ON THE "BANANA"-TRICK OF MARGULIS

ANDREAS WIESER

ABSTRACT. In this short note we would like to present a thickening technique of G.A. Margulis [Mar04] and how it can be used to prove certain equidistribution statements. The reader is assumed to be familiar with the theorem of Howe-Moore, with basic notions concerning linear groups, Haar measures and Haar measures on quotients by lattices and with the hyperbolic plane.

CONTENTS

1. EQUIDISTRIBUTION OF LONG HOROCYCLE ORBITS ON THE MODULAR SURFACE

In this section we would like to examine the asymptotic distribution of the horocycle orbits on

$$X_2 = \mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$$

which one can essentially think of as the "unit tangent bundle" of the modular surface $\mathrm{SL}_2(\mathbb{Z})\setminus\mathbb{H}$. Since the latter is not really a manifold (but rather an orbifold), the precise identification is $X_2 \mathrel{\widehat{=}} \mathrm{SL}_2(\mathbb{Z})\setminus \mathrm{T}^1\mathbb{H}$.

1.1. **Horocycles and their parametrization.** Recall that a **horocycle** through a point $(z, v) \in \mathrm{T}^1\mathbb{H}$ is the set of points in $\mathrm{T}^1\mathbb{H}$, whose orbits under the geodesic flow are asymptotic. The analogous set in $\mathrm{SL}_2(\mathbb{R})$ is given for $g \in \mathrm{SL}_2(\mathbb{R})$ by the set of $h \in \mathrm{SL}_2(\mathbb{R})$ with

$$\mathrm{d}(ga_t^{-1}, ha_t^{-1}) \to 0$$

as $t \to \infty$ where

$$a_t = \begin{pmatrix} \mathrm{e}^{-t/2} & 0 \\ 0 & \mathrm{e}^{t/2} \end{pmatrix}$$

Date: November 29, 2018.

and d denotes the left-invariant Riemannian metric on $\mathrm{SL}_2(\mathbb{R})$ (c.f. [EW11]).

Lemma 1.1. *Let $g \in \mathrm{SL}_2(\mathbb{R})$. Any $h \in \mathrm{SL}_2(\mathbb{R})$ with $\mathrm{d}(ga_t^{-1}, ha_t^{-1}) \to 0$ as $t \to \infty$ is of the form gu_s where*

$$u_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$$

for some $s \in \mathbb{R}$. Conversely, we have $\mathrm{d}(ga_t^{-1}, gu_s a_t^{-1}) \to 0$ as $t \to \infty$ for any $s \in \mathbb{R}$.

The lemma shows that the horocycle orbit through $g.\mathrm{i} \in \mathrm{T}^1\mathbb{H}$ as the orbit $U^-.(g.\mathrm{i})$ where

$$U^- = \{u_s \mid s \in \mathbb{R}\}$$

is the subgroup of unipotent upper triangular matrices.

Proof of the lemma. By using left-invariance of d and replacing h with $g^{-1}h$ we may assume without loss of generality that $g = \mathrm{id}$. Again by left-invariance one sees that

$$\mathrm{d}(a_t^{-1}, ha_t^{-1}) = \mathrm{d}(\mathrm{id}, a_t h a_t^{-1}) = \mathrm{d}(a_t h a_t^{-1}, \mathrm{id})$$

for any $t \in \mathbb{R}$. Writing

$$h = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

we compute

$$a_t h a_t^{-1} = \begin{pmatrix} \mathrm{e}^{-t/2} & 0 \\ 0 & \mathrm{e}^{t/2} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{t/2} & 0 \\ 0 & \mathrm{e}^{-t/2} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12}\mathrm{e}^{-t} \\ a_{21}\mathrm{e}^{t} & a_{22}. \end{pmatrix} \tag{1.1}$$

Therefore, $a_t h a_t^{-1} \to \mathrm{id}$ as $t \to \infty$ if and only if $a_{11} = a_{22} = 1$ and $a_{21} = 0$, that is, if and only if h lies in U^-. □

