Diagonal and unipotent flows

Andreas Strömbergsson
Uppsala University
http://www.math.uu.se/~astrombe

Sarajevo, August 2022

Basic setting

Let G be a Lie group.

Let Γ be a lattice in G.

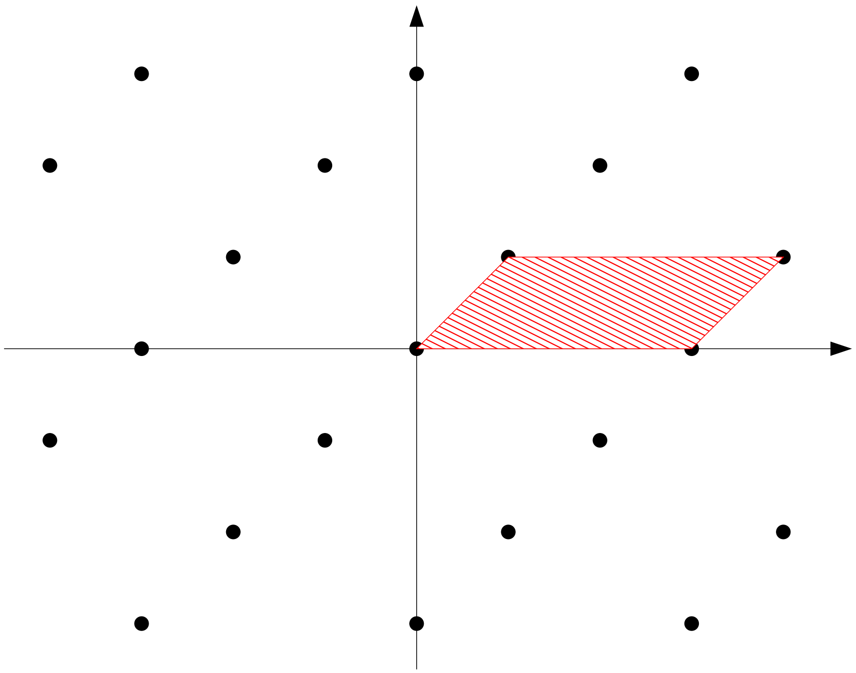
Set $X = \Gamma\backslash G = \{\Gamma g \ : \ g \in G\}$ — our **homogeneous space**.

Let μ be a left Haar measure on G; then μ induces a G-invariant measure on $\Gamma\backslash G$ which we *also* call "μ". From now on we assume that μ is normalized so that $\mu(\Gamma\backslash G) = 1$.

– see $\boxed{\text{Problem 2}}$

Example: $G = \mathbb{R}^d$, $\Gamma = L$, a lattice in $\mathbb{R}^d$.

That is: $L = \{c_1\vec{v}_1 + \cdots + c_d\vec{v}_d \ : \ c_1, \ldots, c_d \in \mathbb{Z}\}$ for some $\vec{v}_1, \ldots, \vec{v}_d \in \mathbb{R}^d$ which form a basis of $\mathbb{R}^d$.



Then $X = \Gamma\backslash G = L\backslash\mathbb{R}^d = \mathbb{R}^d/L$, a *torus*.

μ = normalized Lebesgue measure.

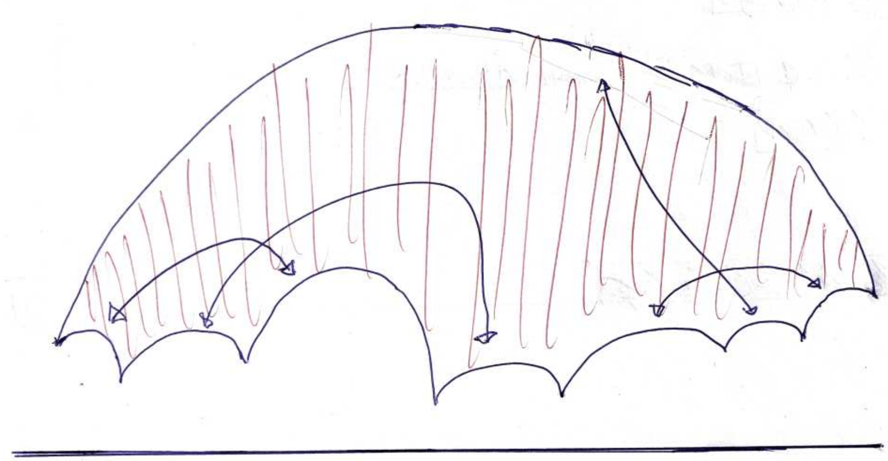
Example: $G = \mathrm{SL}_2(\mathbb{R})$ and Γ a lattice in G with $-I \in \Gamma$, and without elliptic elements.

