Dynamics on homogeneous spaces
and new applications to number theory

Dmitry Kleinbock

Brandeis University

Nachdiplom Lectures, ETH, Zürich

Spring 2022
Anna Chernenko, Dnipro
Художник Аня Черненко
Художник Аня Черненко
Lecture 12, 25/5/22. Other Applications.

Here is a short history of the course so far:

- dynamics on homogeneous spaces;
- some problems in Diophantine approximation;
- their reduction to dynamics on a very specific homogeneous space.

Namely,

\[ X = X_d = \Gamma \backslash G, \]
\[ \text{where } G = \text{SL}_d(\mathbb{R}), \]
\[ \Gamma = \text{SL}_d(\mathbb{Z}); \]

\[ d = m + n, \]
\[ a_t = \text{diag} \left( e^{t/m}, \ldots, e^{t/m}, e^{-t/n}, \ldots, e^{-t/n} \right). \]

Of course this is just one possible illustration of connections between number theory and dynamics. During the remaining two lectures I will try to show a broader picture.
12.1. **Quadratic forms.** Let me briefly describe the set-up which probably contributed the most to the development of connections between number theory and homogeneous dynamics.

Consider a nondegenerate indefinite quadratic form \( f \) on \( \mathbb{R}^d \) of signature \((m, n)\), where \( d = m + n \).

For example, \((m, n) = (2, 1)\):

**Question:** what can one say about the values of \( f \) at integer points?

Of course if \( f \) is proportional to a form with rational coefficients, \( f(\mathbb{Z}^d) \) is a discrete set (boring!) so let’s assume otherwise.

Compare with \( f(p, q) = \|qY - p\|^m \|q\|^n \):

\[ 1 \leq \varepsilon \]
Here is a precise statement along these lines, conjectured in 1929 by Oppenheim and proved in 1987 by Margulis:

**Theorem 12.1** Let $f$ be a real nondegenerate indefinite quadratic form of signature $(m, n)$, $d = m + n > 2$, not proportional to a rational form. Then $\inf_{x \in \mathbb{Z}^d \setminus \{0\}} |f(x)| = 0$.

**Remarks:**

- This is a statement for all forms $f$, not almost all;
- not true if $d = 2$. 

The proof is based on a reduction to dynamics going back to Cassels and Swinnerton-Dyer ('55) and Raghunathan ('70s), which is very similar to the one we used for systems of linear forms.

Namely, we can write $f(x) = \lambda Q_{m,n}(xg)$, where $\lambda \in \mathbb{R}$, $g \in G$, and

$$Q_{m,n}(x_1, \ldots, x_d) = x_1^2 + \cdots + x_m^2 - x_{m+1}^2 - \cdots - x_d^2.$$
\[ f(x) = \lambda Q_{m,n}(xg) \]

\[ f(p, q) = \|qY - p\|^m \|q\|^n \]

\[ Z^d \]

\[ y \]

\[ \Lambda \]

\[ \Lambda \text{ has a small nonzero vector} \]

\[ h \in H := \text{Stab}_G Q_{m,n} \]
Proposition 12.2. Let \( f = \lambda Q_{m,n}(xg) \) be as above, and let \( H = \text{Stab} \mathfrak{g} \cong \text{SO}(m, n) \) and \( \Lambda = \mathbb{Z}^d g \).
Then
\[
|f(x)| \geq \varepsilon \text{ for some } \varepsilon > 0 \text{ and all } x \in \mathbb{Z}^d \setminus \{0\}.
\]

\[\uparrow\]
the orbit \( \Lambda H \) is relatively compact in \( X_d \).

Proof. See the diagram above.

Therefore we can state Theorem 12.1 in an equivalent form:

Theorem 12.1. With \( H \) as above, any relatively compact orbit \( \Lambda H \) is closed.

Explanation: \((\mathbb{Z}^d g)H\) is closed

\[\downarrow \uparrow \text{ (Raghunathan '70s)}\]

\( f \) is proportional to a rational form.
Again, the corresponding statement is false for $n = 2$, since we have seen that the $a_t$-orbits can behave quite wildly.

