SOME MEASURE RIGIDITY RESULTS FOR UNIPOTENT FLOWS ON HOMOGENEOUS SPACES WITH MINIMAL PRE-REQUISITES.

JON CHAIKA AND DONALD ROBERTSON

1. Introduction

The general setting for Ratner's theorems is any quotient G/Γ of a Lie group G by one of its lattices Γ and a one-parameter subgroup $t \mapsto u^t$ of unipotent elements of G. In a series of papers in the early nineties [**MR1075042**; **MR1062971**; **MR1135878**; **MR1106945**] Ratner proved many deep and far-reaching results about the resulting "unipotent flows" on G/Γ. Our goal in these notes is to describe the proof of one of these results – the classification of invariant probability measures – in two concrete cases: the modular surface $\mathsf{SL}(2,\mathbb{R})/\mathsf{SL}(2,\mathbb{Z})$ together with the one-parameter subgroup $\mathsf{u}^t = \left[\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right]$; and the moduli space $\mathsf{SL}(2,\mathbb{R}) \ltimes \mathbb{R}^2/\mathsf{SL}(2,\mathbb{Z}) \ltimes \mathbb{Z}^2$ of doubly marked tori together with the one-parameter subgroup $\tilde{\mathsf{u}}^t = (\left[\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right])$. In this introductory section we present some background on horocycle flows and a brief review of the earlier results before describing Ratner's theorems.

1.1. The horocycle flow on a compact hyperbolic surface.

Modeling the hyperbolic plane as the upper half-space

$$\mathbb{H} = \{x + iy \in \mathbb{C} : y > 0\}$$

equipped with the Riemannian metric

$$\langle v, w \rangle_{(x,y)} = \frac{v_1 w_1 + v_2 w_2}{4y^2}$$

every element of $\mathsf{SL}(2,\mathbb{R})$ defines an isometry of $\mathbb{H}$ by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} z = \frac{az + c}{bz + d}$$

for all $z \in \mathbb{H}$.

Exercise 1. Verify that $t \mapsto e^{2t} i$ is a geodesic for the above metric.

Solution. *The geodesic equations on* $\mathbb{H}$ *are*

$$\ddot{x}y = 2\dot{x}\dot{y}$$
$$\ddot{y}y = \dot{y}^2 - \dot{x}^2$$

and the given curve is easily verified to satisfy them.

The action of $\mathsf{SL}(2,\mathbb{R})$ is transitive and the stabilizer of i is the subgroup $\mathsf{SO}(2,\mathbb{R})$ of $\mathsf{SL}(2,\mathbb{R})$ so we can identify $\mathbb{H}$ with $\mathsf{SO}(2,\mathbb{R})\backslash\mathsf{SL}(2,\mathbb{R})$. In fact the action is transitive even on the unit tangent bundle $\mathsf{T}^1\mathbb{H}$ and the stabilizer of the tangent vector $i \in \mathsf{T}_i\mathbb{H}$ is $\{\pm I\}$ so we can identify $\mathsf{T}^1\mathbb{H}$ with $\mathsf{PSL}(2,\mathbb{R})$. The geodesic flow on $\mathsf{T}^1\mathbb{H}$ is then identical to the flow on $\mathsf{PSL}(2,\mathbb{R})$ of left multiplication by the one-parameter subgroup A parameterized by

$$\mathsf{g}^t = \begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix}$$

for all $t \in \mathbb{R}$.

Every compact, closed Riemann surface X with a metric of constant negative curvature is the quotient $\mathbb{H}/\Gamma$ of the hyperbolic plane $\mathbb{H}$ by a discrete and co-compact subgroup Γ of $\mathsf{PSL}(2,\mathbb{R})$. The unit tangent bundle of X can be identified with the quotient $\mathsf{PSL}(2,\mathbb{R})/\Gamma$ and the geodesic flow on T^1X is then identified with the flow on $\mathsf{PSL}(2,\mathbb{R})/\Gamma$ of left multiplication by the one-paramater subgroup $t \mapsto \mathsf{g}^t$. The Liouville measure on T^1X is identified up to normalization with the natural projection m of Haar measure to the quotient $\mathsf{PSL}(2,\mathbb{R})/\Gamma$. In 1936, E. Hopf [**MR1501848**] proved that the geodesic flow on X is "metrically transitive" or ergodic. That is, the only A invariant subsets of X have m measure either zero or one.

In these notes we are mainly concerned with another flow on the quotient $\mathsf{PSL}(2,\mathbb{R})/\Gamma$ namely the **(unstable) horocycle flow** arising from the subgroup N of $\mathsf{SL}(2,\mathbb{R})$ parameterized by

$$\mathsf{u}^t = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

for $t \in \mathbb{R}$. Unstable horocycle orbits have the following geometric property: all points in any given N orbit have the same "distant past" under the geodesic flow. Precisely, since

$$\mathsf{g}^{-t}\mathsf{u}^s = \mathsf{u}^{se^{-2t}}\mathsf{g}^{-t}$$

for all $s,t \in \mathbb{R}$ we see upon fixing a right-invariant metric D on $\mathsf{PSL}(2,\mathbb{R})$ that

$$\lim_{t\to\infty} \mathsf{D}(\mathsf{g}^{-t}\mathsf{u}^s x, \mathsf{g}^{-t}x) = \lim_{t\to\infty} \mathsf{D}(\mathsf{u}^{se^{-2t}}\mathsf{g}^{-t}x, \mathsf{g}^{-t}x) = \lim_{t\to\infty} \mathsf{D}(\mathsf{u}^{se^{-2t}}, I) = 0$$

for all $s \in \mathbb{R}$ and all $x \in \mathsf{T}^1\mathbb{H}$.

Hedlund [**MR1545946**] proved in 1936 that, whenever Γ is discrete and cocompact the unstable horocycle flow on $\mathsf{PSL}(2,\mathbb{R})/\Gamma$ is "transitive" or minimal, which is to say that the orbit of every point under the flow is dense in the space. Ergodicity of the horocycle flow was proved by Parasyuk [**MR0058883**] in 1953 using work of Gelfand and Fomin [**MR0052701**]. Ergodicity also follows from a later criterion due to Moore [**MR0193188**] expanding upon the Mautner phenomenon [**MR0084823**].

A continuous flow on a compact metric space is **uniquely ergodic** if there is a unique Borel probability measure on the space that is flow invariant. Furstenberg [**MR0393339**] proved in 1973 that the horocycle flow on $\mathsf{PSL}(2,\mathbb{R})/\Gamma$ is uniquely ergodic whenever Γ is discrete and cocompact.

1.2. **Unipotent flows on Lie groups.** The situation of general lattices – those discrete subgroups Γ of $\mathsf{PSL}(2,\mathbb{R})$ for which the quotient $\mathsf{PSL}(2,\mathbb{R})/\Gamma$ supports a finite measure invariant under multiplication by $\mathsf{PSL}(2,\mathbb{R})$ from the left – is more complicated as the quotient need no longer be compact. It is possible in this case for N to have periodic orbits. For example, every point of the form $(x,i)\mathsf{SL}(2,\mathbb{Z})$ in $\mathsf{SL}(2,\mathbb{R})/\mathsf{SL}(2,\mathbb{Z}) \cong \mathsf{PSL}(2,\mathbb{R})/\mathsf{PSL}(2,\mathbb{Z})$ has a periodic N orbit.

More generally, one is interested in the situation where G is a connected Lie group and Γ is a discrete subgroup such that the quotient G/Γ possesses a finite measures that is invariant under the action of G on G/Γ by left multiplication. An example is $G = \mathsf{SL}(2,\mathbb{R})$ and $\Gamma = \mathsf{SL}(2,\mathbb{Z})$. In place of the one-parameter subgroup $t \mapsto \mathsf{u}^t$ one is interested in the action on G/Γ of subgroups U with the property that every element of U is unipotent. An element g of a Lie group G is **unipotent** if $\mathsf{Ad}(g) - \mathsf{id}$ is nilpotent as a linear map of the Lie algebra $\mathfrak{g}$ of G. When G is a matrix group this is the same as nilpotence of the matrix $g - I$.

In 1971 Margulis [**MR0291352**; **MR0470140**] proved that unipotent flows on $\mathsf{SL}(n,\mathbb{R})/\mathsf{SL}(n,\mathbb{Z})$ are never divergent in the following sense: for every one-parameter subgroup $t \mapsto u^t$ of unipotent elements in $\mathsf{SL}(n,\mathbb{R})$, for every $x \in \mathsf{SL}(n,\mathbb{R})$, and for every compact subset K of $\mathsf{SL}(n,\mathbb{R})/\mathsf{SL}(n,\mathbb{Z})$, there is a sequence of times $t_n \nearrow \infty$ for which $u^{t_n} x \mathsf{SL}(n,\mathbb{Z}) \in K$. Dani [**MR758891**] proved the same result for semisimple Lie groups of real rank 1 in 1984, and the general result – for unipotent flows on quotients of semisimple Lie groups by uniform lattices – follows by combining these two works.

Earlier, Dani [**MR629475**] had generalized Furstenberg's work to the non-compact case, proving for lattices Γ in reductive Lie groups G that the left action of any maximal horospherical subgroup N on G/Γ has only algebraic measures as ergodic, invariant measures. A subgroup N of a Lie group G is **horospherical** if

$$N = \{u \in G : g^n u g^{-n} \to I \text{ as } n \to \infty\}$$

for some $g \in G$. In 1986, Dani [**MR835804**] showed the assumption that N be maximal is unnecessary, proving the above conclusion holds for any horospherical subgroup of G.

Exercise 2. Verify that N is a horospherical subgroup of $\mathsf{SL}(2,\mathbb{R})$.

Solution. *Fix $0 < a < 1$. We have*

$$\begin{pmatrix} a^n & 0 \\ 0 & a^{-n} \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} a^{-n} & 0 \\ 0 & a^n \end{pmatrix} = \begin{pmatrix} \alpha & a^{2n}\beta \\ a^{-2n}\gamma & \delta \end{pmatrix}$$

so must have

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

for some $t \in \mathbb{R}$ if the above expression is to converge to I as $n \to \infty$. With $g = \mathsf{g}^a$ we therefore have $U = \mathsf{N}$.

1.3. **Ratner's theorems.** At the beginning of the 1990s Ratner [**MR1075042**; **MR1062971**; **MR1135878**; **MR1106945**], based on methods from her earlier work on horocycle flows [**MR721735**; **MR657240**; **MR717825**] proved the following results on unipotent flows.

Theorem 1.1 (Ratner's measure classification theorem [**MR1135878**])**.** *Let G be a connected Lie group and let U be a subgroup of G every element of which is unipotent. For every discrete subgroup $\Gamma < G$ every U invariant probability measure on G/Γ is algebraic.*

Given a measure μ on G/Γ one can consider the closed subgroup $\Lambda(\mu) = \{g \in G : g\mu = \mu\}$ of G. A measure μ on G/Γ is **algebraic** if there is $x\Gamma \in G/\Gamma$ such that $\mu(\Lambda(\mu)x\Gamma) = 1$. If μ is algebraic then we can identify the orbit $\Lambda(\mu)x\Gamma$ with the quotient $\Lambda(\mu)/x\Gamma x^{-1} \cap \Lambda(\mu)$ to deduce that $x\Gamma x^{-1} \cap \Lambda(\mu)$ is a lattice in $\Lambda(\mu)$.

Theorem 1.2 (Ratner's orbit classification theorem [**MR1106945**])**.** *Let G be a connected Lie group and let U be a unipotent subgroup of G. For every lattice Γ in G and every $x \in G$ there is a closed subgroup $L < G$ containing U such that $xLx^{-1} \cap \Gamma$ is a lattice in L with $\overline{Ux\Gamma} = Lx\Gamma$.*

Theorem 1.3 (Ratner's genericity theorem [**MR1106945**]). *Let G be a connected Lie group and let U be a one-parameter subgroup of G. For every lattice Γ in G and every $x \in G$ there is an algebraic measure ν on G/Γ such that $x\Gamma$ is generic for ν along U. Moreover ν is the unique algebraic measure with support equal to $\overline{Ux\Gamma}$.*

These results were generalized by Shah [**MR1699367**] to (not necessarily connected) subgroups U generated by unipotent elements.

2. $\mathsf{SL}(2,\mathbb{R})/\mathsf{SL}(2,\mathbb{Z})$

In this section we prove Theorem 1.1 in the special case $G = \mathsf{SL}(2,\mathbb{R})$, $\Gamma = \mathsf{SL}(2,\mathbb{Z})$ and $U = \mathsf{N} = \{\mathsf{u}^t : t \in \mathbb{R}\}$.

2.1. **Setup.** Let $G = \mathsf{SL}(2,\mathbb{R})$, $\Gamma = \mathsf{SL}(2,\mathbb{Z})$ and $X = G/\Gamma$. The quotient X can be identified with the space of unit co-volume lattices in $\mathbb{R}^2$. That is, the space of closed, discrete subgroups of $\mathbb{R}^2$ on to whose quotients the Lebesgue measure of $\mathbb{R}^2$ descends with total measure 1.

To see this identification, first note that a discrete subgroup of $\mathbb{R}^2$ must have rank at most 2 as a $\mathbb{Z}$ module, and if the quotient is to have finite area then the rank must equal 2. Any such discrete subgroup can be mapped to any other by some element of $\mathsf{SL}(2,\mathbb{R})$. Indeed $\mathsf{SL}(2,\mathbb{R})$ acts transitively on pairs of vectors that give the columns of a matrix with a fixed determinant. The identification is now complete since $\mathsf{SL}(2,\mathbb{Z})$ is the stabilizer of the distinguished co-volume 1 lattice $\mathbb{Z}^2$. As tori can be identified with the quotient of $\mathbb{R}^2$ by a lattice X is also the space of area 1 tori together with a specific direction.

