Dynamics on homogeneous spaces
and new applications to number theory

Dmitry Kleinbock

Brandeis University

Nachdiplom Lectures, ETH, Zürich

Spring 2022
Bucha. Anton Logov, Kyiv
Lecture 6, 6/4/22. Ergodicity and mixing on $\Gamma \backslash G$

Recall: at the end of the last lecture we were talking about

$$X_d := \Gamma \backslash G \text{ where } G = \text{SL}_d(\mathbb{R}) \text{ and } \Gamma = \text{SL}_d(\mathbb{Z}),$$

the groups of real (resp. integer) $d \times d$ unimodular matrices. In particular we showed that $\Gamma$ is a non-uniform lattice in $G$, that is, $X_d$ is not compact and carries a finite $G$-invariant measure.

The plan now is to look at dynamics on $\Gamma \backslash G$, with $X_d$ being the main example. The reasons are:

- modulo technical difficulties it is quite similar to the general case;
- the space can be visualized nicely;
- there are plenty of applications to number theory.
Our first goal is to prove

**Theorem 6.1 (Moore’s Ergodicity Theorem).**
Let $\Gamma \subset G = \text{SL}_d(\mathbb{R})$ be a lattice, $X = \Gamma \backslash G$ and $\mu = m_X$.
Let $g \in G$ be such that the group generated by $g$ is unbounded in $G$.
Then $R_g : x \mapsto xg^{-1}$ acts ergodically on $(X, \mu)$.

Here I will follow the representation-theoretic, or spectral, approach going back to Gelfand–Fomin and Mautner (1950s).

Recall that ergodicity of $T : X \to X$ is equivalent to
the absence of non-constant $U_T$-invariant functions in $L^2(X)$.

However we will need to consider the action of several elements of $G$ simultaneously. This defines a unitary representation $\rho$ of $G$
on $\mathcal{H} = L^2(X)$: $\rho(g) = U_{R_{g^{-1}}}$, that is, for $f \in L^2(X)$ we have

$$(\rho(g)f)(x) = f(g^{-1} \cdot x) = f(xg).$$

And the ergodicity result will in fact follow from a result about unitary representations.
**Theorem 6.1.** Let $\mathcal{H}$ be a Hilbert space, and let $\rho$ be a unitary representation of $G = \text{SL}_d(\mathbb{R})$ on $\mathcal{H}$. Suppose that $g \in G$ generates an unbounded subgroup. Then any vector $v_0 \in \mathcal{H}$ that is fixed by $\rho(g)$ is fixed by all of $\rho(G)$.

Because $G$ acts transitively on $X$, the only $\rho(G)$-invariant functions are constant; hence Theorem 6.1 follows from Theorem $6.\overline{1}$.

More generally, we have

**Corollary 6.2.** Let $X$ be a locally compact metric space with a Borel probability measure $\mu$, and suppose that $\mu$ is ergodic for a measure-preserving action of $G = \text{SL}_d(\mathbb{R})$. Then any $g \in G$ generating an unbounded subgroup acts ergodically on $X$.

**Examples:**

- $\text{SL}_d(\mathbb{R}) = G \subset L$, a bigger Lie group, and $X = \Gamma \backslash L$, where $\Gamma$ is a lattice in $L$;

- $G = \text{SL}_2(\mathbb{R})$ acting on the moduli space of Riemann surfaces.
Now let us prove Theorem 6.1; the key argument, usually referred to as "Mautner Phenomenon" or "Mautner Lemma".

[Mautner 1957, Geodesic flows on symmetric Riemannian spaces]

**Proposition 6.3.** Let $G$ be a locally compact metric group, and let $\rho$ be a unitary representation of $G$ on a Hilbert space $\mathcal{H}$. Suppose that $v \in \mathcal{H}$ is fixed by $\rho(F)$ for some subgroup $F \subset G$. Then $v$ is also fixed by $\rho(h)$ for any $h \in G$ with the property that

$$hB_G(r) \cap FB_G(r)F \neq \emptyset \text{ for every } r > 0; \quad (*)$$

(equivalently, $h = \lim_{k \to \infty} g_k h g_k'$ for some $g_k, g'_k \in F$ and $h_k \to e$).

\[ \text{Diagram with group elements and orbits}\]
How is it related to Theorem 6.1?