Recall that $\mathrm{SL}_2(\mathbb{R})$ acts by isometries on the hyperbolic upper half space:

$$\mathbf{H} = \{z = x + iy \;:\; x, y \in \mathbb{R},\ y > 0\}, \quad \text{with metric } ds = \frac{\sqrt{dx^2 + dy^2}}{y}.$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} (z) = \frac{az + b}{cz + d}.$$

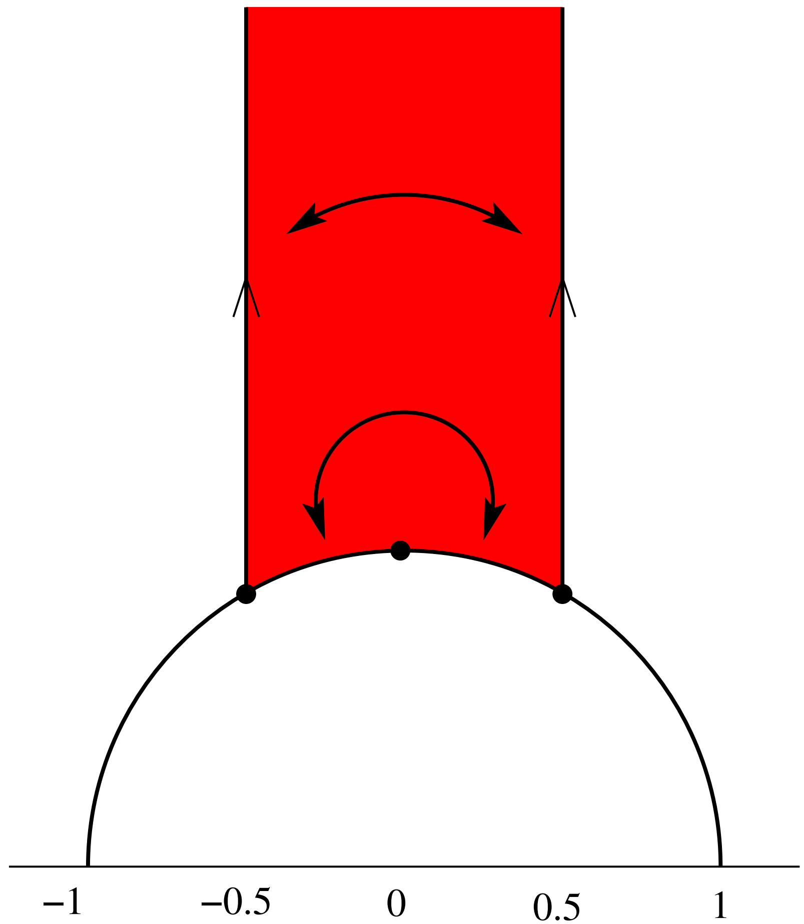
Now $\Gamma\backslash\mathbf{H}$ is a hyperbolic surface of finite area, and $\boxed{X = \Gamma\backslash G = T^1(\Gamma\backslash\mathbf{H})}$, viz., $\Gamma\backslash G$ can be identified with the unit tangent bundle of $\Gamma\backslash\mathbf{H}$.



– see $\boxed{\text{Problem 1}}$

For $G = \mathrm{SL}_2(\mathbb{R})$, $\boxed{\Gamma = \mathrm{SL}_2(\mathbb{Z})}$

$X = \Gamma\backslash G = T^1(\Gamma\backslash \mathbf{H})$, where $\Gamma\backslash \mathbf{H}$ looks as follows:



Example: $G = \mathrm{SL}_d(\mathbb{R})$ and $\Gamma = \mathrm{SL}_d(\mathbb{Z})$ $(d \geqslant 2)$. Then

$X = \Gamma\backslash G = \mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R}) =$ **[the space of lattices in $\mathbb{R}^d$ of covolume 1]**.

