Dynamics on homogeneous spaces
and new applications to number theory

Dmitry Kleinbock

Brandeis University

Nachdiplom Lectures, ETH, Zürich

Spring 2022
Zoya Cherkassky. Kyiv, Ukraine → Israel
Lecture 11, 18/5/22.

Shrinking targets and Dirichlet improvability.

Let us start by recalling the main result of the last lecture.

**Corollary 10.7.** Let

- $G$ be a connected semisimple center-free Lie group without compact factors,
- $\Gamma$ an irreducible lattice in $G$, $X = \Gamma \backslash G$,
- $\mu$ = the $G$-invariant probability measure on $X$,
- $(a_t)$ a diagonalizable one-parameter subgroup of $G$,
- $\Delta$ a DL function on $X$.

Then for $\mu$-a.e. $x \in X$ and any real-valued function $R(\cdot)$ one has

$$\mu(\{x \in X : \Delta(xa_t) \geq R(t) \text{ for } \infty \text{ many } t \in \mathbb{N}\}) = \begin{cases} 0 \\ 1 \end{cases}$$

$$\sum_{t=1}^{\infty} \Phi_\Delta(R(t)) \begin{cases} < \infty \\ = \infty \end{cases}$$
Recall: for a nonnegative function $\Delta$ on $X$ and $z \geq 0$, we denoted

$$\Phi_\Delta(z) := \mu(\Delta^{-1}([z, \infty)))$$

(the tail distribution function of $\Delta$), and said that $\Delta$ is DL if

$\checkmark$ it is uniformly continuous

$\checkmark$ $\Phi_\Delta$ does not decrease too fast:

$$\exists c, \delta > 0 \text{ such that } \Phi_\Delta(z + \delta) \geq c\Phi_\Delta(z) \ \forall \ z \geq 0. \quad (DL)$$

**Corollary 10.7+.** If instead of (DL) we have

$$\forall c < 1 \ \exists \delta > 0 \text{ and } z_0 > 0 \text{ such that } \Phi_\Delta(z + \delta) \geq c\Phi_\Delta(z) \ \forall \ z \geq z_0, \quad (DL+)$$

then for any $R(\cdot)$ with $\sum_{t=1}^{\infty} \Phi_\Delta(R(t)) = \infty$ we have

$$\frac{\#\{1 \leq t \leq N : \Delta(xa_t) \geq R(t)\}}{\sum_{t=1}^{N} \Phi_\Delta(R(t))} \overset{\mu\text{-a.e.}}{\rightarrow} 1.$$
Now recall the special case suitable for Diophantine applications:

- $X = X_d = \Gamma \backslash G$, where $G = \text{SL}_d(\mathbb{R})$ and $\Gamma = \text{SL}_d(\mathbb{Z})$;
- $\Delta : X \rightarrow \mathbb{R}_+$ defined by $\Delta(\Lambda) := \max_{v \in \Lambda \setminus \{0\}} \log \left( \frac{1}{\|v\|} \right)$.

It was mentioned several times already that this function satisfies (DL) or even (DL+), and also that

$$\mu \left( \{ \Lambda \in X : \Delta(\Lambda) \geq R \} \right) \asymp e^{-dR} \iff \mu \left( Q_r^c \right) \asymp r^d.$$

Let us finally discuss it now.
One way to do it is to use **Reduction Theory**, that is, approximation of $X$ by a Siegel domain.

**Recall:** A set $\Sigma_{s,t} := U_s A_t K$, where $0 < t \leq \sqrt{3}/2$, $s \geq 1/2$,

$$K = \text{SO}(d), \quad U_s := \left\{ \begin{pmatrix} 1 \\ * \\ \ddots \\ * \\ * \\ 1 \end{pmatrix} \colon |*| \leq s \right\}$$

and

$$A_t := \left\{ \text{diag}(a_1, \ldots, a_d) : \left| \frac{a_{i+1}}{a_i} \right| \geq t \text{ for } i = 1, \ldots, d - 1 \right\}$$

is called a Siegel domain; we proved (Proposition 5.11) that it is a surjective set (that is, $\pi(\Sigma_{s,t}) = X_d$).
Furthermore, it can be shown that $\Sigma_{s,t}$ has the following Siegel Property:

$$\#\{\gamma \in \Gamma : \gamma \Sigma_{s,t} \cap \Sigma_{s,t} \neq \emptyset\} < \infty.$$ 

