Non-divergence in the space of lattices

Nicolas de Saxcé

January 20, 2022

Abstract

Using Harder-Narasimhan filtrations and Grayson polygons to describe the geometry of the space of lattices, we give a new proof of the Kleinbock-Margulis quantitative non-divergence estimate.

Introduction

Non-divergence estimates were first introduced by Margulis [8] in his study of unipotent flows on the space of lattices. They were later refined by Dani [5] and became a fundamental tool for the study of unipotent orbits in quotients of Lie groups by lattices, which led in particular to Ratner's work on the subject [10, 11]. More recently, these estimates were made quantitative by Kleinbock and Margulis [6] and used to solve long-standing conjectures of Sprindzuk and Baker in diophantine approximation. Since then, they have found many other applications in the field; among those, we only cite three recent articles [1], [2] and [15] where the interested reader can find more references. As far as we know, until now, all the refinements of non-divergence estimates were proved following the strategy of the original paper of Margulis [8] on the subject.

The goal of this note is two-fold: first, we present a statement of the quantitative non-divergence estimates in terms of Grayson polygons and Harder-Narasimhan filtrations, two objects used to describe the geometry of lattices in $\mathbb{R}^d$, and then we give a new proof for them, based on these geometric tools. We hope that the pictures associated to the Grayson polygons can help the reader visualize and better understand the meaning of the technical statement discussed here.

1 Statement of the non-divergence estimate

Given a lattice Δ in $\mathbb{R}^d$, for $k \in \{1, \dots, d\}$, we define

$$\lambda_k(\Delta) = \inf\{\lambda \mid \exists v_1, \dots, v_k \in \Delta, \text{ linearly independent with } \forall i, \|v_i\| \leq \lambda\}.$$

We have of course $\lambda_1(\Delta) \leq \lambda_2(\Delta) \leq \cdots \leq \lambda_d(\Delta)$, and the numbers $\lambda_k(\Delta)$ are called the *successive minima* of the lattice Δ. The successive minima describe the rough position of Δ inside the space of lattices $\Omega = \mathrm{GL}_d(\mathbb{R})/\mathrm{GL}_d(\mathbb{Z})$

endowed with a distance induced by a right-invariant Riemannian metric on $\mathrm{GL}_d(\mathbb{R})$: two lattices with the same successive minima are at bounded distance from each other. Alternatively, the position of a lattice Δ in Ω can be described by the *successive covolumes*, given by

$$\mu_k(\Delta) = \min\{\|v_1 \wedge \cdots \wedge v_k\| \ ; \ v_1, \ldots, v_k \in \Delta \text{ linearly independent}\},$$

where $\|v_1 \wedge \cdots \wedge v_k\|$ denotes the covolume of the sublattice spanned by $v_1, \ldots, v_k$ in its real span. By Minkowski's second theorem, the numbers $\mu_k(\Delta)$ determine the successive minima within multiplicative constants depending only on d:

$$\mu_k(\Delta) \asymp \lambda_1(\Delta) \cdots \lambda_k(\Delta).$$

1.1 The Grayson polygon of a lattice

Let Δ be a lattice in $\mathbb{R}^d$. An important property of the covolume function on the set of sublattices of Δ is that for every sublattices V and W, one has $\|V \cap W\|\|V + W\| \leq \|V\|\|W\|$. Equivalently, if τ denotes the function $\tau : V \mapsto \log\|V\|$ on sublattices of Δ, then τ is *submodular*, i.e. satisfies

$$\tau(V \cap W) + \tau(V + W) \leq \tau(V) + \tau(W).$$

To any such function the so-called "slope formalism" associates two canonical objects: the Grayson polygon and the Harder-Narasimhan filtration. We briefly recall their definition and elementary properties here, and refer the reader to [4, §4] or [3, §1.3] for more details on their construction.

The *Grayson polygon* of a lattice Δ is a convex function $L_\Delta : [0, d] \to \mathbb{R}$ that allows one to understand all the successive covolumes — or the successive minima — together in a nice picture. By definition, it is the maximal convex function on $[0, d]$ such that $L_\Delta(0) = 0$ and for each $k \in \{1, \ldots, d\}$, $L_\Delta(k) \leq \log \mu_k(\Delta)$. An example of a Grayson polygon is given in Figure 2 below, where each point $(k, \log \mu_k(\Delta))$; $k = 0, \ldots, 6$, is marked with a red cross, and the graph of L_Δ is plotted in blue.