Noticing that $\mathsf{SL}(2,\mathbb{Z})$ is not the stabilizer of all such lattices is the same as noticing that it is not a normal subgroup of G and so our points in X are just left cosets. So the coset $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} g\Gamma$ is not necessarily the same as the coset $g\Gamma$.

We will be interested in the left action of $\mathsf{SL}(2,\mathbb{R})$ and some of its one-parameter subgroups on the quotient $X = \mathsf{SL}(2,\mathbb{R})/\mathsf{SL}(2,\mathbb{Z})$ of co-volume one lattices in $\mathbb{R}^2$. For any $t \in \mathbb{R}$ write

$$\mathsf{u}^t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \qquad \mathsf{v}^t = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \qquad \mathsf{g}^t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$$

as parameterizations of the **unstable horocycle** subgroup N, the **stable horocycle** subgroup M, and the **geodesic** subgroup A respectively.

Fix a right-invariant metric D on $\mathsf{SL}(2,\mathbb{R})$ that generates the standard topology on $\mathsf{SL}(2,\mathbb{R})$ and has the property that

$$\mathsf{D}(\mathsf{u}^t, I) \leq |t| \qquad \mathsf{D}(\mathsf{v}^t, I) \leq |t| \qquad \mathsf{D}(\mathsf{g}^t, I) \leq |t| \tag{1}$$

for all $t \in \mathbb{R}$. The quantity

$$\mathsf{d}(x\Gamma, y\Gamma) = \inf\{\mathsf{D}(x, y\gamma) : \gamma \in \Gamma\}$$

defines a metric on X. In particular, we have $\mathsf{d}(x\Gamma, y\Gamma) \leq \mathsf{D}(x, y)$ and

$$\mathsf{d}(x\Gamma, gx\Gamma) \leq \mathsf{D}(g, I) \tag{2}$$

for all $x, y, g \in G$.

Exercise 3. Verify that d is a metric on X and that (2) holds.

Solution. *First we show that it is a metric. It follows from right invariance of* D *that* d *is symmetric. If* $\mathsf{d}(x\Gamma, y\Gamma) = 0$ *then we can find* $\gamma_n \in \Gamma$ *with* $\mathsf{D}(x, y\gamma_n) < \frac{1}{n}$ *for all* $n \in \mathbb{N}$. *Thus* $\gamma_n \to y^{-1}x$ *in* G. *Therefore* $y^{-1}x \in \Gamma$ *giving* $x\Gamma = y\Gamma$.

For the triangle inequality fix $\epsilon > 0$ *and* $\gamma, \eta \in \Gamma$ *with*

$$\mathsf{d}(x\Gamma, y\Gamma) \leq \mathsf{D}(x, y\gamma) \leq \mathsf{d}(x\Gamma, y\Gamma) + \epsilon$$
$$\mathsf{d}(y\Gamma, z\Gamma) \leq \mathsf{D}(y, z\eta) \leq \mathsf{d}(y\Gamma, z\Gamma) + \epsilon$$

both holding. We have

$$\begin{aligned}\mathsf{d}(x\Gamma, z\Gamma) \leq \mathsf{D}(x, z\eta\gamma) &\leq \mathsf{D}(x, y\gamma) + \epsilon + \mathsf{D}(y\gamma, z\eta\gamma) + \epsilon \\ &\leq \mathsf{d}(x\Gamma, y\Gamma) + \mathsf{d}(y\Gamma, z\Gamma) + 2\epsilon\end{aligned}$$

as desired. For (2) *note that* $\mathsf{d}(x\Gamma, gx\Gamma) \leq \mathsf{D}(x, gx) = \mathsf{D}(I, g)$.

2.2. **The horocycle action.** Our main goal in this section is the following striking classification of N invariant probability measures on X.

Theorem 2.1. *Let* μ *be an* N *ergodic and* N *invariant probability measure on* X. *Either:*

(1) *μ is "Haar measure on X"; or*
(2) *μ is supported on a periodic orbit of* N.

By "Haar measure on X" we mean the following: choose a fundamental domain S for the Γ action on G. Let ν be the measure on X defined by $\nu(A) = \mathsf{m}(\{g \in G : g\Gamma \in A\} \cap S)$, where m denotes the usual Haar measure on G normalized so that $\mathsf{m}(S) = 1$.

We will often make use of the following decomposition

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{c}{a} & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \frac{1}{a} \end{pmatrix} \begin{pmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{pmatrix} \tag{3}$$

that allows us to write any element of $\mathsf{SL}(2,\mathbb{R})$ with $a \neq 0$ as a product of stable horocycle, geodesic, and unstable horocycle elements. In particular, we have

$$\mathsf{u}^s\mathsf{v}^t\mathsf{g}^a\mathsf{u}^\ell = \mathsf{v}^{\frac{t}{1+st}}\mathsf{g}^{a+\log(1+st)}\mathsf{u}^{\ell+\frac{s}{1+st}e^{-2a}} \tag{4}$$

for all $s,t,a,\ell \in \mathbb{R}$ with $1+st \neq 0$.

Exercise 4. Verify (4). (First check the decomposition

$$\mathsf{u}^s\mathsf{v}^t = \mathsf{v}^{\frac{t}{1+st}}\mathsf{g}^{\log(1+st)}\mathsf{u}^{\frac{s}{1+st}} \tag{5}$$

of $\mathsf{u}^s\mathsf{v}^t$. Then verify

$$\mathsf{u}^r\mathsf{g}^b = \mathsf{g}^b\mathsf{u}^{re^{-2b}} \tag{6}$$

for all $b,r \in \mathbb{R}$.)

For the remainder of this section, μ will denote an N ergodic and invariant probability measure.

2.3. **Proof of Theorem 2.1.** Fix an N invariant and N ergodic probability measure μ on X. We want to show that if μ is not supported on a periodic orbit then μ is $\mathsf{SL}(2,\mathbb{R})$ invariant. First, we prove that it suffices to show μ is g^a invariant for some non-zero $a \in \mathbb{R}$.

Proposition 2.2. *If μ is N ergodic and N invariant and g^a invariant for some non-zero $a \in \mathbb{R}$ then μ is Haar measure on X.*

Before getting to the proof of the proposition in earnest, we establish some useful general results that motivate our approach.

Our aim now is to find $a \in \mathbb{R}$ non-zero such that $\mathsf{g}^a\mu = \mu$. We begin with some basic results.

Lemma 2.3. *For every $a \in \mathbb{R}$ the measure $\mathsf{g}^a\mu$ is N invariant.*

Proof. Fix $s \in \mathbb{R}$ and $A \subset X$ measurable. Since
$mu(\mathsf{u}^sA) = \mu(A)$ we have

$$(\mathsf{g}^a\mu)(\mathsf{u}^sA) = \mu(\mathsf{g}^{-a}\mathsf{u}^sA) = \mu(\mathsf{u}^{sa^{-2}}\mathsf{g}^{-a}A) = \mu(\mathsf{g}^{-a}A) = (\mathsf{g}^a\mu)(A)$$

so $\mathsf{g}^a\mu$ is N invariant. □

Exercise 5. Prove that each of the measure $\mathsf{g}^a\mu$ is N ergodic probability measure.

Solution. *We first show that it is a probability measure*

$$(\mathsf{g}^a\mu)(X) = \mu(\mathsf{g}^{-a}X) = \mu(X) = 1$$

and so it is a probability measure. We see ergodicity similarly to above: If $(\mathsf{g}^a\mu)(\mathsf{u}^t A \setminus A) = 0$ *then*

$$\mu(\mathsf{g}^{-a}(\mathsf{u}^t A \setminus A)) = \mu\big(\mathsf{g}^{-a}\mathsf{u}_t A \setminus \mathsf{g}^{-a}A\big) = \mu\big(\mathsf{u}^{ta^{-2}}\mathsf{g}^{-a}A \setminus \mathsf{g}^{-a}A\big) = 0$$

and by the N*-ergodicity of* μ *this implies that* $\mu(\mathsf{g}^{-a}A) \in \{0,1\}$*. In turn this gives* $(\mathsf{g}^a\mu)(A) \in \{0,1\}$ *implying the* N *ergodicity of* $\mathsf{g}^a\mu$*.*

Corollary 2.4. For all $a \in \mathbb{R}$ either $\mu = \mathsf{g}^a\mu$ or $\mu \perp \mathsf{g}^a\mu$.

Proof. It is a standard fact from ergodic theory that any two ergodic invariant probability measures are either equal or mutually singular. Indeed if these measures ν_1, ν_2 are different there exists $f \in C_c(X) \subset L^1(\nu_1) \cap L^1(\nu_2)$ so that $\int f \,\mathrm{d}\nu_1 \neq \int f \,\mathrm{d}\nu_2$. Now by the Birkhoff ergodic theorem there exists sets A_1, A_2 of full ν_1, ν_2 measure respectively such that

$$\lim_{T\to\infty} \frac{1}{T} \int\limits_0^T f(U_s x) \,\mathrm{d}s = \int f \,\mathrm{d}\nu_i$$

for all $x \in A_i$. Since the right hand sides of the above formula are different for the two choices of i these sets must be disjoint implying the singularity of the measures. □

Lemma 2.5. *If* μ *is* N *ergodic and* g^a *ergodic for some* $a > 0$ *then* μ *is Haar measure on* X*.*

To prove this lemma we use the notion of generic points:

Defintion 2.6. *Let* μ *be a Borel probability measure and let* $t \mapsto F^t$ *be a measurable flow that preserves* μ*. We say* x *is a* μ ***generic*** *point along* F *if*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(F^t x) \,\mathrm{d}t = \int f \,\mathrm{d}\mu \tag{7}$$

for every $f \in \mathsf{C_c}(X)$*.*

That x is μ generic along F is equivalent to that statement that the family of of measures ξ_T given by

$$\xi_T(A) = \frac{1}{T} \int_0^T 1_A(F^t x) \,\mathrm{d}t$$

converges to μ in the weak* topology.

Exercise 6. The set of points that are μ generic for a flow is of full μ measure if μ is ergodic for the flow.

Solution. *If $f_n \to h$ in $\mathsf{C_c}(X)$ and x is a point such that (7) holds for all f_n in place of f then it also holds with h in place of f. Fix a countable, dense subset $\mathcal{F}$ of $\mathsf{C_c}(X)$. By the Birkhoff ergodic theorem the set of points for which (7) holds for all $f \in \mathcal{F}$ has full μ measure. Since $\mathcal{F}$ is dense all these points are generic.*

Exercise 7. Prove that $\lim_{t\to\infty} \mathsf{d}(\mathsf{g}^t\mathsf{v}^s x, \mathsf{g}^t x)$ for all $s \in \mathbb{R}$.

Solution. *We have $\mathsf{g}^t\mathsf{v}^s = \mathsf{v}^{se^{-2t}}\mathsf{g}^t$ for all s, t so*

$$\mathsf{d}(\mathsf{g}^t\mathsf{v}^s x, \mathsf{g}^t x) \le \mathsf{D}(\mathsf{v}^{se^{-2t}}, I)$$

which, for any fixed non-zero s, converges to zero as $t \to \infty$.

For clarity in the next proof we will consider the points that are g^a generic for μ, not N generic points.

Sublemma: If x is g^a generic for μ then $\mathsf{v}^s x$ is g^a generic for μ for all s.

Proof. From the above exercise with $t = ia$ we have

$$|f(\mathsf{g}^{ia}\mathsf{v}^s x) - f(\mathsf{g}^{ia}x)| \to 0$$

as $i \to \infty$. Therefore,

$$\frac{1}{N}\sum_{i=0}^{N-1} f(\mathsf{g}^{ia}x) \to \int f d\mu$$

implies

$$\frac{1}{N}\sum_{i=0}^{N-1} f(\mathsf{g}^{ia}\mathsf{v}^s x) \to \int f d\mu.$$

So if x is g^a generic for μ then so is $\mathsf{v}^s x$. □

Proof of Lemma. Observe that the sublemma shows that if y is g^a generic for $\mathsf{g}^\ell\mu$ then so is $\mathsf{v}^s y$. Now if x is g^a generic for μ then $\mathsf{g}^\ell x$ is g^a generic for $\mathsf{g}^\ell\mu$. So the set of points that are generic for a measure of the form $\mathsf{g}^b\mu$ is MA invariant. Because μ is N invariant, there exists an x so that $\mathsf{u}^t x$ is g^a generic for μ for Lebesgue almost every t. So there exists x and a full measure set of $\mathbb{R}$, D so that for all $t \in D$, $s, \ell \in \mathbb{R}$ we have that $\mathsf{v}^s\mathsf{g}^\ell\mathsf{u}^t x$ is g^a generic for some measure of the form $\mathsf{g}^b\mu$. The set of such $\mathsf{v}^s\mathsf{g}^\ell\mathsf{u}^t x$ has full measure for Haar measure on X. Since Haar on X is g^a ergodic, we have that one of these points is g^a-generic for Haar on X. So there exists b so that $\mathsf{g}^b\mu$ is the projection of Haar to X. This implies μ is Haar on X. □

The proposition (which generalizes from the case where μ is g^a ergodic to g^a invariant) follows by showing that if μ is N ergodic and g^a invariant then the set of points that are g^a generic for $\mathsf{g}^s\nu$ for some $s \in \mathbb{R}$ and ν in the g^a-ergodic decomposition of μ is M and A invariant. The proof of this is as above.

