dirimptalk2020

Thursday, June 11, 2020 7:45 PM

dirimptalk...

Geometry and dynamics of improvements to Dirichlet's Theorem in Diophantine approximation

Dmitry Kleinbock

Brandeis University

ICERM 2020

Based on joint work with

Anurag Rao, Jinpeng An, Lifan Guan

Starting point: Dirichlet's Theorem

Theorem (Dirichlet): for any $A \in M_{m\times n}(\mathbb{R})$ and any $T > 1$ $\exists$ $\mathbf{p} \in \mathbb{Z}^m$, $\mathbf{q} \in \mathbb{Z}^n \smallsetminus \{0\}$ such that

(1) $$\|A\mathbf{q} - \mathbf{p}\| \le \frac{1}{T^{n/m}} \quad \text{and} \quad \|\mathbf{q}\| < T.$$

Corollary (Dirichlet): for any $A \in M_{m\times n}(\mathbb{R})$ $\exists$ ∞ many $\mathbf{q} \in \mathbb{Z}^n$ such that

(2) $$\|A\mathbf{q} - \mathbf{p}\| < \frac{1}{\|\mathbf{q}\|^{n/m}} \text{ for some } \mathbf{p} \in \mathbb{Z}^m.$$

Here $\|\cdot\|$ stands for the supremum norm on $\mathbb{R}^m$ and $\mathbb{R}^n$.

Most of Metric Theory of Diophantine approximation answers the following

Question: what happens if the RHS of (2) is replaced by a faster decreasing function of $\|\mathbf{q}\|$? (Khinitchine-type Theorems)

Alternatively: can try to replace $\frac{1}{T^{n/m}}$ in the RHS of (1) by a faster decreasing function of T (Improving Dirichlet's Theorem)

Improving Dirichlet's Theorem

In this talk we will deal with only one type of improvement:

replacing $\frac{1}{T^{n/m}}$ by $\frac{c}{T^{n/m}}$, where $c < 1$.

This dates back to Davenport and Schmidt (1969).

Definition: Say that $A \in \widehat{D}^{m,n}$ (Dirichlet-Improvable) if $\exists\ c < 1$ such that for all large enough T $\exists\ \mathbf{p} \in \mathbb{Z}^m$, $\mathbf{q} \in \mathbb{Z}^n \smallsetminus \{0\}$ with

$$(1c) \qquad \|A\mathbf{q} - \mathbf{p}\| < \frac{c}{T^{n/m}} \quad \text{and} \quad \|\mathbf{q}\| < T.$$

Here is what was proved by Davenport and Schmidt:

Theorem DS1: $\widehat{D}^{m,n}$ has Lebesgue measure zero.

Theorem DS2: $\widehat{D}^{m,n}$ has full Hausdorff dimension.

The latter was proved via

Theorem DS2′: The set $BA^{m,n}$ of badly approximable systems of linear forms is contained in $\widehat{D}^{m,n}$.

Note: the complement $\widehat{D}^{m,n} \smallsetminus BA^{m,n}$ is nontrivial unless $m = n = 1$, when it coincides with $\mathbb{Q}$ (contains singular systems of linear forms).

Lattices

Before moving on, let us quickly prove Theorems DS1 and DS2 (or rather DS2′) by introducing dynamics (Dani's Correspondence, which was in fact implicit in the work of Davenport and Schmidt).

Put $d = m + n$ and let

$$X_d \cong \mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$$

be the space of unimodular lattices in $\mathbb{R}^d$.

A general principle: the Diophantine properties of A can be understood via the trajectory $\{g_t\Lambda_A : t \geq 0\}$, where

$$\Lambda_A = \begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix} \mathbb{Z}^d = \left\{ \begin{pmatrix} A\mathbf{q} - \mathbf{p} \\ \mathbf{q} \end{pmatrix} : \mathbf{p} \in \mathbb{Z}^m,\ \mathbf{q} \in \mathbb{Z}^n \right\}$$

and

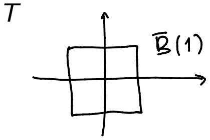
$$g_t = \begin{pmatrix} e^{t/m} I_m & 0 \\ 0 & e^{-t/n} I_n \end{pmatrix}.$$

Dynamics

Indeed: the inequalities $\|A\mathbf{q} - \mathbf{p}\| \le \frac{1}{T^{n/m}}$, $\|\mathbf{q}\| \le T$
have a non-trivial integer solution

$$\Updownarrow$$

$$g_t \Lambda_A \cap \bar{B}(1) \neq \{0\},$$



where $\bar{B}(1)$ = closed ball of radius 1 w.r.t. the supremum norm.

