FLOWS ON S-ARITHMETIC HOMOGENEOUS SPACES AND APPLICATIONS TO METRIC DIOPHANTINE APPROXIMATION

Dmitry Kleinbock and George Tomanov

Abstract. The goal of this work is to prove quantitative nondivergence estimates for flows on S-arithmetic homogeneous spaces, and then apply them to establish S-arithmetic generalization of conjectures of Baker and Sprindžuk in metric Diophantine approximation. In the present paper we develop the general framework and establish results for polynomial maps. The sequel [KT] will contain applications to Diophantine approximation on differentiable S-arithmetic submanifolds.

0. Introduction

0.1. Actions of unipotent subgroups on homogeneous spaces provide examples of important dynamical systems with numerous applications to geometry and number theory. A fundamental phenomemon discovered by G.A. Margulis in 1971 [M] showed that, in sharp contrast to partially hyperbolic flows, orbits of unipotent flows are never divergent. Margulis' paper gave rise to numerous generalizations and applications. In a series of papers [D1–3], S.G. Dani generalized and strengthened this result. In particular, for one-parameter flows he proved that any orbit returns to big compact subsets with high frequency. A very general explicit estimate for this frequency in terms of the size of compact sets was given in 1998 in the paper of Margulis and the first named author [KM]. To state one of the results of that paper, let us recall that the homogeneous space $\mathrm{SL}(m,\mathbb{R})/\mathrm{SL}(m,\mathbb{Z})$ can be interpreted as the space of unimodular lattices in $\mathbb{R}^m$, and, in view of Mahler's Compactness Criterion [Rag, Corollary 10.9], the family of sets

$$Q_\varepsilon \overset{\text{def}}{=} \left\{ g\,\mathrm{SL}(m,\mathbb{Z}) \;\middle|\; g\mathbb{Z}^n \text{ contains no nonzero vector of length } < \varepsilon \right\} \tag{0.1}$$

gives an exhaustion of $\mathrm{SL}(m,\mathbb{R})/\mathrm{SL}(m,\mathbb{Z})$ by compact subsets. Here is a special case of [KM, Theorem 5.2]:

Theorem. *For $m,d \in \mathbb{N}$, let $h : \mathbb{R}^d \to \mathrm{SL}(m,\mathbb{R})$ be a polynomial map, and fix $\Lambda \in \mathrm{SL}(m,\mathbb{R})/\mathrm{SL}(m,\mathbb{Z})$. Then there exists positive numbers α (depending only on m, d and the degree of h) and C (depending only on m, d, the degree of h and Λ) such that for any positive ε and any ball $B \subset \mathbb{R}^d$ centered at 0 one has*

$$\frac{\lambda\left(\left\{x \in B \;\middle|\; h(x)\Lambda \notin Q_\varepsilon\right\}\right)}{\lambda(B)} \le C\varepsilon^\alpha\,.$$

(Here λ stands for the Lebesgue measure on $\mathbb{R}^d$.)

The results of Margulis and Dani played a crucial role in M. Ratner's [Rat1–2] proof of conjectures of Raghunathan, Dani and Margulis on orbit closures and invariant measures of groups generated by one-parameter unipotent subgroups. Note that Dani used his results to prove that any locally finite measure invariant under a unipotent flow is finite [D2, Theorem 4.3].

0.2. In [BP], A. Borel and G. Prasad initiated the study of actions of S-adic unipotent groups on homogeneous spaces. The generalization of Ratner's results to the S-adic setting was achieved in the papers [MT1, Rat3, MT2]. In the present paper we prove the S-adic analogue of explicit quantitative estimates from [KM], and use them to study locally finite unipotent-invariant measures on S-adic homogeneous spaces. Namely, let K be a number field, S a finite set of its normalized valuations containing all the Archimedean ones, $\mathcal{O}_S$ the ring of S-integers of K, and K_S the direct product of completions K_v of K over $v \in S$. Put $\mathrm{GL}(m, K_S) \stackrel{\mathrm{def}}{=} \prod_{v\in S} \mathrm{GL}(m, K_v)$ and

$$\mathrm{GL}^1(m, K_S) \stackrel{\mathrm{def}}{=} \left\{ (g^{(v)})_{v\in S} \in \mathrm{GL}(m, K_S) \,\middle|\, \prod_{v\in S} |\det(g^{(v)})|_v = 1 \right\} .$$

Then one can interprete the homogeneous space $\mathrm{GL}^1(m, K_S)/\mathrm{GL}(m, \mathcal{O}_S)$ as the space of lattices in K_S^m of covolume $(\sqrt{D_K})^m$, where D_K is the discriminant of K, and can define the family of subsets Q_ε of $\mathrm{GL}^1(m, K_S)/\mathrm{GL}(m, \mathcal{O}_S)$ in a way similar to (0.1), see §5.12 and §6.4 for details.

Let us now formulate a special case of our Theorem 6.4. Let $X = \prod_{v\in S} K_v^{d_v}$, where $d_v \in \mathbb{N}$, and let $h = (h_v)_{v\in S} : X \to \mathrm{GL}^1(m, K_S)$ be a polynomial map (see §6.4 for the definition) with $h(0) = e$. We will endow X with the product metric and the measure

$$\lambda \stackrel{\mathrm{def}}{=} \prod_{v\in S} \lambda_v\,, \text{ where } \lambda_v \text{ is a Haar measure on } K_v^{d_v}\,. \tag{0.2}$$

Theorem. *There exists $\alpha > 0$ (depending only on m, d_v and the degrees of the maps h_v) such that for every $\Lambda \in \mathrm{GL}^1(m, K_S)/\mathrm{GL}(m, \mathcal{O}_S)$ one can find $C > 0$ (depending only on m, d_v, the degrees of the maps and Λ) with the following property: for any positive ε and any ball $B \subset X$ centered at 0 one has*

$$\frac{\lambda\left(\{x \in B \mid h(x)\Lambda \notin Q_\varepsilon\}\right)}{\lambda(B)} \le C\varepsilon^\alpha\,.$$

0.3. Theorem 6.4 makes it possible to generalize the main results from [D1–3] to the case of products of real and p-adic algebraic groups. In particular, here is a special case of our Theorem 7.1:

Theorem. *Let $G = \prod_{v\in S} G_v$, where G_v is the group of K_v-rational points of an algebraic group $\mathbf{G}_v$ defined over K_v, let Γ be a lattice in G, d a natural number, and let X and λ be as in §0.2. Then for any $\beta > 0$ there exists a compact $M_\beta \subset G/\Gamma$ such* *that for any $y \in G/\Gamma$ and any parametrization $\phi : X \to U$ of degree $\le d$ of a unipotent algebraic subgroup U of G (see §7.1 for precise definitions) one of the following holds: either*

(i) *there exists $R = R(y, \phi)$ such that*

$$\frac{\lambda\left(\left\{x \in B \mid \phi(x)y \in M_\beta\right\}\right)}{\lambda(B)} \ge 1 - \beta$$

for any ball B in X centered at 0 with radius at least R, or

(ii) *there exists a closed proper subgroup H of G containing U such that the orbit Hy is closed and carries a finite H-invariant Borel measure.*

In the particular case when U is one-parameter, Theorem 0.3 has been announced in [MT1, Theorem 11.4], [Rat3, Theorem 9.1] and [T, Theorem 3.3], and used for the proof of Ratner's uniform distribution theorem [Rat3, Theorem 3] and for other related results.

0.4. The next theorem, which follows from Theorem 7.1, had been proved for real Lie groups by Dani [D1, D2], and for the general case had been announced in [MT1, Theorem 11.5].

Theorem. *Let G and Γ be as in Theorem 0.3, let H be a subgroup of G generated by unipotent one-parameter subgroups of G, and let μ be a locally finite H-invariant measure on G/Γ. Then there exist Borel H-invariant subsets Y_i, $1 \le i < \infty$, such that $\mu(Y_i) < \infty$ for all i and $\mu(G/\Gamma \smallsetminus \bigcup_i Y_i) = 0$. In particular, every locally finite H-invariant ergodic measure on G/Γ is finite.*

0.5. Another motivation for our work comes from applications to number theory. Indeed, using an S-adic generalization of the correspondence between homogeneous dynamics and Diophantine approximation, we establish S-arithmetic versions of several important results in metric Diophantine approximation over $\mathbb{R}$.

Let us recall several basic facts from the theory of simultaneous real Diophantine approximation. We will consider n-tuples $\mathbf{y} = (y_1, \dots, y_n)$ of real numbers viewed as *linear forms*, and will be interested in approximating their values $\mathbf{y}\cdot\mathbf{q} = y_1q_1 + \cdots + y_nq_n$ at integer points $\mathbf{q} = (q_1, \dots, q_n)$ by integers. One says that $\mathbf{y} \in \mathbb{R}^n$ is *very well approximable* (VWA) if for some $\varepsilon > 0$ there are infinitely many $\mathbf{q} \in \mathbb{Z}^n$ such that

$$|q_0 + \mathbf{q}\cdot\mathbf{y}| \le \|\mathbf{q}\|^{-n(1+\varepsilon)} \tag{0.3}$$

for some $q_0 \in \mathbb{Z}$, and that it is *very well multiplicatively approximable* (VWMA) if for some $\varepsilon > 0$ there are infinitely many $\mathbf{q} \in \mathbb{Z}^n$ such that

$$|q_0 + \mathbf{q}\cdot\mathbf{y}| \le \Pi_+(\mathbf{q})^{-(1+\varepsilon)}$$

for some $q_0 \in \mathbb{Z}$, where we put $\Pi_+(\mathbf{q}) \stackrel{\text{def}}{=} \prod_{i=1}^n \max(|q_i|, 1)$.

It is well known and easy to see that the set of VWMA vectors has Lebesgue measure zero and contains the set of VWA vectors. The subject of metric Diophantine approximation with dependent quantities originated with a conjecture of Mahler [M] stating that for almost every $x \in \mathbb{R}$, the point

$$\mathbf{f}(x) = (x, x^2, \dots, x^n) \tag{0.4}$$

is not very well approximable. More generally, following [KLW], we will say that a measure μ on $\mathbb{R}^n$ is *extremal* (resp. *strongly extremal*) if μ-a.e. point of $\mathbb{R}^n$ is not VWA (resp. not VWMA). In these terms, Mahler's conjecture, proved by Sprindžuk in the 1960s (see [Sp]), states that the pushforward $\mathbf{f}_*\lambda$ of Lebesgue measure λ on $\mathbb{R}$ by the map $\mathbf{f}$ as in (0.4) is extremal. On the other hand, the strong extremality of $\mathbf{f}_*\lambda$, conjectured by A. Baker in 1972 [B], has not been proved until the paper [KM], which introduced a dynamical approach to this class of problems.

It will be convenient to use the following terminology: if X is a metric space and μ a measure on X, a map $\mathbf{f} = (f_1, \dots, f_n)$ from X to F^n, where F is a locally compact field, will be called *μ-nonplanar at* $x_0 \in X$ if for any neighborhood B of x_0 the restrictions of $1, f_1, \dots, f_n$ to $B \cap \operatorname{supp} \mu$ are linearly independent over F; in other words, if $\mathbf{f}(B \cap \operatorname{supp} \mu)$ is not contained in any proper affine subspace of F^n. Note that in the special case when X is F^d, λ is Haar measure and $\mathbf{f}$ is polynomial, this condition is independent on x_0. We will omit the dependence on the measure when it is taken to be Lebesgue or Haar.

Here is a special case of [KM, Theorem A], settling Baker's conjecture:

Theorem. *Let* $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^n$ *be a nonplanar polynomial map. Then* $\mathbf{f}_*\lambda$ *is strongly extremal.*

0.6. Since the original work of Sprindžuk, one of the directions of research in metric Diophantine approximation has been studying local fields other than $\mathbb{R}$. In particular, Sprindžuk proved [Sp] that for almost all $x \in \mathbb{Q}_p$ and any $\varepsilon > 0$ there are at most finitely many $\tilde{\mathbf{q}} \stackrel{\text{def}}{=} (q_0, q_1, \dots, q_n) \in \mathbb{Z}^{n+1}$ such that

$$|q_0 + \mathbf{q} \cdot \mathbf{y}| \le \|\tilde{\mathbf{q}}\|^{-(n+1)(1+\varepsilon)}\,, \tag{0.5}$$

where $\mathbf{y}$ is of the form (0.4). We refer the reader to §8.2 of the present paper for the explanation of why the coefficients n in (0.3) and $n+1$ in (0.5) are optimal, and why $\|\mathbf{q}\|$ in the right hand side of (0.3) has to be replaced by $\|\tilde{\mathbf{q}}\|$. To be consistent, let us say that $\mathbf{y} \in \mathbb{Q}_p^n$ is VWA if for some $\varepsilon > 0$ there are infinitely many $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1}$ such that (0.5) holds, and that it is VWMA if for some $\varepsilon > 0$ there are infinitely many $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1}$ such that

$$|q_0 + \mathbf{q} \cdot \mathbf{y}| \le \Pi_+(\tilde{\mathbf{q}})^{-(1+\varepsilon)}\,.$$

As before, say that a measure μ on $\mathbb{Q}_p^n$ is *extremal* (resp., *strongly extremal*) if μ-almost every point of $\mathbb{Q}_p^n$ is not VWA (resp. not VWMA). Thus the aforementioned Sprindžuk's result is the extremality of $\mathbf{f}_*\lambda$, where $\mathbf{f} : \mathbb{Q}_p \to \mathbb{Q}_p^n$ is of the form (0.4) and λ is a Haar measure on $\mathbb{Q}_p$; that is, a p-adic analogue of the conjecture of Mahler. It seems natural to ask about a p-adic analogue of Baker's conjecture, and indeed it has been conjectured by I. Shparlinski [Bernik, private communication] that $\mathbf{f}_*\lambda$ as above is strongly extremal. In this paper we verify this conjecture, and, more generally, prove

Theorem. *Let* $\mathbf{f} : \mathbb{Q}_p^d \to \mathbb{Q}_p^n$ *be a nonplanar polynomial map. Then* $\mathbf{f}_*\lambda$ *is strongly extremal.*

0.7. In fact, the above theorem is a very special case of much more general S-arithmetic results. We postpone giving precise definitions until §8, where we review an S-adic generalization of real/p-adic metric Diophantine approximation, and in particular define extremality and strongly extremality of a measure on $\mathbb{Q}_S^n = \prod_{v\in S}\mathbb{Q}_v^n$. The following can be proved:

Theorem. *Let* S *be a finite set of normailzed valuations of* $\mathbb{Q}$*, let* $\mathbf{f}$ *be of the form* $(\mathbf{f}^{(v)})_{v\in S} : \prod_{v\in S}\mathbb{Q}_v^{d_v} \to \mathbb{Q}_S^n$*, where each* $\mathbf{f}^{(v)}$ *is a nonplanar polynomial map from* $\mathbb{Q}_v^{d_v}$ *to* $\mathbb{Q}_v^n$*, and let* λ *be defined as in* (0.2)*, with* $K = \mathbb{Q}$*. Then* $\mathbf{f}_*\lambda$ *is strongly extremal.*

Note that the paper [Z] considers the case when each $\mathbf{f}^{(v)}$ is of the form (0.4), and proves the extremality of $\mathbf{f}_*\lambda$.

Theorem 0.7 is a special case of Theorem 8.4, which is much more general but requires a certain terminology so we do not state it in the introduction. The advantage of having such a general result is that it is possible to exhibit many non-polynomial maps satisfying the assumptions of Theorem 8.4. This was done in [KM] for so-called smooth *nondegenerate* maps, and we plan to develop an S-adic version in the sequel [KT] to this paper.

Another advantage is that Theorem 8.4 generalizes the main result of [KLW], which, among other things, studies Diophantine properties of generic points on certain fractal subsets of $\mathbb{R}^n$. In fact, following [H], the latter class of sets can be defined in vector spaces over any locally compact field, and Theorem 8.4, combined with certain results of [KLW], implies that natural measures supported on those sets are strongly extremal.

0.8. The structure of the paper is as follows. In §1 and §2 we introduce and discuss the so-called *Besicovitch* metric spaces, *Federer* measures and (C,α)*-good* functions, that is, the language in which our main results are stated. Then in §3 and §4 we generalize a combinatorial estimate of [KM]. In §5 we prove auxiliary results from algebraic number theory important for the S-arithmetic applications. The quantitative S-arithmetic nondivergence is discussed in §6 where, in particular, Theorem 0.2 is proved. §7 is devoted to proving Theorems 0.3 and 0.4. Then in §8 we turn to the S-arithmetic Diophantine approximation, giving all the definitions, stating the most general strong extremality result (Theorem 8.4), and mentioning applications to fractal measures. The proof of Theorem 8.4 breaks into two special cases (when S does or does not contain the Archimedean valuation), which are treated in §9 and §10 respectively. In both cases the argument hinges on a modification of the dynamical approach to real Diophantine approximation as developed in [KM].

Acknowledgements: Part of the work was done during the authors' collaboration at the University of Lyon and at Brandeis University, and during the second-named author's stay at Max Planck Institute; the hospitality of these institutions is gratefully acknowledged. The authors are thankful to G.A. Margulis for his interest in this work and valuable comments. Thanks are also due to Vasily Bernik, Yuri Bilu, Elon Lindenstrauss and Barak Weiss for helpful discussions.

§1. Besicovitch covering property

1.1. For a metric space X, $x \in X$ and $r > 0$, we denote by $B(x,r)$ the open ball of radius r centered at x. For a subset B of X and a function $f : X \to F$, where $(F, |\cdot|)$ is a valued field, we let $\|f\|_B \stackrel{\text{def}}{=} \sup_{x\in B} |f(x)|$. If μ is a locally finite Borel measure on X and B is a subset of X with $\mu(B) > 0$, we define $\|f\|_{\mu,B}$ to be equal to $\|f\|_{B\,\cap\,\mathrm{supp}\,\mu}$, which, in case f is continuous and B is open, is the same as the $L^\infty(\mu)$ norm of $f|_B$, i.e.

$$\|f\|_{\mu,B} = \sup\left\{c \;\middle|\; \mu(\{x \in B : |f(x)| > c\}) > 0\right\}.$$

We will say that a metric space X *is Besicovitch* if there exists a constant N_X such that the following holds: for any bounded subset A of X and for any family $\mathcal{B}$ of nonempty open balls in X such that

$$\forall\, x \in A \text{ is a center of some ball of } \mathcal{B}, \tag{1.1}$$

there is a finite or countable subfamily $\{B_i\}$ of $\mathcal{B}$ with

$$1_A \le \sum_i 1_{B_i} \le N_X, \tag{1.2}$$

i.e. $A \subset \bigcup_i B_i$, and the multiplicity of that subcovering is at most N_X.

Example. Suppose that X is *ultrametric*, that is, the non-Archimedean triangle inequality $\mathrm{dist}(x_1,x_2) \le \max_{i=1,2} \mathrm{dist}(x,x_i)$ holds for all $x,x_1,x_2 \in X$. Then any two balls in X are either disjoint or contain one another (this observation will be repeatedly used throughout the sequel). This implies that any covering of any subset of X by balls has a subcovering of multiplicity 1; thus any separable ultrametric space is Besicovitch with $N_X = 1$ (the separability of X is equivalent to the collection of all its balls being countable).

1.2. Example. The fact that $\mathbb{R}^d$ is Besicovitch is the content of Besicovitch's Covering Theorem [Mat, Theorem 2.7] (the statement "$N_{\mathbb{R}} = 2$" is known as Vitali's Covering Theorem). In fact, Besicovitch's proof, see [Mat, pp. 29–34], can be easily generalized to give the following

1.3. Lemma. *For a metric space X, define*

$$M_X \stackrel{\text{def}}{=} \sup\left\{ k \;\middle|\; \begin{array}{l} \exists \textit{ balls } B_i = B(x_i,r_i),\ 1 \le i \le k, \\ \textit{such that } \bigcap_{i=1}^k B_i \ne \varnothing\,,\ \textit{ and } x_i \notin \bigcup_{\substack{j=1\\ j\ne i}}^k B_j \ \forall\, i \end{array} \right\}, \tag{1.3}$$

and also, for $c > 1$,

$$D_X(c) \stackrel{\text{def}}{=} \sup\left\{ k \;\middle|\; \begin{array}{l} \exists\, x \in X,\ r > 0 \textit{ and pairwise disjoint balls} \\ B_1,\dots,B_k \textit{ of radius } r \textit{ contained in } B(x,cr) \end{array} \right\}. \tag{1.4}$$

Then $N_X \le M_X D_X(8)$; hence X is Besicovitch if M_X and $D_X(8)$ are finite.

Proof. We follow the proof of Theorem 2.6 in [Mat]. Take a bounded $A \subset X$ and a family $\mathcal{B}$ of nonempty open balls in X satisfying (1.1). For each $x \in A$ pick one ball $B\big(x, r(x)\big) \in \mathcal{B}$. As A is bounded, we may assume that

$$R_1 = \sup_{x \in A} r(x) < \infty\,.$$

Choose $x_1 \in A$ with $r(x_1) \ge R_1/2$ and then inductively

$$x_{j+1} \in A \smallsetminus \bigcup_{i=1}^{j} B\big(x_i, r(x_i)\big) \quad \text{with } r(x_{j+1}) \ge R_1/2$$

as long as possible. Since A is bounded, the process terminates and we get a finite sequence $x_1, \dots, x_{k_1}$.

Next let

$$R_2 = \sup\left\{ r(x) \,\middle|\, x \in A \smallsetminus \bigcup_{i=1}^{k_1} B\big(x_i, r(x_i)\big) \right\}.$$

Choose $x_{k_1+1} \in A \smallsetminus \bigcup_{i=1}^{k_1} B\big(x_i, r(x_i)\big)$ with $r(x_{k_1+1}) \ge R_2/2$ and again inductively

$$x_{j+1} \in A \smallsetminus \bigcup_{i=1}^{j} B\big(x_i, r(x_i)\big) \quad \text{with } r(x_{j+1}) \ge R_2/2\,.$$

Continuing this process we obtain an increasing sequence of integers $0 = k_0 < k_1 < k_2 < \dots$, a decreasing sequence of positive numbers R_i with $2R_{i+1} \le R_i$, and a sequence of balls $B_i = B\big(x_i, r(x_i)\big)$ with the following properties. For $j \in \mathbb{N}$, let $I_j = \{k_{j-1}+1, \dots, k_j\}$. Then one has

$$r(x) < R_j/2 \qquad \text{for } j \in \mathbb{N},\ x \in A \smallsetminus \bigcup_{i=1}^{k_j} B_i\,, \tag{1.4a}$$

$$R_j/2 \le r(x_i) \le R_i \qquad \text{for } i \in I_j\,, \tag{1.4b}$$

$$x_{j+1} \in A \smallsetminus \bigcup_{i=1}^{j} B_i \qquad \text{for } j \in \mathbb{N}\,, \tag{1.4c}$$

$$x_k \in A \smallsetminus \bigcup_{l \ne j} \bigcup_{i \in I_l} B_i \qquad \text{for } j \in \mathbb{N},\ k \in I_j\,, \tag{1.4d}$$

$$A \subset \bigcup_{i=1}^{\infty} B_i\,. \tag{1.4e}$$

The first three properties follow immediately from the construction.