This means that the multiplicity of $\pi|_{\Sigma_{s,t}}$ is finite, thus, up to a constant multiple, measures of subsets of $X$ can be approximated by measures of their $\pi$-preimages in the Siegel domain.
In particular, $\pi^{-1}(Q^c_r)$ is (approximately) equal to

$$\{uak \in \Sigma_{1/2, \sqrt{3}/2} : |a_1| < r\}$$

(in other words, $a_1$ is responsible for the length of a minimal vector of the lattice $\mathbb{Z}^d uak$).

Hence the problem reduces to a measure computation in $\Sigma_{1/2, \sqrt{3}/2}$. 
But we want to do better!

The following was proved in [K–Margulis ’99]:

**Proposition 11.1.** There exist positive $C_d, C'_d$ such that

$$C_d r^d \geq \mu(Q_r^c) \geq C_d r^d - C'_d r^{2d} \quad \text{for all } 0 < r < 1.$$ 

The main tool is a generalized Siegel’s summation formula.

Here is a brief synopsis:

- A vector $v$ in a lattice $\Lambda \subset \mathbb{R}^d$ is called **primitive** in $\Lambda$ if it is not a multiple of another element of $\Lambda$; equivalently, if there exists a basis $\{v_1, \ldots, v_k\}$ of $\Lambda$ with $v_1 = v$.

- Denote by $\text{Pr}(\Lambda)$ the set of all primitive vectors in $\Lambda$.

- Given a function $\varphi$ on $\mathbb{R}^d$, define a function $\widehat{\varphi}$ on $X_d$ by

$$\widehat{\varphi}(\Lambda) := \sum_{v \in \text{Pr}(\Lambda)} \varphi(v)$$

(the **primitive Siegel transform** of $\varphi$).

In particular, $\widehat{1}_B(\Lambda) = \#\{\text{Pr}(\Lambda) \cap B\}.$
Here is a result proved by Siegel in 1945:

**Theorem 11.2.** For any $\varphi \in L^1(\mathbb{R}^d)$, one has

$$\int_{\chi_d} \hat{\varphi} \, d\mu = \frac{1}{\zeta(d)} \int_{\mathbb{R}^d} \varphi \, dv$$

**Sketch of Proof.** This is essentially a Fubini Theorem: if $H := \text{Stab}_G(e_1)$, then

$$H \backslash G \cong \mathbb{R}^d \setminus \{0\} \quad \text{and} \quad \text{Pr}(\mathbb{Z}^d) \cong (\Gamma \cap H) \backslash \Gamma$$

$$\downarrow$$

the left hand side of Siegel’s formula $= \int_{\Gamma \backslash G} \int_{(\Gamma \cap H) \backslash \Gamma}$

while the right hand side $= \int_{H \backslash G} \int_{(\Gamma \cap H) \backslash H}$.

\[ \square \]
Sketch of Proof of Proposition 11.1.

Observe that the above theorem immediately implies the upper estimate: take \( r > 0 \) and let

\[ \varphi := \text{the characteristic function of } B(r). \]

Then:

- \( \Lambda \notin Q_r \implies B(r) \) contains at least two primitive vectors (\( v \) and \( -v \)) of \( \Lambda \);
- \( \Lambda \in Q_r \implies \hat{\varphi}(\Lambda) = 0. \)

Therefore, with \( c_d = \text{the volume of the unit ball in } \mathbb{R}^d \),

\[
\frac{1}{\zeta(d)} c_d r^d = \frac{1}{\zeta(d)} \int_{\mathbb{R}^d} \varphi \, dv \\
= \int_{\chi_d} \hat{\varphi} \, d\mu = \int_{Q_r^c} \hat{\varphi} \, d\mu \geq 2\mu(Q_r^c).
\]

And for the lower estimate one needs to show that lattices \( \Lambda \) with \( \hat{\varphi}(\Lambda) > 2 \) contribute to \( \int_{Q_r^c} \hat{\varphi} \, d\mu \) rather insignificantly.
Namely, say that an ordered pair \((v_1, v_2) \subset \Lambda\) is primitive if it is extendable to a basis of \(\Lambda\), and denote by \(\text{Pr}^2(\Lambda)\) the set of all such pairs.

