Comment. Math. Helv. 82 (2007), 519–581 **Commentarii Mathematici Helvetici**
© Swiss Mathematical Society

Flows on *S*-arithmetic homogeneous spaces and applications to metric Diophantine approximation

Dmitry Kleinbock and George Tomanov

Abstract. The main goal of this work is to establish quantitative nondivergence estimates for flows on homogeneous spaces of products of real and p-adic Lie groups. These results have applications both to ergodic theory and to Diophantine approximation. Namely, earlier results of Dani (finiteness of locally finite ergodic unipotent-invariant measures on real homogeneous spaces) and Kleinbock–Margulis (strong extremality of nondegenerate submanifolds of $\mathbb{R}^n$) are generalized to the S-arithmetic setting.

Mathematics Subject Classification (2000). Primary 11J83; Secondary 11J54, 11J61, 37A45, 37D40.

Keywords. Lattices, homogeneous flows, quantitative nondivergence, ultrametric local fields, Diophantine approximation on manifolds, extremality and strong extremality of measures.

1. Introduction

1.1. Actions of unipotent subgroups on homogeneous spaces of real Lie groups provide examples of important dynamical systems with numerous applications to geometry and number theory. A fundamental phenomenon discovered by G. A. Margulis in 1971 [M] showed that, in sharp contrast to partially hyperbolic flows, orbits of unipotent flows are never divergent. Then in the papers [D1]–[D3] S. G. Dani generalized and strengthened this result. In particular, for one-parameter flows he proved that any unipotent 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]. In fact it was done in a bigger generality, i.e. for a large class of maps from $\mathbb{R}^d$ into $\mathrm{SL}(n,\mathbb{R})$, making it possible to derive important applications to metric Diophantine approximation on manifolds. The nondivergence theorem of Dani and Margulis was used in M. Ratner's proof [Rt1], [Rt2] of Raghunathan's topological conjecture, and in Dani's proof [D2] of finiteness of locally finite ergodic measures invariant under a unipotent flow.

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], [MT3] and [Rt3]–[Rt4]. In the present paper we prove an S-adic analogue of explicit quantitative estimates from [KM]. As one of the applications, we show the finiteness of locally finite unipotent-invariant ergodic measures on S-adic homogeneous spaces. We also derive applications to number theory. Namely, generalizing the correspondence between homogeneous dynamics and Diophantine approximation over $\mathbb{R}$, we prove p-adic, and more generally, S-arithmetic versions of conjectures of A. Baker and V. Sprindžuk, in particular answering a question posed by I. Shparlinski.

In order to state our results, let us introduce some notation. Let S be a finite set of normalized valuations of $\mathbb{Q}$ containing the Archimedean one, $\mathbb{Z}_S$ the ring of S-integers of $\mathbb{Q}$, and $\mathbb{Q}_S$ the direct product of completions $\mathbb{Q}_v$ of $\mathbb{Q}$ over $v \in S$. Put $\mathrm{GL}(m, \mathbb{Q}_S) \stackrel{\text{def}}{=} \prod_{v\in S} \mathrm{GL}(m, \mathbb{Q}_v)$ and

$$\mathrm{GL}^1(m, \mathbb{Q}_S) \stackrel{\text{def}}{=} \big\{(g^{(v)})_{v\in S} \in \mathrm{GL}(m, \mathbb{Q}_S) \mid \textstyle\prod_{v\in S} |\det(g^{(v)})|_v = 1\big\}.$$

Then one can interpret the homogeneous space

$$\Omega^1_{S,m} \stackrel{\text{def}}{=} \mathrm{GL}^1(m, \mathbb{Q}_S)/\mathrm{GL}(m, \mathbb{Z}_S)$$

as the space of lattices Λ in $\mathbb{Q}_S^m$ of covolume 1 (see §8.8 for details). For any $v \in S$, the valuation $|\cdot|_v$ induces the norm $\|\cdot\|_v$ on $\mathbb{Q}_v^m$, and we let

$$c(x) \stackrel{\text{def}}{=} \prod_{v\in S} |x^{(v)}|_v \quad \text{for } x = (x^{(v)})_{v\in S} \in \mathbb{Q}_S, \tag{1.1}$$

and $c(\boldsymbol{x}) \stackrel{\text{def}}{=} \prod_{v\in S} \|\boldsymbol{x}^{(v)}\|_v$ for $\boldsymbol{x} = (\boldsymbol{x}^{(v)})_{v\in S} \in \mathbb{Q}_S^m$. The latter function plays the same role as the usual norm in the case $S = \{\infty\}$, see Lemma 8.6.