To prove (1.4d), take $j \in \mathbb{N}$, $k \in I_j$, $l \ne j$ and $i \in I_l$. If $l < j$, then $x_k \notin B_i$ by (1.4b). If $l > j$, then $x_i \notin B_k$ by (1.4b), and also $r(x_i) < r(x_k)$ by the construction, hence $x_k \notin B_i$. Finally, to verify (1.4e), observe that since $R_j \to 0$ as $j \to \infty$, (1.4a) forces $r(x)$ to be equal to zero for any $x \in A \setminus \bigcup_{i=1}^{\infty} B_i$.

Clearly (1.4e) proves the first inequality in (1.2). To establish the second inequality, assume that a point $x \in X$ belongs to k balls B_i, say,

$$x \in \bigcup_{i=1}^{k} B_{m_i} .$$

Using (1.4d) and (1.3), we see that the indices m_i can belong to at most M_X different blocks I_j. We now claim that for any $j \in \mathbb{N}$,

$$\#\big(I_j \cap \{m_i \mid i = 1, \ldots, k\}\big) \le D_X(8) .$$

Indeed, fix $j \in \mathbb{N}$ and write $\{n_1, \ldots, n_l\} \overset{\text{def}}{=} I_j \cap \{m_i \mid i = 1, \ldots, k\}$. By (1.4b) and (1.4c), the balls $B\big(x_{n_i}, R_j/4)\big)$, $i = 1, \ldots, l$, are disjoint and contained in $B(x, 2R_j))$, so the claim follows from (1.4). This proves that X is Besicovitch, with $N_X \le M_X D_X(8)$. $\square$

1.4. We will use the above lemma to prove that the products of $\mathbb{R}^d$ and certain ultrametric spaces are Besicovitch. Here and hereafter, the product of two metric spaces (X, dist_X) and (Y, dist_Y) will always be supplied with the product metric

$$\mathrm{dist}\,\big((x_1, y_1), (x_2, y_2)\big) = \max\big(\,\mathrm{dist}_X(x_1, x_2), \mathrm{dist}_Y(y_1, y_2)\big)\,,$$

so that balls in $X \times Y$ are products of balls in X and in Y. In particular, this convention forces the product of two ultrametric spaces to be ultrametric as well.

Lemma. *If Y is ultrametric, one has $M_{X\times Y} = M_X$.*

Proof. Assume $M_{X\times Y} > M_X$, and choose $k > M_X$ and balls $B_i = B(x_i, r_i)$, $1 \le i \le k$, in $X \times Y$ such that $\bigcap_{i=1}^{k} B_i \neq \varnothing$ and x_i is not in $\bigcup_{j=1,\, j\neq i}^{k} B_j$ for each i. Write $B_i = E_i \times F_i$, where E_i and F_i are projections of B_i onto X and Y respectively. Without loss of generality suppose that the sequence $\{r_i\}$ is non-increasing. Since Y is ultrametric and $\bigcap_{i=1}^{k} F_i \neq \varnothing$, one has $F_k \subset F_i$ for all i. On the other hand, since $M_X < k$, the center of E_k mush lie in E_i for some $i < k$, therefore x_k mush lie in B_i, a contradiction.

The converse inequality is straightforward (and not needed for our purposes). $\square$

1.5. Corollary. *If Y is ultrametric, one has*

$$N_{X\times Y} \le M_X D_X(8) D_Y(8)\,; \tag{1.5}$$

in particular, $X \times Y$ is Besicovitch if the three constants in the right hand side of (1.5) *are finite.*

Proof. It suffices to check that $D_{X\times Y}(8) \le D_X(8) D_Y(8)$, which is straightforward, as, since Y is ultrametric, any ball in Y of radius $8r$ is a disjoint union of at most $D_Y(8)$ balls of radius r. $\square$

1.6. Example. Let F be a locally compact field with a non-trivial ultrametric valuation, and let p be the number of elements in the residue class field of F, that is, the number of representatives in the closed unit ball (*the ring of integers of* F) $\mathcal{O} \stackrel{\text{def}}{=} B(0,1^+)$ modulo the open unit ball (*the valuation ideal of* F) $\mathcal{P} \stackrel{\text{def}}{=} B(0,1)$. Without loss of generality we can, and will from now on, *normalize* the valuation so that the diameter of $\mathcal{P}$ is equal to $1/p$. (If $F = \mathbb{Q}_p$, this way one gets $\mathcal{P} = \mathbb{Z}_p$, $\mathcal{P} = p\mathbb{Z}_p$, and $|\cdot| = |\cdot|_p$, the standard p-adic valuation.) Then it is easy to see that for any $c \geq 1$, any ball in F of radius cr is a disjoint union of at most $p^{[\log_p c]+1}$ balls of radius r. Therefore Corollary 1.5, in particular, implies that the metric space

$$X = \mathbb{R}^{d_0} \times F_1^{d_1} \times \cdots \times F_l^{d_\ell} \tag{1.6}$$

is Besicovitch for any ultrametric locally compact fields $F_1, \ldots, F_\ell$ and any $d_0, d_1, \ldots, d_\ell \in \mathbb{N}$.

1.7. We close the section with a measure-theoretic counterpart of the Besicovitch property. Namely, say that a locally finite Borel measure μ on X is *uniformly Federer* if there exists $D > 0$ such that

$$\sup_{r>0} \frac{\mu\big(B(x,3r)\big)}{\mu\big(B(x,r)\big)} < D \quad \text{for all } x \in \operatorname{supp}\mu\,. \tag{1.7}$$

Equivalently, one can replace "3" in (1.7) by any $c > 1$. In other words, μ is uniformly Federer if and only if for all $c > 1$ one has

$$D_\mu(c) \stackrel{\text{def}}{=} \sup_{\substack{x \in \operatorname{supp}\mu \\ r>0}} \frac{\mu\big(B(x,cr)\big)}{\mu\big(B(x,r)\big)} < \infty\,.$$

To simplify notation, we are going to write D_μ instead of $D_\mu(3)$. Note that if μ is a uniformly Federer measure on X with $\operatorname{supp}\mu = X$, for all $c > 1$ one automatically has $D_X(c) \leq D_\mu(2c)$, or even $\leq D_\mu(c)$ if X is ultrametric.

It will also be useful to have a non-uniform version of the above definition: following [KLW], we will say that μ as above is *Federer*[1] if for μ-a.e. $x \in X$ there exists a neighborhood U of x such that $\mu|_U$ is uniformly Federer.

Example. Let X be as in (1.6), and denote by p_i the number of elements in the residue class field of F_i, $i = 1, \ldots, \ell$. It is clear that any Haar measure λ on X is uniformly Federer, with $D_\lambda(c) \leq c^{d_0} \prod_{i=1}^{\ell} (cp_i)^{d_i}$.

§2. (C,α)-GOOD FUNCTIONS

2.1. Let X be a metric space, μ a Borel measure on X, and $(F, |\cdot|)$ a valued field. For a subset U of X and $C, \alpha > 0$, say that a function $f : X \to F$ is (C,α)-*good on* U *with respect to* μ if for any open ball $B \subset U$ centered in $\operatorname{supp}\mu$ one has

$$\forall\, \varepsilon > 0 \quad \mu\big(\{x \in B \mid |f(x)| < \varepsilon\}\big) \leq C \left(\frac{\varepsilon}{\|f\|_{\mu,B}}\right)^{\alpha} \mu(B)\,. \tag{2.1}$$

[1] See [S] and [KLW, §6] for a weaker non-uniform version.

Whenever it does not cause confusion, we will omit the reference to the measure μ, and simply say that the functions are (C, α)-good on U. In fact, in all the applications of our results, the metric space X will be the normed ring as in (1.6), and μ will be chosen to be a Haar measure λ on X. In particular, μ will be positive on open sets, so it will be always possible to replace $\|f\|_{\mu,B}$ in (2.1) by $\|f\|_B$ and not pay attention to the restriction of the center of B lying in $\operatorname{supp}\mu$. The above definition generalizes the one from [KM], which involved functions on $\mathbb{R}^d$, with μ being Lebesgue measure. See however [KLW] where measures supported on proper subsets of $\mathbb{R}^d$ are considered.

The following properties are immediate from the definition:

Lemma. *Let metric spaces X, Y, $U \subset X$, $V \subset Y$ and $C, \alpha > 0$ be given.*

(a) *f is (C, α)-good on U $\Leftrightarrow$ so is $|f|$;*
(b) *f is (C, α)-good on U $\Rightarrow$ so is cf $\forall\, c \in F$;*
(c) *f_i, $i \in I$, are (C, α)-good on U $\Rightarrow$ so is $f = \sup_{i\in I} |f_i|$;*
(d) *If f is (C, α)-good on U and $c_1 \le \frac{|f(x)|}{|g(x)|} \le c_2$ for all $x \in U$, then g is $\big(C(c_2/c_1)^\alpha, \alpha\big)$-good on U;*

2.2. The next lemma will be useful in dealing with functions on products of metric spaces:

Lemma. *Let metric spaces X, Y with measures μ, ν be given. Suppose f is a continuous function on $U \times V$, where $U \subset X$ and $V \subset Y$ are open subsets, and suppose C, D, α, β are positive constants such that*

$$\begin{aligned}&\textit{for all } y \in V \cap \operatorname{supp}\nu, \textit{ the function } x \mapsto f(x, y)\\ &\qquad \textit{is } (C, \alpha)\textit{-good on } U \textit{ with respect to } \mu\,,\end{aligned} \tag{2.2a}$$

and

$$\begin{aligned}&\textit{for all } x \in U \cap \operatorname{supp}\mu, \textit{ the function } y \mapsto f(x, y)\\ &\qquad \textit{is } (D, \beta)\textit{-good on } V \textit{ with respect to } \nu\,.\end{aligned} \tag{2.2b}$$

Then f is (E, γ)-good on $U \times V$ with respect to $\mu \times \nu$, where

$$\gamma = \frac{\alpha\beta}{\alpha + \beta} \quad \textit{and} \quad E = \left(\left(\frac{C}{\beta}\right)^\beta \left(\frac{D}{\alpha}\right)^\alpha \left(\alpha^{\alpha+\beta} + \beta^{\alpha+\beta}\right)\right)^{\frac{1}{\alpha+\beta}}. \tag{2.3}$$

Proof. Fix a ball in $U \times V$ of the form $A \times B$, where A and B are balls in X and Y, and denote $\|f\|_{\mu\times\nu, A\times B}$ by s. Also denote by φ the function defined on B by

$$\varphi(y) \overset{\text{def}}{=} \sup_{x\in A\,\cap\,\operatorname{supp}\mu} |f(x, y)|\,, \tag{2.4}$$

and for any $\eta > 0$ consider the set

$$B_\eta \overset{\text{def}}{=} \big\{y \in B \,\big|\, \varphi(y) \le \eta s\big\}\,.$$

It follows from Lemma 2.1(c) and (2.2b) that φ is (D,β)-good on V with respect to ν, therefore

$$\nu(B_\eta) \le D\left(\frac{\eta s}{\|\varphi\|_{\nu,B}}\right)^\beta \nu(B) = D\eta^\beta \nu(B)\,.$$

On the other hand, for $y \in (B \cap \operatorname{supp}\nu) \smallsetminus B_\eta$ one can write

$$\nu\big(\{x \in A \mid |f(x,y)| \le \varepsilon\}\big) \underset{(2.2a)\text{ and }(2.4)}{\le} C\left(\frac{\varepsilon}{\varphi(y)}\right)^\alpha \mu(A) \le C\left(\frac{\varepsilon}{\eta s}\right)^\alpha \mu(A)\,.$$

Therefore

$$\begin{aligned}(\mu\times\nu)\big(\{(x,y)\in A\times B \mid |f(x,y)| \le \varepsilon\}\big) \le\ & \mu(A)\nu(B_\eta) + C\left(\frac{\varepsilon}{\eta s}\right)^\alpha \mu(A)\nu(B)\\ \le\ & \left(C\left(\frac{\varepsilon}{\eta s}\right)^\alpha + D\eta^\beta\right)\mu(A)\nu(B)\,.\end{aligned}$$

An elementary computation shows that the expression in the right hand side is minimized by choosing $\eta = (C\alpha/D\beta)^{\frac{1}{\alpha+\beta}}(\varepsilon/s)^{\frac{\alpha}{\alpha+\beta}}$, and the minimum is equal to $E\,(\varepsilon/s)^\gamma\,\mu(A)\nu(B)$, with E and γ as in (2.3). $\square$

2.3. Corollary. *For $j = 1,\dots,\ell$, let X_j be a metric space, μ_j a measure on X_j, $U_j \subset X_j$ open, f_j a function which is (C_j,α_j)-good on U_j with respect to μ_j. Then $f : (x_1,\dots,x_\ell) \mapsto f_1(x_1)\cdot\ldots\cdot f_\ell(x_\ell)$ is (C,α)-good on $U_1\times\cdots\times U_\ell$ with respect to $\mu_1\times\cdots\times\mu_\ell$, with C and α explicitly computed in terms of C_j, α_j.*

2.4. The papers [KM] and [BKM] describe various classes of real valued functions on open subsets of $\mathbb{R}^d$ which are (C,α)-good with respect to Lebesgue measure. For example, the fact that polynomials in one real variable of degree $\le k$ have that property (with $\alpha = 1/k$) follows easily from Lagrange's interpolation, see [DM, Lemma 4.1] and [KM, Lemma 3.2]. Similarly, following [T], one can consider polynomials over ultrametric fields:

Lemma. *Let F be a locally compact ultrametric valued field. Then for any $d,k\in\mathbb{N}$, any polynomial $f\in F[x_1,\dots,x_d]$ of degree not greater than k is $(C_{d,k},1/dk)$-good on F^d with respect to λ, where $C_{d,k}$ is a constant depending only on d and k.*

Proof. The case $k=1$ is proved in [T, Lemma 4.1], and the general case can be easily derived from the case $k=1$ and Lemma 2.2 by induction on d. $\square$

2.5. We will need the following terminology: if $(F,|\cdot|)$ is a valued field, X a metric space, μ a measure on X and $\mathbf{f}$ a map from X to F^n, we will say that $\mathbf{f}$ is *μ-good at* $x_0\in X$ (cf. [K2]) if there exists a neighborhood V of x_0 and positive C,α such that any linear combination of $1, f_1,\dots,f_n$ is (C,α)-good on V with respect to μ. Again, the reference to the measure will be omitted when $\mu=\lambda$. For example, it follows from the aforementioned results that polynomial maps are good at every point with respect to Lebesgue measure.

It has also been shown in [KM] that the (C,α)-good property can be established for differentiable $\mathbb{R}$-valued functions on subsets of $\mathbb{R}^d$ with non-vanishing derivatives of some order. In particular, it follows from the results of [KM, §3] that for any function f which is real analytic near $\mathbf{x}_0 \in \mathbb{R}^d$ there exist C, α and a neighborhood U of $\mathbf{x}_0$ such that f is (C,α)-good on U with respect to Lebesgue measure; consequently, any real analytic map $\mathbb{R}^d \to \mathbb{R}^n$ is good.

It is possible to prove analogues of the aforementioned results for certain appropriately defined classes of differentiable functions (and in particular for analytic ones) on F^d, where F is a locally compact ultrametric field. This produces numerous non-Archimedean examples of (C,α)-good functions and maps. More detail will appear in [KT].

§3. Maps of posets into spaces of good functions

3.1. The goal of this section is to generalize a construction described in [KM] in order to make it work for functions defined on arbitrary metric spaces[2]. More precisely, we will work with mappings of partially ordered sets (*posets*) into spaces of functions on a metric space X with a measure μ. Given such a mapping, we will *mark* certain points (see the definition below), and prove an upper estimate (Theorem 3.1) for the measure of the set of "unmarked" points.

For a poset $\mathfrak{P}$, we will denote by $l(\mathfrak{P})$ the *length* of $\mathfrak{P}$, i.e. the number of elements in a maximal linearly ordered subset of $\mathfrak{P}$. If $\mathfrak{S}$ is a subset of $\mathfrak{P}$, we let $\mathfrak{P}(\mathfrak{S})$ be the poset of elements of $\mathfrak{P} \smallsetminus \mathfrak{S}$ comparable with any element of $\mathfrak{S}$. Note that one always has

$$l\big(\mathfrak{P}(\mathfrak{S})\big) \le l(\mathfrak{P}) - l(\mathfrak{S})\,. \tag{3.1}$$

We will fix a metric space X, and consider posets $\mathfrak{P}$ together with a mapping ψ from $\mathfrak{P}$ to the space $C(B)$ of $\mathbb{R}$-valued continuous functions on some subset B of X, to be denoted by $s \mapsto \psi_s$. Given such a mapping and positive numbers $\varepsilon \le \rho$, we will say that a point $z \in B$ is (ε,ρ)-*marked relative to* $\mathfrak{P}$ if there exists a linearly ordered subset $\mathfrak{S}_z$ of $\mathfrak{P}$ such that

(M1) $\varepsilon \le |\psi_s(z)| \le \rho \quad \forall\, s \in \mathfrak{S}_z$;
(M2) $|\psi_s(z)| \ge \rho \quad \forall\, s \in \mathfrak{P}(\mathfrak{S}_z)$.

We will denote the set of all such points by $\Phi(\varepsilon,\rho,\mathfrak{P})$. When it does not cause confusion, we will omit the reference to either $\mathfrak{P}$ or (ε,ρ), and will simply say that z is (ε,ρ)-marked, or marked relative to $\mathfrak{P}$.

Theorem. *Let X be a Besicovitch metric space, μ a uniformly Federer measure on X, $m \in \mathbb{Z}_+$ and $C,\alpha,\rho > 0$. Suppose that we are given a poset $\mathfrak{P}$, a ball $B = B(x,r)$ in X, and a mapping $\psi : \mathfrak{P} \to C(\tilde{B})$, where $\tilde{B} \stackrel{\text{def}}{=} B\big(x, 3^m r\big)$, such that the following holds:*

(A0) *$l(\mathfrak{P}) \le m$;*
(A1) *$\forall\, s \in \mathfrak{P}, \quad \psi_s$ is (C,α)-good on $\tilde{B}$ with respect to μ;*

[2] A possibility of such a generalization is mentioned in [KM, §6.1]. The paper [KLW] contains a slightly different presentation of the same argument, written in the special case of $\mathfrak{P}$ being the poset of nonzero rational subspaces of $\mathbb{R}^m$.

(A2) $\forall\, s \in \mathfrak{P}, \quad \|\psi_s\|_{\mu,B} \ge \rho$;
(A3) $\forall\, y \in \tilde{B} \cap \operatorname{supp}\mu, \quad \#\{s \in \mathfrak{P} \mid |\psi_s(y)| < \rho\} < \infty$.

Then $\forall\, \varepsilon \le \rho$ one has

$$\mu\big(B \smallsetminus \Phi(\varepsilon,\rho,\mathfrak{P})\big) \le mC\big(N_X D_\mu^2\big)^m \left(\frac{\varepsilon}{\rho}\right)^{\alpha} \mu(B)\,.$$

Proof. We proceed by induction on m. If $m = 0$, the poset $\mathfrak{P}$ is empty, and for any $z \in B$ one can take $\mathfrak{S}_z = \varnothing$ and check that (M1) and (M2) are satisfied for all ε, ρ; thus all points of B are marked. Now take $m \ge 1$ and suppose that the claim is proved for all smaller values of m.

Fix $C, \alpha, \rho, b, \mathfrak{P}, B = B(x,r)$ and ψ as in the formulation of the theorem. For any $y \in B \cap \operatorname{supp}\mu$ define

$$H(y) \stackrel{\text{def}}{=} \{s \in \mathfrak{P} \mid |\psi_s(y)| < \rho\}\,;$$

this is a finite subset of $\mathfrak{P}$ in view of (A3). If $H(y)$ is empty, y is clearly (ε,ρ)-marked for any positive ε: indeed, since $|\psi_s(y)| \ge \rho$ for all $s \in \mathfrak{P}$, one can again take $\mathfrak{S}_y$ to be the empty set and check that (M1) and (M2) are satisfied. Thus one only needs to consider points y from the set

$$E \stackrel{\text{def}}{=} \{y \in B \cap \operatorname{supp}\mu \mid H(y) \ne \varnothing\} = \{y \in B \cap \operatorname{supp}\mu \mid \exists\, s \in \mathfrak{P} \text{ with } |\psi_s(y)| < \rho\}\,.$$

Take $y \in E$ and $s \in H(y)$, and define

$$r_{s,y} \stackrel{\text{def}}{=} \sup\{t > 0 \mid \|\psi_s\|_{\mu,B(y,t)} < \rho\}\,. \tag{3.2}$$

It follows from the continuity of functions ψ_s that for small enough positive t one has $\|\psi_s\|_{\mu,B(y,t)} < \rho$, hence $r_{s,y} > 0$. Denote $B(y, r_{s,y})$ by $B_{s,y}$. From (A1) it is clear that $B_{s,y}$ does not contain B; therefore one has $r_{s,y} < 2r$. Note also that (3.2) immediately implies that

$$\|\psi_s\|_{\mu,B_{s,y}} \le \rho\,. \tag{3.3}$$

Now for any $y \in E$ choose an element s_y of $H(y)$ such that $r_{s_y,y} \ge r_{s,y}$ for all $s \in H(y)$ (this can be done since $H(y)$ is finite). For brevity let us denote $r_{s_y,y}$ by r_y and $B_{s_y,y} = \bigcup_{s\in H(y)} B_{s,y}$ by B_y. Also let us denote the poset $\mathfrak{P}(\{s_y\})$ by $\mathfrak{P}_y$.

3.1.1. The next lemma allows one to show a point $z \in B_y$ to be marked relative to $\mathfrak{P}$ once it is marked relative to $\mathfrak{P}_y$. It is proved by a verbatim repetition of the proof of [KM, Lemma 4.6], yet we do it in full detail here to make the argument self-contained.