Identification map: $\Gamma g \mapsto \mathbb{Z}^d g = \{\vec{x}g \,:\, \vec{x} \in \mathbb{Z}^d\} \quad (g \in G)$.

X is **noncompact.**

Mahler's criterion: A sequence $\Gamma g_1, \Gamma g_2, \ldots$ in $\Gamma\backslash G$ *diverges*
($\overset{\text{def}}{\iff}$ leaves every compact subset of $\Gamma\backslash G$)
if and only if

$$m(\Gamma g_j) := \min\{\|\vec{x}g_j\| \,:\, \vec{x} \in \mathbb{Z}^d\backslash\{\vec{0}\}\} \quad \to 0 \quad \text{as } j \to \infty.$$

Next: DYNAMICS on the homogeneous space $X = \Gamma \backslash G$

Let $(h_t)_{t \in \mathbb{R}}$ be a *1-parameter subgroup of G.*

(That is, the map $t \mapsto h_t$ is a Lie group homomorphism. Recall that in this situation there exists a unique element $X \in \mathfrak{g}$ such that $h_t = \exp(tX)$, $\forall t \in \mathbb{R}$.)

Objective: Understand properties of the **"homogeneous flow"**

$$\Phi_t(\Gamma g) := \Gamma g h_t \qquad \text{on } X = \Gamma \backslash G.$$

(Note that this flow preserves μ.)

FLOWS (a basic notion in dynamical systems)

A **flow** on a set X is a map $\Phi : \mathbb{R} \times X \to X$ such that

$$\forall x \in X,\ t, s \in \mathbb{R}: \qquad \Phi(0, x) = x \quad \text{and} \quad \Phi(s, \Phi(t, x)) = \Phi(s + t, x).$$

We will write $\boxed{\Phi_t(x)}$ in place of $\Phi(t, x)$. Then:

$$\boxed{\forall x \in X,\ t, s \in \mathbb{R}: \qquad \Phi_0(x) = x \quad \text{and} \quad \Phi_s(\Phi_t(x)) = \Phi_{s+t}(x).}$$

Note: For each $t \in \mathbb{R}$, Φ_t is a bijection $X \xrightarrow{\sim} X$.

Compare with a (bijective) **map** $T : X \to X$; then study $T^{\circ n} : X \to X$ for $n \in \mathbb{Z}$ — for a **flow** we take instead "$n \in \mathbb{R}$".

Usually X has extra structure which Φ preserves.

Ex: X a topological space, and $\Phi : \mathbb{R} \times X \to X$ *continuous.*

Ex: X a C^∞ manifold, and $\Phi : \mathbb{R} \times X \to X$ is C^∞.

We say that the flow Φ **preserves a measure** $\mu \in P(X)$ if $\Phi_{t*}(\mu) = \mu$ for all $t \in \mathbb{R}$ (that is, $\mu(\Phi_{-t}(A)) = \mu(A)$ for every measurable $A \subset X$).

In this situation, μ **is ergodic** if, for every measurable subset $A \subset X$ satisfying $\Phi_t(A) = A$, $\forall t \in \mathbb{R}$, one has $\mu(A) = 0$ or 1.

Birkhoff's **Pointwise Ergodic Theorem** for flows: Let $(X, \mathcal{B}, \mu)$ be a probability space; and let $\Phi : \mathbb{R} \times X \to X$ be a measurable flow preserving μ. Let $f : X \to \mathbb{C}$ be measurable and $\int_X |f| \, d\mu < \infty$. Set

$$A_T^f(x) := \frac{1}{T} \int_0^T f(\Phi_t(x)) \, dt \qquad (T > 0).$$

Then the function A_T^f converges μ-almost everywhere, and also in L^1-norm, to a function $F : X \to \mathbb{C}$ with $F \circ \Phi_t \equiv F$ $(\forall t \in \mathbb{R})$.

DYNAMICS on the homogeneous space $X = \Gamma\backslash G$

Let $(h_t)_{t\in\mathbb{R}}$ be a *1-parameter subgroup of* G.

Objective: Understand properties of the **"homogeneous flow"**

$$\Phi_t(\Gamma g) := \Gamma g h_t \qquad \text{on } X = \Gamma\backslash G.$$

(Note that this flow preserves μ.)

Examples: Let $G = \mathrm{SL}_2(\mathbb{R})$.