We now turn to proving that μ is g^a invariant for some $a \in \mathbb{R}$.

Proposition 2.7. *If μ is not supported on a single periodic N orbit then μ is A invariant as well.*

Strategy: If, for some non-zero $a \in \mathbb{R}$, the ergodic measures $\mathsf{g}^a\mu$ and μ are distinct then 99% of the points that are statistically "typical" for μ have the property that their images under g^a are typical for $\mathsf{g}^a\mu$ and are a fixed positive distance from our 99% of the μ typical points. By the Poincaré recurrence theorem we may find x and a sequence of points x_i converging to x that are all typical for μ along their respective u orbits. Since the x_i are converging to x we can write

$$x_i = \mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x$$

with $t_i \neq 0$ and with all $t_i, a_i, \ell_i \to 0$ as $i \to \infty$. For each i we consider the unstable horocycle orbits of x and x_i. Upon calculating

$$\mathsf{u}^s(\mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x) = \mathsf{v}^{\frac{t_i}{1+st_i}}\mathsf{g}^{a_i+\log(1+st_i)-a}\mathsf{u}^{e^{2a}\ell_i}\mathsf{g}^a\mathsf{u}^{\frac{s}{1+st_i}e^{-2a_i}}x \tag{8}$$

and taking i large it will be possible to find s so that the approximation

$$\mathsf{v}^{\frac{t_i}{1+st_i}}\mathsf{g}^{a_i+\log(1+st_i)-a}\mathsf{u}^{e^{2a}\ell_i}\mathsf{g}^a\mathsf{u}^{\frac{s}{1+st_i}e^{-2a_i}}x \approx \mathsf{g}^a\mathsf{u}^{\frac{s}{1+st_i}e^{-2a_i}}x$$

holds. If we can choose s to have the additional property that the points $\mathsf{u}^s(\mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x)$ and $\mathsf{u}^{\frac{s}{1+st_i}e^{-2a_i}}x$ are among the 99% of μ typical points then the latter point will be at once both $\mathsf{g}^a\mu$ typical (by equivariance) and very close to the μ typical point $\mathsf{u}^s(\mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x)$ (by our approximation). This contradicts the fixed positive distance introduced in the beginning.

In the remainder of this section we make this argument precise. First of all, what do we mean by a point that is statistically "typical" for μ? As before we will use generic points.

Lemma 2.8. *If μ is not A invariant then there exists K compact, $a > 0$ and $c > 0$ so that all of the following hold:*

(1) every $x \in K$ is generic for μ along N;
(2) $\mu(K) > \frac{99}{100}$;
(3) $\mathsf{d}(\mathsf{g}^aK, K) > c$.

Proof. Since μ is not A invariant we have $\mathsf{g}^a\mu \neq \mu$ for some non-zero $a \in \mathbb{R}$ whence $\mathsf{g}^a\mu \perp \mu$ by Corollary 2.4. By the inner regularity of Borel measures we can find a compact set K of points that are N generic for μ with $\mu(K) > \frac{99}{100}$. By equivariance, the points in $\mathsf{g}^a K$ are all N generic for $\mathsf{g}^a\mu$ so the compact set $\mathsf{g}^a K$ is disjoint from K. Because two disjoint compact sets are at a positive distance, we have the lemma. □

We now take a brief digression to prove some results on the subsets of G that do not change μ.

Lemma 2.9. *The set* $\{t \in \mathbb{R} : \mathsf{g}^a\mu = \mu\}$ *is closed in* $\mathbb{R}$ *for any Borel probability measure* μ *on* X.

Exercise 8. Prove the lemma. (Hint: Recall that the weak$*$ topology on measures is the one with respect to which $\mu \mapsto \int f \,\mathrm{d}\mu$ is continuous for every $f \in \mathsf{C_c}(X)$.)

Solution. *It suffices to prove that the map* $t \mapsto \mathsf{g}^t\mu$ *is continuous where we equip the set* $\mathcal{M}(X)$ *of Borel probability measures on* X *with the weak* topology coming from* $\mathsf{C_c}(X)$. *But for any* $f \in \mathsf{C_c}(X)$ *the map*

$$t \mapsto \int f(\mathsf{g}^t x)\,\mathrm{d}\mu(x)$$

is continuous because the action of $\mathsf{SL}(2,\mathbb{R})$ *on* X *is continuous. This is because the weak* topology asks for pointwise convergence not uniform convergence.*

Exercise 9. Strengthen Lemma 2.9 to if μ is a N invariant probability measure on X then $\{g \in G : g\mu = \mu\}$ is closed.

Solution. *Idea: for all* $f \in C_c(X)$ *and* $h \in \overline{\{g \in G : g\mu = \mu\}}$ *we have that* $\int f \circ g d\mu = \int f d\mu$.

Exercise 10. Show that if there exists $a \neq 0 \neq b$ so that $F^a\mu = \mu$, $F^b\mu = \mu$ and $\frac{a}{b} \notin \mathbb{Q}$ then μ is F^t invariant.

We now state our goal and an issue with it: $\mathsf{u}^{\tilde{s}}x \in K$ and $\mathsf{u}^s\mathsf{v}^{t_i}\mathsf{g}^{a_i}x \in K$ and the distance between $\mathsf{g}^a\mathsf{u}^{\tilde{s}}x$ (which is in $\mathsf{g}^a K$) and $\mathsf{u}^s\mathsf{v}^{t_i}\mathsf{g}^{a_i}x$ is less than $\frac{c}{2}$.

Issue: How can we be sure these points will be in K? At the very least there probably is some of the measure of μ that is not in K.

The next lemma addresses this issue.

Lemma 2.10. *For all $\epsilon > 0$ there is $T_0 > 0$ and $E \subset X$ with $\mu(E) > 1 - \epsilon$ so that*
$$\left| \frac{1}{sT} \int_T^{T+sT} \chi_K(\mathsf{u}^s x)\,\mathrm{d}s - \mu(K) \right| < \epsilon$$
for all $T > T_0$ and $s \geq \epsilon$.

Exercise 11. Prove the lemma.

Hint: Show that the Birkhoff ergodic theorem formally implies that if F^t is ν ergodic then for any $f \in L^1(\nu)$ and $\epsilon > 0$ we have
$$\lim_{T \to \infty} \frac{1}{\epsilon T} \int_T^{T+\epsilon T} f(F^t x) dt = \int f d\nu$$
for μ a.e. x.

Corollary 2.11. If μ is not supported on a N periodic orbit then there exists $p_i \to x$ with $p_i, x \in E \cap K$ with $p_i = \mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x$ where $t_i, a_i, \ell_i \to 0$ and at least one of $t_i, a_i \neq 0$ for all i. Moreover, if μ is not A invariant then p_i we may assume $t_i \neq 0$ for all i.

Proof. The first condition follows from Poincaré recurrence because $E \cap K$ has positive measure. Indeed, we may choose $s_i \to \infty$ so that $\mathsf{u}^{s_i}x \to x$ and $\mathsf{u}^{s_i}x \in E \cap K$ for all i. (See the exercise below.) Because $\mathsf{u}^{s_i}x \to x$ we have that $\mathsf{u}^{s_i}x = \mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x$ with $t_i, a_i, \ell_i \to 0$. If x is not N periodic then one of $t_i, a_i \neq 0$ for all i so that s_i is large enough. This implies the first sentence of the corollary. For the second sentence, if $x, \mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x \in K$ then $\mathsf{g}^{a_i}\mu = \mu$. So there exists $a_i \to 0$ so that $\mathsf{g}^{a_i}\mu = \mu$. So $\{a : \mathsf{g}^a\mu = \mu\}$ is ϵ-dense in $\mathbb{R}$ for all ϵ and so it is all of $\mathbb{R}$. This contradicts our assumption that μ is not A invariant. □

Exercise 12. Let Z be a metric space, let ν be a Borel probability measure on Z, and let $T : Z \to Z$ be Borel measurable and measure-preserving. Use the Poincaré recurrence theorem to show that for any $A \in \mathcal{B}$ with $\nu(A) > 0$ for almost every $z \in A$ there exists $n_i \to \infty$ so that $T^{n_i}z \to z$ and $T^{n_i}z \in A$ for all i.

Solution. *Let $B_1, \ldots$ be a countable set of metric balls in Z so that for all $\epsilon > 0$ the union of the B_i with radius at most ϵ has full measure. Let*
$$E_i = \{z \in B_i \cap A : \exists n > 0 \text{ so that } T^n z \in B_i \cap A\} \cup (B_i \cap A)^c$$
and by the Poincaré recurrence theorem, for every i we have $\nu(E_i) = 1$. If there does not exist $n_1, \ldots$ so that $T^{n_j}z \to z$ and $T^{n_j}z \in A$ for all j and z is in the full measure set so that for every $\epsilon > 0$ there exists i

with $z \in B_i$ and the radius of B_i is at most ϵ then $z \notin \cap_i E_i$. Because $\nu(\cap_i E_i) = 1$ the exercise follows.

Proof of Proposition 2.7. We proceed by contradiction, assuming that μ is not A invariant. Let K be the compact set given by Lemma 2.8 and $1 > c > 0$, $a > 0$ the corresponding constants. Fix $\epsilon < \min\{\frac{c}{4}, \frac{1}{9}, \frac{\frac{c}{4}}{2e^{2a}}\}$. Let $T_0 > 0$ and $E \subset X$ be as in Lemma 2.8.

Since μ is not supported on a single N orbit, the Poincaré recurrence theorem implies that for μ almost every $x \in E$ there is a sequence $p_i \to \infty$ in $\mathbb{R}$ with $\mathsf{u}^{p_i}x \to x$ and $\mathsf{u}^{p_i}x \in E$ for all $i \in \mathbb{N}$.

We can write $\mathsf{u}^{p_i}x = \mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x$ so that $(t_i, a_i) \neq (0,0)$ for all i large enough and $(t_i, a_i, \ell_i) \to (0,0,0)$. By Corollary 2.11 we may further assume that for all large enough i, $t_i \neq 0$. For simplicity we assume $t_i > 0$ and $a_i > 0$. Now choose i so large that

- $|t_i|e^{a_i - a} + e^{2a}|\ell_i| < \frac{c}{4}$.
- $|t_i| < \frac{c}{8}\frac{1}{T_0}$.
- $\tau := \frac{e^{a-a_i}-1}{t_i} > T_0$.
- $\frac{\frac{c}{4}\tau}{(1+\tau t_i)(1+(1+\frac{c}{4})\tau t_i} > \frac{\frac{c}{4}\tau}{2e^{2a}} > \epsilon\frac{\tau}{1+\tau t_i}$
- $\frac{\tau}{1+\tau t_i}e^{-2a} \geq \frac{\tau}{e^a}e^{-2a} > T_0$.

Performing some arithmetic, we see that

$$\mathsf{u}^s(\mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x) = \mathsf{v}^{\frac{t_i}{1+st_i}}\mathsf{g}^{a_i+\log(1+st_i)-a}\mathsf{u}^{e^{2a}\ell_i}\mathsf{g}^a\mathsf{u}^{\frac{s}{1+st_i}e^{-2a_i}}x$$

as in (8). We estimate that

$$\begin{aligned} &\mathsf{d}(\mathsf{u}^s\mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x, \mathsf{g}^a\mathsf{u}^{\frac{s}{1+st_i}e^{-2a_i}}x) \\ &\leq \mathsf{D}(\mathsf{v}^{\frac{t_i}{1+st_i}}\mathsf{g}^{a_i+\log(1+st_i)-a}\mathsf{u}^{e^{2a}\ell_i}, I) \\ &\leq \mathsf{D}(\mathsf{v}^{\frac{t_i}{1+st_i}}, I) + \mathsf{D}(\mathsf{g}^{a_i+\log(1+st_i)-a}, I) + \mathsf{D}(\mathsf{u}^{e^{2a}\ell_i}, I) \\ &\leq \left|\frac{t_i}{1+st_i}\right| + |a_i + \log(1+st_i) - a| + e^{2a}|\ell_i| \end{aligned} \tag{9}$$

using (1) and (2).

If $s \in [\tau, \tau(1+\frac{c}{4})]$ then

$$e^{a-a_i} \leq 1 + st_i \leq e^{a-a_i}(1+\tfrac{c}{4}) - \tfrac{c}{4}$$

whence $a - a_i \leq \log(1+st_i) \leq a - a_i + \frac{c}{4}$ and $\frac{1}{1+st_i} \leq e^{a_i-a}$. We therefore have

$$\mathsf{d}(\mathsf{u}^s\mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x, \mathsf{g}^a\mathsf{u}^{\frac{s}{1+st_i}e^{-2a_i}}x) \leq \tfrac{c}{2}$$

from our choices together with (9).