Now, given a function \(\Phi\) on \(\mathbb{R}^{2d}\), define a function \(\hat{\Phi}^2\) on \(X_d\) by

\[
\hat{\Phi}^2(\Lambda) := \sum_{(v_1,v_2) \in \text{Pr}^2(\Lambda)} \varphi(v_1, v_2).
\]

Then one has

**Theorem 11.3.** For any \(\Phi \in L^1(\mathbb{R}^{2d})\), one has

\[
\int_{X_d} \hat{\Phi}^2 \, d\mu = \frac{1}{\zeta(d)\zeta(d-1)} \int_{\mathbb{R}^{2d}} \Phi \, dv_1 \, dv_2.
\]

**Sketch of Proof.** Also a Fubini-type theorem. \(\square\)
Now observe that:

- whenever there exist at least two linearly independent vectors in $\Lambda \cap B(r)$, for any $v_1 \in \Pr(\Lambda)$ one can find $v_2 \in \Lambda \cap B$ such that

  $$(v_1, v_2) \in \Pr^2(\Lambda) \text{ and } (v_1, -v_2) \in \Pr^2(\Lambda).$$

- Consequently, one has

  $$\hat{\varphi}(\Lambda) = \#(\Pr(\Lambda) \cap B(r)) \leq \frac{1}{2} \# \left( \Pr^2(\Lambda) \cap (B(r) \times B(r)) \right)$$

whenever $\hat{\varphi}(\Lambda) > 2$.

- The right hand side is equal to $\frac{1}{2} \hat{\Phi}^2(\Lambda)$, where $\Phi = 1|_{B(r) \times B(r)}$. 
Therefore

\[
\frac{1}{\zeta(d)} c_d r^d = \int_{X_d} \hat{\varphi} \, d\mu = \int_{\{\hat{\varphi}(\Lambda) = 2\}} \hat{\varphi} \, d\mu + \int_{\{\hat{\varphi}(\Lambda) > 2\}} \hat{\varphi} \, d\mu \\
\leq 2\mu(\{\hat{\varphi}(\Lambda) = 2\}) + \frac{1}{2} \int_{\{\hat{\varphi}(\Lambda) > 2\}} \hat{\Phi}^2 \, d\mu \\
\leq 2\mu(Q_r^c) + \frac{1}{2} \int_{X_d} \hat{\Phi}^2 \, d\mu \\
= 2\mu(Q_r^c) + \frac{1}{2\zeta(d)\zeta(d-1)} (c_d r^d)^2.
\]
Now that we have mastered several tricks involving shrinking targets on homogeneous spaces, let us recall the connection to number theory.

**The starting point:**
the set of $Y \in M_{n \times m}(\mathbb{R})$ such that certain systems of inequalities have nontrivial integer solutions coincides with the set of $Y \in M_{n \times m}(\mathbb{R})$ such that the $a_t$-trajectory of the lattice $\Lambda_Y$ in $X_d$ behaves in a certain way. Here

- $X_d = \Gamma \backslash G$, where $G = \text{SL}_d(\mathbb{R})$, $\Gamma = \text{SL}_d(\mathbb{Z})$;
- $d = m + n$, $v_Y = \begin{pmatrix} I_m & 0 \\ Y & I_n \end{pmatrix}$;
- $\Lambda_Y = \mathbb{Z}^d v_Y = \{(qY - p, q) : p \in \mathbb{Z}^m, q \in \mathbb{Z}^n\}$;
- $a_t = \text{diag}(e^{t/m}, \ldots, e^{t/m}, e^{-t/n}, \ldots, e^{-t/n})$. 