Let $a_t \equiv \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$: $\quad \Phi_t(\Gamma g) := \Gamma g a_t$ is the **geodesic flow** on $T^1(\Gamma\backslash\mathbf{H})$.

Let $u_t \equiv \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$: $\quad \Phi_t(\Gamma g) := \Gamma g u_t$ is the **horocycle flow** on $T^1(\Gamma\backslash\mathbf{H})$.

– see $\boxed{\text{Problem 1}}$

For general $\Gamma\backslash G$ and (h_t), the behavior of the flow $\Phi_t(\Gamma g) = \Gamma g h_t$ can differ hugely, depending on (h_t) (and on G)! Two important, very different, cases:

$$\boxed{(h_t)\ \textbf{unipotent}} \overset{\text{def}}{\iff} \Big[\ \mathrm{Ad}_{h_t} \text{ is unipotent } (\forall t \in \mathbb{R})\ \Big].$$

Examples: The horocycle flow: $G = \mathrm{SL}_2(\mathbb{R})$ and $u_t \equiv \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$.

More generally, $G = \mathrm{SL}_d(\mathbb{R})$ and any $u_t \equiv \begin{pmatrix} 1 & * & * & \cdots & * \\ 0 & 1 & * & \cdots & * \\ 0 & 0 & 1 & & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$.

Also: $G = \mathbb{R}^d$ ($\Rightarrow X$ a torus, and all homogeneous flows are *linear*).

$\boxed{(h_t)\ \textbf{(}\mathbb{R}\textbf{-)diagonal}} \overset{\text{def}}{\Longleftrightarrow} \left[\begin{array}{l} \mathrm{Ad}_{h_t} \text{ is diagonalizable over } \mathbb{R}\ (\forall t \in \mathbb{R}) \text{, and } G \text{ is a} \\ \textit{simple} \text{ Lie group with finite center.} \end{array}\right]$

((Or more generally, G *semisimple* + extra conditions.))

Examples: The geodesic flow: $G = \mathrm{SL}_2(\mathbb{R})$ and $a_t \equiv \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$.

More generally, $G = \mathrm{SL}_d(\mathbb{R})$ and $a_t \equiv \begin{pmatrix} e^{c_1 t} & 0 & \cdots & 0 \\ 0 & e^{c_2 t} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{c_d t} \end{pmatrix}$.

Summary of properties

Diagonal flows are **“chaotic”**:

— they are *Bernoulli*,
have *positive entropy*,
non-zero Lyapunov exponents ($\Rightarrow$ are *partially hyperbolic*),
and are *exponentially mixing.*

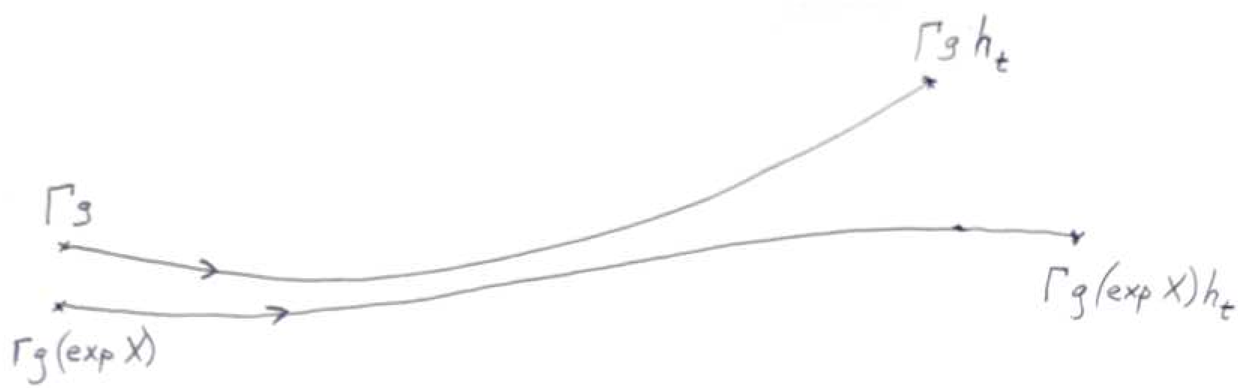
Unipotent behave in a much more controlled way:

— they have *entropy zero* ($\Rightarrow$ they are **“deterministic”**),
they have *all Lyapunov exponents* $= 0$,
they are *at most polynomially mixing* (see Problem 4(c)),
and they exhibit *measure ridigity* (Ratner’s Theorem).

