EQUIDISTRIBUTION OF LARGE HOROCYCLES

RÉKA RÁCZ, YIXUAN LIU, ANNINA SIEGRIST
SEMINAR ON COUNTING PROBLEMS HS23, ETHZ

Our goal in this talk is to examine the asymptotic (i.e. limiting) distribution of the horocycle orbits on the quotient space

$$X_2 \coloneqq \mathrm{PSL}_2(\mathbb{Z})\backslash\mathrm{PSL}_2(\mathbb{R}) \cong \mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R}).$$

In this talk, we will focus on $\mathrm{SL}_2(\mathbb{R})$.

1. Recap on the Geodesic Flow, Horocycles and their Parametrizations

In our last talk, we defined horocycles in the Hyperbolic plane visually as "circles with one point at infinity". In the Poincaré disk model, these were simply the circles touching the boundary of the disk. Similarly, in the half-plane model, these were either circles touching the x-axis *or* lines parallel to the x-axis (see Figure 1). Intuitively, one can switch from the disk model to the half-plane model by doing the following: Cut open the boundary of the disk at point x, so that the disk unfolds into the half-plane. Now you can take that point x and define it to be infinity. Note that horocycles that touch the boundary of the disk in point x turn into horizontal lines in the half-plane model. Both models are useful, but the disk model is "more honest" in a way since there is no special infinity point and all horocycles look the same.

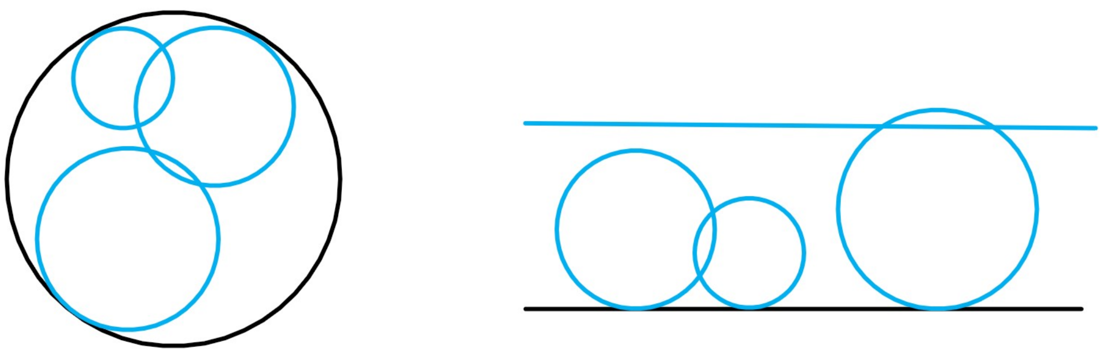

Figure 1. Horocycles in the disk model and in the halfplane model of $\mathbb{H}$.

Also, remember that we can identify any element $g \in \mathrm{SL}_2(\mathbb{R})$ (or $\mathrm{PSL}_2(\mathbb{R})$) with a unique element of the unit tangent bundle of the hyperbolic plane $(z, v) \in \mathrm{T}^1\mathbb{H}$ by acting with g on i by Möbius transformations. So this means that z represents the "point" and v the "little arrow".

Furthermore, we introduced the matrix

$$a_t = \begin{pmatrix} e^{-t/2} & 0 \\ 0 & e^{t/2} \end{pmatrix} \quad \text{with inverse} \quad a_t^{-1} = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$$

and the *geodesic flow* on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ by

$$g_t : x \longmapsto x a_t = a_t \cdot x.$$

We can now give a more abstract definition of a horocycle:

Definition/Proposition 1.1 (Horocycle defined by $T^1\mathbb{H}$)**.** A *horocycle* through a point $(z, v) \in T^1\mathbb{H}$ is the set of points in $T^1\mathbb{H}$ whose orbits under the geodesic flow are asymptotic.
The analogous set in $SL_2(\mathbb{R})$ is given for $g \in SL_2(\mathbb{R})$ by the set of $h \in SL_2(\mathbb{R})$ with

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

where we use the usual left-invariant Riemannian metric on $SL_2(\mathbb{R})$ as introduced in the talks before.