Now define

$$Q_\varepsilon \stackrel{\text{def}}{=} \big\{\Lambda \in \Omega^1_{S,m} \mid c(x) \geq \varepsilon \text{ for all nonzero } x \in \Lambda\big\}. \tag{1.2}$$

It follows from the generalized Mahler's Compactness Criterion (Theorem 8.8) that the sets Q_ε are compact.

Let us now formulate a special case of our Theorem 9.4.

Theorem 1.1. *Consider the space $X = \prod_{v\in S} \mathbb{Q}_v^{d_v}$, $d_v \in \mathbb{N}$, endowed with the product metric* (*see* (2.6)) *and the measure*

$$\lambda \stackrel{\text{def}}{=} \prod_{v\in S} \lambda_v, \ \textit{where } \lambda_v \textit{ is a Haar measure on } \mathbb{Q}_v^{d_v}, \tag{1.3}$$

and let $h = (h_v)_{v\in S}\colon X \to \mathrm{GL}^1(m, \mathbb{Q}_S)$ *be a polynomial map* (*see* §9.4 *for the definition*) *with* $h(0) = e$. *Then there exists* $\alpha > 0$ *such that for every* $\Lambda \in \Omega^1_{S,m}$ *one can find* $C > 0$ *with the following property: for any positive* ε *and any ball* $B \subset X$ *centered at* 0 *one has*

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

We remark that Theorem 9.4 is in its turn a special case of a more general result, Theorem 9.3, where the restrictions on X, λ and h are significantly relaxed.

1.2. Theorem 9.4 makes it possible to generalize the main results from [D1]–[D3] to the case of products of real and p-adic algebraic groups. In particular, here is a special case of our Theorem 10.1:

Theorem 1.2. *Let* $G = \prod_{v\in S} G_v$, *where* G_v *is the group of* $\mathbb{Q}_v$*-rational points of an algebraic group* $\boldsymbol{G}_v$ *defined over* $\mathbb{Q}_v$, *let* Γ *be a lattice in* G, *and let* X *and* λ *be as in Theorem* 1.1. *Then for any* $d \in \mathbb{N}$ *and* $\beta > 0$ *there exists a compact* $M \subset G/\Gamma$ *such that for any* $y \in G/\Gamma$ *and any parametrization* $\phi\colon X \to U$ *of degree* $\le d$ *of a unipotent algebraic subgroup* U *of* G (*see* §10.1 *for precise definitions*) *one of the following holds: either*

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

$$\frac{\lambda\big(\{x \in B \mid \phi(x)y \in M\}\big)}{\lambda(B)} \ge 1 - \beta$$

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 1.2 has been announced, with indication of the proof, in [MT3, Theorem 11.4], [Rt4, Theorem 9.1] and [To, Theorem 3.3], and used for the proof of Ratner's uniform distribution theorem [Rt4, Theorem 3] and for other related results.

1.3. The next theorem, which follows from Theorem 10.1, has been proved for real Lie groups by Dani [D2, Theorem 4.3], and for the general case has been announced in [MT3, Theorem 11.5].

Theorem 1.3. *Let* G *and* Γ *be as in Theorem* 1.2, *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* $\mu(G/\Gamma \setminus \bigcup_i Y_i) = 0$. *In particular, every locally finite* H*-invariant ergodic measure on* G/Γ *is finite.*

1.4. Now let us turn to applications to number theory. Let S be a finite set consisting of distinct normalized valuations of $\mathbb{Q}$, with or without the infinite valuation. We will interpret elements

$$\boldsymbol{y} = (\boldsymbol{y}^{(v)})_{v\in S} = (y_1, \dots, y_n)$$

of $\mathbb{Q}_S^n = \prod_{v\in S} \mathbb{Q}_v^n$, where $\boldsymbol{y}^{(v)} = (y_1^{(v)}, \dots, y_n^{(v)}) \in \mathbb{Q}_v^n$ and $y_i = (y_i^{(v)}) \in \mathbb{Q}_S$, as linear forms on $\mathbb{Q}_S^n$, and will study their values $\boldsymbol{y}\cdot\boldsymbol{q} = y_1q_1 + \cdots + y_nq_n$ at integer points $\boldsymbol{q} = (q_1, \dots, q_n)$. Denote by ℓ the cardinality of S, and say that $\boldsymbol{y} \in \mathbb{Q}_S^n$ is *very well approximable*, or *VWA*, if for some $\varepsilon > 0$ there are infinitely many solutions $\tilde{\boldsymbol{q}} = (q_0, q_1, \dots, q_n) \in \mathbb{Z}^{n+1}$ to

$$|q_0 + \boldsymbol{q}\cdot\boldsymbol{y}|^\ell \le \begin{cases} \|\tilde{\boldsymbol{q}}\|^{-(n+1)(1+\varepsilon)} & \text{if } \infty \notin S, \\ \|\boldsymbol{q}\|^{-n(1+\varepsilon)} & \text{if } \infty \in S. \end{cases} \tag{1.4}$$