Let's see it in the case

$$G = \text{SL}_2(\mathbb{R}) \text{ and } g = a_t := \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}. $$

The crucial point is the relationship between $a_t$ and two unipotent subgroups of $G$: $u_s := \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$ and $v_s := \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix}$, namely,

$$a_t u_s a_{-t} = \text{ and } a_t v_s a_{-t} = \text{.}$$
Let $F = g^\mathbb{Z} = \{g^n : n \in \mathbb{Z}\}$. Then

$$FB_G(r)F \supset \bigcup_{k \in \mathbb{N}} a_{kt}B_G(r)a_{-kt}$$

will contain every $h$ of the form $u_s, s \in \mathbb{R}$, and likewise

$$FB_G(r)F \supset \bigcup_{k \in \mathbb{N}} a_{-kt}B_G(r)a_{kt} \supset \{v_s : s \in \mathbb{R}\}.$$

Hence any $g$-invariant vector will also be $u_s$- and $v_s$-invariant for any $s \in \mathbb{R}$, and $\{u_s : s \in \mathbb{R}\}$ and $\{v_s : s \in \mathbb{R}\}$ generate $G$

$\implies$ Theorem 6.1 follows for this case.
Proof of Proposition 6.3.

Recall that we are given \( v \in \mathcal{H} \) fixed by \( \rho(F) \) for \( F \subset G \). Without loss of generality we may assume that \( \|v\| = 1 \).

Define the auxiliary function, so-called matrix coefficient

\[
\phi : G \rightarrow \mathbb{R}, \quad \phi(h) := \langle \rho(h)v, v \rangle.
\]

Notice that \( \phi \) is continuous by the continuity requirement in the definition of a unitary representation.

Also, for \( g_1, g_2 \in F \) and \( h \in G \) we have

\[
\phi(g_1 hg_2) = \langle \rho(g_1 hg_2)v, v \rangle = \langle \rho(h)\rho(g_2)v, \rho(g_1^{-1})v \rangle
\]

\[
= \langle \rho(h)v, v \rangle = \phi(h),
\]

that is, \( \phi \) is bi-\( \rho(F) \)-invariant.
Now let $h \in G$ be as in the statement of the proposition, and, using $(\ast)$, choose sequences $h_k \to e$ and $g_k, g'_k \in F$ such that

$$g_k h_k g'_k \to h \text{ as } k \to \infty.$$  

Then

$$\langle \rho(h)v, v \rangle = \phi(h) \leftarrow \phi(g_k h_k g'_k) = \phi(h_k) \to \phi(e) = \langle v, v \rangle = 1$$

as $k \to \infty$. Since $\|\rho(h)v\| = \|v\| = 1$, this (and the equality case in the Cauchy-Schwartz inequality) implies that $\rho(h)v = v$. 

[Diagram of a triangle with arrows and labels $\phi(h)$]
Proof of Theorem 6.1 for $G = \mathrm{SL}_2(\mathbb{R})$.
Already know it for $g = a_t$; so fix $s \in \mathbb{R} \setminus \{0\}$ and let $g = u_s$ (the remaining case for the unbounded subgroup of $G$).

Suppose that $f \in L^2(X, \mu)$ is invariant under $\rho(g)$.