[Diagram of a trajectory on a graph with a labeled coordinate system.]
More specifically, we considered the sets

\[ S_{n,m}(\psi, T) = \left\{ Y \in M_{n \times m}(\mathbb{R}) \mid \begin{array}{l}
\text{the system } \begin{cases}
\|qY - p\|_2^m \leq \psi(T) \\
\|q\|_2^n \leq T
\end{cases}
\text{ has a nontrivial integer solution}
\end{array} \right\} \]

\[ = \left\{ Y \in M_{n \times m}(\mathbb{R}) : \Delta(\Lambda Y a_t) \geq R(t) \right\}, \]

where

\[ \psi\left(e^{t-nR(t)}\right) = e^{-t-nR(t)} \]

\(\psi(\cdot)\) and \(R(\cdot)\) are related to each other via the Dani Correspondence (Lemma 9.1)

\(t\) is related to \(T\) via \(T = e^{t-nR(t)}\)

\(\Delta(\Lambda) = \max_{v \in \Lambda \setminus \{0\}} \log \left( \frac{1}{\|v\|} \right)\).
Consequently, we have (in the set-up of asymptotic approximation)

\[ A_{n,m}(\psi) = \limsup_{T \to \infty} S_{n,m}(\psi, T) \]

\[ = \begin{cases} 
Y \in M_{n \times m}(\mathbb{R}) & \exists \text{ an unbounded set of } T \\
\text{such that the system} & \left\{ \begin{array}{l}
\|qY - p\|^n \leq \psi(T) \\
\|q\|^n \leq T
\end{array} \right. \\
\text{has a nontrivial integer solution} & \end{cases} \]

\[ = \{ Y \in M_{n \times m}(\mathbb{R}) : \Delta(\Lambda Y a_t) \geq R(t) \text{ for an unbounded set of } t > 0 \}. \]

Compare with Corollary 10.7:

\[ \mu \left( \Lambda \in X \mid \Delta(\Lambda a_t) \geq R(t) \text{ for } \infty \text{ many } t \in \mathbb{N} \right) = \begin{cases} 
0 & \iff \sum_{t=1}^{\infty} e^{-dR(t)} \left\{ \begin{array}{l}
< \infty \\
= \infty
\end{array} \right. \\
1
\end{cases} \]

and find two differences.
Indeed, we need to pass

- from “\(\Lambda_Y\), where \(Y \in M_{n \times m}(\mathbb{R})\)” to “\(\Lambda \in X\)”;
- from “for \(\infty\) many \(t \in \mathbb{N}\)” to “for an unbounded set of \(t > 0\)”.

The first one is standard: any \(\Lambda \in X_d\) can be written as

\[
\Lambda = \Lambda_Y \begin{pmatrix} Y_1 & Y_2 \\ 0 & Y_3 \end{pmatrix}
\]

\[
\Lambda a_t = (\Lambda_Y a_t) a_{-t} \begin{pmatrix} Y_1 & Y_2 \\ 0 & Y_3 \end{pmatrix} a_t = \Lambda_Y a_t \begin{pmatrix} Y_1 & e^{-(t/m + t/n)} Y_2 \\ 0 & Y_3 \end{pmatrix}
\]

\[\implies \text{dist}(\Lambda a_t, \Lambda_Y a_t) \leq \text{const} \implies |\Delta(\Lambda a_t) - \Delta(\Lambda_Y a_t)| \leq \text{const}.\]
The second one is also OK under some regularity assumptions on $R(\cdot)$. For example, say that $R(\cdot)$ is **quasi-increasing** if

$$\exists D \in \mathbb{R} \text{ such that } R(t) \geq R(s) - D \text{ whenever } s \leq t \leq s + 1.$$ 

Then:

$$\Delta(\Lambda a_t) \geq R(t) \text{ for } \infty \text{ many } t \in \mathbb{N}$$

$$\Downarrow$$

$$\Delta(\Lambda a_t) \geq R(t) \text{ for an unbounded set of } t > 0$$

$$\Delta(\Lambda a_k a_s) \geq R(k) - D, \quad t = k + s$$

$$\Downarrow$$

$$\Delta(\Lambda a_k) \geq R(k) - D \text{ for } \infty \text{ many } k \in \mathbb{N} \text{ (taking } k = \lfloor t \rfloor).$$

↑ maybe a different constant
This way we can extract a Diophantine corollary (0–∞ law for the sets $A_{n,m}(\psi) \subset M_{n \times m}(\mathbb{R})$, which leads to a continuous-time problem) from our discrete-time dynamical Borel–Cantelli lemma on $X$.