Lemma. *For $\varepsilon \le \rho$ and $y \in E$, let $z \in B_y \cap \operatorname{supp}\mu \cap \Phi(\varepsilon,\rho,\mathfrak{P}_y)$ be such that $|\psi_{s_y}(z)| \ge \varepsilon$; then z belongs to $\Phi(\varepsilon,\rho,\mathfrak{P})$. Equivalently,*

$$(B_y \cap \operatorname{supp}\mu) \smallsetminus \Phi(\varepsilon,\rho,\mathfrak{P}) \subset \big(B_y \smallsetminus \Phi(\varepsilon,\rho,\mathfrak{P}_y)\big) \cup \{z \in B_y \mid |\psi_{s_y}(z)| < \varepsilon\}\,. \tag{3.4}$$

Proof. By definition of $\Phi(\varepsilon,\rho,\mathfrak{P}_y)$, there exists a linearly ordered subset $\mathfrak{S}_{y,z}$ of $\mathfrak{P}_y$ such that

$$\varepsilon \le |\psi_s(z)| \le \rho \quad \forall\, s \in \mathfrak{S}_{y,z} \tag{3.5}$$

and

$$|\psi_s(z)| \ge \rho \quad \forall\, s \in \mathfrak{P}_y(\mathfrak{S}_{y,z})\,. \tag{3.6}$$

Put $\mathfrak{S}_z \overset{\text{def}}{=} \mathfrak{S}_{y,z} \cup \{s_y\}$. Then $\mathfrak{P}(\mathfrak{S}_z) = \mathfrak{P}_y(\mathfrak{S}_{y,z})$; therefore (M2) immediately follows from (3.6), and, in view of (3.5), it remains to check (M1) for $s = s_y$. The latter is straightforward: $|\psi_{s_y}(z)|$ is not less than ε by the assumption and is not greater than ρ in view of (3.3). $\square$

3.1.2. Note that one clearly has $r_y < 2r$, which in particular implies that $B_y \subset B(x,3r)$. We are going to fix some r'_y strictly between r_y and $\min(2r,3r_y)$, and denote $B(y,r'_y)$ by B'_y. Clearly one has

$$\|\psi_s\|_{\mu,B'_y} \ge \rho \quad \text{for any } y \in E \text{ and } s \in \mathfrak{P}\,. \tag{3.7}$$

(Indeed, the definition of r_y and (3.2) imply the above inequality for any $s \in H(y)$, and it obviously holds if $s \notin H(y)$.)

Now observe that $\mathfrak{P}_y$, B'_y and $\tilde B'_y \overset{\text{def}}{=} B(y,3^{m-1}r'_y)$ satisfy properties

- (A0) with m replaced by $m-1$ — in view of (3.1);
- (A2) — in view of (3.7);
- (A1) and (A3) — since

$$\tilde B'_y = B(y,3^{m-1}r'_y) \subset B(x,3^{m-1}r'_y + r) \subset B\big(x,(2\cdot 3^{m-1}+1)r\big) \subset B(x,3^m r) = \tilde B\,.$$

Therefore one has

$$\begin{aligned}\mu\big(B_y \smallsetminus \Phi(\varepsilon,\rho,\mathfrak{P}_y)\big) &\le \mu\big(B'_y \smallsetminus \Phi(\varepsilon,\rho,\mathfrak{P}_y)\big)\\ &\le (m-1)C\big(N_X D_\mu^2\big)^{m-1}\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\mu(B'_y)\\ &\le D_\mu(m-1)C\big(N_X D_\mu^2\big)^{m-1}\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\mu(B_y)\end{aligned} \tag{3.8}$$

by the induction assumption and the Federer property of μ. On the other hand, in view of ψ_{s_y} being (C,α)-good on $\tilde B \supset B'_y$, one can write

$$\begin{aligned}\mu\left(\{z \in B_y \mid |\psi_{s_y}(z)| < \varepsilon\}\right) &\le \mu\left(\{z \in B'_y \mid |\psi_{s_y}(z)| < \varepsilon\}\right)\\ &\le C\left(\frac{\varepsilon}{\|\psi_{s_y}\|_{\mu,B'_y}}\right)^{\alpha}\mu(B'_y) \underset{(3.7)}{\le} C\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\mu(B'_y)\\ &\le CD_\mu\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\mu(B_y)\,.\end{aligned} \tag{3.9}$$

Recall that we need to estimate the measure of $E \smallsetminus \Phi(\varepsilon, \rho, \mathfrak{P})$. For any $y \in E$, in view of (3.4), (3.8) and (3.9) one has

$$
\begin{aligned}
\mu\big(B_y \smallsetminus \Phi(\varepsilon, \rho, \mathfrak{P})\big) &\le C\left((m-1)N_X^{m-1}D_\mu^{2m-1} + D_\mu\right)\left(\frac{\varepsilon}{\rho}\right)^\alpha \mu(B_y) \\
&\le mCN_X^{m-1}D_\mu^{2m-1}\left(\frac{\varepsilon}{\rho}\right)^\alpha \mu(B_y)\,.
\end{aligned} \tag{3.10}
$$

Now consider the covering $\{B_y \mid y \in E\}$ of E, choose a countable subset Y of E such that the multiplicity of the subcovering $\{B_y \mid y \in Y\}$ is at most N_X, and write

$$
\sum_{y\in Y} \mu(B_y) \le N_X \mu\Big(\bigcup_{y \in Y} B_y\Big) \le N_X \mu\big(B(x, 3r)\big) \le N_X D_\mu \mu(B)\,. \tag{3.11}
$$

Therefore the measure of $E \smallsetminus \Phi(\varepsilon, \rho, \mathfrak{P})$ is bounded from above by

$$
\begin{aligned}
\sum_{y\in Y} \mu\big(B_y \smallsetminus \Phi(\varepsilon, \rho, \mathfrak{P})\big) &\underset{(3.10)}{\le} mCN_X^{m-1}D_\mu^{2m-1}\left(\frac{\varepsilon}{\rho}\right)^\alpha \sum_{y\in Y}\mu(B_y) \\
&\underset{(3.11)}{\le} mC\big(N_X D_\mu^2\big)^m \left(\frac{\varepsilon}{\rho}\right)^\alpha \mu(B)\,. \qquad \square
\end{aligned}
$$

4. Primitive submodules of $\mathcal{D}^m$

4.1. We start this section by assuming that

- $\mathcal{D}$ is an integral domain, that is, a commutative ring with 1 and without zero divisors;
- K is the quotient field of $\mathcal{D}$;
- $\mathcal{R}$ is a commutative ring containing K as a subring.

We need the following elementary lemma:

Lemma. *Let $k, m \in \mathbb{N}$, $k \le m$, and let $\gamma_1, \dots, \gamma_k \in K^m$ be linearly independent over K. Then they are linearly independent over $\mathcal{R}$.*

Proof. Let A be the $m \times k$-matrix with columns given by $\gamma_1, \dots, \gamma_k$. Then there exists at least one $k \times k$ minor B of A with $\det(B)$ being a nonzero element of K, hence invertible in $\mathcal{R}$. By Cramer's rule, for any solution $\beta = (\beta_1, \dots, \beta_k) \in \mathcal{R}^k$ of $A\beta = 0$ one must have $\det(B)\beta_i = 0$ for every i, hence $\beta = 0$. $\square$

4.2. If Δ is an $\mathcal{D}$-submodule of $\mathcal{R}^m$, let us denote by $K\Delta$ (resp. $\mathcal{R}\Delta$) its K- (resp. $\mathcal{R}$-) linear span inside $\mathcal{R}^m$, and define the *rank* $\operatorname{rk}\Delta$ of Δ by

$$
\operatorname{rk}(\Delta) \overset{\text{def}}{=} \dim_K(K\Delta)\,. \tag{4.1}
$$

For example, one has $\operatorname{rk}(\mathcal{D}^m) = m$ for any $m \in \mathbb{N}$. If Λ is an $\mathcal{D}$-submodule of $\mathcal{R}^m$ and Δ is a submodule of Λ, say that Δ is *primitive in* Λ if any submodule of Λ of rank equal to $\operatorname{rk}(\Delta)$ and containing Δ is equal to Δ. It is clear that the set of nonzero primitive submodules of a fixed $\mathcal{D}$-submodule Λ of $\mathcal{R}^m$ is a partially ordered set (with respect to inclusion) of length equal to $\operatorname{rk}(\Lambda)$.

The next lemma characterizes primitive submodules of $\mathcal{D}^m$:

Lemma. *The following are equivalent for a submodule Δ of $\mathcal{D}^m$:*
(i) Δ *is primitive;*
(ii) $\Delta = K\Delta \cap \mathcal{D}^m$*;*
(iii) $\Delta = \mathcal{R}\Delta \cap \mathcal{D}^m$ *for any commutative ring $\mathcal{R}$ containing K as a subring.*

Proof. If $\Delta = \{0\}$, the claim is trivial. Otherwise, it is obvious that (iii)⇒(ii)⇒(i). Assuming (i) and taking $\gamma \in \mathcal{R}\Delta \cap \mathcal{D}^m$, let $\gamma_1, \dots, \gamma_k \in \Delta$ be a basis of $K\Delta$, with $k = \mathrm{rk}(\Delta)$. Then $\gamma, \gamma_1, \dots, \gamma_k$ are linearly dependent over $\mathcal{R}$, hence, in view of Lemma 4.1, over K. But $\gamma_1, \dots, \gamma_k$ are linearly independent over K, thus γ belongs to $K\Delta$, therefore the $\mathcal{D}$-module Δ' generated by Δ and γ has rank k. By primitivity of Δ, $\Delta' = \Delta$, i.e. $\gamma \in \Delta$. □

In fact, Lemma 4.2 implies that for any $\Delta' \subset \mathcal{D}^m$ there exists a unique primitive $\Delta \supset \Delta'$ of the same rank, namely, $\Delta = K\Delta' \cap \mathcal{D}^m$.

4.3. Remark. Note that if $\mathcal{D}$ is not a principal ideal domain, then there exist primitive submodules which are not free. For example, let $\Delta' = \mathcal{D}\mathbf{a}$, where $\mathbf{a} = (a_1, ..., a_m) \in \mathcal{D}^m$. Then the primitive submodule $\Delta = K\Delta' \cap \mathcal{D}^m$ is free if and only if the ideal I in $\mathcal{D}$ generated by $a_1, ..., a_m$ is principal. Indeed, take $x \in K$. Then $x\mathbf{a} \in \Delta \iff xa_i \in \mathcal{D}$ for all $i \iff x \in I^{-1}$. Therefore $\Delta = I^{-1}\mathbf{a}$. But $I^{-1} \cong \Delta$ as $\mathcal{D}$-modules. Hence Δ is free $\iff I^{-1}$ is principal $\iff I$ is principal, which proves the claim.

Let us now assume in addition that $\mathcal{R}$ is a topological ring, and consider the topological group $\mathrm{GL}(m, \mathcal{R})$ of automorphisms of $\mathcal{R}^m$, i.e. the group of $m \times m$ invertible matrices with entries in $\mathcal{R}$. Any $g \in \mathrm{GL}(m, \mathcal{R})$ maps $\mathcal{D}$-submodules of $\mathcal{R}^m$ to $\mathcal{D}$-submodules of $\mathcal{R}^m$, preserving their rank and the inclusion relation. Let us introduce the following notation

$$\mathfrak{M}(\mathcal{R}, \mathcal{D}, m) \stackrel{\text{def}}{=} \{g\Delta \mid g \in \mathrm{GL}(m, \mathcal{R}),\ \Delta \text{ is a submodule of } \mathcal{D}^m\}, \tag{4.2}$$

and

$$\mathfrak{P}(\mathcal{D}, m) \stackrel{\text{def}}{=} \text{ the set of all nonzero primitive submodules of } \mathcal{D}^m.$$

Note that the inclusion relation makes $\mathfrak{P}(\mathcal{D}, m)$ a poset of length m.

4.4. We would like to have a way to measure "size" of submodules from the above collection. Specifically, let us say that a function $\nu : \mathfrak{M}(\mathcal{R}, \mathcal{D}, m) \mapsto \mathbb{R}_+$ is *norm-like* if the following three properties hold:

(N1) For any $\Delta, \Delta' \in \mathfrak{M}(\mathcal{R}, \mathcal{D}, m)$ with $\Delta' \subset \Delta$ and $\mathrm{rk}(\Delta') = \mathrm{rk}(\Delta)$ one has $\nu(\Delta') \geq \nu(\Delta)$;
(N2) there exists $C_\nu > 0$ such that for any $\Delta \in \mathfrak{M}(\mathcal{R}, \mathcal{D}, m)$ and any $\gamma \notin \mathcal{R}\Delta$ one has $\nu(\Delta + \mathcal{D}\gamma) \leq C_\nu \cdot \nu(\Delta)\nu(\mathcal{D}\gamma)$;
(N3) for every submodule Δ of $\mathcal{D}^m$, the function $\mathrm{GL}(m, \mathcal{R}) \to \mathbb{R}_+$, $g \mapsto \nu(g\Delta)$, is continuous.

If ν is as above and $\gamma \in \mathcal{R}^m$, we will define $\nu(\gamma)$ to be equal to $\nu(\mathcal{D}\gamma)$. The model example is given by taking $\mathcal{D} = \mathbb{Z}$, $K = \mathbb{Q}$, $\mathcal{R} = \mathbb{R}$. Then the set $\mathfrak{M}(\mathcal{R}, \mathcal{D}, m)$ coincides with the set of all discrete subgroups of $\mathbb{R}^n$, and one can define $\nu(\Delta)$ to be the covolume of Δ in $\mathbb{R}\Delta$, with $\nu(\mathbf{v})$ being equal to the Euclidean norm of a vector

$\mathbf{v} \in \mathbb{R}^m$; in that case one can easily check that (N1)–(N3) are satisfied, with $C_\nu = 1$. In the next section we will do this in a more general context, when $\mathcal{R}$ is not a field anymore.

Now we can apply Theorem 3.1 to the poset $\mathfrak{P}(\mathcal{D}, m)$.

Theorem. *Let X be a Besicovitch metric space, μ a uniformly Federer measure on X, and let $\mathcal{D} \subset K \subset \mathcal{R}$ be as above, $\mathcal{R}$ being a topological ring. For $m \in \mathbb{N}$, let a ball $B = B(x_0, r_0) \subset X$ and a continuous map $h : \tilde{B} \to \mathrm{GL}(m, \mathcal{R})$ be given, where $\tilde{B}$ stands for $B(x_0, 3^m r_0)$. Also let ν be a norm-like function on $\mathfrak{M}(\mathcal{R}, \mathcal{D}, m)$. For any $\Delta \in \mathfrak{P}(\mathcal{D}, m)$ denote by ψ_Δ the function $x \mapsto \nu\big(h(x)\Delta\big)$ on $\tilde{B}$. Now suppose for some $C, \alpha > 0$ and $0 < \rho < 1/C_\nu$ one has*

(i) *for every $\Delta \in \mathfrak{P}(\mathcal{D}, m)$, the function ψ_Δ is (C, α)-good on $\tilde{B}$ with respect to μ;*
(ii) *for every $\Delta \in \mathfrak{P}(\mathcal{D}, m)$, $\|\psi_\Delta\|_{\mu, B} \geq \rho$;*
(iii) $\forall\, x \in \tilde{B} \cap \operatorname{supp} \mu, \quad \#\big\{\Delta \in \mathfrak{P}(\mathcal{D}, m) \bigm| \psi_\Delta(x) < \rho\big\} < \infty.$

Then for any positive $\varepsilon \leq \rho$ one has

$$\mu\left(\left\{x \in B \,\middle|\, \begin{array}{l} \nu\big(h(x)\gamma\big) < \varepsilon \textit{ for} \\ \textit{some } \gamma \in \mathcal{D}^m \smallsetminus \{0\} \end{array}\right\}\right) \leq mC\big(N_X D_\mu^2\big)^m \left(\frac{\varepsilon}{\rho}\right)^\alpha \mu(B)\,. \tag{4.3}$$

Proof. For simplicity let us denote $\mathfrak{P}(\mathcal{D}, m)$ by $\mathfrak{P}$. As was observed above, the length of $\mathfrak{P}$ is equal to m, and one immediately verifies that conditions (i)–(iii) imply that $\mathfrak{P}$, B and $\tilde{B}$ satisfy properties (A1)–(A3) of Theorem 3.1. Thus it suffices to prove that for any positive $\varepsilon \leq \rho$ one has

$$\Phi(\varepsilon, \rho, \mathfrak{P}) \subset \big\{x \in B \mid \nu\big(h(x)\gamma\big) \geq \varepsilon \text{ for all } \gamma \in \mathcal{D}^m \smallsetminus \{0\}\big\}\,. \tag{4.4}$$

Take an (ε, ρ)-marked point $x \in B$, and let $\{0\} = \Delta_0 \subsetneq \Delta_1 \subsetneq \cdots \subsetneq \Delta_l = \mathcal{D}^m$ be all the elements of $\mathfrak{S}_x \cup \big\{\{0\}, \mathcal{D}^m\big\}$. Pick any $\gamma \in \mathcal{D}^m \smallsetminus \{0\}$. Then there exists i, $1 \leq i \leq l$, such that $\gamma \in \Delta_i \smallsetminus \Delta_{i-1}$. From the primitivity of Δ_{i-1} and Lemma 4.2 it follows that $\gamma \notin \mathcal{R}\Delta_{i-1}$, hence $g\gamma \notin g\mathcal{R}\Delta_{i-1} = \mathcal{R}g\Delta_{i-1}$ for any $g \in \mathrm{GL}(m, \mathcal{R})$. Therefore, if one defines $\Delta' \stackrel{\text{def}}{=} \mathcal{D}\Delta_{i-1} + \mathcal{D}\gamma$, in view of (N2) one has

$$\nu\big(h(x)\Delta'\big) \leq C_\nu \nu\big(h(x)\Delta_{i-1}\big)\nu\big(h(x)\gamma\big)\,. \tag{4.5}$$

Further, let $\Delta \stackrel{\text{def}}{=} K\Delta' \cap \mathcal{D}^m$. It is a primitive submodule containing Δ' and of rank equal to $\mathrm{rk}(\Delta')$, so, by (N1),

$$\nu\big(h(x)\Delta\big) \leq \nu\big(h(x)\Delta'\big)\,. \tag{4.6}$$

Moreover, it is also contained in Δ_i, since

$$\Delta = K\Delta' \cap \mathcal{D}^m = K\Delta \cap \mathcal{D}^m \subset K\Delta_i \cap \mathcal{D}^m = \Delta_i\,.$$

Therefore it is comparable to any element of $\mathfrak{S}_x$, i.e. belongs to $\mathfrak{S}_x \cup \mathfrak{P}(\mathfrak{S}_x)$. Then one can use properties (M1) and (M2) to deduce that

$$|\psi_\Delta(x)| = \nu\big(h(x)\Delta\big) \geq \min(\varepsilon, \rho) = \varepsilon\,,$$

and then, in view of (4.5) and (4.6), conclude that

$$\nu\big(h(x)\gamma\big) \geq \nu\big(h(x)\Delta\big)/C_\nu \nu\big(h(x)\Delta_{i-1}\big) \geq \varepsilon/C_\nu \rho \geq \varepsilon\,.$$

This shows (4.4) and completes the proof of the theorem. $\square$

5. Discrete submodules of K_S^m

5.1. The goal of this section is to describe a certain class of triples $\mathcal{D} \subset K \subset \mathcal{R}$ and construct a norm-like function on $\mathfrak{M}(\mathcal{R}, \mathcal{D}, m)$ which is important in applications to both dynamics and Diophantine approximation.

We take K to be a *number field*, that is, a finite extension of $\mathbb{Q}$, and let S be a finite set of places of K containing all the Archimedean ones. We recall that every completion K_v of K is isomorphic to a finite extension of either $\mathbb{R}$ or $\mathbb{Q}_p$. We wil denote by $|\cdot|_v$ the *normalized* valuation on K_v. In other words, $|\cdot|_v$ is the usual absolute value if v is real, the square of the absolute value if v is complex, and is defined as in Example 1.6 if v is ultrametric.

We let K_S be the direct product of all the completions K_v, $v \in S$, in which K is diagonally imbedded. Let $\infty \subset S$ stand for the subset of all Archimedean valuations, and let us denote $S \smallsetminus \infty$ by S_f. Also let

$$K_\infty \stackrel{\text{def}}{=} \prod_{v \in \infty} K_v, \quad \mathcal{O}_f \stackrel{\text{def}}{=} \prod_{v \in S_f} \mathcal{O}_v \quad \text{and} \quad K_f \stackrel{\text{def}}{=} K_\infty \times \mathcal{O}_f\,,$$

where for every $v \in S$ we denote by $\mathcal{O}_v$ the ring of integers of K_v. We also let $\mathcal{O}$ stand for the ring of integers of K, and denote by

$$\mathcal{O}_S \stackrel{\text{def}}{=} \big\{x \in K_S \bigm| |x^{(v)}|_v \le 1 \text{ for all } v \notin S\big\}$$

the ring of S-integers of K.

Denote by λ_v the *normalized* (see [W, I.V.4]) Haar measure on K_v, and by $\lambda_S = \prod_{v \in S} \lambda_v$ the product measure on K_S. Elements of K_S will be denoted as $x = (x^{(v)})_{v \in S}$ or simply $x = (x^{(v)})$, where $x^{(v)} \in K_v$. For x of this form, we define the *content* $c(x)$ of x to be the product of all $|x^{(v)}|_v$, $v \in S$. Since all the valuations are normalized, one has

$$\lambda_S(xM) = c(x)\lambda_S(M)\,. \tag{5.1}$$

where $x \in K_S$ and M is a measurable subset of K_S.

If m is a natural number, we preserve the same notations λ_v and λ_S to denote the product measure on K_v^m and K_S^m, respectively. Elements $\mathbf{x} = (x_1, \dots, x_m)$ of K_S^m will be denoted as $\mathbf{x} = (\mathbf{x}^{(v)})$, where $\mathbf{x}^{(v)} = (x_1^{(v)}, \dots, x_m^{(v)}) \in K_v^m$. We denote by $\|\cdot\|_v$ the square of the usual (Hermitian) norm on K_v^m if v is complex, the usual (Euclidean) norm on K_v^m if v is real, and the l^∞ norm defined by

$$\|(x_1^{(v)}, \dots, x_m^{(v)})\|_v = \max_i |x_i^{(v)}|_v$$

if v is non-Archimedean. For $\mathbf{x} = (\mathbf{x}^{(v)})$ in K_S^m we define the *content* $c(\mathbf{x})$ of $\mathbf{x}$ as the product of all the numbers $\|(\mathbf{x}^{(v)})\|_v$, $v \in S$. The group $\mathrm{GL}(m, K_S) = \prod_{v \in S} \mathrm{GL}(m, K_v)$ acts naturally on K_S^m, and one has

$$\lambda_S(gM) = c\big(\det(g)\big)\lambda_S(M),$$

where $M \subset K_S^m$ is any measurable subset of K_S^m, $g = (g^{(v)}) \in \mathrm{GL}(m, K_S)$, and $\det(g) = \big(\det(g^{(v)})\big) \in K_S^*$.