Similarly: $A \in \widehat{D}^{m,n}$, i.e. $\exists\ c < 1$ such that the inequalities

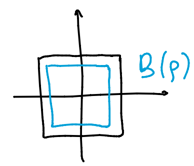
(1c) $\qquad \|A\mathbf{q} - \mathbf{p}\| < \frac{c}{T^{n/m}}$ and $\|\mathbf{q}\| < T$.

have a non-trivial integer solution for all large enough T

$$\Updownarrow$$

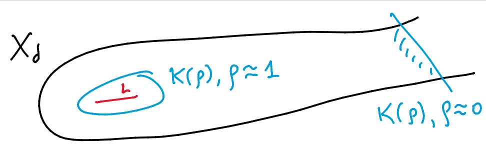
$$\exists\ \rho < 1 \text{ such that } g_t \Lambda_A \cap B(\rho) \neq \{0\}$$

for all large enough t



(here $B(\rho)$ = open ball of radius ρ
with respect to the supremum norm).

Critical Locus



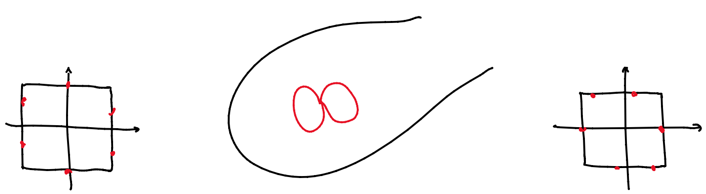
Given $\rho > 0$, **define**

$$K(\rho) := \{\Lambda \in X_d : \Lambda \cap B(\rho) = \{0\}\},$$

compact sets whose union over all $\rho > 0$ exhausts X_d.
But we are interested in the intersection over all ρ less than 1:

$$\bigcap_{\rho<1} K(\rho) = K(1) =: L,$$

the critical locus for the supremum norm.



Geometry and dynamics of improvements to Dirichlet's Theorem in Diophantine approximation
Kleinbock
Background
Generalizations
New results and conjectures
Thanks

The correspondence

Conclusion: $A \in \widehat{D}^{m,n}$

$$\Updownarrow$$

$$\exists\ \rho < 1 \text{ such that } g_t\Lambda_A \cap B(\rho) \neq \{0\} \text{ for all large enough } t$$

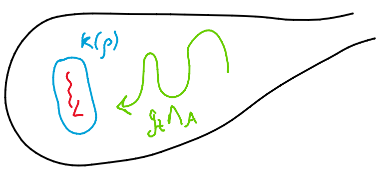
$$\Updownarrow$$

$$\exists\ \rho < 1 \text{ such that } g_t\Lambda_A \notin K(\rho) \text{ for all large enough } t$$

$$\Updownarrow$$

$g_{\mathbb{R}_+}\Lambda_A$ eventually avoids L

($\exists$ a neighborhood $U \supset L$ such that $g_t\Lambda_A \notin U$ for all large enough t)



Proof of Davenport–Schmidt Theorems

A restatement: $A \notin \widehat{D}^{m,n}$

$$\Updownarrow$$

$$\exists \text{ a sequence } t_k \to \infty \text{ such that } \lim_{k\to\infty} g_{t_k}\Lambda_A \in L.$$

Proof of Theorem DS1 uses ergodicity of the g_t-action on X_d (Moore's Ergodicity Theorem) and the fact that $\{\Lambda_A : A \in M_{m\times n}(\mathbb{R})\}$ is an expanding horosphere with respect to the g_t-action. □

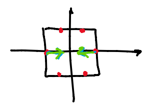
Proof of Theorem DS2′ uses the structure of the critical locus L (Hajos–Minkowski Theorem). Namely,

