}. \tag{2.3}$$

Equation (2.3) has a solution if and only if the following three conditions on the variables  $y, s, t$  are satisfied (note that we may assume  $s, t \in (0, 1]$ ): (i)  $y = k$  is a positive integer; (ii)  $t = \ell/k$  for some  $1 \leq \ell \leq k$  with  $\text{gcd}(\ell, k) = 1$ ; and (iii)  $s = \ell^*/k$ , where  $\ell^*$  is the number  $1 \leq \ell^* \leq k$  satisfying  $\ell \cdot \ell^* = 1 \pmod k$ .

---

<sup>2</sup> Significant progress towards the suggested conjecture was obtained recently in [AES14, AES15].

To summarize, if (2.1) holds then  $y = k$  for some positive integer  $k$  and the set in (2.2) equals

$$\{(\ell/k, \ell^*/k) \in (\mathbb{R}/\mathbb{Z})^2 : 1 \leq \ell \leq k, \gcd(\ell, k) = 1\}. \tag{2.4}$$

Taking into account both descriptions (2.2), (2.4) and our identification  $(\mathbb{R}/\mathbb{Z})^2 \cong Ux_0 \times Vx_0$ , we see that this set is exactly  $\tilde{a}(k)\tilde{\mathcal{P}}_k$  which appears as the support of the measure in Theorem 1.3. If we denote the normalized counting measure on (2.4) by  $\theta_k$  then the statement of Theorem 1.3 is interpreted as saying that  $\theta_k$  equidistributes in  $(\mathbb{R}/\mathbb{Z})^2$  as  $k \rightarrow \infty$ . This equidistribution statement translates to estimating Kloosterman sums

$$K(a, b, k) = \sum_{\substack{1 \leq x \leq k \\ \gcd(x, k) = 1}} e^{2\pi i(ax + bx^*)} = \phi(k) \int_{(\mathbb{R}/\mathbb{Z})^2} e^{2\pi i(at + bs)} d\theta_k, \tag{2.5}$$

where  $\phi(k)$  is the Euler function. The known estimates for Kloosterman sums (see, for example, [Iwa02, p. 48, Equation (2.25)]), imply that for any choice of  $(a, b) \in \mathbb{Z}^2$  not both 0,  $\phi(k)^{-1}K(a, b, k) \rightarrow 0$  as  $k \rightarrow \infty$ , which establishes the desired equidistribution of  $\theta_k$ .

This establishes the case  $m = n = 1$  in Theorem 1.3. The main objective of this paper is to establish the case  $m \geq 2$  using techniques from homogeneous dynamics.

### 2.2 A basic observation and the structure of the proof

Similarly to (1.1) and (1.2), let us define

$$\mathcal{P}_k \stackrel{\text{def}}{=} \{x_{k^{-1}\mathbf{u}} : \mathbf{u} \text{ is } k\text{-primitive}\} \subset Ux_0 \subset X_d, \tag{2.6}$$

$$\mu_k \stackrel{\text{def}}{=} \frac{1}{|\mathcal{P}_k|} \sum_{x \in \mathcal{P}_k} \delta_x. \tag{2.7}$$

Using the fact that the  $a(y)$ -action is mixing on  $X_d$ , one can show that the pushed periodic orbit  $a(y)Ux_0$  equidistributes in  $X_d$  when  $y \rightarrow \infty$ , as  $U$  is the expanding horospherical subgroup of  $a(y)$  (see, for example, [Mar04]). On the other hand, the collection  $\mathcal{P}_k \subset Ux_0$  is composed of rational points  $x$  for which the trajectories  $\{a(y)x\}_{y>1}$  are divergent in  $X_d$ . It is therefore natural to investigate the tension between these two facts and to analyze the distribution of  $a(y_k)\mathcal{P}_k$  in  $X_d$  for various choices of sequences  $y_k \rightarrow \infty$ . Setting  $y_k = k$  as in Theorem 1.3 is natural because of the following lemma (which is in some sense the starting point of our discussion).<sup>3</sup>

LEMMA 2.1 (Basic lemma). *For any positive integer  $k$ ,  $a(k)\mathcal{P}_k \subset Hx_0$ .*

We prove Lemma 2.1 towards the end of this section. It shows that we cannot expect the sequence  $\tilde{a}(k)_*\tilde{\mu}_k$  to equidistribute in  $Ux_0 \times X_d$  as any limit point of this sequence is clearly a measure<sup>4</sup> supported in  $Ux_0 \times Hx_0$ .

The space  $X_m = \text{SL}_m(\mathbb{R})/\text{SL}_m(\mathbb{Z})$  is naturally embedded in  $Hx_0 \subset X_d$ , simply by identifying  $\text{SL}_m(\mathbb{R})$  as a subgroup of  $H$  in the obvious way. Throughout we alternate between thinking of  $X_m$  as a subset of  $X_d$  and as the space of  $m$ -dimensional unimodular lattices in  $\mathbb{R}^m$ . When thinking of  $X_m$  as the space of  $m$ -dimensional lattices, the identity coset  $x_0$  corresponds to the lattice  $\mathbb{Z}^m$ . There is a natural projection  $\pi_3 : Hx_0 \rightarrow X_m$  defined by

$$\text{for } x = \begin{pmatrix} h & \mathbf{v} \\ 0 & I \end{pmatrix} x_0 \in Hx_0, \quad \pi_3(x) \stackrel{\text{def}}{=} \begin{pmatrix} h & 0 \\ 0 & I \end{pmatrix} \text{SL}_d(\mathbb{Z}),$$

<sup>3</sup> See Remark 2.6 for other natural choices of  $y_k$ .

<sup>4</sup> A priori the limit measure is not even known to be a probability measure.

which corresponds to the  $m$ -dimensional unimodular lattice  $h\mathbb{Z}^m$ . Let

$$\nu_k \stackrel{\text{def}}{=} (\pi_3)_* a(k)_* \mu_k, \tag{2.8}$$

and consider the diagram with natural projection maps

$$\begin{array}{ccc}
 & (Ux_0 \times Hx_0, \tilde{a}(k)_* \tilde{\mu}_k) & \\
 \swarrow \pi_1 & & \searrow \pi_2 \\
 (Ux_0, \mu_k) & & (Hx_0, a(k)_* \mu_k) \\
 & & \downarrow \pi_3 \\
 & & (X_m, \nu_k)
 \end{array} \tag{2.9}$$

We describe the structure of the proof of Theorem 1.3 using diagram (2.9).

- Step 1. Establish the convergence  $\mu_k \xrightarrow{w^*} \mathbf{m}_{Ux_0}$ , which takes place on the left-hand side of diagram (2.9).
- Step 2. Show the convergence  $\nu_k \xrightarrow{w^*} (\pi_3)_* \theta$ , which equals  $\mathbf{m}_{X_m}$  in the case  $m > n$ , and the Dirac measure  $\delta_{x_0}$  in the case  $m = n$ . This convergence takes place on the right-hand side at the bottom of diagram (2.9).
- Step 3. Use Step 2 to establish the convergence  $a(k)_* \mu_k \xrightarrow{w^*} \theta$  taking place on the right-hand side of diagram (2.9). Many of the ideas appearing in the argument of Step 3 already appear in simplified versions in the proof of Step 1.
- Step 4. Combine Steps 1 and 3 and use a disjointness argument to prove the convergence  $\tilde{a}(k)_* \tilde{\mu}_k \xrightarrow{w^*} \mathbf{m}_{Ux_0} \times \theta$  taking place at the top of diagram (2.9).

We comment here that one could prove Step 1 quite easily using Fourier transform arguments. We choose a different approach that is more compatible with the proofs of the other steps (see the footnote following Lemma 4.8).

### 2.3 The method of proof

From this point on we will assume  $m \geq 2$ . As described above, in the course of the proof of Theorem 1.3 we will frequently need to establish a convergence  $\eta_n \xrightarrow{w^*} \eta$  of probability measures. It will turn out that all the measures involved are  $\Lambda$ -invariant under natural actions of the group<sup>5</sup>

$$\Lambda \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} : \delta_1 \in \text{SL}_m(\mathbb{Z}), \delta_2 \in \text{SL}_n(\mathbb{Z}) \right\}.$$

The following is the strategy we will use for proving such a claim. One starts by *classifying* the  $\Lambda$ -ergodic measures and shows that there are only countably many such. Say  $\{\sigma_i\}_{i=0}^\infty$ , with  $\eta = \sigma_0$ . Then, given an accumulation point  $\eta'$  of  $\eta_n$ , the ergodic decomposition of  $\eta'$  is given by  $\eta' = \sum_{i=0}^\infty c_i \sigma_i$  with  $c_i \geq 0$  and  $\sum c_i = 1$ . One then appeals to a *non-accumulation* result to show that the only possibility of  $c_i$  being positive is that  $\text{supp}(\eta_n) \subset \text{supp}(\sigma_i)$  for infinitely many values of  $n$ . One then verifies that the only  $i$  for which such an inclusion is possible is  $i = 0$ .

<sup>5</sup> Note that the assumption that  $m \geq 2$  is equivalent to  $\Lambda$  being non-trivial.

The above strategy is constructed from (i) a non-accumulation result and (ii) a measure classification result. In § 3 we prove the non-accumulation result. In § 4 we prove various measure classification results. In § 6 we use the above strategy to prove Theorem 1.3 through Steps 1–4 which are described after (2.9).

### 2.4 Elementary divisors

In this subsection we gather together some further preliminaries that will be needed throughout the paper. We shall use the following theorem, the proof of which can be found in [Swi01, Appendix Lemma A2, Theorem A1].

**THEOREM 2.2 (Elementary divisors).** *Let  $\Sigma \neq \{0\}$  be a subgroup of  $\mathbb{Z}^m$ . Then there exist an integer  $1 \leq r \leq m$  and positive integers  $\ell_1, \dots, \ell_r$  such that  $\ell_i | \ell_{i+1}$  and such that one can find a basis  $v_1, \dots, v_m$  of  $\mathbb{Z}^m$  for which  $\ell_1 v_1, \dots, \ell_r v_r$  form a basis for  $\Sigma$ . Furthermore, the numbers  $r, \ell_i$  with this property are unique and are called the rank and the elementary divisors of  $\Sigma$  with respect to  $\mathbb{Z}^m$ , respectively.*

The following restatement of Theorem 2.2 will be more convenient for us.

**LEMMA 2.3.** *Consider the action of  $\begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} \in \Lambda$  on  $\mathbf{u} \in \text{Mat}_{n \times m}(\mathbb{Z})$  given by  $\begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} \mathbf{u} = \delta_2 \mathbf{u} \delta_1^{-1}$ . Each  $\Lambda$ -orbit in  $\text{Mat}_{n \times m}(\mathbb{Z})$  contains an element of the form*

$$\mathbf{u} = \begin{pmatrix} \ell_1 & 0 & \dots & 0 \\ 0 & \ddots & 0 & \dots & 0 \\ 0 & & \ell_n & \dots & 0 \end{pmatrix} \tag{2.10}$$

such that the  $\ell_i$  are integers satisfying  $\ell_1 | \ell_2 | \dots | \ell_n$ . Moreover, the integers  $\ell_i$  are unique up to sign.

*Proof.* The lemma follows from an application of Theorem 2.2 to the group  $\Sigma$  generated by the rows of  $\mathbf{u}$  in  $\mathbb{Z}^m$ . Note that if  $\text{rank}(\Sigma) = r < n$  then the tuple  $(\ell_1, \dots, \ell_n)$  ends with  $n - r$  zeros. Note also that the only case where we cannot require that  $\ell_i \geq 0$  is when  $n = m$  and  $\mathbf{u}$  is an invertible matrix with negative determinant. □

Let us refer to the tuple  $(\ell_1, \dots, \ell_n)$  attached to the  $\Lambda$ -orbit of  $\mathbf{u}$  as its *elementary divisors tuple*.

Since  $k$ -primitivity as defined in Definition 1.2 is invariant under the  $\Lambda$ -action, it is clear that  $\mathbf{u}$  is  $k$ -primitive if and only if  $\ell_n$  is coprime to  $k$ . For the proof of Lemma 2.1 we will use the following characterization of  $k$ -primitivity.

**LEMMA 2.4.** *The matrix  $\mathbf{u} \in \text{Mat}_{n \times m}(\mathbb{Z})$  is  $k$ -primitive if and only if there exists a matrix  $\gamma \in \text{SL}_d(\mathbb{Z})$  whose bottom  $n$  rows coincide with the  $n \times d$  matrix  $(\mathbf{u} \ kI_n)$ .*

*Proof.* It is straightforward to show that the property described in the statement of the lemma is also invariant under the action of  $\Lambda$ . It follows that it is enough to verify the validity of the statement for matrices  $\mathbf{u}$  in the form (2.10). Clearly if the elementary divisor  $\ell_n$  is not coprime to  $k$ , that is, if  $\mathbf{u}$  is not  $k$ -primitive, then the bottom row of  $(\mathbf{u} \ kI_n)$  is not a primitive vector in  $\mathbb{Z}^d$  and so the existence of  $\gamma$  as in the statement is ruled out. If, on the other hand,  $\mathbf{u}$  is  $k$ -primitive then the elementary divisors  $\ell_i$  are all coprime to  $k$ . It follows that there are integers  $e_i, f_i$  such that  $\det \begin{pmatrix} e_i & f_i \\ \ell_i & k \end{pmatrix} = 1$ . The determinant of the  $d \times d$  matrix

$$\gamma \stackrel{\text{def}}{=} \begin{pmatrix} \text{diag}(e_1 \dots e_n) & 0 & \text{diag}(f_1 \dots f_n) \\ 0 & I_{n-m} & 0 \\ \text{diag}(\ell_1 \dots \ell_n) & 0 & \text{diag}(k \dots k) \end{pmatrix} \tag{2.11}$$

equals  $\prod_1^n \det \begin{pmatrix} e_i & f_i \\ \ell_i & k \end{pmatrix} = 1$  (because after conjugating by a suitable permutation matrix, the matrix  $\gamma$  transforms into a block-diagonal matrix with diagonal blocks being equal to the  $\begin{pmatrix} e_i & f_i \\ \ell_i & k \end{pmatrix}$  complemented by  $m - n$  1's).  $\square$

**2.5 Proof of the basic lemma**

*Proof of Lemma 2.1.* Note that in order to show that a point  $x = gx_0 \in X_d$  lies in  $Hx_0$ , one needs to show that there exists  $\gamma \in \Gamma$  such that  $g\gamma \in H$ . Let  $\mathbf{u} \in \text{Mat}_{n \times m}(\mathbb{Z})$  be  $k$ -primitive and  $a(k)x_{k^{-1}\mathbf{u}} \in a(k)\mathcal{P}_k$  be the corresponding point. By Lemma 2.4,  $\mathbf{u}$  is  $k$ -primitive if and only if there exists a matrix  $\gamma^{-1} \in \Gamma$  whose bottom  $n$  rows are given by the rows of the  $n \times d$  matrix  $(\mathbf{u} \ kI_n)$ . It follows that if we denote by  $A_\gamma, B_\gamma, C_\gamma, D_\gamma$  the block components of  $\gamma$ , then

$$a(k) \begin{pmatrix} I & 0 \\ k^{-1}\mathbf{u} & I \end{pmatrix} \gamma = \begin{pmatrix} k^{-n/m}I & 0 \\ \mathbf{u} & kI \end{pmatrix} \begin{pmatrix} A_\gamma & B_\gamma \\ C_\gamma & D_\gamma \end{pmatrix} = \begin{pmatrix} k^{-n/m}A_\gamma & k^{-n/m}B_\gamma \\ 0 & I \end{pmatrix}. \tag{2.12}$$

The above equation shows that the point  $a(k)x_{k^{-1}\mathbf{u}}$  belongs to  $Hx_0$  as desired.  $\square$

*Remark 2.5.* Equation (2.12) (which is analogous to (2.3)), is fundamental for our discussion and deserves some attention. We note two things.