5.2. Our next goal is to define the normalized Haar measure on free K_S-submodules of K_S^m. Let L be a free K_S-module of rank r generated by $\mathbf{a}_1, \dots, \mathbf{a}_r \in K_S^m$. Then one can write $L = \prod_{v \in S} L_v$, where $L_v = K_v \mathbf{a}_1^{(v)} \oplus \cdots \oplus K_v \mathbf{a}_r^{(v)}$. Let us fix a basis $(\mathbf{b}_1, \dots, \mathbf{b}_r)$ of L with the following properties:

$$\text{if } v \in S_f, \quad \text{then } (\mathbf{b}_1^{(v)}, \dots, \mathbf{b}_r^{(v)}) \text{ is an orthonormal basis of } L_v\,; \tag{5.2a}$$

$$\text{if } v \in \infty, \text{ then } L_v \cap \mathcal{O}_v^m = \mathcal{O}_v \mathbf{b}_1^{(v)} + \cdots + \mathcal{O}_v \mathbf{b}_r^{(v)}\,. \tag{5.2b}$$

Then consider the K_S-linear map sending the standard basis $(\mathbf{e}_1, \dots, \mathbf{e}_r)$ of K_S^r to $(\mathbf{b}_1, \dots, \mathbf{b}_r)$, and define the *volume* on L as the pushforward of λ_S (the normalized Haar measure on K_S^r) by this map. When it does not lead to confusion, we will denote this measure on L by λ_S as well.

Note that the existence of the above basis in the case (5.2b) easily follows from the fact that the ring $\mathcal{O}_v$ is a principal ideal domain. Note also that the above definition does not depend on the choice of the basis $(\mathbf{b}_1, \dots, \mathbf{b}_r)$ satisfying (5.2ab), because if $(\mathbf{b}'_1, \dots, \mathbf{b}'_r)$ is another such basis and h is an isomorphism of L sending the first one to the second one, then $c\big(\det(h)\big) = 1$, which implies that h is measure-preserving.

For any $r = 1, \dots, m$ we will also consider the r-th exterior power

$$\textstyle\bigwedge^r K_S^m \simeq \bigoplus_{v \in S} \bigwedge^r K_v^m$$

of K_S^m, which is a free K_S-module with the standard basis

$$\{\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_r} \mid 1 \le i_1 < \cdots < i_r \le m\}\,,$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ is the standard basis of K_S^m. We will keep the notations λ_v, λ_S, $\|\cdot\|_v$ and $c(\cdot)$ to denote the measures, the norms and the content on the exterior powers $\bigwedge^r K_v^m$ and $\bigwedge^r K_S^m$, respectively.

5.3. Recall (cf. [W]) that $\mathcal{O}_S$ is a lattice in K_S with covolume $\sqrt{D_K}$, where D_K is the *discriminant* of K. Any finitely generated $\mathcal{O}_S$-submodule of K is called a *fractional* ideal of K (with respect to $\mathcal{O}_S$). The (fractional) ideal I is called an *integer ideal* (respectively, a *principal ideal*) if $I \subset \mathcal{O}_S$ (respectively, $I = \mathcal{O}_S \alpha$ for some $\alpha \in K$). Recall that the set $P_S(K)$ (respectively, $P_S(K)^0$) of all non-zero fractional (respectively, principal) ideals is an abelian group with respect to multiplication of fractional ideals in K, with $\mathcal{O}_S$ being the neutral element of this group. The group $P_S(K)$ is generated by the principal ideals of $\mathcal{O}_S$. The quotient group $\mathrm{Cl}_S(K) = P_S(K)/P_S(K)^0$ is finite and is called the *group of ideal classes*. If I is a non-zero integer ideal, then I is a subgroup of finite index of the abelian group $\mathcal{O}_S$; this index is called the *norm* of I and is denoted by $N(I)$. The map $N(\cdot)$ can be extended to a group homomorphism $N : P_S(K) \to \mathbb{Q}^+$. (The above well-known facts are proved in [W] in the case when S consists of Archimedean valuations, and a generalization for any S can be easily deduced from this particular case.)

Since $\mathcal{O}_S$ is a Dedekind domain, in view of [CR, 22.5] we have the following

Proposition. *Let Δ be a finitely generated torsion-free module over $\mathcal{O}_S$ of rank r (defined as in* (4.1)*). Then there exist fractional ideals $I_1, \dots, I_r$ such that $\Delta \cong I_1 \oplus \dots \oplus I_r$.*

We also have

5.4. Proposition. *Let Δ be a discrete $\mathcal{O}_S$-submodule of K_S^m of rank r. Then it is finitely generated, torsion-free, and there exist linear independent over K elements $\mathbf{a}_1, \dots, \mathbf{a}_r \in K_S^m$ and integer ideals $I_1, \dots, I_r$ such that*

$$\Delta = I_1\mathbf{a}_1 \oplus \dots \oplus I_r\mathbf{a}_r\,. \tag{5.3}$$

Furthermore, there exists $g \in \mathrm{GL}(m, K_S)$ such that Δ is contained in $g\mathcal{O}_S^m$.

Proof. Let Δ be a discrete $\mathcal{O}_S$-submodule, and denote the intersection of Δ with K_f^m by Δ_0. Since the natural projection of K_f^m onto K_∞^m has compact kernel, Δ_0 is discrete in K_f^m, it follows that the projection of Δ_0 into K_∞^m is a discrete subgroup in a real vector space and, therefore, it is a finitely generated abelian group. Also, because K_f^m is torsion-free, Δ_0 is isomorphic to its projection. If $\mathbf{x}$ is any element in Δ, then there exists an S-adic unit ξ such that $\xi\mathbf{x} \in \Delta_0$. Thus Δ is a finitely generated $\mathcal{O}_S$-module. Also Δ is without torsion because it is contained in a K-vector space and $\mathcal{O}_S \subset K$. Now (5.3), with $I_1, \dots, I_r$ being fractional ideals, follows from Proposition 5.3. Note that one can choose $n \in \mathbb{Z}$ such that $nI_i \subset \mathcal{O}_S$ for all i; thus, replacing $\mathbf{a}_i$ with $\frac{1}{n}\mathbf{a}_i$ we can assume that $I_1, \dots, I_r$ are integer ideals. It remains to note that the discreteness of Δ forces $\mathbf{a}_1^{(v)}, \dots, \mathbf{a}_r^{(v)}$ to be linearly independent over K_v^m for any $v \in S$; thus one can complete $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ to a basis $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$ of the free K_S-module K_S^m, and take g to be the element sending the standard basis of K_S^m to $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$. □

5.5. Remark. Since $\mathrm{Cl}_S(K)$ is a finite group, there exists a finite set of integer ideals $\{J_1, \dots, J_k\}$, where k is the number of elements in $\mathrm{Cl}_S(K)$, representing the ideal classes of $\mathrm{Cl}_S(K)$. So, if Δ is as in Proposition 5.4, there exist $\alpha_i \in K$ such that $I_i\alpha_i \in \{J_1, \dots, J_k\}$ for all i. Therefore, arguing as in the proof of Proposition 5.4, one can see that the ideals $I_1, \dots, I_r$ in (5.3) can be chosen from the apriori fixed set $\{J_1, \dots, J_k\}$.

5.6. It follows from Proposition 5.4 that the set of discrete $\mathcal{O}_S$-submodules of K_S^m can be identified with $\mathfrak{M}(K_S, \mathcal{O}_S, m)$ as defined in (4.2). It also follows that any $\Delta \in \mathfrak{M}(K_S, \mathcal{O}_S, m)$ is a lattice in $K_S\Delta$. Let us now define $\mathrm{cov}(\Delta)$ to be the covolume of Δ in $K_S\Delta$ with respect to the volume on $L = K_S\Delta$ normalized as in §5.2.

The following lemma shows how one can explicitly compute covolumes:

Lemma. *Let $\Delta = I_1\mathbf{a}_1 \oplus \dots \oplus I_r\mathbf{a}_r \in \mathfrak{M}(K_S, \mathcal{O}_S, m)$, where $I_1, \dots, I_r \in P_S(K)$ and $\mathbf{a}_1, \dots, \mathbf{a}_r \in K_S^m$. Then*

$$\mathrm{cov}(I_1\mathbf{a}_1 \oplus \dots \oplus I_r\mathbf{a}_r) = (\sqrt{D_K})^r N(I_1 \cdots I_r) c(\mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_r)\,.$$

Proof. First consider the special case $\Delta = \mathcal{O}_S\mathbf{a}_1 \oplus \dots \oplus \mathcal{O}_S\mathbf{a}_r$. Put $L = K_S\Delta$, define a basis $(\mathbf{b}_1, \dots, \mathbf{b}_r)$ of L as in (5.2ab), and then complete it to a basis $(\mathbf{b}_1, \dots, \mathbf{b}_m)$ of the whole space K_S^m. Also let $h_v \in \mathrm{GL}(m, K_v)$ be such that $h_v \mathbf{b}_i^{(v)} = \mathbf{e}_i^{(v)}$ for all $1 \le i \le m$, where $\mathbf{e}_1^{(v)}, \dots, \mathbf{e}_m^{(v)}$ is the standard basis of K_v^m. It follows from the definition of the measure on L that the map $h : L \to h(L)$, where $h = (h_v) \in \mathrm{GL}(m, K_S)$, is measure preserving. Since the map $\bigwedge^r h : \bigwedge^r K_S^m \to \bigwedge^r K_S^m$ preserves the content c on $\bigwedge^r K_S^m$, we may reduce the problem to the case $\mathbf{b}_i = \mathbf{e}_i$ for all $1 \le i \le m$. Let $\varphi \in \mathrm{GL}(r, K_S)$ be such that $\varphi(\mathbf{e}_i) = \mathbf{a}_i$ for all $1 \le i \le r$. Then since $\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_r = \det(\varphi)\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_r$, we get

$$\operatorname{cov}(\Delta) = \operatorname{cov}(\varphi(\mathcal{O}_S^r)) = c(\det \varphi) \operatorname{cov}(\mathcal{O}_S^r) = (\sqrt{D_K})^r c(\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_r) .$$

For arbitrary $I_1, \dots, I_r$, one can choose an element $\alpha \in \mathcal{O}_S$ such that $\alpha I_i \subset \mathcal{O}_S$ for all i. Since $[\mathcal{O}_S : \alpha I_i] = N(\alpha I_i)$, one has

$$[(\mathcal{O}_S \mathbf{a}_1 \oplus \cdots \oplus \mathcal{O}_S \mathbf{a}_r) : (I_1 \alpha \mathbf{a}_1 \oplus \cdots \oplus I_r \alpha \mathbf{a}_r)] = N(\alpha^r I_1 \cdots I_r) = N(\alpha \mathcal{O}_S)^r N(I_1 \cdots I_r).$$

Also, in view of (5.1),

$$c(\alpha) = \frac{\operatorname{cov}(\alpha \mathcal{O}_S)}{\operatorname{cov}(\mathcal{O}_S)} = N(\alpha \mathcal{O}_S) ,$$

Therefore

$$\begin{aligned}
\operatorname{cov}(I_1 \mathbf{a}_1 \oplus \cdots \oplus I_r \mathbf{a}_r) &= \frac{1}{c(\alpha)^r} \operatorname{cov}(I_1 \alpha \mathbf{a}_1 \oplus \cdots \oplus I_r \alpha \mathbf{a}_r) \\
&= \frac{1}{c(\alpha)^r} N(\alpha \mathcal{O}_S)^r N(I_1 \cdots I_r) \operatorname{cov}(\mathcal{O}_S \mathbf{a}_1 \oplus \cdots \oplus \mathcal{O}_S \mathbf{a}_r) \\
&= (\sqrt{D_K})^r N(I_1 \cdots I_r)\, c(\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_r) ,
\end{aligned}$$

as desired. □

5.7. Lemma 5.6 immediately implies

Corollary. *For every $\Delta \in \mathfrak{M}(K_S, \mathcal{O}_S, m)$, the function $\mathrm{GL}(m, \mathcal{O}_S) \to \mathbb{R}_+$, $g \mapsto \operatorname{cov}(g\Delta)$, is continuous.*

5.8. Corollary. *If $\Delta,\ \Delta' \in \mathfrak{M}(K_S, \mathcal{O}_S, m)$ are such that $K\Delta \cap K\Delta' = \{0\}$, then*

$$\operatorname{cov}(\Delta + \Delta') \le \operatorname{cov}(\Delta) \operatorname{cov}(\Delta').$$

Proof. In view of Proposition 5.3 there exist linearly independent over K elements $\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{b}_1, \dots, \mathbf{b}_s$ and fractional ideals $I_1, \dots I_r, J_1, \dots, J_s$ such that $\Delta = I_1 \mathbf{a}_1 \oplus \cdots \oplus I_r \mathbf{a}_r$ and $\Delta' = J_1 \mathbf{b}_1 \oplus \cdots \oplus J_r \mathbf{b}_r$.

Since

$$K_S \Delta + K_S \Delta' = K_S \Delta \oplus K_S \Delta',$$

in view of Lemma 5.6, it is enough to prove that

$$c(\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_r \wedge \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_s) \le c(\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_r) c(\mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_s),$$

which is easy to verify using the definition of content and the basic properties of the exterior product. □

5.9. In the remaining part of the section we investigate metric properties of discrete submodules of K_S^m. For $\mathbf{x} = (\mathbf{x}^{(v)}) \in K_S^m$ let us define the *norm* of $\mathbf{x}$ by $\|\mathbf{x}\| \stackrel{\text{def}}{=} \sup_v \|(\mathbf{x}^{(v)})\|_v$. If $m = 1$ this norm will be denoted by $|\cdot|$ and will be called the *absolute value*.

Let us now state the following S-arithmetic version of the classical Minkowski's Lemma:

Lemma. *Let $\Delta \in \mathfrak{M}(K_S, \mathcal{O}_S, m)$ be of rank r, and let B be a closed ball in $K_S\Delta$ (with respect to the norm $\|\cdot\|$) centered at 0 such that $\lambda_S(B) \geq 2^{dr} \operatorname{cov}(\Delta)$, where d is the degree of the number field K. Then $\Delta \cap B \neq \{0\}$.*

Proof. Since the volume λ_S on $L = K_S\Delta$ was defined by identifying L with K_S^r via the basis (5.2ab), without loss of generality we can assume that Δ is a lattice in K_S^r. Then we can write $B = B_\infty \times B_f$, where $B_\infty \stackrel{\text{def}}{=} B \cap K_\infty^r$ and $B_f \stackrel{\text{def}}{=} B \cap \left(\prod_{v \in S_f} K_v\right)^r$. Note that

$$\lambda_S(\tfrac{1}{2}B_\infty \times B_f) = \frac{1}{2^{dr}}\lambda_S(B) \geq \operatorname{cov}(\Delta).$$

Since $\frac{1}{2}B_\infty \times B_f$ is closed, the above implies that there exist $\mathbf{x}, \mathbf{y} \in \frac{1}{2}B_\infty \times B_f$, $\mathbf{x} \neq \mathbf{y}$, such that $\mathbf{x} - \mathbf{y} \in \Delta$. This finishes the proof, since clearly $\mathbf{x} - \mathbf{y} \in B$ as well. $\square$

5.10. We will also need the following result:

Lemma. *There exists a constant $A > 1$ depending only on K and S such that if $\mathbf{x} \in K_S^m$ and $c(\mathbf{x}) \neq 0$ then there exists an S-adic unit ξ with*

$$A^{-1}c(\mathbf{x})^{1/\ell} \leq \|\xi\mathbf{x}\| \leq Ac(\mathbf{x})^{1/\ell},$$

where ℓ is the cardinality of S.

Proof. Let

$$H = \{(a_1, \ldots, a_\ell) \in \mathbb{R}_+^\ell \mid a_1 \cdots a_\ell = 1\}.$$

If $S = \{v_1, \ldots, v_\ell\}$, we define the map $i : \mathcal{O}_S^* \to H$, $\xi \to (|\xi|_{v_1}, \ldots, |\xi|_{v_\ell})$. By the proof of the Dirichlet Unit Theorem (see [L, Ch. V, §1]) $i(\mathcal{O}_S^*)$ is a cocompact lattice in the multiplicative group H. Therefore there exists a constant $A > 1$ such that for any $(a_i) \in H$ one can find $\xi \in \mathcal{O}_S^*$ with

$$A^{-1} \leq |\xi|_{v_i} a_i \leq A \tag{5.5}$$

for all i.

Let $\mathbf{x} = (\mathbf{x}^{(v_i)}) \in K_S^m$ and $c(\mathbf{x}) \neq 0$. Note that the vector $\left(\frac{\|\mathbf{x}\|_{v_i}}{c(\mathbf{x})^{1/\ell}}\right)$ is in H, and one has

$$\|\xi\mathbf{x}^{(v_i)}\|_{v_i} = |\xi|_{v_i}\|\mathbf{x}^{(v_i)}\|_{v_i},$$

and

$$c(\mathbf{x}) = c(\xi\mathbf{x})$$

for all $\xi \in \mathcal{O}_S^*$. This, in view of (5.5), implies the claim. $\square$

5.11. Let us put

$$\Omega_{S,m} \stackrel{\text{def}}{=} \{g\mathcal{O}_S^m \mid g \in \mathrm{GL}(m, K_S)\}\,. \tag{5.5}$$

Since $\mathcal{O}_S^m$ is a lattice in K_S^m, every $\Lambda \in \Omega_{S,m}$ is a free $\mathcal{O}_S$-module of rank m and a lattice in K_S^m.

Corollary. *For any $\Lambda \in \Omega_{S,m}$ and any $\rho > 0$, the number of submodules of Λ with covolume $\leq \rho$ is finite.*

Proof. It follows from Lemma 5.6 and Remark 5.5 that it is enough to prove that the number of *free* $\mathcal{O}_S$-submodules of Λ with covolume $< \rho$ is finite. If Δ is a free $\mathcal{O}_S$-submodule with $\mathrm{rk}(\Delta) = r$, then $\bigwedge^r \Delta$ is a free $\mathcal{O}_S$-submodule of rank 1 of the lattice $\bigwedge^r \Lambda$ in $\bigwedge^r K_S^m$, and, in view of Lemma 5.6,

$$\mathrm{cov}(\Delta) = (\sqrt{D_K})^{r-1}\,\mathrm{cov}(\textstyle\bigwedge^r \Delta).$$

Therefore it is enough to prove that the number of rank-one free $\mathcal{O}_S$-submodules of Λ is finite. If $\Delta = \mathcal{O}_S\mathbf{a}$ is such a submodule, then in view of Lemma 5.10 the generator $\mathbf{a}$ can be chosen in such a way that

$$\|\mathbf{a}\| \leq A\left(\frac{\rho}{\sqrt{D_K}}\right)^{1/\ell}.$$

Since Λ is discrete in K_S^m, the set of all $\mathbf{a} \in \Lambda$ satisfying the above inequality is finite, which proves the corollary. $\square$

5.12. Because of lack of an appropriate reference we will prove the S-adic version of Mahler compactness criterion. Consider the group

$$\mathrm{GL}^1(m, K_S) \stackrel{\text{def}}{=} \left\{g \in \mathrm{GL}(m, K_S) \;\middle|\; c\big(\det(g)\big) = 1\right\}, \tag{5.6}$$

consisting of λ_S-preserving linear automorphisms of K_S^m. Let

$$\Omega_{S,m}^1 \stackrel{\text{def}}{=} \left\{\Lambda \in \Omega_{S,m} \mid \mathrm{cov}(\Lambda) = (\sqrt{D_K})^m\right\}.$$

Note that $\mathcal{O}_S^m$ is an element of $\Omega_{S,m}^1$, and its stabilizer in $\mathrm{GL}(m, K_S)$ coincides with $\mathrm{GL}(m, \mathcal{O}_S)$ which is understood to be diagonally imbedded in $\mathrm{GL}(m, K_S)$. Note also that it follows from Artin's Product Formula [C, Theorem 10.2.1] that $\mathrm{GL}(m, \mathcal{O}_S)$ is contained in $\mathrm{GL}^1(m, K_S)$. Thus $\Omega_{S,m}^1$ is naturally identified with the homogeneous space $\mathrm{GL}^1(m, K_S)/\mathrm{GL}(m, \mathcal{O}_S)$. Since $\mathrm{SL}(m, \mathcal{O}_S)$ is a lattice in $\mathrm{SL}(m, K_S)$ and $\mathcal{O}_S^*$ is a lattice in $\mathrm{GL}(1, K_S)$, $\mathrm{GL}(m, \mathcal{O}_S)$ is a lattice in $\mathrm{GL}^1(m, K_S)$.

Let us say that a set Q of lattices in $\mathcal{R}^m$, where $\mathcal{R}$ is a topological ring, is *separated from* 0 if there exists a nonempty neighborhood B of 0 in $\mathcal{R}^m$ such that $\Lambda \cap B = \{0\}$ for all $\Lambda \in Q$.