You can check that the visual definition in Figure 1 coincides with the one given in Definition 1.1: The horocycles parallel to the x-axis all correspond to the points with an arrow pointing upwards (see Figure 2). Intuitively, you can think of the geodesic flow with respect to time, so point (z_3, v_3) is always at a non-zero positive distance from point (z_2, v_2) and they will never meet.
In general, all points on the same horocycle are exactly the points of which the geodesic flow "meets at the same point at the same time" at infinity (see Figure 3). Note that this is only the case because they start at the same "height", i.e. the same horocycle.

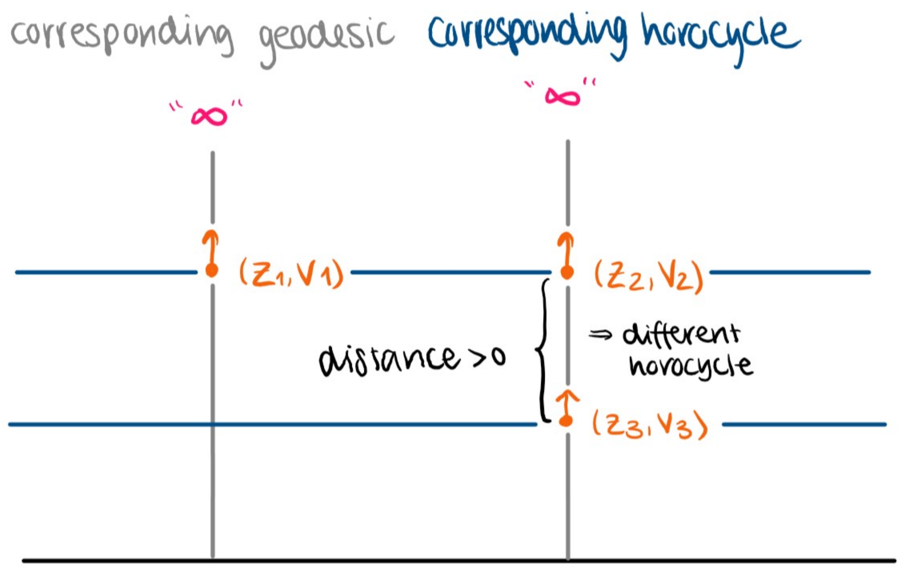

Figure 2. Horocycles horizontal to the x-axis in the half-plane model.

Lemma 1.2. *Let $g \in SL_2(\mathbb{R})$. Any $h \in SL_2(\mathbb{R})$ with $d(ga_t^{-1}, ha_t^{-1}) \to 0$ 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 $d(ga_t^{-1}, gu_s a_t{-}1) \to 0$ as $t \to \infty$ for any $s \in \mathbb{R}$.

In other words, this shows that the horocycle orbit through $g \cdot i \in T^1\mathbb{H}$ as the orbit $U^- \cdot (g \cdot i)$, where $U^- = \{u_s \mid s \in \mathbb{R}\}$ is the subgroup of unipotent upper triangular matrices.

Proof of Lemma 1.2. By using the left-invariance of d under $SL_2(\mathbb{R})$ and replacing h with $g^{-1}h$ we may assume without loss of generality that $g = \mathrm{id}$. Again, by left-invariance, we get that