We next claim our choices imply there exists $r \in [\tau, \tau(1+\frac{c}{4})]$ so that $\mathsf{u}^r(\mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x)$ and $\mathsf{u}^{\frac{r}{1+rt_i}e^{-2a_i}}x$ both belong to K. Indeed, first note that

$$|\{s \in [\tau, \tau(1+\tfrac{c}{4})] : \mathsf{u}^s(\mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i})x \in K\}| > \tfrac{8}{9}\tau\tfrac{c}{4} \tag{10}$$

because $\mathsf{v}^{t_i}\mathsf{g}^{a_i}\mathsf{u}^{\ell_i}x \in E$. Similarly, $\frac{\tau}{1+\tau t_i}e^{-2a_i} > T_0$ and $\frac{(1+\frac{c}{4})\tau}{1+(1+\frac{c}{4})\tau}e^{-2a_i} - \frac{\tau}{1+\tau t_i}e^{-2a_i} > \epsilon\frac{\tau}{1+\tau t_i}e^{-2a_i}$ gives

$$|\{s \in [\tau, \tau(1+\frac{c}{4})] : \mathsf{u}^{\frac{s}{1+st_i}}x \in K\}| > \frac{5}{9}\tau\frac{c}{4}. \tag{11}$$

To see this, let $f(s) = \frac{s}{1+st_i}e^{-2a_i}$ and because

$$|\{s \in [f(\tau), f(\tau(1+\frac{c}{4})] : \mathsf{u}^s x \notin K\}| < \epsilon(f(\tau(1+\frac{c}{4})) - f(\tau))$$

we have

$$|\{s \in [\tau, \tau(1+\frac{c}{4})] : \mathsf{u}^{f(s)}x \notin K\}| < \frac{\max\{f'(s) : s \in [\tau, \tau(1+\frac{c}{4})]\}}{\min\{f'(s) : s \in [\tau, \tau(1+\frac{c}{4})]\}}\epsilon\tau\frac{c}{4}.$$

Now

$$\frac{\max\{f'(s) : s \in [\tau, \tau(1+\frac{c}{4})]\}}{\min\{f'(s) : s \in [\tau, \tau(1+\frac{c}{4})]\}} = \frac{\frac{1}{1+\tau t_i}}{\frac{1}{1+(1+\frac{c}{4})\tau}} < 4$$

implying (11).

So $\mathsf{g}^a\mathsf{u}^r x \in \mathsf{g}^a K$. But $r \in [\tau, \tau(1+\frac{c}{4})]$ so $\mathsf{d}(\mathsf{g}^a K, K) < \frac{c}{2}$ by (9). This contradicts the properties of K in Lemma 2.8. □

Proof of Theorem 2.1. If μ is not supported on a periodic orbit then by Proposition 2.7 it is A invariant. By Proposition 2.2 this implies that it is Haar measure on X. □

2.4. **Concluding remarks.**

Remark on the periodic case. In the case where μ is supported on a single N orbit, there exists $x \in X$, $s \in \mathbb{R}$ so that $\mathsf{u}^s x = x$ and

$$\mathsf{supp}(\mu) = \{\mathsf{u}^t x : t \in \mathbb{R}\} = \{\mathsf{u}^t x : t \in [0, s)\}$$

and $\mu(A) = |\{0 \leq \ell < s : \mathsf{u}^\ell x \in A\}|$. We have that x corresponds to a lattice with a horizontal vector (which is fixed by u^t). Any such lattice will give a periodic u^t invariant measure. Indeed, if w and h are two vectors that generate the lattice, and h is horizontal then there exists s so that $\mathsf{u}^s w = w + h$. However, $\mathsf{u}^s h = h$ so that lattice generated by $\mathsf{u}^s w$ and $\mathsf{u}^s h$ is the same as the lattice generated by w and h.

Exercise 13. Show that if $\mathsf{u}^b x = x$ for some $b \neq 0$ then the lattice corresponding to x has a horizontal vector.

Solution. *Let $\{u, v\}$ be a generating set for the lattice corresponding to x. Since $\mathsf{u}^b x = x$ we have that $\mathsf{u}^b u$ and $\mathsf{u}^b v$ generate the lattice corresponding to x as well. So there exist $c, d \in \mathbb{Z}$ relatively prime with $cu_2 + dv_2 = u_2$. Thus $(c-1)u + dv$ is a horizontal vector in the lattice corresponding to x.*

Remark on the aperiodic case. On the topological side the orbit closure $\overline{\{\mathsf{u}^s x : s \in \mathbb{R}\}}$ is either a single periodic N orbit or is all of X. On the measure side more is true than what we proved: either x corresponds to a lattice with a horizontal vector, or x equidistributes according to Haar measure on X.

Remarks on the proof. Lemma 2.3 only uses that A normalizes N. Also the first paragraph of Corollary 2.4 is an important fact in ergodic theory (and our proof). Furthermore, in the proof of Proposition 2.7 we were aided in getting invariance under A by the fact that $\mathsf{g}^a\mu$ is ergodic. If this were not the case, there would be no reason that a point in $\mathsf{g}^a K$ would need to be generic. Generic points are useful for our arguments, because two measures sharing a generic point must be the same.

3. $\mathsf{SL}(2,\mathbb{R}) \ltimes \mathbb{R}^2/\mathsf{SL}(2,\mathbb{Z}) \ltimes \mathbb{Z}^2$

3.1. **Setup.** Let $\tilde{G} = \mathsf{SL}(2,\mathbb{R}) \ltimes \mathbb{R}^2$. Thus $\tilde{G}$ is the semi-direct product of $\mathsf{SL}(2,\mathbb{R})$ and the additive group $\mathbb{R}^2$ under the operation

$$(A, v) \cdot (B, w) = (AB, v + Aw)$$

for all $A, B \in \mathsf{SL}(2,\mathbb{R})$ and all $v, w \in \mathbb{R}^2$. We can identify $\tilde{G}$ with the subgroup

$$\left\{ \begin{bmatrix} A & v \\ 0 & 1 \end{bmatrix} \in \mathsf{SL}(3,\mathbb{R}) : A \in \mathsf{SL}(2,\mathbb{R}), v \in \mathbb{R}^2 \right\}$$

of $\mathsf{SL}(3,\mathbb{R})$ if we think of elements of $\mathbb{R}^2$ as column vectors.

Let $\tilde{\Gamma} = \mathsf{SL}(2,\mathbb{Z}) \ltimes \mathbb{Z}^2$ and $\tilde{X} = \tilde{G}/\tilde{\Gamma}$. We can identify $\tilde{X}$ with the space of **affine lattices** in $\mathbb{R}^2$ i.e. as the space of all cosets of all unimodular lattices in $\mathbb{R}^2$. It can also be thought of as the moduli space of tori with two marked points and a preferencial direction, since homogeneity allows us to move any one marked point to the origin. In this guise, it can be used to help understand 3-IETs or the space of 2 identitical tori with glued along a slit.

Let

$$\tilde{\mathsf{u}}^t = \left(\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right)$$

for all $t \in \mathbb{R}$. The set $\tilde{\mathsf{N}} = \{\tilde{\mathsf{u}}^t : t \in \mathbb{R}\}$ is a one-parameter subgroup of $\tilde{G}$. As in the previous section, here we are interested in classifying the $\tilde{\mathsf{N}}$ invariant probability measures on $\tilde{X}$. In doing so we will make use of the following one-parameter subgroups

$$\tilde{\mathsf{g}}^t = \left(\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) \quad \tilde{\mathsf{v}}^t = \left(\begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right)$$

of G comparable to those in Section 2 and the one-parameter subgroups

$$\mathsf{h}^t = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} t \\ 0 \end{pmatrix} \right) \qquad \mathsf{w}^t = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ t \end{pmatrix} \right)$$

corresponding to flows in the $\mathbb{R}^2$ direction. Let $\tilde{\mathsf{A}}$, $\tilde{\mathsf{M}}$, H and W be the one-parameter subgroups defined by these parameterizations respectively.

As in Section 2 write $X = \mathsf{SL}(2,\mathbb{R})/\mathsf{SL}(2,\mathbb{Z})$ and write $\pi : \tilde{X} \to X$ for the projection on the first coordinate. We will classify $\tilde{\mathsf{N}}$ ergodic and invariant probability measures on $\tilde{X}$ and our proof will use the description of $\tilde{X}$ as a fiber bundle over X with torus fibers.

Exercise 14. Verify that the fiber of π over a lattice Λ in X can be identified with the quotient $\mathbb{R}^2/\Lambda$.

Solution. *The map $(A, v)\tilde{\Gamma} \mapsto \Lambda + v$ is a well-defined isomorphism from the fiber $\pi^{-1}\Lambda$ to $\mathbb{R}^2/\Lambda$.*

If μ is a $\tilde{\mathsf{N}}$ invariant probability on $\tilde{X}$ then the push-forward $\pi\mu$ is a probability measure on X that is invariant under the one-parameter subgroup N from Section 2. From Theorem 2.1 we know that $\pi\mu$ is either Haar measure on X or supported on a periodic N orbit.

Theorem 3.1. *If μ is an $\tilde{\mathsf{N}}$ ergodic probability measure then there exists a closed subgroup L of G containing $\tilde{\mathsf{N}}$, $x \in \tilde{G}$ such that $L \cap x\tilde{\Gamma}x^{-1}$ is a lattice in L and μ is supported on an L orbit and μ is a scaled projection of Haar measure on L.*

These possibilities are:

- $\pi\mu$ is Haar on X and μ is Haar on $\tilde{X}$.
- $\pi\mu$ is Haar on X. There exists $x \in \tilde{X}$ and $\mathsf{h}_a \in \mathsf{H}$ so that μ is supported on $\mathsf{h}^{-s}\mathsf{SL}(2,\mathbb{R})\mathsf{h}^s x$.
- $\pi\mu$ is supported on a periodic N orbit, μ is supported on an orbit of the subgroup generated by $\tilde{\mathsf{N}}$ and H, and μ is the projection of the Haar measure on this subgroup to the orbit.

- $\pi\mu$ is supported on a periodic N orbit and μ is supported on a periodic $\tilde{\mathsf{N}}$ orbit.

Remark 3.2. Note that $\mathsf{h}^{-s}\mathsf{SL}(2,\mathbb{R})\mathsf{h}^{s}$ is a closed subgroup of $\tilde{G}$ that contains $\tilde{\mathsf{N}}$ and that it has a conjugate (probably different from itself) that meets $\tilde{\Gamma}$ in a lattice.

3.2. **Preliminaries.**

Proposition 3.3. *If $\pi\mu$ is supported on a closed horocycle then μ is the projection of Haar measure on a closed subgroup of $\tilde{G}$ to $\tilde{X}$.*

Lemma 3.4. *It suffices to treat the case where $\pi\mu$ is supported on $\mathsf{N}\Gamma$.*

Exercise 15. Prove this.

Solution. *A lattice in X has a periodic horocycle orbit if and only if it has a horizontal vector. So we may assume that $\pi\mu$ is supported on the set of lattices with a horizontal vector of a given length, say e^L. If Λ has a horizontal vector of length e^L then $\mathsf{g}^{-L}\Lambda$ has a horizontal vector of length 1 and $\mathsf{g}^{-L}\pi\mu = \pi\tilde{\mathsf{g}}^{-L}\mu$ is supported on the set of lattices having a horizontal vector of length 1. If the proposition is true for this measure, applying $\tilde{\mathsf{g}}^L$ to $\tilde{\mathsf{g}}^{-L}\mu$ gives the result for our original measure.*

Proof of Proposition 3.3. By the previous lemma we may assume that $\pi\mu$ is supported on $\mathsf{N}\Gamma$. We will prove that there are two possibilities.

(1) There exists an irrational y_0 so that the support of μ is

$$\left\{\left(\begin{pmatrix}1 & s\\ 0 & 1\end{pmatrix}, \begin{pmatrix}x\\ y_0\end{pmatrix}\right)\tilde{\Gamma} : s, x \in [0,1)\right\}$$

in which case μ is the projection to $\tilde{X}$ of the Haar measure on $L = \langle \tilde{\mathsf{N}}, \mathsf{H}\rangle$. Note that

$$L \cap \tilde{\Gamma} = \left\{\left(\begin{pmatrix}1 & n\\ 0 & 1\end{pmatrix}, \begin{pmatrix}a\\ 0\end{pmatrix}\right) : n, a \in \mathbb{Z}\right\}$$

is a lattice in L.

(2) There exists $y_0 \in [0,1)$ rational and $p_1, ..., p_k \in [0,1)$ so that μ is supported on

$$\left\{\left(\begin{pmatrix}1 & s\\ 0 & 1\end{pmatrix}, \begin{pmatrix}p_i\\ y_0\end{pmatrix}\right)\tilde{\Gamma} : s \in [0,1), 1 \le i \le k\right\}$$

in which case μ is supported on a periodic $\tilde{\mathsf{N}}$ orbit.

Every generic point for μ has the form

$$\left(\begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix} \right) \tilde{\Gamma}$$

for some $s \in \mathbb{R}$ and some $w \in \mathbb{R}^2$ because $\pi\mu$ lives on $\mathsf{N}\Gamma$. Since the $\tilde{\mathsf{N}}$ orbit of a point generic for μ along $\tilde{\mathsf{N}}$ consists entirely of points generic for μ along $\tilde{\mathsf{N}}$, we can find a point of the form

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right) \tilde{\Gamma}$$

that is generic for μ along $\tilde{\mathsf{N}}$. Now

$$\tilde{\mathsf{u}}^1 \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right) = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} x + y \\ y \end{pmatrix} \right)$$

so if $y \notin \mathbb{Q}$ then

$$\tilde{\mathsf{u}}^n \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right) = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} x + ny \\ y \end{pmatrix} \right)$$

equidistributes along

$$\left\{ \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} z \\ y \end{pmatrix} \right) : z \in [0,1) \right\}$$

and we have the first case, where as if y is rational then

$$\left\{ \tilde{\mathsf{u}}^n \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right) : n \in \mathbb{Z} \right\}$$

is a finite number of points in $\tilde{X}$ and we have the second case. □

The next lemma implies that $\mathsf{h}^s\mu$ is $\tilde{\mathsf{N}}$ ergodic for every $s \in \mathbb{R}$, a fact that is used freely in the next section.

