Ergodic and mixing properties of horocycle flows and their time-changes

Simons Semester – Lecture 5

Davide Ravotti

University of Vienna

31 May 2023

Plan for today

Today we will talk about some rigidity problems. In particular, we will

- discuss the isomorphism problem for time-changes of horocycle flows,
- take a look at Ratner's Rigidity Theorems.

Isomorphisms of time-changes of horocycle flows

The setting

- $G = \mathrm{SL}_2(\mathbb{R})$ and $\Gamma \leq G$ is discrete,
- $M = \Gamma \backslash G$ is compact,
- vol is the Haar measure on M, normalized so that $\mathrm{vol}(M) = 1$.
- $\mathbf{u} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathfrak{g}$ generates the horocycle flow

$$h_t = \varphi_t^{\mathbf{u}} \colon \Gamma g \mapsto \Gamma g \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

- $\mathbf{x} = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \in \mathfrak{g}$ generates the geodesic flow

$$\varphi_t^{\mathbf{x}} \colon \Gamma g \mapsto \Gamma g \begin{pmatrix} e^{\frac{t}{2}} & 0 \\ 0 & e^{-\frac{t}{2}} \end{pmatrix}.$$

The setting

- $\alpha \colon M \to \mathbb{R}_{>0}$ is a smooth positive function, normalized so that $\int_M \alpha \, \mathrm{d}\mathrm{vol} = 1$,
- $(h_t^\alpha)_{t\in\mathbb{R}}$ is the time-change induced by α, namely the flow generated by the vector field $U_\alpha := \alpha^{-1} U$,
- $\tau(t,p)$ is the additive cocycle for h_t^α such that

$$h_t^\alpha(p) = h_{\tau(t,p)}(p),$$

- vol^α is the measure equivalent to vol with density α, which is invariant for $(h_t^\alpha)_{t\in\mathbb{R}}$.

The isomorphism problem

A natural question is whether a given time-change $(h_t^\alpha)_{t\in\mathbb{R}}$ is isomorphic to the standard horocycle flow $(h_t)_{t\in\mathbb{R}}$; namely, whether there exists a measurable automorphism $\psi\colon M \to M$ such that

$$\psi_*(\mathrm{vol}^\alpha) = \mathrm{vol},$$

and

$$\psi \circ h_t^\alpha(p) = h_t \circ \psi(p), \qquad \text{for all } t \in \mathbb{R} \text{ and } p \in M.$$

The isomorphism is said to be continuous, $\mathscr{C}^k$, smooth, if the automorphism ψ is continuous, $\mathscr{C}^k$, smooth.

Coboundaries

Definition

- A function $f\colon M \to \mathbb{R}$ is a (measurable) *coboundary* for the horocycle flow if there exists a measurable function g, called the transfer function, such that $f = Ug$.
- A function $f\colon M \to \mathbb{R}$ is a (measurable) *almost coboundary* for the horocycle flow if there exists a measurable function g, called the transfer function, and a constant c such that $f - c = Ug$.

As an exercise, one can verify the following facts:

- If f is an almost coboundary, then, in the definition above, we have $c = \int_M f \, \mathrm{d}\mathrm{vol}$,
- If f is a coboundary, any two transfer functions differ by a constant.

Almost coboundaries

Lemma

If the function α inducing the time-change $(h_t^{\alpha})_{t\in\mathbb{R}}$ is an almost coboundary, then $(h_t^{\alpha})_{t\in\mathbb{R}}$ and $(h_t)_{t\in\mathbb{R}}$ are isomorphic. The regularity of the isomorphism is the same as the regularity of any transfer function.

Proof.

Let us assume that $\alpha - 1 = U\beta$ for some function $\beta \colon M \to M$. We define

$$\psi(p) = h_{\beta(p)}(p),$$

and we now verify it is an isomorphism between $(h_t^{\alpha})_{t\in\mathbb{R}}$ and $(h_t)_{t\in\mathbb{R}}$.