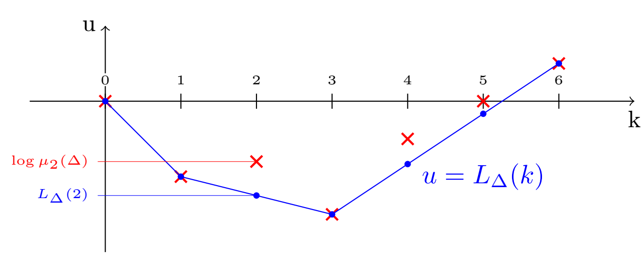

Figure 1: Graph of $k \mapsto L_\Delta(k)$, for $k \in [0, 6]$

At each index $k \in \{1, \ldots, d\}$ where the graph of L_Δ has an angle, i.e.

$$L_\Delta(k) - L_\Delta(k-1) < L_\Delta(k+1) - L_\Delta(k),$$

there is a unique primitive sublattice Δ_k of rank k in Δ such that $\mu_k(\Delta) = \|\Delta_k\|$. Moreover, if the angles of L_Δ occur at the indices $k_1 < k_2 < \cdots < k_s$, then the sublattices Δ_k form a partial flag

$$\{0\} < \Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_s} < \Delta$$

called the *Harder-Narasimhan* filtration of Δ.

Example. When the Harder-Narasimhan filtration of Δ is trivial, or equivalently when the Grayson polygon L_Δ is constant equal to zero, one says that Δ is *stable*, or *semistable*. It is worth noting that when the space Ω is endowed with the Haar probability measure, the measure of the set of stable lattices tends to one as the dimension goes to infinity [14].

1.2 Quantitative non-divergence

The goal of this note is to describe the behavior of the Grayson polygon and of the Harder-Narasimhan filtration a random lattice $\Delta = x\mathbb{Z}^d$, where $x \in \mathrm{GL}_d(\mathbb{R})$ is a random element distributed according to some probability ν on $\mathrm{GL}_d(\mathbb{R})$ satisfying natural regularity conditions. To state the result, it is convenient to extend the definition on a Grayson polygon to arbitrary compact subsets $S \subset \mathrm{GL}_d(\mathbb{R})$. For each $k \in \{1, \ldots, d\}$, we write $\mathcal{W}_k(\mathbb{Z})$ for the set of non-zero pure k-vectors in $\wedge^k \mathbb{Z}^d$ and set

$$\mu_k(S) = \begin{cases} 1 & \text{if } k = 0 \\ \inf_{\mathbf{w} \in \mathcal{W}_k(\mathbb{Z})} \sup_{x \in S} \|x\mathbf{w}\| & \text{if } 1 \leq k \leq d. \end{cases}$$

As before, we then define $L_S : [0, d] \to \mathbb{R}$ to be the maximal convex function such that for each k, $L_S(k) \leq \log \mu_k(S)$. When $S = \{x\}$ is reduced to a singleton, one recovers the Grayson polygon of the lattice $x\mathbb{Z}^d$; we then simply write L_x instead of $L_{\{x\}}$. One should note the following two properties:

1. If L_S has an angle at k, then there exists a sublattice $\mathbf{w}_k$ of rank k in $\mathbb{Z}^d$ such that for all $x \in S$, $\|x\mathbf{w}_k\| \leq e^{L_S(k)} = \mu_k(S)$. (If S is not reduced to a singleton, this sublattice may not be unique.)

2. For all $k \in \{1, \ldots, d\}$ and all $x \in S$, there exists a sublattice $\mathbf{w}_{k,x}$ of rank k in $\mathbb{Z}^d$ such that $\|x\mathbf{w}_{k,x}\| \leq Ke^{L_S(k)}$, for some constant K depending only on d.

The first assertion is clear by definition of L_S and $\mu_k(S)$, and the second follows from the first, using Minkowski's second theorem on successive minima. Here is another way to understand this second property: For every $x \in S$, the graph of $k \mapsto L_x(k)$, which up to bounded additive constants represents the minimal covolume of a sublattice of rank k in $x\mathbb{Z}^d$, for $k = 1, \ldots, d$, lies below the graph of L_S.