Theorem (Mahler compactness criterion)**.** *A subset $Q \subset \Omega^1_{S,m}$ is bounded if and only if it is separated from* 0*.*

Proof. The implication ($\Longrightarrow$) is trivial. In order to prove the converse, note that

$$\mathrm{GL}^1(m, K_S) = K^1_S \ltimes \mathrm{SL}(m, K_S) \quad \text{and} \quad \mathrm{GL}(m, \mathcal{O}_S) = \mathcal{O}^*_S \ltimes \mathrm{SL}(m, \mathcal{O}_S)\,,$$

where $K^1_S = \{x \in K_S \mid c(x) = 1\}$. Since $\mathcal{O}^*_S$ is a cocompact lattice in K^1_S, it is enough to prove the theorem with $\Omega^1_{S,m}$ replaced by $\mathrm{SL}(m, K_S)/\,\mathrm{SL}(m, \mathcal{O}_S)$, i.e. with the set

$$\{g\mathcal{O}^m_S \mid g \in \mathrm{SL}(m, \mathcal{O}_S)\}\,.$$

It follows from the strong approximation theorem for classical groups [Kn] that

$$\mathrm{SL}(m, K_S) = \mathrm{SL}(m, K_f)\,\mathrm{SL}(m, \mathcal{O}_S)\,.$$

Thus every $g \in \mathrm{SL}(m, K_S)$ can be represented as $g = g_f g_l$, where $g_f \in \mathrm{SL}(m, K_f)$ and $g_l \in \mathrm{SL}(m, \mathcal{O}_S)$. One has

$$g\mathcal{O}^m_S \cap K^m_f = g_f\left(\mathcal{O}^m_S \cap K^m_f\right) = g_f\mathcal{O}^m\,. \tag{5.7}$$

Let $\tilde{Q} \subset \mathrm{SL}(m, K_S)$ be such that the set of lattices $\{g\mathcal{O}^m_S \mid g \in \tilde{Q}\}$ is separated from 0. It follows from (5.7) that

$$\{g_f\mathcal{O}^m \mid g \in \tilde{Q}\} \text{ is separated from } 0 \text{ in } K^m_f\,. \tag{5.8}$$

Note that $\mathrm{SL}(m, K_f) = \mathrm{SL}(m, K_\infty) \times \mathrm{SL}(m, \mathcal{O}_f)$, and, therefore, every $g_f \in \mathrm{SL}(m, K_f)$ can be written as $g_f = g_\infty g_c$, where $g_\infty \in \mathrm{SL}(m, K_\infty)$ and g_c belongs to the compact group $\mathrm{SL}(m, \mathcal{O}_f)$. It follows from (5.8) that $\{g_\infty\mathcal{O}^m \mid g \in \tilde{Q}\}$ is separated from 0. This reduces the proof to the case when $S = \infty$. In this case, using Weil's restriction of scalars functor $R_{K/\mathbb{Q}}$ [Zi, §6.1], one can identify $\mathrm{SL}(m, K_\infty)$ with the group of $\mathbb{R}$-points G of a $\mathbb{Q}$-algebraic subgroup of $\mathrm{SL}(N, \mathbb{C})$ and $\mathrm{SL}(m, \mathcal{O})$ with $G \cap \mathrm{SL}(N, \mathbb{Z})$. Now the theorem follows from Mahler's compactness criterion for $\mathrm{SL}(N, \mathbb{R})/\,\mathrm{SL}(N, \mathbb{Z})$ (see [R, Corollary 10.9]). □

6. S-arithmetic quantitative nondivergence

6.1. In this section we apply Theorem 4.4 to the triple $(\mathcal{D}, K, \mathcal{R})$ of a particular type. Namely, as in §5, we let K be a number field, choose a finite set S of valuations $|\cdot|_v$ of K containing all the Archimedean ones, and for the rest of this section take $\mathcal{D} = \mathcal{O}_S$ and $\mathcal{R} = K_S$. Then for any $m \in \mathbb{N}$, the set $\mathfrak{M}(K_S, \mathcal{O}_S, m)$, as defined in (4.2), is equal to the set of all submodules of all lattices $\Lambda \in \Omega_{S,m}$, where $\Omega_{S,m}$ is as defined in (5.5). Let us now state the following

Lemma. *The function $\nu : \mathfrak{M}(K_S, \mathcal{O}_S, m) \to \mathbb{R}_+$ given by $\nu(\Delta) = \mathrm{cov}(\Delta)$, with $\mathrm{cov}(\cdot)$ as defined in §5.6, is norm-like, with $C_\nu = 1$.*

Proof. Property (N1) is straightforward since $\Delta' \subset \Delta$ and $K_S\Delta' = K_S\Delta$ implies that Δ is a subgroup of Δ' of finite idex and $\mathrm{cov}(\Delta') = [\Delta : \Delta']\,\mathrm{cov}(\Delta)$. Property (N2) with $C_\nu = 1$ follows from Corollary 5.8 with $\Delta' = \mathcal{O}_S\gamma$. Finally, (N3) has already been mentioned as Corollary 5.7. □

6.2. Now define a function $\delta : \Omega_{S,m} \to \mathbb{R}_+$ by

$$\delta(\Lambda) \stackrel{\text{def}}{=} \min\left\{c(\mathbf{x}) \mid \mathbf{x} \in \Lambda \smallsetminus \{0\}\right\}.$$

Note that the minimum is well defined due to Lemma 5.10 and every $\Lambda \in \Omega_{S,m}$ being discrete in K_S^m. We will use the following

Lemma. *There exists a constant $A > 0$ depending only on K, S and m such that the following holds: for $\rho > 0$ and $\Lambda \in \Omega_{S,m}$ suppose there exists a submodule Δ of Λ with $\operatorname{cov}(\Delta) \le \rho$; then $\delta(\Lambda) \le A\rho^{1/m}$.*

Proof. Take $\varepsilon > 0$ and let B be a ball in $K_S\Delta$ centered at 0 of radius ε (with respect to the norm $\|\cdot\|$ introduced in §5.9). Then one has $\lambda_S(B) \le \text{const} \cdot \varepsilon^{r\ell}$, where const depends only on K, S and m, $r = \operatorname{rk}(\Delta)$, and ℓ is the cardinality of S. By Lemma 5.9, Δ has a nontrivial intersection with B whenever $\text{const} \cdot \varepsilon^{r\ell} \ge 2^{dr}\rho$. This shows how one can choose A such that Δ (and hence Λ) is guaranteed to contain a nonzero vector $\mathbf{x}$ with $\|\mathbf{x}\| \le A^{1/\ell}\rho^{1/r\ell}$, which clearly implies $C(\mathbf{x})A\rho^{1/r} \le A\rho^{1/m}$. $\square$

6.3. As in §4, let us use the notation $\mathfrak{P}(\mathcal{O}_S, m)$ for the set of all nonzero primitive submodules of $\mathcal{O}_S^m$.

Theorem. *Let X be a Besicovitch metric space, μ a uniformly Federer measure on X, and let K and S be as above. For $m \in \mathbb{N}$, let a ball $B = B(x_0, r_0) \subset X$ and a continuous map $h : \tilde{B} \to \mathrm{GL}(m, K_S)$ be given, where $\tilde{B}$ stands for $B(x_0, 3^m r_0)$. Now suppose that for some $C, \alpha > 0$ and $0 < \rho < 1$ one has*

(i) *for every $\Delta \in \mathfrak{P}(\mathcal{O}_S, m)$, the function $\operatorname{cov}\big(h(\cdot)\Delta\big)$ is (C, α)-good on $\tilde{B}$ with respect to μ;*

(ii) *for every $\Delta \in \mathfrak{P}(\mathcal{O}_S, m)$, $\|\operatorname{cov}\big(h(\cdot)\Delta\big)\|_{\mu,B} \ge \rho$.*

Then for any positive $\varepsilon \le \rho/\sqrt{D_K}$ one has

$$\mu\left(\left\{x \in B \mid \delta\big(h(x)\mathcal{O}_S^m\big) < \varepsilon\right\}\right) \le mC\big(N_X D_\mu^2\big)^m \left(\frac{\varepsilon\sqrt{D_K}}{\rho}\right)^\alpha \mu(B)\,. \tag{6.1}$$

Note that [KLW, Theorem 4.3] is a special case (corresponding to $K = \mathbb{Q}$ and $S = \{\infty\}$) of the above theorem.

Proof. To apply Theorem 4.4, one uses Lemma 6.1 which guarantees the norm-like property of $\operatorname{cov}(\cdot)$, and Corollary 5.11 which implies condition (iii) of Theorem 4.4. To derive (6.1) from (4.3) it remains to observe that $\delta\big(h(x)\mathcal{O}_S^m\big) < \varepsilon$ implies the existence of a vector $\mathbf{x} \in h(x)\mathcal{O}_S^m \smallsetminus \{0\}$ with

$$\operatorname{cov}(\mathcal{O}_S\mathbf{x}) = \sqrt{D_K} \cdot c(\mathbf{x}) \le \sqrt{D_K} \cdot \varepsilon\,. \qquad \square$$

6.4. In order to interprete the above result, let us assume, as it will be the case in many applications, that the function h takes values in the group $\mathrm{GL}^1(m, K_S)$. Then for any $x \in B$, $h(x)\mathcal{O}_S^m$ belongs to $\Omega^1_{S,m}$. It follows from Lemma 5.10 and Theorem 5.12 that for any positive ε, the set

$$Q_\varepsilon \stackrel{\text{def}}{=} \{\Lambda \in \Omega^1_{S,m} \mid \delta(\Lambda) \geq \varepsilon\}$$

is compact, and clearly $\Omega^1_{S,m}$ is the union of Q_ε over all $\varepsilon > 0$. This way, Theorem 6.3 estimates, in terms of ε, the relative measure of points $x \in B$ which are mapped, by $x \mapsto h(x)\mathcal{O}_S^m$, to the complement of Q_ε in $\Omega^1_{S,m}$.

As an application, let us take X, μ and h of a special form. Namely, for every $v \in S$ choose $d_v \in \mathbb{N}$, and consider $X = \prod_{v\in S} K_v^{d_v}$, $\mu = \lambda$ as defined in (0.2), and a map $h = (h_v)_{v\in S} : X \to \mathrm{GL}^1(m, K_S)$, where each h_v is a map from $K_v^{d_v}$ to $\mathrm{GL}(m, K_v)$. We say that h is *polynomial* (or *regular*) if for every v all matrix coefficients of h_v and its inverse are polynomials (equivalently, if every h_v is the restriction of a regular map of algebraic varieties $\bar{K}_v^{d_v} \to \mathrm{GL}(m, \bar{K}_v)$, where $\bar{K}_v$ is the algebraic closure of K_v).

Theorem. *Let X and h be defined as above. Then there exists $\alpha > 0$ (depending only on m, d_v and the degrees of the maps) such that for every compact set $L \subset \Omega^1_{S,m}$ one can find positive C_0 and τ (depending only on m, d_v, the degrees of the maps h_v, and L) such that $Q_\tau \supset L$, and the following property holds: for any positive ε and any ball $B \subset X$ one has*

$$\lambda\left(\left\{x \in B \;\middle|\; h(x)\mathcal{O}_S^m \notin Q_\varepsilon\right\}\right) \leq C_0\varepsilon^\alpha\lambda(B) \tag{6.2}$$

whenever $h(B)\mathcal{O}_S^m \cap Q_\tau \neq \varnothing$. Furthermore, if $h(X)\mathcal{O}_S^m \cap Q_\tau = \varnothing$, then there exists a proper $\Delta \in \mathfrak{P}(\mathcal{O}_S, m)$ such that $h(x)(K_S\Delta) = h(0)(K_S\Delta)$ for all $x \in X$.

Proof. It was mentioned in Example 1.6 that X is Besicovitch, and in Example 1.7 that λ is uniformly Federer. Using the exterior power representation of Lemma 5.6, one can easily show that for every $\Delta \in \mathfrak{P}(\mathcal{O}_S, m)$ the function $\mathrm{cov}\big(h(\cdot)\Delta\big)$ has the form $\prod_{v\in S}\|f_v\|_v$, where each f_v is a polynomial map from $K_v^{d_v}$ to another vector space over K_v. Every such function is (C', α')-good with uniform C' and α' due to Lemma 2.4, Lemma 2.1(cd) and Corollary 2.3. Thus condition (i) of Theorem 6.3 is satisfied with some $C, \alpha > 0$.

Now let us take $\tilde{C} \stackrel{\text{def}}{=} \max\big(mC\big(N_X D_\lambda^2\big)^m, 1\big)$. It follows from Lemma 6.2 and Theorem 5.12 that there exists $\tau = \tau(L) > 0$ such that for any submodule Δ of any $\Lambda \in L$ one has $\mathrm{cov}\left(\Delta\right) \geq \tau\sqrt{D_K}$. Without loss of generality we can assume that $\tau < \frac{1}{(2\tilde{C})^{1/\alpha}\sqrt{D_K}}$. Note that L is contained in Q_τ, since by definition of τ one has $c(\mathbf{x}) = \frac{1}{\sqrt{D_K}}\mathrm{cov}(\mathcal{O}_S\mathbf{x}) \geq \tau$ for any nonzero element $\mathbf{x}$ of any $\Lambda \in L$.

If B is such that $h(B)\mathcal{O}_S^m \cap Q_\tau \neq \varnothing$, it follows from Lemma 6.2 that condition (ii) of Theorem 6.3 is satisfied with $(\tau/A)^m$ in place of ρ (where A is as in Lemma 6.2). Therefore one has

$$\lambda\left(\left\{x \in B \;\middle|\; h(x)\mathcal{O}_S^m \notin Q_\varepsilon\right\}\right) \leq \tilde{C}\left(\frac{\varepsilon\sqrt{D_K}A^m}{\tau^m}\right)^\alpha \lambda(B)$$

for all $\varepsilon \le \tau^m / \sqrt{D_K} A^m$. Replacing, if necessary, $\tilde{C}$ by a larger number C, we conclude that (6.2) is valid for all positive ε.

Now assume that $h(x)\mathcal{O}_S^m \notin Q_\tau$ for all $x \in X$. Take $\rho \stackrel{\text{def}}{=} (2\tilde{C})^{1/\alpha}\sqrt{D_K} \cdot \tau < 1$, write $X = \bigcup_{i=1}^\infty B_i$ where B_i are balls centered at 0 with $B_i \subset B_{i+1}$ for all i, and consider

$$\mathfrak{P}_i \stackrel{\text{def}}{=} \left\{\Delta \in \mathfrak{P}(\mathcal{O}_S, m) \mid \| \operatorname{cov}\big(h(\cdot)\Delta\big)\|_{B_i} < \rho\right\}.$$

Then clearly $\mathfrak{P}_{i+1} \subset \mathfrak{P}_i$ for all i, and all these sets are finite due to Corollary 5.11. We claim that $\bigcap_{i=1}^\infty \mathfrak{P}_i$ must be nonempty. Indeed, otherwise one obtains a nonempty ball B such that $\| \operatorname{cov}\big(h(\cdot)\Delta\big)\|_B \ge \rho$ for every $\Delta \in \mathfrak{P}(\mathcal{O}_S, m)$. Thus Theorem 6.3 can be applied, and one can conclude that

$$\lambda(B) = \lambda\left(\left\{x \in B \mid h(x)\mathcal{O}_S^m \notin Q_\tau\right\}\right) \le \tilde{C}\left(\frac{\tau\sqrt{D_K}}{\rho}\right)^\alpha \lambda(B) = \frac{1}{2}\lambda(B),$$

a contradiction.

Consequently, there exists a proper $\Delta \in \mathfrak{P}(\mathcal{O}_S, m)$ such that $\operatorname{cov}\big(h(x)\Delta\big) < \rho$ for all $x \in X$. It follows that each of the polynomials f_v in the aforementioned representation for $\operatorname{cov}\big(h(\cdot)\Delta\big)$ is bounded. Therefore $f_v \equiv \text{const}$ for each v, which implies that $h(x)(K_S\Delta)$ does not depend on x. $\square$

We note that to derive Theorem 0.2 from the above theorem, one needs to take $L = \{\Lambda\}$ and h of the form $x \mapsto h(x)g$, where $g \in \mathrm{GL}^1(m, K_S)$ is such that $\Lambda = g\mathcal{O}_S^m$, and observe that $h(B)\mathcal{O}_S^m \cap Q_\tau$ is nonempty whenever B contains 0.

7. Invariant locally finite measures for actions of unipotent groups on homogeneous spaces

7.1. Theorem 6.4 implies results closely related to the structure of orbits and invariant measures on S-adic homogeneous spaces under the action of subgroups generated by unipotent elements (see [Rat3], [MT1–2] and [T]).

Let us recall some definitions from [MT1]. Let $\mathcal{T}$ be a finite set, and for every $v \in \mathcal{T}$ let K_v be a non-discrete locally compact field of characteristic 0. Denote by $K_\mathcal{T}$ the direct product of all K_v, $v \in \mathcal{T}$. By a *$K_\mathcal{T}$-algebraic group* $\mathbf{G}$ we mean a (formal) direct product $\prod_{v\in\mathcal{T}} \mathbf{G}_v$ of K_v-algebraic groups $\mathbf{G}_v$. The group $\prod_{v\in\mathcal{T}} \mathbf{G}_v(K_v)$ will be denoted by $\mathbf{G}(K_\mathcal{T})$ and called *the group of $K_\mathcal{T}$-rational points* of $\mathbf{G}$. We will also use the simpler notations G for $\mathbf{G}(K_\mathcal{T})$ and G_v for $\mathbf{G}_v(K_v)$. If $\mathbf{H}$ is another $K_\mathcal{T}$-algebraic group then a homomorphism $\varphi : \mathbf{G} \to \mathbf{H}$ is called *$K_\mathcal{T}$-homomorphism* if φ is a product of K_v-homomorphisms of algebraic groups $\varphi_v : \mathbf{G}_v \to \mathbf{H}_v$, $v \in \mathcal{T}$. We preserve the same terminology for the restriction map $\varphi : G \to H$. By the *Zariski topology* on $\mathbf{G}$ (respectively, G) we mean the formal product of the Zariski topologies on $\mathbf{G}_v$ (respectively, G_v). By a *$K_\mathcal{T}$-algebraic subgroup* of G (or simply an algebraic subgroup of G) we mean a Zariski closed subgroup of G. An element $g = (g_v)$ in G is *unipotent* if each component $g_v \in G_v$ is unipotent. A subgroup of G consisting of unipotent elements is called *unipotent*.

Let U be a unipotent algebraic subgroup of G. Then $U = \prod_{v\in\mathcal{T}} U_v$, where U_v are unipotent algebraic subgroups of G_v. Given $v \in \mathcal{T}$, we denote by $\exp_v : \mathrm{Lie}(U_v) \to U_v$

the exponential map and by $\log_v = \exp_v^{-1}$ the logarithmic map. Also we denote by $\mathrm{Lie}(U)$ the direct product of the Lie algebras $\mathrm{Lie}(U_v)$ of U_v, $v \in \mathcal{T}$ and by $\exp : \mathrm{Lie}(U) \to U$ the direct product of the maps $\exp_v$. By a (rational) *parametrization* of U we mean a product $\phi = (\phi_v)_{v\in\mathcal{T}}$ of surjective maps $\phi_v : K_v^{d_v} \to U_v$, $v \in \mathcal{T}$, such that for every $v \in \mathcal{T}$ the map $\log_v \circ \phi_v : K_v^{d_v} \to \mathrm{Lie}(U_v)$ is polynomial and $\phi_v(0) = e$. If d_v is the degree of $\log \circ \phi_v$ then $d = \max\{d_v | v \in \mathcal{T}\}$ is called the *degree* of the parametrization ϕ. Clearly exp is a parametrization of U which we call trivial. We will denote by λ the Haar measure on $X = \prod_{v\in\mathcal{T}} K_v^{d_v}$.

Up to the end of this section we will denote by Γ a lattice in G. Any subgroup of G acts on the homogeneous space G/Γ by left translations.

The following theorem generalizes earlier results for one-parameter real groups proved in [D1, D2] and for one-parametric ultrametric groups announced, with indications of the proof, in [MT1, Theorem 11.4] and [Rat, Theorem 9.1].

Theorem. *Let G and Γ be as above, let d be a natural number, and let L be a compact subset of G/Γ. Then L is contained in a compact L_0 with the following property: given $\beta > 0$ there exists a compact $M_\beta \subset G/\Gamma$ such that for any $y \in G/\Gamma$ and any parametrization $\phi : X \to U$ of degree $\leq d$ of a unipotent algebraic subgroup U of G one of the following holds:*

(i) *If $Uy \cap L_0 \neq \varnothing$ and B is a ball in X with $\phi(B)y \cap L_0 \neq \varnothing$ then*

$$\frac{\lambda\left(\left\{x \in B \mid \phi(x)y \in M_\beta\right\}\right)}{\lambda(B)} \geq 1 - \beta\,. \tag{7.1}$$

In particular, (7.1) *is satisfied if $y \in L$ and B contains the origin.*

(ii) *If $Uy \cap L_0 = \varnothing$, then there exists a proper closed subgroup H of G containing U such that the orbit Hy is closed and carries a finite H-invariant Borel measure.*

Proof. Applying Weil's restriction of scalars functor, we may assume without loss of generality that each K_v is a completion of the field of rationals $\mathbb{Q}$, so that we have $\mathcal{T} = S$ and $K_\mathcal{T} = \mathbb{Q}_S$. Let $R(G)$ be the direct product of the solvable radicals $R(G_v)$ of G_v, $v \in S$. It is easy to see using Auslander's theorem ([Rag, Theorem 8.24], [MT2, Lemma 1.11] and [Rag, Theorem 3.1]) that $R(G) \cap \Gamma$ is a cocompact lattice in $R(G)$. Let N be a maximal subgroup in the class of all normal algebraic subgroups of G containing $R(G)$ as a cocompact subgroup. Let $\mathbf{N}$ be the Zariski closure of N in $\mathbf{G}$. By the general structure theory of algebraic groups [Bo], $\mathbf{H} = \mathbf{G}/\mathbf{N}$ is a $\mathbb{Q}_S$-algebraic group, $\mathbf{H} = \prod_{v\in S} \mathbf{H}_v$ where each $\mathbf{H}_v$ is a semisimple group and $\mathbf{H}_v(K_v)$ has no compact factors, and there exists a $\mathbb{Q}_S$-homomorphism $\varphi : G \to H$ (where $H = \mathbf{H}(\mathbb{Q}_S)$) such that $\varphi(G)$ has finite index in H. Denote $\varphi(\Gamma)$ by Σ. Since $N \cap \Gamma$ is cocompact in N, Σ is a lattice in H and the natural map $\widetilde{\varphi} : G/\Gamma \to H/\Sigma$ is proper. Also note that $\varphi \circ \phi$ is a rational parametrization for $\varphi(U)$ of degree depending only on the degree of ϕ and the homomorphism φ. Therefore without loss of generality we can reduce the proof to the case when every G_v is a semisimple group without compact factors. Furthermore, we may assume that Γ is an irreducible lattice in G. Let $\mathrm{rank}_S\, G = \sum_{p\in S} \mathrm{rank}_{\mathbb{Q}_v} G_v$ be the S-rank of G. If $\mathrm{rank}_S\, G = 1$, then either G is a real rank-one semisimple group and the theorem is proved in [D2] (see also [D3, Remark 3.7]), or G is a p-adic Lie group and, therefore, Γ is a cocompact lattice [Ta] in

G and there is nothing to prove. It remains to consider the case $\operatorname{rank}_S G > 1$. In view of Margulis' Arithmeticity Theorem [Zi, Theorem 10.1.12] we may assume that $\mathbf{G}$ is a $\mathbb{Q}$-algebraic subgroup of $\mathbf{SL}(m)$, $\Gamma = G \cap \mathrm{SL}(m, \mathcal{O}_S)$ and, after eventually replacing $\mathbf{G}$ by its image under a $\mathbb{Q}$-irreducible representation, we may also assume that $\mathbf{G}(\mathbb{Q})$ acts irreducibly on $\mathbb{Q}^m$.

Writing y in the form $g\Gamma$ for some $g \in G$ and applying Theorem 6.4 with the map $h(x) = \phi(x)g$, we get a compact $L_0 = Q_\tau \supset L$ and constants $C, \alpha > 0$ such that for any $\tau > 0$ and any ball $B \subset X$ one has

$$\lambda\left(\left\{x \in B \mid \phi(x)g\mathcal{O}_S^m \in Q_\varepsilon\right\}\right) \geq (1 - C\varepsilon^\alpha)\lambda(B),$$

whenever $\phi(B)y \cap L_0 \neq \varnothing$. Choosing ε such that $\beta \geq C\varepsilon^\alpha$, we get a compact $M_\beta \overset{\text{def}}{=} Q_\varepsilon$ satisfying (7.1).