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

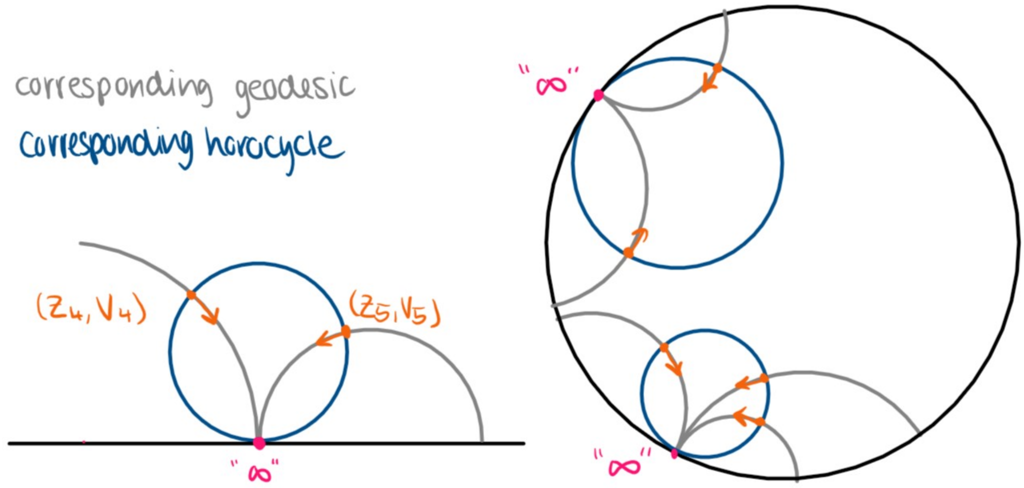

Figure 3. Other horocycles in the half-plane model and the disk model.

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

$$h = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$$

we compute

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

Thus, $a_t h a_t^{-1} \to \mathrm{id}$ 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^-. □

It is also interesting to consider the group

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

which is called the *unstable horocycle subgroup*. On the other hand, U^- is then the *stable horocycle subgroup*. Here, an element u_s can be understood as the Möbius transformation sending $z \in \mathbb{H}$ to $z + s$. So it simply translates the point z horizontally by s. They are both normalized by the diagonal subgroup $A = \{a_t \mid t \in \mathbb{R}\}$. Furthermore, we define the following subgroup:

Definition 1.3 (Borel subgroup)**.** The *Borel subgroup* is defined as

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

You can think of B as the set of elements g of $\mathrm{SL}_2(\mathbb{R})$ for which $d(ga_t^{-1}, ha_t^{-1})$ stays bounded as $t \to -\infty$.

2. Local Coordinates and Haar measure

Now these stable and unstable horocycle groups provide local coordinates on $\mathrm{SL}_2(\mathbb{R})$: Consider the following **Lie-algebras**

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

of U^-, U^+ and A respectively. Then it holds $\mathfrak{u}^- \oplus \mathfrak{u}^+ \oplus \mathfrak{a} = \mathfrak{sl}_2(\mathbb{R})$. Also, it holds $\mathfrak{u}^- \oplus \mathfrak{a} = \mathfrak{b}$ is the Lie algebra of the Borel subgroup B.

Lemma 2.1 (Local coordinates)**.** *The map*

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

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 $t \to -\infty$).

Proof of Lemma 2.1. 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 Φ restricted on $\mathcal{O}'$ is a diffeomorphism. □

The Lemma 2.1 allows us to consider neighborhoods of the identity in $\mathrm{SL}_2(\mathbb{R})$ and other groups that satisfy certain natural properties for the conjugation action. For instance, the map

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

is also a local diffeomorphism around 0. This yields the following construction/definition:

Definition 2.2 (Rectangular Neighborhood)**.** Given two small enough neighborhoods $\mathcal{O}_{\mathfrak{a}}$, respectively $\mathcal{O}_{\mathfrak{u}^+}$ of 0 in $\mathfrak{a}$, respectively $\mathfrak{u}^+$, the image under the above map yields a *rectangular neighborhood* $\mathcal{O}_B$ of the identity in B, which satisfies that

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

for any $t \leqslant 0$.

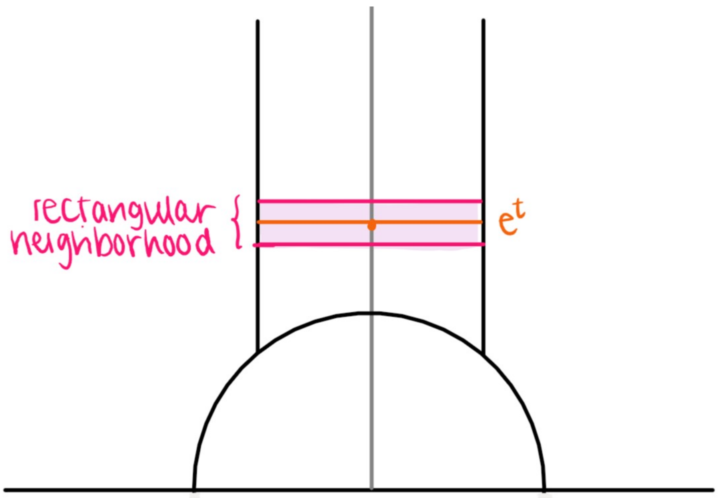

Figure 4. Example of a rectangular neighborhood in the standard fundamental domain.