Immediate consequences of the definitions (unipotent/diagonal)

Note, for any $g \in G$ and $X \in \mathfrak{g}$:

$$\Gamma g(\exp X)h_t = \Gamma g h_t \exp\Big(\mathrm{Ad}_{h_t}^{-1}(X)\Big).$$



Therefore,

$$(h_t)\begin{Bmatrix}\text{unipotent}\\ \text{diagonal}\end{Bmatrix} \quad\Rightarrow\quad \begin{Bmatrix}\textbf{polynomial}\\ \textbf{exponential}\end{Bmatrix}\ \textbf{divergence of trajectories.}$$

– see $\boxed{\text{Problem 4}}$

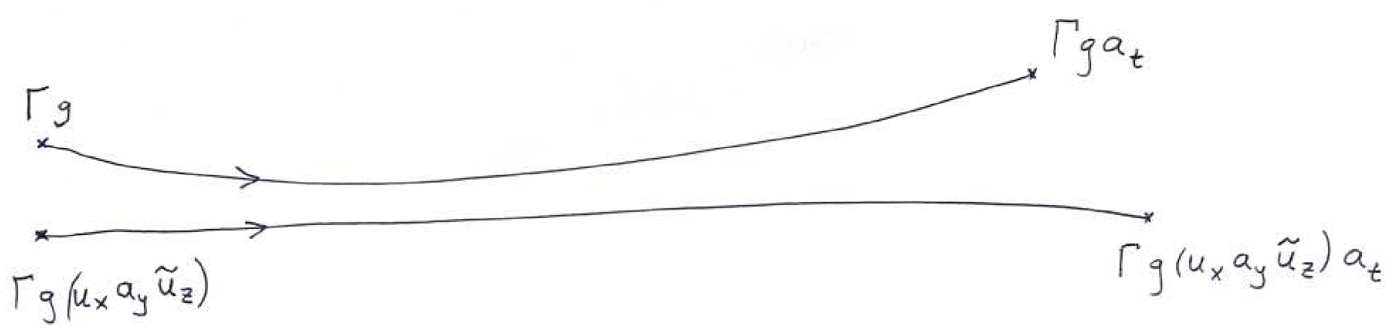
More concretely, and explicitly, for $G = \mathbf{SL}_2(\mathbb{R})$

Write $a_y = \begin{pmatrix} e^{y/2} & 0 \\ 0 & e^{-y/2} \end{pmatrix}$, $\quad u_x = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$, $\quad \widetilde{u}_z = \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix}$.

Then $\langle x, y, z\rangle \mapsto u_x a_y \widetilde{u}_z$, $\mathbb{R}^3 \to G$, is a C^∞ parametrization of
$G_+ := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G \ : \ d > 0 \right\}$. $\qquad$ (Note also $\langle 0,0,0\rangle \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.)

Key relations: $\boxed{u_x a_t = a_t u_{xe^{-t}}}$ and $\boxed{\widetilde{u}_z a_t = a_t u_{ze^t}}$

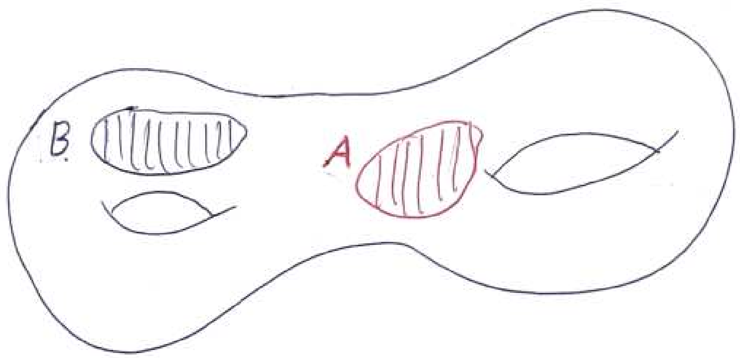
It follows that $\qquad \Gamma g \left(u_x a_y \widetilde{u}_z\right) a_t \equiv \Gamma g \, a_t \left(u_{xe^{-t}} a_y \widetilde{u}_{ze^t}\right)$.