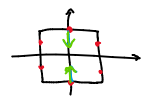
$A \notin \widehat{D}^{m,n} \Rightarrow g_{t_k}\Lambda_A \to \Lambda \in L$

$\Rightarrow \Lambda \ni \mathbf{e}_k$ for some $k = 1, \ldots, d$

$\Rightarrow$ either $g_{\mathbb{R}_+}\Lambda$ or $g_{\mathbb{R}_-}\Lambda$ is divergent

$\Rightarrow g_{\mathbb{R}_+}\Lambda_A$ is unbounded. □





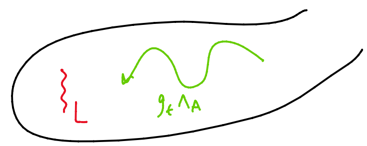
The two ingredients

We now see that the definition of the set

$$\widehat{D}^{m,n} = \{A \in M_{m\times n}(\mathbb{R}) : g_{\mathbb{R}_+}\Lambda_A \text{ eventually avoids } L\}$$

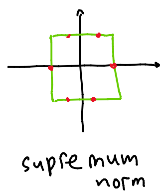
hinges on two ingredients:

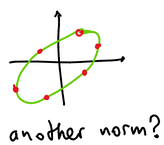
- the acting group $\{g_t\}$
- the critical locus $L = K(1)$



The latter, in its turn, depends on the choice of the supremum norm $\|\cdot\|$ on $\mathbb{R}^d$: indeed the choice of $\rho = 1$ comes from the fact that

$$1 = \sup\left\{\rho : \Lambda \cap B(\rho) = \{0\} \text{ for some } \Lambda \in X_d\right\}.$$





Weights

Question 1: what if $g_t = \begin{pmatrix} e^{t/m} I_m & 0 \\ 0 & e^{-t/n} I_n \end{pmatrix}$ is replaced by a more general one-parameter subgroup, for example

$$g_t^{\mathbf{r},\mathbf{s}} := \operatorname{diag}(e^{r_1 t}, \dots e^{r_m t}, e^{-s_1 t}, \dots e^{-s_n t}),$$

where $r_i, s_j > 0$ and $\sum_i r_i = \sum_j s_j = 1$?

Answer: this is called Diophantine approximation with weights.
One can similarly define

$$\widehat{D}^{\mathbf{r},\mathbf{s}} := \left\{ A \in M_{m\times n}(\mathbb{R}) : g_{\mathbb{R}_+}^{\mathbf{r},\mathbf{s}} \Lambda_A \text{ eventually avoids } L \right\}$$

$$= \left\{ A \in M_{m\times n}(\mathbb{R}) \;\middle|\; \begin{array}{c} \exists\ c < 1 \text{ such that for all large enough } T \\ \text{there exists a non-trivial integer solution} \\ \text{of } |A_i \mathbf{q} - p_i| < \dfrac{c}{T^{r_i}},\ |q_j| < T^{s_j} \end{array} \right\}.$$

Theorems DS1, DS2, DS2′ of Davenport–Schmidt (and their proofs) extend to this generality in a (more or less) straightforward way.

Norms

Question 2: what if the supermum norm on $\mathbb{R}^d$ is replaced by another norm ν?

Define the Hermite constant of ν as

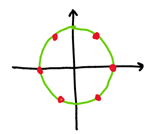
$$\gamma_\nu := \max_{\Lambda \in X_d} \min_{\mathbf{x} \in \Lambda \smallsetminus \{0\}} \nu(\mathbf{x})^2,$$

i.e. the square of the radius of the biggest ν-ball with no nonzero vectors of some lattice $\Lambda \in X_d$.

(It so happens that $\gamma_\nu = 1$ when $\nu = \| \cdot \|_\infty$.)

Note: in many cases the value γ_ν is not even known.

For example if $\nu = \| \cdot \|_2$, the Euclidean norm on $\mathbb{R}^d$, it is only known when $d = 1, 2, \ldots, 8$ and 24.



Nevertheless, we can prove theorems about it!