Now let ν be a measure on $\mathrm{GL}_d(\mathbb{R})$, and $S = \operatorname{Supp} \nu$. The above observations show that for ν-almost every x, all points on the graph of L_x are under the graph of L_S. Under some regularity assumptions on ν, the next theorem gives a converse statement: the ν-measure of points such that L_x contains one point below $L_S - \log\frac{1}{\varepsilon}$ goes to zero as $\varepsilon > 0$ goes to zero. Recall that given $D > 0$, a finite measure ν on a metric space X is D-*doubling* if for every $x \in \operatorname{Supp}\nu \cap X$ and every $r > 0$, $\nu(B(x, 2r) \leq D\nu(B(x, r))$, and that a function $f : X \to \mathbb{R}$ is (C, α)-*good* on X with respect to the measure ν if we have, for every ball $B = B(x, r) \subset X$ centered at $x \in \operatorname{Supp}\nu$,

$$\nu(\{x \in B \mid |f(x)| \leq \varepsilon\}) \leq C \left(\frac{\varepsilon}{\|f\|_{\nu,B}}\right)^{\alpha} \nu(B),$$

where $\|f\|_{\nu,B} = \sup_{x\in B\cap\operatorname{Supp}\nu} |f(x)|$. In the sequel, we endow $\mathrm{GL}_d(\mathbb{R})$ with the metric induced by any Euclidean metric on the space $M_d(\mathbb{R})$ of $d \times d$ matrices with real entries.

Theorem 1 (Non-divergence with Grayson polygons)**.** *Given positive constants $D, C_0, \alpha_0 > 0$, there exist $C, \alpha > 0$ such that the following holds. Let B be a ball in $\mathrm{GL}_d(\mathbb{R})$ and ν a finite measure on $\mathrm{GL}_d(\mathbb{R})$, D-doubling on $5B$. Assume that for every $k \in \{1, \dots, d\}$, for every non-zero $\mathbf{w} = v_1 \wedge \cdots \wedge v_k$ in $\wedge^k\mathbb{Z}$, the map $x \mapsto \|x\mathbf{w}\|$ is (C_0, α_0)-good on $5B$ with respect to ν, and let $S = B \cap \operatorname{Supp}\nu$. For $\varepsilon \in (0, 1)$, consider the set $\mathcal{B}_\varepsilon$ of points $x \in \mathrm{GL}_d(\mathbb{R})$ satisfying*

$$\exists k \in \{1, \dots, d\} : \ L_x(k) \leq L_S(k) + \log\varepsilon$$

Then,

$$\nu(\mathcal{B}_\varepsilon \cap B) \leq C\varepsilon^\alpha \nu(B).$$

In other words, if ν is a probability measure satisfying the assumptions of the theorem and $\Delta = x\mathbb{Z}^d$, with x chosen randomly in B according to the law ν, then the probability that the graph of L_Δ lies between that of L_S and $L_S + \log\varepsilon$ is bounded below by $1 - C\varepsilon^\alpha$.

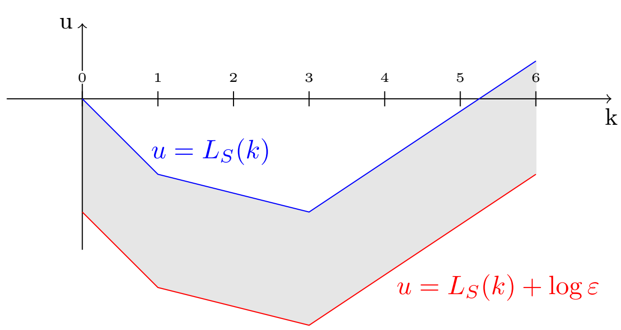

Figure 2: With probability $\geq 1 - C\varepsilon^\alpha$ the graph of L_Δ lies in the gray area.

One consequence of the theorem is that in some sense the Harder-Narasimhan filtration of $\Delta = x\mathbb{Z}^d$ is constant with very high probability. This is the content of the following corollary.

Corollary 1 (Partial flag associated to a good measure)**.** *There exist constants $C', \alpha' > 0$ such that under the assumptions of Theorem 1, the following holds. Let*

$$J_S(\varepsilon) = \{k \in \{1, \ldots, d-1\} \mid L_S(k+1) + L_S(k-1) - 2L_S(k) \geq -\log \varepsilon\}$$

and write $J_S(\varepsilon) = \{k_1 < \cdots < k_s\}$. For $\varepsilon > 0$ small enough, there exists a unique partial flag

$$0 < V_{k_1}^S < \cdots < V_{k_s}^S < \mathbb{Z}^d$$

such that for each $i = 1, \ldots, s$

$$\nu(\{x \in B \mid \forall i,\ \mu_{k_i}(x\mathbb{Z}^d) \text{ is attained on } V_{k_i}^S\}) \geq 1 - C'\varepsilon^{\alpha'}.$$

Remark 1. It formally follows from the statement of the corollary that, provided $\varepsilon > 0$ is small enough, the subspaces $V_{k_i}^S$ do not depend on ε. But the set of indices $J_S(\varepsilon) = \{k_1, \ldots, k_s\}$ can decrease as ε goes to zero.