**Corollary 10.7’.** Let $R(\cdot)$ be a quasi-increasing function. Then

$$\text{Leb} \left( \left\{ Y \in M_{n \times m}(\mathbb{R}) \, \left| \begin{array}{c}
\Delta(\Lambda_Y a_t) \geq R(t) \\
\text{for an unbounded set of } t > 0
\end{array} \right\} \right) = \left\{ \begin{array}{c} 0 \\
\infty
\end{array} \right. \right.$$  

\[ \downarrow \]

$$\int_0^\infty e^{-dR(t)} \, dt \begin{cases} < \infty \\ = \infty \end{cases}$$

And our previous work on the Dani Correspondence brings us back to the Khintchine–Groshev Theorem.
With this in mind, let us turn to the uniform approximation, which is the featured topic of today's lecture.

There will be more questions and conjectures than results. The reason is that the set

\[ D_{n,m}(\psi) = \liminf_{T \to \infty} S_{n,m}(\psi, T) \]

\[ \Delta \geq R(t) \]

\[ = \left\{ \begin{array}{l} Y \in M_{n \times m}(\mathbb{R}) \quad \forall \text{sufficiently large } T \text{ the system} \\
\quad \left\{ \begin{array}{l} \|qY - p\|^m \leq \psi(T) \\
\|q\|^n \leq T \\
\text{has a nontrivial integer solution} \end{array} \right. \end{array} \right. \]

\[ = \left\{ Y \in M_{n \times m}(\mathbb{R}) : \Delta(\Lambda Y a_t) \geq R(t) \text{ for all sufficiently large } t > 0 \right\} \]

is a liminf set, which is difficult to handle.
On the other hand, any liminf set is a complement of a limsup set!

In other words, one can consider

\[ D_{n,m}(\psi)^c = \limsup_{T \to \infty} S_{n,m}(\psi, T)^c \]

\[
= \left\{ Y \in M_{n \times m}(\mathbb{R}) \mid \begin{array}{l}
\text{for an unbounded set of } T \text{ the system}
\begin{cases}
\|qY - p\|^m \leq \psi(T) \\
\|q\|^n \leq T
\end{cases}
\text{has no nontrivial integer solutions}
\end{array} \right\}.
\]

But this complement to \( S_{n,m}(\psi, T) \)

is quite painful to work with!

(Complement of lots of neighborhoods of affine subspaces)

This explains why classical metric number theory is powerless here.
However the game changes (somewhat) if one brings in the correspondence with dynamics:

\[ \mathcal{D}_{n,m}(\psi) = \left\{ Y \in M_{n \times m}(\mathbb{R}) \middle| \forall \text{ sufficiently large } t > 0 \right\} \]

\[ \mathcal{D}_{n,m}(\psi)^c = \left\{ Y \in M_{n \times m}(\mathbb{R}) \middle| \Delta(\Lambda Y a_t) < R(t) \text{ for an unbounded set of } t > 0 \right\} \]

So the problem reduces to studying trajectories eventually visiting (or not visiting) shrinking neighborhoods of the critical locus \( \mathcal{L}_\infty \).
Recall: this situation corresponds to the approximation function $\psi$ of the form

$$\psi(T) = \frac{1 - h(T)}{T}, \text{ where } h(T) \to 0 \text{ as } T \to \infty$$

and by our previous work on the Dani Correspondence (see Lecture 9) we can take $R(t) \approx h(e^t)$. 
**Note:** one can play the same game of improving Dirichlet’s Theorem with the sup norm \( \| \cdot \| = \| \cdot \|_\infty \) replaced by an arbitrary norm \( \nu \).

This set-up, as we have seen before, is very sensitive to the choice of the norm. This way one can define the set of \( \psi \)-Dirichlet matrices with respect to \( \nu \):

\[
D_{n,m}^\nu(\psi) := \left\{ Y \in M_{n \times m}(\mathbb{R}) \left| \begin{array}{c}
\Delta_\nu(\Lambda_Y a_t) \geq R(t) \\
\forall \text{ sufficiently large } t > 0
\end{array} \right. \right\},
\]

where again \( \psi(\cdot) \) and \( R(\cdot) \) are related to each other via the Dani Correspondence, and

\[
\Delta_\nu(\Lambda) := \max_{\nu \in \Lambda \setminus \{0\}} \log \left( \frac{r_\nu}{\nu(\nu)} \right)
\]

is nonnegative everywhere and vanishes on the critical locus \( \mathcal{L}_\nu \).