Lemma 3.5. *If F is an ergodic and measure-preserving flow on $(X, \mathcal{B}, \mu)$ and $S : X \to X$ commutes with F^t for all $t \in \mathbb{R}$ then F is also ergodic and measure-preserving on $(X, \mathcal{B}, S\mu)$.*

Exercise 16. Prove this.

Solution. *This is a consequence of symbol pushing. First, if we have $(S\mu)(F^tA \triangle A) > 0$ then $\mu(S^{-1}(F^tA) \triangle S^{-1}(A)) > 0$. But since S commutes with F^t this is $\mu(F^t(S^{-1}A) \triangle (S^{-1}A))$. So the ergodicity of μ and $S\mu$ are equivalent. Similarly*

$$(S\mu)(F^tA) = \mu(S^{-1}F^tA) = \mu(F^tS^{-1}A) = \mu(S^{-1}A) = (S\mu)(A)$$

proves that $S\mu$ is F invariant.

3.3. **A key step.** It remains to prove Theorem 3.1 when $\pi\mu$ is Haar measure on X. A key technical step in this part of the argument is to be able to write μ as an integral of probability measures on the fibers of π. Theorem A.8 in the appendix makes precise the sense in which this can be done.

Fix a disintegration P of μ coming from the map π. We get, for almost every $\Lambda \in X$ a probability measure $P(\Lambda) = \mu_\Lambda$ on $\pi^{-1}(\Lambda)$ which we identify with a measure on $\mathbb{R}^2/\Lambda$. For each $t \in \mathbb{R}$ we can consider the almost-surely defined maps $P'(\Lambda) = \tilde{\mathsf{u}}^t\mu_\Lambda$ and $P''(\Lambda) = \mu_{\mathsf{u}^t\Lambda}$. They are both disintegrations of μ so must be equal $\pi\mu$ almost-surely. We therefor have, for each $t \in \mathbb{R}$ the equivariance

$$\tilde{\mathsf{u}}^t\mu_\Lambda = \mu_{\mathsf{u}^t\Lambda}$$

for $\pi\mu$ almost every Λ.

The main result in this section is the following lemma.

Lemma 3.6. *Either μ_Λ has finite support for $\pi\mu$ almost every Λ or it has uncountable support for $\pi\mu$ almost every Λ. If the support is finite, then the cardinality is $\pi\mu$ almost everywhere constant. If the support is infinite (and thus uncountable) then μ_Λ has no atoms $\pi\mu$ almost everywhere.*

Proof. First observe that $\pi\mu$ is N ergodic. The cardinality of the support of μ_Λ is a N invariant quantity. It is a Borel measurable quantity because the map on probability measures endowed with the weak* topology to itself given by sending a measure to its atomic part is Borel measurable. (Note that this is a non-trivial but not too hard exercise to prove directly.) Ergodicity implies that the cardinality of the support of μ_Λ is constant $\pi\mu$ almost everywhere.

We now study the atoms of μ_Λ. Let $Q_s \subset \tilde{X}$ be the set of points (Λ, w) so that $\mu_\Lambda(\{(\Lambda, w)\}) < (\tilde{\mathsf{u}}^s\mu_\Lambda)(\{(\Lambda, w)\})$. Note that $\mu(Q_s) < \mu(\tilde{\mathsf{u}}^s Q_s)$ unless $\mu(Q_s) = 0$. Similarly if Q'_s is the set of points (Λ, w) so that $\mu_\Lambda(\{(\Lambda, w)\}) > (\tilde{\mathsf{u}}^s\mu_\Lambda)(\{(\Lambda, w)\})$ then $\mu(Q'_s) > \mu(\tilde{\mathsf{u}}^s Q'_s)$ unless $\mu(Q'_s) = 0$. Because μ is $\tilde{\mathsf{N}}$ invariant this implies that both $\mu(Q_s)$ and $\mu(Q'_s)$ are 0. So $f(\Lambda, w) = \mu_\Lambda(\{(\Lambda, w)\})$ is almost everywhere $\tilde{\mathsf{N}}$ invariant and so it is constant μ almost everywhere. That is, if a μ positive measure subset of points (Λ, w) are atoms of μ_Λ then the size is the same for μ almost every point. The reciprocal of this size is the cardinality of the support of μ_Λ for $\pi\mu$ almost every Λ. Since this is a finite number, if there are atoms, the cardinality of the support of μ_Λ is finite $\pi\mu$ almost surely. Similarly if the support of μ_Λ is infinite, there can not be atoms, so it must be uncountable. □

Exercise 17. Show that if (Y, d) is a σ-compact metric space and ϕ maps Borel probability measures to their atomic part, it is a Borel map from Borel measures with total variation at most 1 with the weak* topology to Borel measures with total variation at most 1 with the weak* topology.

Solution. *Let $\mathcal{P}_1, \ldots$ be a sequence of countable, measurable partitions of Y so that $\lim_{n\to\infty} \max\{diam(P) : P \in \mathcal{P}_n\} = 0$.*

Fix $\epsilon > 0$ and $n \in \mathbb{N}$. Write $Q_1, Q_2, \ldots$ for the members of $\mathcal{P}_n$. For each ν the set $K(\nu) = \{k \in \mathbb{N} : \nu(Q_k) \geq \epsilon\}$ is finite or empty and $Q(E) = \cup_{k\in E} P_k$. Let $\nu_{n,\epsilon}$ be $\nu|_{Q(E)}$ and we wish to prove that $\Psi(\nu) = \nu_{n,\epsilon}$ is a Borel map. Fixing $E \subset \mathbb{N}$ finite, it suffices to prove for every continuous function $f : Y \to \mathbb{R}$ of compact support that

$$\nu \mapsto \int f\, 1_{Q(E)}\, \mathrm{d}\nu \tag{12}$$

is measurable on the set $\{\nu : K(\nu) = E\}$. Indeed, (12) *shows if U is an open in the set of measures with total variation at most 1 with the weak* topology then $\Psi^{-1}(U)$ is measurable.*

To prove that, first note that

$$\mu \mapsto \int f 1_A\, \mathrm{d}\nu \tag{13}$$

is measurable whenever A is open and has compact closure by approximating 1_A by continuous function. Then note that the collection of sets A for which (13) *is measurable forms a σ algebra.*

Let ν_ϵ be a weak limit of $\nu_{n,\epsilon}$ as n goes to infinity and observe that it is ν restricted to its atoms of measure at least ϵ (and so the limit actually exists). This is measurable because it is a pointwise limit of measurable maps. Now let ν_{atom} be a weak* limit of $\nu_{\frac{1}{k}}$ as k goes to infinity. This is ν restricted to its atoms (and so the limit exists and the map is measurable).*

We now have a dichotomy when $\pi\mu$ is Haar measure on X. Either the disintegration of μ over π is atomic almost surely, or has uncountable support and no atoms almost surely. In the next two sections we describe what μ is in both possibilities.

3.4. **Uncountable fibers.** In this section we assume $\pi\mu$ is the projection of Haar measure on G to X, and that almost every measure μ_Λ in the disintegration of μ over π has uncountable support. Our goal is to prove that μ is Haar measure on $\tilde{X}$.

Proposition 3.7. *If μ is a $\tilde{\mathsf{N}}$ ergodic and $\tilde{\mathsf{N}}$ invariant measure on $\tilde{X}$ so that $\pi\mu$ is the projection of Haar measure on G to X, and if μ_Λ has uncountable support for $\pi\mu$ almost every Λ then μ is Haar measure on $\tilde{X}$.*

First, we observe that it suffices to prove μ is H invariant. We will then argue as in Section 2 that this is the case.

Lemma 3.8. *If $\pi\mu$ is the projection of Haar on G to X and μ is H invariant then μ is the projection of Haar on $\tilde{G}$ to $\tilde{X}$.*

Proof. Consider the H ergodic decomposition of μ. The subgroup H preserves the fibers of π and every fiber $\pi^{-1}(\Lambda)$ of π can be identified with the quotient $\mathbb{R}^2/\Lambda$ so almost every measure in the ergodic decomposition is supported on some quotient $\mathbb{R}^2/\Lambda$. For $\pi\mu$ almost every Λ in X any H ergodic measure on $\mathbb{R}^2/\Lambda$ is Lebesgue measure. This is because, on almost every lattice quotient, every point in generic for Lebesgue measure along H. Since Haar measure on $\tilde{X}$ disintegrates over $\pi\mu$ as Lebesgue measure on almost every fiber of π the lemma follows from uniqueness of disintegrations. $\square$

Note: the same proof gives the following: if $\pi\mu$ is the projection of Haar on G to X and there exists $a \neq 0$ so that $\mathsf{h}^a\mu = \mu$ then μ is the projection of Haar on G to X.

To conclude μ is the projection of Haar measure on $\tilde{G}$ to $\tilde{X}$ it remains to prove that μ is H invariant.

Strategy: Suppose μ is not H invariant. As in Lemma 2.8 we can find $K \subset X$ compact consisting of points that are generic for μ along $\tilde{\mathsf{N}}$, and $a > 0$ and $c > 0$ with $\mu(K) > \frac{99}{100}$ and $\mathsf{d}(\mathsf{h}^a K, K) > c$. Our goal is to find nearby points whose $\tilde{\mathsf{N}}$ orbits will diverge in the direction of H. We cannot appeal to Poincaré recurrence to choose nearby points as in Section 2 because we wish our nearby points to be on the same fiber of π. (If they are not on the same fiber then there will be some divergence in X in the A direction, which we don't need at the moment.) Now, if points of the form $(\Lambda, \left(\begin{smallmatrix} x_i \\ y_i \end{smallmatrix}\right))$ are converging to $(\Lambda, \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right))$ then, upon choosing i large enough and t appropriately, we can arrange for

$$\tilde{\mathsf{u}}^t \left(\Lambda, \begin{pmatrix} x \\ y \end{pmatrix}\right) = \mathsf{h}^{x-x_i}\mathsf{w}^{y-y_i}\mathsf{h}^{t(y-y_i)}\tilde{\mathsf{u}}^t \left(\Lambda, \begin{pmatrix} x_i \\ y_i \end{pmatrix}\right) \approx \mathsf{h}^a \tilde{\mathsf{u}}^t \left(\Lambda, \begin{pmatrix} x_i \\ y_i \end{pmatrix}\right)$$

with $\tilde{\mathsf{u}}^t(\Lambda, \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right))$ and $\tilde{\mathsf{u}}^t(\Lambda, \left(\begin{smallmatrix} x_i \\ y_i \end{smallmatrix}\right))$ both in K contradicting separation of K and $\mathsf{h}^a K$.

Lemma 3.9. *If μ_Λ has uncountable support for a set of Λ with $\pi\mu$ of measure at least $\frac{1}{2}$ then we have for a set of $(\Lambda, \binom{x}{y})$ of μ measure $\frac{1}{2}$ that there exists $(\Lambda, \binom{x_i}{y_i}) \to (\Lambda, \binom{x}{y})$ so that $y_i \neq y$ for all i and $(\Lambda, \binom{x_i}{y_i}) \in K$ for all i.*

Sublemma 1. *We may assume there exists $s_0 > 0$ so that $K \cap \mathsf{h}^s K = \emptyset$ for all $0 < |s| < s_0$.*

Proof. If $x \in \mathsf{h}^s K$ then x is $\tilde{\mathsf{u}}^t$ generic for $\mathsf{h}^s\mu$. So if $x \in K \cap \mathsf{h}^s K$ then it is generic for $\mathsf{h}^s\mu$ and μ. So these are the same measures. So if this is not true there exists $s_i \neq 0$, but $s_i \to 0$ so that $\mathsf{h}^{s_i}\mu = \mu$. Since $\{s : \mathsf{h}^s_*\mu = \mu\}$ is a closed subgroup of $\mathbb{R}$, this implies μ is h^s invariant, which by Lemma 3.8 completes the proof of the proposition. □

Proof of Lemma 3.9. We prove this by contradiction. Observe that we may assume the support of μ_Λ is contained in an at most countable set of horizontal lines for a set of Λ with $\pi\mu$ measure $\frac{1}{2}$ as otherwise we could find points as requested by the conclusion of Lemma 3.9. This implies that K intersected with some horizontal line segment has positive measure. Next, observe that since the support of μ_Λ is uncountable, it is non atomic and so K intersected with this horizontal line segment is uncountable. So it contains an accumulation point, contradicting the sublemma. □

We are now ready for the proof of Proposition 3.7.

Proof of Proposition 3.7. As in Lemma 2.10 there exists $T_0 > \max\{a, 1\}$ so that the μ measure of

$$E = \left\{ (\Lambda, \tbinom{x}{y}) : \frac{9}{cT} \int_T^{T+\frac{cT}{9}} 1_K(\tilde{\mathsf{u}}^t(\Lambda, \tbinom{x}{y}))\, \mathrm{d}t > \frac{98}{100} \text{ for all } T > T_0 \right\}$$

is at least $\frac{99}{100}$.