The difference is that $H$ as above is generated by unipotent elements. Theorem 12.1, proved by Margulis in 1987, is a special case of Raghunathan’s Conjectures. Later it evolved into fundamental theorems established by Ratner (’90).
**Theorem R1.** (Ratner’s Orbit Closure Theorem).
Let $G$ be any Lie group, $\Gamma$ a lattice in $G$, and $U = \{ u_s \}$ a unipotent one-parameter subgroup.
Then the closure of every $U$-orbit in $\Gamma \backslash G$ is **homogeneous**, that is, it is a closed orbit $xL$ of a Lie group $L$ with $U \subset L \subset G$ carrying a finite $L$-invariant measure.

![Diagram](image)

**Theorem R2.** (Ratner’s Measure Classification Theorem).
Let $G, \Gamma, U$ be as above.
Then every ergodic $U$-invariant probability measure on $\Gamma \backslash G$ is **homogeneous**, that is, it is the unique $L$-invariant probability measure on $xL$ with $L$ as above.

**Theorem R3.** (Ratner’s Equidistribution Theorem).
Let $G, \Gamma, U$ be as above.
Then then every $U$-orbit is equidistributed in its closure with respect to the measure discussed in Theorem R2.
These theorems have numerous applications to Diophantine approximation, worthy of a separate lecture course.

For example: Dani and Margulis ('93), for $f$ as above, an interval $I \subset \mathbb{R}$ and $T > 0$, studied the quantity

$$N(f, \varepsilon, T) := \# \{x \in \mathbb{Z}^d \cap B(T) : |f(x)| < \varepsilon \}$$

in comparison with

$$V(f, \varepsilon, T) := \text{Leb}(\{x \in \mathbb{Z}^d \cap B(T) : |f(x)| < \varepsilon \}),$$

and showed that

$$\liminf_{T \to \infty} \frac{N(f, \varepsilon, T)}{V(f, \varepsilon, T)} \geq 1$$

for any $f$ not proportional to a rational form.

**Tools:** Ratner’s Equidistribution Theorem + linearization method.
The complementary upper estimate \( \limsup_{T \to \infty} \frac{N(f, \varepsilon, T)}{V(f, \varepsilon, T)} \leq 1 \) were proved by Eskin–Margulis–Mozes ('98) but only when \( d \geq 5 \); for signatures \((2, 1)\) and \((2,2)\) there exist counterexamples.

This is done via a method of integral inequalities on homogeneous spaces, which also deserves a special lecture course.

**Remark:** the above methods are not effective: given \( T > 0 \), they do not give information on the smallest \( \varepsilon = \psi(T) \) for which there exists \( x \in \mathbb{Z}^d \) with

- \( \|x\| \leq T \) and \( |f(x)| < \psi(T) \) (uniform approximation), or
- \( |f(x)| < \psi(\|x\|) \) (asymptotic approximation).

Effectivization of Ratner-type results has become a hot topic during the last decade, with a lot of developments. For example, in [Lindenstrauss–Margulis ’14] for ternary indefinite quadratic forms it is shown that

- under some Diophantine condition on the coefficients of \( f \), for any large enough \( T \) there exists \( x \in \mathbb{Z}^3 \) with

\[
0 < \|x\| < T \quad \text{and} \quad |f(x)| < (\log T)^{-c},
\]

where \( c \) is an absolute constant.
**Question:** what if we want to get results for almost all lattices (almost all quadratic forms of a given signature) and develop a theory of asymptotic/uniform approximation?

Let us cook up a few definitions.

Given functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and $\psi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{>0}$, say that $f$ is

- **$\psi$-approximable** if

  $$\# \{ x \in \mathbb{Z}^d : |f(x)| \leq \psi(\|x\|) \} = \infty;$$

- **uniformly $\psi$-approximable** if

  $\forall$ sufficiently large $T > 0$ there exists $x \in \mathbb{Z}^d \setminus \{0\}$ with

  $$\|x\| \leq T \text{ and } |f(x)| \leq \psi(T).$$

Note: these definitions depend on the choice of the norm, but if $\psi$ does not change too rapidly, the dependence is mild.
Recently there were some results in this direction via two different methods: using

- effective mean ergodic theorem and strong spectral gap [Ghosh–Gorodnik–Nevo ’20];
- Rogers’ second moment estimates for Siegel transforms [Athreya–Margulis ’18, Kelmer–Yu ’20].