Apply Proposition 6.3 with

$$F = g^Z \text{ and } h_k = v_{1/k} = \begin{pmatrix} 1 & 0 \\ 1/k & 1 \end{pmatrix}. $$

For $m, n \in \mathbb{Z}$, write

$$g^n v_{1/k} g^m = \begin{pmatrix} 1 & ns \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1/k & 1 \end{pmatrix} \begin{pmatrix} 1 & ms \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 + \frac{ns}{k} & ns \\ 1/k & 1 \end{pmatrix} \begin{pmatrix} 1 & ms \\ 0 & 1 \end{pmatrix}. $$

$$2 = \frac{1 + \frac{ns}{k}}{1/k} \left(1 + \left(1 + \frac{ns}{k}\right) ms + ns \right) \rightarrow y \rightarrow 1/2,$$
For large enough $k$ we may choose

- $n = n(k)$ such that $1 + \frac{ns}{k} \to 2$ as $k \to \infty$;
- $m = m(k)$ with $(1 + \frac{ns}{k}) ms + ns$ uniformly bounded, so one can choose a subsequence for which

$$
(1 + \frac{ns}{k}) ms + ns \to y
$$

for some $y \in \mathbb{R}$;

- then $ms/k$ has to converge to $1/2$.

This way we conclude that $g^n v_{1/k} g^m$ converges to $h := 
\begin{pmatrix} 2 & y \\ 0 & 1/2 \end{pmatrix}$.

By Proposition 6.3 we see that $f$ is invariant under $\rho(h)$. Since this $h$ is conjugate to $a_t$ for some $t$, the theorem follows from the previous case.

The case $d > 2$ follows with some more work using many copies of $SL_2(\mathbb{R})$s inside $SL_d(\mathbb{R})$. \qed
It turns out that the representation-theoretic approach to actions on $\Gamma \backslash G$ does not stop at ergodicity. Using similar ideas and building up on Theorems 6.1 and 6.1, it is possible to prove

**Theorem 6.4.** Let $\Gamma \subset G = \text{SL}_d(\mathbb{R})$ be a lattice, $X = \Gamma \backslash G$, and $\mu = m_X$. Then the action of $G$ on $(X, \mu)$ is mixing.

Recall that it means the following:

$$\langle U_{g_n} \varphi, \psi \rangle \rightarrow \int_X \varphi \, d\mu \int_X \psi \, d\mu$$

for any $\varphi, \psi \in L^2(X, \mu)$ and any $(g_n) \subset G$ with $g_n \to \infty$ as $n \to \infty$

(that is, eventually leaving any compact subset of $G$)

\[\mapsto\]

$$\langle U_{g_n} \varphi, \psi \rangle \to 0 \ \forall \varphi, \psi \in L^2_0(X, \mu) \text{ and } (g_n) \subset G \text{ as above.}$$
Similarly to what we have done about ergodicity, the above theorem follows from an analogous statement involving unitary representations of $G$:

**Theorem 6.4.** Let $\mathcal{H}$ be a Hilbert space, and let $\rho$ be a unitary representation of $G = \text{SL}_d(\mathbb{R})$ on $\mathcal{H}$ without nonzero $\rho(G)$-invariant vectors. Then $\forall \, v, w \in \mathcal{H}$ the matrix coefficients $\langle \rho(\cdot)v, w \rangle$ vanish at $\infty$:

$$\langle \rho(g_n)v, w \rangle \to 0 \text{ for any } (g_n) \subset G \text{ with } g_n \to \infty \text{ as } n \to \infty.$$  

(Equivalently, $\rho(g_n)v \to 0$ in the weak topology on $\mathcal{H}$.)
**Remark:** It actually suffices to prove the theorem for $g_n$ being diagonal matrices with positive entries.

The reason is **Cartan decomposition**: any $g \in G$ can be written as $kak'$, where $a$ is diagonal and $k, k' \in K = SO(d)$.

Indeed, suppose that

$$g_n = k_n a_n k_n'$$

$$G = K A T K$$

$$|\langle \rho(k_n) \rho(a_n) \rho(k_n') v, w \rangle| = |\langle \rho(k_n a_n k_n') v, w \rangle| \geq \varepsilon$$

$$| \langle \rho(a_n k_n') v, \rho(k_n a_n') w \rangle | \geq \varepsilon$$

$$| \langle \rho(k_n') v, \rho(a_n) \rho(k_n')^{-1} w \rangle | \geq \frac{\varepsilon}{2}$$

$$| \langle \rho(a_n) \rho(k_n') v, \rho(k_n')^{-1} w \rangle | \geq \frac{\varepsilon}{4}$$
Proof of Theorem 6.4 for $d = 2$. Suppose for some $\nu \in \mathcal{H}$ the sequence $\rho(a_{t_n})\nu$ weakly converges to $\nu_0 \in \mathcal{H}$, where, as before, $a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$.