(1) Let  $\mathbf{u}$  be  $k$ -primitive and suppose  $\gamma$  solves (2.12). By considering determinants we see that  $A_\gamma \in \text{Mat}_m(\mathbb{Z})$  must have determinant  $k^n$ . This means that when considered as a lattice in  $\mathbb{R}^m$ ,  $\pi_3(a(k)x_{k^{-1}\mathbf{u}})$  equals  $k^{-n/m}A_\gamma\mathbb{Z}^m$ , and up to homothety equals  $A_\gamma\mathbb{Z}^m$  which is a subgroup of index  $k^n$  of  $\mathbb{Z}^m$ . In Lemma 6.1 we show that the collection  $\{\pi_3(a(k)x_{k^{-1}\mathbf{u}}) : \mathbf{u} \text{ is } k\text{-primitive}\}$  consists of all such lattices of a given *Hecke-type* (see Definition 5.1 for terminology).

(2) Assume again that  $\mathbf{u}$  is  $k$ -primitive and that  $\gamma$  solves (2.12). The first  $m$  columns of  $\gamma$  form a basis for the discrete group of rank  $m$  which we denote by  $\Lambda_{(\mathbf{u},k)}$ , consisting of integer vectors in the orthocomplement of the linear space spanned by the rows of the  $n \times d$  matrix  $(\mathbf{u} \ kI)$ . It follows that  $\pi_3(a(k)x_{k^{-1}\mathbf{u}})$ , as a lattice in  $\mathbb{R}^m$ , is (up to homothety) the projection of  $\Lambda_{(\mathbf{u},k)}$  onto the copy of  $\mathbb{R}^m$  given by the first  $m$ -coordinates. This is what furnishes the link with Schmidt's theorem and its strengthening given in Theorem 1.4.

*Remark 2.6.* Let  $y_k \rightarrow \infty$  be given. By Lemma 2.1,  $a(y_k)\mathcal{P}_k = a(y_k/k)a(k)\mathcal{P}_k \subset a(y_k/k)Hx_0$ . Thus, if we set  $y_k = k^\alpha$  for some positive  $\alpha$ , then we obtain that  $a(k^\alpha)\mathcal{P}_k \subset a(k^{\alpha-1})Hx_0$ . In the case  $\alpha > 1$ , the collection  $a(k^\alpha)\mathcal{P}_k$  is therefore contained in the uniformly divergent periodic orbit  $a(k^{\alpha-1})Hx_0$ , and in particular, the sequence of measures  $a(k^\alpha)_*\mu_k$  converges to the zero measure on  $X_d$ . In the case  $\alpha < 1$ , the collection  $a(k^\alpha)\mathcal{P}_k$  is therefore contained in the equidistributing periodic orbit  $a(k^{\alpha-1})Hx_0$ . It turns out that one can prove an analogue of Theorem 1.3 in this context. This analysis is non-trivial and we plan on elaborating on it in a future paper.

**3. Non-accumulation**

Our goal in this section is to prove a certain non-accumulation result (Theorem 3.1) that will be used in various steps in the proof of Theorem 1.3. Although we do not aim at greatest generality, we still choose to state and prove the results in a somewhat abstract setting (if only to isolate the necessary features that are needed for the result to hold). To this end, in this section only, we abandon the notation introduced so far and make the following assumptions. Let  $G$  be a real Lie group, let  $\Gamma < G$  be a lattice, and let  $\Lambda, L < G$  be closed subgroups with  $\Lambda$  being discrete and generated by finitely many Ad-unipotent elements. Assume, furthermore, that there is a decomposition

$$\text{Lie}(G) = \text{Lie}(L) \oplus W \tag{3.1}$$



such that  $W$  is invariant under the action of  $\Lambda$  via the adjoint representation. Let  $X = G/\Gamma$  and  $z \in X$  be a point such that the orbit  $Lz$  is periodic and  $\Lambda$ -invariant.

**THEOREM 3.1 (Non-accumulation).** *Let  $G, \Gamma, \Lambda, L, z, W, X$  be as above and assume the  $\Lambda$ -representation on  $W$  does not contain any fixed vectors. Let  $P_k \subset X$  be a sequence of finite  $\Lambda$ -invariant sets and  $\mu_k$  the normalized counting measure on  $P_k$ . If  $P_k \cap Lz = \emptyset$  for all  $k$  then any weak\* accumulation point  $\sigma$  of  $\{\mu_k\}_{k=1}^\infty$  satisfies  $\sigma(Lz) = 0$ .*

We use the absolute value symbol  $|\cdot|$  to denote the usual absolute value of a real number as well as the Lebesgue measure of a set and the cardinality of a finite set. This should not cause any confusion. In the proof of Theorem 3.1 we will need to use some elementary properties of polynomials which we now recall. The following lemma may be found in [KSS02, Proposition 3.2.2].

**LEMMA 3.2.** *For any degree  $d$  there exists a constant  $c_d > 0$  such that for any polynomial  $p : \mathbb{R} \rightarrow \mathbb{R}$  of degree bounded by  $d$ , for any interval  $I \subset \mathbb{R}$ , we have that if  $\rho = \max\{|p(x)| : x \in I\}$  then for any  $0 < \epsilon \leq \rho$ ,*

$$\frac{|\{x \in I : |p(x)| \leq \epsilon\}|}{|I|} \leq c_d \left(\frac{\epsilon}{\rho}\right)^{1/d}. \tag{3.2}$$

We deduce the following integer-value version of this lemma.

**LEMMA 3.3.** *For any degree  $d$  there exists a constant  $c_d > 0$  such that for any polynomial  $p : \mathbb{N} \rightarrow \mathbb{R}$  of degree bounded by  $d$ , for any interval  $J \subset \mathbb{N}$ , we have that if  $\rho = \max\{|p(n)| : n \in J\}$  then for any  $0 < \epsilon \leq \rho$ ,*

$$\frac{|\{n \in J : |p(n)| \leq \epsilon\}|}{|J|} \leq c_d \left(\frac{\epsilon}{\rho}\right)^{1/d} + \frac{d}{|J|}. \tag{3.3}$$

*Proof.* Let  $I \subset \mathbb{R}$  denote the real interval defined as the convex hull of  $J$ . For  $\epsilon > 0$  let  $I_\epsilon \stackrel{\text{def}}{=} \{x \in I : |p(x)| \leq \epsilon\}$  and  $J_\epsilon \stackrel{\text{def}}{=} \{n \in J : |p(n)| \leq \epsilon\}$ . The set  $I_\epsilon$  is a disjoint union of finitely many closed intervals  $I_{\epsilon,\ell}$  for  $\ell = 1 \dots k$ . We have that

$$\begin{aligned} |J_\epsilon| &= |I_\epsilon \cap \mathbb{Z}| = \sum_\ell |I_{\epsilon,\ell} \cap \mathbb{Z}| \\ &= \sum_{\ell: |I_{\epsilon,\ell} \cap \mathbb{Z}|=1} 1 + \sum_{\ell: |I_{\epsilon,\ell} \cap \mathbb{Z}|>1} |I_{\epsilon,\ell} \cap \mathbb{Z}| \\ &\leq d + \sum_{\ell: |I_{\epsilon,\ell} \cap \mathbb{Z}|>1} 2|I_{\epsilon,\ell}| \leq d + 2I_\epsilon, \end{aligned}$$

where we used that if  $I_{\epsilon,\ell}$  contains more than one integer then

$$|I_{\epsilon,\ell} \cap \mathbb{Z}| \leq 2|I_{\epsilon,\ell}|,$$

and also that the number of  $\ell$  for which  $I_{\epsilon,\ell}$  contains a single integer is bounded by the degree  $d$  of the polynomial because between each pair of such intervals there must be a zero of the derivative.

Inequality (3.3) now follows from (3.2) (with a slightly bigger constant  $c_d$ ). □

*Proof of Theorem 3.1.* Let  $\sigma$  be a weak\* accumulation point as in the statement. Let  $\Omega_1 \subset X$  be a compact set. We will show that for  $K_1 \stackrel{\text{def}}{=} \Omega_1 \cap Lz$  one has  $\sigma(K_1) = 0$ . This is enough as  $Lz$  is a countable union of such sets. Choose some norm on  $\text{Lie}(G)$  and denote by  $B_\epsilon^W$  the ball of radius  $\epsilon$  around 0 in  $W$ . Choose an open set  $\tilde{\Omega}_1$  and a compact set  $\Omega_2$  such that (i)  $\Omega_1 \subset \tilde{\Omega}_1 \subset \Omega_2$ , (ii)  $\tilde{\Omega}_1 \cap (\exp B_1^W \cdot (X \setminus \Omega_2)) = \emptyset$ ; in other words,  $\Omega_2$  is big enough so that one cannot reach  $\tilde{\Omega}_1$  by acting on points outside of  $\Omega_2$  by elements of the form  $\exp w$ , where  $w \in W$  is of norm less than or equal to 1. Let  $\tilde{K}_1 \stackrel{\text{def}}{=} \tilde{\Omega}_1 \cap Lz$  and  $K_2 \stackrel{\text{def}}{=} \Omega_2 \cap Lz$  and for any  $\epsilon > 0$  any subset  $F \subset Lz$  denote by  $\mathcal{T}(\epsilon, F) \stackrel{\text{def}}{=} \{\exp(w)x : x \in F, w \in W, \text{ and } \|w\| \leq \epsilon\}$  the  $\epsilon$ -tube around  $F$ .

There exists  $0 < \epsilon_0 < 1/2$  small enough so that the map  $(w, x) \mapsto \exp(w)x$  from  $B_{\epsilon_0}^W \times K_2 \rightarrow \mathcal{T}(\epsilon_0, K_2)$  is a homeomorphism onto its image. This gives a natural coordinate system on the  $\epsilon_0$ -tube around  $K_2$ ; we denote for  $y \in \mathcal{T}(\epsilon_0, K_2)$  by  $w_y \in B_{\epsilon_0}^W$  the  $W$ -coordinate and by  $x_y \in Lz$  the orbit-coordinate so that the identity  $y = \exp(w_y)x_y$  holds for  $y \in \mathcal{T}(\epsilon_0, K_2)$ .

Let  $u_1, \dots, u_r \in \Lambda$  be Ad-unipotent elements that generate  $\Lambda$ . Let  $0 \leq \epsilon < \epsilon_0$  and set

$$\mathcal{S}_j(\epsilon, \tilde{K}_1) \stackrel{\text{def}}{=} \{y \in \mathcal{T}(\epsilon, \tilde{K}_1) : \text{Ad}_{u_j}(w_y) \neq w_y\},$$

$j = 1, \dots, r$ , so that  $\mathcal{T}(\epsilon, \tilde{K}_1) \setminus \tilde{K}_1 = \bigcup_1^r \mathcal{S}_j(\epsilon, \tilde{K}_1)$ . The inclusion  $\supset$  is clear. The other inclusion holds because  $\Lambda$  is generated by the  $u_j$  and  $W$  contains no  $\Lambda$ -fixed non-zero vectors. We will find a function  $\psi(\epsilon) \rightarrow_{\epsilon \rightarrow 0} 0$  such that for any  $j$ ,

$$\frac{|P_k \cap \mathcal{S}_j(\epsilon, \tilde{K}_1)|}{|P_k|} \leq \psi(\epsilon) \quad \text{for any } k. \tag{3.4}$$

Since  $P_k \cap Lz = \emptyset$ , this implies that

$$\frac{|P_k \cap \mathcal{T}(\epsilon, \tilde{K}_1)|}{|P_k|} \leq \sum_1^r \frac{|P_k \cap \mathcal{S}_j(\epsilon, \tilde{K}_1)|}{|P_k|} \leq r\psi(\epsilon) \xrightarrow{\epsilon \rightarrow 0} 0.$$

In turn, this implies that  $\sigma(K_1) = 0$  as desired because  $\mathcal{T}(\epsilon, \tilde{K}_1)$  is an open set containing the compact set  $K_1$ .

To this end, fix  $1 \leq j \leq r$  and denote, for each  $y \in \mathcal{S}_j(\epsilon, \tilde{K}_1)$ ,

$$p_y(n) \stackrel{\text{def}}{=} \|\text{Ad}_{u_j}^n(w_y)\|^2.$$

Note that as  $u_j$  is Ad-unipotent it follows that (for an appropriate choice of a norm  $\|\cdot\|$ ),  $p_y(n)$  is a non-constant polynomial in  $n$  of degree less than or equal to  $d$  for some integer  $d$  depending on  $\dim G$  only. Let us use the following notation:

- $n_y \stackrel{\text{def}}{=} \max\{n \geq 0 : p_y(k) \leq \epsilon_0^2 \text{ for } 0 \leq k \leq n\}$ ;
- $J_y \stackrel{\text{def}}{=} [0, n_y] \cap \mathbb{Z}$ ;
- $V_y \stackrel{\text{def}}{=} \{u_j^n y : n \in J_y\}$ ;
- $M \stackrel{\text{def}}{=} \max_{j=1 \dots r} \|\text{Ad}_{u_j}^{\pm 1}\|$ .

Observe that  $n_y$  is finite as  $p_y(n)$  is non-constant. We will shortly show that the following properties hold for  $y, y_1, y_2 \in \mathcal{S}_j(\epsilon, \tilde{K}_1)$ :

- (1)  $V_{y_1} \cap V_{y_2} \neq \emptyset \implies V_{y_1} \subset V_{y_2} \text{ or } V_{y_2} \subset V_{y_1}$ ;
- (2)  $V_y \cap \mathcal{S}_j(\epsilon, \tilde{K}_1) \subset \{u_j^n y : p_y(n) \leq \epsilon^2\}$ ;
- (3)  $p_y(n_y) \geq (\epsilon_0/M)^2$ ;
- (4)  $|J_y| \geq \log(\epsilon_0/\epsilon)/\log M$ .

We now conclude the proof using properties (1)–(4). Given any  $k$ , we choose a finite collection  $\{y_i\} \subset P_k \cap \mathcal{S}_j(\epsilon, \tilde{K}_1)$  so that the  $V_{y_i}$  are maximal with respect to inclusion among  $\{V_y : y \in P_k \cap \mathcal{S}_j(\epsilon, \tilde{K}_1)\}$ , and such that  $P_k \cap \mathcal{S}_j(\epsilon, \tilde{K}_1) \subset \bigcup_i V_{y_i}$ . By property (1) we deduce that the  $V_{y_i}$  are disjoint. Since  $V_{y_i} \subset P_k$ , we deduce that  $\sum_i |J_{y_i}|/|P_k| \leq 1$ . It follows that

$$\begin{aligned} \frac{|P_k \cap \mathcal{S}_j(\epsilon, \tilde{K}_1)|}{|P_k|} &= \frac{|\bigcup_i V_{y_i} \cap \mathcal{S}_j(\epsilon, \tilde{K}_1)|}{|P_k|} \leq \sum_i \frac{|J_{y_i}|}{|P_k|} \cdot \frac{|V_{y_i} \cap \mathcal{S}_j(\epsilon, \tilde{K}_1)|}{|J_{y_i}|} \\ &\stackrel{\text{by (2)}}{\leq} \sum_i \frac{|J_{y_i}|}{|P_k|} \cdot \frac{|\{n \in J_{y_i} : p_{y_i}(n) \leq \epsilon^2\}|}{|J_{y_i}|} \\ &\stackrel{\text{by (3), (4) and (3.3)}}{\leq} c_d \left(\frac{\epsilon}{\epsilon_0/M}\right)^{2/d} + \frac{d}{\log(\epsilon_0/\epsilon)/\log M}. \end{aligned}$$

As this last expression goes to 0 as  $\epsilon \rightarrow 0$  we conclude that (3.4) holds with this last expression as  $\psi(\epsilon)$ . It remains to verify the validity of (1)–(4).