If $Uy \cap L_0 = \varnothing$, then in view of Theorem 6.4 there exists a proper non-zero vector subspace $V \subset \mathbb{Q}^m$ such that $K_S V$ is invariant under the action of $g^{-1}Ug$. (We consider $\mathbb{Q}^m$ diagonally imbedded in $\mathbb{Q}_S^m$, which justifies the expression $K_S V$.) Let $\mathbf{P}$ be the Zariski closure of the stabilizer of V in $\mathbf{G}(\mathbb{Q})$ under the natural action of $\mathbf{G}(\mathbb{Q})$ on $\mathbb{Q}^m$. Since $\mathbf{G}(\mathbb{Q})$ acts irreducibly on $\mathbb{Q}^m$, $\mathbf{P}$ is a proper $\mathbb{Q}$-algebraic subgroup of $\mathbf{G}$. Therefore so is the subgroup $\mathbf{P}_u$ of $\mathbf{P}$ spanned by all unipotent elements of $\mathbf{P}$. It is easy to see that $g^{-1}Ug \subset \mathbf{P}_u(\mathbb{Q}_S)$. By Borel–Harish-Chandra theorem $\mathbf{P}_u(\mathcal{O}_S)$ is a lattice in $\mathbf{P}_u(\mathbb{Q}_S)$. Therefore the group $H = g\mathbf{P}_u(\mathbb{Q}_S)g^{-1}$ satisfies the requirements of the formulation of the theorem. $\square$

We remark that to derive Theorem 0.3 from the above theorem it suffices to take $L = \{y\}$ and, in the case $Uy \cap L_0 \neq \varnothing$, choose R such that the intersection of $\phi\big(B(0,R)\big)$ with L_0 is nonempty,

7.2. Note that in the above theorem the measure λ can be replaced by a Haar measure σ on U after appropriately choosing a metric on U. Indeed, given a bijective parametrization $\phi : X \to U$, let $\phi_*\lambda$ be the pushforward of λ to U via ϕ. It is enough to find a ϕ such that $\sigma = \phi_*\lambda$. If U is abelian, ϕ can be taken to be the trivial parametrization (because exp is an isomorphism of locally compact groups). In the general case, since U is unipotent and algebraic, there exist abelian algebraic subgroups $U_1, \ldots, U_s$ of U such that the map $\psi : U_1 \times \cdots \times U_s \to U$, $\psi(u_1, \ldots, u_s) = u_1 \cdots u_s$, is bijective, and for any i, $1 \leq i < s$, one knows that $U_1 \cdots U_i$ is a normal subgroup of U and $U_1 \cdots U_{i+1}/U_1 \cdots U_i$ is a central subgroup of $U/U_1 \cdots U_i$. For any j let ϕ_j be the trivial parametrization of the abelian group U_j. Then a simple computation shows that $\phi = \psi \circ (\phi_1, \ldots, \phi_s)$ is such that $\sigma = \phi_*\lambda$, proving the claim. Using the bijective map ϕ we translate the metric from X to U. In view of the preceding discussions, Theorem 7.1 immediately implies the following

Theorem. *Let G and Γ be as in Theorem 7.1. Then every compact $L \subset G/\Gamma$ is contained in a compact L_0 with the following property: given $\beta > 0$ there exists a compact subset M_β of G/Γ such that for any $y \in G/\Gamma$ and any unipotent algebraic subgroup U of G one of the following is satisfied:*

(i) *If $Uy \cap L_0 \neq \varnothing$ and B is a ball in U such that $\phi(B)y \cap L_0 \neq \varnothing$, then*

$$\frac{\sigma\left(\left\{u \in B \mid uy \in M_\beta\right\}\right)}{\sigma(B)} \geq 1 - \beta\,, \tag{7.2}$$

where σ is the Haar measure on U. In particular, (7.2) *holds if $y \in L$ and $e \in B$.*

(ii) *If $Uy \cap L_0 = \varnothing$, then there exists a closed proper subgroup H of G containing U such that the orbit Hy is closed and carries a finite H-invariant Borel measure.*

7.3. We can now proceed with the

Proof of Theorem 0.4. In order to prove the theorem, as in the real case considered in [D1], it is enough to find a function $h \in L^1(G/\Gamma, \mu)$ which is G-invariant and $h(y) > 0$ for μ-a.e. $y \in G/\Gamma$. Indeed, if h is such a function, then the sets

$$Y_i = \left\{ y \in G/\Gamma \mid h(y) \geq \tfrac{1}{i} \right\}$$

satisfy the conditions in the formulation of the theorem.

Among the unipotent algebraic subgroups in H we fix a maximal one and denote it by U. It is well known that the minimal normal subgroup of H containing U coincides with H itself (see [Bo]). In view of the Mautner phenomenon for products of real and p-adic Lie groups [MT2, Proposition 2.1], if $f \in L^1(G/\Gamma, \mu)$ is U-invariant, then there exists an G-invariant μ-integrable function on G/Γ which coincides with f almost everywhere. Therefore it is enough to prove the theorem for $H = U$.

Let $\mu = \int_{(A,\rho)} \mu_a \, d\rho$ be the decomposition of μ into U-invariant ergodic locally finite measures, where (A, ρ) is a measure space parametrizing the ergodic components μ_a. For almost every $y \in G/\Gamma$ there exists a well defined ergodic component $\mu_{a(y)}$, where $a(y) \in A$, whose support contains y. Further on, when it does not lead to confusion, we will identify U with its Lie algebra via exp. Let f be a positive continuous μ-integrable function on G/Γ. In view of [MT1, Proposition 7.1, Lemma 7.2 and formula (4) in §5] and [Tem, Ch. 6], the following version of Birkhoff individual ergodic theorem holds: there exists a sequence of linear parametrizations $\phi_i : U \to U$ such that for any relatively compact measurable neighborhood B of e in U we have $U = \bigcup_i B_i$, where $B_i = \phi_i(B)$, and for almost all $y \in G/\Gamma$

$$\lim_{n \to \infty} \frac{1}{\sigma(B_n)} \int_{B_n} f(gy) \, d\sigma(g) = \int_{G/\Gamma} f(z) \, d\mu_{a(y)}(z) \,,$$

where σ is the Haar measure on U. Furthermore, the limit function

$$f^*(y) \stackrel{\text{def}}{=} \int_{G/\Gamma} f(z) \, d\mu_{a(y)}(z)$$

is μ-integrable and U-invariant. Note that

$$\frac{1}{\sigma(B_n)} \int_{B_n} f(gy) \, d\sigma(g) = \frac{1}{\sigma(B)} \int_B f\big(\phi_n(g) y\big) \, d\sigma(g) \,.$$

It follows from Theorem 7.1 and the fact that $\sigma = \phi_* \lambda$ for some rational parametrization ϕ of U that there exists a compact $M \subset G/\Gamma$ such that

$$\frac{\sigma\big(\{ g \in B \mid \phi_i(g) y \in M \}\big)}{\sigma(B)} > \frac{1}{2}$$

for all i. Since f is positive, the above formula implies that $f^*(y) > 0$ μ-a.e., which, in view of the preceding discussion, completes the proof of the theorem. $\square$

7.4. Remark. Using the methods from [MT2] it is easy to see that Theorems 7.1, 7.2 and 0.4 remain valid for the larger class of so-called *almost linear* groups (that is, when G is a finite direct product of a connected real Lie group and finite central extensions of closed linear p-adic groups) as well as for the class of groups treated in [Rat3].

8. S-arithmetic Diophantine approximation

8.1. In this section we state the main Diophantine result of the present paper, of which Theorems 0.5, 0.6 and 0.7 are special cases. Here and for the rest of the paper we fix a set S of cardinality ℓ consisting of distinct normalized valuations of $\mathbb{Q}$, denote by $\mathbb{Q}_v$ the completion of $\mathbb{Q}$ with respect to $v \in S$ (so that $\mathbb{Q}_v = \mathbb{Q}_p$ if v is the p-adic valuation, and $\mathbb{Q}_v = \mathbb{R}$ if v is the Archimedean one), and let $\mathbb{Q}_S \stackrel{\text{def}}{=} \prod_{v\in S} \mathbb{Q}_v$. We remark that similarly, with some modifications, it is possible to treat the case of an arbitrary finite extension K of $\mathbb{Q}$; however we chose to only discuss the case $K = \mathbb{Q}$ for the sake of making our terminology and results more coincise and transparent.

We will interprete elements

$$\mathbf{y} \;=\; (\mathbf{y}^{(v)})_{v\in S} \;=\; (y_1, \ldots, y_n)$$

of $\mathbb{Q}_S^n$, where $\mathbf{y}^{(v)} = (y_1^{(v)}, \ldots, y_n^{(v)}) \in \mathbb{Q}_v^n$ and $y_i = (y_i^{(v)}) \in \mathbb{Q}_S$, as linear forms on $\mathbb{Q}_S^n$, and will study their values $\mathbf{y}\cdot\mathbf{q} = y_1q_1 + \cdots + y_nq_n$ at integer points $\mathbf{q} = (q_1, \ldots, q_n)$. The approximation properties of our interest will be related to these values being close (in terms of the S-adic absolute value $|\cdot|$ on $\mathbb{Q}_S$) to integers. Alternatively, one could consider a dual case when one approximates (in terms of the S-adic norm) $\mathbf{y} \in \mathbb{Q}_S^n$ by rational vectors. See [KLW] and [KM, Remark 6.2] for a discussion of this set-up in the real case.

For $\mathbf{y}$, $\mathbf{q}$ as above and for $q_0 \in \mathbb{Z}$, it will be convenient to use the notation

$$\tilde{\mathbf{y}} \stackrel{\text{def}}{=} (1, y_1, \ldots, y_n) \quad \text{and} \quad \tilde{\mathbf{q}} \stackrel{\text{def}}{=} (q_0, q_1, \ldots, q_n)\,,$$

so that $q_0 + y_1q_1 + \cdots + y_nq_n = q_0 + \mathbf{y}\cdot\mathbf{q}$ is written as $\tilde{\mathbf{y}}\cdot\tilde{\mathbf{q}}$. Also, by the absolute value $|\cdot|$ of integers and the norm $\|\cdot\|$ of integer vectors we will always mean those coming from the infinite valuation. Hopefully it will cause no confusion.

8.2. A natural starting point in the theory of simultaneous Diophantine approximation is usually a Dirichlet-principle-type result. Let us work it out. The goal is to find the optimal exponent β such that for any $\mathbf{y} \in \mathbb{Q}_S^n$ and any $N > 0$ one is guaranteed to have two different integer vectors $\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2$ of norm $\le N$ such that $\tilde{\mathbf{y}}\cdot\tilde{\mathbf{q}}_1$ and $\tilde{\mathbf{y}}\cdot\tilde{\mathbf{q}}_2$ are at most $\mathrm{const}(\mathbf{y})N^{-\beta}$ apart.

It turns out that the answer depends on whether or not S contains the Archimedean valuation $v = \infty$. Indeed, note that one has

$$|\tilde{\mathbf{y}}^{(v)}\cdot\tilde{\mathbf{q}}|_v \le \max(\|\mathbf{y}^{(v)}\|_v, 1) \tag{8.1}$$

for any ultrametric v and any integer vectors $\tilde{\mathbf{q}}$. Therefore:

- If all the valuations in S are ultrametric, the $(2N+1)^{n+1}$ values of $\tilde{\mathbf{y}}\cdot\tilde{\mathbf{q}}$ for all $\tilde{\mathbf{q}}$ of norm $\le N$ are in the ball of radius $\max(\|\mathbf{y}\|, 1)$ in $\mathbb{Q}_S$. The latter can be partitioned into const $\cdot N^{n+1}$ balls of radius $\mathrm{const}(\mathbf{y})N^{-\frac{n+1}{\ell}}$. Thus for any $\mathbf{y} \in \mathbb{Q}_S^n$ and any $N > 0$ one can find $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1} \smallsetminus \{0\}$ with $\|\tilde{\mathbf{q}}\| \le N$ and

$$|\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}| \le \mathrm{const}(\mathbf{y}) \cdot N^{-\frac{n+1}{\ell}} \,. \tag{8.2}$$

- If $v = \infty$, there is clearly no universal upper bound similar to (8.1). However one can for any given $\mathbf{q} \in \mathbb{Z}^n$ choose $q_0 \in \mathbb{Z}$ such that

$$|\tilde{\mathbf{y}}^{(\infty)} \cdot \tilde{\mathbf{q}}|_\infty = |q_0 + \mathbf{y}^{(\infty)} \cdot \mathbf{q}|_\infty \le 1 \,. \tag{8.1_∞}$$

  Thus, taking all $\mathbf{q}$ of norm $\le N$, one is only guaranteed to have $(2N+1)^n$ values of $\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}$ in the ball of radius $\max(\|\mathbf{y}\|, 1)$ in $\mathbb{Q}_S$. Partitioning it into const $\cdot N^n$ balls of radius $\mathrm{const}(\mathbf{y})N^{-\frac{n}{\ell}}$, one gets a nonzero integer vector $\tilde{\mathbf{q}}$ with $\|\mathbf{q}\| \le N$ and

$$|\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}| = |q_0 + \mathbf{y} \cdot \mathbf{q}| \le \mathrm{const}(\mathbf{y})N^{-\frac{n}{\ell}} \quad \text{for some } q_0 \in \mathbb{Z} \,. \tag{8.2_∞}$$

  Note that the absolute value of q_0 above, and hence the norm of $\tilde{\mathbf{q}}$, is bounded from above by $\mathrm{const}(\mathbf{y})N$.

It will be convenient to define

$$i_S \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } \infty \notin S \\ 0 & \text{if } \infty \in S \,. \end{cases}$$

Then it follows that for any $\mathbf{y} \in \mathbb{Q}_S^n$ the supremum of $w > 0$ for which there exist infinitely many $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1}$ with

$$|\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}|^\ell \le \|\tilde{\mathbf{q}}\|^{-w}$$

is not less than $n + i_S$. On the other hand, it can be easily shown using the Borel-Cantelli Lemma that the above supremum is equal to $n + i_S$ for almost every $\mathbf{y} \in \mathbb{Q}_S^n$ (with respect to Haar measure λ_S on $\mathbb{Q}_S^n$). Similarly to the real/p-adic situation, we say that $\mathbf{y} \in \mathbb{Q}_S^n$ is *very well approximable*, or *VWA*, if the above supremum is strictly bigger than $n + i_S$; in other words, if for some $\varepsilon > 0$ there are infinitely many solutions $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1}$ to

$$|\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}|^\ell \le \|\tilde{\mathbf{q}}\|^{-(n+i_S)(1+\varepsilon)} \,. \tag{8.3}$$

Note that in the case when $\infty \in S$, any solution $\tilde{\mathbf{q}}$ of (8.3) automatically satisfies

$$|q_0| \le 1 + n\|\mathbf{y}^{(\infty)}\|_\infty \|\mathbf{q}\| \,, \tag{8.3$'$}$$

and hence (8.3) can be replaced by

$$|\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}|^\ell \le \|\mathbf{q}\|^{-n(1+\varepsilon)} \,, \tag{8.3_∞}$$

agreeing with the way it was written in the introduction for $S = \{\infty\}$.

8.3. The next step is to define *very well multiplicatively approximable*, or *VWMA*, vectors $\mathbf{y} \in \mathbb{Q}_S^n$. To do this, first let us replace the left hand side of (8.3) by the product of norms of all the components of $\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}$; that is, write

$$c(\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}) \leq \|\tilde{\mathbf{q}}\|^{-(n+i_S)(1+\varepsilon)} . \tag{8.4}$$

Even though the latter inequality is weaker than (8.3), one can still show that for any positive ε it has at most finitely many solutions for almost every $\mathbf{y}$. However, if one wants to weaken it even further by replacing the norm of $\tilde{\mathbf{q}}$ in (8.4) with the geometric mean of its coordinates, one needs to be careful and keep in mind the dichotomy in the Dirichlet-principle argument.

Namely, if $\infty \notin S$ one can indeed replace $\|\tilde{\mathbf{q}}\|^{n+1}$ by $\Pi_+(\tilde{\mathbf{q}})$, and thus define $\mathbf{y} \in \mathbb{Q}_S^n$ to be *VWMA* if for some $\varepsilon > 0$ there are infinitely many solutions $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1}$ to

$$c(\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}) \leq \Pi_+(\tilde{\mathbf{q}})^{-(1+\varepsilon)} , \tag{8.5}$$

as was done in the introduction for $S = \{p\}$.

On the other hand, if $\infty \in S$ it seems tempting to mimic the situation for $S = \{\infty\}$ and define $\mathbf{y}$ to be *VWMA* if for some $\varepsilon > 0$ there are infinitely many $\mathbf{q}$ such that

$$c(\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}) = c(q_0 + \mathbf{y} \cdot \mathbf{q}) \leq \Pi_+(\mathbf{q})^{-(1+\varepsilon)} \tag{8.5'}$$

holds for some $q_0 \in \mathbb{Z}$. However it is not hard to see that, whenever the cardinality of S is at least 2, for any $\varepsilon > 0$ the set of $\mathbf{y} \in \mathbb{Q}_S^n$ for which (8.5′) admits infinitely many solutions has full measure. Indeed, the trouble here comes from the fact that an upper estimate for $c(\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}})$ does not imply a bound similar to (8.3′), that is, a bound on q_0 in terms of $\mathbf{q}$. And one can easily show that for any fixed $\mathbf{q}$ the set of $\mathbf{y} \in \mathbb{Q}_S^n$ for which there exists q_0 saitsfying (8.5′) has full measure.

It follows that in order to achieve a multiplicative analogue of (8.4) in the case $\{\infty\} \subsetneq S$, one needs to take special precautions in the case when $|q_0|$ is much bigger than the norm of $\mathbf{q}$. Here is a suggested way to do it. For any $\mathbf{y} \in \mathbb{Q}_S^n$ with $\infty \in S$ and any $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1}$ let us define

$$\Psi_{\mathbf{y}}(\tilde{\mathbf{q}}) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } |q_0| \leq (1 + n\|\mathbf{y}^{(\infty)}\|_\infty) \cdot \|\mathbf{q}\| \\ |q_0| & \text{otherwise}, \end{cases}$$

and use this function to make a modification to (8.5′). Namely, in the case $\infty \in S$ we will define $\mathbf{y} \in \mathbb{Q}_S^n$ to be *VWMA* if for some $\varepsilon > 0$ there are infinitely many solutions $\mathbf{q}$ to

$$c(\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}) \leq \Pi_+(\mathbf{q})^{-(1+\varepsilon)} \Psi_{\mathbf{y}}(\tilde{\mathbf{q}})^{-\varepsilon} . \tag{8.5$_\infty$}$$

Several remarks are in order. First, as was observed before, if $S = \{\infty\}$, then the second case in the definition of $\Psi_{\mathbf{y}}(\tilde{\mathbf{q}})$ never occurs, thus the above definition is equivalent to the one given in the introduction. Second, it is easily seen that infinitely many solutions to (8.4) imply infinitely many solutions to (8.5$_\infty$), perhaps with a slightly different value of ε. And yet, VWMA as defined above happens to be a zero measure condition. This can be shown directly using a Borel-Cantelli argument, and it will also be an implication of Theorem 8.4 below.

8.4. Generalizing definitions given in the introduction, let us say that a measure μ on $\mathbb{Q}_S^n$ is *extremal* (resp., *strongly extremal*) if μ-almost every point of $\mathbb{Q}_S^n$ is not VWA (resp., not VWMA). Here is the main theorem of the section:

Theorem. *For $v \in S$, let X_v be a metric space with a measure μ_v such that $X = \prod_{v\in S} X_v$ is Besicovitch and $\mu = \prod_{v\in S}\mu_v$ is Federer, and let $\mathbf{f} = (\mathbf{f}^{(v)})_{v\in S}$, where $\mathbf{f}^{(v)}$ are continuous maps from X_v to $\mathbb{Q}_v^n$ which are μ_v-good and μ_v-nonplanar at μ_v-almost every point of X_v. Then $\mathbf{f}_*\mu$ is strongly extremal.*

It is clear from Examples 1.6 and 1.7, Lemma 2.4 and Corollary 2.3 that Theorem 0.7, and hence Theorems 0.5 and 0.6, follow from the above result.

8.5. Let us also remark that Theorem 8.4 generalizes the main result of [KLW]. Indeed, in the latter paper a certain class of measures on $\mathbb{R}^n$ was introduced, and it was proved that measures from that class are strongly extremal. Specifically, following [KLW] let us say that a measure μ on F^n, where F is a locally compact field, is

- *nonplanar* if $\mu(L) = 0$ for any proper affine subspace of F^n;
- *decaying* if for for μ-a.e. $\mathbf{y} \in F^n$ there exist a neighborhood V of $\mathbf{y}$ and $C, \alpha > 0$ such that all affine functions are (C,α)-good on V with respect to μ;
- *friendly* if it is Federer, nonplanar and decaying.

Comparing this with §2.5, one easily observes that μ is decaying if and only if the identity map $F^n \to F^n$ is μ-good at μ-almost every point. It is also not hard to see that the nonplanarity of μ forces the aforementioned identity map to be μ-nonplanar at μ-almost every point (converse is true under the additional assumption that μ is decaying). It is now clear that Theorem 8.4 immediately implies

Corollary. *Let $\mu = \prod_{v\in S}\mu_v$, where μ_v is a friendly measure on $\mathbb{Q}_v^n$ for every $v \in S$. Then μ is strongly extremal.*

Thus [KLW, Theorem 1.1] is a special case of Theorem 8.4. (As was mentioned before, our proof is also a generalization of the argument from [KLW], which, in turn, generalizes the one from [KM].)

8.6. It is not hard to see that many examples of friendly measures on $\mathbb{R}^n$ exhibited in [KLW] can be constructed on a vector space over arbitrary locally compact valued field F. For instance, fix a valuation $|\cdot|$ on F and a norm $\|\cdot\|$ on F^n, and say that a map $\mathbf{h} : F^n \to F^n$ is a *contracting similitude* with *contraction rate* ρ if $\mathbf{h}$ is affine, $0 < \rho < 1$ and

$$\|\mathbf{h}(\mathbf{y})\| = \rho\|\mathbf{y}\| \quad \forall\, \mathbf{y} \in F^n\,.$$

It is known, see [H, §3.1], that for any finite family $\mathbf{h}_1, \dots, \mathbf{h}_m$ of contracting similitudes there exists a unique nonempty compact set Q, called the *attractor* of the family, such that

$$Q = \bigcup_{i=1}^{m} \mathbf{h}_i(Q).$$

Say that $\mathbf{h}_1, \dots, \mathbf{h}_m$ as above satisfy the *open set condition* if there exists an open subset $U \subset F^n$ such that

$$\mathbf{h}_i(U) \subset U \text{ for all } i = 1, \dots, m\,,$$

and

$$i \neq j \Longrightarrow \mathbf{h}_i(U) \cap \mathbf{h}_j(U) = \varnothing\,.$$

J. Hutchinson [H, §5.3] proved[3] that if $\mathbf{h}_i$, $i = 1, \dots, m$, are contracting similitudes with contraction rates ρ_i satisfying the open set condition, and if $s > 0$ is the unique solution of $\sum_i \varrho_i^s = 1$, called the *similarity dimension* of the family $\{\mathbf{h}_i\}$, then the s-dimensional Hausdorff measure $\mathcal{H}^s$ of Q is positive and finite. Let us also say that the family $\{\mathbf{h}_i\}$ is *irreducible* if there does not exist a finite $\{\mathbf{h}_i\}$-invariant collection of proper affine subspaces of F^n. The proof of [KLW, Theorem 2.3] applies verbatim and yields

Proposition. *For any completion $\mathbb{Q}_v$ of $\mathbb{Q}$, let $\{\mathbf{h}_1, \dots, \mathbf{h}_m\}$ be an irreducible family of contracting similitudes of $\mathbb{Q}_v^n$ satisfying the open set condition, s its similarity dimension, μ the restriction of $\mathcal{H}^s$ to its attractor. Then μ is friendly (and hence strongly extremal).*

Measures on $\mathbb{R}^n$ obtained via the above construction have been thoroughly studied; perhaps the simplest example is given by the $\frac{\log 2}{\log 3}$-dimensional Hausdorff measure on the Cantor ternary set. Similarly one can consider ultrametric analogues of the Cantor set, for example let

$$Q = \left\{ \sum_{k=0}^{\infty} a_k 3^k \,\middle|\, a_k = 1, 2 \right\} \subset \mathbb{Z}_3\,.$$

It is a 3-adic version of the Cantor ternary set, which also has Hausdorff dimension $s = \frac{\log 2}{\log 3}$, and it follows that almost all numbers in Q (with respect to the s-dimensional Hausdorff measure) are not VWA.