Here $\|\cdot\|$ is any norm on $\mathbb{R}^{n+1}$ (nothing will depend on the choice of the norm and we will use the Euclidean norm), and $|\cdot|$ is the *S-adic absolute value*, $|\boldsymbol{x}| = \max_{v\in S} |x^{(v)}|_v$. Also, say that $\boldsymbol{y}$ is *very well multiplicatively approximable*, or *VWMA*, if for some $\varepsilon > 0$ there are infinitely many $\tilde{\boldsymbol{q}} \in \mathbb{Z}^{n+1}$ such that

$$c(q_0 + \boldsymbol{q}\cdot\boldsymbol{y}) \le \begin{cases} \Pi_+(\tilde{\boldsymbol{q}})^{-(1+\varepsilon)} & \text{if } \infty \notin S, \\ \Pi_+(\boldsymbol{q})^{-(1+\varepsilon)} |q_0|_+^{-\varepsilon} & \text{if } \infty \in S, \end{cases} \tag{1.5}$$

where we put

$$|x|_+ \stackrel{\text{def}}{=} \max(|x|, 1) \ \text{ for } x \in \mathbb{R}, \quad \Pi_+(\boldsymbol{x}) \stackrel{\text{def}}{=} \prod_{i=1}^n |x_i|_+ \ \text{ for } \boldsymbol{x} \in \mathbb{R}^n,$$

and define $c(\cdot)$ as in (1.1). This unifies the standard definitions in the real and p-adic set-ups; we refer the reader to §11 for motivation, in particular for an explanation of the term $|q_0|^{-\varepsilon}$ in the second line of (1.5), and for a uniform way to write down the expressions on the right-hand sides of (1.4) and (1.5). It is easy to check using the Borel–Cantelli Lemma that the set of VWMA vectors, and hence the smaller set of VWA vectors, has zero Haar measure.

The subject of metric Diophantine approximation with dependent quantities originated with a conjecture of Mahler proved by Sprindžuk in the 1960s (see [Sp1]), stating that for almost every $x \in \mathbb{R}$, the vector

$$\boldsymbol{f}(x) = (x, x^2, \dots, x^n) \tag{1.6}$$

is not VWA. On the other hand, a similar statement with VWA replaced by VWMA, conjectured by A. Baker in 1972 [B], had not been proved until the paper [KM], which introduced a dynamical approach to this class of problems. (See also [K1] for

a survey.) Note that Sprindžuk had also proved the p-adic counterpart of Mahler's conjecture, and the problem of establishing its multiplicative version, that is, proving that the vector (1.6) is not VWMA for λ-a.e. $x \in \mathbb{Q}_p$, was recently posed by I. Shparlinski (V. Bernik, private communication).

More generally, following [KLW], let us say that a measure μ on $\mathbb{Q}_S^n$ is *extremal* (resp. *strongly extremal*) if μ-almost every vector in $\mathbb{Q}_S^n$ is not VWA (resp., not VWMA). In these terms, the conjectures of Mahler (resp. Baker) state that the pushforward $\boldsymbol{f}_*\lambda$ of Lebesgue measure λ on $\mathbb{R}$ by the map $\boldsymbol{f}$ as in (1.6) is extremal (resp. strongly extremal).

An important property of the curve (1.6) is that it does not belong to any proper affine subspace of $\mathbb{R}^n$. More generally, consider a C^k map $\boldsymbol{f}\colon U \to F^n$, where F is any locally compact valued field and U is an open subset of F^d, and say that $\boldsymbol{f}$ is *nondegenerate* at $\boldsymbol{x}_0 \in U$ if the space F^n is spanned by partial derivatives of $\boldsymbol{f}$ at $\boldsymbol{x}_0$ up to some finite order. (Note that the definition of C^k functions of an ultrametric variable is more involved than in the real case; see §4 for details.) One can view this condition as an infinitesimal version of not lying in any proper affine hyperplane, i.e. of the linear independence of $1, f_1, \dots, f_n$ over F (see §5 and [KM, §1] for further discussion).