(1) Assume  $y_1, y_2 \in \mathcal{S}_j(\epsilon, K_1)$  are such that  $V_{y_1} \cap V_{y_2} \neq \emptyset$ ; that is, there exists  $n_i \in J_{y_i}$  such that  $u_j^{n_1} y_1 = u_j^{n_2} y_2$ . Assume without loss of generality that  $n_2 \leq n_1$  and so  $u_j^{n_1-n_2} y_1 = y_2$  and  $n_1 - n_2 \in J_{y_1}$ . Following the definitions we see that  $V_{y_2} \subset V_{y_1}$  as desired.

(2) Let  $y \in \mathcal{S}_j(\epsilon, \tilde{K}_1)$  and assume  $n \in J_y$  is such that  $p_y(n) > \epsilon^2$  so that we know that  $\epsilon < \|\text{Ad}_{u_j^n}^n(w_y)\| \leq \epsilon_0$ . We need to show that  $u_j^n y \notin \mathcal{S}_j(\epsilon, \tilde{K}_1)$ . We have that

$$u_j^n y = u_j^n \exp(w_y) u_j^{-n} u_j^n x_y = \exp(\text{Ad}_{u_j^n}^n(w_y)) u_j^n x_y. \tag{3.5}$$

If  $u_j^n x_y \notin \Omega_2$  then (3.5) implies that  $u_j^n y \in \exp B_{\epsilon_0}^W(X \setminus \Omega_2)$ , which is disjoint from  $\mathcal{T}(\epsilon_0, \tilde{K}_1)$  by choice of  $\Omega_2$ . So, in particular,  $u_j^n y \notin \mathcal{S}_j(\epsilon, \tilde{K}_1)$ . If on the other hand  $u_j^n x_y \in \Omega_2$ , then as  $u_j^n x_y \in Lz$  (because  $Lz$  is  $\Lambda$ -invariant), we deduce that  $u_j^n x_y \in K_2$  which in turn implies by (3.5) that  $u_j^n y \in \mathcal{T}(\epsilon_0, K_2)$  and the orbit and  $W$  coordinates of  $u_j^n y$  are given by  $\text{Ad}_{u_j^n}^n(w_y)$  and  $u_j^n x_y$ , respectively. By the lower bound on the  $W$ -coordinate we deduce that  $u_j^n y \notin \mathcal{T}(\epsilon, K_2)$  and, in particular,  $u_j^n y \notin \mathcal{S}_j(\epsilon, \tilde{K}_1)$ .

(3) We have that  $\epsilon_0 \leq \|\text{Ad}_{u_j^{n_y+1}}^{n_y+1}(w_y)\| \leq M\sqrt{p_y(n_y)}$ .

(4) Similarly,  $\epsilon_0 \leq \|\text{Ad}_{u_j^{n_y+1}}^{n_y+1}(w_y)\| \leq M^{n_y+1}\epsilon$ . □

**COROLLARY 3.4.** *The conclusion of Theorem 3.1 remains valid if the assumption  $P_k \cap Lz = \emptyset$  is relaxed to  $|P_k \cap Lz|/|P_k| \rightarrow 0$  as  $k \rightarrow \infty$ .*

*Proof.* We split  $\mu_k = (1 - \alpha_k)\mu_k^1 + \alpha_k\mu_k^2$  with  $\mu_k^i$ ,  $i = 1, 2$ , being the normalized counting measures on  $P_k \setminus Lz$  and  $P_k \cap Lz$ , respectively. In this case  $\alpha_k = |P_k \cap Lz|/|P_k| \rightarrow 0$  as  $k \rightarrow \infty$  by assumption and so the accumulation points of  $\mu_k$  are the same as those of  $\mu_k^1$  for which Theorem 3.1 applies. □

#### 4. $\Lambda$ -invariance and measure classifications

In this section we will show that all the measures appearing in our discussions are invariant under a certain group  $\Lambda$ . We will then classify all the  $\Lambda$ -invariant and ergodic probability measures in certain situations. This will serve us in the proof of Theorem 1.3 along the lines described in § 2.3.

**4.1 Invariance**

We revert to the notation introduced in §§ 1 and 2. In particular, recall that the subgroup  $\Lambda \simeq \text{SL}_m(\mathbb{Z}) \times \text{SL}_n(\mathbb{Z}) < \text{SL}_d(\mathbb{Z})$  is defined by

$$\Lambda \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} : \delta_1 \in \text{SL}_m(\mathbb{Z}), \delta_2 \in \text{SL}_n(\mathbb{Z}) \right\}. \tag{4.1}$$

Further, let  $\Lambda_\Delta < G \times G$  denote the diagonal embedding of  $\Lambda$  in  $G \times G$ .

LEMMA 4.1. *The following periodic orbits in either  $X_d$  or  $X_d \times X_d$  and probability measures supported on them are  $\Lambda$ -invariant or  $\Lambda_\Delta$ -invariant, respectively:*

- (1) the periodic orbit  $Ux_0$  and the measures  $\mathbf{m}_{Ux_0}, \{\boldsymbol{\mu}_k : k \in \mathbb{Z}_{>0}\}$ ;
- (2) the periodic orbit  $Hx_0$  and the measures  $\mathbf{m}_{Hx_0}, \{a(k)_* \boldsymbol{\mu}_k : k \in \mathbb{Z}_{>0}\}$ ;
- (3) the periodic orbit  $Vx_0$  and the measure  $\mathbf{m}_{Vx_0}$ ;
- (4) the periodic orbit  $X_m$  and the measures  $\mathbf{m}_{X_m}, \{\boldsymbol{\nu}_k : k \in \mathbb{Z}_{>0}\}$ ;
- (5) the periodic orbit  $Ux_0 \times Hx_0$  and the measures  $\mathbf{m}_{Ux_0} \times \mathbf{m}_{Hx_0}, \{\tilde{a}(k)_* \tilde{\boldsymbol{\mu}}_k : k \in \mathbb{Z}_{>0}\}$ ;
- (6) the periodic orbit  $Ux_0 \times Vx_0$  and the measure  $\mathbf{m}_{Ux_0} \times \mathbf{m}_{Vx_0}$ .

*Proof.* The  $\Lambda$ -actions on  $Ux_0, Hx_0$  and  $X_m$  are given respectively by

$$\begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} \begin{pmatrix} I & 0 \\ \mathbf{u} & I \end{pmatrix} x_0 = \begin{pmatrix} I & 0 \\ \delta_2 \mathbf{u} \delta_1^{-1} & I \end{pmatrix} x_0, \tag{4.2}$$

$$\begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} \begin{pmatrix} h & \mathbf{v} \\ 0 & I \end{pmatrix} x_0 = \begin{pmatrix} \delta_1 h & \delta_1 \mathbf{v} \delta_2^{-1} \\ 0 & I \end{pmatrix} x_0, \tag{4.3}$$

$$\begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} \begin{pmatrix} h & 0 \\ 0 & I \end{pmatrix} x_0 = \begin{pmatrix} \delta_1 h & 0 \\ 0 & I \end{pmatrix} x_0.$$

The  $\Lambda$ -action on  $Vx_0$  is given by a similar formula. This shows that the corresponding periodic orbits are indeed  $\Lambda$ -invariant. Also, as conjugation by  $\Lambda$  fixes the volume form on the groups giving rise to these periodic orbits, the Haar measures on these periodic orbits are preserved.

The measures  $\tilde{\boldsymbol{\mu}}_k$  are  $\Lambda_\Delta$ -invariant because it follows from (4.2) and Definition 1.2 that  $\tilde{\mathcal{P}}_k$  is  $\Lambda_\Delta$ -invariant. In turn, because the  $\Lambda_\Delta$ -action on  $X_d \times X_d$  commutes with that of  $\tilde{a}(k)$ , we conclude that  $\tilde{a}(k)_* \tilde{\boldsymbol{\mu}}_k$  is  $\Lambda_\Delta$ -invariant. Similarly,  $a(k)_* \boldsymbol{\mu}_k$  is  $\Lambda$ -invariant. The invariance of the measure  $\boldsymbol{\nu}_k$  now follows from that of  $\boldsymbol{\mu}_k$  as the projection  $\pi_3$  in (2.9) intertwines the  $\Lambda$ -actions on  $Hx_0$  and  $X_m$ . □

**4.2 Rationality issues**

We will need the following lemmas in order to establish various rationality statements when classifying measures. These rationality statements are important to us because they imply the countability of the measures we classify in each discussion. This countability is used later along the lines described in § 2.3.

LEMMA 4.2. *Let  $N$  be an integer and let  $\lambda \in \text{SL}_N(\mathbb{Z})$  be matrix acting naturally on the torus  $\mathbb{T}^N = \mathbb{R}^N / \mathbb{Z}^N$ . Assume that all the eigenvalues of  $\lambda$  are not roots of unity. Then, if  $\mathbf{w} \in \mathbb{T}^N$  has a finite  $\lambda$ -orbit then  $\mathbf{w}$  is a rational point (that is, any vector representing it is rational).*

*Proof.* Assume that  $\lambda^j \mathbf{w} = \mathbf{w}$ . If  $w \in \mathbb{R}^N$  projects to  $\mathbf{w}$  then this means that there is an integer vector  $e$  such that  $\lambda^j w = w + e$  or, equivalently,  $(\lambda^j - I)w = e$ . By assumption  $\lambda^j - I$  is invertible and its inverse is a rational matrix, so  $w = (\lambda^j - I)^{-1}e$  is rational as well. □

We recall the definition of the commensurator group. Let  $G'$  be a topological group and  $\Gamma' < G'$  a closed subgroup. Let  $\text{comm}_{G'}(\Gamma') \stackrel{\text{def}}{=} \{g \in G' : g\Gamma'g^{-1}, \Gamma' \text{ are commensurable}\}$ .

LEMMA 4.3. *Let  $G'$  be a topological group and  $\Gamma', \Lambda'$  closed subgroups.*

- (1) *Let  $g \in G'$  and  $q \in \text{comm}_{G'}(\Gamma')$ . Then, if the orbit  $\Lambda g\Gamma' \subset G'/\Gamma'$  is finite, then so is  $\Lambda gq\Gamma'$ .*
- (2) *If  $\Gamma' < G'$  is a lattice then  $q \in \text{comm}_{G'}(\Gamma')$  if and only if the orbit of  $\Gamma'g\Gamma' \subset G'/\Gamma'$  is finite.*

*Proof.* (1) Note that  $\Lambda'g\Gamma' \subset G'/\Gamma'$  is finite if and only if  $\text{Stab}_{\Lambda'} g\Gamma' = \Lambda' \cap g\Gamma'g^{-1} < \Lambda'$  is of finite index. As  $q \in \text{comm}_{G'}(\Gamma')$  we deduce that the intersection  $gq\Gamma'q^{-1}g^{-1} \cap g\Gamma'g^{-1}$  is of finite index in both groups. In particular,  $\Lambda' \cap gq\Gamma'q^{-1}g^{-1} < \Lambda'$  is of finite index. Arguing in reverse this, implies now that  $\Lambda'gq\Gamma' \subset G'/\Gamma'$  is finite.

(2) One direction of implication follows by applying part (1) to  $\Lambda' = \Gamma'$ . In the other direction, if the orbit  $\Gamma'q\Gamma'$  is finite then, as before, this implies that the group  $\text{Stab}_{\Gamma'} q\Gamma' = \Gamma' \cap q\Gamma'q^{-1}$  is of finite index in  $\Gamma'$ . In particular, it is a lattice in  $G'$ , which forces its index in  $q\Gamma'q^{-1}$  to be finite as well; that is,  $q \in \text{comm}_{G'}(\Gamma')$ . □

LEMMA 4.4. *If  $\begin{pmatrix} g & \mathbf{v} \\ 0 & I \end{pmatrix}x_0$  has a finite  $\Lambda$ -orbit then*

$$g \in \text{comm}_{\text{SL}_m(\mathbb{R})}(\text{SL}_m(\mathbb{Z})) \quad \text{and} \quad \mathbf{v} \text{ is rational.}$$

*A similar statement holds for  $\begin{pmatrix} g & 0 \\ \mathbf{u} & I \end{pmatrix}x_0$ .*

*Proof.* We prove the first statement. The second statement follows by applying the involution on  $X_d$  induced by the transpose inverse operation. Assume  $\Lambda \begin{pmatrix} g & \mathbf{v} \\ 0 & I \end{pmatrix}x_0$  is finite. Projecting to  $X_m \cong \text{SL}_m(\mathbb{R})/\text{SL}_m(\mathbb{Z})$ , we deduce by Lemma 4.3(2) that

$$g \in \text{Comm}_{\text{SL}_m(\mathbb{R})}(\text{SL}_m(\mathbb{Z})).$$

In turn, this implies that  $q = \begin{pmatrix} g & 0 \\ 0 & I \end{pmatrix} \in \text{comm}_G(\text{SL}_d(\mathbb{Z}))$ . It now follows from Lemma 4.3(1) that  $\Lambda \begin{pmatrix} g & \mathbf{v} \\ 0 & I \end{pmatrix}q^{-1}x_0 = \Lambda \begin{pmatrix} I & \mathbf{v} \\ 0 & I \end{pmatrix}x_0$  is finite as well. Thus we have reduced the discussion to the situation where  $g = I$ .

We now identify  $Vx_0$  with the torus  $\mathbb{R}^N/\mathbb{Z}^N$  with  $N = m \cdot n$  and apply Lemma 4.2 to conclude that  $\mathbf{v}$  is rational. Indeed it is straightforward to verify the existence of  $\lambda \in \Lambda$  that acts on  $\mathbb{R}^N$  without roots of unity as eigenvalues. □

### 4.3 Measure classifications in general

THEOREM 4.5. *Let  $\mu$  be a  $\Lambda$ -invariant and ergodic probability measure on  $X_d$ . Then there exist an intermediate subgroup  $\Lambda < L < G$  and a periodic  $L$ -orbit  $Lx \subset X_d$ , such that  $\mu$  is the  $L$ -invariant probability measure  $\mathbf{m}_{Lx}$ .*

This theorem is a particular case of a more general measure classification by Shah [Sha98] which uses and generalizes Ratner’s measure classification theorem [Rat91a, Rat91b]. It is applicable since  $\Lambda = \text{SL}_m(\mathbb{Z}) \times \text{SL}_n(\mathbb{Z})$  is generated by unipotent elements. In fact, because  $\Lambda$  is a lattice in a semisimple Lie subgroup of  $G$  with no compact factors, using the suspension technique [Wit94, Corollary 5.8] it is straightforward to deduce Theorem 4.5 directly from Ratner’s measure classification theorem for the actions of semisimple groups without compact factors (for a simplified proof for this case, see [Ein06]).