Then we have to deal with shrinking neighborhoods of \( \mathcal{L}_\nu \).
Four reasons why this problem is harder than the asymptotic one, and how to overcome those, d’apres [K–Strömbergsson–Yu ’21]:

(1) The passage

\[ \Lambda_Y \in \mathbb{Z}^d \mathcal{V} \leftrightarrow \Lambda \in \mathcal{X} \]

becomes problematic.

Solution: instead of mixing, apply

effective equidistribution of \( a_t \)-translates of \( \{ \Lambda_Y \} \)

[K–Margulis ’96, Bjorklund–Gorodnik ’19].
**Note:** with the notation

\[ T : \mathcal{X} \rightarrow \mathcal{X}, \Lambda \mapsto \Lambda a_1 \]

and for \( m \leq n \), before we needed to estimate

\[ \mu(T^{-m}(B_m) \cap T^{-n}(B_n)) = \mu(B_m \cap T^{-(n-m)}(B_n)) \]

But now we need

\[ \text{Leb}\left( \{ Y \in M_{n \times m}(\mathbb{R}) : \Lambda Y \in T^{-m}(B_m) \cap T^{-n}(B_n) \} \right) = ??? \]

**Solution:** use effective double equidistribution of \( a_t \)-translates of \( \Lambda \mathcal{V} \) [K–Shi–Weiss ’17, Bjorklund–Gorodnik ’19].
The passage

for an unbounded set of $t > 0 \rightarrow$ for $\infty$ many $t \in \mathbb{N}$

is also problematic. Indeed, if

$$\Lambda a_t \in B_t := \Delta^{-1}([0, R(t))]$$

and $k = \lfloor t \rfloor$, $\Lambda a_k$ can be quite far from $B_k$

(at a bounded distance, but this is not enough!)
Solution: discretize the continuous-time shrinking target problem by introducing the following thickened targets:

\[ \tilde{B}_k := \bigcup_{0 \leq s < 1} B_{k+s} a_{-s} \]
\[ = \bigcup_{0 \leq s < 1} \Delta^{-1}\left([0, R(k + s))\right) a_{-s}. \]

Then for any \( \Lambda \in X_d \),

\[ \Lambda a_k \in \tilde{B}_k \iff \Lambda a_t \in B_t \text{ for some } t \in [k, k + 1). \]

In particular,

\[ Y \notin \mathcal{D}_{n,m}(\psi) \iff \Lambda_Y a_t \in B_t \text{ for an unbounded set of } t > 0 \]
\[ \iff \Lambda_Y a_k \in \tilde{B}_k \text{ for } \infty \text{ many } k \in \mathbb{N}, \]

and the expected convergence/divergence condition is

\[ \sum_{k=1}^{\infty} \mu(\tilde{B}_k), \text{ not } \sum_{k=1}^{\infty} \mu(B_k). \]
(3) How to compute $\mu(B_t)$ and $\mu(\tilde{B}_k)$?

This is actually quite tricky.

For small $R > 0$ we need the measure of

$$\Delta^{-1}([0, R]) = \{ \Lambda \in X_d : \Lambda \cap \bar{B}(0, e^{-R}) = \{0\} \}$$

$$= Q_{e^{-R}} \approx Q_{1-R},$$

or rather its thickening in the $a_t$-direction.

Here is a picture for $d = 2$:
**Theorem 11.5.** [K–Strömbergsson–Yu ’21]

Let $\kappa_d = \frac{d^2 + d - 4}{2}$ and $\lambda_d = \frac{d(d-1)}{2}$. Then

$$\mu(Q_1 - R) \preceq_d R^{\kappa_d + 1} \log \lambda_d \left( \frac{1}{R} \right) \quad \text{as } R \to 0^+.$$ 

When $d = 2$: $\kappa_2 = \lambda_2 = 1$ and we are getting [K–Yu ’20]

$$\mu(Q_1 - R) \preceq R^2 \log \left( \frac{1}{R} \right) \quad \text{as } R \to 0^+. $$
As for the thickening
\[
\tilde{B}_k = \bigcup_{0 \leq s < 1} B_{k+s}a_{-s}
\]
of \(B_k\), it lies between the thickenings of \(B_k\) and \(B_{k+1}\) by \(\{a_{-s} : 0 \leq s < 1\}\), so one can use the measure of those sets for upper and lower estimates.