Let us use notation \( \psi_s(T) = \frac{1}{T^s} \) as before.

**Theorem 12.3.** [Ghosh–Gorodnik–Nevo, Kelmer–Yu]

Let \( d = m + n > 2 \) and \( 0 < s < d - 2 \).

Then almost every quadratic form \( f \) of signature \((m, n)\), i.e. of the form \( f(x) = \lambda Q_{m,n}(xg) \) for Haar-a.e. \( g \in G = \text{SL}_d(\mathbb{R}) \), is uniformly \( \psi_s \)-approximable.

(Recall that it means: \( \forall \) sufficiently large \( T > 0 \) \( \exists x \in \mathbb{Z}^d \setminus \{0\} \) with \( \|x\| \leq T \) and \( |f(x)| \leq 1/T^s \).)

Also note that \( f \) is uniformly \( \psi \)-approximable and \( 0 \not\in f(\mathbb{Z}^d \setminus \{0\}) \) \( \implies \) \( f \) is \( \psi \)-approximable.
Given our past experience, this naturally motivates two more questions:

1. Is it possible to get a criterion for almost every / almost no function in the $G$-orbit of $f$ being $\psi$-approximable? If yes, it is easy to guess what it should be:

$$\text{Leb} \left( \{ \mathbf{x} \in \mathbb{R}^d : |f(\mathbf{x})| \leq \psi(\|\mathbf{x}\|) \} \right) \asymp \int_1^\infty \psi(z)z^{d-3} \, dz \begin{cases} < \infty \\ = \infty \end{cases}$$

(consistent with the critical exponent $s = d - 2$).
(2) Maybe it is possible to use dynamics similarly to the way it was used in the case of systems of linear forms?

It was a beautiful idea and we started doing it with my then graduate student Mishel Skenderi. But then...
...it turned out that the method using Rogers’ formulas was so neat, that it could be adapted to a much more general setting.

What is it about? Recall that for a function \( \varphi \) on \( \mathbb{R}^d \) and a \( \text{SL}_d(\mathbb{Z}) \)-invariant subset \( \mathcal{P} \) of \( \mathbb{Z}^d \), we defined the \( \mathcal{P} \)-Siegel transform of \( \varphi \):

\[
\hat{\varphi}(\mathbb{Z}^d g) := \sum_{x \in \mathcal{P}} \varphi(x)
\]

(actually we did it for \( \mathcal{P} = \text{Pr}(\mathbb{Z}^d) \), but \( \mathcal{P} = \mathbb{Z}^d \setminus \{0\} \) is also a good choice.)

**Siegel’s Formula** (Theorem 11.2) says that

\[
\int_{X_d} \hat{\varphi} \, d\mu = c_{\mathcal{P}} \int_{\mathbb{R}^d} \varphi \, dx;
\]

in fact \( c_{\mathcal{P}} \) is equal to

\[
\begin{cases} 
\frac{1}{\zeta(d)} & \text{if } \mathcal{P} = \text{Pr}(\mathbb{Z}^d), \\
1 & \text{if } \mathcal{P} = \mathbb{Z}^d \setminus \{0\}. 
\end{cases}
\]
Now here is what Rogers proved in 1956:

**Theorem 12.4.** Take $\mathcal{P} = \begin{cases} \Pr(\mathbb{Z}^d) & \text{if } d \geq 2, \\ \mathbb{Z}^d \setminus \{0\} & \text{if } d \geq 3. \end{cases}$

Then there exists a constant $C = C(d) > 0$ such that for any bounded Borel measurable $E \subset \mathbb{R}^d$ one has

$$\text{Var}(\hat{1}_E) \leq C \cdot \text{Leb}(E).$$

(Recall: $\hat{1}_E(\mathbb{Z}^d g) = \# \{ x \in \mathcal{P} : gx \in E \}$.)