We need to prove that $\nu_0 = 0$.

In view of Theorem 6.1, it is enough to prove that

$\nu_0$ is invariant by $\rho(u_s)$,

where, as before, $u_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$.

\[ \text{ergodicity} \implies \text{mixing} \]
Of course we are again going to use the fact that

$$a_{-t} u_s a_t \to e \text{ as } t \to +\infty.$$

Take any $w \in \mathcal{H}$, and write

$$\langle v_0, w \rangle = \lim_{n \to \infty} \langle \rho(a_{t_n}) v, w \rangle = \lim_{n \to \infty} \langle v, \rho(a_{-t_n}) w \rangle$$

and also, for any $s \in \mathbb{R}$

$$\langle \rho(u_s) v_0, w \rangle = \langle v_0, \rho(u_{-s}) w \rangle$$

$$= \lim_{n \to \infty} \langle \rho(a_{t_n}) v, \rho(u_{-s}) w \rangle$$

$$= \lim_{n \to \infty} \langle \rho(u_s a_{t_n}) v, w \rangle$$

$$= \lim_{n \to \infty} \langle \rho(a_{-t_n} u_s a_{t_n}) v, \rho(a_{-t_n}) w \rangle.$$
Hence
\[
\left| \langle \rho(u_s)v_0, w \rangle - \langle v_0, w \rangle \right| = \lim_{n \to \infty} \left| \langle \rho(a_{-t_n}u_s a_{t_n})v - v, \rho(a_{-t_n})w \rangle \right|
\leq \lim_{n \to \infty} \| \rho(a_{-t_n}u_s a_{t_n})v - v \| \| w \| = 0.
\]

Since this holds for all \( w \in \mathcal{H} \), we get \( \rho(u_s)v_0 = v_0 \) as claimed.

**Remark.** The scheme of the argument is quite general and is applicable to more general groups; again, the case of \( \text{SL}_2(\mathbb{R}) \) is a building block for a proof of the general result. All that is needed is

- Cartan decomposition;
- existence of a nontrivial element \( "u_s" \) contracted by the conjugation with the acting elements \( "a_t" \).

Then the absence of mixing for \( "a_t" \) will force an invariant vector for \( "u_s" \), that is, the absence of ergodicity. In other words, the proof basically says that ergodicity implies mixing.
**Example.** Consider \( G = \text{SL}_d(\mathbb{R}) \), \( \Gamma \) a lattice in \( G \), and look at the action of

\[
a_t := \text{diag} \left( e^{t/m}, \ldots, e^{t/m}, e^{-t/n}, \ldots, e^{-t/n} \right) = \begin{pmatrix}
    e^{t/m}I_m & 0 \\
    0 & e^{-t/n}I_n
\end{pmatrix}
\]

on \( \Gamma \setminus G \), where \( d = m + n \),

kind of a higher-dimensional analogue
of the geodesic flow on \( \Gamma \setminus \text{SL}_2(\mathbb{R}) \).

Then upper-triangular (lower-triangular) unipotent matrices
are contracted (expanded) by the map \( g \mapsto a_{-t}ga_t \).

Specifically we can define

\[
U := \{ u_Y : Y \in M_{m \times n}(\mathbb{R}) \}, \text{ where } u_Y := \begin{pmatrix}
    I_m & Y \\
    0 & I_n
\end{pmatrix}
\]

(contracting horospherical subgroup with respect to \( a_1 \))

and

\[
V := \{ v_Y : Y \in M_{n \times m}(\mathbb{R}) \}, \text{ where } v_Y := \begin{pmatrix}
    I_m & 0 \\
    Y & I_n
\end{pmatrix}
\]

(expanding horospherical subgroup with respect to \( a_1 \)).
Why so much emphasis on mixing? Turns out that it is quite useful for discovering many other wonderful properties of the action. Here is a simple picture.