The computation in (1.1) also suggest that one might be interested in considering the group

$$U^+ = \left\{ \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix} : s \in \mathbb{R} \right\},$$

which (by (1.1)) consists of those $h \in \mathrm{SL}_2(\mathbb{R})$ with $a_t h a_t^{-1} \to \mathrm{id}$ as $t \to -\infty$. The directions given by U^- resp. U^+ are usually referred to as the stable resp. unstable directions for the geodesic flow. Correspondingly one calls U^- (resp. U^+) the **stable horocycle subgroup** (resp. unstable horocycle subgroup). Note that (by (1.1)) both these subgroups are normalized by the diagonal subgroup

$$A = \{a_t \mid t \in \mathbb{R}\}.$$

Furthermore, the **Borel subgroup**

$$B = U^+A = AU^+ = \left\{ \begin{pmatrix} a & 0 \\ b & a^{-1} \end{pmatrix} : a \in \mathbb{R}^\times, b \in \mathbb{R} \right\} < \mathrm{SL}_2(\mathbb{R}).$$

admits a similar interpretation as U^+, U^-. In fact, in analogy to Lemma 1.1 one can think of B as the set of elements g of $\mathrm{SL}_2(\mathbb{R})$ for which $\mathrm{d}(ga_t^{-1}, gu_s a_t^{-1})$ stays bounded as $t \to -\infty$.

1.2. **Local coordinates and Haar measures.** The stable and unstable directions provide local coordinates on $\mathrm{SL}_2(\mathbb{R})$ which we record here as they will be of great use to us later.

1.2.1. *Interpretation in terms of Lie-algebras.* Consider the Lie algebras

$$\mathfrak{u}^- = \left\{ \begin{pmatrix} 0 & s \\ 0 & 0 \end{pmatrix} : s \in \mathbb{R} \right\}, \quad \mathfrak{u}^+ = \left\{ \begin{pmatrix} 0 & 0 \\ s & 0 \end{pmatrix} : s \in \mathbb{R} \right\}$$
$$\mathfrak{a} = \left\{ \begin{pmatrix} t & 0 \\ 0 & -t \end{pmatrix} : t \in \mathbb{R} \right\}$$

of U^-, U^+ and A respectively and observe that $\mathfrak{u}^- \oplus \mathfrak{u}^+ \oplus \mathfrak{a} = \mathfrak{sl}_2(\mathbb{R})$. This is exactly the decomposition of $\mathfrak{sl}_2(\mathbb{R})$ into eigenspaces for the action of A on $\mathfrak{sl}_2(\mathbb{R})$ via the adjoint representation. For instance,

$$\mathrm{Ad}_{a_t} \begin{pmatrix} 0 & s \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & s\mathrm{e}^{-t} \\ 0 & 0 \end{pmatrix} = \mathrm{e}^{-t} \begin{pmatrix} 0 & s \\ 0 & 0 \end{pmatrix}$$

by (1.1). Note that the Lie algebra $\mathfrak{b}$ of B is simply the direct product $\mathfrak{u}^+ \oplus \mathfrak{a}$. The behaviour of the adjoint representation reflects the asymptotic property of points on the same horocycle and in fact yields a quantitative approaching speed. We will mainly use the Lie-algebras in order to parametrize $\mathrm{SL}_2(\mathbb{R})$ in a neighborhood of the identity as follows.

Lemma 1.2 (Local coordinates)**.** *The map*

$$\mathfrak{u}^- \oplus \mathfrak{b} \to \mathrm{SL}_2(\mathbb{R}), \quad (X, Y) \mapsto \exp(X)\exp(Y)$$

is a local diffeomorphism around 0*.*

Note that the map in the lemma is just a slightly adapted version of the exponential map, which respects the decomposition of $\mathfrak{sl}_2(\mathbb{R})$ into expanding and non-expanding directions (as $t \to -\infty$). For the definition and properties of the exponential map

$$\exp : \mathfrak{sl}_2(\mathbb{R}) = \{X \in \mathrm{Mat}_2(\mathbb{R}) : \mathrm{Tr}(X) = 0\} \to \mathrm{SL}_2(\mathbb{R})$$

we refer to [EW11, Section 9.3.1].