This is precisely what one needs to apply Lemma 3.11 (quasi-independent Borel–Cantelli) to a sequence of functions $1_{E_k}$ with

$$E_k = \{ x \in B(k) : |f(x)| \leq \psi(\|x\|) \}.$$  

(here $f = Q_{m,n}$)
and conclude that

$$\text{Leb}\left(\{ x \in \mathbb{R}^d : |f(x)| \leq \psi(\|x\|)\}\right) \begin{cases} < \infty \\ = \infty \end{cases}$$

$$\downarrow$$

$$\#\{ v \in \Lambda : |f(v)| \leq \psi(\|v\|)\} \begin{cases} < \infty \\ = \infty \end{cases} \quad \text{for } \mu\text{-a.e. } \Lambda \in X_d$$

$$\downarrow$$

$$\#\{ x \in \mathbb{Z}^d : |f(gx)| \leq \psi(\|gx\|)\} \begin{cases} < \infty \\ = \infty \end{cases} \quad \text{for Haar-a.e. } g \in G$$

However note that to prove that almost every / almost no function in the $G$-orbit of $f$ is $\psi$-approximable we need to look at

$$\#\{ x \in \mathbb{Z}^d : |f(gx)| \leq \psi(\|x\|)\},$$

which is different!
To pass from $\psi(\|g\mathbf{x}\|)$ to $\psi(\|\mathbf{x}\|)$, we decided to impose an additional assumption on $\psi$:

**Definition.** Say that $\psi$ is regular if for some $a > 1$ and $b > 0$ one has

$$\psi(az) \geq b\psi(z) \text{ for all } z > 0.$$

Then multiplication by $g$ will amount to replacing $\psi$ by $b\psi$ (and we get independence of a choice of the norm as an added bonus).

However, now we need to make sure that there is no big difference between

$$\text{Leb}\left( \{ |f(\mathbf{x})| \leq \psi(\|\mathbf{x}\|) \} \right)$$

and

$$\text{Leb}\left( \{ |f(\mathbf{x})| \leq b\psi(\|\mathbf{x}\|) \} \right).$$
This is magically taken care of by the following

**Definition.** Say that $f$ is **subhomogeneous** if there exists a constant $c > 0$ such that

$$|f(tx)| \leq t^c |f(x)| \quad \forall \ t \in (0, 1) \text{ and } \forall \ x \in \mathbb{R}^n.$$ 

Finally here is a (special case of the) theorem we proved NOT using dynamics:

**Theorem 12.5.** [K–Skenderi ’21] Let $f : \mathbb{R}^n \to \mathbb{R}$ be **subhomogeneous**, and let $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ be **regular** and nonincreasing. Then

$$\text{Leb} \left( \{ x \in \mathbb{R}^d : |f(x)| \leq \psi(\|x\|) \} \right) \asymp \int_{1}^{\infty} \psi(z) z^{d-3} \, dz \begin{cases} < \infty \\ = \infty \end{cases}$$

\[ \uparrow \]

then $f \circ g$ is $\psi$-approximable for Haar-

\[ \begin{cases} \text{almost no} \\ \text{almost every} \end{cases} \quad g \in G. \]

(Also a sufficient condition for uniform $\psi$-approximability.)
12.2. Quantitative non-divergence. This is another theme related to unipotent flows with lots of recent applications to number theory. Here is a timeline:

- [Margulis ’71]: if \( \{u_x\}_{x \in \mathbb{R}} \) is a one-parameter subgroup of \( G = \text{SL}_d(\mathbb{R}) \) consisting of unipotent matrices, then \( \{\Lambda u_x\} \) is non-divergent for any \( \Lambda \in X_d \).

\[ d = 2 : \text{horocycles come down after going up} \]

\[ d > 2 : \text{need an inductive procedure} \]
[Dani '79]: strengthened Margulis' result by showing that such orbits return into a suitably chosen compact set with positive frequency. More precisely, Dani proved that there are $0 < \varepsilon, \eta < 1$ such that for any interval $[0, t] \subset [0, +\infty)$ one has

$$\text{Leb}\left(\{x \in [0, t] : \Lambda u_x \not\in Q_\varepsilon\} \right) < \eta \cdot t. \quad (*)$$

[Shah '94]: generalized Dani's result to polynomial maps $h : \mathbb{R}^k \rightarrow G$ in place of $x \mapsto u_x$. 