Generalized Dirichlet's Theorem

For example, it follows immediately from the definition of γ_ν that

$$\Lambda \cap \bar{B}_\nu(\sqrt{\gamma_\nu}) \neq \{0\} \text{ for any } \Lambda \in X_d,$$

in particular for Λ of the form $g_t \Lambda_A$ or $g_t^{\mathbf{r},\mathbf{s}} \Lambda_A$.

Hence we immediately get a

Dirichlet–Minkowski Theorem for an arbitrary norm:

for any $A \in M_{m\times n}(\mathbb{R})$ and any $T > 1$ $\exists$ a non-trivial solution to

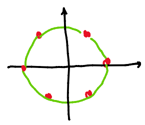
$$\begin{pmatrix} T^{n/m}(A\mathbf{q} - \mathbf{p}) \\ T^{-1}\mathbf{q} \end{pmatrix} \in \bar{B}_\nu(\sqrt{\gamma_\nu}).$$

(similarly one can write down a more general weighted version)

Example: $(m = n = 1,\ \nu = \|\cdot\|_2,\ \gamma_\nu = \frac{2}{\sqrt{3}})$

for any $\alpha \in \mathbb{R}$ and $T > 1$ $\quad \exists\ p \in \mathbb{Z},\ q \in \mathbb{Z} \smallsetminus \{0\}$

such that $\quad T(\alpha q - p)^2 + \dfrac{q^2}{T} \ \leq \dfrac{2}{\sqrt{3}}.$



Generalized Dirichlet-improvement

Say that $A \in \widehat{D}_\nu^{m,n}$ if $\exists\ \rho < \sqrt{\gamma_\nu}$ such that for large enough T there exists a nontrivial integer solution to

$$\begin{pmatrix} T^{n/m}(A\mathbf{q} - \mathbf{p}) \\ T^{-1}\mathbf{q} \end{pmatrix} \in B_\nu(\rho).$$

(similarly one can define $\widehat{D}_\nu^{\mathbf{r},\mathbf{s}}$, the weighted version)

Example: $(m = n = 1,\ \nu = \|\cdot\|_2)$ $\alpha \in \widehat{D}_\nu$ if for some $c < 1$ and large enough T the inequality

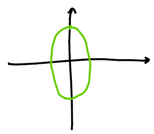
$$T(\alpha q - p)^2 + \frac{q^2}{T} \leq \frac{2}{\sqrt{3}}c$$

has a non-trivial integer solution.

The work of Andersen–Duke

Andersen and Duke (arXiv:1905.05236) introduced this set (in their notation, 'the set of numbers for which Minkowski's approximation theorem can be improved') in the case $m = n = 1$, and, among other things, proved

Theorem AD: Suppose that the norm ν on $\mathbb{R}^2$ is strongly symmetric, that is, satisfies



$$\nu(x, y) = \nu(|x|, |y|) \text{ for all } (x, y) \in \mathbb{R}^2.$$

Then $\widehat{D}_\nu$ is uncountable and has Lebesgue measure zero.

This was done using continued fractions of a special type, defined with the help of the norm ν.

The goal of this talk is to reprove their results, and extend to a much more general set-up

A dynamical restatement

A weighted normed version: $A \in \widehat{D}_\nu^{\mathbf{r},\mathbf{s}}$

$$\Updownarrow$$

$$\exists\ \rho < \sqrt{\gamma_\nu} \text{ such that } g_t^{\mathbf{r},\mathbf{s}}\Lambda_A \cap B_\nu(\rho) \neq \varnothing \text{ for all large enough } t$$

$$\Updownarrow$$

$$\exists\ \rho < \sqrt{\gamma_\nu} \text{ such that } g_t^{\mathbf{r},\mathbf{s}}\Lambda_A \notin K_\nu(\rho) \text{ for all large enough } t$$

$$\Updownarrow$$

$$g_{\mathbb{R}_+}^{\mathbf{r},\mathbf{s}}\Lambda_A \text{ eventually avoids } L_\nu$$

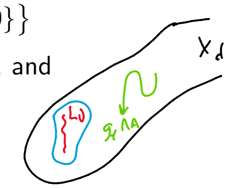
($\exists$ a neighborhood $U \supset L_\nu$ such that $g_t^{\mathbf{r},\mathbf{s}}\Lambda_A \notin U$ for all large enough t),

where $$K_\nu(\rho) := \big\{\Lambda \in X_d : \Lambda \cap B_\nu(\rho) = \{0\}\big\}$$

(compact, and with non-empty interior if $\rho < \sqrt{\gamma_\nu}$), and

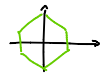
$$L_\nu := K_\nu(\sqrt{\gamma_\nu}) = \bigcap_{\rho < \sqrt{\gamma_\nu}} K(\rho)$$

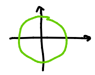
(the critical locus for the norm ν).