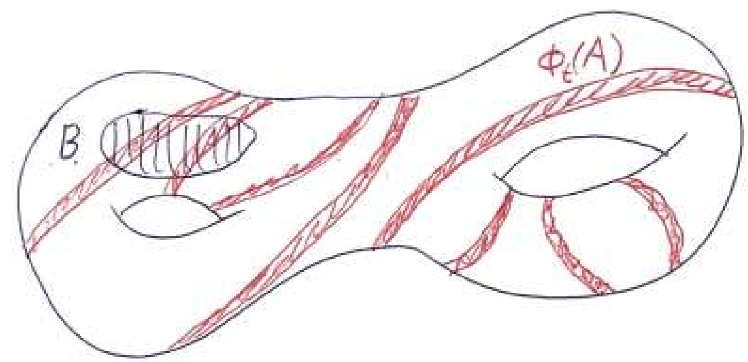
Mixing (a basic notion in dynamical systems)

Definition: A measure-preserving flow (Φ_t) on a probability space (X, μ) is said to be **(strongly) mixing** if

(i) for any $A, B \subset X$, $\lim_{t\to\infty} \mu\big(\Phi_t(A) \cap B\big) = \mu(A)\mu(B)$,



$\cdots \longrightarrow$



or equivalently

(ii) $\forall f_1, f_2, \in L^2(X, \mu)$: $\lim_{t\to\infty}\langle f_1 \circ \Phi_t^{-1}, f_2\rangle = \mu(f_1)\mu(f_2)$.

(Proof of (i)$\Leftrightarrow$(ii): "Simple functions are dense in L^2".)

Note: **mixing** $\Rightarrow$ **ergodic**

Diagonal flows are exponentially mixing

Theorem (work by *many* people):

Let G be a simple Lie group with finite center, and let Γ be a lattice in G. (Example: $G = \mathrm{SL}_d(\mathbb{R})$, Γ any lattice in G.)

Set $X = \Gamma\backslash G$, and let (a_t) be a non-trivial, $\mathbb{R}$-**diagonal** one-parameter subgroup of G. Then $\exists \ell \in \mathbb{Z}^+$, $C > 0$, $\eta > 0$ such that

$\forall f_1, f_2 \in C_c^\infty(X)$, $t \geqslant 0$:

$$\left| \int_X f_1(xa_{-t}) f_2(x)\, d\mu(x) - \mu(f_1)\mu(f_2) \right| \leqslant C \cdot S_{2,\ell}(f_1) S_{2,\ell}(f_2) \cdot e^{-\eta t}.$$

(On the other hand, a unipotent flow can be *at most polynomially mixing* — see Problem 4(c).)

Application: Equidistribution of expanding horocycles in $\mathbf{SL}_2(\mathbb{R})$ ("Margulis' thickening technique")

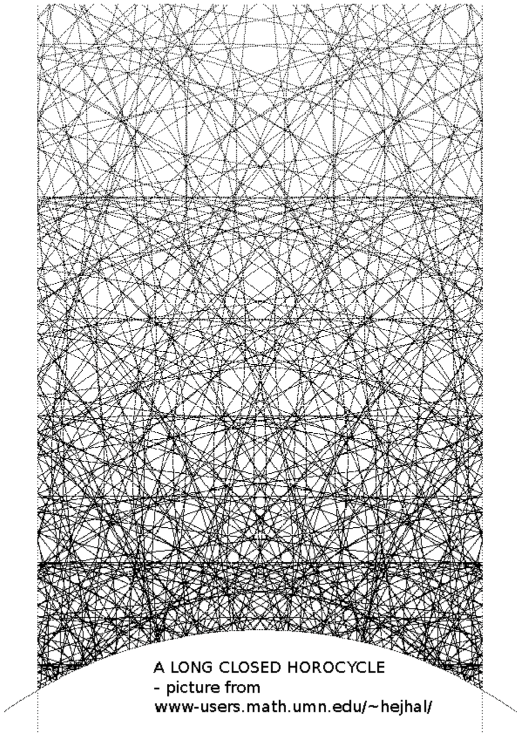
Theorem: Let $G = \mathrm{SL}_2(\mathbb{R})$, let Γ be a lattice in G, and set $X = \Gamma\backslash G$. Then for any $p_0 \in X$ and $f \in C_c(X)$:

$$\int_0^1 f(p_0 u_s a_t)\, ds \to \mu(f) \qquad \text{as } t \to -\infty.$$

Remarks:

• For $p_0 = \Gamma \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and Γ with $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma$:
Equidistribution of long closed horocycles,
very classical (Selberg, Zagier, Sarnak, Hejhal, ...)