Proof. Let Φ be the map in the lemma. The differential of Φ at zero is the identity. Thus, there is a neighborhood $\mathcal{O}'$ of zero so that $\Phi|_{\mathcal{O}'}$ is a diffeomorphism. □

The lemma and its proof allows one to consider neighborhoods of the identity in $\mathrm{SL}_2(\mathbb{R})$ and other groups which satisfy certain natural properties for the conjugation action. For instance, the map

$$\mathfrak{b} = \mathfrak{a} \oplus \mathfrak{u}^+ \to B, \quad (X, Y) \mapsto \exp(X)\exp(Y)$$

is also a local diffeomorphism around 0. Given two small enough neighborhoods $\mathcal{O}_{\mathfrak{a}}$ resp. $\mathcal{O}_{\mathfrak{u}^+}$ of 0 in $\mathfrak{a}$ resp. $\mathfrak{u}^+$ the image under the above map yields a "rectangular" neighborhood $\mathcal{O}_B$ of the identity in B, which satisfies that

$$a_t \mathcal{O}_B a_{-t} \subset \mathcal{O}_B$$

for any $t \leq 0$.

Exercise 1.3. Give an explicit construction of such a neighborhood in B.

1.2.2. *Haar measures.* We will need to relate the Haar measure on $\mathrm{SL}_2(\mathbb{R})$ to the Haar measures on B and U^-. When talking about a Haar measure of B one needs some amount of caution:

Exercise 1.4. Show that the Borel subgroup B is not unimodular and compute its modular function.

For this, we recall that the following lemma, which is a corollary of [EW11, Lemma 11.31] since by Lemma 1.2 the product U^-B contains an open neighborhood of the identity in $\mathrm{SL}_2(\mathbb{R})$.

Lemma 1.5 (Decomposition of the Haar measure)**.** *Let $m_B^{(r)}$ be a right Haar measure on B and let m_{U^-} be a left Haar measure on U^-. Then any left Haar measure on $\mathrm{SL}_2(\mathbb{R})$ restricted to U^-B is proportional to the pushforward $\phi_*(m_{U^-} \times m_B^{(r)})$ where $\phi : U^- \times B \to \mathrm{SL}_2(\mathbb{R})$, $(u, b) \mapsto ub$.*

The Haar measure we will use on U^- is the pushforward of the Lebesgue-measure under $s \in \mathbb{R} \mapsto u_s \in U^-$. Furthermore, we will choose the Haar measure $m_{\mathrm{SL}_2(\mathbb{R})}$ on $\mathrm{SL}_2(\mathbb{R})$ so that the fundamental domain for $\mathrm{SL}_2(\mathbb{Z}) < \mathrm{SL}_2(\mathbb{R})$ has volume 1 i.e. so that the natural measure m_{X_2} on the quotient space X_2 is a probability measure. We then choose a right Haar measure $m_B^{(r)}$ on B so that equality in Lemma 1.5 is satisfied.

1.3. **Parametrization of periodic horocycle orbits.** In this section we would like to understand the periodic orbits of stable horocycle subgroup U on X_2. Here, a point $x \in X_2$ is called **periodic** if there is $s \in \mathbb{R}$ so that $u_s.x = x$. In this case, the smallest such s is called the period and the orbit $U.x$ is also called periodic.

Lemma 1.6 (A collection of periodic orbits)**.** *For any $t \in \mathbb{R}$ the orbit*

$$U^-.(\mathrm{SL}_2(\mathbb{Z})a_t) = \mathrm{SL}_2(\mathbb{Z})a_tU^- = \mathrm{SL}_2(\mathbb{Z})U^-a_t$$

is periodic with period e^t.

Proof. We first note that the orbit of the identity coset $U.(\mathrm{SL}_2(\mathbb{Z})\,\mathrm{id}) = \mathrm{SL}_2(\mathbb{Z})U$ in X_2 is periodic of period 1. Indeed, the point $\mathrm{SL}_2(\mathbb{Z})u_s$ is $\mathrm{SL}_2(\mathbb{Z})\,\mathrm{id}$ if and only if $u_s \in \mathrm{SL}_2(\mathbb{Z})$ or in other words if and only if $s \in \mathbb{Z}$.

Now let $t \in \mathbb{R}$. Then $\mathrm{SL}_2(\mathbb{Z})a_tu_s = \mathrm{SL}_2(\mathbb{Z})a_t$ if and only if (see (1.1))

$$a_tu_sa_t^{-1} = \begin{pmatrix} 1 & s\mathrm{e}^{-t} \\ 0 & 1 \end{pmatrix} = u_{s\mathrm{e}^{-t}} \in \mathrm{SL}_2(\mathbb{Z}).$$

This shows that the point $\mathrm{SL}_2(\mathbb{Z})a_t$ is periodic with period e^t as desired. □

One can show that the periodic orbits of U are in fact all of the form as in the lemma above.