[Dani '86]: improved his previous result by showing that for any $\Lambda \in X_d$ and any $\eta > 0$ there exists $\varepsilon > 0$ such that $(*)$ holds for any unipotent $\{u_x\}_{x \in \mathbb{R}} \subset G$ and any $T > 0$;
The essence of the main result of [K–Margulis ’98] is basically an explicit dependence of $\eta$ on $\varepsilon$ in $(*)$, applicable to a very general class of maps $\phi$ of several variables that do not have to be polynomial, let alone the orbits of unipotent subgroups. Below is a version for real analytic maps.

**Theorem 12.6.** For any $k, d \in \mathbb{N}$ there exist $C, \alpha > 0$ with the following property. Suppose a ball $B \subset \mathbb{R}^k$, a real analytic function $h : \mathbb{R}^k \to \text{SL}_d(\mathbb{R})$ and $\rho > 0$ be given. Assume that

$$\sup_{x \in B} \|w h(x)\| \geq \rho^\ell \quad \forall \ell = 1, \ldots, d - 1 \text{ and } \forall w \in \wedge^\ell (\mathbb{Z}^d) \setminus \{0\}. \quad (**)$$

Then for any $\varepsilon > 0$ one has

$$\text{Leb}\left(\left\{x \in B : \mathbb{Z}^d h(x) \notin Q_\varepsilon\right\}\right) \leq C \left(\frac{\varepsilon}{\rho}\right)^\alpha \text{Leb}(B).$$

An application to unipotent flows:

if $h(x) = gu_x$ for $0 \leq x \leq t$, then

$$\sup_{x \in B} \|w h(x)\| \geq \|w h(0)\| = \|wg\|.$$
An application to Diophantine approximation:
we take \( h(x) = v_{\phi(x)} a_t \), where \( \phi : \mathbb{R}^k \to M_{n \times m}(\mathbb{R}) \) is real analytic, \( d = m + n \), and \( v_Y, a_t \) are as before:

\[
  v_Y = \begin{pmatrix}
  I_m & 0 \\
  Y & I_n
\end{pmatrix}, \quad a_t = \begin{pmatrix}
  e^{t/m} I_m & 0 \\
  0 & e^{-t/n} I_n
\end{pmatrix}.
\]

The upshot is that often it is possible to check (***) uniformly for all \( t \).
This way, for any \( \lambda > 0 \) we get a measure estimate for \( \varepsilon = e^{-\lambda t} \)

\[
  \{ x \in B : \Lambda_{\phi(x)} a_t \notin Q_{e^{-\lambda t}} \} = \{ x \in B : \Delta(\Lambda_{\phi(x)} a_t) \geq \lambda t \}
\]
decaying exponentially with \( t \).
Hence, by the Borel–Cantelli Lemma + the Dani Correspondence, \( Y = \phi(x) \) is not \( \psi_s \)-approximable for any \( s > 1 \) \( \text{and a.e. } x \in \mathbb{R}^k \)

\[
\downarrow
\]
can develop Diophantine approximation for generic points of the submanifold \( \{ Y = \phi(x) \} \) of \( M_{n \times m}(\mathbb{R}) \).
Example. When $m = 1$ (or, dually, $n = 1$), (**) holds when the image of $\phi$ is not contained in any affine hyperplane of $\mathbb{R}^n$.

This proved Sprindžuk’s 1980 conjecture [K–Margulis ’98].

The case $\phi(x) = (x, \ldots, x^n)$ corresponds to Mahler’s Conjecture proved by Sprindžuk in 1964.

Furthermore, with lots of additional work it was possible to prove the full version of the Khintchine–Groshev Theorem for smooth submanifolds of $\mathbb{R}^n$ [Bernik–K–Margulis ’01, Beresnevich–Bernik–K–Margulis ’02, . . . , Beresnevich–L.Yang ’21]

When $\min(m, n) > 1$ the corresponding non-degeneracy condition is trickier, worked out by [Aka–Breuillard–Rosenzweig–de Saxcé ’18], with the notion of constraining pencils in $M_{n \times m}(\mathbb{R})$ replacing affine hyperplanes of $\mathbb{R}^n$. Other applications: [Beresnevich–K ’22]
12.3. Equidistribution of translates of subsets of unstable leaves.