3. Parametrization of periodic Horocycle Orbits

In this section, we are discussing the periodic orbits of the stable horocycle subgroup U^- on the quotient space X_2. We define the following:

Definition 3.1 (Periodic points and orbits)**.** A point $x \in X_2$ is called *periodic* if there is $s \in \mathbb{R}\backslash\{0\}$ such 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 3.2 (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 $\Phi : U^- \times B \to \mathrm{SL}_2(\mathbb{R})$, $(u, b) \mapsto ub$.*

Proof. For the proof see [3, Lemma 1.5, page 4]. □

Lemma 3.3 (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.

Intuitively, Lemma 3.3 gives us an example of periodic orbits: These are simply the ones that are "horizontal" horocycles in the quotient space X_2. Because $\mathrm{SL}_2(\mathbb{Z})$ is represented by (i, i) in X_2 (see Figure 5 for visualization). If we let a_t act on it we will get the geodesic flow which moves $(i.i)$ along the y-axis. Then letting U^- act on it, we will have the horocycle flow which is parallel to the x-axis. One might ask why the periods of periodic orbits are different when they look like they have the same length in the upper-half plane model, but notice that we are using the hyperbolic distance, not the Euclidean distance.

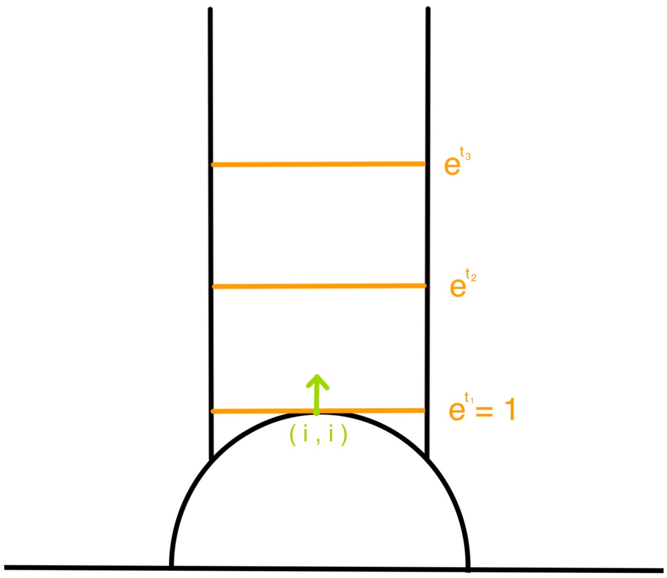

Figure 5. Periodic orbits in the half-plane model of $\mathbb{H}$.

Proof of Lemma 3.3. We first claim 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, take any point $\mathrm{SL}_2(\mathbb{Z})u_s$ in $\mathrm{SL}_2(\mathbb{Z})U^-$.

It is in $\mathrm{SL}_2(\mathbb{Z})\,\mathrm{id}$ if and only if $u_s \in \mathrm{SL}_2(\mathbb{Z})$, which is if and only if $s \in \mathbb{Z}$. Thus, the smallest non-zero $s \in \mathbb{Z}$ (i.e. the period) is 1.
Now let $t \in \mathbb{R}$. Then $\mathrm{SL}_2(\mathbb{Z})a_t u_s = \mathrm{SL}_2(\mathbb{Z})a_t$ if and only if (see Lemma 1.2)

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

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

However, we can do even better: The next proposition states that the periodic orbits in Lemma 3.3 are the only periodic orbits in X_2!