Proposition 1.7 (One-parameter family of periodic orbits)**.** *Let $x \in X_2$ be a periodic point for U^-. Then there is some $t \in \mathbb{R}$ so that $U^-.x = U^-.(\mathrm{SL}_2(\mathbb{Z})a_t)$.*

Proof. We first claim that $a_t.x \to \infty$ as $t \to \infty$. By this we mean that for any compact set $K \subset X_2$ there is some $T_K > 0$ so that $a_t.x \notin K$ for all $t \geq T_K$. Let S be the period of x. Then $a_t.x$ is also periodic for U^- and has period $S\mathrm{e}^{-t}$ as

$$u_sa_t.x = a_tu_{s\mathrm{e}^t}.x = a_t.x$$

if and only if $s\mathrm{e}^t$ is a multiple of S.

Now suppose that $a_t.x \not\to \infty$ as $t \to \infty$. Then there is a compact set K and a sequence (t_n) with $t_n \to \infty$ for $n \to \infty$ so that $a_{t_n}.x \in K$ for all $n \in \mathbb{N}$. Let $r > 0$ be a uniform injectivity radius on K (c.f. [EW11, Prop.9.14]). For any $u_s \in U^- \cap B_r(\mathrm{id})$ and any $n \in \mathbb{N}$ we therefore have

$$u_s.(a_{t_n}.x) = a_{t_n}.x \implies s = 0.$$

However, since the period of the elements $a_{t_n}.x$ goes to zero, we know that arbitrarily small, non-zero $s_n \in \mathbb{R}$ with $u_{s_n}.(a_{t_n}.x) = a_{t_n}.x$ exist. A contradiction, which proves the claim.

To see how the claim implies the proposition, recall that orbits of the geodesic flow (i.e. geodesics) are either vertical lines or circles centered on the real axis. Let $(z_t, v_t) \in F$ be the point corresponding to $a_t.x$, where $F \subset \mathrm{T}^1\mathbb{H}$ is the standard fundamental domain. Let K' be the set of points in F with imaginary part ≤ 1 and note that the image K of K' in $\mathrm{SL}_2(\mathbb{Z}) \setminus \mathrm{T}^1\mathbb{H}$ is compact. Therefore, let $T_K > 0$ so that $a_t.x \notin K$ for all $t \geq T_K$ by the claim. This implies that v_{T_K} is a multiple of i. Indeed, if this were not the case, the geodesic through (z_{T_K}, v_{T_K}) would be a half circle and would therefore z_t would reach imaginary part ≤ 1 for some $t > T_K$.

Applying some $u \in U^-$ we obtain that the point in F corresponding to $u.(a_{T_K}.x)$ lies on the imaginary axis and has a vector pointing north. Therefore, there is some $t' \in \mathbb{R}$ so that $a_{t'}.(ua_{T_K}.x) = \mathrm{SL}_2(\mathbb{Z})\,\mathrm{id}$ (i.e. transporting back to $(\mathrm{i}, \mathrm{i}) \in F$). In particular,

$$U.x = U.(\mathrm{SL}_2(\mathbb{Z})a_{t'}ua_{T_K}) = U.(\mathrm{SL}_2(\mathbb{Z})a_{t'+T_K})$$

as in the proposition. □

1.4. **Equidistribution of long periodic horocycle orbits.** Notice that any periodic U^--orbit gives rise to a natural probability measure on the orbit. Indeed, if $x \in X_2$ is periodic of period T then

$$\frac{1}{T}\int_0^T f(u_s.x)\,\mathrm{d}s$$

for $f \in C_c(X_2)$ defines a linear functional (and hence a measure) with the required properties. Alternatively, the periodic orbit measure on $U.x$ is given by the push-forward of the normalized Lebesgue measure on $[0, T]$ under the map $s \mapsto u_s.x$.