• the theorem is proved using the fact that (a_t) is mixing. Using **exponential** mixing, one also obtains an **error term.**



Outline of proof: (See $\boxed{\text{Problem 5}}$ for details.)

Naively, we want to apply the fact that (a_t) is **mixing**, with

$$f_2 \equiv f \quad \text{and} \quad f_1 = \big[\text{ 1-dim Lebesgue measure along } \{p_0 u_s \,:\, 0 \leqslant s \leqslant 1\}\,\big].$$

However, that f_1 is *not permitted!*

To fix up, instead take f_1 to be the *characteristic function of*

$$H_{p_0,\varepsilon} := \big[\ \varepsilon\text{-neighbourhood of } \big\{p_0 u_s \,:\, 0 \leqslant s \leqslant 1\big\}\big].$$

Using the fact that (a_t) is **mixing**, we get:

$$\lim_{t\to-\infty} \int_X f_1(xa_{-t}) f_2(x)\, d\mu(x) = \mu(f_1)\mu(f_2)$$

$$\Rightarrow \qquad \lim_{t\to-\infty} \frac{1}{\mu(f_1)} \int_X f_1(x) f_2(xa_t)\, d\mu(x) = \mu(f_2)$$

$$\Rightarrow \qquad \lim_{t\to-\infty} \frac{1}{\mu(H_{p_0,\varepsilon})} \int_{H_{p_0,\varepsilon}} f_2(xa_t)\, d\mu(x) = \mu(f_2).$$

In the last integral, use the fact that $H_{p_0,\varepsilon}\, a_t$ is contained in an ε-neighbourhood of

$$\left\{ p_0 u_s a_t \;:\; 0 \leqslant s \leqslant 1 \right\},$$

because of

$$p_0 u_s \left(a_y \widetilde{u}_z \right) a_t = p_0 u_s a_t \left(a_y \widetilde{u}_{ze^{-t}} \right).$$

(Recall $t \to -\infty$.)

Therefore,

$$\frac{1}{\mu(H_{p_0,\varepsilon})} \int_{H_{p_0,\varepsilon}} f_2(x a_t)\, d\mu(x) \approx \int_0^1 f_2(p_0 u_s a_t)\, ds,$$

and we are done. □

Measure rigidity for unipotent flows (Ratner)

Theorem (Ratner, 1991): Let G be a Lie group and Γ a lattice in G, and let $\Phi_t(\Gamma g) = \Gamma g u_t$ be a **unipotent** flow on $X = \Gamma\backslash G$. Then every

$$\Phi_t\text{-invariant ergodic } \nu \in P(X)$$

is **"homogeneous"** ($\Leftrightarrow$ **"algebraic"**),

meaning that there exist $x \in X$ and a closed connected subgroup $S \subset G$ such that $\{u_t\} \subset S$, $xS = \overline{\{\Phi_t(x) \;:\; t \in \mathbb{R}\}}$, $\nu(xS) = 1$ and ν is S-invariant.

In the above situation, it follows that ν is the *unique* S-invariant probability measure on xS; and also that $\Phi_{\mathbb{R}}(x)$ is *equidistributed* in xS.

Also, xS is a *closed regular submanifold* of X. Explicitly, take $g \in G$ such that $x = \Gamma g$, and set $\widetilde{S} = gSg^{-1}$, $\widetilde{\Gamma} = \Gamma \cap \widetilde{S}$ and $\widetilde{X} = \widetilde{\Gamma}\backslash\widetilde{S}$. Then $\widetilde{\Gamma}$ is a *lattice* in $\widetilde{S}$, and the map

$$J : \widetilde{X} \to X, \qquad J(\widetilde{\Gamma}\widetilde{s}) := \Gamma\widetilde{s}g \qquad (\widetilde{s} \in \widetilde{S})$$

is a C^∞ diffeomorphism of $\widetilde{X}$ onto xS, mapping the $\widetilde{S}$-invariant measure of $\widetilde{X}$ onto ν.