For a closed subgroup  $L < G$  we denote by  $L^\circ$  the connected component of the identity of  $L$ . We have the following corollary which will be more convenient for us.

**COROLLARY 4.6.** *Let  $\mu$  be a  $\Lambda$ -invariant and ergodic probability measure on  $X_d$  and let  $L, x$  be the group and the point that arise by applying Theorem 4.5 so that  $\mu = \mathbf{m}_{Lx}$ . Then there exist  $x_1, \dots, x_N \in Lx$  such that  $Lx = \bigsqcup_{i=1}^N L^\circ x_i$ , each orbit  $L^\circ x_i$  is periodic,  $\Lambda$  acts transitively by permuting the collection of orbits  $\{L^\circ x_i\}$ , and  $\mu = (1/N) \sum_1^N \mathbf{m}_{L^\circ x_i}$ .*

*Proof.* As  $L^\circ \leq L$  is open and closed, the orbit  $Lx$  decomposes into a union of (relatively) open and closed  $L^\circ$ -orbits. As  $\mu$  is finite, this decomposition is finite,  $Lx = \bigsqcup_{i=1}^N L^\circ x_i$ . As  $L^\circ$  is a normal subgroup of  $L$ ,  $\Lambda$  acts on the  $L^\circ$ -orbits by permuting them. Moreover, the ergodicity assumption implies that this action is transitive. It follows that  $\mu(L^\circ x_i) = N^{-1}$  and the formula  $\mu = (1/N) \sum_1^N \mathbf{m}_{L^\circ x_i}$  follows.  $\square$

For convenience of reference we also state the following elementary lemma whose proof we omit.

**LEMMA 4.7.** *Let  $F_1, F_2$  be closed subgroups of  $G$  and  $x \in X_d$ . If  $F_1x \subset F_2x$  then  $F_1^\circ \subset F_2$ .*

The rest of this section is devoted to classifying  $\Lambda$ -invariant and ergodic measures in various situations that will be encountered in the course of the proof of Theorem 1.3.

#### 4.4 Measures supported in $Ux_0$

Let us define

$$\text{Tor}_k(Ux_0) \stackrel{\text{def}}{=} \{x_{k^{-1}\mathbf{u}} \in Ux_0 : \mathbf{u} \in \text{Mat}_{n \times m}(\mathbb{Z})\}. \tag{4.4}$$

**LEMMA 4.8.** *The ergodic  $\Lambda$ -invariant probability measures on  $Ux_0$  are exactly the normalized counting measures on finite  $\Lambda$ -orbits and  $\mathbf{m}_{Ux_0}$ . Moreover, any finite  $\Lambda$ -orbit is contained in  $\text{Tor}_k(Ux_0)$  for some positive integer  $k$ .*

*Proof.* Let<sup>6</sup>  $\mu$  be an ergodic  $\Lambda$ -invariant measure supported in the orbit  $Ux_0$ . Applying Theorem 4.5, we conclude the existence of a closed subgroup  $\Lambda < L < G$  such that  $\mu$  is the  $L$ -invariant probability measure supported on a periodic  $L$ -orbit. Applying Lemma 4.7 we conclude that  $L^\circ < U$ . Viewing  $L^\circ$  as a subspace of  $\text{Mat}_{n \times m}(\mathbb{R}) \cong U$ , the fact that  $\Lambda$  normalizes  $L^\circ$  translates into this subspace being invariant under the linear representation of  $\Lambda$  on  $\text{Mat}_{n \times m}(\mathbb{R})$  (here  $\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \in \text{SL}_m(\mathbb{R}) \times \text{SL}_n(\mathbb{R})$  acts on  $\mathbf{u} \in \text{Mat}_{n \times m}(\mathbb{R})$  by  $D\mathbf{u}A^{-1}$ ). As this representation of  $\text{SL}_m(\mathbb{R}) \times \text{SL}_n(\mathbb{R})$  is irreducible and  $\Lambda$  is Zariski dense in the former group, we deduce that  $L^\circ$  is either the trivial group or  $U$ .

If  $L^\circ = U$  then clearly  $\mu = \mathbf{m}_{Ux_0}$ . If  $L^\circ$  is trivial, an application of Corollary 4.6 gives that  $\mu$  is the normalized counting measure on a finite  $\Lambda$ -orbit  $\Lambda x_{\mathbf{u}} \subset Ux_0$ . Lemma 4.4 implies now that  $\mathbf{u}$  is in fact a rational matrix. If  $k$  is a common denominator for its entries then  $x_{\mathbf{u}} \in \text{Tor}_k(Ux_0)$  as desired.  $\square$

#### 4.5 Measures supported in $Hx_0$

Our next objective is to classify the  $\Lambda$ -invariant ergodic probability measures supported in  $Hx_0$ . By Theorem 4.5 and Corollary 4.6 such a measure is always of the form  $(1/N) \sum \mathbf{m}_{Lx_i}$ , where  $L < G$  is a closed connected subgroup normalized by  $\Lambda$ . As  $Lx_i \subset Hx_0$  and  $L$  is connected, we conclude from Lemma 4.7 that  $L < H$ . We have the following classification of groups.

---

<sup>6</sup> Lemma 4.8 could be proved using Fourier arguments, but we choose to appeal to Theorem 4.5 as this is more compatible with the later arguments. Furthermore, this is by no means a new result and the proof is included for expository reasons. For a similar argument, see, for example, [Wit94, Example 5.9].

LEMMA 4.9. *Let  $L < H$  be a closed connected subgroup normalized by  $\Lambda$ . Then there are four possibilities:*

- (1)  $L = \{e\}$ ;
- (2)  $L = V = \left\{ \begin{pmatrix} I & \mathbf{v} \\ 0 & I \end{pmatrix} : \mathbf{v} \in \text{Mat}_{m \times n}(\mathbb{R}) \right\}$ ;
- (3)  $L = \text{SL}_m(\mathbb{R}) = \left\{ \begin{pmatrix} h & 0 \\ 0 & I \end{pmatrix} : h \in \text{SL}_m(\mathbb{R}) \right\}$ ;
- (4)  $L = H$ .

*Proof.* Consider the projection of  $L$  in the simple group  $\text{SL}_m(\mathbb{R}) = H/V$ . As  $L$  is normalized by  $\Lambda$ , this projection is a connected subgroup normalized by  $\text{SL}_m(\mathbb{Z})$ , and by Zariski density we deduce that it is a connected normal subgroup. It therefore follows that this projection is either trivial or  $\text{SL}_m(\mathbb{R})$ . In the first case,  $L < V$ . As the adjoint action of  $\Lambda$  on the Lie algebra of  $V$  (which is isomorphic to  $V$  itself) is irreducible, it follows that either  $L = V$  or  $L$  is the trivial group.

Assume then that the projection of  $L$  is onto and consider the subgroup  $V' = L \cap V$  which is normalized by  $\Lambda$ . The same irreducibility argument as before implies that either  $V' = V$  or  $V'$  is trivial. In the first case  $L = H$ . In the second case the projection of  $L$  onto  $H/V$  must also be injective so there exists a map  $\psi : \text{SL}_m(\mathbb{R}) \rightarrow \text{Mat}_{m \times n}(\mathbb{R})$  such that  $L = \left\{ \begin{pmatrix} g & \psi(g) \\ 0 & I \end{pmatrix} : g \in \text{SL}_m(\mathbb{R}) \right\}$ . If  $\lambda = \begin{pmatrix} \delta & 0 \\ 0 & I \end{pmatrix} \in \Lambda$  then  $\lambda \begin{pmatrix} \delta & \psi(\delta) \\ 0 & I \end{pmatrix} \lambda^{-1} = \begin{pmatrix} \delta & \delta \psi(\delta) \\ 0 & I \end{pmatrix} \in L$  so that  $\delta \psi(\delta) = \psi(\delta)$ . By Zariski density we deduce that this formula holds for any  $\delta \in \text{SL}_m(\mathbb{R})$ . In turn, if 1 is not an eigenvalue of  $\delta$ , this implies that the columns of  $\psi(\delta)$  must be zero. We conclude by continuity that  $\psi$  vanishes and so  $L$  is the standard copy of  $\text{SL}_m(\mathbb{R})$  in  $H$ . □

In light of Lemma 4.9 (and for the sake of the current discussion), the following definition makes sense.

DEFINITION 4.10. We say that a  $\Lambda$ -invariant and ergodic measure on  $Hx_0$  is of *type* (1)–(4) depending on which one of the four groups that appear in Lemma 4.9 is attached to it.

COROLLARY 4.11. *The following is a classification of the  $\Lambda$ -invariant and ergodic probability measures on  $Hx_0$ .*

(1) *Measures of type (1) are simply normalized counting measures on a (necessarily finite)  $\Lambda$ -orbit of a point of the form  $\begin{pmatrix} h & \mathbf{v} \\ 0 & I \end{pmatrix} x_0$ , in which  $h \in \text{comm}_{\text{SL}_m(\mathbb{R})}(\text{SL}_m(\mathbb{Z}))$  and  $\mathbf{v}$  is rational. If such a measure projects to  $\delta_{x_0}$  under  $(\pi_3)_*$  then we can choose  $h = I$ .*

(2) *Measures of type (2) are of the form  $(1/N) \sum_{i=1}^N \mathbf{m}_{Vx_i}$ , where  $\{x_i\}$  is the  $\Lambda$ -orbit of  $x_1 = \begin{pmatrix} h & 0 \\ 0 & I \end{pmatrix} x_0$ , in which  $h \in \text{comm}_{\text{SL}_m(\mathbb{R})}(\text{SL}_m(\mathbb{Z}))$ . There is only one such measure that projects to  $\delta_{x_0}$ , and that is  $\mathbf{m}_{Vx_0}$ .*

(3) *Measures of type (3) are of the form  $(1/N) \sum_{i=1}^N \mathbf{m}_{\text{SL}_m(\mathbb{R})x_i}$ , where  $x_i = \begin{pmatrix} I & \mathbf{v}_i \\ 0 & I \end{pmatrix} x_0$  and the  $\mathbf{v}_i$  are rational. All such measures project under  $(\pi_3)_*$  to  $\mathbf{m}_{X_m}$ .*

(4) *There is only one measure of type (4), namely  $\mathbf{m}_{Hx_0}$ , and it projects under  $(\pi_3)_*$  to  $\mathbf{m}_{X_m}$ .*

*Proof.* Let  $\mu$  denote a  $\Lambda$ -invariant and ergodic probability measure on  $Hx_0$  and present it as  $\mu = (1/N) \sum_{i=1}^N \mathbf{m}_{Lx_i}$  with  $L$  connected and determining the type as explained above. We prove statements (1)–(3), as (4) is clear.

(1) If  $\mu$  is of type (1) we deduce from the ergodicity that  $\{x_i\}_1^N$  forms a finite  $\Lambda$ -orbit and Lemma 4.4 gives the commensurability and rationality statements. The statement regarding the projection is clear.

(2) Suppose  $\mu$  is of type (2) and write  $x_i = \begin{pmatrix} h_i & \mathbf{v}_i \\ 0 & I \end{pmatrix} x_0$ . Using the  $V$ -invariance we may assume without loss of generality that  $\mathbf{v}_i = 0$  and so  $\pi_3(x_i) = x_i$ . Then  $(\pi_3)_*\mu = (1/N) \sum_1^N \delta_{x_i}$  is a finitely supported  $\Lambda$ -invariant and ergodic measure and so the  $x_i$  form an orbit. Now, by Lemma 4.4,  $h_1 \in \text{comm}_{\text{SL}_m(\mathbb{R})}(\text{SL}_m(\mathbb{Z}))$  as required.

(3) Assume  $\mu$  is of type (3). As the periodic orbits  $\text{SL}_m(\mathbb{R})x_i$  are closed and the  $\text{SL}_m(\mathbb{R})$ -orbits are transverse to the  $V$ -orbits (all of which are periodic in  $Hx_0$ ), we deduce that the intersection  $\text{supp}(\mu) \cap Vx_0$ , which is  $\Lambda$ -invariant and closed, is finite. Lemma 4.4 now implies that the points  $x_i$ , which without loss of generality might be assumed to belong to  $\text{supp}(\mu) \cap Vx_0$ , have rational representatives. □

**4.6 Measures supported in  $Ux_0 \times Hx_0$**

**THEOREM 4.12.** *The only  $\Lambda_\Delta$ -invariant probability measure on  $Ux_0 \times Hx_0$  which projects under  $\pi_1$  and  $\pi_2$  to  $\mathbf{m}_{Ux_0}$ ,  $\mathbf{m}_{Hx_0}$ , respectively, is the product measure  $\mathbf{m}_{Ux_0} \times \mathbf{m}_{Hx_0}$ .*

*Proof.* Let  $u \in \Lambda$  be a unipotent element that belongs to the subgroup  $\text{SL}_m(\mathbb{Z}) < \Lambda$  (such an element always exists as  $m \geq 2$ ). Note the following two facts.

(1) The action of  $u$  on  $(Hx_0, \mathbf{m}_{Hx_0})$  is mixing by the Howe–Moore theorem applied to the action of  $\text{SL}_m(\mathbb{R})$ .

(2) The action of  $u$  on the torus  $\mathbb{T}^{n \times m} \cong Ux_0$  is by a unipotent automorphism and so the ergodic components of  $\mathbf{m}_{Ux_0}$  with respect to the action of  $u$  are minimal rotation on compact abelian groups.

The theorem now follows from the following two lemmas. □

**LEMMA 4.13** (cf. [Gla03, Theorem 6.27]). *Let  $(X, \mu, u), (Y, \nu, u)$  be two dynamical systems. If  $(X, \mu, u)$  is a minimal rotation on a compact abelian group and  $(Y, \nu, u)$  is mixing, then the two systems are disjoint.*

**LEMMA 4.14.** *Let  $(X, \mu, u), (Y, \nu, u)$  be two dynamical systems such that  $\nu$  is ergodic and let  $\mu = \int \mu_x d\mu(x)$  be the ergodic decomposition of  $\mu$ . Then, if for  $\mu$ -almost every  $x$  the systems  $(X, \mu_x, u), (Y, \nu, u)$  are disjoint, then  $(X, \mu, u), (Y, \nu, u)$  are disjoint as well.*

We give both proofs for the sake of completeness.