Proposition 3.4 (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 prove the following claim:
Claim: It holds that $a_t.x \to \infty$ as $t \to \infty$. As seen in our first presentation, this is equivalent to saying that for any compact set $K \subseteq X_2$, there exists some T_K such that $a_t.x \notin K$ for all $t \geqslant T_K$ (i.e. the sequence leaves any compact set at some time). Moreover, let S be the period of x. Then $a_t.x$ is also periodic for U^- and has period Se^{-t} because

$$u_s a_t.x = a_t(a_t^{-1} u_s a_t).x = a_t u_{se^t}.x = a_t.x$$

if and only if se^t is a multiple of the period S.
Proof of the Claim: By contradiction, suppose that $a_t.x \not\to \infty$ as $t \to \infty$. Then there exists a compact set K and a sequence $(t_n)_n$ with $t_n \to \infty$ for $n \to \infty$ such that $a_t.x \in K$ for all $n \in \mathbb{N}$. Let now $r > 0$ be a uniform injectivity radius on K (we have seen that this exists in our last talk). So for any $u_s \in U^- \cap B_r(\mathrm{id})$ and any $n \in \mathbb{N}$ we therefore have

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

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. This is a contradiction, which ends the proof of the claim.
To see how the claim implies the proposition, remember that orbits of the geodesic flow (i.e. the geodesics) are either vertical lines (parallel to the y-axis) or half-circles centered on the x-axis. Let $(z_t, v_t) \in F$ be the point corresponding to $a_t.x$, where $F \subseteq \mathrm{T}^1\mathbb{H}$ is the standard fundamental domain in X_2 (see our last talk). This means we are simply looking at what happens to the corresponding point in the fundamental domain.
Let K' be the set of points in F with imaginary part $\leqslant 1$. Note that the image K of K' in $\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{T}^1\mathbb{H}$ is compact. Thus, let $T_K > 0$ so that $a_t.x \notin K$ for all $t \geqslant T_K$ as in the claim. We claim that 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 thus z_t would reach the imaginary part $\leqslant 1$ for some $t > T_K$ (see Fig6).

Applying any $u \in U^-$ 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 $a_{t'}.(ua_{T_K}.x) = \mathrm{SL}_2(\mathbb{Z})id$ (i.e. transporting back to $(i,i) \in F$). 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.

□

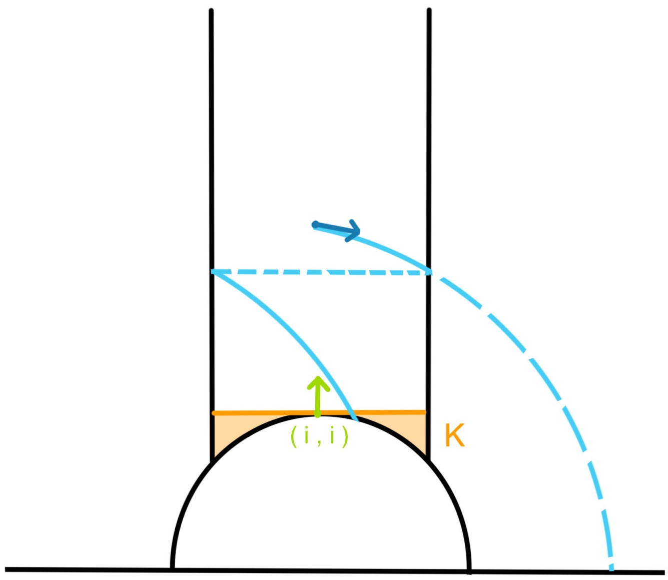

Figure 6. Orbit of a point (z_{T_K}, v_{T_K}) with the error v_{T_K} not a multiple of i in the upper half-plane model

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 $\frac{1}{T}\int_0^T f(u_s.x)ds$ for $f \in C_c(X_2)$ defines a linear functional (and hence a measure) with the required properties.

Theorem 4.1 (Sarnak, 1981 [2])**.** *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 $n \to \infty$.*

The main idea of the proof of Theorem 4.1 is to "thicken" the periodic orbits which we have seen in Proposition 3.4. This happens by choosing some ε "rectangular neighborhoods" around each of these orbits, making them "thicker" (see Figure 7 for intuition). In the end, we let them go to a line (i.e. the ε-neighborhoods to zero), finalizing the proof.