While this note was still being written, Lindenstrauss, Margulis, Mohammadi and Shah informed us that they had independently observed that Margulis's proof also allowed to derive the two results above, a fact that they used in their work on the quantitative behavior of unipotent flows [7, Theorem 5.3]. Our motivation to prove Theorem 1 was another application to diophantine approximation, described in a paper written in collaboration with Emmanuel Breuillard [3]. One advantage of the proof given here is that it can be generalized to obtain an intrinsic statement for the non-divergence estimate in an arbitrary semi-simple $\mathbb{Q}$-group G, without having to embed G in some linear group GL_d. For more details on this aspect, and for other applications of such estimates, the reader is referred to [12].

2 Proofs

The argument used in this note starts with an estimate for the probability of having a small sublattice of fixed rank k, Theorem 2 below, a statement that bears its own interest. Before we turn to its proof, we recall a useful interpretation of the successive covolumes $\mu_k(\Delta)$ of a lattice Δ using exterior powers. If Δ is a lattice in $\mathbb{R}^d$, then $\wedge^k\Delta$ is a lattice in $\wedge^k\mathbb{R}^d \simeq \mathbb{R}^{\binom{d}{k}}$ and, its successive minima are determined by those of Δ by the following formula, due to Mahler. See [13, Theorem 7A, page 109].

Lemma 1 (Successive minima of exterior powers)**.** *Let Δ be a lattice in $\mathbb{R}^d$. Within multiplicative constants depending only on d, the successive minima of $\wedge^k\Delta$ are given by*

$$\lambda_\tau(\wedge^k\Delta) \asymp \lambda_{\tau_1}(\Delta)\lambda_{\tau_2}(\Delta)\ldots\lambda_{\tau_k}(\Delta); \quad \tau = \{\tau_1, \ldots, \tau_k\} \subset \{1, \ldots, d\}.$$

The order on the set of indices τ depends on the values of $\lambda_1(\Delta), \ldots, \lambda_d(\Delta)$, but the first and second minima of $\wedge^k\Delta$ are always given by $\{1, \ldots, k\}$ and

$\{1, \ldots, k-1, k+1\}$*, so that*

$$\lambda_{\{1,\ldots,k\}}(\wedge^k \Delta) \asymp \mu_k(\Delta) \quad \text{and} \quad \lambda_{\{1,\ldots,k-1,k+1\}}(\wedge^k \Delta) \asymp \frac{\lambda_{k+1}(\Delta)}{\lambda_k(\Delta)}\mu_k(\Delta).$$

2.1 Sublattices of rank k

For non-divergence estimates as in [6], one is given a map $h : X \to \mathrm{GL}_d(\mathbb{R})$ and a ball $B \subset X$ such that for every sublattice $\mathbf{w} \subset \mathbb{Z}^d$ there exists a point $x \in B$ such that

$$\|h(x)\mathbf{w}\| \geq \rho, \tag{1}$$

and concludes that the proportion of points in B such that $\lambda_1(h(x)\mathbb{Z}^d) \leq \varepsilon\rho$ is bounded above by $C\varepsilon^\alpha$, for some constants $C, \alpha > 0$ depending on some regularity properties of h. Of course, it is readily seen by considering the map

$$\begin{array}{ccc} [0,1] & \to & \mathrm{GL}_2(\mathbb{R}) \\ x & \mapsto & \begin{pmatrix} x & 1 \\ x^2 - \varepsilon & x \end{pmatrix} \end{array}$$

that it is not enough to assume (1) only for vectors in $\mathbb{Z}^d$. Yet, the main point of this section is to explain what can be concluded from such an assumption, or more generally, what can be said when one only assumes (1) for sublattices $\mathbf{w}$ of fixed rank k.