"apple - banana argument"
And here is what can be proved.

The set-up, again:

- $G = \text{SL}_d(\mathbb{R})$, $\Gamma$ is a lattice in $G$, $X = \Gamma \backslash G$, $d = m + n$;
- $a_t = \begin{pmatrix} e^{t/m} I_m & 0 \\ 0 & e^{-t/n} I_n \end{pmatrix}$, $\nu_Y = \begin{pmatrix} 1 \\ Y \\ 1 \end{pmatrix}$, $V = \{ \nu_Y \}$;
- $a_{-t} \nu_Y a_t = \nu_{e^{(\frac{1}{m} + \frac{1}{n})t} Y}$.

**Theorem 6.5.** Let $f$ be a bounded measurable function on $V$ with compact support, and let $\psi \in L^2(X)$ be uniformly continuous. Then for any compact subset $Q$ of $X$ and any $\varepsilon > 0$ there exists $T > 0$ such that

$$\left| \int_V f(\nu) \psi(xva_t) \, dm_V - \int_V f(\nu) \, dm_V \int_X \psi \, d\mu \right| \leq \varepsilon$$

for all $x \in Q$ and all $t \geq T$. 
In other words:

- expanding translates of $V$-orbits by the $a_t$-action become equidistributed in $X$;

- in particular, for any $x \in X_d$ and almost every $v \in V$, the trajectory $\{xva_t : t \geq 0\}$ is dense in $X_d$;

- the rate of equidistribution is uniform over the initial point lying in a fixed compact subset of $X$. 
Sketch of proof:

- Since \( \text{supp}(f) \) and \( Q \) are compact, we can write

\[
f = \sum_{j=1}^{N} f_j
\]

with \( \pi_y \) injective on \( \text{supp}(f_j) \) for all \( y \in Q \) and for each \( j \).

Hence one can without loss of generality assume that the maps \( \pi_y \) are injective on \( \text{supp}(f) \) for all \( y \in Q \).
- **thicken** \( f \) to create functions \( \phi \) on \( G \) and \( \varphi_y := \phi \circ \pi_y^{-1} \) on \( X \) for any \( y \in Q \);

- **apply mixing:**

\[
\int_X \varphi_y(x) \psi(xa_t) \, d\mu(x) \to \int_X \varphi_y \, d\mu \int_X \psi \, d\mu;
\]
\[
\int_X \varphi_y \, d\mu \text{ with } \int_V f \, dm_V
\]

and
\[
\int_X \varphi_y(x) \psi(xa_t) \, d\mu(x) \text{ with } \int_V f(v) \psi(yva_t) \, dm_V.
\]
Need to integrate on $G$ in coordinates coming from subgroups $U$, $V$ and $A := \{\text{block-diagonal matrices}\}$.

Since the Lie algebras of these subgroups span the Lie algebra of $G$, the multiplication map $V \times A \times U \to G$ is a diffeomorphism on a small neighborhood of the identity;

The Haar measure on $G$ decomposes according to Lemma 5.10, with $W = AU$, and Haar measure on $W$ can be written similarly to what was done in the proof of Claim 5.13.
**Conclusion 1:** \(\forall \varepsilon > 0 \ \exists T > 0 \text{ such that } t \geq T\)

\[
\left| \int_V f(v)\psi(yva_t) \, dm_V - \int_X \varphi_y(x)\psi(xa_t) \, d\mu(x) \right|
\]

\[
= \left| \int_G \varphi(vw)\psi(yva_t) \, dm_G(vw) - \int_G \varphi(vw)\psi(yvwa_t) \, dm_G(vw) \right|
\]

\[
= \left| \int_G \varphi(vw)(\psi(yva_t) - \psi(yvwa_t)) \, dm_G(vw) \right| \leq \varepsilon/2
\]

\[
\psi(yva_t a_t^{-1} wa_t)
\]

\[\epsilon\]
We also need to apply mixing to functions \( \varphi_y \) uniformly in \( y \in Q \):

**Proposition 6.6.** Let \( \mathcal{H} \) be a Hilbert space, and let \( \rho \) be a unitary representation of a locally compact group \( G \) on \( \mathcal{H} \).