Proof of the lemma

- From the property of $\tau(t,p)$, we have

$$t = \int_0^{\tau(t,p)} \alpha \circ h_r(p)\,\mathrm{d}r = \int_0^{\tau(t,p)} (1 + U\beta) \circ h_r(p)\,\mathrm{d}r$$
$$= \tau(t,p) + \int_0^{\tau(t,p)} \frac{\mathrm{d}}{\mathrm{d}r}\beta \circ h_r(p)\,\mathrm{d}r = \tau(t,p) + \beta(h_t^\alpha(p)) - \beta(p).$$

- For any $t \in \mathbb{R}$ and $p \in M$, we deduce

$$\psi \circ h_t^\alpha(p) = h_{\beta(h_t^\alpha(p))}(h_t^\alpha(p)) = h_{\beta(h_t^\alpha(p)) + \tau(t,p)}(p)$$
$$= h_{t+\beta(p)}(p) = h_t \circ \psi(p).$$

Proof of the lemma

- Thinking of $(h_t)_{t\in\mathbb{R}}$ as a time-change of $(h_t^\alpha)_{t\in\mathbb{R}}$ induced by α^{-1}, and repeating the same argument, we obtain

$$\psi^{-1}(p) = h^\alpha_{-\beta(p)}(p).$$

Hence ψ is a measurable automorphism of M.

- By unique ergodicity, for every $f \in \mathscr{C}(M)$ and $p \in M$,

$$\begin{aligned}\mathrm{vol}(f) &= \lim_{t\to\infty} \frac{1}{t}\int_0^t f \circ h_r(\psi(p))\,\mathrm{d}r = \lim_{t\to\infty} \frac{1}{t}\int_0^t (f\circ\psi)\circ h_r^\alpha(p)\,\mathrm{d}r \\ &= \mathrm{vol}^\alpha(f\circ\psi) = [\psi_*(\mathrm{vol}^\alpha)](f),\end{aligned}$$

which implies $\psi_*(\mathrm{vol}^\alpha) = \mathrm{vol}$.

Rigidity results

Ratner's Rigidity Theorem

Recall $M = \Gamma \backslash G$, where $G = \mathrm{SL}_2(\mathbb{R})$.
For any $\mathbf{n} \in N_G(\Gamma)$, the map $\Psi_{\mathbf{n}} \colon M \to M$ given by

$$\Psi_{\mathbf{n}}(\Gamma \mathbf{g}) = \Gamma \, \mathbf{n} \, \mathbf{g}$$

is well-defined, since $\mathbf{n} \cdot \Gamma \cdot \mathbf{n}^{-1} = \Gamma$.

Theorem (Ratner 1986)

Assume that there exists a measurable isomorphism $\psi \colon M \to M$ between $(h_t^{\alpha})_{t \in \mathbb{R}}$ and $(h_t)_{t \in \mathbb{R}}$. Then, there exists $\mathbf{n} \in N_G(\Gamma)$ and there exists a measurable function $\beta \colon M \to \mathbb{R}$ such that

$$\psi(p) = h_{\beta(\Psi_{\mathbf{n}}(p))}(\Psi_{\mathbf{n}}(p)), \qquad \text{for all } p \in M.$$

In particular, α is an almost coboundary.

A $\mathscr{C}^1$-rigidity result

For the moment, let us try to prove the following, more modest, result.

Proposition

Assume that there exists a $\mathscr{C}^1$-isomorphism $\psi \colon M \to M$ between $(h_t^\alpha)_{t\in\mathbb{R}}$ and $(h_t)_{t\in\mathbb{R}}$. Then, α is a continuous almost coboundary, namely there exists a continuous function $\beta \colon M \to \mathbb{R}$ such that $\alpha - 1 = U\beta$.

Since we are assuming $\psi \circ h_t^\alpha = h_t \circ \psi$, by the chain rule it follows that

$$D\psi(h_t^\alpha(p)) \cdot Dh_t^\alpha(p) = Dh_t(\psi(p)) \cdot D\psi(p), \qquad \text{for all } p \in M.$$

We will apply the differentials above to the vector field X generating the geodesic flow.

Proof of the proposition

For the sake of notation, we suppress the dependence on the points in the tangent vectors.
We make the following observations:

- by assumption, $D\psi(p)(U_\alpha) = U$ for every $p \in M$;
- for every $W \in \{V, X, U\}$, if we write

$$D\psi(p)(W) = f_{W,V}(p)\, V + f_{W,X}(p)\, X + f_{W,U}(p)\, U,$$

  then $|f_{W,V}|, |f_{W,X}|$, and $|f_{W,U}|$ are uniformly bounded;
- from the first lecture, we recall that $Dh_t \simeq \mathrm{Ad}(\exp(t\mathbf{u}))$ so that

$$Dh_t(X) = X + tU, \qquad \text{and} \qquad Dh_t(V) = V - 2tX - t^2 U.$$

Proof of the proposition

We now compute

$$\begin{aligned}[D(h_t \circ \psi)(p)](X) &= Dh_t(\psi(p)) \cdot D\psi(p)(X) \\ &= Dh_t(\psi(p)) \cdot (f_{X,V}(p)\,V + f_{X,X}(p)\,X + f_{X,U}(p)\,U) \\ &= f_{X,V}(p)(V - 2tX - t^2 U) + f_{X,X}(p)(X + tU) + f_{X,U}(p)U \\ &= f_{X,V}(p)\,V + (-2tf_{X,V}(p) + f_{X,X}(p))\,X \\ &\qquad + (-t^2 f_{X,V}(p) + tf_{X,X}(p) + f_{X,U}(p)\,U.\end{aligned}$$

The expression above is equal to $[D(\psi \circ h_t^{\alpha})(p)](X)$, which we now study.