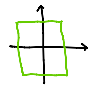
Examples of critical loci

- finite subsets of X_d





- smooth closed curves in X_2
- the union of two closed horocycles



- new examples: any closed subset of S^1 is isometric to some critical locus

[K-Rao-Sathiamurthy]

Main results

Theorem 1 [K-Rao]: For any norm ν on $\mathbb{R}^d$ and any weights $\mathbf{r}, \mathbf{s}$, the set $\widehat{D}_\nu^{\mathbf{r},\mathbf{s}}$ has Lebesgue measure zero.

Same proof as for Theorem DS1:

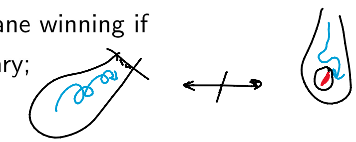
- by ergodicity in the unweighted case;
- by equidistribution of translates $g_t^{\mathbf{r},\mathbf{s}}\{\Lambda_A : A \in M_{m\times n}(\mathbb{R})\}$ in the weighted case [K-Weiss '08, K-Margulis '12].

Conjecture: For any norm ν on $\mathbb{R}^d$ and any weights $\mathbf{r}, \mathbf{s}$, the set $\widehat{D}_\nu^{\mathbf{r},\mathbf{s}}$ is hyperplane winning.

($\Longrightarrow$ of full Hausdorff dimension, stable under countable intersections)

Theorem 2 [K-Rao]: $\widehat{D}_\nu^{\mathbf{r},\mathbf{s}}$ is hyperplane winning if

- ν is Euclidean on $\mathbb{R}^d$, $\mathbf{r}, \mathbf{s}$ arbitrary;
- $m = n = 1$, ν any norm on $\mathbb{R}^2$.



Note: there is no Theorem 2′ in this generality,

boundedness of the orbit $\not\leftrightarrow$ eventually avoiding the critical locus L_ν.

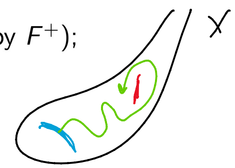
Exceptional subsets of G/Γ

Recall:

$$\widehat{D}_{\nu}^{\mathbf{r},\mathbf{s}} = \left\{A \in M_{m\times n}(\mathbb{R}) \colon \{g_t^{\mathbf{r},\mathbf{s}}\Lambda_A : t \geq 0\} \text{ eventually avoids } L_\nu\right\}.$$

Thus Theorem 2 gets to be a special case of a more general set-up:

- G a connected Lie group, $\Gamma \subset G$ discrete, $X = G/\Gamma$;
- $F^+ = \{g_t : t \geq 0\}$ a one-parameter (Ad-diagonalizable) semigroup;
- $H \subset G$ a connected subgroup (normalized by F^+);
- $Z \subset X$ a (compact) C^1 submanifold;
- $x \in X$ an arbitrary point.



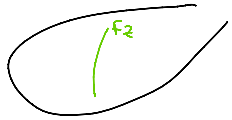
Question: what conditions are sufficient for the existence of
(a lot of) $h \in H$ such that F^+hx eventually avoids Z?
$\uparrow$
(full Hausdorff dimension, winning, hyperplane winning)

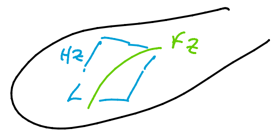
(F, H)-transversality

Definition: Let Z be a (compact) C^1 submanifold of X,
H a connected subgroup of G,
F a one-parameter subgroup of G.