Theorem 2 (Non-divergence for k-sublattices)**.** *Given positive constants D, C_0 and $\alpha_0 > 0$, there exist $C_1, \alpha_1 > 0$ such that the following holds for any $k \in \{1, \ldots, d\}$ and any choice of parameters $M > 1$ and $\rho_k > 0$. Let B be a ball in $\mathrm{GL}_d(\mathbb{R})$, let ν be a finite measure on $\mathrm{GL}_d(\mathbb{R})$, doubling on $5B$, and set $S = B \cap \operatorname{Supp} \nu$. Assume that for every non-zero $\mathbf{w} = v_1 \wedge \cdots \wedge v_k$ in $\wedge^k \mathbb{Z}$:*

1. *The map $x \mapsto \|x\mathbf{w}\|$ is (C_0, α_0)-good on $5B$ with respect to ν.*
2. $\sup_{x \in S} \|x\mathbf{w}\| \geq \rho_k$

For $\varepsilon \in (0, 1)$, consider the set

$$\mathcal{A}_\varepsilon^{(k)} = \left\{ x \in B \;\middle|\; \begin{array}{c} \varepsilon^M \rho_k \leq \mu_k(x\mathbb{Z}^d) \leq \varepsilon\rho_k \\ \lambda_k(x\mathbb{Z}^d) < \varepsilon\lambda_{k+1}(x\mathbb{Z}^d) \end{array} \right\}.$$

Then,

$$\nu(\mathcal{A}_\varepsilon^{(k)}) \leq C_1 M \varepsilon^{\alpha_1} \nu(B).$$

Proof. To simplify notation, we shall write $\mu_k(x)$ and $\lambda_k(x)$ for $\mu_k(x\mathbb{Z}^d)$ and $\lambda_k(x\mathbb{Z}^d)$, respectively. Note also that it is enough to prove the proposition for ε smaller than an arbitrary constant.

<u>First case:</u> $1 < M \leq \frac{4}{3}$.

For $x \in S$, choose a pure k-vector $\mathbf{w}_x = v_1^x \wedge \cdots \wedge v_k^x$ such that $\mu_k(x) = \|x\mathbf{w}_x\|$. Then, if $x \in \mathcal{A}_\varepsilon^{(k)} \cap S$, let $R_x > 0$ be maximal such that

$$\forall y \in B(x, R_x) \cap \operatorname{Supp} \nu, \quad \|y\mathbf{w}_x\| \leq \varepsilon^{\frac{1}{2}} \rho_k.$$

Let $B_x = B(x, R_x)$. Note that $B_x \subset 3B$ and $2B_x \subset 5B$ because $\sup_{y\in S} \|y\mathbf{w}_x\| \geq \rho_k$. Moreover, by maximality of R_x,

$$\sup_{y\in 2B_x} \|y\mathbf{w}_x\| \geq \varepsilon^{\frac{1}{2}} \rho_k. \tag{2}$$

We claim that for any $y \in \mathcal{A}_\varepsilon^{(k)} \cap B_x$,

$$\mathbf{w}_y = \mathbf{w}_x. \tag{3}$$

In words, up to sign, $\mathbf{w}_x$ is the only non-zero vector in $\wedge^k \mathbb{Z}^d$ satisfying $\|y\mathbf{w}_x\| = \mu_k(y)$. Indeed, if $y \in \mathcal{A}_\varepsilon^{(k)}$, then $\lambda_{k+1}(y) > \varepsilon^{-1}\lambda_k(y)$, so, by Lemma 1, for any $\mathbf{w} \in \wedge^k \mathbb{Z}^d$ that is not collinear to $\mathbf{w}_y$, for small enough $\varepsilon > 0$,

$$\|y\mathbf{w}\| \gg \varepsilon^{-1}\mu_k(y) \geq \varepsilon^{\frac{1}{3}} \rho_k,$$

and since $\|y\mathbf{w}_x\| \leq \varepsilon^{\frac{1}{2}}\rho_k$, we must have (3). Thus,

$$y \in \mathcal{A}_\varepsilon^{(k)} \cap B_x \quad \Rightarrow \quad \|y\mathbf{w}_x\| \leq \varepsilon\rho_k.$$

Using that $y \mapsto \|y\mathbf{w}_x\|$ is (C_0, α_0)-good on $2B_x \subset 5B$ and satisfies (2), we find

$$\nu(\mathcal{A}_\varepsilon^{(k)} \cap B_x) \leq C_0 \varepsilon^{\frac{\alpha_0}{2}} \nu(2B_x) \leq DC_0 \varepsilon^{\frac{\alpha_0}{2}} \nu(B_x).$$