It was conjectured by Sprindžuk in 1980 [Sp2, Conjecture H_2], and proved in [KM] in 1998 that $\boldsymbol{f}_*\lambda$ is strongly extremal for $\boldsymbol{f}\colon U \to \mathbb{R}^n$, $U \subset \mathbb{R}^d$, which is nondegenerate at λ-a.e. point of U. Much less has been known for other fields. For example, the extremality of $\boldsymbol{f}_*\lambda$ was shown by E. Kovalevskaya [Ko1], [Ko2] for $\boldsymbol{f}\colon \mathbb{Z}_p \to \mathbb{Z}_p^3$ which is normal in the sense of Mahler (a subclass of p-adic analytic functions) and nondegenerate λ-a.e.

In this paper we are able to prove much more general results. The following theorem, which we derive from Theorem 9.3, yields an S-arithmetic version of Sprindžuk's Conjecture H_2, in particular answering Shparlinski's question.

Theorem 1.4. *Let S be a finite set of normalized valuations of $\mathbb{Q}$, for any $v \in S$ take $k_v, d_v \in \mathbb{N}$ and an open subset $U_v \subset \mathbb{Q}_v^{d_v}$, and let λ be defined as in* (1.3)*. Suppose that $\boldsymbol{f}$ is of the form $(\boldsymbol{f}^{(v)})_{v\in S}$, where each $\boldsymbol{f}^{(v)}$ is a C^{k_v} map from U_v into $\mathbb{Q}_v^n$ which is nondegenerate at λ_v-a.e. point of U_v. Then $\boldsymbol{f}_*\lambda$ is strongly extremal.*

Note that the paper [Z] considers the case when each $\boldsymbol{f}^{(v)}$ is of the form (1.6), and proves the extremality of $\boldsymbol{f}_*\lambda$.

Theorem 1.4 is a special case of Theorem 11.1, which requires a certain terminology so we do not state it in the introduction. In fact, Theorem 11.1 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 particular, the so-called *self-similar open set condition* fractals (see §11.6 for details). Following [H], those can be considered in vector spaces over arbitrary locally compact fields; Theorem 11.1,

combined with certain results of [KLW], implies that natural measures supported on them are strongly extremal.

1.5. The structure of the paper is as follows. In §2 and §3 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. The most important examples of good functions are given by linear combinations of coordinate functions of smooth nondegenerate maps $F^d \to F^n$. This has been established in [KM, §3] for $F = \mathbb{R}$, and in §4 and §5 we develop a similar theory in the ultrametric setting. The key new ingredient of the proof (Proposition 4.2), which hinges on combinatorics of higher order difference quotients of C^k functions of an ultrametric variable, makes it possible to bypass the use of the Mean Value Theorem. Then in §6 and §7 we generalize a measure estimate of [KM, §4] to the setting of functions on a Besicovitch metric space X which are good with respect to a Federer measure on X. In §8 we prove auxiliary results about discrete $\mathbb{Z}_S$-submodules of $\mathbb{Q}_S^m$. The quantitative S-arithmetic nondivergence is discussed in §9 where, in particular, Theorem 1.1 is proved. §10 is devoted to proving Theorems 1.2 and 1.3. Then in §11 we turn to the S-arithmetic Diophantine approximation, giving all the definitions, stating the most general strong extremality result (Theorem 11.1), and mentioning applications to fractal measures. The proof of Theorem 11.1 breaks into two special cases (when S does or does not contain the Archimedean valuation), which are treated in §12 and §13 respectively. In both cases the argument is based on a modification of the dynamical approach to real Diophantine approximation as developed in [KM]. The last section of the paper lists several possible generalizations and open questions.

Acknowledgements. Part of the work was done during the authors' collaboration at the University of Lyon 1, Brandeis University and the Max Planck Institute; the hospitality of these institutions is gratefully acknowledged. A preliminary version of this paper [KT] appeared as a preprint of the MPI. 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, and to the referee for useful remarks. The first named author was supported in part by NSF Grants DMS-0072565 and DMS-0232463.

2. Besicovitch covering property

2.1. For a metric space X, $x \in X$ and $r > 0$, we denote by $B(x, r)$ the open ball $B(x, r) \stackrel{\text{def}}{=} \{y \in X \mid \operatorname{dist}(x, y) < r\}$ of radius r centered at x, and by $B(x, r^+)$ the closed ball $B(x, r^+) \stackrel{\text{def}}{=} \{y \in X \mid \operatorname{dist}(x, y) \le r\}$. (Note that $B(x, r^+)$ in general does not have to coincide with the closure $\overline{B(x, r)}$ of $B(x, r)$.) For a subset B of X and a

function $f: B \to F$, 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 $\|f\|_{B\cap \operatorname{supp}\mu}$, which, in case f is continuous and B is open, is the same as the $L^\infty(\mu)$ norm of $f|_B$, i.e.

$$\|f\|_{\mu,B} = \sup\big\{c \mid \mu(\{x \in B : |f(x)| > c\}) > 0\big\}.$$

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

$$x \text{ is the center of some ball of } \mathcal{B} \text{ for every } x \in A, \tag{2.1}$$

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

$$1_A \le \sum_i 1_{B_i} \le N_X, \tag{2.2}$$

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

Example 2.1. Suppose that X is *ultrametric*, that is, the non-Archimedean triangle inequality $\operatorname{dist}(x_1, x_2) \le \max_{i=1,2} \operatorname{dist}(x, x_i)$ 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).