Again recall our main set-up responsible for applications to Diophantine approximation:

- $X = X_d = \Gamma \backslash G$, where $G = \text{SL}_d(\mathbb{R})$, $\Gamma = \text{SL}_d(\mathbb{Z})$;
- $d = m + n$, $a_t = \text{diag}(e^{t/m}, \ldots, e^{t/m}, e^{-t/n}, \ldots, e^{-t/n})$;
- $\nu_Y = \begin{pmatrix} I_m & 0 \\ Y & I_n \end{pmatrix}$, so that $V = \{ \nu_Y : Y \in M_{n \times m}(\mathbb{R}) \}$ is the expanding horospherical subgroup with respect to $a_1$;
- then we worked with $\Lambda_Y = \mathbb{Z}^d \nu_Y = \{(qY - p, q) : p \in \mathbb{Z}^m, q \in \mathbb{Z}^n \}$.

An informal summary:

dynamical properties of ‘lots of’ $\Lambda \in X$ \implies similar properties of ‘lots of’ $\Lambda_Y$
(full measure of dense orbits, full dimension of exceptional orbits).
Another useful feature of the above example is that expansion of the elements of \( V \) by conjugation is uniform (conformal):

\[
\Lambda_Y a_t = Z^d v_Y a_t = Z^d a_t (a_{-t} v_Y a_t) = Z^d a_t \cdot v e^{\frac{m \cdot t}{n} + r}
\]

Of course this is due to the fact that \( a_t \) as above is of a rather special form. On the other hand, ergodic theory works quite well for arbitrary diagonal actions. So what happens if instead of \( a_t \) as above we take a diagonal subgroup of \( G \) ‘in a general position’? say, with real eigenvalues. In a suitable basis, it has the form

\[
g_t := g_t^{r,s} = \text{diag}(e^{r_1 t}, \ldots, e^{r_m t}, e^{-s_1 t}, \ldots, e^{-s_n t}),
\]

where the weights \( r = (r_i) \) and \( s = (s_j) \) are such that

\[
0 < r_i, s_j \quad \text{and} \quad \sum_{i=1}^{m} r_i = 1 = \sum_{j=1}^{n} s_j.
\]

Our previous choice obviously corresponds to equal weights

\[
r = m := (\frac{1}{m}, \ldots, \frac{1}{m}) \quad \text{and} \quad s = n := (\frac{1}{n}, \ldots, \frac{1}{n}).
\]
If the weights are ordered so that $r_1 \geq \cdots \geq r_m$ and $s_1 \leq \cdots \leq s_n$, the expanding horospherical subgroup with respect to $g_1$ is contained in the group of unipotent lower-triangular matrices (coincides with it if all the weights are different).

But for Diophantine purposes it still makes sense to study $g_t$-trajectories of $\Lambda_Y$. This corresponds to

Diophantine approximation with weights

where one ‘pays more attention’ to some of the components of $q$ and linear forms comprising $Y$, and less to others.
**Recall:** For $T > 0$ and a non-increasing function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$, we defined

$$S_{n,m}(\psi, T) := \left\{ Y \in M_{n \times m}(\mathbb{R}) \mid \begin{aligned} &\text{the system} \left\{ \begin{aligned} \| qY - p \|^m &\leq \psi(T) \\ \| q \|^n &\leq T \end{aligned} \right\} \\ &\text{has a nontrivial integer solution} \end{aligned} \right\}.$$  

Now replace the system

$$\left\{ \begin{aligned} |qY_i - p_i|^m &\leq \psi(T), \quad i = 1, \ldots, m \\ |q_j|^n &\leq T, \quad j = 1, \ldots, n \end{aligned} \right\}$$

(here $Y_i$ are the columns of $Y$) with

$$\left\{ \begin{aligned} |qY_i - p_i|^{1/r_i} &\leq \psi(T), \quad i = 1, \ldots, m \\ |q_j|^{1/s_j} &\leq T, \quad j = 1, \ldots, n \end{aligned} \right\} \quad \text{(add weights)}$$
The latter can be shortened to

\[
\begin{align*}
\|qY - p\|_r & \leq \psi(T) \\
\|q\|_s & \leq T
\end{align*}
\]

if for \( u = (u_1, \ldots, u_k) \in \mathbb{R}^k_+ \), \( k \in \mathbb{N} \),
we define the \( w \)-quasinorm \( \| \cdot \|_u \) on \( \mathbb{R}^k \) by

\[
\|x\|_u := \max_{1 \leq i \leq k} |x_i|^{1/u_i}.
\]