*Proof of Lemma 4.13.* Let  $\eta$  be a joining of  $\mu, \nu$ . As  $\eta$  is a joining, the projections  $p_X : (X \times Y, \eta) \rightarrow (X, \mu), p_Y : (X \times Y, \eta) \rightarrow (Y, \nu)$  induce injections of the  $L^2$ -spaces; that is, the Hilbert spaces  $\mathcal{H}_X \stackrel{\text{def}}{=} L^2(X, \mu), \mathcal{H}_Y \stackrel{\text{def}}{=} L^2(Y, \nu)$  are isometrically embedded in  $\mathcal{H} \stackrel{\text{def}}{=} L^2(X \times Y, \eta)$ . We denote by  $\mathcal{H}^0, \mathcal{H}_X^0, \mathcal{H}_Y^0$  the subspaces orthogonal to the constant functions in each of the spaces. Showing that  $\eta$  is the product measure is the same as showing that the two subspaces  $\mathcal{H}_X^0, \mathcal{H}_Y^0$  of  $\mathcal{H}^0$  are orthogonal. For this, note that the mixing assumption for the  $u$ -action on  $(Y, \nu)$  is equivalent to saying that for each  $v \in \mathcal{H}_Y^0$  the sequence  $T_u^n v$  converges weakly to 0, where we write  $T_u$  for the unitary operator on  $\mathcal{H}$  induced by the  $u$ -action. If the two subspaces are not orthogonal, then there exists a vector  $0 \neq v \in \mathcal{H}_Y^0$  that has a non-trivial projection to  $\mathcal{H}_X^0$ . Let us denote this projection of  $v$  by  $w \in \mathcal{H}_X^0$ . We claim that  $T_u^n w$  must converge weakly to 0 as well. This simply follows from the fact that the projection commutes with  $T_u$  as the subspace we project on is  $T_u$ -invariant and  $T_u$  is unitary. However, as  $(X, \mu, u)$  is isomorphic to a minimal rotation on a compact group there exists an integer sequence  $n_k \rightarrow \infty$  such that  $T_u^{n_k} w \rightarrow w$  as  $k \rightarrow \infty$ . This contradiction implies the lemma. □



*Proof of Lemma 4.14.* Let  $\eta$  be a joining of  $\mu, \nu$  and let

$$\eta = \int \eta_{(x,y)} d\eta(x,y)$$

be the ergodic decomposition of  $\eta$ . Let  $p_X, p_Y$  denote the projections from  $X \times Y$  to  $X, Y$  respectively. For  $\eta$ -almost every  $(x, y)$  we have that  $(p_Y)_*\eta_{(x,y)}$  is an ergodic  $u$ -invariant probability measure on  $Y$  and as  $\eta$  projects to  $\nu$  we deduce that  $\nu = \int (p_Y)_*\eta_{(x,y)} d\eta(x,y)$ . Since  $\nu$  is ergodic we conclude that  $\nu = (p_Y)_*\eta_{(x,y)}$  for  $\eta$ -almost any  $(x, y)$ . By a similar reasoning one can argue that for  $\eta$ -almost every  $(x, y)$ ,  $(p_X)_*\eta_{(x,y)} = \mu_x$ . It follows that for  $\eta$ -almost any  $(x, y)$  the ergodic component  $\eta_{(x,y)}$  is a joining of  $\mu_x$  and  $\nu$  and therefore by our disjointness assumption we have  $\eta_{(x,y)} = \mu_x \times \nu$  for  $\eta$ -almost any pair  $(x, y)$ . This implies the lemma.  $\square$

### 4.7 Measures supported in $Ux_0 \times Vx_0$

We will use the following notation in the special case where  $m = n = 2$ . Let us denote by  $\mathbf{u} \mapsto \mathbf{u}^\#$  the linear isomorphism of  $\text{Mat}_2(\mathbb{R})$  given by  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^\# = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$ . For coprime integers  $p, q$  we denote

$$L_{p,q} \stackrel{\text{def}}{=} \left\{ \left( \begin{pmatrix} I & 0 \\ p\mathbf{u} & I \end{pmatrix}, \begin{pmatrix} I & q\mathbf{u}^\# \\ 0 & I \end{pmatrix} \right) : \mathbf{u} \in \text{Mat}_2(\mathbb{R}) \right\} < U \times V. \tag{4.5}$$

**THEOREM 4.15.** *Assume  $n = m$ . Let  $\eta$  be a  $\Lambda_\Delta$ -invariant and ergodic probability measure supported on  $Ux_0 \times Vx_0$  and suppose  $(\pi_1)_*\eta = \mathbf{m}_{Ux_0}$ ,  $(\pi_2)_*\eta = \mathbf{m}_{Vx_0}$ . Then  $\eta = \mathbf{m}_{Ux_0} \times \mathbf{m}_{Vx_0}$ , or  $n = 2$  and  $\eta = (1/N) \sum_1^N \mathbf{m}_{L_{p,q}(x_i, y_i)}$ , where  $p, q$  are coprime integers and the points  $x_i, y_i$  may be chosen to have rational matrix representatives.*

*Proof.* In an abuse of notation we write  $\Lambda$  for  $\Lambda_\Delta$ . We identify  $Ux_0 \times Vx_0$  with the product torus  $\mathbb{T}^{n^2} \times \mathbb{T}^{n^2}$  and recall that the  $\Lambda = \text{SL}_n(\mathbb{Z}) \times \text{SL}_n(\mathbb{Z})$ -action on it is induced by the  $\Lambda$ -representation on  $\text{Mat}_n(\mathbb{R}) \times \text{Mat}_n(\mathbb{R})$  given by (see (4.2), (4.3)),

$$(\delta_1, \delta_2)(\mathbf{u}, \mathbf{v}) = (\delta_2 \mathbf{u} \delta_1^{-1}, \delta_1 \mathbf{v} \delta_2^{-1}). \tag{4.6}$$

As  $\Lambda$  is generated by unipotents, a suitable application of Theorem 4.5 implies that there exist a  $\Lambda$ -invariant subspace  $W \subset \text{Mat}_n(\mathbb{R}) \times \text{Mat}_n(\mathbb{R})$  and finitely many periodic orbits  $\{W(\mathbf{u}_i, \mathbf{v}_i)\}_1^N \subset \mathbb{T}^{n^2} \times \mathbb{T}^{n^2}$  such that  $\eta = (1/N) \sum_1^N \mathbf{m}_{W(\mathbf{u}_i, \mathbf{v}_i)}$ . From our assumption that  $\eta$  is a joining of  $\mathbf{m}_{Ux_0}$  and  $\mathbf{m}_{Vx_0}$  we conclude that  $W$  projects onto  $\text{Mat}_n(\mathbb{R})$  under both left and right projections.

We will need the following general representation-theoretic lemma. Its proof is straightforward and left to the reader.

**LEMMA 4.16.** *Let  $\rho_i$ , for  $i = 1, 2$ , be two irreducible representations of a group  $\Lambda$  on the vector spaces  $V_i$  and let  $\rho = \rho_1 \oplus \rho_2$ . If  $\{0\} \subsetneq W \subsetneq V_1 \oplus V_2$  is a  $\Lambda$ -invariant subspace, then either  $W = V_1 \times \{0\}$ , or  $W = \{0\} \times V_2$ , or there is an isomorphism  $\varphi : (V_1, \rho_1) \rightarrow (V_2, \rho_2)$  of  $\Lambda$ -representations such that  $W = \{(v, \varphi(v)) : v \in V_1\}$ .*

We apply this lemma with  $V_1 = V_2 = \text{Mat}_n(\mathbb{R})$ , with the representations  $\rho_i$  of  $\Lambda$  which are given by restricting to left and right coordinates of the formula (4.6), and with  $W$  being the subspace identified above which is attached to  $\eta$ . If  $W$  is the whole space then  $\eta = \mathbf{m}_{Ux_0} \times \mathbf{m}_{Vx_0}$ . Otherwise, we deduce from the lemma (and the fact that  $W$  projects onto each factor) the existence of the isomorphism  $\varphi$ . Then, applying (4.6) to the diagonal copy of  $\text{SL}_n(\mathbb{Z})$  in  $\Lambda$ , we deduce that the image of the line  $\{sI : s \in \mathbb{R}\} \subset V_1$  must be stable under conjugation by every element of  $\text{SL}_n(\mathbb{Z})$ , which implies that this line must be mapped to itself. Therefore, there exists

a scalar  $\rho_\varphi$  such that  $\varphi(I) = \rho_\varphi I$ . It then follows from (4.6) that the restriction of  $\varphi$  to the set  $\{s\delta : s \in \mathbb{R}, \delta \in \text{SL}_n(\mathbb{Z})\}$  is given by the formula

$$\varphi(s\delta) = \rho_\varphi s\delta^{-1}. \tag{4.7}$$

For  $n > 2$  this is not a linear map and so the existence of  $\varphi$  is ruled out, and so indeed  $\eta = \mathbf{m}_{Ux_0} \times \mathbf{m}_{Vx_0}$  as claimed.

In the case  $n = 2$ , on the other hand, formula (4.7) is given by

$$\mathbf{u} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow{\varphi} \rho_\varphi \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \rho_\varphi \mathbf{u}^\#, \tag{4.8}$$

which is linear, and by Zariski density this must be the formula for  $\varphi$  on  $V_1 = \text{Mat}_2(\mathbb{R})$ . By the above lemma we have that

$$W = \{(\mathbf{u}, \rho_\varphi \mathbf{u}^\#) : \mathbf{u} \in \text{Mat}_n(\mathbb{R})\},$$

and since it has periodic orbits in  $\mathbb{T}^{n^2} \times \mathbb{T}^{n^2}$ ,  $\rho_\varphi$  must be rational, say  $\rho_\varphi = p/q$  for some coprime integers  $p, q$ . Under our identifications  $W$  corresponds to the subgroup  $L_{p,q}$  from (4.5).

Finally, we need to justify the rationality of the points  $(\mathbf{u}_i, \mathbf{v}_i)$ . Consider the action of  $\Lambda$  on the quotient torus  $\mathbb{T}^{n^2} \times \mathbb{T}^{n^2}/W$  which we identify with the standard torus  $\mathbb{T}^N$  for a suitable  $N$ . The measure  $\eta$  projects there to a finitely supported  $\Lambda$ -invariant measure. It is straightforward to show the existence of  $\lambda \in \Lambda$  which acts on  $\mathbb{R}^N \cong \text{Mat}_n \times \text{Mat}_n/W$  without roots of unity as eigenvalues, and therefore by Lemma 4.2 we conclude the rationality of the images of  $(\mathbf{u}_i, \mathbf{v}_i)$  in  $\mathbb{T}^N$  and in turn the desired rationality before projecting (because  $W$  is a rational space).  $\square$

### 5. Hecke friends

For any positive integer  $\ell$  and an  $m$ -dimensional lattice  $x \in X_m$  we define the set of  $\ell$ -Hecke friends of  $x$  to be the collection of lattices

$$\{\ell^{-1/m}\Delta : \Delta \text{ is a subgroup of index } \ell \text{ of } x\} \subset X_m.$$

Clearly, it is enough to understand the collection of  $\ell$ -Hecke friends of  $\mathbb{Z}^m$ , for if  $x = h\mathbb{Z}^m$  for some  $h \in \text{SL}_m(\mathbb{R})$ , then the collection of  $\ell$ -Hecke friends of  $x$  is simply the image under  $h$  of the corresponding collection for  $\mathbb{Z}^m$ .

Given a subgroup  $\Delta < \mathbb{Z}^m$  of index  $\ell$ , we choose  $\mathbf{u} \in \text{Mat}_m(\mathbb{Z})$  whose columns form a basis for  $\Delta$  and with  $\det \mathbf{u} > 0$ . By Lemma 2.3 (and its proof) there is a unique positive elementary divisors tuple  $(\ell_1, \dots, \ell_m)$  attached to it. Consequently,  $\ell = \prod_1^m \ell_i$ . It is straightforward to show that the tuple does not depend on the choice of  $\mathbf{u}$  and so we make the following definition.

DEFINITION 5.1. Given an  $\ell$ -Hecke friend of  $\mathbb{Z}^m$ , we define its Hecke type to be the corresponding  $m$ -tuple of elementary divisors.

The above discussion could be restated as follows.

LEMMA 5.2. The collection of  $\ell$ -Hecke friends of  $\mathbb{Z}^m$  is partitioned into  $\text{SL}_m(\mathbb{Z})$ -orbits in the following way

$$\{\ell\text{-Hecke friends of } \mathbb{Z}^m\} = \bigsqcup_{(\ell_1, \dots, \ell_m)} \text{SL}_m(\mathbb{Z}) \text{diag}(\ell_1, \dots, \ell_m)\mathbb{Z}^m,$$

where the union is taken over the  $\ell$ -Hecke types; that is, all positive tuples  $(\ell_1, \dots, \ell_m)$  such that  $\ell_i | \ell_{i+1}$  and  $\prod_1^m \ell_i = \ell$ .

The following equidistribution result will be needed in the proof of Theorem 1.3.

**THEOREM 5.3.** *Let  $\{(\ell_{i1}, \dots, \ell_{im})\}_{i=1}^\infty$  be a sequence of types of  $\ell_i$ -Hecke friends of  $\mathbb{Z}^m$ , where  $\ell_i \stackrel{\text{def}}{=} \prod_{j=1}^m \ell_{ij}$ , and let  $\nu_i$  denote the normalized counting measure supported on the collection of  $\ell_i$ -Hecke friends of  $\mathbb{Z}^m$  of type  $(\ell_{i1}, \dots, \ell_{im})$ . Then  $\nu_i \xrightarrow{w^*} \mathbf{m}_{X_m}$  if and only if  $\ell_{im}/\ell_{i1} \rightarrow \infty$ .*

*Proof.* This theorem follows from [COU01]. More conveniently, it is a special case of [EO06, Theorem 1.2]. This result states that for  $q_i \in \text{GL}_m(\mathbb{Q})$ , the normalized counting measure on the (finite) orbit  $1/(\det q_i)^{1/m} \text{SL}_m(\mathbb{Z})q_i\mathbb{Z}^m$  equidistributes to  $\mathbf{m}_{X_m}$  once  $\deg q_i \rightarrow \infty$ , where  $\deg q_i$  is the size of the corresponding orbit. We apply this for  $q_i = \text{diag}(\ell_{i1}, \dots, \ell_{im})$ . We need to explain why the condition  $\deg(\text{diag}(\ell_{i1}, \dots, \ell_{im})) \rightarrow \infty$  is equivalent to the requirement  $\ell_{im}/\ell_{i1} \rightarrow \infty$ .

Without loss of generality we may assume on scaling by  $1/\ell_{i1}$  that  $\ell_{i1} = 1$ . If  $\ell_{im}$  is bounded along some subsequence, then clearly the degrees do not diverge to  $\infty$ . On the other hand, if  $(1, \dots, \ell_m)$  is a Hecke type and  $\ell_m > M$ , then there exists some  $1 \leq r < m$  such that  $\ell_{r+1}/\ell_r > M^{1/m}$ . Then the orbit  $\text{SL}_m(\mathbb{Z}) \text{diag}(1, \dots, \ell_m)\mathbb{Z}^m$  contains an embedded copy of the orbit  $\text{SL}_2(\mathbb{Z}) \begin{pmatrix} 1 & 0 \\ 0 & \ell_{r+1}/\ell_r \end{pmatrix} \mathbb{Z}^2$  which contains at least  $p^{e-1}(p+1)$  points where  $p^e$  is the largest prime power dividing  $\ell_{r+1}/\ell_r$ . As this prime power goes to  $\infty$  with  $M$ , we are done.  $\square$

### 6. Proof of Theorem 1.3

We follow the scheme presented after (2.9), dividing the proof into four steps.