Proof of Theorem 4.1. 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 $\mathrm{SL}_2(\mathbb{R}) \longrightarrow X_2$ is 1-Lipschitz, there is a $\delta > 0$ so that

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

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 $\mathcal{O}_B \subseteq B \cap B_\delta^{\mathrm{SL}_2(\mathbb{R})}(\mathrm{id})$ be a rectangular neighborhood of the identity as introduced in Definition 2.2 so that

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

for all $t \geqslant 0$. (You can check this inclusion if you insert the matrices and calculate them, $\mathcal{O}$ is of the form B.) Moreover, let

$$\tilde{P}_0 = \mathcal{O}.P_0$$

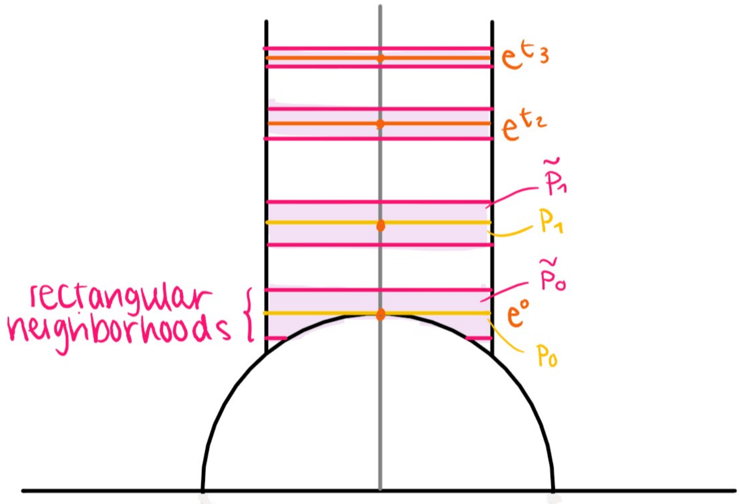

Figure 7. The "thickened orbits" used in the proof of Theorem 4.1.

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 thinner in the unstable direction as $t \to \infty$ and do not get thicker in any direction. For convenience we also define

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

These sets do not have a geometric visualization, they are simply practical for computations. Note also 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 a 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

$$\left| e^t \int_0^{e^t} f(\mathrm{SL}_2(\mathbb{Z})a_t u_s) ds - \frac{e^{-t}}{m_B^{(r)}(a_{-t}\mathcal{O}_B a_t)} \int_{a_{-t}\mathcal{O}_B a_t} \int_0^{e^t} f(\mathrm{SL}_2(\mathbb{Z})a_t u_s b) ds dm_B^{(r)}(b) \right|$$
$$\leqslant \frac{e^{-t}}{m_B^{(r)}(a_{-t}\mathcal{O}_B a_t)} \int_{a_{-t}\mathcal{O}_B a_t} \int_0^{e^t} |f(\mathrm{SL}_2(\mathbb{Z})a_t u_s) - f(\mathrm{SL}_2(\mathbb{Z})a_t u_s b)| ds dm_B^{(r)}(b) < \varepsilon \quad (3)$$

since $a_{-t}\mathcal{O}_B a_t \subseteq \mathcal{O}_B$ for any $t > 0$ by the choice of the neighborhood $\mathcal{O}_B$ (according to the uniform continuity). By Lemma 3.2 the normalized integral

$$\frac{e^{-t}}{m_B^{(r)}(a_{-t}\mathcal{O}_B a_t)} \int_{a_{-t}\mathcal{O}_B a_t} \int_0^{e^t} f(\mathrm{SL}_2(\mathbb{Z})a_t u_s b) ds dm_B^{(r)}(b) \quad (4)$$

is equal to

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

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 (4) is equal to

$$\frac{1}{m_{\mathrm{SL}_2(\mathbb{R})}(S_0)} \int_{S_0} f(\mathrm{SL}_2(\mathbb{Z}) g a_t) dm_{\mathrm{SL}_2(\mathbb{R})}(g). \tag{5}$$

Integral over the 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 $\tilde{P}_0$.