Example 2.2. The fact that $\mathbb{R}^d$ is Besicovitch is the content of Besicovitch's Covering Theorem [Mt, Theorem 2.7]. In fact, Besicovitch's proof, see [Mt, pp. 29–34], can be easily generalized to give the following

Lemma 2.3. *For a metric space X, define*

$$\begin{aligned} M_X \stackrel{\text{def}}{=} \sup\big\{k \mid &\text{there are balls } B_i = B(x_i, r_i),\ 1 \le i \le k,\\ &\text{such that } \textstyle\bigcap_{i=1}^k B_i \ne \varnothing,\ \text{and } x_i \notin \bigcup_{\substack{j=1\\ j\ne i}}^k B_j \text{ for all } i\big\} \end{aligned} \tag{2.3}$$

and also, for $c > 1$,

$$\begin{aligned} D_X(c) \stackrel{\text{def}}{=} \sup\big\{k \mid &\text{there are } x \in X,\ r > 0 \text{ and pairwise disjoint}\\ &\text{balls } B_1, \dots, B_k \text{ of radius } r \text{ contained in } B(x, cr)\big\}. \end{aligned} \tag{2.4}$$

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 [Mt]. Take a bounded $A \subset X$ and a family $\mathcal{B}$ of nonempty open balls in X satisfying (2.1). 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 $r(x_1) \geq R_1/2$ and then inductively

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

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

$$R_2 = \sup \big\{ r(x) \mid x \in A \setminus \textstyle\bigcup_{i=1}^{k_1} B\big(x_i, r(x_i)\big) \big\}.$$

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

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

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 $2R_{i+1} \leq R_i$, 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

$$r(x) < R_j/2 \quad \text{for } j \in \mathbb{N},\ x \in A \setminus \bigcup_{i=1}^{k_j} B_i, \tag{2.5a}$$

$$R_j/2 \leq r(x_i) \leq R_i \quad \text{for } i \in I_j, \tag{2.5b}$$

$$x_{j+1} \in A \setminus \bigcup_{i=1}^{j} B_i \quad \text{for } j \in \mathbb{N}, \tag{2.5c}$$

$$x_k \in A \setminus \bigcup_{l \neq j} \bigcup_{i \in I_l} B_i \quad \text{for } j \in \mathbb{N},\ k \in I_j, \tag{2.5d}$$

$$A \subset \bigcup_{i=1}^{\infty} B_i. \tag{2.5e}$$

The first three properties follow immediately from the construction.

To prove (2.5d), take $j \in \mathbb{N}$, $k \in I_j$, $l \neq j$ and $i \in I_l$. If $l < j$, then $x_k \notin B_i$ by (2.5b). If $l > j$, then $x_i \notin B_k$ by (2.5b), and also $r(x_i) < r(x_k)$ by the construction,

hence $x_k \notin B_i$. Finally, to verify (2.5e), observe that since $R_j \to 0$ as $j \to \infty$, (2.5a) forces $r(x)$ to be equal to zero for any $x \in A \setminus \bigcup_{i=1}^{\infty} B_i$.

Clearly (2.5e) proves the first inequality in (2.2). 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 (2.5d) and (2.3), we see that the indices m_i can belong to at most M_X different blocks I_j. We now claim that for any $j \in \mathbb{N}$,

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

Indeed, fix $j \in \mathbb{N}$ and write $\{n_1, \dots, n_l\} \stackrel{\text{def}}{=} I_j \cap \{m_i \mid i = 1, \dots, k\}$. By (2.5b) and (2.5c), the balls $B\big(x_{n_i}, R_j/4\big)\big), i = 1, \dots, l$, are disjoint and contained in $B(x, 2R_j)\big)$, so the claim follows from (2.4). This proves that X is Besicovitch, with $N_X \le M_X D_X(8)$. □

2.2. 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 metric spaces (X_i, dist_i), $i = 1, \dots, k$, will always be supplied with the product metric

$$\mathrm{dist}\,\big((x_1, \dots, x_k), (y_1, \dots, y_k)\big) = \max_i \big(\mathrm{dist}_i(x_i, y_i)\big), \tag{2.6}$$

so that balls in $\prod_i X_i$ are products of balls in X_i. In particular, this convention forces the product of ultrametric spaces to be ultrametric as well.