#### 6.1 Step 1

In this step we establish the convergence  $\mu_k \xrightarrow{w^*} \mathbf{m}_{Ux_0}$ . Let  $\sigma$  be a weak\* limit of the sequence  $\{\mu_k\}$ . By Lemma 4.1(1),  $\sigma$  is a  $\Lambda$ -invariant probability measure. By Lemma 4.8 there are only countably many  $\Lambda$ -invariant ergodic probability measures on  $Ux_0$ . We let  $\sigma_0 \stackrel{\text{def}}{=} \mathbf{m}_{Ux_0}$  and let  $\{\sigma_i\}_{i=1}^\infty$  be any enumeration of the measures supported on finite  $\Lambda$ -orbits. By the ergodic decomposition we may write  $\sigma = \sum_{i=0}^\infty c_i \sigma_i$ , where  $c_i \geq 0$  and  $\sum_i c_i = 1$ . The proof of Step 1 will be concluded once we show that  $c_0 = 1$ . By Lemma 4.8, each  $\sigma_i$ ,  $i > 0$ , is supported in the torsion points  $\text{Tor}_{k_i}(Ux_0)$  for some integer  $k_i$ . It is straightforward to show that for  $k > k_i$  we have that  $\mathcal{P}_k \cap \text{Tor}_{k_i}(Ux_0) = \emptyset$ , and so for all large enough  $k$  the support of  $\mu_k$  is disjoint from the orbit on which  $\sigma_i$  is supported. We now apply the non-accumulation theorem (Theorem 3.1) and deduce that  $c_i = 0$  for  $i > 0$  as desired. Theorem 3.1 is applied with the following choices:  $G = \Lambda \ltimes U$ ,  $\Gamma = \Lambda \ltimes \text{Mat}_{n \times m}(\mathbb{Z})$ ,  $L = \Lambda$ ,  $z$  being one of the points in the support of  $\sigma_i$ . The space  $W$ , being a linear complement of  $\text{Lie}(L) = \{0\}$ , equals  $\text{Lie}(U) = \text{Mat}_{n \times m}(\mathbb{R})$  and indeed there are no  $\Lambda$ -fixed vectors in  $W$  other than zero.

#### 6.2 Step 2

In this step we establish the convergence  $\nu_k \xrightarrow{w^*} (\pi_3)_* \theta$ . This follows from the following lemma.

**LEMMA 6.1.** *The measure  $\nu_k \stackrel{\text{def}}{=} (\pi_3)_* a(k)_* \mu_k$  is the normalized counting measure on the collection of  $k^n$ -Hecke friends of  $\mathbb{Z}^m$  of type  $(\underbrace{1 \dots 1}_{m-n} \underbrace{k \dots k}_n)$ . In particular, if  $m > n$ ,  $\nu_k$  equidistributes to  $\mathbf{m}_{X_m}$  and if  $m = n$  then  $\nu_k$  is the Dirac measure supported on  $x_0$ .*

*Proof.* The last sentence in the statement simply follows from Theorem 5.3 (and the observation that  $\mathbb{Z}^m$  is the only  $k^n$ -Hecke friend of type  $(k \dots k)$  of itself). We are thus left to verify the first sentence of the statement.

The set  $\mathcal{P}_k$  decomposes into  $\Lambda$ -orbits and each such orbit is mapped by  $\pi_3 \circ a_k$  to a single  $\Lambda$ -orbit in  $X_m$ . Let  $\Omega \subset \mathcal{P}_k$  be such an orbit. By Lemma 2.3 there exists  $x_{k^{-1}\mathbf{u}} \in \Omega$  with  $\mathbf{u}$  of the form (2.10). By Lemma 5.2, if we show that  $\pi_3(a_k x_{k^{-1}\mathbf{u}})$  is a Hecke friend of  $\mathbb{Z}^m$  of the type prescribed in the statement, it will follow that the normalized counting measure on  $\Omega$  is pushed by  $(\pi_3)_*(a_k)_*$  to the normalized counting measure on the set of Hecke friends on  $\mathbb{Z}^m$  of this type.<sup>7</sup> This implies the same statement regarding  $\nu_k$ .

To this end, let  $x_{k^{-1}\mathbf{u}} \in \mathcal{P}_k$  with  $\mathbf{u}$   $k$ -primitive of the form (2.10). We now calculate  $\pi_3(a(k)x_{k^{-1}\mathbf{u}})$ . Let  $\gamma = \begin{pmatrix} A_\gamma & B_\gamma \\ C_\gamma & D_\gamma \end{pmatrix}$  be a matrix solving (2.12) so that

$$\pi_3(a(k)x_{k^{-1}\mathbf{u}}) = k^{-n/m} A_\gamma \mathbb{Z}^m. \tag{6.1}$$

It follows from the proofs of Lemmas 2.1 and 2.4 that a possible choice for  $\gamma$  is the inverse of the matrix  $\delta$  appearing in (2.11). A short computation shows that (using the formula for the inverse of a  $2 \times 2$  matrix) this inverse is

$$\delta^{-1} = \begin{pmatrix} \text{diag}(k \dots k) & 0 & -\text{diag}(f_1 \dots f_n) \\ 0 & I_{n-m} & 0 \\ -\text{diag}(\ell_1 \dots \ell_n) & 0 & \text{diag}(e_1 \dots e_n) \end{pmatrix}. \tag{6.2}$$

In other words,  $\pi_3(a(k)x_{k^{-1}\mathbf{u}})$  is the  $m$ -dimensional unimodular lattice

$$k^{-n/m} \text{diag}(\underbrace{k, \dots, k}_n, \underbrace{1, \dots, 1}_{m-n}) \mathbb{Z}^m$$

which is a Hecke friend of  $\mathbb{Z}^m$  of the desired type. □

### 6.3 Step 3

In this step we establish the convergence  $a(k)_* \mu_k \xrightarrow{w^*} \theta$ . Before turning to the proof, we need notation and a lemma. The orbit  $Hx_0$  breaks into periodic  $V$ -orbits which are the  $\pi_3$ -fibers. In order to state the next lemma we need to define the notions of  $k$ -torsion and  $k$ -primitive points in such a fiber. For  $x \in Hx_0$ , we write  $x = \begin{pmatrix} g & \mathbf{v} \\ 0 & I \end{pmatrix} x_0$  and set

$$\text{Tor}_k(Vx) \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} g & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & k^{-1}\mathbf{v} \\ 0 & I \end{pmatrix} x_0 \in Vx : \mathbf{v} \in \text{Mat}_{m \times n}(\mathbb{Z}) \right\}, \tag{6.3}$$

$$\text{Tor}_k^{\text{prim}}(Vx) \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} g & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & k^{-1}\mathbf{v} \\ 0 & I \end{pmatrix} x_0 \in Vx : \mathbf{v}^t \text{ is } k\text{-primitive} \right\}. \tag{6.4}$$

Note that although  $g$  is not well defined, the coset  $g \text{SL}_m(\mathbb{Z})$  is well defined and therefore  $\text{Tor}_k(Vx)$  and  $\text{Tor}_k^{\text{prim}}(Vx)$  are well defined as well.

LEMMA 6.2. For any positive integer  $k$ ,

$$a_k \mathcal{P}_k \subset \bigcup_{x \in X_m} \text{Tor}_k^{\text{prim}}(\pi_3^{-1}(x)). \tag{6.5}$$

*Proof.* We first observe that the sets on both sides of (6.5) are  $\Lambda$ -invariant. Regarding  $a_k \mathcal{P}_k$ , this follows from Lemma 4.1(2). Regarding the set on the right, given  $\lambda = \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} \in \Lambda$  and a point  $\begin{pmatrix} g & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & k^{-1}\mathbf{v} \\ 0 & I \end{pmatrix} x_0 \in \text{Tor}_k^{\text{prim}}(\pi_3^{-1}(x))$ , where  $x = \begin{pmatrix} g & 0 \\ 0 & I \end{pmatrix} x_0$  and  $\mathbf{v}^t$   $k$ -primitive, then

$$\lambda \begin{pmatrix} g & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & k^{-1}\mathbf{v} \\ 0 & I \end{pmatrix} x_0 = \begin{pmatrix} \delta_1 g \delta_1^{-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & k^{-1} \delta_1 \mathbf{v} \delta_2^{-1} \\ 0 & I \end{pmatrix} x_0,$$

which belongs to  $\text{Tor}_k^{\text{prim}}(\pi_3^{-1}(\lambda x))$  because  $(\delta_1 \mathbf{v} \delta_2^{-1})^t$  is  $k$ -primitive.

---

<sup>7</sup> Note that there follows from this a stronger statement than the one sought in Step 2 which is the reason for the strengthening of Theorem 1.3 discussed in § 7.

It now follows from Lemma 2.3 that it is enough to show that  $a_k x_{k^{-1}\mathbf{u}}$  is in the right-hand side of (6.5) for  $\mathbf{u}$  as in (2.10). For such a choice of  $\mathbf{u}$  we have seen in the course of the proof of Lemma 6.1 that the matrix  $\gamma$  solving (2.12) may be chosen to be as in (6.2). Thus (2.12) now becomes

$$a_k x_{k^{-1}\mathbf{u}} = \begin{pmatrix} k^{-m/n} A_\gamma & k^{-m/n} B_\gamma \\ 0 & I \end{pmatrix} x_0 = \begin{pmatrix} k^{-m/n} A_\gamma & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & A_\gamma^{-1} B_\gamma \\ 0 & I \end{pmatrix} x_0, \tag{6.6}$$

and indeed the concrete form of  $\gamma$  (6.2) gives by a straightforward calculation that

$$k A_\gamma^{-1} B_\gamma = \begin{pmatrix} -\text{diag}(f_1 \dots f_n) \\ 0_{m-n \times m-n} \end{pmatrix}, \tag{6.7}$$

so indeed  $(k A_\gamma^{-1} B_\gamma)^t$  is  $k$ -primitive. □

In particular, we record the following corollary for future reference.

**COROLLARY 6.3.** *For any  $k > k_0$  and any  $x \in X_m$ ,  $\text{supp}(a(k)_* \mu_k) \cap \text{Tor}_{k_0}(\pi_3^{-1}(x)) = \emptyset$ .*

We now turn to the proof of Step 3. Let  $\sigma$  be a weak\* accumulation point of  $\{a(k)_* \mu_k\}_{k=1}^\infty$ . By Lemmas 2.1 and 4.1,  $\sigma$  is a  $\Lambda$ -invariant measure supported in  $Hx_0$ . By Step 2 we conclude that  $\sigma$  is a probability measure that projects under  $\pi_3$  to either  $\mathbf{m}_{X_m}$  or  $\delta_{x_0}$  according to whether  $m > n$  or  $m = n$ , respectively. By Corollary 4.11, there are only countably many ergodic  $\Lambda$ -invariant probability measures supported in  $Hx_0$ , and so let  $\{\sigma_i\}_{i=0}^\infty$  be an enumeration of them. By the ergodic decomposition we may represent  $\sigma = \sum_{i=0}^\infty c_i \sigma_i$ , where  $c_i \geq 0$  and  $\sum_{i=0}^\infty c_i = 1$ .

*Assume  $m = n$ .* Let us denote  $\sigma_0 = \mathbf{m}_{V_{x_0}}$ . With this notation we will conclude this case once we show that  $c_i = 0$  for  $i > 0$ . Fixing  $i > 0$  and using the terminology of Definition 4.10,  $\sigma_i$  is of type (1)–(4). According to the type we will show that  $c_i = 0$ .

It is impossible that  $\sigma_i$  is of type (2) because, since  $(\pi_3)_* \sigma = \delta_{x_0}$  by Step 2, it would follow from Corollary 4.11(2) that  $\sigma_i = \mathbf{m}_{V_{x_0}} = \sigma_0$ . Similarly, Corollary 4.11 implies that if  $\sigma_i$  is of type (3) or (4) then  $c_i = 0$ . Finally, if  $i$  is such that  $\sigma_i$  is of type (1) then by Corollary 6.3 we have that the finite  $\Lambda$ -orbit on which  $\sigma_i$  is supported is disjoint, for large values of  $k$ , from the support of  $a(k)_* \mu_k$ . By Theorem 3.1 we deduce that  $c_i = 0$ . Note that, similarly to the discussion at the end of §6.1, the application of Theorem 3.1 was done with the following choices:  $G = \Lambda \ltimes V$ ,  $\Gamma = \Lambda \ltimes \text{Mat}_{m \times n}(\mathbb{Z})$ ,  $L = \Lambda$ ,  $z$  being a point in the support of  $\sigma_i$  and  $W = \text{Lie}(V) = \text{Mat}_{m \times n}(\mathbb{R})$  which indeed has no non-zero  $\Lambda$ -fixed vectors. This concludes the proof of Step 3 for the case  $n = m$ .

*Assume  $m > n$ .* Let us denote  $\sigma_0 = \mathbf{m}_{Hx_0}$ , the only measure of type (4). With this notation we will conclude this case once we show that  $c_i = 0$  for all  $i > 0$ . We do this by considering the possible types. By Step 2,  $(\pi_3)_* \sigma = \mathbf{m}_{X_m}$ , so we conclude from Corollary 4.11 that for any  $i$  such that  $\sigma_i$  is of type (1) or (2) we have that  $c_i = 0$ . Let  $i > 0$  be such that  $\sigma_i$  is of type (3). Note that by Corollary 4.11, the intersection of the support of a measure of type (3) with any  $\pi_3$ -fiber  $\pi_3^{-1}(x)$  is contained in  $\text{Tor}_{k_0}(\pi_3^{-1}(x))$  for a fixed  $k_0$  that does not depend on the choice of  $x \in X_m$ . By Corollary 6.3 we conclude that for  $k > k_0$  the support of  $a(k)_* \mu_k$  is disjoint from the periodic  $\Lambda \cdot \text{SL}_m(\mathbb{R})$ -orbit on which  $\sigma_i$  is supported. It now follows from Theorem 3.1 that  $c_i = 0$ . Theorem 3.1 is applied with the following choices:  $G = \Lambda \cdot H$ ,  $\Gamma = \Lambda \ltimes \text{Mat}_{m \times n}(\mathbb{Z})$ ,  $L = \Lambda \cdot \text{SL}_m(\mathbb{R})$ ,  $z$  any point in the support of  $\sigma_i$ , and  $W = \text{Lie}(V) = \text{Mat}_{m \times n}(\mathbb{R})$ . This concludes the proof of Step 3.

**6.4 Step 4**

In this step we conclude the proof of Theorem 1.3; that is, we establish the convergence  $\tilde{a}(k)_* \tilde{\mu}_k \xrightarrow{w^*} \mathbf{m}_{Ux_0} \times \boldsymbol{\theta}$ . At some point close to the end of the proof we will need the following lemma.<sup>8</sup>

LEMMA 6.4. *In the case<sup>9</sup>  $m = n$ ,*

$$\tilde{a}_k \tilde{\mathcal{P}}_k = \left\{ \left( \begin{pmatrix} I & 0 \\ k^{-1}\mathbf{u} & I \end{pmatrix} x_0, \begin{pmatrix} I & k^{-1}\mathbf{u}^* \\ 0 & I \end{pmatrix} x_0 \right) : \mathbf{u} \text{ is } k\text{-primitive} \right\},$$

where  $\mathbf{u}^* \in \text{Mat}_n(\mathbb{Z})$  is an inverse of  $\mathbf{u}$  modulo  $k$ .