By Lemma 3.9 we may choose $(\Lambda, \binom{x_i}{y_i}) \to (\Lambda, \binom{x}{y})$ so that

- $y_i \neq y$ for all i;
- $(\Lambda, \binom{x_i}{y_i}) \in E$ for all i;
- $(\Lambda, \binom{x}{y}) \in E$;

all hold. For simplicity we assume $y > y_i$ for all $i \in \mathbb{N}$. In place of (8) we have

$$\tilde{\mathsf{u}}^t \left(\Lambda, \begin{pmatrix} x \\ y \end{pmatrix} \right) = \mathsf{h}^{x-x_i} \mathsf{w}^{y-y_i} \mathsf{h}^{t(y-y_i)} \tilde{\mathsf{u}}^t \left(\Lambda, \begin{pmatrix} x_i \\ y_i \end{pmatrix} \right) \tag{14}$$

for all $t \in \mathbb{R}$.

Sublemma 2. *If j is such that*

- $|y_j - y| < \frac{ac}{10T_0}$
- $|x_j - x| < \frac{c}{8}$
- $(\Lambda, \binom{x_j}{y_j}) \in E$
- $(\Lambda, \binom{x}{y}) \in E$

all hold then there exists t so that

(a) $\tilde{\mathsf{u}}^t(\Lambda, \binom{x_j}{y_j}) \in K$
(b) $\tilde{\mathsf{u}}^t(\Lambda, \binom{x}{y}) \in K$
(c) $\mathsf{d}(\tilde{\mathsf{u}}^t(\Lambda, \binom{x}{y}), \mathsf{h}^a \tilde{\mathsf{u}}^t(\Lambda, \binom{x_j}{y_j}) < c$

all hold.

Exercise 18. Prove the sublemma.

Solution. *Consider the interval $I = [\frac{a}{y_j - y}, \frac{a}{y_j - y} + \frac{c}{2(y_j - y)}]$. We claim there exists $t \in I$ satisfying the conclusions (a), (b) and (c). First observe that by* (14) *and our choice of j, any $t \in I$ will satisfy (c). Second, because $(\Lambda, \binom{x_j}{y_j})$ and $(\Lambda, \binom{x}{y})$ are both in E we have that*

$$|\{t \in I : \tilde{\mathsf{u}}^t(\Lambda, \binom{x_j}{y_j}) \in K\}| > \frac{99}{100}|I|$$

and

$$|\{t \in I : \tilde{\mathsf{u}}^t(\Lambda, \binom{x_j}{y_j}) \in K\}| > \frac{99}{100}|I|$$

both hold. Thus the set of $t \in I$ satisfying all three conclusions has measure at least $\frac{98}{100}|I|$.

Condition (a), (b) and (c) together contradict the separation of K and $\mathsf{h}^a K$ concluding the proof of Proposition 3.7. □

3.5. **Atomic fibers.** In this section we conclude the proof of Theorem 3.1 by proving the following result.

Proposition 3.10. *If $\pi\mu$ is Haar measure on X but μ is not Haar measure on $\tilde{X}$ then μ is supported on a single $\mathsf{h}^{-s}\mathsf{SL}(2,\mathbb{R})\mathsf{h}^s$ orbit.*

Lemma 3.11. *It suffices to show that μ is invariant under a one parameter subgroup of $\mathsf{H}\tilde{\mathsf{A}}$ other than H.*

This result is involved and we break the proof into three steps.

Step 1. If μ is invariant under a one parameter subgroup of $\mathsf{H}\tilde{\mathsf{A}}$ other than H then there exists $s \in \mathbb{R}$ so that $\mathsf{h}^s\mu$ is invariant under $\tilde{\mathsf{A}}$.

Sublemma 3. *Any 1-parameter subgroup L of $\mathsf{H}\tilde{\mathsf{A}}$ is either H or of the form $\mathsf{h}^{-s}\tilde{\mathsf{A}}\mathsf{h}^s$ for some $s \in \mathbb{R}$.*

Exercise 19. Prove this.

Solution. *Every one parameter subgroup is of the form $t \mapsto \exp(Xt)$ for some X in the Lie algebra of $\mathsf{H}\tilde{\mathsf{A}}$. But*

$$\exp\left(\begin{pmatrix} a & 0 & b \\ 0 & -a & 0 \\ 0 & 0 & 0 \end{pmatrix} t\right) = \begin{pmatrix} e^{at} & 0 & \frac{b}{a}(e^{at}-1) \\ 0 & e^{-at} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

so, up to a time scaling, all one-parameter subgroups of $\mathsf{H}\tilde{\mathsf{A}}$ distinct from H are of the form $t \mapsto \mathsf{h}^{s(e^t-1)}\tilde{\mathsf{g}}_t = \mathsf{h}^{-s}\tilde{\mathsf{g}}^t\mathsf{h}^s$ for some $s \in \mathbb{R}$.

Sublemma 4. *If for some $s \in \mathbb{R}$ the measure μ is invariant under the one-parameter subgroup $\mathsf{h}^{-s}\tilde{\mathsf{A}}\mathsf{h}^s$ then $\mathsf{h}^s\mu$ is invariant under $\tilde{\mathsf{A}}$.*

Proof. One verifies that

$$(\mathsf{h}^s\mu)(\tilde{\mathsf{g}}^t A) = \mu(\mathsf{h}^{-s}\tilde{\mathsf{g}}^t A) = \mu(\mathsf{h}^{-s}\tilde{\mathsf{g}}^t\mathsf{h}^s\mathsf{h}^{-s}A) = \mu(\mathsf{h}^{-s}A) = (\mathsf{h}^s\mu)(A)$$

for all Borel sets A. □

Step 2. If $\mathsf{h}^s\mu$ is $\tilde{\mathsf{A}}$ invariant then μ is supported on a single G orbit.

Write ν for $\mathsf{h}^s\mu$. The proof of Proposition 2.2 does not generalize to our current setting, because we don't know that ν sits inside a G orbit that is the support of an $\tilde{\mathsf{N}}$ ergodic measure. The proof in this case is beyond the scope of these notes, but we give an outline.

Idea: We want to show that if ν is $\tilde{\mathsf{g}}^a$ invariant and $\tilde{\mathsf{N}}$ invariant then it is $\tilde{\mathsf{M}}$ invariant. We assume a is positive, the other case being similar. The transformation $\tilde{\mathsf{g}}^a$ expands the $\tilde{\mathsf{N}}$ flow by the maximal rate possible. Indeed

$$\tilde{\mathsf{g}}^a \bigcup_{s=0}^{t} \tilde{\mathsf{u}}^s x = \bigcup_{s=0}^{e^{2a}t} \tilde{\mathsf{u}}^s \tilde{\mathsf{g}}^a x$$

for all $t \in \mathbb{R}$. For ν to be invariant under $\tilde{\mathsf{g}}^a$ there has to be a balancing contraction. Such a contraction can only come from a conjugate of $\tilde{\mathsf{M}}$. To make this heuristic precise, one uses a notion, *entropy*, which goes beyond the scope of this minicourse.

Sublemma 5. *The measure ν is supported on a single G orbit.*

Proof. Let $(\Lambda, w) \in \tilde{X}$ be a point in the support of ν. Since ν is $\tilde{\mathsf{M}}$, $\tilde{\mathsf{A}}$ and $\tilde{\mathsf{N}}$ invariant the point $(B\Lambda, Bw)$ is in the support of ν for all $B \in \mathsf{SL}(2,\mathbb{R})$. So $\nu_{B\Lambda}(\{(B\Lambda, Bw)\}) = c > 0$ for all $B \in \mathsf{SL}(2,\mathbb{R})$ by Lemma 3.6. Disintegration of measures the implies that the ν measure of the $\mathsf{SL}(2,\mathbb{R})$ orbit of (Λ, w) is at least c. This is a $\tilde{\mathsf{N}}$ invariant set and so by the ergodicity of ν (which is a consequence of the $\tilde{\mathsf{N}}$ ergodicity of μ) it has full ν measure. □

The sublemma establishes that $\mu = \mathsf{h}^{-s}\nu$ is supported on a single $\mathsf{h}^{-s}\mathsf{SL}(2,\mathbb{R})\mathsf{h}^s$ orbit.

Step 3. The measure μ is invariant under a one-parameter subgroup of $\mathsf{H}\tilde{\mathsf{A}}$ other than H.

Lemma 3.12. *The measure μ is invariant under a one parameter subgroup of $\mathsf{H}\tilde{\mathsf{A}}$ other than H.*

Sublemma 6. *It suffices to show for all $\epsilon > 0$ that there is $g \in \mathsf{H}\tilde{\mathsf{A}}\backslash\{I\}$ with $\mathsf{D}(g, I) < \epsilon$ and $g\mu = \mu$.*

Proof. Suppose we can find for every $\epsilon > 0$ some $g \in \mathsf{H}\tilde{\mathsf{A}} \setminus \{I\}$ with $\mathsf{D}(g, I) < \epsilon$ and $g\mu = \mu$. By Lemma 3.8 we get that the subgroup can not be H. Thus if for all $\epsilon > 0$ there exists $g \in \mathsf{H}\tilde{\mathsf{A}} \setminus \{I\}$ so that $\mathsf{D}(g, I) < \epsilon$ and $g\mu = \mu$ then for all $\delta > 0$ and all t there exist s, a with $(\mathsf{g}^s, \left(\begin{smallmatrix} a \\ 0 \end{smallmatrix}\right))\mu = \mu$ and $|s - t| < \delta$. Let $L = \{g \in \tilde{G} : g\mu = \mu\}$ which, just as in Exercise 9, is a closed subgroup of $\tilde{G}$. So, for each t there exists a with $(\mathsf{g}^t, \left(\begin{smallmatrix} a \\ 0 \end{smallmatrix}\right)) \in L$. If it is not unique then μ is $\mathsf{H}\tilde{\mathsf{A}}$ invariant. This gives the sublemma. □

To complete the proof of Proposition 3.10 it remains to verify the condition in the statement of the sublemma.

In analogy with (8) and (14) we calculate that

$$\begin{aligned} \tilde{\mathsf{u}}^t \mathsf{h}^a \mathsf{w}^b \tilde{\mathsf{v}}^s &= \mathsf{h}^{a+tb} \mathsf{w}^b \tilde{\mathsf{v}}^{\frac{s}{1+ts}} \tilde{\mathsf{g}}^{\log(1+ts)} \tilde{\mathsf{u}}^{\frac{t}{1+ts}} \\ &= \tilde{\mathsf{v}}^{\frac{s}{1+ts}} \mathsf{w}^{b - \frac{s}{1+ts}(a+tb)} \mathsf{h}^{a+tb} \tilde{\mathsf{g}}^{\log(1+ts)} \tilde{\mathsf{u}}^{\frac{t}{1+ts}} \end{aligned} \tag{15}$$

whenever $st \neq -1$. Once again, fix K a compact subset of points generic for μ along $\tilde{\mathsf{N}}$ with measure at least $\frac{99}{100}$. Let $E \subset X$ and $T_0 > 1$ be so that $\mu(E) > \frac{99}{100}$ and that

$$|\{0 \leq s \leq T : \tilde{\mathsf{u}}^s x \in K\}| > \tfrac{98}{100} T$$

for all $T > T_0$ and all $x \in E$. Because the projection of μ to X equals the projection of Haar measure on G to X, we can find $x \in E$, and sequences $s_i, a_i, b_i \to 0$ so that $\mathsf{h}^{a_i} \mathsf{w}^{b_i} \tilde{\mathsf{v}}^{s_i} x \in E$ for all i.

Suppose there is $0 < \epsilon < 1$ such that $g\mu \neq \mu$ whenever $\mathsf{D}(g, I) \leq \epsilon$. Then whenever $\mathsf{D}(g, I) < \epsilon$ we have $\mathsf{d}(K, gK) > 0$. Since $g \mapsto \mathsf{d}(K, gK)$ is continuous there is $0 < \delta < \frac{\epsilon}{9}$ such that

$$\tfrac{\epsilon}{100} \leq \mathsf{D}(g, I) \leq \epsilon \Rightarrow \mathsf{d}(K, gK) > \delta \tag{16}$$

holds. By shrinking ϵ if necessary we may assume that

$$\mathsf{D}(\mathsf{h}^a\tilde{\mathsf{g}}^b, I) > \tfrac{1}{2}\sqrt{a^2+b^2}, \tag{17}$$

holds, which is the other side of (2).

Using (15) we will obtain a contradiction as in the proof of Proposition 2.7. For convenience we assume $t_i > 0$, $a_i \geq 0$ and $b_i \geq 0$ for all $i \in \mathbb{N}$. Choose i so that

- $x, \mathsf{h}^{a_i}\mathsf{w}^{b_i}\tilde{\mathsf{v}}^{t_i}x \in E$;
- $s_i \neq 0$;
- $\max\{|b_i|, |s_i|, |a_i|\} < \frac{\delta}{32T_0}$

all hold.

Let $\sigma = \min\{s : \mathsf{D}(\mathsf{h}^{a_i+sb_i}\tilde{\mathsf{g}}^{\log(1+st_i)}, I) = \frac{\epsilon}{8}\}$ and consider the interval $J = [\sigma, \frac{9\sigma}{8}]$. Note that by (17) we have

$$|a_i + \sigma b_i| \text{ and } |\log(1+\sigma t_i)| \text{ are at most } \tfrac{\epsilon}{4} \tag{18}$$

for our choice of i.