Lemma 2.4. *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 $\bigcap_{i=1}^{k} B_i \ne \varnothing$ and x_i is not in $\bigcup_{j=1,\, j\ne i}^{k} B_j$ 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 $\bigcap_{i=1}^{k} F_i \ne \varnothing$, one has $F_k \subset F_i$ for all i. On the other hand, since $M_X < k$, the center of E_k must lie in E_i for some $i < k$, therefore x_k must lie in B_i, a contradiction.

The converse inequality is straightforward (and not needed for our purposes). □

Corollary 2.5. *If Y is ultrametric, one has*

$$N_{X\times Y} \le M_X D_X(8) D_Y(8); \tag{2.7}$$

in particular, $X \times Y$ is Besicovitch if the three constants on the right-hand side of (2.7) *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. □

Example 2.6. Let F be a locally compact field with a nontrivial 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 $\mathcal{O} = \mathbb{Z}_p$, $\mathcal{P} = p\mathbb{Z}_p$, and $|\cdot| = |\cdot|_p$, the standard p-adic valuation.) Then it is easy to see that for any $c \ge 1$, 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 2.5, in particular, implies that the metric space

$$X = \mathbb{R}^{d_0} \times F_1^{d_1} \times \cdots \times F_\ell^{d_\ell} \tag{2.8}$$

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

2.3. 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

$$\sup_{r>0} \frac{\mu\big(B(x, 3r)\big)}{\mu\big(B(x, r)\big)} < D \quad \text{for all } x \in \operatorname{supp}\mu. \tag{2.9}$$

Equivalently, one can replace "3" in (2.9) 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 $D_X(c) \le D_\mu(2c)$, or even $\le D_\mu(c)$ if X is ultrametric.

It is often 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 2.7. Let X be as in (2.8), and denote by p_i the number of elements in the residue class field of F_i, $i = 1, \dots, \ell$. It is clear that any Haar measure λ on X is uniformly Federer, with $D_\lambda(c) \le c^{d_0} \prod_{i=1}^{\ell} (cp_i)^{d_i}$.

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

3. (C, α)-good functions

3.1. Roughly speaking, a function is said to be *good* if the set of points where it takes small values has small measure. To simplify notation, it will be convenient to introduce a special symbol for a set of points x in a set B such that the value of a function f at x has norm less than ε. Namely, define

$$B^{f,\varepsilon} \stackrel{\text{def}}{=} \{x \in B \mid |f(x)| < \varepsilon\}$$

for any $f\colon B \to F$, where $(F, |\cdot|)$ is a valued field.

Now let X be a metric space and μ a Borel measure on X. For a subset U of X and $C, \alpha > 0$, say that a Borel measurable function $f\colon U \to F$ 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

$$\mu\big(B^{f,\varepsilon}\big) \le C\left(\frac{\varepsilon}{\|f\|_{\mu,B}}\right)^{\alpha} \mu(B) \quad \text{for all } \varepsilon > 0. \tag{3.1}$$

In all the applications of our results, the metric space X will be the normed ring as in (2.8), and μ will be chosen to be a Haar measure λ on X, in which case we will omit the reference to the measure and will simply say that the functions are (C, α)-good on U. In particular, μ will be positive on open sets, so it will be always possible to replace $\|f\|_{\mu,B}$ in (3.1) 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 3.1. *Let a metric space X, a measure μ on X, a subset U of X and $C, \alpha > 0$ be given.*

(a) *f is (C, α)-good on U with respect to $\mu \Leftrightarrow$ so is $|f|$;*

(b) *f is (C, α)-good on U with respect to $\mu \Rightarrow$ so is cf for all $c \in F$;*

(c) *f_i, $i \in I$, are (C, α)-good on U with respect to μ and the function $f = \sup_{i\in I} |f_i|$ is Borel measurable $\Rightarrow$ f is also (C, α)-good on U with respect to μ;*

(d) *f is (C, α)-good on U with respect to μ, and $c_1 \le \frac{|f(x)|}{|h(x)|} \le c_2$ for all $x \in U \Rightarrow$ h is $(C(c_2/c_1)^{\alpha}, \alpha)$-good on U with respect to μ.*