Claim 4.2. *If δ is small enough, the set $S_0 = \{u_s b \mid s \in [0,1), b \in \mathcal{O}_B\}$ is injective.*

Therefore, the Haar measure on S_0 equals the Haar measure on $\tilde{P}_0$. We will look at the proof of the Claim a bit later. However, with it, the integral in (5) is equal to

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

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 as we have seen in the two previous presentations. That is, for any $f_1, f_2 \in L^2(X_2)$ we have

$$\langle a_t.f_1, f_2 \rangle \longrightarrow \int_{X_2} f_1 dm_{X_2} \int_{X_2} f_1 dm_{X_2}$$

as $t \longrightarrow \pm\infty$. In particular

$$\langle a_t.f, f_0 \rangle \longrightarrow \int_{X_2} f dm_{X_2} \int_{X_2} f_0 dm_{X_2} = \int_{X_2} f dm_{X_2}$$

as $t \longrightarrow \infty$. Tracing back our arguments (i.e. combining all equations (3), (4) and (5)), we can deduce that the integral

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

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

$$e^{-t} \int_0^{e^t} f(\mathrm{SL}_2(\mathbb{Z}) a_t u_s) ds \longrightarrow \int_{X_2} f dm_{X_2}$$

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

Proof of the Claim 4.2. 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 $b = id$. This shows that $s_1 - s_2 \in \mathbb{Z}$ and thus $s_1 m = s_2$ as $s_1, s_2 \in [0,1)$

□

Proposition 4.3. *Every hyperbolic circle (i.e. a boundary of a ball in the hyperbolic plane) is an Euclidean circle.*

If you are interested, in [3, page 8] they do a much more technical proof. Here we show that also every Euclidean circle is a hyperbolic circle.

Proof of Proposition 4.3. The key fact you need is that Möbius transformations of $\mathbb{C}\cup\{\infty\}$ preserve all circles and lines. You can prove this by the fact that all orientation preserving Möbius transformations are generated by $N(z) = \frac{1}{z}$, $T_b(z) = z + b$, $M_a(z) = az$ for $a, b \in \mathbb{C}$, with $a \neq 0$ which all preserve lines and circles.
So the group of Möbius transformations of the upper half-plane (denoted H^+) preserves the set of Euclidean circles in H^+ (the only lines in H^+ are horizontal and transform to themselves or to Euclidean circles tangent to the real line, none of which is a Euclidean circle entirely contained in H^+). Also, hyperbolic circles are preserved under that action, since that action is the same as the group of orientation-preserving isometries of the hyperbolic metric on H^+.
There exists a Möbius transformation that transforms the Poincaré disc to the upper half plane, which is called the *Cayley transformation*:

$$\phi : B^1 \longrightarrow H^+, \quad z \longmapsto \frac{z-i}{z+i}$$

This Cayley transformation takes Euclidean circles to Euclidean circles (as there are no lines in the Poincaré disc), and it takes hyperbolic circles to hyperbolic circles (it is an isometry between the hyperbolic metrics on the Poincaré disc and the upper half plane). So, we have reduced the problem to showing that in the Poincaré disc with the hyperbolic metric, hyperbolic circles are the same as Euclidean circles. The group of Möbius transformations on the Poincaré disc equals the group of orientation-preserving isometries of the hyperbolic metric on the Poincaré disc, and this action is transitive on points. Hence, for all $r > 0$ the action is transitive on the set of hyperbolic circles of hyperbolic radius r. Thus, it suffices to find, for each $r > 0$, a single example of a hyperbolic circle of hyperbolic radius r which is simultaneously a Euclidean circle: There is a Euclidean circle centered at the origin which is a circle of hyperbolic radius r. □

Note that in the next theorem, the paper [3, page 9] switched the coset. We leave it this way, but just so you notice.

Theorem 4.4. *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*

$$\frac{1}{m_Y(Y)} \int_Y f(a_t.y)\, dm_Y(y) \longrightarrow \frac{1}{m_X(X)} \int_X f(x)\, dm_X(x)$$

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 $t \longrightarrow \infty$. The statement for $t \longrightarrow -\infty$ is the same just with arrows pointing inwards. See Figure 8 for intuition.