Then the following are equivalent:

1. \( \forall v, w \in \mathcal{H} \) the matrix coefficients \( \langle \rho(\cdot)v, w \rangle \) vanish at \( \infty \):

   \[ \langle \rho(g_n)v, w \rangle \to 0 \text{ for any } (g_n) \subset G \text{ with } g_n \to \infty \text{ as } n \to \infty; \]

2. for any two subsets \( \mathcal{V}, \mathcal{W} \) of \( \mathcal{H} \) which are compact in the norm topology on \( \mathcal{H} \) and \( \forall \varepsilon > 0 \) there exists a compact subset \( L \) of \( G \) such that

   \[ v \in \mathcal{V}, \ w \in \mathcal{W}, \ g \notin L \implies |\langle \rho(g)v, w \rangle| \leq \varepsilon. \]
Proof. \((2) \implies (1)\) is clear;
not \((2) \implies \exists\) compact \(\mathcal{V}, \mathcal{W} \subset \mathcal{H}, \nu_n \in \mathcal{V}, \omega_n \in \mathcal{W}\)
and \(g_n \to \infty\) in \(G\) such that

\[
|\langle \rho(g_n)\nu_n, \omega_n \rangle| \geq \varepsilon
\]

\[
Downarrow
\]

\[
|\langle \rho(g_n)\nu_n\omega_n^{-1} \rangle| > \frac{\varepsilon}{2}
\]

\[
|\langle \nu_n, \rho(g_n^{-1})\omega_n \rangle| > \frac{\varepsilon}{2}
\]

\[
\implies \text{contradiction to (1)}
\]

Conclusion 2: for any \(\varepsilon > 0\) there exists \(T > 0\) such that

\[
\left| \int_X \varphi_y(x)\psi(xa_t) \, d\mu(x) - \int_X \varphi_y \, d\mu \int_X \psi \, d\mu \right| \leq \varepsilon/2
\]

for all \(y \in Q\) and all \(t \geq T\).
Combining the two conclusions, for any \( \varepsilon > 0 \) we obtain \( T > 0 \) such that \( y \in Q \) and \( t \geq T \)

\[
\Downarrow
\left| \int_V f(v) \psi(y v a_t) \, dm_V - \int_X \varphi_y \, d\mu \int_X \psi \, d\mu \right|
= \left| \int_V f(v) \psi(y v a_t) \, dm_V - \int_V f \, dm_V \int_X \psi \, d\mu \right| \leq \varepsilon. \quad \square
\]

**Remarks.**

- The argument can be traced back to the Ph.D. Thesis of Margulis (1969, unpublished until 1990s), where similar results were proved for geodesic flows in variable negative curvature.

- Even though the proof is presented for \( G = \text{SL}_d(\mathbb{R}) \), the argument, together with the main principle

  “mixing \( \implies \) equidistribution of unstable leaves”

works for arbitrary homogeneous spaces.
Here is one interesting consequence of Theorem 6.5.

**Theorem 6.7.** Suppose that $\Gamma$ is a uniform lattice in $G$, and let $V$ be a horospherical subgroup of $G$ (expanding horospherical for some $a_t$). Then the $V$-action on $X = \Gamma \backslash G$ is **uniquely ergodic**. That is, $m_X$ is the unique $V$-invariant probability measure on $X$.

In particular, the horocycle flow on the unit tangent bundle of a compact quotient of $\mathbb{H}$ is uniquely ergodic (Furstenberg, 1970s).

*Proof.*