*Proof.* The proof is a further inspection of the arguments in Lemmas 6.1 and 6.2, so we keep it short. We need to show that  $a_k x_{k^{-1}\mathbf{u}} = \begin{pmatrix} I & k^{-1}\mathbf{u}^* \\ 0 & I \end{pmatrix} x_0$ , and by acting with  $\Lambda$  we see that it is enough to show this for  $\mathbf{u}$  as in (2.10). For such  $\mathbf{u}$  this follows from (6.6) and (6.7), taking into account the concrete form of  $\gamma$  given by (6.2). □

We now turn to the proof of Step 4. Let  $\tilde{\sigma}$  be a weak\* accumulation point of the sequence  $\{\tilde{a}(k)_* \tilde{\mu}_k\}$ . By Steps 1 and 3 we know that  $(\pi_1)_* \tilde{a}(k)_* \tilde{\mu}_k \xrightarrow{w^*} \mathbf{m}_{Ux_0}$  and  $(\pi_2)_* \tilde{a}(k)_* \tilde{\mu}_k \xrightarrow{w^*} \boldsymbol{\theta}$ . As  $\pi_2$  is proper we deduce from the fact that  $\boldsymbol{\theta}$  is a probability measure that  $\tilde{\sigma}$  is a probability measure. This in turn implies that  $(\pi_1)_* \tilde{\sigma} = \mathbf{m}_{Ux_0}$  and  $(\pi_2)_* \tilde{\sigma} = \boldsymbol{\theta}$ . Also, by Lemma 4.1(5),  $\tilde{\sigma}$  is  $\Lambda_\Delta$ -invariant. Hence, the dynamical system  $(Ux_0 \times Hx_0, \tilde{\sigma}, \Lambda_\Delta)$  is a joining of  $(Ux_0, \mathbf{m}_{Ux_0}, \Lambda)$  and  $(Hx_0, \boldsymbol{\theta}, \Lambda)$ . By Theorems 4.12 and 4.15, we deduce that if  $(m, n) \neq (2, 2)$  then  $\tilde{\sigma} = \mathbf{m}_{Ux_0} \times \boldsymbol{\theta}$  and the proof is concluded. We are thus left to deal with the case  $n = m = 2$ .

6.4.1 *The case  $n = m = 2$ .* As the projections  $\mathbf{m}_{Ux_0}$  and  $\mathbf{m}_{Vx_0}$  are  $\Lambda$ -ergodic we conclude that (almost) any ergodic component of  $\tilde{\sigma}$  is a joining of  $\mathbf{m}_{Ux_0}$  and  $\mathbf{m}_{Vx_0}$  as well. By Theorem 4.15 there are only countably many such ergodic components and we will be done once we prove the following lemma in which we use the notation of (4.5).

LEMMA 6.5. *Let  $p, q$  be coprime integers and let  $(x, y) \in Ux_0 \times Vx_0$ . Then  $\tilde{\sigma}(L_{p,q}(x, y)) = 0$ .*

*Proof.* By Corollary 3.4 we deduce that it is enough to prove that

$$\frac{|\tilde{a}_k \tilde{\mathcal{P}}_k \cap L_{p,q}(x, y)|}{|\tilde{\mathcal{P}}_k|} \rightarrow 0 \quad \text{as } k \rightarrow \infty. \tag{6.8}$$

Here Corollary 3.4 is applied with the following choices:  $G = \Lambda_\Delta \ltimes (U \times V)$ ,  $\Gamma = \Lambda_\Delta \ltimes (U(\mathbb{Z}) \times V(\mathbb{Z}))$ ,  $L = L_{p,q}$ , and  $z = (x, y)$ . The existence of the  $\Lambda_\Delta$ -invariant linear complement of  $\text{Lie}(L_{p,q})$  follows from the semisimplicity of the group  $\text{SL}_2(\mathbb{R}) \times \text{SL}_2(\mathbb{R})$  and the fact that  $\text{Lie}(L_{p,q})$  is an invariant subspace of  $\text{Mat}_2(\mathbb{R}) \times \text{Mat}_2(\mathbb{R})$  under the representation (4.6).

For  $\mathbf{v} \in \text{Mat}_2(\mathbb{R})$  let us denote  $y_{\mathbf{v}} \stackrel{\text{def}}{=} \begin{pmatrix} I & \mathbf{v} \\ 0 & I \end{pmatrix} x_0 \in Vx_0$ , and for  $\mathbf{u} \in \text{GL}_2(\mathbb{Z}/k\mathbb{Z})$  we write (using the notation of Lemma 6.4)

$$\Phi(\mathbf{u}) \stackrel{\text{def}}{=} (x_{k^{-1}\mathbf{u}}, y_{k^{-1}\mathbf{u}^*}). \tag{6.9}$$

Lemma 6.4 gives that  $\Phi$  is a bijection between  $\text{GL}_2(\mathbb{Z}/k\mathbb{Z})$  and

$$\tilde{a}_k \tilde{\mathcal{P}}_k = \{\Phi(\mathbf{u}) : \mathbf{u} \in \text{GL}_2(\mathbb{Z}/k\mathbb{Z})\}. \tag{6.10}$$

<sup>8</sup> This lemma is needed only in the case  $m = n = 2$ .

<sup>9</sup> Compare this with (2.4).

The left and right actions of  $\mathrm{GL}_2(\mathbb{Z}/k\mathbb{Z})$  on itself induce (via  $\Phi$ ) actions of it on  $\tilde{\mathcal{P}}_k$ . Fixing  $\mathbf{u} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}/k\mathbb{Z})$ , we will establish (6.8) by analyzing the orbit  $\Phi(\{g_\ell \mathbf{u} : \ell \in (\mathbb{Z}/k\mathbb{Z})^\times\})$ , where  $g_\ell \stackrel{\text{def}}{=} \begin{pmatrix} \ell & 0 \\ 0 & 1 \end{pmatrix}$ , which is of size  $\phi(k)$  (here  $\phi$  is the Euler function), and proving that

$$\frac{|\{\ell : \Phi(g_\ell \mathbf{u}) \text{ is in the same } L_{p,q}\text{-orbit of } \Phi(\mathbf{u})\}|}{\phi(k)} \xrightarrow{k \rightarrow \infty} 0. \tag{6.11}$$

Following the definition, we see that  $\Phi(\mathbf{u})$  and  $\Phi(g_\ell \mathbf{u})$  are in the same  $L_{p,q}$ -orbit if and only if the difference  $(g_\ell \mathbf{u}, \mathbf{u}^* g_\ell^*) - (\mathbf{u}, \mathbf{u}^*)$  (thought of as an element of  $\mathrm{Mat}_2(\mathbb{R}) \times \mathrm{Mat}_2(\mathbb{R})$ ) is in  $\mathrm{Lie}(L_{p,q}) + \mathrm{Mat}_2(k\mathbb{Z}) \times \mathrm{Mat}_2(k\mathbb{Z})$ . In other words, if we let  $j = \det \mathbf{u}$  and recall that  $\mathbf{u}^* = j^* \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ ,  $g_\ell^* = \ell^* \begin{pmatrix} 1 & 0 \\ 0 & \ell \end{pmatrix}$ , and the definition of  $L_{p,q}$  in (4.5), we see that  $\Phi(\mathbf{u})$  and  $\Phi(g_\ell \mathbf{u})$  are in the same  $L_{p,q}$ -orbit if and only if

$$p \begin{pmatrix} 0 & (1 - \ell)b \\ 0 & (\ell - 1)a \end{pmatrix} = q \begin{pmatrix} ((\ell^* - 1)j^* d & 0 \\ (1 - \ell^*)j^* c & 0 \end{pmatrix} \pmod{k\mathbb{Z}}. \tag{6.12}$$

We claim that (6.12) implies  $\ell = 1 \pmod{k\mathbb{Z}}$ . For that purpose we split  $k = k_1 k_2$  with  $k_1$  coprime to  $p$  and  $k_2$  coprime to  $q$ . Using the second column in (6.12), we get  $(\ell - 1)b \in k_1\mathbb{Z}$  and  $(\ell - 1)a \in k_1\mathbb{Z}$ . However, since  $a, b$  form the first row of an invertible matrix modulo  $k$ , we see that the greatest common divisor of  $a, b, k_1$  is 1 and we conclude  $\ell = 1 \pmod{k_1\mathbb{Z}}$ . Using the first column of (6.12) and invertibility of  $j^*$  modulo  $k_2$  in the same way, we get  $\ell^* = 1 \pmod{k_2\mathbb{Z}}$ . Together this shows  $\ell = 1 \pmod{k\mathbb{Z}}$  and so (6.11) (since  $\phi(k) \rightarrow \infty$ ). This concludes the proof the lemma and thus the proof of Theorem 1.3. □

### 7. A strengthening of Theorem 1.3

Before we turn to the closing section of this paper, we observe that the proof of Theorem 1.3 presented above actually establishes more than what is stated. So far we have looked at (various images of) the counting measure on  $\tilde{\mathcal{P}}_k$  and used its  $\Lambda_\Delta$ -invariance, but  $\tilde{\mathcal{P}}_k$  splits into  $\Lambda_\Delta$ -orbits and we could restrict our attention to single orbits. Let

$$I_k \stackrel{\text{def}}{=} \{\vec{\ell} = (\ell_1, \dots, \ell_n) \in \mathbb{Z}^n : \ell_i | \ell_{i+1}, \gcd(\ell_i, k) = 1\}, \tag{7.1}$$

and denote by  $[I_k]$  the image of  $I_k$  in  $(\mathbb{Z}/k\mathbb{Z})^n$ . We refer to the elements of  $[I_k]$  as *k-primitive elementary divisors tuples* and sometimes, in an abuse of notation, we do not distinguish between elements in  $[I_k]$  and their integer representatives. It follows from Lemma 2.3 that a  $\Lambda_\Delta$ -orbit in  $\tilde{\mathcal{P}}_k$  is a set of the form

$$\mathcal{P}_{k, \vec{\ell}} \stackrel{\text{def}}{=} \{(x_{k^{-1}\mathbf{u}}, x_{k^{-1}\mathbf{u}}) : \text{the elementary divisor tuple of } \mathbf{u} \text{ equals } \vec{\ell} \pmod{k}\}.$$

Let  $\tilde{\mu}_{k, \vec{\ell}}$  be the normalized counting measure on  $\tilde{\mathcal{P}}_{k, \vec{\ell}}$  so that  $\tilde{\mu}_k$  is an average of the  $\tilde{\mu}_{k, \vec{\ell}}$ .

**THEOREM 7.1.** *If  $(m, n) \neq (2, 2), (1, 1)$ . Then, for any choice of a sequence of k-primitive elementary divisor tuples  $\vec{\ell}_k \in [I_k]$  (as  $k \rightarrow \infty$ ) we have that  $\tilde{a}(k) * \tilde{\mu}_{k, \vec{\ell}_k} \xrightarrow{w^*} \mathbf{m}_{Ux_0} \times \theta$ .*

*Proof.* We use the  $\Lambda_\Delta$ -invariance of  $\tilde{\mu}_{k, \vec{\ell}_k}$  and follow Steps 1–4 as presented in §2.2. The proof of Theorem 1.3 presented above carries over verbatim, replacing  $\tilde{\mu}_k$  with  $\tilde{\mu}_{k, \vec{\ell}_k}$ . □

It is interesting to note that the statement of Theorem 7.1 is false in the cases  $(m, n) = (1, 1), (2, 2)$ . Indeed, for  $(n, m) = (1, 1)$  the measures  $\tilde{\mu}_{k, \vec{\ell}_k}$  are Dirac masses. For  $(n, m) = (2, 2)$  we note that the argument presented above establishing Theorem 1.3 in this case used in § 6.4.1, in addition to the  $\Lambda_\Delta$ -invariance, the invariance of  $\tilde{\mu}_k$  under a  $\text{GL}_2(\mathbb{Z}/k\mathbb{Z})$ -action which is no longer present when considering the  $\tilde{\mu}_{k, \vec{\ell}}$ . In fact, if  $n = m = 2$  and we choose for any  $k$  the  $k$ -primitive elementary divisors tuple to be  $\vec{\ell} = (1, 1)$  then the set  $\text{supp}(\tilde{a}(k)_* \tilde{\mu}_{k, \vec{\ell}}) = \tilde{a}(k) \tilde{\mathcal{P}}_{k, \vec{\ell}}$ , which is the  $\Lambda_\Delta$ -orbit of  $\Phi\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$  ( $\Phi$  as in (6.9)), is contained in the fixed orbit  $L_{1,1}(x_0, x_0) \subset Ux_0 \times Vx_0$  (notation being as in (4.5)). In particular,  $\tilde{a}(k)_* \tilde{\mu}_{k, \vec{\ell}}$  cannot be expected to converge to  $\mathbf{m}_{Ux_0} \times \mathbf{m}_{Vx_0}$ . This shows that the statement of Theorem 7.1 is simply false in this case.

8. Proof of a strengthening of Theorem 1.4

Below we state and prove a generalization of Theorem 1.4. Recall the notation of § 1.2 and, in particular, that  $n = 1$  and so  $d = m + 1$ . Theorem 7.1 takes on a rather concrete meaning: the set  $[I_k]$  from (7.1) is simply  $(\mathbb{Z}/k\mathbb{Z})^\times$  and the elementary divisor tuple of a vector  $\mathbf{u}$  is simply  $\text{gcd}(\mathbf{u})$ , so that for  $c_k \in (\mathbb{Z}/k\mathbb{Z})^\times$ , the measure  $\tilde{\mu}_{k, c_k}$  is simply the normalized counting measure on the diagonal embedding of the collection  $\{x_{k^{-1}\mathbf{u}} : \text{gcd}(\mathbf{u}) = c_k\}$  in  $Ux_0 \times Ux_0$ . Theorem 7.1 states that  $\tilde{a}(k)_* \tilde{\mu}_{k, c_k} \xrightarrow{w^*} \mathbf{m}_{Ux_0} \times \mathbf{m}_{Hx_0}$  for any choice of  $c_k < k$  coprime to  $k$ . In particular, it follows that if  $\mu_{k, c_k} \stackrel{\text{def}}{=} (\pi_2)_* \tilde{\mu}_{k, c_k}$ , then  $a(k)_* \mu_{k, c_k} \rightarrow \mathbf{m}_{Hx_0}$ . We shall need the following corollary of Theorem 7.1.

COROLLARY 8.1. *Let  $F \subset Ux_0$  be a subset satisfying (i)  $\mathbf{m}_{Ux_0}(F) > 0$ , (ii)  $\mathbf{m}_{Ux_0}(\partial F) = 0$  (where  $\partial F$  denotes the boundary of  $F$  in  $Ux_0$ ), and let  $f \stackrel{\text{def}}{=} (1/\mathbf{m}_{Ux_0}(F))\chi_F$ . For any sequence  $c_k$  with  $1 \leq c_k \leq k$  and  $\text{gcd}(k, c_k) = 1$ , we have  $a(k)_* f d\mu_{k, c_k} \xrightarrow{w^*} \mathbf{m}_{Hx_0}$ .*

*Proof.* Given a test function  $\varphi \in C_c(X_d)$ , we consider the function  $(f \times \varphi)$  on  $Ux_0 \times X_d$  given by  $(f \times \varphi)(x, y) = f(x)\varphi(y)$ . Although  $(f \times \varphi)$  is not continuous, assumption (ii) implies that the points of discontinuity are of  $\mathbf{m}_{Ux_0} \times \mathbf{m}_{Hx_0}$ -measure zero. Using this, the weak\* convergence  $\tilde{a}(k)_* \tilde{\mu}_{k, c_k} \rightarrow \mathbf{m}_{Ux_0} \times \mathbf{m}_{Hx_0}$  from Theorem 7.1 implies the convergence

$$\begin{aligned} \int_{X_d \times X_d} (f \times \varphi) d\tilde{a}(k)_* \tilde{\mu}_{k, c_k} &\rightarrow \int_{X_d \times X_d} (f \times \varphi) d(\mathbf{m}_{Ux_0} \times \mathbf{m}_{Hx_0}) \\ &= \int_{X_d} \varphi d\mathbf{m}_{Hx_0}. \end{aligned}$$

We conclude that indeed  $a(k)_* f d\mu_{k, c_k} \xrightarrow{w^*} \mathbf{m}_{Hx_0}$ , as desired. □

In the course of proving Theorem 1.4 it is clearly harmless to assume that the set  $\mathcal{F}$  is contained in the face

$$\{(u_1, \dots, u_m, 1) \in \mathbb{R}^d : |u_i| < 1\} \subset \partial B_\infty.$$

Recall that  $\eta_k$  is the normalized counting measure on the set

$$\{[\Lambda_v] : v \in k\mathcal{F}, v = (\mathbf{u}, k), \text{gcd}(\mathbf{u}, k) = 1\}.$$

We split this set (and  $\eta_k$ ) according to the  $\Lambda$ -orbit of the various  $\mathbf{u}$ : for  $1 \leq c \leq k$  with  $\text{gcd}(c, k) = 1$ , we denote by  $\eta_{k, c}$  the normalized counting measure on  $\{[\Lambda_v] : v \in k\mathcal{F}, v = (\mathbf{u}, k), \text{gcd}(\mathbf{u}) = c\}$ ; then  $\eta_k$  is a convex combination of the  $\eta_{k, c}$ . We conclude that Theorem 1.4 is implied by the following theorem.