We would like to know the behaviour of these periodic orbit measures (for the orbits from Lemma 1.6) as the period goes to infinity.[1]

Theorem 1.8 (Sarnak [Sar81])**.** *Let x_n be a sequence of U^--periodic points whose period goes to infinity as $n \to \infty$. Then the periodic orbit measures on $U^-.x_n$ equidistribute to the normalized Haar measure m_{X_2} on X_2 as $n \to \infty$.*

At this point we should remark that Sarnak's result from [Sar81] is in fact much stronger than what is stated above. In fact, Sarnak shows that there is some positive exponent $\alpha > 0$, a constant $C > 0$ and an L^2-Sobolev Norm $\mathcal{S}$ on $C_c^\infty(X_2)$ so that

$$\left|\mathrm{e}^{-t}\int_0^{\mathrm{e}^t} f(\mathrm{SL}_2(\mathbb{Z})a_t u_s)\,\mathrm{d}s - \int_{X_2} f\,\mathrm{d}m_{X_2}\right| \leq C\mathrm{e}^{-\alpha t}\mathcal{S}(f)$$

[1]When the period goes to zero, the periodic orbit measures converge to the zero measure as is quite directly seen.

for any $t \in \mathbb{R}$ and any function $f \in C_c^\infty(X_2)$. It is remarkable that obtaining the optimal exponent in the above estimate is equivalent to the Riemann Hypothesis – this is a result of Zagier [Zag81].

The argument for Theorem 1.8 we will present here does not follow Sarnak's methods but will be based on more soft arguments, namely on mixing properties of the geodesic flow. The idea of this proof presumably dates back to the thesis of Margulis. As we will see, the trick will be to *thicken* the periodic orbits along the unstable and the geodesic direction so as to apply mixing, which is why this trick is referred to (by Margulis and others) as the "banana" trick.

Proof of Theorem 1.8. Let $f \in C_c(X_2)$ and let $\varepsilon > 0$. As the function f has compact support, it is uniformly continuous. As the projection $\mathrm{SL}_2(\mathbb{R}) \to X_2$ is 1-Lipschitz, there is a $\delta > 0$ so that

$$\mathrm{d}(g, \mathrm{id}) < \delta \implies |f(xg) - f(x)| < \varepsilon$$

for any $g \in \mathrm{SL}_2(\mathbb{R})$ and $x \in X_2$. Denote by $P_0 = \mathrm{SL}_2(\mathbb{Z})U^-$ the periodic orbit of period 1. As P_0 is compact, there is a uniform injectivity radius on P_0. By shrinking δ if necessary we may assume that δ itself is an injectivity radius on P_0.

Definition of the thickening: Let $\mathcal{O}_B \subset B \cap B_\delta^{\mathrm{SL}_2(\mathbb{R})}(\mathrm{id})$ be a rectangular neighborhood of the identity as in Section 1.2.1 so that

$$a_{-t}\mathcal{O}_B a_t \subset \mathcal{O}_B$$

for all $t \geq 0$. Moreover, let $\tilde{P}_0 = \mathcal{O}_B.P_0$ be the thickening of the orbit P_0 given by $\mathcal{O}_B$ and denote by P_t the orbit of period e^t and by

$$\tilde{P}_t = a_{-t}.\tilde{P}_0 = (a_{-t}\mathcal{O}_B a_t).P_t$$

the induced thickening. Notice that the neighborhoods $a_{-t}\mathcal{O}_B a_t$ get thiner in the unstable direction as $t \to \infty$ and do not get thicker in any direction. For convenience we also define

$$S_t = \left\{ u_s b \mid s \in [0, \mathrm{e}^t),\, b \in a_{-t}\mathcal{O}_B a_t \right\}.$$

Note that $S_t = a_{-t} S_1 a_t$ and that

$$\tilde{P}_t = \{\mathrm{SL}_2(\mathbb{Z}) a_t g \mid g \in S_t\} = \{\mathrm{SL}_2(\mathbb{Z}) g a_t \mid g \in S_1\}.$$