Proof. We restrict our attention to the case $t \longrightarrow \infty$ as the case $t \longrightarrow -\infty$ is analogous. Let $f \in C_c(X)$ abd let $\varepsilon > 0$. Denote $G = \mathrm{SL}_2(\mathbb{R})$ and $K = \mathrm{SO}(2)$. Let $O \in U^- A$ be an open (rectangular) neighborhood of the identity with $a_t O a_{-t} \subset O$ for any $t > 0$ (a

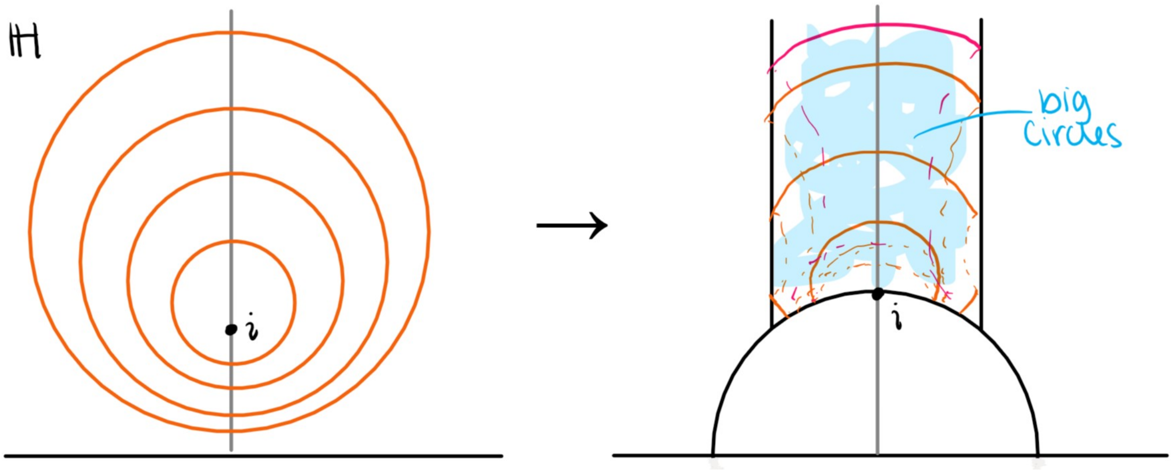

Figure 8. Intuition for Theorem 4.4: Taking the integral of a function over a very large horocycle is approximately the same as taking the integral over the whole fundamental domain.

contracted neighborhood) so that $f(g.x)$ is ε-close to $f(x)$ for any $g \in O$ and any $x \in X$. By these choices the integral $\frac{1}{m_Y(Y)} \int_Y f(a_t.y)\, dm_Y(y)$ is ε-close to

$$I_t = \frac{1}{m_{UA}(O)} \frac{1}{m_Y(Y)} \int_O \int_Y f(a_t.g.y)\, dm_Y(y)\, dm_{U^-}(g)$$

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

$$\frac{1}{m_{UA}(O)} \frac{1}{m_K(k)} \int_O \int_F f(a_t.gk\Gamma)\, dm_K(k)\, dm_{U^-}(g)$$

by definition of the Haar measure ob Y. By Lemma 3.2 applied to the Iwasawa decomposition (with U^-A and K)

$$I_t = \frac{1}{m_G(OF)} \int_{OF} f(a_t.g\Gamma)\, dm_G(g),$$

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

References

[1] Manfred Einsiedler and Thomas Ward. *Ergodic theory with a view towards number theory*, volume 259 of *Graduate Texts in Mathematics*. Springer-Verlag London, Ltd., London, 2011. Chapters 9.3/9.4.

[2] Peter Sarnak. Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series. *Comm. Pure Appl. Math.*, 34(6):719–739, 1981.

[3] Andreas Wieser. On the "Banana"-Trick of Margulis. Find on `https://metaphor.ethz.ch/x/2018/fs/401-3370-67L/sc/banana.pdf`, 2018.