**THEOREM 8.2.** *For any positive integer  $k$ , choose  $1 \leq c_k \leq k$  with  $\gcd(k, c_k) = 1$ . Let  $\mathcal{F} \subset \{(u_1, \dots, u_m, 1) \in \mathbb{R}^d : |u_i| < 1\}$  be a measurable set of positive measure with boundary in  $\partial B_\infty$  of measure 0 with respect to the  $m$ -dimensional Lebesgue measure on  $\partial B_\infty$ . Then  $\eta_{k,c_k} \xrightarrow{w^*} \mathbf{m}_{Z_m}$  as  $k \rightarrow \infty$ .*

*Proof.* Given coprime integers  $c, k$ , we will use the abbreviation  $\widehat{\mathbb{Z}}_{k,c}^d \stackrel{\text{def}}{=} \{(\mathbf{u}, k) \in \mathbb{Z}^d : \gcd(\mathbf{u}) = c\}$ . We need to show that for any  $f \in C_c(Z_m)$ ,

$$\frac{1}{|\widehat{\mathbb{Z}}_{k,c_k}^d \cap k\mathcal{F}|} \sum_{v \in \widehat{\mathbb{Z}}_{k,c_k}^d \cap k\mathcal{F}} f([\Lambda_v]) \rightarrow \int_{Z_m} f d\mathbf{m}_{Z_m} \quad \text{as } k \rightarrow \infty. \tag{8.1}$$

Here and below we use the notation  $[\Lambda]$  to denote the projection of a lattice  $\Lambda \in X_m$  to  $Z_m$ . Fix a function  $f \in C_c(Z_m)$ ,  $\epsilon > 0$ , and a subset  $F \subset \mathcal{F}$  as in Corollary 8.1. Below we think of  $X_m$  (respectively,  $Z_m$ ) as the space of lattices (respectively,  $\text{SO}_m(\mathbb{R})$ -orbits of such) up to homothety in the hyperplane  $\mathbb{R}^m$  spanned by the first  $m$  coordinate axes in  $\mathbb{R}^d$ .

We identify  $\{(u_1, \dots, u_m, 1) \in \mathbb{R}^d : 0 \leq u_i < 1\}$  with  $Ux_0 \subset X_d$  in the obvious manner; that is,  $(\mathbf{u}, 1)$  corresponds to  $x_{\mathbf{u}}$ . This way we may think of  $F$  as a subset of  $X_d$  or  $\mathbb{R}^d$  at our convenience. It follows from the discussion in Remark 2.5(2) that if we denote by  $p : \mathbb{R}^d \rightarrow \mathbb{R}^m$  the projection onto the first  $m$  coordinates, then the measure  $(\pi_3)_* a(k)_*(\chi_F d\mu_{k,c_k})$  on  $X_m$  (when normalized to be a probability measure) is the normalized counting measure supported on the collection

$$\{p(\Lambda_v) \in X_m : v \in \widehat{\mathbb{Z}}_{k,c_k}^d \cap kF\}$$

(here we think of  $F$  as a subset of  $\partial B_\infty$  and of  $kF$  as a subset of  $\mathbb{R}^d$ ). Since we assume  $F$  has positive measure, Corollary 8.1 tells us that these measures converge to  $\mathbf{m}_{X_m}$  as  $k \rightarrow \infty$ , that is,

$$\frac{1}{|\widehat{\mathbb{Z}}_{k,c_k}^d \cap kF|} \sum_{v \in \widehat{\mathbb{Z}}_{k,c_k}^d \cap kF} \delta_{p(\Lambda_v)} \xrightarrow{w^*} \mathbf{m}_{X_m} \quad \text{as } k \rightarrow \infty. \tag{8.2}$$

The convergence in (8.2) is not too far from implying (8.1); we only need to deal with the distortion that the orthogonal projection  $p$  brings into the picture. We will see that this distortion is controllable by splitting  $\mathcal{F}$  into finitely many sets  $F \subset \mathcal{F}$  whose diameters are sufficiently small, say smaller than  $\delta > 0$ .

Choose a finite partition  $\mathcal{F} = \bigsqcup_{i=1}^\ell F_i$ , where the sets  $F_i$  are measurable, have boundary of Lebesgue measure zero (relative to  $\partial B_\infty$ ), and have diameter less than or equal to  $\delta$ . In order to conclude the proof we will use (8.2) to show that once  $\delta$  is chosen small enough, for any large enough  $k$  and for any  $1 \leq i \leq \ell$ ,

$$\left| \int_{Z_m} f d\mathbf{m}_{Z_m} - \frac{1}{|\widehat{\mathbb{Z}}_{k,c_k}^d \cap kF_i|} \sum_{v \in \widehat{\mathbb{Z}}_{k,c_k}^d \cap kF_i} f([\Lambda_v]) \right| \leq \epsilon. \tag{8.3}$$

Since  $\epsilon$  is arbitrary, this implies (8.1).

To this end, fix  $1 \leq i \leq \ell$  and denote  $F = F_i$ . Choose a reference point  $v_0 \in F$  and let  $S$  be the inverse of the restriction of the projection  $p$  to the  $m$ -dimensional linear space  $\{v_0\}^\perp$ . Let  $k_{v_0} \in \text{SO}_d(\mathbb{R})$  be chosen as described in § 1.2 (i.e. such that  $k_{v_0}\{v_0\}^\perp = \mathbb{R}^m$ ), and consider the function  $\varphi \in C_c(X_m)$  defined by  $\varphi(\Lambda) \stackrel{\text{def}}{=} f([k_{v_0}S(\Lambda)])$ . Note that this definition depends on

the choice of  $v_0$  but is independent of the choice of  $k_{v_0}$ , and also that  $\int_{X_m} \varphi \, d\mathbf{m}_{X_m} = \int_{Z_m} f \, d\mathbf{m}_{Z_m}$ . Applying (8.2) to the function  $\varphi$ , we conclude that

$$\lim_k \left| \frac{1}{|\widehat{\mathbb{Z}}_{k,c_k}^d \cap kF|} \sum_{v \in \widehat{\mathbb{Z}}_{k,c_k}^d \cap kF} f([k_{v_0}S(p(\Lambda_v))]) - \int_{Z_m} f \, d\mathbf{m}_{Z_m} \right| = 0. \tag{8.4}$$

Let  $\text{Cone}(F) \stackrel{\text{def}}{=} \{tF : t > 0\}$ . Let  $d_{Z_m}(\cdot, \cdot)$  denote a distance function on  $Z_m$  and assume for the moment that it satisfies

$$\sup_{v \in \mathbb{Z}^d \cap \text{Cone}(F)} d_{Z_m}([\Lambda_v], [k_{v_0}S(p(\Lambda_v))]) \rightarrow 0 \text{ as } \text{diam}(F) \rightarrow 0. \tag{8.5}$$

Using (8.5) and the uniform continuity of  $f$ , we see that once  $\delta > 0$  is chosen small enough,  $|f([k_{v_0}S(p(\Lambda_v))]) - f([\Lambda_v])| \leq \epsilon$ . Plugging this into (8.4) gives (8.3) as desired.

We are left to explain the validity of (8.5). We let  $d_{\text{SL}_m(\mathbb{R})}(\cdot, \cdot)$  denote a right  $\text{SL}_m(\mathbb{R})$ -invariant metric on  $\text{SL}_m(\mathbb{R})$ . With it we induce metrics  $d_{X_m}, d_{Z_m}$  on  $X_m, Z_m$  respectively by the formulas

$$\begin{aligned} d_{X_m}(g \text{SL}_m(\mathbb{Z}), h \text{SL}_m(\mathbb{Z})) &= \inf_{\gamma \in \text{SL}_m(\mathbb{Z})} d_{\text{SL}_m(\mathbb{R})}(g\gamma, h), \\ d_{Z_m}([x], [y]) &= \inf_{k_1, k_2 \in \text{SO}_m(\mathbb{R})} d_{X_m}(k_1x, k_2y) \end{aligned}$$

for  $g, h \in \text{SL}_m(\mathbb{Z})$  and  $x, y \in X_m$ . The validity of (8.5) now follows as

$$\begin{aligned} d_{Z_m}([\Lambda_v], [k_{v_0}S(p(\Lambda_v))]) &\leq d_{X_m}(k_v\Lambda_v, k_{v_0}Spk_v^{-1}k_v\Lambda_v) \\ &\leq d_{\text{SL}_d(\mathbb{R})}(e, k_{v_0}Spk_v^{-1}) \end{aligned}$$

and as  $\text{diam}(F) \rightarrow 0$ , if the matrix  $k_v$  is chosen correctly,

$$k_{v_0}Spk_v^{-1}|_{\mathbb{R}^m} : \mathbb{R}^m \rightarrow \mathbb{R}^m$$

approaches the identity. □

ACKNOWLEDGEMENTS

We would like to thank Han Li and Jens Marklof for valuable discussions. We further thank the Erwin Schrödinger Institute in Vienna and the Forschungsinstitut Mathematik at the ETH Zurich for their hospitality. Finally, we thank the referee for critical comments that improved the presentation significantly.

REFERENCES

AES14 M. Aka, M. Einsiedler and U. Shapira, *Integer points on spheres and their orthogonal grids*, J. Lond. Math. Soc. (2), to appear, [arXiv:1411.1272](https://arxiv.org/abs/1411.1272).  
 AES15 M. Aka, M. Einsiedler and U. Shapira, *Integer points on spheres and their orthogonal lattices (with an appendix by Ruixiang Zhang)*, Preprint (2015), [arXiv:1502.04209](https://arxiv.org/abs/1502.04209).  
 COU01 L. Clozel, H. Oh and E. Ullmo, *Hecke operators and equidistribution of Hecke points*, Invent. Math. **144**(2) (2001), 327–351; [MR 1827734](https://doi.org/10.1007/s002220000044) (2002m:11044).  
 Ein06 M. Einsiedler, *Ratner’s theorem on  $\text{SL}(2, \mathbb{R})$ -invariant measures*, Jahresber. Dtsch. Math.-Ver. **108**(3) (2006), 143–164; [MR 2265534](https://doi.org/10.1007/s00033-006-0048-1) (2008b:37048).  
 EO06 A. Eskin and H. Oh, *Ergodic theoretic proof of equidistribution of Hecke points*, Ergodic Theory Dynam. Systems **26**(1) (2006), 163–167; [MR 2201942](https://doi.org/10.1007/s10231-006-0068-1) (2006j:11068).

- Gla03 E. Glasner, *Ergodic theory via joinings*, Mathematical Surveys and Monographs, vol. 101 (American Mathematical Society, Providence, RI, 2003); [MR 1958753](#) (2004c:37011).
- Iwa02 H. Iwaniec, *Spectral methods of automorphic forms*, Graduate Studies in Mathematics, vol. 53, second edition (American Mathematical Society, Providence, RI, 2002); [MR 1942691](#) (2003k:11085).
- KSS02 D. Kleinbock, N. Shah and A. Starkov, *Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory*, in *Handbook of dynamical systems, Vol. 1A* (North-Holland, Amsterdam, 2002), 813–930; [MR 1928528](#) (2004b:22021).
- Mar04 G. A. Margulis, *On some aspects of the theory of Anosov systems*, Springer Monographs in Mathematics (Springer, Berlin, 2004); with a survey by Richard Sharp: periodic orbits of hyperbolic flows, translated from the Russian by Valentina Vladimirovna Szulikowska; [MR 2035655](#) (2004m:37049).
- Mar10 J. Marklof, *The asymptotic distribution of Frobenius numbers*, *Invent. Math.* **181**(1) (2010), 179–207; [MR 2651383](#) (2011e:11133).
- MS13 J. Marklof and A. Strömbergsson, *Diameters of random circulant graphs*, *Combinatorica* **33**(4) (2013), 429–466; [MR 3133777](#).
- MS95 S. Mozes and N. Shah, *On the space of ergodic invariant measures of unipotent flows*, *Ergodic Theory Dynam. Systems* **15**(1) (1995), 149–159.
- Rat91a M. Ratner, *On Raghunathan’s measure conjecture*, *Ann. of Math. (2)* **134**(3) (1991), 545–607; [MR 1135878](#) (93a:22009).
- Rat91b M. Ratner, *Raghunathan’s topological conjecture and distributions of unipotent flows*, *Duke Math. J.* **63**(1) (1991), 235–280; [MR 1106945](#) (93f:22012).
- Sch98 W. M. Schmidt, *The distribution of sublattices of  $\mathbf{Z}^m$* , *Monatsh. Math.* **125**(1) (1998), 37–81; [MR 1485976](#) (99c:11083).
- Swi01 H. P. F. Swinnerton-Dyer, *A brief guide to algebraic number theory*, London Mathematical Society Student Texts, vol. 50 (Cambridge University Press, Cambridge, 2001); [MR 1826558](#) (2002a:11117).
- Sha98 N. A. Shah, *Invariant measures and orbit closures on homogeneous spaces for actions of subgroups generated by unipotent elements*, in *Lie groups and ergodic theory* (Tata Institute for Fundamental Research, Mumbai, 1998), 229–271; [MR 1699367](#) (2001a:22012).
- Wit94 D. Witte, *Measurable quotients of unipotent translations on homogeneous spaces*, *Trans. Amer. Math. Soc.* **345**(2) (1994), 577–594; [MR 1181187](#) (95a:22005).

Manfred Einsiedler [manfred.einsiedler@math.ethz.ch](mailto:manfred.einsiedler@math.ethz.ch)  
ETH Zürich, CH-8092 Zürich, Switzerland

Shahar Mozes [mozes@math.huji.ac.il](mailto:mozes@math.huji.ac.il)  
Institute of Mathematics, The Hebrew University of Jerusalem,  
Givat Ram, 91904 Jerusalem, Israel

Nimish Shah [shah@math.osu.edu](mailto:shah@math.osu.edu)  
Ohio State University, Columbus 43210-1174, USA

Uri Shapira [ushapira@tx.technion.ac.il](mailto:ushapira@tx.technion.ac.il)  
Technion - Israel Institute of Technology, 32000 Haifa, Israel