8.7. We conclude this section with the following modification of Theorem 8.4:

Theorem. *For every $v \in S$, let X_v be a metric space with a measure μ_v such that $X = \prod_{v \in S} X_v$ is Besicovitch and $\mu = \prod_{v \in S} \mu_v$ is uniformly Federer, and let $\mathbf{f} = (\mathbf{f}^{(v)})_{v \in S}$, where $\mathbf{f}^{(v)}$ are continuous maps from X_v to $\mathbb{Q}_v^n$ such that for μ_v-a.e. $x_v \in X_v$ one can find a ball $B_v = B(x_v, r) \subset X_v$ with the following properties:*

$$\text{(8.6)} \qquad \begin{gathered} \textit{for some } C_v, \alpha_v > 0, \textit{ any linear combination of } 1, f_1^{(v)}, \dots, f_n^{(v)} \\ \textit{is } (C_v, \alpha_v)\textit{-good on } B(x_v, 3^{n+1} r) \textit{ with respect to } \mu_v\,, \end{gathered}$$

and

$$\text{(8.7)} \qquad \begin{gathered} \textit{the restrictions of } 1, f_1^{(v)}, \dots, f_n^{(v)} \textit{ to } B_v \cap \operatorname{supp} \mu_v \\ \textit{are linearly independent over } \mathbb{Q}_v\,. \end{gathered}$$

Then $\mathbf{f}_* \mu$ is strongly extremal.

Reduction of Theorem 8.4 to Theorem 8.7. Let X and $\mathbf{f}$ be as in Theorem 8.4. First note that, replacing X_v by appropriate neighborhoods of its μ_v-generic points for each

[3] Hutchinson stated his results for the case $F = \mathbb{R}$, but the proofs apply verbatim to the case of arbitrary locally compact valued field.

v, one can without loss of generality assume that μ is uniformly Federer. Then for any v, since $\mathbf{f}^{(v)}$ is μ_v-good at μ_v-a.e. point, one can for μ_v-a.e. $x_v \in X_v$ choose a neighborhood U_v of x_v and $C_v, \alpha_v > 0$ such that any linear combination of $1, f_1^{(v)}, \dots, f_n^{(v)}$ is (C_v, α_v)-good on U_v with respect to μ_v. Further, since $\mathbf{f}^{(v)}$ is μ_v-nonplanar μ_v-almost everywhere, one can (after throwing away points from a null set) take a ball $B_v = B(x_v, r)$ such that $B(x_v, 3^{n+1}r) \subset U_v$ and (8.7) holds, and the conclusion follows. $\square$

In the next two sections we present the proof of Theorem 8.7, separately considering the cases of S containing or not containing the Archimedean valuation. In both cases the core of the proof is a generalization of the correspondence between real Diophantine approximation and dynamics on real homogeneous spaces.

9. Proof of Theorem 8.7 for $\infty \notin S$

9.1. In order to prove Theorem 8.7, we are going to dynamically interprete the approximation properties of S-adic vectors defined in the previous section, similarly to the approach of [KM]. In this section we suppose that all the valuations in S are ultrametric, that is, $S = \{p_1, \dots, p_\ell\}$ where $p_1, \dots, p_\ell$ are distinct primes. Up to the end of this section we will work with

$$S^* = S \cup \{\infty\}, \quad \mathcal{R} \stackrel{\text{def}}{=} \mathbb{Q}_{S^*} = \mathbb{Q}_S \times \mathbb{R}, \quad \text{and} \quad \mathcal{D} \stackrel{\text{def}}{=} \mathcal{O}_{S^*} = \mathbb{Z}[\tfrac{1}{p_1}, \dots, \tfrac{1}{p_\ell}]\,.$$

Then to any $\mathbf{y} \in \mathbb{Q}_S^n$ we associate a lattice $u_{\mathbf{y}}\mathcal{D}^{n+1}$ in $\mathcal{R}^{n+1}$, where $u_{\mathbf{y}} \in \mathrm{GL}^1(n+1, \mathcal{R})$ is defined by

$$u_{\mathbf{y}}^{(p_j)} = \begin{pmatrix} 1 & \mathbf{y}^{(p_j)} \\ 0 & I_n \end{pmatrix}, \ j = 1, \dots, \ell, \quad u_{\mathbf{y}}^{(\infty)} = I_{n+1}\,,$$

with I_k standing for the $k \times k$ identity matrix. Note that the p_j-adic components of vectors from $u_{\mathbf{y}}\mathcal{D}^{n+1}$ are of the form $\begin{pmatrix} \tilde{\mathbf{y}}^{(p_j)} \cdot \tilde{\mathbf{q}} \\ \mathbf{q} \end{pmatrix}$, where $\tilde{\mathbf{q}} \in \mathcal{D}^{n+1}$.

We need to introduce some more notation. For a vector $\tilde{\mathbf{t}} \stackrel{\text{def}}{=} (t_0, t_1, \dots, t_n) \in \mathbb{R}^{n+1}$ we denote $(t_1, \dots, t_n)$ by $\mathbf{t}$, and let

$$\tilde{t} = \sum_{i=0}^{n} t_i \quad \text{and} \quad t = \sum_{i=1}^{n} t_i \tag{9.1}$$

(this convention will be used throughout the next two sections, so that whenever t and $\mathbf{t}$, or $\tilde{t}$ and $\tilde{\mathbf{t}}$, appear in the same context, (9.1) will be assumed). Then, given $\tilde{\mathbf{t}}$ as above and another vector $\mathbf{s} = (s_1, \dots, s_\ell) \in \mathbb{Z}_+^\ell$, define $g_{\mathbf{s},\tilde{\mathbf{t}}} \in \mathrm{GL}(n+1, \mathcal{R})$ by

$$(g_{\mathbf{s},\tilde{\mathbf{t}}})^{(p_j)} = \begin{pmatrix} p_j^{-s_j} & 0 \\ 0 & I_n \end{pmatrix}, \ j = 1, \dots, \ell, \quad (g_{\mathbf{s},\tilde{\mathbf{t}}})^{(\infty)} = \mathrm{diag}(e^{-t_0}, e^{-t_1}, \dots, e^{-t_n})\,. \tag{9.2}$$

The next lemma shows how a good approximation for $\mathbf{y}$ in the sense of (8.4) gives rise to a translation of $u_{\mathbf{y}}\mathcal{D}^{n+1}$ by $g_{\mathbf{s},\tilde{\mathbf{t}}}$ for some $\mathbf{s}, \tilde{\mathbf{t}}$, so that $\delta(g_{\mathbf{s},\tilde{\mathbf{t}}}u_{\mathbf{y}}\mathcal{D}^{n+1})$ is small. This allows one to use Theorem 6.3 to derive the needed measure estimate.

Lemma. *Let $\varepsilon > 0$, $\mathbf{y} \in \mathbb{Q}_S^n$ and $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1}$ be such that (8.5) holds. For $i = 0, 1, \ldots, n$ define $t_i > 0$ by*

$$|q_i|_+ = \Pi_+(\tilde{\mathbf{q}})^{-\frac{\varepsilon}{n+1}} e^{t_i}\,, \tag{9.3a}$$

and let

$$\gamma = \frac{\varepsilon}{(n+1)(1+\varepsilon)}\,. \tag{9.3b}$$

Then there exists $\mathbf{s} = (s_1, \ldots, s_\ell) \in \mathbb{Z}_+^\ell$ such that

$$\delta(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{y}} \mathcal{D}^{n+1}) \le \sqrt{n+1} e^{-\gamma \tilde{t}} \tag{9.3c}$$

and

$$\prod_{j=1}^{\ell} p_j^{s_j} \le e^{\tilde{t}} < \prod_{j=1}^{\ell} p_j^{s_j+1} \tag{9.3d}$$

Proof. Multiplying equalities (9.3a), we get

$$e^{\tilde{t}} = \Pi_+(\tilde{\mathbf{q}})^{1+\varepsilon} \tag{9.3e}$$

and

$$e^{-t_i}|q_i| \le e^{-t_i}|q_i|_+ \underset{(9.3a)}{\le} \Pi_+(\tilde{\mathbf{q}})^{-\frac{\varepsilon}{n+1}} \underset{(9.3be)}{=} e^{-\gamma\tilde{t}},\ i = 0, 1, \ldots, n\,,$$

hence $\|(g_{\mathbf{s},\tilde{\mathbf{t}}})^{(\infty)} u_{\mathbf{y}}^{(\infty)} \tilde{\mathbf{q}}\|_\infty = \|(g_{\mathbf{s},\tilde{\mathbf{t}}})^{(\infty)} \tilde{\mathbf{q}}\|_\infty \le \sqrt{n+1} e^{-\gamma\tilde{t}}$.

Now let us define s_j, $j = 1, \ldots, \ell$, inductively by

$$p_j^{s_j} \le \min\left(\frac{e^{\tilde{t}}}{\prod_{i=1}^{j-1} p_i^{s_i}}, \frac{1}{|\tilde{\mathbf{y}}^{(p_j)} \cdot \tilde{\mathbf{q}}|_{p_j}}\right) < p_j^{s_j+1} \tag{9.4}$$

(where if $j = 1$ we set $\prod_{i=1}^{j-1} p_i^{s_i} = 1$). This, in particular, implies that

$$|p_j^{-s_j} \tilde{\mathbf{y}}^{(p_j)} \cdot \tilde{\mathbf{q}}|_{p_j} = p_j^{s_j} |\tilde{\mathbf{y}}^{(p_j)} \cdot \tilde{\mathbf{q}}|_{p_j} \le 1$$

for each j. Taking into account that $|q_i|_{p_j} \le 1$ for all i and j, one concludes that $c(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{y}} \tilde{\mathbf{q}}) \le \sqrt{n+1} e^{-\gamma\tilde{t}}$.

It remains to check that inequalities (9.3d) are satisfied. Taking $j = \ell$ in (9.4) immediately implies the lower estimate. To prove the upper estimate, let us consider two cases:

- If for some j the minimum in (9.4) is equal to the first of the quantities compared, then clearly

$$e^{\tilde{t}} < p_j^{s_j+1} \prod_{i=1}^{j-1} p_i^{s_i} \le \prod_{j=1}^{\ell} p_j^{s_j+1}\,.$$

- Otherwise, it follows that $|\tilde{\mathbf{y}}^{(p_j)} \cdot \tilde{\mathbf{q}}|_{p_j} > p_j^{-(s_j+1)}$ for all j, and to derive the desired estimate it remains to notice that (8.5), in view of (9.3e), can be rewritten as $c(\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}) = \prod_{j=1}^{\ell} |\tilde{\mathbf{y}}^{(p_j)} \cdot \tilde{\mathbf{q}}|_{p_j} \le e^{-\tilde{t}}$. $\square$

9.2. Corollary. *Assume that* $\mathbf{y} \in \mathbb{Q}_S^n$ *is VWMA. Then for some* $c, \gamma > 0$ *there are infinitely many* $\tilde{\mathbf{t}} \in \mathbb{Z}_+^{n+1}$ *and* $\mathbf{s} \in \mathbb{Z}_+^\ell$ *such that*

$$e^{-(n+1)} \prod_{j=1}^{\ell} p_j^{s_j} \le e^{\tilde{t}} < \prod_{j=1}^{\ell} p_j^{s_j+1} \tag{9.5a}$$

and

$$\delta(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{y}} \mathcal{D}^{n+1}) \le c e^{-\gamma \tilde{t}}\,. \tag{9.5b}$$

Proof. By definition, for some $\varepsilon > 0$ there are infinitely many solutions $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1}$ of (8.5). Therefore, by the above lemma and with γ as in (9.3b), there exists an unbounded set of $\tilde{\mathbf{t}} \in \mathbb{R}_+^{n+1}$ such that (9.3c) holds for some $\mathbf{s} \in \mathbb{Z}_+^\ell$ satisfying (9.3d). Denote by $[\tilde{\mathbf{t}}]$ the vector consisting of integer parts of t_i, then clearly the ratio of $\delta(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{y}} \mathcal{D}^{n+1})$ and $\delta(g_{\mathbf{s},[\tilde{\mathbf{t}}]} u_{\mathbf{y}} \mathcal{D}^{n+1})$ is bounded from above by some uniform constant. Thus, replacing $\tilde{\mathbf{t}}$ by $[\tilde{\mathbf{t}}]$, for some $c > 0$ one gets infinitely many solutions $\tilde{\mathbf{t}} \in \mathbb{Z}_+^{n+1}$ of (9.5b), with $e^{\tilde{t}}$ being smaller than before by at most a factor of e^{n+1}, hence (9.5a). $\square$

9.3. Corollary. *Let* X *be a Besicovitch metric space and* μ *a uniformly Federer measure on* X. *Suppose we are given a continuous map* $\mathbf{f} : X \to \mathbb{Q}_S^n$ *such that for* μ*-a.e.* $x_0 \in X$ *there exist a ball* $B = B(x_0, r)$ *and constants* C, α, ρ *with the following property: for any* $\Delta \in \mathfrak{P}(\mathcal{D}, n+1)$ *and any* $\mathbf{s} \in \mathbb{Z}_+^\ell$, $\tilde{\mathbf{t}} \in \mathbb{Z}_+^{n+1}$ *satisfying* (9.5a), *one has*

$$\text{the function } x \mapsto \operatorname{cov}\big(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \Delta\big) \text{ is } (C,\alpha)\text{-good on } B(x_0, 3^{n+1} r) \text{ w.r.t. } \mu\,, \tag{9.6a}$$

and

$$\sup_{x \in B \,\cap\, \operatorname{supp} \mu} \operatorname{cov}\big(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \Delta\big) \ge \rho\,. \tag{9.6b}$$

Then $\mathbf{f}_* \mu$ *is strongly extremal.*

Proof. Applying Theorem 6.3, with $h(x) = g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)}$ and $m = n+1$, we conclude that

$$\mu\Big(\big\{x \in B \;\big|\; \delta(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathcal{D}^{n+1}) < c e^{-\gamma \tilde{t}}\big\}\Big) \le (n+1) C \big(N_X D_\mu^2\big)^{n+1} \left(\frac{c e^{-\gamma \tilde{t}}}{\rho}\right)^{\alpha} \mu(B)$$

whenever $c e^{-\gamma \tilde{t}} \le \rho$ and (9.5a) holds. Note that for fixed $\tilde{\mathbf{t}}$, the number of different $\mathbf{s} \in \mathbb{Z}_+^\ell$ satisfying (9.5a) is at most const $\cdot \tilde{t}^{\,\ell-1}$. Therefore the sum (over all integer $\mathbf{s}, \tilde{\mathbf{t}}$ for which (9.5a) holds) of measures of sets $\big\{x \in B \;\big|\; \delta(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathcal{D}^{n+1}) < c e^{-\gamma \tilde{t}}\big\}$ is finite for every $c, \gamma > 0$. An application of the Borel-Cantelli Lemma shows that for every $c, \gamma > 0$ and μ-a.e. $x \in B$, and hence for μ-a.e. $x \in B$ and all $c, \gamma > 0$, there are at most finitely many integer solutions $\mathbf{s}, \tilde{\mathbf{t}}$ to (9.5ab). Corollary 9.2 then implies that $\mathbf{f}(x)$ is not VWMA for μ-a.e. $x \in B$. $\square$

9.4. We are now ready for the

Proof of Theorem 8.7, the case $\infty \notin S$. Recall that we are given the balls $B_v \subset X_v$, $v \in S$, which will be referred to as $B_1, \dots, B_\ell$, and measures μ_v on X_v, which we will call $\mu_1, \dots, \mu_\ell$. We will take B to be equal to $\prod_{j=1}^{\ell} B_j$ (recall that we are using the product metric on X) and show that it satisfies the assumptions of Corollary 9.3. Thus we need to have explicit expressions for functions $x \mapsto \operatorname{cov}\big(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \Delta\big)$.

Recall that any nonzero submodule $\Delta \subset \mathcal{D}^{n+1}$ of rank r can be written in the form (5.3), and, in view of Lemma 5.6, $D_{\mathbb{Q}} = 1$ and $\mathcal{D}$ being a principal ideal ring, we know that $\operatorname{cov}(\Delta) = c(\mathbf{w})$, where $\mathbf{w} \stackrel{\text{def}}{=} \mathbf{a}_1 \wedge \dots \wedge \mathbf{a}_r$ and $\mathbf{a}_1, \dots, \mathbf{a}_r$ are nonzero elements of $\mathcal{D}$. It will be convenient to use the standard basis $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_n$ of $\mathcal{R}^{n+1}$, where

$$\mathbf{e}_i = \left(\mathbf{e}_i^{(v)}\right)_{v \in S^*} = \left(\mathbf{e}_i^{(p_1)}, \dots, \mathbf{e}_i^{(p_\ell)}, \mathbf{e}_i^{(\infty)}\right)$$

for each $i = 0, 1, \dots, n$. Similarly, we will use the standard basis $\big\{\mathbf{e}_I \mid I \subset \{0, 1, \dots, n\}\big\}$ of $\bigwedge \mathcal{R}^{n+1}$, where we let $\mathbf{e}_I \stackrel{\text{def}}{=} \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_r} \in \bigwedge^r(\mathcal{R}^{n+1})$ for $I = \{i_1, \dots, i_r\} \subset \{0, \dots, n\}$, $i_1 < i_2 < \dots < i_r$. Thus we can write $\mathbf{w}$ as above in the form $\mathbf{w} = \sum_{I \subset \{0,\dots,n\}} w_I \mathbf{e}_I$, where $w_I \in \mathcal{D}$.

Now let us see how the coordinates of $\mathbf{w}$ as above change under the action of $g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)}$. Note that:

- $u_{\mathbf{f}(x)}^{(\infty)}$ is trivial, and each $\mathbf{e}_I^{(\infty)}$ is an eigenvector of $(g_{\mathbf{s},\tilde{\mathbf{t}}})^{(\infty)}$ with eigenvalue e^{-t_I}, where $t_I \stackrel{\text{def}}{=} \sum_{i \in I} t_i$;
- the action of $u_{\mathbf{f}(x)}^{(p_j)}$ leaves $\mathbf{e}_0^{(p_j)}$ invariant and sends $\mathbf{e}_i^{(p_j)}$ to $\mathbf{e}_i^{(p_j)} + f_i^{(p_j)}(x)\mathbf{e}_0^{(p_j)}$, $i = 1, \dots, n$, and each $\mathbf{e}_I^{(p_j)}$ is an eigenvector of $(g_{\mathbf{s},\tilde{\mathbf{t}}})^{(p_j)}$ with eigenvalue 1 if $0 \notin I$ and p^{-s_j} otherwise; in other words,

$$(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathbf{e}_I)^{(p_j)} = \begin{cases} p_j^{-s_j} \mathbf{e}_I^{(p_j)} \text{ if } 0 \in I \\ \mathbf{e}_I^{(p_j)} + p_j^{-s_j} \sum_{i \in I} \pm f_i^{(p_j)}(x) \mathbf{e}_{I \cup \{0\} \smallsetminus \{i\}}^{(p_j)} \text{ otherwise}. \end{cases} \tag{9.7}$$

Therefore one has $(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathbf{w})^{(\infty)} = \sum_I e^{-t_I} w_I \mathbf{e}_I^{(\infty)}$ and

$$(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathbf{w})^{(p_j)} = \sum_{0 \notin I} w_I \mathbf{e}_I^{(p_j)} + p_j^{-s_j} \sum_{0 \in I} \Big(w_I + \sum_{i \notin I} \pm w_{I \cup \{i\} \smallsetminus \{0\}} f_i^{(p_j)}(x)\Big) \mathbf{e}_I^{(p_j)} \tag{9.8}$$

for $j = 1, \dots, \ell$.

In particular, real components of all the coordinates of $g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathbf{w}$ are constant, and p_j-adic components are linear combinations of $1, f_1^{(p_j)}, \dots, f_n^{(p_j)}$. Condition (9.6a) then immediately follows from Lemma 2.1(bc), (8.6) and Corollary 2.3. On the other hand, for any $j = 1, \dots, \ell$ one can use (8.7) and the compactness of the unit sphere in $\mathbb{Q}_{p_j}^{n+1}$ to find $\rho_j > 0$ such that for any $\mathbf{a} = (a_0, a_1, \dots, a_n) \in \mathbb{Q}_{p_j}^{n+1}$ one has

$$\sup_{x \in B_j \cap \operatorname{supp} \mu_j} |a_0 + a_1 f_1^{(p_j)}(x) + \dots + a_n f_n^{(p_j)}(x)|_{p_j} \geq \rho_j \|\mathbf{a}\|_{p_j}\,. \tag{9.9}$$

It remains to notice that all the components of $\mathbf{w}$ necessarily appear in the second sum in (9.8) (that is, the sum of terms with $0 \in I$). Therefore (9.8) and (9.9) imply

$$\sup_{x \in B_j \cap \operatorname{supp} \mu_j} \|(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathbf{w})^{(p_j)}\|_{p_j} \geq \rho_j {p_j}^{s_j} \max_I |w_I|_{p_j} \,, \tag{9.10}$$

and hence

$$\begin{aligned} \sup_{x \in B \cap \operatorname{supp} \mu} c(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathbf{w}) &\geq \left(\prod_{j=1}^{\ell} \rho_j {p_j}^{s_j} \max_I |w_I|_{p_j} \right) \max_I e^{-t_I} |w_I|_\infty \\ &\geq \left(\prod_{j=1}^{\ell} \rho_j \right) e^{-\tilde{t}} \left(\prod_{j=1}^{\ell} {p_j}^{s_j} \right) \max_I c(w_I) \underset{(9.5a)}{\geq} \prod_{j=1}^{\ell} \frac{\rho_j}{p_j} \,. \end{aligned}$$

Condition (9.6b) is thus established, and the theorem follows. $\square$

10. Proof of Theorem 8.7 for $\infty \in S$

10.1. In this section we suppose that $S = \{p_1, \dots, p_{\ell-1}, \infty\}$, where $p_1, \dots, p_{\ell-1}$ are distinct primes. In this case there is no need to artificially add the infinite valuation to S; that is, we will now work with

$$\mathcal{R} \stackrel{\text{def}}{=} \mathbb{Q}_S \quad \text{and} \quad \mathcal{D} \stackrel{\text{def}}{=} \mathcal{O}_S = \mathbb{Z}[\tfrac{1}{p_1}, \dots, \tfrac{1}{p_{\ell-1}}] \,.$$

The element $u_{\mathbf{y}} \in \mathrm{GL}^1(n+1, \mathcal{R})$ can be simply defined by $u_{\mathbf{y}} \stackrel{\text{def}}{=} \begin{pmatrix} 1 & \mathbf{y} \\ 0 & I_n \end{pmatrix}$, so that

$$u_{\mathbf{y}} \mathcal{D}^{n+1} = \left\{ \begin{pmatrix} \tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}} \\ \mathbf{q} \end{pmatrix} \middle| \, \tilde{\mathbf{q}} \in \mathcal{D}^{n+1} \right\} .$$

The definition (9.2) of the diagonal element $g_{\mathbf{s},\tilde{\mathbf{t}}} \in \mathrm{GL}(n+1, \mathcal{R})$ given in the previous section will still be valid, except for $\mathbf{s}$ now having $\ell - 1$ components.