Integral over thickened neighborhood in the group. First, we would like to replace the integral of f along the orbit P_t by the integral over a larger neighborhood in $\mathrm{SL}_2(\mathbb{R})$. Observe first that

$$\begin{aligned}
&\Bigg| \mathrm{e}^{-t} \int_0^{\mathrm{e}^t} f(\mathrm{SL}_2(\mathbb{Z}) a_t u_s)\,\mathrm{d}s \\
&\qquad - \tfrac{1}{m_B^{(r)}(a_{-t}\mathcal{O}_B a_t)} \mathrm{e}^{-t} \int_{a_{-t}\mathcal{O}_B a_t} \int_0^{\mathrm{e}^t} f(\mathrm{SL}_2(\mathbb{Z}) a_t u_s b)\,\mathrm{d}s d m_B^{(r)}(b) \Bigg| \\
&\leq \tfrac{\mathrm{e}^{-t}}{m_B^{(r)}(a_{-t}\mathcal{O}_B a_t)} \int_{a_{-t}\mathcal{O}_B a_t} \int_0^{\mathrm{e}^t} |f(\mathrm{SL}_2(\mathbb{Z}) a_t u_s) - f(\mathrm{SL}_2(\mathbb{Z}) a_t u_s b)|\,\mathrm{d}s d m_B^{(r)}(b) \\
&< \varepsilon
\end{aligned}$$

since $a_{-t}\mathcal{O}_B a_t \subset \mathcal{O}_B$ for any $t > 0$ by the choice of the neighborhood $\mathcal{O}_B$ (according to the uniform continuity). By Lemma 1.5 the normalized integral

$$\frac{1}{m_B^{(r)}(a_{-t}\mathcal{O}_B a_t)} e^{-t} \int_{a_{-t}\mathcal{O}_B a_t} \int_0^{e^t} f(\mathrm{SL}_2(\mathbb{Z}) a_t u_s b) \, \mathrm{d}s dm_B^{(r)}(b). \tag{1.2}$$

is equal to

$$\frac{1}{m_{\mathrm{SL}_2(\mathbb{R})}(S_t)} \int_{S_t} f(\mathrm{SL}_2(\mathbb{Z}) a_t g) \, \mathrm{d}m_{\mathrm{SL}_2(\mathbb{R})}(g).$$

Since $\mathrm{SL}_2(\mathbb{R})$ is unimodular, $m_{\mathrm{SL}_2(\mathbb{R})}(S_t) = m_{\mathrm{SL}_2(\mathbb{R})}(S_0)$ and by replacing $a_t g a_t^{-1}$ with g the integral in (1.2) is equal to

$$\frac{1}{m_{\mathrm{SL}_2(\mathbb{R})}(S_0)} \int_{S_0} f(\mathrm{SL}_2(\mathbb{Z}) g a_t) \, \mathrm{d}m_{\mathrm{SL}_2(\mathbb{R})}(g). \tag{1.3}$$

Integral over thickened orbit: Note that the image $\{\mathrm{SL}_2(\mathbb{Z})g \mid g \in S_0\}$ under the projection of S_0 to X_2 is simply $\tilde{P}_0$. Therefore, the Haar measure on S_0 is equal to the Haar measure on $\tilde{P}_0$ if we assume the following claim.

Claim 1.9. *If δ is small enough, the set $S_0 = \{u_s b \mid s \in [0,1), b \in \mathcal{O}_B\}$ is injective.*

Let us postpone the proof for the moment. Then the integral in (1.3) is equal to

$$\frac{1}{m_{X_2}(\tilde{P}_0)} \int_{\tilde{P}_0} f(x a_t) \, \mathrm{d}m_{X_2}(x) = \langle a_{-t}.f, f_0 \rangle$$

where $f_0 = \frac{1}{m_{X_2}(\tilde{P}_0)} \chi_{\tilde{P}_0}$.

Applying the mixing property Recall that the geodesic flow X_2 is mixing on X_2. That is, for any $f_1, f_2 \in L^2(X_2)$ we have

$$\langle a_t.f_1, f_2 \rangle \to \int_{X_2} f_1 \, \mathrm{d}m_{X_2} \int_{X_2} f_1 \, \mathrm{d}m_{X_2}$$

as $t \to \pm\infty$. In particular,

$$\langle a_{-t}.f, f_0 \rangle \to \int_{X_2} f \, \mathrm{d}m_{X_2} \int_{X_2} f_0 \, \mathrm{d}m_{X_2} = \int_{X_2} f \, \mathrm{d}m_{X_2}$$

as $t \to \infty$. Tracing back our arguments, we can deduce that the integral

$$e^{-t} \int_0^{e^t} f(\mathrm{SL}_2(\mathbb{Z}) a_t u_s) \, \mathrm{d}s$$

over the orbit P_0 is always within ε of a convergent expression with limit $\int_{X_2} f \, \mathrm{d}m_{X_2}$ and therefore for large enough t within 2ε of the limit itself. Thus

$$e^{-t} \int_0^{e^t} f(\mathrm{SL}_2(\mathbb{Z}) a_t u_s) \, \mathrm{d}s \to \int_{X_2} f \, \mathrm{d}m_{X_2}$$

as $t \to \infty$ as claimed in the proposition.