It suffices to prove that the conditions

(1) $\sigma > T_0$;
(2) $\frac{\epsilon}{100} < \mathsf{D}(\mathsf{h}^{a_i+sb_i}\tilde{\mathsf{g}}^{\log(1+st_i)}, I) < \epsilon$ for all $s \in J$;
(3) $\mathsf{D}(\tilde{\mathsf{u}}^s\mathsf{h}^{a_i}\mathsf{w}^{b_i}\tilde{\mathsf{v}}^{t_i}, \mathsf{h}^{a_i+sb_i}\tilde{\mathsf{g}}^{\log(1+st_i)}\tilde{\mathsf{u}}^{\frac{s}{1+st_i}}) < \frac{\delta}{2}$ for all $s \in J$;

all hold. Indeed conditions (1) and (2) imply there exists $s \in [\sigma, \frac{4}{3}\sigma]$ so that $\tilde{\mathsf{u}}^{\frac{s}{1+t_is}}x$ and $\tilde{\mathsf{u}}^s\mathsf{h}^{a_i}\mathsf{w}^{b_i}\tilde{\mathsf{v}}^{t_i}x$ both belong to K. Thus, by condition (3) we have $\mathsf{d}(K, \mathsf{h}^{a_i+sb_i}\tilde{\mathsf{g}}^{\log(1+st_i)}K) < \delta$. This contradicts (16) completing the proof.

We outline the proof of condition (2) and leave the rest as an exercise. First, estimate that

$$\begin{aligned}
&\mathsf{D}(\mathsf{h}^{a_i+\frac{9}{8}\sigma b_i}\tilde{\mathsf{g}}^{\log(1+\frac{9}{8}\sigma t_i)}, \mathsf{h}^{a_i+\sigma b_i}\tilde{\mathsf{g}}^{\log(1+\sigma t_i)}) \\
=&\mathsf{D}(\mathsf{h}^{a_i+\frac{9}{8}\sigma b_i}\tilde{\mathsf{g}}^{\log(1+\frac{9}{8}\sigma t_i)-\log(1+\sigma t_i)}\mathsf{h}^{-a_i-\sigma b_i}, I) \\
=&\mathsf{D}(\mathsf{h}^{a_i+\frac{9}{8}\sigma b_i+\frac{1+\frac{9}{8}\sigma t_i}{1+\sigma t_i}(a_i+\sigma b_i)}\tilde{\mathsf{g}}^{\log(1+\frac{9}{8}\sigma t_i)-\log(1+\sigma t_i)}, I)
\end{aligned}$$

because D is right invariant.

Now as $\log(1+\frac{9}{8}\sigma t_i) - \log(1+\sigma t_i) < \frac{1}{8}\log(1+\sigma t_i)$ and $1 < \frac{1+\frac{9}{8}\sigma t_i}{1+\sigma t_i} < \frac{9}{8}$ our bounds (2) and (18) together with the triangle inequality imply

$$\mathsf{D}(\mathsf{h}^{a_i+\frac{9}{8}\sigma b_i+\frac{1+\frac{9}{8}\sigma t_i}{1+\sigma t_i}(a_i+\sigma b_i)}\tilde{\mathsf{g}}^{\log(1+\frac{9}{8}\sigma t_i)-\log(1+\sigma t_i)}, I) < \epsilon(\tfrac{1}{9}\tfrac{1}{4} + \tfrac{1}{9}\tfrac{1}{4})$$

holds. Note that our argument here gives

$$\mathsf{D}(\mathsf{h}^{a_i+\sigma b_i}\tilde{\mathsf{g}}^{\log(1+\sigma t_i)}, \mathsf{h}^{a_i+sb_i}\tilde{\mathsf{g}}^{\log(1+st_i)}) < \tfrac{\epsilon}{18}$$

for all $s \in J$. By the triangle inequality and the fact that σ is defined so that $\mathsf{D}(\mathsf{h}^{a_i+\sigma b_i}\tilde{\mathsf{g}}^{\log(1+\sigma t_i)}, I) = \frac{\epsilon}{8}$ this gives condition (2).

3.6. Concluding remarks.

Remarks on the proof. We repeatedly used the fact that if two measures share a generic point then they are equal.

4. Dani's proof

In this section let m_X denotes Haar measure on X, with the scalings chosen so that m_X is a probability measure.

Here we reproduce Dani's result tha nayt horospherical invariant, ergodic probability measures on $SL(2, \mathbb{R}/\mathsf{SL}(2, \mathbb{Z})$ is either supported on a closed orbit or is m_X. Our proof is based on Alex Eskin's notes from clay summer school Section 2. For a slightly different (and probably slicker) proof see [**MR1248925**].

Theorem 4.1. *Any* N *ergodic and* N *invariant Borel probability measure on* X *is either Haar measure to* X *or supported on a periodic* N *orbit.*

For every $\eta > 0$ define

$$\mathcal{K}_\eta = \{x \in X : \text{ the shortest vector in } x \text{ has length at least } \epsilon\}$$

which, by Mahler's compactness criterion, is compact.

Lemma 4.2. *The set* $\mathcal{K}_\eta$ *is compact for all* $\eta > 0$.

Exercise 20. Prove this special case of Mahler's compactness criterion directly.

Solution. *Fix* $\eta > 0$. *We show that* $\mathcal{K}_\eta$ *is sequentially compact. Let* x_i *be a sequence in* $\mathcal{K}_\eta$. *For each* i *let* v_i *be a vector of minimal positive length in* x_i. *Note that every lattice of unit co-volume contains a vector of length at most 6, so the* v_i *all lie in a compact subset of* $\mathbb{R}^2 \setminus B(0, \eta)$. *Thus there exists indices* i_j *so that* v_{i_j} *converges to* $v_\infty \in \mathbb{R}^2 \setminus B(0, \eta)$. *Now for each* i_j *we consider the vectors that together with* v_{i_j} *generate* x_{i_j}. *Note that if* w_{i_j} *is such a vector then* $w_{i_j} - kv_{i_j}$ *is too for all* $k \in \mathbb{Z}$. *So if we consider* $w_{i_j} = a_{i_j} + b_{i_j}$ *where* a_{i_j} *is parallel to* v_{i_j} *and* b_{i_j} *is perpendicular to* v_{i_j} *then* $|a_{i_j}| < |v_{i_j}| \leq 6$. *Now, the rectangle with sides* w_{i_j} *and* v_{i_j} *has area 1 (because our lattice has covolume 1) and so the* w_{i_j} *have a convergent sequence as well.*

Lemma 4.3. *There exists $\eta > 0$ so that for all $x \in X$ that do not contain a horizontal vector, there are arbitrarily large $d > 0$ such that $\mathsf{g}^{-d}x$ belongs to $\mathcal{K}_\eta$.*

Proof. If a lattice x has a horizontal vector then so does $\mathsf{g}^s x$ for all s. It therefore sufficies to exhibit $\eta > 0$ so that, for every $y \in X$ without a horizontal vector, there exists $d > 0$ so that $\mathsf{g}^{-d}y \in \mathcal{K}_\eta$.

Fix $y \in X$ without a horizontal vector. Since y is unimodular the parallelogram spanned by any two vectors generating y has area 1. This implies there exists a constant $\eta > 0$ so that, in any spanning set for any lattice without a horizontal vector, at least one of the vectors must have length at least η.

Suppose that v is a shortest non-zero vectors in y and that $\|v\| < \eta$. We claim that $\mathsf{g}^{-d}y \in \mathcal{K}_\eta$ for some $t > 0$. Since v is not horizontal $\|\mathsf{g}^{-s}v\| \to \infty$ as $s \to \infty$ and so there exists a smallest time d so that $\|\mathsf{g}^{-d}v\| = \eta$. By the previous paragraph, for any $0 \le s < d$, any vector w such that $\{w, \mathsf{g}^{-s}v\}$ generates $\mathsf{g}^{-s}y$ satisfies $\|w\| \ge \eta$. We conclude that every non-zero vector in $\mathsf{g}^{-d}y$ has length at least η, establishing the lemma. □

Lemma 4.4. *For all $\eta > 0$ there exists $R, A, K > 0$ so that for all $x \in \mathcal{K}_\eta$ one has*

$$\mathsf{v}^r\mathsf{g}^a\mathsf{u}^k x \neq \mathsf{v}^s\mathsf{g}^b\mathsf{u}^l x$$

for all $(r, a, k) \neq (s, b, l)$ with $0 \le r, s < R$, $0 \le a, b < A$ and $0 \le k, l < K$. In other words, the map $(r, a, k) \mapsto \mathsf{v}^r\mathsf{g}^a\mathsf{u}^k x$ is injective on the cuboid $[0, R) \times [0, A) \times [0, K)$ for all $x \in \mathcal{K}_\eta$.

Proof. Fix $\eta > 0$. Since Γ is discrete we can find for every $x \in X$ quantities $R, A, K > 0$ so that $\mathsf{v}^r\mathsf{g}^a\mathsf{u}^k x \neq \mathsf{v}^s\mathsf{g}^b\mathsf{u}^l x$ whenever $(r, a, k) \neq (s, b, l)$ with $0 \le r, s < R$, $0 \le a, b < A$ and $0 \le k, l < K$. The constants R, A, K clearly change continuously as x does so by Lemma 4.2 we can choose these constants uniformly on $\mathcal{K}_\eta$. □

The main step in the proof of Theorem 4.1 is the following proposition.

Proposition 4.5. *If x has no horizontal vector then*

$$\liminf_{R \to \infty} \left| \frac{1}{R} \int_0^R f(\mathsf{u}^t x)\, \mathrm{d}t - \int f \,\mathrm{d}\mathsf{m}_X \right| = 0$$

whenever $f \in \mathsf{C_c}(X)$ is 1-Lipschitz.

Exercise 21. Prove Theorem 4.1 assuming Proposition 4.5.

Solution. *Assume μ is an N ergodic and N invariant probability measure on X that is distinct from Haar measure. Let $x \in X$ be a generic point for μ along N. Fix $f \in \mathsf{C_c}(X)$ so that $\int f \,\mathrm{d}\mu \neq \int f \,\mathrm{d}\mathsf{m}_X$. Since Lipschitz functions are dense in $\mathsf{C_c}(X)$ for the supremum norm and therefore for the L^1 norm, we may assume f is Lipschitz. We may scale f to be 1-Lipschitz and observe that since $\int cf \,\mathrm{d}\mu \neq \int cf \,\mathrm{d}\mathsf{m}_X$ the proposition implies x must have a horizontal vector. This establishes the theorem.*

It is not hard to see that this proposition actually establishes more than Ratner's measure classification theorem: it establishes Ratner's orbit closure theorem and a weak form of Ratner's genericity theorem.

Write λ for Lebesgue measure on $\mathbb{R}^3$.

Lemma 4.6. *Define $\phi : \mathbb{R}^3 \to \mathsf{SL}(2,\mathbb{R})$ by $\phi(r,a,k) = \mathsf{v}^r \mathsf{g}^a \mathsf{u}^k$. There exists $\sigma > 0$ and $\delta > 0$ so that*

$$\sigma \leq \frac{d(\phi\lambda)}{d\mathsf{m}_G}(\alpha,\beta,\gamma) \leq 2\sigma$$

whenever $0 \leq |r|, |a|, |k| < \delta$.

Proof. The map ϕ is a diffeomorphism between a neighbourhood of 0 in $\mathbb{R}^3$ and a neighbourhood of the identity in $\mathsf{SL}(2,\mathbb{R})$. There its Jacobian is non-singular, so $\phi\lambda$ and m_G are equivalent on a neighbourhood of the identity in $\mathsf{SL}(2,\mathbb{R})$. □

We turn now to the proof of Proposition 4.5. Fix $f : X \to \mathbb{R}$ non-zero that is 1-Lipschitz and fix $\epsilon > 0$. Let η be as in Lemma 4.3. Fix R, A, K as in Lemma 4.4 applied to our value of η. Let δ and σ be as in Lemma 4.6. Fix $L < \min\{R, A, K, \delta, \eta, \epsilon/2\}$ such that

$$\|f\|_{\sup} \left| \frac{e^{2L}-1}{2L} - 1 \right| < \epsilon \tag{19}$$

holds.

For every $S > 0$ define the set

$$E(S) = \left\{ y \in X : \left| \frac{1}{T} \int_0^T f(\mathsf{u}^t y)\,\mathrm{d}t - \int f \,\mathrm{d}\mathsf{m}_X \right| < \epsilon \text{ for all } T \geq S \right\}$$

and note that $\mathsf{m}_X(E(S)) \to 1$ as $S \to \infty$ because m_X almost every $y \in X$ is generic for m_X along N by Theorem A.11.

Finally, fix $x \in X$ corresponding to a lattice that does not contain a horizontal vector. Choose $S > 0$ such that

$$\mathsf{m}_X(E(S)) > 1 - \frac{\epsilon L^3}{\|f\|_{\sup}\sigma}$$

holds. Fix $T \geq S$ and choose $d > 0$ such that

- $\mathsf{g}^{-d}x \in \mathcal{K}_\eta$
- $Le^{-2d} < \frac{\epsilon}{2}$
- $(1 + T + T^2)Le^{-2d} < \epsilon$
- $T\|f\|_{\sup} < Le^{2d}\epsilon$

all hold. This is possible by Lemma 4.3.