Proof of the proposition

We start from $Dh_t^\alpha(p)(X)$. Recall that

$$\frac{\mathrm{d}}{\mathrm{d}t} Dh^\alpha_{-t}(p)(X) = [\mathscr{L}_{\alpha^{-1}U}(X)] \circ h^\alpha_{-t}(p)$$
$$= [\alpha^{-1}\mathscr{L}_U(X) - X(\alpha^{-1})U] \circ h^\alpha_{-t}(p)$$
$$= -(\alpha^{-1} + X(\alpha^{-1})) \circ h^\alpha_{-t}(p)\, U$$
$$= -\left(1 - \frac{X\alpha}{\alpha}\right) \circ h^\alpha_{-t}(p)\, U_\alpha.$$

If we write

$$Dh_t^\alpha(p)(X) = v_t(h_t^\alpha(p))\, V + x_t(h_t^\alpha(p))\, X + u_t(h_t^\alpha(p))\, U_\alpha,$$

then we have

$$[v_t(h_t^\alpha(p))]'\, V + [x_t(h_t^\alpha(p))]'\, X + [u_t(h_t^\alpha(p))]'\, U_\alpha = \left(1 - \frac{X\alpha}{\alpha}\right) \circ h_t^\alpha(p)\, U_\alpha.$$

Proof of the proposition

The solution of the previous ODE with the initial conditions $v_0(p) = 0$, $x_0(p) = 1$, and $u_0(p) = 0$ gives us

$$Dh_t^\alpha(p)(X) = X + \left(\int_0^t \left(1 - \frac{X\alpha}{\alpha}\right) \circ h_r^\alpha(p)\, \mathrm{d}r \right) U_\alpha = X + A_t(p) U_\alpha,$$
$$\text{where} \qquad A_t(p) := \int_0^t \left(1 - \frac{X\alpha}{\alpha}\right) \circ h_r^\alpha(p)\, \mathrm{d}r.$$

Therefore, we obtain

$$\begin{aligned}[D(\psi \circ h_t^\alpha)(p)](X) &= D\psi(h_t^\alpha(p))(X + A_t(p)U_\alpha) \\ &= f_{X,V}(h_t^\alpha(p))\, V + f_{X,X}(h_t^\alpha(p))\, X + f_{X,U}(h_t^\alpha(p))\, U + A_t(p)\, U.\end{aligned}$$

Proof of the Proposition

We now equate the two expressions we obtained:

$$\begin{aligned} f_{X,V}(h_t^\alpha(p))\,V + f_{X,X}(h_t^\alpha(p))\,X + f_{X,U}(h_t^\alpha(p))\,U + A_t(p)\,U \\ = f_{X,V}(p)\,V + (-2tf_{X,V}(p) + f_{X,X}(p))\,X \\ + (-t^2 f_{X,V}(p) + tf_{X,X}(p) + f_{X,U}(p))\,U. \end{aligned}$$

Comparing the coefficients yields $f_{X,V}(p) = 0$ and

$$|A_t(p) - tf_{X,X}(p)| \leq K,$$

for some constant K independent of p. In particular,

$$f_{X,X}(p) = \lim_{t\to\infty} \frac{A(p)}{t} = \lim_{t\to\infty} \frac{1}{t} \int_0^t \left(1 - \frac{X\alpha}{\alpha}\right) \circ h_r^\alpha(p)\,\mathrm{d}r = 1.$$

Proof of the Proposition

We deduced that there exists $K > 0$ such that

$$\sup_{p\in M}\sup_{t\in\mathbb{R}} |A_t(p) - t| = \sup_{p\in M}\sup_{t\in\mathbb{R}} \left| \int_0^t \frac{X\alpha}{\alpha} \circ h_r^\alpha(p)\, \mathrm{d}r \right| \leq K.$$

Since

$$\int_0^t \frac{X\alpha}{\alpha} \circ h_r^\alpha(p)\, \mathrm{d}r = \int_0^t X\alpha \circ h_{\tau(r,p)}(p) \frac{\mathrm{d}}{\mathrm{d}r}\tau(r,p)\, \mathrm{d}r = \int_0^t X\alpha \circ h_r(p)\, \mathrm{d}r,$$

we conclude that the ergodic integrals of $X\alpha$ are uniformly bounded:

$$\sup_{p\in M}\sup_{t\in\mathbb{R}} \left| \int_0^t X\alpha \circ h_r(p)\, \mathrm{d}r \right| \leq K.$$

Proof of the Proposition

In order to conclude, it suffices to prove the following lemma.