To conclude, we use Besicovitch's covering theorem [9, page 30] in the Euclidean space $M_d(\mathbb{R}) \simeq \mathbb{R}^{d^2}$: take a cover of $\mathcal{A}_\varepsilon^{(k)} \cap B$ by balls B_{x_i}, $i \in \mathbb{N}$ with intersection multiplicity bounded by the Besicovitch constant K of $M_d(\mathbb{R})$, and bound

$$\begin{aligned}
\nu(\mathcal{A}_\varepsilon^{(k)} \cap B) &\leq \sum_{i\in\mathbb{N}} \nu(\mathcal{A}_\varepsilon^{(k)} \cap B_{x_i}) \\
&\leq DC_0 \varepsilon^{\frac{\alpha_0}{2}} \sum_{i\in\mathbb{N}} \nu(B_{x_i}) \\
&\leq C_0 D^3 K \varepsilon^{\frac{\alpha_0}{2}} \nu(B).
\end{aligned}$$

This is the desired result, with $C_1 = C_0 D^3 K$ and $\alpha_1 = \frac{\alpha_0}{2}$.
General case: $M \geq \frac{4}{3}$
We cover the interval $[\varepsilon^M \rho_k, \varepsilon\rho_k]$ with intervals of the form $[\varepsilon^{\frac{j+1}{3}}\rho_k, \varepsilon^{\frac{j}{3}}\rho_k]$, for $j = 3, 4, \ldots, 3M$. For each j, we may apply the first part of the proof with ρ_k replaced by $\varepsilon^{\frac{j}{3}-1}\rho_k$ and find that the ν-measure of points in B satisfying

$$\begin{cases} \varepsilon^{\frac{j+1}{3}}\rho_k \leq \mu_k(x\mathbb{Z}^d) \leq \varepsilon^{\frac{j}{3}}\rho_k \\ \lambda_k(x\mathbb{Z}^d) < \varepsilon\lambda_{k+1}(x\mathbb{Z}^d) \end{cases}$$

is bounded above by $C_1\varepsilon^{\alpha_1}\nu(B)$. To conclude, it suffices to sum all these inequalities; we obtain

$$\nu(\mathcal{A}_\varepsilon^{(k)} \cap B) \leq 3C_1 M \varepsilon^{\alpha_1} \nu(B).$$

$\square$

2.2 Full flags

In this final paragraph, we derive Theorem 1 and its corollary. The proof of Theorem 1 is based on Theorem 2, applied for some well-chosen $k \in \{1, \dots, d\}$. The derivation of Theorem 1 from Theorem 2 is based on the following elementary observation: there exists a constant $\tau > 0$ depending only on d such that if $A > 0$ is some parameter and $L_x \le L_S$ any two convex functions on $\{0, \dots, d\}$ such that $L_x(i) \le L_h(i) - A$ for some $i \in \{0, \dots, d\}$, then one may find some $k \in \{0, \dots, d\}$ such that $L_x(k) \le L_S(k) - \tau A$ and such that the angle at k of the function L_x is at least τA (or $k = 0$ or d). The formal argument is given below.

Proof of Theorem 1. Of course, it is enough to prove the estimate for ε smaller than an arbitrarily small constant depending on d, and in particular, we may assume $\varepsilon \in (0, \frac{1}{2})$. Moreover, if $\mathcal{B}_\varepsilon^0$ denotes the set of points $x \in \mathrm{GL}_d(\mathbb{R})$ satisfying

$$\begin{cases} \exists i \in \{1, \dots, d\} : \ \log \mu_i(x) \le L_S(i) + \log \varepsilon \\ \forall i \in \{1, \dots, d\}, \ \log \mu_i(x) \ge L_S(i) + 2 \log \varepsilon \end{cases}$$

then it is enough to show that

$$\nu(\mathcal{B}_\varepsilon^0 \cap B) \le C \varepsilon^\alpha \nu(B).$$

Indeed, from there, using the inclusion $\mathcal{B}_\varepsilon \subset \bigcup_{n=0}^\infty \mathcal{B}^0_{\varepsilon^{2^n}}$, we find

$$\nu(\mathcal{B}_\varepsilon \cap B) \le C \nu(B) \sum_{n=0}^{\infty} \varepsilon^{2^n \alpha} \le C' \varepsilon^\alpha \nu(B).$$

Let $\tau = \frac{1}{2d^2}$ and $\alpha = \tau \alpha_1$, where C_1, α_1 are given by Theorem 2. Since $x \mapsto \det x$ is (C_0, α_0)-good with respect to ν, we have

$$\nu(\{x \in B \mid \mu_d(x) \le \varepsilon^\tau e^{L_S(d)}\}) \le C_0 \varepsilon^{\alpha_0 \tau} \nu(B).$$