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

Lemma 3.2. *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*

$$\begin{gathered}\text{for all } y \in V \cap \operatorname{supp}\nu, \text{ the function } x \mapsto f(x,y)\\ \text{is } (C,\alpha)\text{-good on } U \text{ with respect to } \mu,\end{gathered} \tag{3.2a}$$

and

$$\begin{gathered}\text{for all } x \in U \cap \operatorname{supp}\mu, \text{ the function } y \mapsto f(x,y)\\ \text{is } (D,\beta)\text{-good on } V \text{ with respect to } \nu.\end{gathered} \tag{3.2b}$$

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

$$\gamma = \frac{\alpha\beta}{\alpha+\beta} \quad \text{and} \quad E = (\alpha+\beta)\left(\left(\frac{C}{\beta}\right)^{\beta}\left(\frac{D}{\alpha}\right)^{\alpha}\right)^{\frac{1}{\alpha+\beta}}. \tag{3.3}$$

Proof. Fix a ball in $U \times V$ of the form $A \times B$, where A and B are balls in X and Y intersecting the supports of μ and ν respectively. Without loss of generality let us rescale $\mu|_A$, $\nu|_B$ and f so that $\mu(A) = \nu(B) = \|f\|_{\mu\times\nu, A\times B} = 1$. Take an arbitrary $\varepsilon > 0$; we need to demonstrate that

$$(\mu \times \nu)\big((A \times B)^{f,\varepsilon}\big) \le E\varepsilon^{\gamma}. \tag{3.4}$$

For $y \in B$ let us denote by f_y the function $x \mapsto f(x,y)$. Also denote by φ the function defined on B by $\varphi(y) \stackrel{\text{def}}{=} \|f_y\|_{\mu,A}$; note that φ is Borel measurable, and $\|\varphi\|_{\nu,B} = 1$ because of the assumed normalization of f. In view of (3.2a), for any $y \in B \cap \operatorname{supp}\nu$ one has

$$\mu(A^{f_y,\varepsilon}) \le C\left(\frac{\varepsilon}{\varphi(y)}\right)^{\alpha} \iff \varphi(y) \le \left(\frac{C}{\mu(A^{f_y,\varepsilon})}\right)^{1/\alpha}\varepsilon. \tag{3.5}$$

Take an arbitrary $t > 0$ (to be fixed later), and denote

$$B_t \stackrel{\text{def}}{=} \{y \in B \mid \mu(A^{f_y,\varepsilon}) \ge t\}.$$

In view of (3.5), $y \in B_t$ implies that $\varphi(y)$ is not bigger than $(C/t)^{1/\alpha}\,\varepsilon$. Since it follows from Lemma 3.1(c) and (3.2b) that φ is (D,β)-good on V with respect to ν, one can write

$$\nu(B_t) \le \nu\big(\{y \in B \mid \varphi(y) \le (C/t)^{1/\alpha}\,\varepsilon\}\big) \le D\left((C/t)^{1/\alpha}\,\varepsilon\right)^{\beta}. \tag{3.6}$$

Now observe that one has $\mu\big(\{x \in A \mid (x,y) \in (A\times B)^{f,\varepsilon}\}\big) < t$ whenever $y \notin B_t$, therefore, by Fubini,

$$\begin{aligned}(\mu \times \nu)\big((A \times B)^{f,\varepsilon}\big) &< (\mu \times \nu)\big(A \times B_t\big) + t \cdot \nu(B \setminus B_t)\\ &\le \nu(B_t) + t \underset{(3.6)}{\le} t + (DC^{\beta/\alpha}\varepsilon^{\beta}) \cdot t^{-\beta/\alpha}.\end{aligned} \tag{3.7}$$

The function on the right-hand side of (3.7) attains its minimum when

$$t = \left(DC^{\beta/\alpha}\varepsilon^{\beta}\tfrac{\beta}{\alpha}\right)^{\frac{\alpha}{\alpha+\beta}} = \left(C^{\beta}\left(\tfrac{D\beta}{\alpha}\right)^{\alpha}\right)^{\frac{1}{\alpha+\beta}}\varepsilon^{\frac{\alpha\beta}{\alpha+\beta}};$$

substituting it into (3.7), one easily obtains (3.4) with E and γ given by (3.3). □