Lemma

If the ergodic integrals of $X\alpha$ are uniformly bounded, then α is a continuous coboundary.

Proof of the Lemma.

- Without loss of generality, we can assume that $\Box\alpha = \mu\alpha$ for some $\mu \in \mathbb{R}$. We focus on the case $\mu > 0$.
- In Lecture 2, we proved that there exist two continuous functions $\mathcal{D}^{\pm}_{\mu}(X\alpha)$ such that

$$\int_0^t X\alpha \circ h_r(p)\,\mathrm{d}r = t^{\frac{1\mp\nu}{2}}\mathcal{D}^{\pm}_{\mu}(X\alpha)(\phi^{\mathbf{x}}_{\log t}(p)) + O(1),$$

  where $\nu \in \mathbb{R}_{\geq 0} \cup \imath\mathbb{R}_{>0}$ satisfies $1 - \nu^2 = 4\mu$.

Proof of the Proposition

- Then, the assumption of the lemma implies that $\mathcal{D}^{\pm}_{\mu}(X\alpha) = 0$.
- From the expressions we found in Lecture 2, one can prove that $\mathcal{D}^{\pm}_{\mu}\alpha$ are a linear combination of $\mathcal{D}^{\pm}_{\mu}(X\alpha)$, so that $\mathcal{D}^{\pm}_{\mu}\alpha = 0$ as well.
- This implies that the ergodic integrals of α are also uniformly bounded.
- Since $(h_t)_{t\in\mathbb{R}}$ is minimal, the Gottschalk-Hedlund Theorem implies that α is a continuous coboundary.

This proves the lemma and completes the proof of the proposition.

One idea behind Ratner's Rigidity Theorem

Ratner's Theorem

Theorem (Ratner 1986)

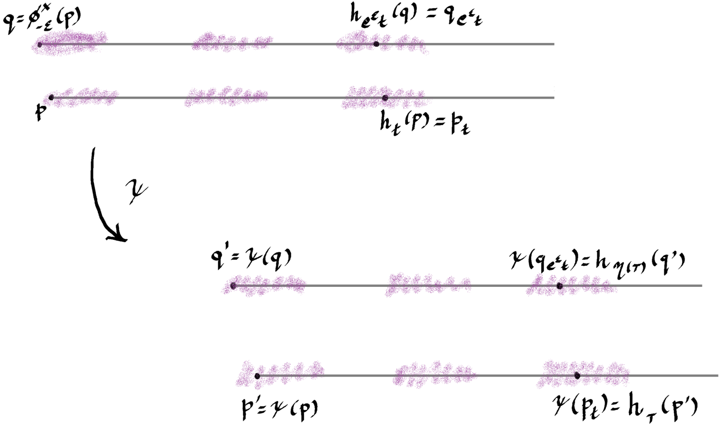
Assume that there exists a measurable isomorphism $\psi \colon M \to M$ between $(h_t^\alpha)_{t\in\mathbb{R}}$ and $(h_t)_{t\in\mathbb{R}}$. Then, there exists $\mathbf{n} \in N_G(\Gamma)$ and there exists a measurable function $\beta \colon M \to \mathbb{R}$ such that

$$\psi(p) = h_{\beta(\Psi_{\mathbf{n}}(p))}(\Psi_{\mathbf{n}}(p)), \qquad \text{for all } p \in M.$$

In particular, α is an almost coboundary.