Proof of Claim 1.9. Assume that there are $s_1, s_2 \in [0,1)$ and $b_1, b_2 \in \mathcal{O}_B$ with

$$\mathrm{SL}_2(\mathbb{Z}) u_{s_1} b_1 = \mathrm{SL}_2(\mathbb{Z}) u_{s_2} b_2.$$

Setting $b = b_1 b_2^{-1}$ and rearranging we have

$$u_{s_1} b u_{s_2}^{-1} \in \mathrm{SL}_2(\mathbb{Z}).$$

Write $b = \begin{pmatrix} \alpha & 0 \\ \beta & \alpha^{-1} \end{pmatrix}$. Then

$$u_{s_1} b u_{s_2}^{-1} = \begin{pmatrix} \alpha + \beta s_1 & \alpha^{-1} s_1 \\ \beta & \alpha^{-1} \end{pmatrix} \begin{pmatrix} 1 & -s_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha + \beta s_1 & \alpha^{-1} s_1 - \alpha s_2 - \beta s_1 s_2 \\ \beta & \alpha^{-1} + \beta s_2. \end{pmatrix}$$

However, if δ is small enough, $b \in \mathcal{O}_B$ must be close to the identity. Since β is by the above an integer, it must be zero. Hence, $\alpha, \alpha^{-1} \in \mathbb{Z}$ and they are both close to the identity. We conclude that $b = \mathrm{id}$. This shows that $s_1 - s_2 \in \mathbb{Z}$ and thus $s_1 = s_2$ as $s_1, s_2 \in [0, 1)$. □

□

2. Equidistribution of large hyperbolic circles

2.1. Hyperbolic circles.

Proposition 2.1. *Every hyperbolic circle (i.e. a boundary of a ball in the hyperbolic plane) is a Euclidean circle.*

Note that also the converse is true, which we will not prove here.

Proof. Note first that for any $(z, v) \in \mathrm{T}^1\mathbb{H}$ the hyperbolic ball of radius $t > 0$ around z is given by the projection onto the base points of the set $g_t(K.(z, v))$, where $K < \mathrm{SL}_2(\mathbb{R})$ is the stabilizer of z under the action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathbb{H}$ by Möbius transformations and where g_t denotes the geodesic flow for time t. If $g \in \mathrm{SL}_2(\mathbb{R})$ satisfies $g.(\mathrm{i}, \mathrm{i}) = (z, v)$ then

$$g_t(K.(z, v)) = g\,\mathrm{SO}(2) a_{-t}.(\mathrm{i}, \mathrm{i}).$$

In particular, the stabilizer of a point z acts transitively on every circle around the point.

For simplicity we will assume for now that $z = \mathrm{i}$ and since v above is arbitrary that $v = \mathrm{i}$. Then $K = \mathrm{SO}(2)$ and the circle $C = \pi_{\mathrm{base}}(K a_t.(\mathrm{i}, \mathrm{i}))$ intersects the y-axis at exactly the points $\mathrm{e}^t\mathrm{i}$, $\mathrm{e}^{-t}\mathrm{i}$ so that by symmetry the natural candidate for a Euclidean center of C is

$$\frac{\mathrm{e}^t\mathrm{i} + \mathrm{e}^{-t}\mathrm{i}}{2} = \cosh(t)\mathrm{i}$$

and the Euclidean radius is $\frac{\mathrm{e}^t - \mathrm{e}^{-t}}{2} = \sinh(t)$. Denote this Euclidean circle by C_{eucl}.

To see that in fact $C = C_{\mathrm{eucl}}$ a computation shows that for any $k \in K$

$$|\, k.(\mathrm{e}^t\mathrm{i}) - \cosh(t)\mathrm{i} \,| = \sinh(t).$$

This proves that $C \subset C_{\mathrm{eucl}}$ by transitivity of the K-action on C. But in this case the reverse inclusion also has to hold so we conclude equality.