Then all the standard Diophantine notions get their weighted analogues by simply replacing

\( \|qY - p\|^m \) with \( \|qY - p\|_r \) and \( \|q\|^n \) with \( \|q\|_s \),
and the corresponding dynamics will involve \( g^{r,s}_t \) instead of \( a_t \).

For example we can define \((r,s)\)-badly approximable matrices by

\[
\text{BA}_{r,s} = \left\{ Y \in M_{n \times m}(\mathbb{R}) : \inf_{(p,q) \in \mathbb{Z}^d \setminus \{0\}} \|qY - p\|_r \|q\|_s > 0 \right\}
\]

\[
= \left\{ Y \in M_{n \times m}(\mathbb{R}) : \{\Lambda Y g^{r,s}_t : t > 0\} \text{ is bounded in } X \right\}.
\]
Now comes a natural

**Question:** can one extract some Diophantine consequences from dynamics, given that \( V \) is now much smaller than the expanding horospherical subgroup with respect to \( g_1 \)?

**Answer:** yes, in this and some other cases it is possible to show that generic dynamical properties are inherited by proper subgroups of the expanding horospherical subgroup.

This particular case was handled in [K–Margulis '12].

**Theorem 12.7.** There exists \( \lambda > 0 \) such that for any \( f \in C_c^\infty(V) \), \( \varphi \in C_c^\infty(X) \) and for any compact \( Q \subset X \) there exists \( C = C(f, \varphi, Q) \) such that for all \( x \in Q \) and \( t > 0 \) one has

\[
\left| \int_V f(v) \varphi(xvg_t^r,s) \, dm_V(v) - \int_V f \, dm_V \int_X \varphi \, d\mu \right| \leq \tilde{C} e^{-(\min_{i,j}\{r_i,s_j\}) \lambda t}.
\]
Remarks:

- Can be used to establish full Hausdorff dimension of

\[ BA_{r,s} = \{ Y \in M_{n \times m}(\mathbb{R}) : \{ \wedge Y g_t^{r,s} : t > 0 \} \text{ is bounded} \} \]

following the argument of [K–Margulis '96]

(see [K–Weiss '10] for a stronger ‘winning’ property; the case \( n = 1 \) was considered in [Pollington–Velani '02] following a method due to Davenport).
An ineffective version of Theorem 12.7 (using Ratner’s theorem) was proved earlier in [K–Weiss ’08].

More general results: [Dabbs–Kelly–Li ’16], [Shi ’21] (the notion of expanding cone of directions in which equidistribution takes place).

Extended to double and multiple equidistribution in [K–Shi–Weiss ’17], [Bjorklund–Gorodnik ’19].

Can be used to study sets $A_{r,s}(\psi)$ and $D_{r,s}(\psi)$ (weighted Khintchine–Groshev, [K–Strömbergsson–Yu ’21]).
Sketch of Proof of Theorem 12.7. Two ingredients:

- **effective equidistribution of** $a_t$-translates;
- **quantitative non-divergence of** $h_s$-translates, where $h_s = a_t g_t^{-1}$.

(Using the fact that $V$ is the expanding horospherical subgroup relative to $a_t$.)
A related topic: equidistribution of translates of analytic/smooth submanifolds of $V$ (with or without weights)

- [Shah ’09]: when $\min(m, n) = 1$ and the image of $\phi$ is not contained in any affine hyperplane of $\mathbb{R}^n$, its $a_t$-translates become equidistributed (by Ratner’s Theorems and Dani–Margulis linearization method)

$\downarrow$

$\phi(x)$ is not Dirichlet-Improvable for almost all $x$

- [Shah ’10]: the same with weights

- [L.Yang ’16, Shah–L.Yang ’20, P.Yang ’20]: the same for $\min(m, n) > 1$, with natural conditions on $\phi$ guaranteeing equidistribution