**Theorem 11.6.** [K–Strömbergsson–Yu ’21]
Let \(\kappa_d = \frac{d^2+d-4}{2}\) and \(\lambda_d = \frac{d(d-1)}{2}\) be as in Theorem 11.5. Then
\[
\mu \left( \bigcup_{0 \leq s < 1} Q_{1-R}a_{-s} \right) \asymp_d R^{\kappa_d} \log^{\lambda_d} \left( \frac{1}{R} \right) \quad \text{as } R \to 0^+.
\]

When \(d = 2\): \(\kappa_2 = \lambda_2 = 1\) and we are getting [K–Yu ’20]
\[
\mu \left( \bigcup_{0 \leq s < 1} Q_{1-R}a_{-s} \right) \asymp R \log \left( \frac{1}{R} \right) \quad \text{as } R \to 0^+.
\]
Now we know what result we should hope for:

\[ \text{Leb} \left( \left\{ Y \in M_{n \times m}(\mathbb{R}) \mid \Delta(Y_t \Lambda) < R(t) \text{ for an unbounded set of } t > 0 \right\} \right) \]

\[ = \text{Leb} \left( \left\{ Y \in M_{n \times m}(\mathbb{R}) : \Lambda \Lambda_k \in \tilde{B}_k \text{ for } \infty \text{ many } k \in \mathbb{N} \right\} \right) = \begin{cases} 0 \\ \infty \end{cases} \]

\[ \uparrow \]

\[ \sum_{k=1}^{\infty} R(k)^{\frac{a_d}{\log \lambda_d}} \left( \frac{1}{R(k)} \right) \left\{ \frac{1}{R(k)} \right\} = \infty \]

\[ \uparrow \]

\[ \sum_{t=1}^{\infty} h(e^t)^{\frac{a_d}{\log \lambda_d}} \log \lambda_d \left( \frac{1}{h(e^t)} \right) \left\{ \frac{1}{h(e^t)} \right\} = \infty \]

(Cauchy condensation) \[ \uparrow \]

\[ \sum_{k=1}^{\infty} \frac{h(k)^{\frac{a_d}{\log \lambda_d}}}{k} \log \lambda_d \left( \frac{1}{h(k)} \right) \left\{ \frac{1}{h(k)} \right\} = \infty \]
But now even the convergence case is not straightforward, since we only know an approximate value of

$$\text{Leb}(\{ Y \in M_{n \times m}(\mathbb{R}) : \Lambda Y a_k \in \tilde{B}_k \})$$

(using effective equidistribution).

And for the divergence case one would need to prove quasi-independence of preimages of sets $\tilde{B}_k$
(also only using approximate values for their measures).
We now come to the final and the most serious obstruction:

(4) Here we have a family of targets \( \{ \tilde{B}_k \} \) zooming in on a compact subset of \( X \).

As a result, if one approximates characteristic functions of these sets by smooth functions, their derivatives become big compared with the measures of the sets.

(Even when the norm \( \nu \) is Euclidean and the sets \( B_t \) are \( K \)-invariant, which lets them mix up really fast, their thickenings \( \tilde{B}_k \) are not!)
Solution: follow the same method of approximating the characteristic functions of
\[
\tilde{B}_k = \bigcup_{0 \leq s < 1} \Delta^{-1} \left( [0, R(k+s)) \right) a_{-s}
\]
by smooth functions.

In fact, it is enough to work with sets of the form
\[
\tilde{B}' = \bigcup_{0 \leq s < 1} \Delta^{-1} \left( [0, R(k)) \right) a_{-s}
\]

Key fact: the norms of the derivatives of the corresponding smooth functions \( f \) grow at most polynomially in \( \frac{1}{R(k)} \).