Say that Z is (F, H)-transversal if for every $z \in Z$ one has

$$T_z(Fz) \not\subset T_z Z \quad \text{and} \quad T_z(Hz) \not\subset T_z Z \oplus T_z(Fz).$$





Theorem 3 [K '99]: Let H be the expanding horospherical subgroup relative to F^+, and let $Z \subset X$ be a compact (F, H)-transversal C^1 submanifold. Then for any $x \in X$ the set

$$\{h \in H : F^+ hx \text{ eventually avoids } Z\}$$

has full Hausdorff dimension.

A weaker version of Theorem 2

Note: when $F = \left\{ g_t = \begin{pmatrix} e^{t/m} I_m & 0 \\ 0 & e^{-t/n} I_n \end{pmatrix} \right\}$
and $F^+ = g_{\mathbb{R}_+}$,

the expanding horospherical subgroup relative to F^+

is precisely $H = \left\{ \begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix} : A \in M_{m \times n}(\mathbb{R}) \right\}$.

Hence Theorem 3 immediately implies

Corollary: With F, H as above and a norm ν on $\mathbb{R}^d$, suppose that the critical locus L_ν is (F, H)-transversal. Then $\widehat{D}_\nu^{m,n}$ has full Hausdorff dimension.

(an unweighted version of Theorem 2 without the winning property)

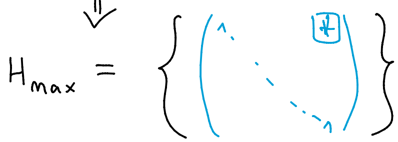
To include weights and upgrade to hyperplane winning, need recent work of An–Guan–K.

$(F, H_{\max})$-transversality

There we introduced a subgroup $H_{\max}$ of H:
maximally expanding horospherical subgroup relative to F^+.

$r_1 > \dots > r_m\,,\ s_1 < \dots < s_n$

$\Downarrow$



Theorem 4 [An–Guan–K]: Let $H_{\max}$ be as above,
and let $Z \subset X$ be a $(F, H_{\max})$-transversal C^1 submanifold.
Then for any $x \in X$ the set

$$\{h \in H : F^+hx \text{ eventually avoids } Z\}$$

is hyperplane winning.

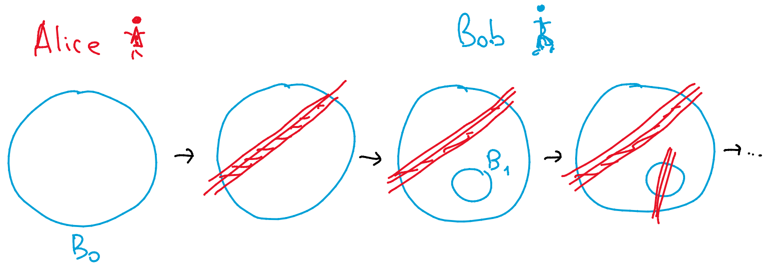
Background

Generalizations

New results and conjectures

Thanks

Hyperplane winning



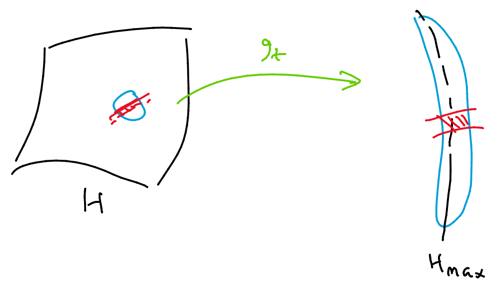
S is hyperplane winning

if Alice has a strategy guaranteeing that

$$\bigcap_{i=0}^{\infty} B_i \cap S \neq \emptyset$$

Background
Generalizations
New results and conjectures
Thanks

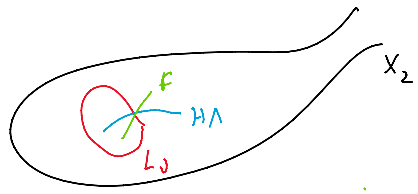
Winning using transversailty



Checking the transversality of L_ν

When $d = 2$: use the work of Mahler on critical lattices for convex symmetric irreducible domains.

(L_ν, F and H are in "general position")



When ν is Euclidean: use the results of Korkine and Zolotarev (L_ν is a finite union of SO(d)-orbits).

Thank you for your attention!

PS: Anurag Rao (on the job market next year)