Next, we claim that if $x \in \operatorname{Supp} \nu$ is such that $\mu_d(x) \ge \varepsilon^\tau e^{L_S(d)}$ and for some i, $\log \mu_i(x) \le L_S(i) + \log \varepsilon$, then there must exist some k such that

$$\mu_k(x) \le \varepsilon^\tau e^{L_S(k)} \quad \text{and} \quad \lambda_k(x) \le \varepsilon^\tau \lambda_{k+1}(x). \tag{4}$$

From there, one concludes using Theorem 2 with ε^τ in place of ε, and $M = \frac{2}{\tau}$:

$$\begin{aligned} \nu(\mathcal{B}_\varepsilon^0 \cap B) &\le \nu(\{x \in B \mid \mu_d(x) \le \varepsilon^\tau e^{L_S(d)}\}) + \sum_{k=1}^{d-1} \nu(\mathcal{A}^{(k)}_{\varepsilon^\tau} \cap B) \\ &\le C_0 \varepsilon^{\alpha_0 \tau} \nu(B) + (d-1) C_1 \varepsilon^{\tau \alpha_1} \nu(B) \\ &\le C \varepsilon^\alpha \nu(B). \end{aligned}$$

Now we turn to the justification of the existence of k satisfying (4). We argue by contrapositive: assuming

$$\begin{cases} \mu_d(x) \ge \varepsilon^\tau e^{L_S(d)} \\ \forall k = 1, \dots, d, \ \mu_k(x) \ge \varepsilon^\tau e^{L_S(k)} \quad \text{or} \quad \lambda_k(x) \ge \varepsilon^\tau \lambda_{k+1}(x) \end{cases} \tag{5}$$

we want to show that for all $i = 1, \ldots, d$, $\log \mu_i(x) \geq L_S(i) + \log \varepsilon$. Using the equivalence between Grayson polygons and successive covolumes, one sees that for $\varepsilon > 0$ small enough assumption (5) implies

$$\begin{cases} L_x(d) \geq L_S(d) + \tau \log \varepsilon \\ \forall k = 1, \ldots, d,\ L_x(k) \geq L_S(k) + 2\tau \log \varepsilon \quad \text{or} \quad \log \frac{\lambda_{k+1}(x)}{\lambda_k(x)} \leq 2\tau |\log \varepsilon|. \end{cases} \tag{6}$$

Now let $i \in \{0, \ldots, d\}$ be arbitrary, and choose a minimal interval $[k_1, k_2]$ (possibly reduced to $\{i\}$) containing i such that

$$L_x(k_j) \geq L_S(k_j) + \tau \log \varepsilon \qquad j = 1, 2.$$

Because of assumption (6), all the angles of L_x on the segment $[k_1, k_2]$ are bounded above by $2\tau|\log \varepsilon|$, and the distance from L_x to L_S at k_1 and k_2 is bounded above by $2\tau|\log \varepsilon|$. But L_x and L_S are convex functions and $L_x \leq L_S$, so we may bound the maximal distance from L_x to L_S on the whole segment $[k_1, k_2]$ by $2\tau|\log \varepsilon| + \tau(k_1 - k_2)^2|\log \varepsilon|$. In particular,

$$L_x(i) \geq L_S(i) + [2 + (k_1 - k_2)^2]\tau \log \varepsilon \geq L_S(i) + 2d^2\tau \log \varepsilon.$$

Now by definition, $\log \mu_i(x) \geq L_x(i)$, so that recalling our choice $\tau = \frac{1}{2d^2}$, this proves the desired inequality. □

To conclude, we prove Corollary 1 on the partial flag associated to the measure ν.

Proof of Corollary 1. We write $\|f\| = \sup_{t \in [0,d]} |f|$ for the supremum norm on real-valued functions on the segment $[0, d]$. By Theorem 1,

$$\nu(\{x \in B \mid \|L_x - L_S\| \geq -\log \varepsilon\}) \leq C\varepsilon^\alpha \nu(B).$$

Given a constant $C_1 > 0$, if $\varepsilon > 0$ is sufficiently close to 0, this inequality implies in particular that there exists $x_1 \in S$ such that

$$\|L_{x_1} - L_S\| \leq -\frac{\log \varepsilon}{C_1}.$$

Then, provided C_1 has been chosen large enough, for each $k = 1, \ldots, d-1$,

$$\left| \log \frac{\lambda_{k+1}(x_1)}{\lambda_k(x_1)} - L_S(k+1) - L_S(k-1) + 2L_S(k) \right| \leq -\frac{\log \varepsilon}{2}.$$