Then we are left with two cases

- \( R(k) \) is not too small: then the polynomial growth of \( \Omega^\ell f \) is beaten by exponential decay of matrix coefficients \( \implies \) exponential rate of equidistribution of \( a_\ell \)-translates of \( \{ \Lambda_Y \} \);

- \( R(k) \) is very small: if it happens often enough, it contributes to a set of measure zero by Borel–Cantelli.
This strategy is enough to settle the convergence case, while the divergence case holds only under an additional technical condition (again coming from the technique of estimating the variance of $\frac{S_N}{E_N}$).

**Theorem 11.7** (K–Strömbergsson–Yu, 2021). Let $\psi(t) = \frac{1-h(t)}{t}$ be non-increasing, with $h(t) = 1 - t\psi(t)$ also non-increasing.

Let $d = m + n$, and let $\kappa_d = \frac{d^2 + d - 4}{2}$ and $\lambda_d = \frac{d(d-1)}{2}$ be as before. If the series

$$
\sum_{k=1}^{\infty} \frac{h(k)\kappa_d}{k} \log \lambda_d \left( \frac{1}{h(k)} \right)
$$

converges, then $\mathcal{D}_{n,m}(\psi)$ is of full Lebesgue measure. Conversely, if (*) diverges and some ugly technical condition holds, then $\mathcal{D}_{n,m}(\psi)$ is of zero Lebesgue measure.

K–Wadleigh (2018): The case $m = n = 1$, without the extra condition (i.e. a perfect 0–1 law), using continued fractions.

$$
\sum \frac{h(k)}{k} \log \left( \frac{1}{h(k)} \right) \quad \kappa_d = \lambda_d = 1
$$
Examples.

Let

\[ h(t) = \frac{1 - c (\log t)^{-\tau}}{t} \]

for some \( c > 0 \) and \( \tau \geq 0 \). In this case (\(*)\) diverges \( \iff \tau \leq \frac{1}{\kappa_d} \), and the “ugly condition” is satisfied whenever \( \tau \leq \frac{1}{\kappa_d} \).

Hence for such \( \psi \), \( \mathcal{D}_{n,m}(\psi) \) is of full measure if \( \tau > \frac{1}{\kappa_d} \), and of zero measure if \( \tau \leq \frac{1}{\kappa_d} \).

Let

\[ h(t) = \frac{1 - c (\log t)^{-1/\kappa_d} (\log \log t)^{-\tau}}{t} \]

for some \( c > 0 \) and \( \tau \in \mathbb{R} \). Then (\*) diverges \( \iff \tau \leq \frac{\lambda_d + 1}{\kappa_d} \), and the “ugly condition” is satisfied whenever \( \tau < \frac{\lambda_d}{\kappa_d} \).

Hence for such \( \psi \), \( \mathcal{D}_{n,m}(\psi) \) is of full measure if \( \tau > \frac{\lambda_d + 1}{\kappa_d} \), and of zero measure if \( \tau < \frac{\lambda_d}{\kappa_d} \); for \( \tau \) in the range \( \frac{\lambda_d}{\kappa_d} \leq \tau \leq \frac{\lambda_d + 1}{\kappa_d} \) — no information!
P.S. A small remark on uniform mixing as introduced by [Fernández, Melián and Pestana '07].

**Definition.** Let \((X, d, \mathcal{B}, \mu, T)\) be a metric probability measure preserving system. Say that \(T\) is uniformly mixing at \(x_0 \in X\) if there exists a positive sequence \((\alpha_n)\) with \(\sum_{n=1}^{\infty} \alpha_n < \infty\), and \(\forall n \in \mathbb{N}\) and for each pair of balls \(A, B\) centered at \(x_0\) one has

\[
|\mu(A \cap T^{-n}B) - \mu(A)\mu(B)| \leq \alpha_n \mu(B).
\]

**Observation.** Suppose \(T\) is continuous, invertible, and \(x_0\) is a non-recurrent point for \(T^{-1}\). (For example, \(X\) is not compact and \(T^{-n}x_0 \to \infty\).) Then there is no uniform mixing.

**Proof.**

\[
\forall n \exists B \ni x_0 \text{ such that } T^{-n}(B) \cap A = \emptyset
\]

\[
\Rightarrow |\mu(A \cap T^{-n}B) - \mu(A)\mu(B)| = \mu(A)\mu(B)
\]

\[
\leq \alpha_n \mu(B)
\]

\[\square\]