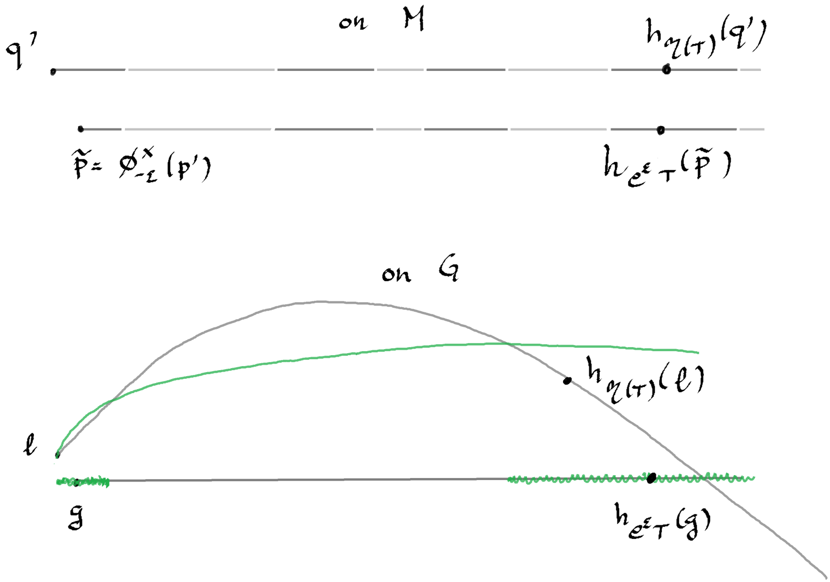
Ratner's Rigidity Theorem (as well as, in a simpler way, the result we proved before) relies on the polynomial divergence of horocycle orbits. The proof is rather technical. We discuss in an informal way one key idea behind the result.

An idea

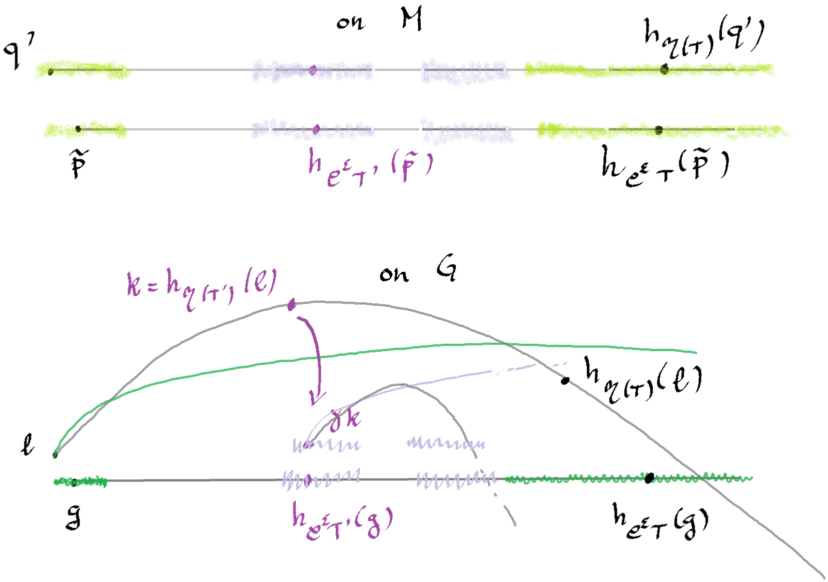


and $\quad \eta(T) = e^{\varepsilon} T + O(T^{1-\gamma})$

An idea


on M
q'
$h_{\eta(T)}(q')$
$\tilde{p} = \phi^X_{-\varepsilon}(p')$
$h_{e^\varepsilon T}(\tilde{p})$
on G
$h_{\eta(T)}(\ell)$
ℓ
g
$h_{e^\varepsilon T}(g)$

An idea



An idea

- The blocks are sub-level sets of quadratic polynomials, where the coefficients depend on the "initial points" (the quadratic term depends on the distance in the V-direction and the linear term depends on the distance in the X-direction).
- By the properties of polynomials, two blocks B_1, B_2 of different colors satisfy

$$d(B_1, B_2) > \min\{|B_1|, |B_2|\}^{1+\delta}, \qquad \text{for some } \delta > 0.$$

- If two blocks C_1, C_2 of the same color are close together, namely $d(C_1, C_2) \leq |C_1|^{1+\delta}$, then we replace them by their convex hull. The coefficients of the defining polynomial can be bounded in terms of the length of the new "superblock".

An idea

- The collection of blocks cover 99% of the interval $[0, T]$ for any sufficiently large $T > 0$.
- A covering lemma tells us that at least one of them has to have length proportional to T.
- In our case, this means that the points q' and $\widetilde{p}$ are at a "sublinear distance".
- This is possible only if q' and $\widetilde{p}$ are on the same U-orbit, which implies

$$\psi(\varphi^{\mathbf{x}}_{-\varepsilon}(p)) = q' = h_\beta(\widetilde{p}) = h_\beta \circ \varphi^{\mathbf{x}}_{-\varepsilon}(\psi(p)).$$