Since $\mathsf{g}^{-d}x$ belongs to $\mathcal{K}_\eta$ the map

$$(r, a, k) \mapsto \mathsf{v}^r\mathsf{g}^a\mathsf{u}^k\mathsf{g}^{-d}x$$

is injective on $[0, L]^3$ by our choice of L. Moreover, its image has measure at least $L^3/2\sigma$ by Lemma 4.6. Now g^d is a measure-preserving homeomorphism of X so the map

$$(r, a, k) \mapsto \mathsf{g}^d\mathsf{v}^r\mathsf{g}^a\mathsf{u}^k\mathsf{g}^{-d}x$$

is also injective on $[0, L]^3$ and its image too has measure at least $L^3/2\sigma$. Our choices $T \geq S > 0$ are therefore such that

$$\left| \frac{1}{TL^3} \int_0^T \int_0^L \int_0^L \int_0^L f(\mathsf{u}^t\mathsf{g}^d\mathsf{v}^r\mathsf{g}^a\mathsf{u}^k\mathsf{g}^{-d}x)\,\mathrm{d}k\,\mathrm{d}a\,\mathrm{d}r\,\mathrm{d}t - \int f\,\mathrm{d}\mathsf{m}_X \right| < 2\epsilon$$

holds. This estimate is the same as

$$\left| \frac{1}{TL^3} \int_0^T \int_0^L \int_0^L \int_0^L f(\mathsf{u}^t\mathsf{v}^{re^{-2d}}\mathsf{g}^a\mathsf{u}^{ke^{2d}}x)\,\mathrm{d}k\,\mathrm{d}a\,\mathrm{d}r\,\mathrm{d}t - \int f\,\mathrm{d}\mathsf{m}_X \right| < 2\epsilon$$

by some matrix manipulation.

For all $0 \leq k, a, r < L$ and all $0 \leq t < T$ we have the estimate

$$\begin{aligned} &\mathsf{d}(\mathsf{u}^t\mathsf{v}^{re^{-2d}}\mathsf{g}^a\mathsf{u}^{ke^{2d}}x, \mathsf{u}^{te^{-2a}}\mathsf{u}^{ke^{2d}}x) \\ &\leq \mathsf{D}(\mathsf{u}^t\mathsf{v}^{re^{-2d}}\mathsf{g}^a\mathsf{u}^{-te^{-2a}}, I) \\ &= \mathsf{D}(\mathsf{u}^t\mathsf{v}^{re^{-2d}}\mathsf{u}^{-t}\mathsf{g}^a, I) \\ &\leq \mathsf{D}(\mathsf{u}^t\mathsf{v}^{re^{-2d}}\mathsf{u}^{-t}, I) + a \end{aligned}$$

from (2). Now

$$\mathsf{u}^t \mathsf{v}^{re^{-2d}} \mathsf{u}^{-t} = \mathsf{v}^{\frac{re^{-2d}}{1+tre^{-2d}}} \mathsf{g}^{\log(1+tre^{-2d})} \mathsf{u}^{\frac{t}{1+tre^{-2d}}} \mathsf{u}^{-t}$$
$$= \mathsf{v}^{\frac{re^{-2d}}{1+tre^{-2d}}} \mathsf{g}^{\log(1+tre^{-2d})} \mathsf{u}^{\frac{-t^2 re^{-2d}}{1+tre^{-2d}}}$$

from (5). Thus

$$\begin{aligned} \mathsf{d}(\mathsf{u}^t \mathsf{v}^{re^{-2d}} \mathsf{g}^a \mathsf{u}^{ke^{2d}} x, \mathsf{u}^{te^{-2a}} \mathsf{u}^{ke^{2d}} x) &< a + re^{-2d} + tre^{-2d} + t^2 re^{-2d} \\ &\leq L + L(1+T+T^2)e^{-2d} < 2\epsilon \end{aligned}$$

holds from $L < \epsilon$ and our choice of d.

Combined with the fact that f is 1-Lipschitz, we see that

$$\frac{1}{TL^3} \int_0^T \int_0^L \int_0^L \int_0^L f(\mathsf{u}^t \mathsf{v}^{re^{-2d}} \mathsf{g}^a \mathsf{u}^{ke^{2d}} x) \, \mathrm{d}k \, \mathrm{d}a \, \mathrm{d}r \, \mathrm{d}t$$

and

$$\frac{1}{TL^2} \int_0^T \int_0^L \int_0^L f(\mathsf{u}^{te^{-2a}} \mathsf{u}^{ke^{2d}} x) \, \mathrm{d}k \, \mathrm{d}a \, \mathrm{d}t$$

are within 2ϵ of each other.

The two paragraphs above combine to give

$$\left| \int f \, \mathrm{d}\mathsf{m}_X - \frac{1}{TL^2 e^{2d}} \int_0^T \int_0^L \int_0^{Le^{2d}} f(\mathsf{u}^{te^{-2a}} \mathsf{u}^k x) \, \mathrm{d}k \, \mathrm{d}a \, \mathrm{d}t \right| < 4\epsilon \tag{20}$$

after a change of variables.

Claim 4.7. For all $0 \leq a \leq L$ the estimate

$$\left| \frac{1}{TLe^{2d}} \int_0^T \int_0^{Le^{2d}} f(\mathsf{u}^{te^{-2a}} \mathsf{u}^k x) \, \mathrm{d}k \, \mathrm{d}t - \frac{e^{2a}}{Le^{2d}} \int_0^{Le^{2d}} f(\mathsf{u}^t x) \, \mathrm{d}t \right| \leq \frac{T}{Le^{2d}} \|f\|_{\mathsf{u}}$$

holds.

Proof. Fix $0 \leq a \leq L$. Upon comparing the square with vertices $(0,0)$, $(Le^{2d}, 0)$, $(0, T)$, (Le^{2d}, T) with the parallelogram with vertices $(0,0)$, $(Le^{2d}, 0)$, $(-T, T)$, $(Le^{2d} - T, T)$ one obtains

$$\left| \int_0^T \int_0^{Le^{2d}} f(\mathsf{u}^{te^{-2a}} \mathsf{u}^k x) \, \mathrm{d}k \, \mathrm{d}t - Te^{2a} \int_0^{Le^{2d}} f(\mathsf{u}^t x) \, \mathrm{d}t \right| \leq T^2 \|f\|_{\sup}$$

after a change of variables. Divide through by TLe^{2d}. ◻

It follows from integrating the claim over $0 \leq a \leq L$ that

$$\left| \frac{1}{TL^2 e^{2d}} \int_0^T \int_0^L \int_0^{Le^{2d}} f(\mathsf{u}^{te^{-2a}} \mathsf{u}^k x) \,\mathrm{d}k \,\mathrm{d}a \,\mathrm{d}t - \frac{e^{2L}-1}{2L} \frac{1}{Le^{2d}} \int_0^{Le^{2d}} f(\mathsf{u}^t x) \,\mathrm{d}t \right|$$

is at most $T\|f\|_{\mathsf{u}}/Le^{2d}$ so, together with (20) and our choice of $d > 0$, we get

$$\left| \int f \,\mathrm{d}\mathsf{m}_X - \frac{e^{2L}-1}{2L} \frac{1}{Le^{2d}} \int_0^{Le^{2d}} f(\mathsf{u}^t x) \,\mathrm{d}t \right| < 5\epsilon$$

and then

$$\left| \int f \,\mathrm{d}\mathsf{m}_X - \frac{1}{Le^{2d}} \int_0^{Le^{2d}} f(\mathsf{u}^t x) \,\mathrm{d}t \right| < 6\epsilon$$

follows from (19). Our choice of d can be arbitrarily large by Lemma 4.3 so the preceding equation gives

$$\liminf_{d \to \infty} \left| \frac{1}{Le^{2d}} \int_0^{Le^{2d}} f(\mathsf{u}^t x) \,\mathrm{d}t - \int f \,\mathrm{d}\mathsf{m}_X \right| < 6\epsilon$$

and since $\epsilon > 0$ was arbitrary this establishes Proposition 4.5.

Appendix A. Results from ergodic theory

Inner regularity. Let X be a metric space. Every Borel measure μ on X has the property that

$$\mu(E) = \sup\{\mu(K) : E \supset K \text{ compact}\}$$

for every Borel set $E \subset X$.

Poincaré recurrence.

Lemma A.1 (Poincaré recurrence Theorem)**.** *Let $(X, \mathcal{B}, \mu)$ be a finite measure space. If $T : X \to X$ is μ-measure preserving then for any measurable set A we have that for μ almost every $x \in A$ there exists $n(x) > 0$ so that $T^{n(x)}(x) \in A$.*

Birkhoff ergodic theorem.

Theorem A.2 (Birkhoff ergodic theorem)**.** *Let $(X, \mathcal{B}, \mu)$ be a probability space. Let $t \mapsto F^t$ be a measurable flow with respect to which μ is F invariant and F ergodic. Fix $f \in L^1(X, \mathcal{B}, \mu)$. Then*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(F^t x)\,\mathrm{d}t = \int f\,\mathrm{d}\mu$$

for μ almost every x. Moreover, if

$$s_N(x) = \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)$$

then s_N converges in $L^1(\mu)$ to $\int f\,\mathrm{d}\mu$.

Corollary A.3. Let (X, F^t, μ) be an ergodic probability measure preserving system and $f \in L^1(\mu)$. For every $\epsilon > 0$ there exists T_0 so that

$$\mu(\{x \in X : \exists T \geq T_0 \text{ with } \big|\frac{1}{T}\int_0^T f(F^t x)dt - \int f d\mu\big| \geq \epsilon\}) < \epsilon.$$

Defintion A.4. *Let μ be an F^t invariant measure. We say x is F^t* forward generic *for μ if for all $f \in C_c(X)$ we have*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(F^s x) ds = \int f d\mu.$$

We similarly define backwards generic points. A point is called generic *for μ if it is both forward and backward generic.*

Corollary A.5. If μ is an ergodic measure then μ almost every point is generic for μ. If μ and ν are different invariant probability measures then the set of their generic points are disjoint.

Change of variables.

Theorem A.6 (Change of variables)**.** *If $f : [a, b] \to \mathbb{R}$ is differentiable and $U \subset [a, b]$ is measurable then*

$$\lambda(f(U)) = \int_U f'\,\mathrm{d}\lambda$$

where λ is Lebesgue measure.

Note that $f(U)$ will not generally be Borel measurable but is analytic and hence Lebesgue measurable.

Corollary A.7. If $f : [a,b] \to \mathbb{R}$ is differentiable and $U \subset [a,b]$ is measurable then

$$\frac{\inf\{f'(s) : s \in (a,b)\}}{\sup\{f'(s) : s \in (a,b)\}} \leq \frac{\lambda(f(U))}{\lambda(f([a,b] \setminus U))} \leq \frac{\sup\{f'(s) : s \in (a,b)\}}{\inf\{f'(s) : s \in (a,b)\}}$$

Disintegrations. Given separable metric spaces X, Y and a map $\pi : X \to Y$ the space X is of course the disjoint union of the fibers $\pi^{-1}(y)$ of π. Write $\mathcal{M}(X)$ for the set of Borel probability measures on X. Equip it with the weak* topology induced by the space $\mathsf{C_c}(X)$ of continuous, compactly supported functions on X. Disintegration of measures allows us to decompose any member of $\mathcal{M}(X)$ as an integral of probability measures supported on the fibres of π. The following theorem makes this precise.

Theorem A.8 (Disintegration of measures)**.** *Let X, Y be separable metric spaces. Let $\pi : X \to Y$ be Borel measurable and let μ be a Borel probability measure on X. Then there exists a μ almost-everywhere defined map $P : Y \to \mathcal{M}(X)$ with the following properties:*

1. *for $\pi\mu$ almost every y the measure P_y gives full measure to the fiber $\pi^{-1}(y)$;*
2. *for every Borel set $B \subset X$ the map $y \mapsto P_y(B)$ is measurable;*
3. *for every Borel set $B \subset X$ the equality*

$$\mu(B) = \int P_y(B) \, \mathrm{d}(\pi\mu)(y)$$

holds.

Moreover, this map is unique in the sense that if $P' : Y \to \mathcal{M}(X)$ also satisfies the above properties then $P'_y = P_y$ for $\pi\mu$ almost every y.

If we have in addition a measurable flow $t \mapsto F^t$ on X that descends to a measurable flow $t \mapsto E^t$ on Y in the sense that $\pi(F^t x) = E^t(\pi x)$ for all $t \in \mathbb{R}$ and all $x \in X$ then we can moreover require of P that for every $t \in \mathbb{R}$ the set

$$\{y \in Y : F^t P_y = P_{E^t y}\}$$

has full $\pi\mu$ measure. Indeed, one can verify that for any fixed $t \in \mathbb{R}$ the maps $P'_y = F^t P_y$ and $P''_y = P_{E^t y}$ satisfy statements (1) through (3) in the disintegration theorem and must therefore be equal almost everywhere. Note that we have for every $t \in \mathbb{R}$ a full-measure set of y on which equivariance holds. The order of quantifiers is important.

The ergodic decomposition.

Theorem A.9 (Ergodic decomposition)**.** *If $t \mapsto F^t$ is a continuous flow on a separable metric space X that preserves a probability measure μ then there is a μ almost-surely defined measurable map $P : X \to \mathcal{M}(X)$ such that*

$$\mu(B) = \int P_x(B)\,\mathrm{d}\mu(x)$$

for all Borel sets $B \subset X$ and P_x is an F invariant and F ergodic probability measure for μ almost all x.

Properties of $\mathsf{SL}(2,\mathbb{R})$. Let X denote $\mathsf{SL}(2,\mathbb{R})/\mathsf{SL}(2,\mathbb{Z})$ and ν denote the projection of Haar measure to X.

Theorem A.10. $\nu(X) < \infty$.

Theorem A.11. *The actions of the one-parameter subgroups*

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \qquad \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$$

on X are all ergodic and invariant for ν.

Theorem A.12 (KAN decomposition)**.** *Every element in $\mathsf{SL}(2,\mathbb{R})$ can be written uniquely in the form*

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$$

for some $t, s, \theta \in \mathbb{R}$.