So, for $k \in J_S(\varepsilon)$, we must have

$$\frac{\lambda_{k+1}(x_1)}{\lambda_k(x_1)} \geq \varepsilon^{-\frac{1}{2}}.$$

By Lemma 1, this implies that for $k \in J_S(\varepsilon)$, there exists a unique sublattice $\mathbf{w}_k$ of rank k in $\mathbb{Z}^d$ such that $\mu_k(x_1) = \|x_1 \mathbf{w}_k\|$, and that for every $\mathbf{w} \in \wedge^k \mathbb{Z}^d$ linearly independent with $\mathbf{w}_k$,

$$\|x_1 \mathbf{w}\| \gg \varepsilon^{-\frac{1}{2}} \mu_k(x_1) \geq \varepsilon^{-\frac{1}{2} + \frac{1}{C_1}} \mu_k(S).$$

Thus, for each $k \in J_S(\varepsilon)$, the sublattice $\mathbf{w}_k$ is the unique sublattice of rank k satisfying

$$\sup_{x \in S} \|x\mathbf{w}_k\| = \mu_k(S).$$

But by construction, the subspace V_k corresponding to $\mathbf{w}_k$ is the k-dimensional term in the Harder-Narasimhan filtration of $x_1\mathbb{Z}^d$, so that if $J_S(\varepsilon) = \{k_1 < \cdots < k_s\}$, then

$$\{0\} < V_{k_1} < \cdots < V_{k_s} < \mathbb{Z}^d.$$

The reasoning made above for the chosen point x_1 is valid for any other x such that $\|L_x - L_S\| \leq -\frac{\log \varepsilon}{C_1}$ so that for any such x, the subspaces $V_{k_1} < \cdots < V_{k_s}$ occur in the Harder-Narasimhan filtration of $x\mathbb{Z}^d$, i.e. $\mu_{k_i}(x) = \|x\mathbf{w}_{k_i}\|$. This finishes the proof of the corollary, because

$$\nu(\{x \in B \mid \|L_x - L_S\| \geq -\frac{\log \varepsilon}{C_1}\}) \leq C'\varepsilon^{\alpha'}\nu(B).$$

$\square$

Acknowledgements: I am indebted to Barak Weiss for suggesting to prove non-divergence estimates using Grayson polygons. It is a pleasure to thank him for his insightful remarks on the subject, and for his comments on an earlier version of this note. I also thank Uri Shapira and Emmanuel Breuillard for motivating discussions, and the anonymous referee for suggestions that helped improve the presentation of the paper.

References

[1] Menny Aka, Emmanuel Breuillard, Lior Rosenzweig, and Nicolas de Saxcé. Diophantine approximation on matrices and Lie groups. *Geom. Funct. Anal.*, 28(1):1–57, 2018.

[2] Victor Beresnevich, Erez Nesharim, and Lei Yang. Winning property of badly approximable points on curves. *arXiv preprint*, 2020.

[3] Emmanuel Breuillard and N. de Saxcé. A subspace theorem for manifolds. *to appear in J. Eur. Math. Soc. (JEMS)*, 2021.

[4] Bill Casselman. Stability of lattices and the partition of arithmetic quotients. *Asian J. Math.*, 8(4):607–637, 2004.

[5] S. G. Dani. Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. *J. Reine Angew. Math.*, 359:55–89, 1985.

[6] D. Y. Kleinbock and G. A. Margulis. Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. of Math. (2)*, 148(1):339–360, 1998.

[7] Lindenstrauss, Margulis, Mohammadi, and Shah. Quantitative behavior of unipotent flows and an effective avoidance principle. *arXiv preprint*, 2019.

[8] G. A. Margulis. On the action of unipotent groups in the space of lattices. In *Lie groups and their representations (Proc. Summer School, Bolyai, János Math. Soc., Budapest, 1971)*, pages 365–370. Halsted, New York, 1975.

[9] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces*. Cambridge university press, 1995.

[10] Ratner. On Raghunathan's measure conjecture.

[11] Ratner. Raghunathan's topological conjecture and distributions of unipotent flows.

[12] Nicolas de Saxcé. Groupes arithmétiques et approximation diophantienne. *mémoire d'habilitation*, 2020.

[13] Wolfgang M. Schmidt. *Diophantine approximation*, volume 785 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980.

[14] B. Weiss and U. Shapira. A volume estimate for the set of stable lattices.

[15] Pengyu Yang. Equidistribution of expanding translates of curves and Diophantine approximation on matrices. *Invent. Math.*, 220(3):909–948, 2020.