To see the proposition for an arbitrary center observe that we may consider any g with $g.\mathrm{i} = z$ (v was arbitrary). The Iwasawa decomposition and the fact that $\mathrm{SL}_2(\mathbb{R})$ acts by isometries on $\mathbb{H}$ then show that we only need to prove that the image of a Euclidean circle under transformations of the form

$$z \mapsto z + a, \quad z \mapsto \alpha z$$

for $a \in \mathbb{R}$ and $\alpha > 0$ is a Euclidean circle. This is indeed the case. □

2.2. **Equidistribution of large circles.**

Theorem 2.2. *Let* Γ *be a lattice in* $\mathrm{SL}_2(\mathbb{R})$ *and let* $X = \mathrm{SL}_2(\mathbb{R})/\Gamma$. *Denote by* m_Y *any Haar measure on the orbit* $Y = \mathrm{SO}(2)\Gamma$ *and by* m_X *any Haar measure on* X. *Then for any* $f \in C_c(X)$ *we have*

$$\frac{1}{m_Y(Y)} \int_Y f(a_t.y)\,\mathrm{d}m_Y(y) \to \frac{1}{m_X(X)} \int_X f(x)\,\mathrm{d}m_X(x)$$

as $t \to \pm\infty$.

Note that up to a switch from left- to right-quotients the theorem essentially states that the circle of radius t with arrows pointing outwards folded up under Γ equidistributes as $t \to \infty$. The statement for $t \to -\infty$ is the same just with arrows pointing inwards.

Proof. We restrict our attention to the case $t \to \infty$ as the case $t \to -\infty$ is analogous[2]. Let $f \in C_c(X)$ and let $\varepsilon > 0$. Denote $G = \mathrm{SL}_2(\mathbb{R})$ and $K = \mathrm{SO}(2)$. Let $\mathcal{O} \subset U^-A$ be an open (rectangular) neighborhood of the identity with $a_t\mathcal{O}a_{-t} \subset \mathcal{O}$ for any $t > 0$ (a contracted neighborhood) so that $f(g.x)$ is ε-close to $f(x)$ for any $g \in \mathcal{O}$ and any $x \in X$. By these choices the integral $\frac{1}{m_Y(Y)} \int_Y f(a_t.y)\,\mathrm{d}m_Y(y)$ is ε-close to

$$I_t = \frac{1}{m_{UA}(\mathcal{O})} \frac{1}{m_Y(Y)} \int_{\mathcal{O}} \int_Y f(a_t.g.y)\,\mathrm{d}m_Y(y)\,\mathrm{d}m_{U^-A}(g)$$

If F is a fundamental domain for $K \to Y$ then the above is equal to

$$\frac{1}{m_{UA}(\mathcal{O})} \frac{1}{m_K(k)} \int_{\mathcal{O}} \int_F f(a_t.gk\Gamma)\,\mathrm{d}m_K(k)\,\mathrm{d}m_{U^-A}(g)$$

by definition of the Haar measure on Y. By Lemma 1.5 applied to the Iwasawa decomposition (with U^-A and K)

$$I_t = \frac{1}{m_G(\mathcal{O}F)} \int_{\mathcal{O}F} f(a_t.g\Gamma)\,\mathrm{d}m_G(g),$$

which converges by the mixing property of the geodesic flow to the desired limit as $t \to \infty$. □

References

[EW11] M. Einsiedler and T. Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, vol. 259, Springer, 2011.

[Mar04] G.A. Margulis, *On some aspects of the theory of ANOSOV systems*, Springer Monographs in Mathematics, Springer, 2004, Translated from the Russian by Szulikowska Valentina Vladimirovna.

[Sar81] P. Sarnak, *Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series*, Communications on Pure and Applied Mathematics **34** (1981), no. 6, 719–739.

[Zag81] D. Zagier, *Eisenstein series and the riemann zeta-function*, pp. 275–301, Springer Berlin Heidelberg, Berlin, Heidelberg, 1981.

Departement Mathematik, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

[2]The decomposition one uses in this case is not the Iwasawa decomposition $G = U^-AK$ but rather $G = U^+AK$. This can be obtained by applying the homomorphism $g \mapsto (g^{-1})^t$.