Now let us recall that the definition of VWMA vectors is more complicated in this case than the one used in the previous section; in fact, the set of VWMA vectors can be split into two parts, according to the two alternatives in the definition of $\Psi_{\mathbf{y}}$. It will be convenient to say that $\mathbf{y} \in \mathbb{Q}_S^n$ is *VWMA*$_\leq$ if for some positive ε there are infinitely many solutions $\tilde{\mathbf{q}}$ to

$$c(\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}) \leq \Pi_+(\mathbf{q})^{-(1+\varepsilon)} \quad \text{and} \quad |q_0| \leq \left(1 + n\|\mathbf{y}^{(\infty)}\|_\infty\right) \|\mathbf{q}\| \,, \tag{10.1$\leq$}$$

and that it is *VWMA*$_>$ if for some ε there are infinitely many solutions to

$$c(\tilde{\mathbf{y}} \cdot \tilde{\mathbf{q}}) \leq \Pi_+(\mathbf{q})^{-(1+\varepsilon)} |q_0|^{-\varepsilon} \quad \text{and} \quad |q_0| > \left(1 + n\|\mathbf{y}^{(\infty)}\|_\infty\right) \|\mathbf{q}\| \,. \tag{10.1$>$}$$

Clearly $\mathbf{y}$ is VWMA if and only if it is either VWMA$_\leq$ or VWMA$_>$. Our strategy will be as follows: we will modify the dynamical approach of the previous section to treat the first case, and use the conclusion of the "$\infty \notin S$" case of Theorem 8.7 to take care of the second case.

Here is a replacement for Lemma 9.1.

Lemma. *Let $\varepsilon > 0$, $\mathbf{y} \in \mathbb{Q}_S^n$ and $\tilde{\mathbf{q}} \in \mathbb{Z}^{n+1}$ be such that (10.1≤) holds. For $i = 1, \dots, n$ define $t_i > 0$ by*

$$|q_i|_+ = \Pi_+(\mathbf{q})^{-\frac{\varepsilon}{n+1}} e^{t_i}\,, \tag{10.2a}$$

and let

$$\gamma = \frac{\varepsilon}{n+1+n\varepsilon}\,. \tag{10.2b}$$

Then there exist $\mathbf{s} = (s_1, \dots, s_{\ell-1}) \in \mathbb{Z}_+^{\ell-1}$ and $t_0 \in \mathbb{R}$ such that

$$\delta(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{y}} \mathcal{D}^{n+1}) \le \sqrt{n+1} e^{-\gamma t}\,, \tag{10.2c}$$

$$-t \le t_0 \le t + \ln\left(1 + 2n\|\mathbf{y}^{(\infty)}\|_\infty\right)\,, \tag{10.2d}$$

and

$$\prod_{j=1}^{\ell-1} p_j^{s_j} \le e^{\tilde{t}} < \prod_{j=1}^{\ell-1} p_j^{s_j+1}\,. \tag{10.2e}$$

Proof. As in the proof of Lemma 9.1, we consider the product of equalities (10.2a), namely

$$e^t = \Pi_+(\mathbf{q})^{1+\frac{n}{n+1}\varepsilon}\,, \tag{10.2f}$$

and then write

$$e^{-t_i}|q_i| \le e^{-t_i}|q_i|_+ \underset{(10.2a)}{\le} \Pi_+(\mathbf{q})^{-\frac{\varepsilon}{n+1}} \underset{(10.2bf)}{=} e^{-\gamma t},\ i = 1, \dots, n\,. \tag{10.2g}$$

After that define t_0 by

$$e^{-t_0} \stackrel{\text{def}}{=} \min\left(e^t, \frac{e^{-\gamma t}}{|\tilde{\mathbf{y}}^{(\infty)} \cdot \tilde{\mathbf{q}}|_\infty}\right)\,. \tag{10.3}$$

It follows that $|e^{-t_0}\tilde{\mathbf{y}}^{(\infty)} \cdot \tilde{\mathbf{q}}|_\infty \le e^{-\gamma t}$, hence $\|(g_{\mathbf{s},\tilde{\mathbf{t}}})^{(\infty)} u_{\mathbf{y}}^{(\infty)} \tilde{\mathbf{q}}\|_\infty \le \sqrt{n+1}e^{-\gamma t}$. The lower estimate in (10.2d) is immediate from (10.3), while the upper estimate clearly holds if the minimum in (10.3) is equal to e^t, and otherwise one has

$$e^{-t_0} = \frac{e^{-\gamma t}}{|q_0 + \mathbf{y}^{(\infty)} \cdot \mathbf{q}|_\infty} \underset{(10.1\le)}{\ge} \frac{e^{-\gamma t}}{\left(1 + 2n\|\mathbf{y}^{(\infty)}\|_\infty\right)\|\mathbf{q}\|} \underset{(10.2g)}{\ge} \frac{e^{-t}}{1 + 2n\|\mathbf{y}^{(\infty)}\|_\infty}\,.$$

Now that all the components of $\tilde{\mathbf{t}}$ are chosen, we can define s_j, $j = 1, \dots, \ell - 1$, as in (9.4). After that one can verify, following the lines of the proof of Lemma 9.1, that

$$\|(g_{\mathbf{s},\tilde{\mathbf{t}}})^{(p_j)} u_{\mathbf{y}}^{(p_j)} \tilde{\mathbf{q}}\|_{p_j} = \max\left(p_j^{s_j}|\tilde{\mathbf{y}}^{(p_j)} \cdot \tilde{\mathbf{q}}|_{p_j}, 1\right) \le 1$$

for each j, so that $c(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{y}} \tilde{\mathbf{q}}) \le \sqrt{n+1}e^{-\gamma t}$, and that inequalities (10.2e) are satisfied. □

10.2. Corollary. *Assume that* $\mathbf{y} \in \mathbb{Q}_S^n$ *is VWMA*$_\le$*. Then for some* $c_0, c, \gamma > 0$ *there are infinitely many* $\tilde{\mathbf{t}} \in \mathbb{Z} \times \mathbb{Z}_+^n$ *and* $\mathbf{s} \in \mathbb{Z}_+^{\ell-1}$ *satisfying*

$$-t - 1 \le t_0 \le t + c_0 \,, \tag{10.4a}$$

$$e^{-(n+1)} \prod_{j=1}^{\ell-1} p_j^{s_j} \le e^{\tilde{t}} < \prod_{j=1}^{\ell-1} p_j^{s_j+1} \,, \tag{10.4b}$$

and

$$\delta(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{y}} \mathcal{D}^{n+1}) \le ce^{-\gamma t} \,. \tag{10.4c}$$

Proof. By definition, for some $\varepsilon > 0$ there are infinitely many solutions $\tilde{\mathbf{q}}$ to (10.1$\le$). Therefore, by the above lemma and with γ as in (10.2b), there exists an unbounded set of $\mathbf{t} \in \mathbb{R}_+^n$ such that inequalities (10.2cde) hold for some $t_0 \in \mathbb{R}$ and $\mathbf{s} \in \mathbb{Z}_+^{\ell-1}$. The rest of the proof of Corollary 9.2 applies verbatim. $\square$

10.3 Corollary. *Let* X *be a Besicovitch metric space and* μ *a uniformly Federer measure on* X*. Suppose we are given a continuous map* $\mathbf{f} : X \to \mathbb{Q}_S^n$ *with the follolwing property: for* μ*-a.e.* $x_0 \in X$ *there exist a ball* $B = B(x_0, r)$ *and constants* C, α, ρ, c_0 *such that conditions* (9.6ab) *hold for any* $\Delta \in \mathfrak{P}(\mathcal{D}, n+1)$ *and any* $\mathbf{s} \in \mathbb{Z}_+^{\ell-1}$*,* $\mathbf{t} \in \mathbb{Z} \times \mathbb{Z}_+^n$ *satisfying* (10.4ab)*. Then* $\mathbf{f}(x)$ *is not VWMA*$_\le$ *for* μ*-a.e.* $x \in X$*.*

Proof. An application of Theorem 6.3, again with $h(x) = g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)}$ and $m = n+1$, yields

$$\mu\left(\left\{x \in B \mid \delta(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathcal{D}^{n+1}) < ce^{-\gamma t}\right\}\right) \le (n+1)C\big(N_X D_\mu^2\big)^{n+1} \left(\frac{ce^{-\gamma t}}{\rho}\right)^\alpha \mu(B)$$

whenever $ce^{-\gamma t} \le \rho$ and (10.4ab) hold. Now observe that for fixed $\mathbf{t}$, the number of different $t_0 \in \mathbb{Z}$ and $\mathbf{s} \in \mathbb{Z}_+^{\ell-1}$ satisfying (10.4ab) is at most const $\cdot t^{\ell-1}$. Therefore for any $c, \gamma > 0$ the sum (over all integers $\mathbf{s}, \tilde{\mathbf{t}}$ for which inequalities (10.4ab) hold) of measures of sets $\left\{x \in B \mid \delta(g_{\mathbf{s},\tilde{\mathbf{t}}} u_{\mathbf{f}(x)} \mathcal{D}^{n+1}) < ce^{-\gamma t}\right\}$ converges. As before, an application of the Borel-Cantelli Lemma shows that for μ-a.e. $x \in B$ there are at most finitely many integer solutions $\mathbf{s}, \tilde{\mathbf{t}}$ to (10.4abc) for any $c, \gamma > 0$. Corollary 10.2 then implies that $\mathbf{f}(x)$ is not VWMA$_\le$ for μ-a.e. $x \in B$. $\square$

10.4. Now let us state a lemma showing that $\mathbf{y}$ being VWMA$_>$ has some implications to the Diophantine properties of its "finite part" $\big(\mathbf{y}^{(v)}\big)_{v \in S_f}$.

Lemma. *Assume that* $\mathbf{y} = \big(\mathbf{y}^{(v)}\big)_{v \in S}$ *is VWMA*$_>$*; then* $\big(\mathbf{y}^{(v)}\big)_{v \in S_f}$ *is VWMA.*

Proof. Note that the second inequality in (10.1>) implies

$$\begin{aligned}
|\tilde{\mathbf{y}}^{(\infty)} \cdot \tilde{\mathbf{q}}|_\infty &= |q_0 + \mathbf{y}^{(\infty)} \cdot \mathbf{q}|_\infty \ge |q_0| - n\|\mathbf{y}^{(\infty)}\|_\infty \|\mathbf{q}\| \\
&\ge |q_0| - \frac{n\|\mathbf{y}^{(\infty)}\|_\infty}{1 + n\|\mathbf{y}^{(\infty)}\|_\infty}|q_0| = \frac{1}{1 + n\|\mathbf{y}^{(\infty)}\|_\infty}|q_0| \,.
\end{aligned}$$

Therefore one has

$$\prod_{v\in S_f} |\tilde{\mathbf{y}}^{(v)}\cdot\tilde{\mathbf{q}}|_v = \frac{c(\tilde{\mathbf{y}}\cdot\tilde{\mathbf{q}})}{|\tilde{\mathbf{y}}^{(\infty)}\cdot\tilde{\mathbf{q}}|_\infty} \le \Pi_+(\mathbf{q})^{-(1+\varepsilon)}|q_0|^{-\varepsilon}\cdot\Big(1+n\|\mathbf{y}^{(\infty)}\|_\infty\Big)|q_0|^{-\varepsilon}$$
$$\le \Big(1+n\|\mathbf{y}^{(\infty)}\|_\infty\Big)\Pi_+(\mathbf{q})^{-(1+\varepsilon)}. \qquad \square$$

10.5. Finally we are ready for the

Proof of Theorem 8.7, the case $\infty\in S$. Applying the case "$\infty\notin S$" of Theorem 8.7 to the map $\big(\mathbf{f}^{(v)}\big)_{v\in S_f}$, we obtain that the pushforward of $\prod_{v\in S_f}\mu_v$ by $\big(\mathbf{f}^{(v)}\big)_{v\in S_f}$ must be strongly extremal, which, in view of Lemma 10.4, implies that $\mathbf{f}(x)$ is not VWMA$_>$ for μ-a.e. $x\in X$. Thus, as before, it suffices to take

$$B=\prod_{v\in S}B_v=\prod_{j=1}^{\ell-1}B_j\times B_\infty \quad\text{and}\quad \mu=\prod_{v\in S}\mu_v=\prod_{j=1}^{\ell-1}\mu_j\times\mu_\infty\,,$$

and check that the assumptions of Corollary 10.3 are satisfied by means of writing down explicit expressions for functions $x\mapsto \operatorname{cov}\big(g_{\mathbf{s},\tilde{\mathbf{t}}}u_{\mathbf{f}(x)}\Delta\big)$. This again boils down to the computation of components of $g_{\mathbf{s},\tilde{\mathbf{t}}}u_{\mathbf{f}(x)}\mathbf{w}$, where $\mathbf{w}=\sum_{I\subset\{0,\dots,n\}}w_I\mathbf{e}_I\in\bigwedge^r(\mathcal{D}^{n+1})$.

Since there was no change in the ultrametric components of $g_{\mathbf{s},\tilde{\mathbf{t}}}$ and $u_{\mathbf{f}(x)}$, formula (9.7) is still valid. Furthermore, an expression for the Archimedean components turns out to be similar to (9.7):

$$(g_{\mathbf{s},\tilde{\mathbf{t}}}u_{\mathbf{f}(x)}\mathbf{e}_I)^{(\infty)}=\begin{cases} e^{-t_I}\mathbf{e}_I^{(\infty)} \text{ if } 0\in I\\ e^{-t_I}\mathbf{e}_I^{(\infty)}+\sum_{i\in I}\pm e^{-t_{I\cup\{0\}\smallsetminus\{i\}}}f_i^{(\infty)}(x)\mathbf{e}^{(\infty)}_{I\cup\{0\}\smallsetminus\{i\}} \text{ otherwise}.\end{cases}$$

Therefore one has (9.8) for $j=1,\dots,\ell-1$, and, in addition,

$$(g_{\mathbf{s},\tilde{\mathbf{t}}}u_{\mathbf{f}(x)}\mathbf{w})^{(\infty)}=\sum_{0\notin I}w_Ie^{-t_I}\mathbf{e}_I^{(\infty)}+\sum_{0\in I}e^{-t_I}\Big(w_I+\sum_{i\notin I}\pm w_{I\cup\{i\}\smallsetminus\{0\}}f_i^{(\infty)}(x)\Big)\mathbf{e}_I^{(\infty)}\,.$$

We see that real (resp. p_j-adic) components of all the coordinates of $g_{\mathbf{s},\tilde{\mathbf{t}}}u_{\mathbf{f}(x)}\mathbf{w}$ are linear combinations of $1, f_1^{(\infty)},\dots,f_n^{(\infty)}$ (resp. $1,f_1^{(p_j)},\dots,f_n^{(p_j)}$). Condition (9.6a) then immediately follows from Lemma 2.1(bc), (8.6) and Corollary 2.3. On the other hand, an argument identical to that of the previous section shows that for every $j=1,\dots,\ell-1$ there exists $\rho_j>0$ such that (9.10) holds, and also that there exists $\rho_\infty>0$ such that

$$\sup_{x\in B_\infty\cap\operatorname{supp}\mu_\infty}\|(g_{\mathbf{s},\tilde{\mathbf{t}}}u_{\mathbf{f}(x)}\mathbf{w})^{(\infty)}\|_\infty\ge\rho_\infty\min_{0\in I}e^{-t_I}\max_I|w_I|_\infty\ge\rho_\infty e^{-\tilde{t}}\max_I|w_I|_\infty\,.$$

Therefore

$$\sup_{x\in B\cap\operatorname{supp}\mu}c(g_{\mathbf{s},\tilde{\mathbf{t}}}u_{\mathbf{f}(x)}\mathbf{w})\ge\left(\prod_{j=1}^{\ell-1}\rho_j{p_j}^{s_j}\max_I|w_I|_{p_j}\right)\rho_\infty e^{-\tilde{t}}\max_I|w_I|_\infty$$
$$\ge\rho_\infty\left(\prod_{j=1}^{\ell-1}\rho_j\right)e^{-\tilde{t}}\left(\prod_{j=1}^{\ell-1}{p_j}^{s_j}\right)\max_I c(w_I)\underset{(10.4\text{b})}{\ge}\rho_\infty\prod_{j=1}^{\ell-1}\frac{\rho_j}{p_j}\,.$$

This implies (9.6b) and shows that $\mathbf{f}(x)$ is not VWMA$_\le$ for μ-a.e. $x\in X$, thus finishing the proof of the theorem. $\square$

References

[B] A. Baker, *Transcendental number theory*, Cambridge Univ. Press, Cambridge, 1975.

[BD] V. Bernik and M. M. Dodson, *Metric Diophantine approximation on manifolds*, Cambridge Univ. Press, Cambridge, 1999.

[BKM] V. Bernik, D. Kleinbock and G. A. Margulis, *Khintchine-type theorems on manifolds: the convergence case for standard and multiplicative versions*, Internat. Math. Res. Notices (2001), no. 9, 453–486.

[Bo] A. Borel, *Linear Algebraic Groups*, Springer, Berlin/Heidelberg/New York, 1991.

[BP] A. Borel and G. Prasad,, *Values of isotropic quadratic forms at S-integral points*, Compositio Math. **83** (1992), 347–372.

[C] J. W. S. Cassels, *Local fields*, London Mathematical Society Student Texts, vol. 3, Cambridge University Press, Cambridge, 1986.

[CR] C. W. Curtis and I. Reiner, *Representation theory of finite groups and associative algebras*, Interscience Publishers, a division of John Wiley & Sons, New York, 1962.

[D1] S. G. Dani, *On invariant measures, minimal sets, and a lemma of Margulis*, Invent. Math. (1979), no. 51, 239–260.

[D2] ———, *On orbits of unipotent flows on homogeneous spaces*, Ergod. Th. Dynam. Sys. (1984), no. 4, 25–34.

[D3] ———, *On orbits of unipotent flows on homogeneous spaces, II*, Ergod. Th. Dynam. Sys. (1986), no. 6, 167–182.

[DM] S. G. Dani and G. A. Margulis, *Limit distributions of orbits of unipotent flows and values of quadratic forms*, in: Adv. in Soviet Math., vol. 16, Part 1, Amer. Math. Soc., Providence, RI, 1993, pp. 91–137.

[H] J. E. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. **30** (1981), 713–747.

[K1] D. Kleinbock, *Some applications of homogeneous dynamics to number theory*, in: Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999), Proc. Symp. Pure Math., vol. 68, Amer. Math. Soc., Providence, RI, 2001, pp. 639–660.

[K2] ———, *Extremal subspaces and their submanifolds*, Geom. Funct. Anal. **13** (2003), no. 2, 437–466.

[KM] D. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. Math. **148** (1998), 339–360.

[KLW] D. Kleinbock, E. Lindenstrauss, and B. Weiss, *On fractal measures and Diophantine approximation*, FIM preprint, ETH, Zurich (2003).

[KT] D. Kleinbock and G. Tomanov, *Diophantine approximation on ultrametric smooth manifolds*, in preparation.

[Kn] M. Kneser, *Starke Approximation in algebraischen Gruppen. I.*, J. Reine Angew. Math. **218** (1965), 190–203.

[L] S. Lang, *Algebraic Numbers*, Addison-Wesley, London, 1964.

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

[MT1] G. A. Margulis and G. Tomanov, *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, Invent. Math. **116** (1994), 347–392.

[MT2] ———, *Measure rigidity for almost linear groups and its applications*, J. Anal. Math. **69** (1996), 25–54.

[Mat] P. Mattila, *Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability*, Cambridge Studies in Advanced Mathematics, 44, Cambridge University Press, Cambridge, 1995.

[Rag] M. S. Raghunathan, *Discrete subgroups of Lie groups*, Springer-Verlag, Berlin and New York, 1972.

[Rat1] M. Ratner, *Raghunathan's topological conjecture and distributions of unipotent flows*, Duke Math. J. **63** (1991), 235–280.

[Rat2] ———, *Invariant measures and orbit closures for unipotent actions on homogeneous spaces*, Geom. Funct. Anal. **4** (1994), 236–257.

[Rat3] ———, *Raghunathan's conjectures for cartesian products of real and p-adic Lie groups,*, Duke Math. J. **77** (1995), 275–382.

[S] H. Sato, *Global density theorem for a Federer measure*, Tohoku Math. J. **44** (1992), no. 4, 581–595.

[Sp] V. G. Sprindžuk, *Mahler's problem in metric number theory*, Translations of Mathematical Monographs, Amer. Math. Soc., Providence, R.I., 1969.

[Tam] T. Tamagawa, *Arithmetic Algebraic Geometry*, Shilling, Harper and Row, New York, 1965.

[Tem] A. Tempelman, *Ergodic Theorems for Group Actions*, Kluwer Academic Publishers, Bordrecht/Boston/London, 1992.

[T] G. Tomanov, *Orbits on homogeneous spaces of arithmetic origin and approximations*, Advanced Studies in Pure Mathematics **26** (2000), 265–297.

[W] A. Weil, *Basic Number Theory*, Springer-Verlag, New York, 1995.

[Z] F. F. Želudevič, *Simultane diophantische Approximationen abhängiger Grössen in mehreren Metriken*, Acta Arith. **46** (1986), no. 3, 285–296. (German)

[Zi] R. Zimmer, *Ergodic theory and semisimple groups*, Birkhäuser, Boston, 1984.

Dmitry Kleinbock, Department of Mathematics, Brandeis University, Waltham, MA 02454-9110, USA

E-mail address: kleinboc@brandeis.edu

George Tomanov, Institut Girard Desargues, Université Claude Bernard – Lyon 1, 69622 Villeurbanne cedex, France

E-mail address: Georges.Tomanov@desargues.univ-lyon1.fr