3.3. Applying the above lemma repeatedly, one easily obtains

Corollary 3.3. *For $j = 1, \dots, d$, let X_j be a metric space, μ_j a measure on X_j, $U_j \subset X_j$ open, $C_j, \alpha_j > 0$, and let f be a function on $U_1 \times \cdots \times U_d$ such that for any $j = 1, \dots, d$ and any $x_i \in U_i$ with $i \neq j$, the function*

$$y \mapsto f(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_d) \tag{3.8}$$

is (C_j, α_j)-good on U_j with respect to μ_j. Then

$$f \text{ is } (\tilde{C}, \tilde{\alpha})\text{-good on } U_1 \times \cdots \times U_d \text{ with respect to } \mu_1 \times \cdots \times \mu_d, \tag{3.9}$$

where $\tilde{C}$ and $\tilde{\alpha}$ are explicitly computable in terms of C_j, α_j. In particular, if each of the functions (3.8) *is (C, α)-good on U_j with respect to μ_j,* (3.9) *holds with*

$$\tilde{\alpha} = \alpha/d \quad \text{and} \quad \tilde{C} = dC.$$

3.4. 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 $\leq k$ 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 [To], one can consider polynomials over other locally compact fields:

Lemma 3.4. *Let F be either $\mathbb{R}$ or 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 $(C, 1/dk)$-good on F^d with respect to λ, where C is a constant depending only on d and k.*

Proof. For ultrametric F the case $d = 1$ is proved in [To, Lemma 4.1], and the general case immediately follows from the one-dimensional case and Corollary 3.3. Likewise, one can use [KM, Lemma 3.2] and Corollary 3.3 to establish the claim for real polynomials of several variables. □

3.5. Another result of [KM], which can be thought of as a generalization of the real case of the previous lemma, is that, roughly speaking, a smooth real-valued function is good on an open subset of $\mathbb{R}^d$ (with respect to Lebesgue measure λ) provided that its partial derivatives of some order do not vanish.

Lemma 3.5 ([KM, Lemma 3.3]). *Given $k \in \mathbb{N}$ and an open subset V of $\mathbb{R}^d$, let $f \in C^k(V)$ be such that for some constants $0 < a \le A$ one has*[2]

$$a \le |\partial_i^k f(\boldsymbol{x})| \le A \quad \text{for all } \boldsymbol{x} \in V,\ i = 1, \dots, d. \tag{3.10}$$

Then f is $\big(dC_k\,(A/a)^{1/k}\,,\,1/dk\big)$-good on V, where $\mathbb{R}^d$ is understood to be equipped with the l^∞ metric (induced by the norm $\|\boldsymbol{x}\| = \max_{i=1}^d |x_i|$, i.e. the product metric on $\mathbb{R} \times \cdots \times \mathbb{R}$ in the sense of (2.6)*), and C_k is a constant dependent on k only (and explicitly estimated in* [KM]).

Our goal in the next section is to describe the class of ultrametric C^k functions and prove a non-Archimedean analogue (Theorem 4.1) of Lemma 3.5.

4. Ultrametric C^k functions

4.1. In this section we state and prove the ultrametric analogue of the "$d = 1$" case of Lemma 3.5. We start by introducing certain terminology, most of which is borrowed from [Sc]. Here and until the end of the section F is a complete field with a nontrivial ultrametric valuation $|\cdot|$, and f an F-valued function on a subset U of F without isolated points. The *first difference quotient* $\Phi^1 f$ of f is the function of two variables given by

$$\Phi^1 f(x, y) \stackrel{\text{def}}{=} \frac{f(x) - f(y)}{x - y} \quad (x, y \in U,\ x \ne y),$$

defined on

$$\nabla^2 U \stackrel{\text{def}}{=} \{(x, y) \in U \times U \mid x \ne y\}.$$

We say that f is C^1 *at* a if the limit

$$\lim_{(x,y)\to(a,a)} \Phi^1 f(x, y)$$

exists, and that $f \in C^1(U)$ if f is C^1 at every point of U.

More generally, for $k \in \mathbb{N}$ set

$$\nabla^k U \stackrel{\text{def}}{=} \{(x_1, \dots, x_k) \in U^k \mid x_i \ne x_j \text{ for } i \ne j\},$$

and define the *k-th order difference quotient* $\Phi^k f \colon \nabla^{k+1} U \to F$ of f inductively by $\Phi^0 f \stackrel{\text{def}}{=} f$ and

$$\Phi^k f(x_1, x_2, \dots, x_{k+1}) \stackrel{\text{def}}{=} \frac{\Phi^{k-1}(x_1, x_3, \dots, x_{k+1}) - \Phi^{k-1}(x_2, x_3, \dots, x_{k+1})}{x_1 - x_2}.$$

[2]The upper estimate in [KM] was stronger than stated here, but in fact our weaker condition suffices for the proof.

Note that one could equivalently take any pair of variables in place of (x_1, x_2), and that $\Phi^k f$ is a symmetric function of its $k+1$ variables. Then say that f is C^k at a if the limit

$$\lim_{(x_1,\dots,x_{k+1})\to(a,\dots,a)} \Phi^k f(x_1,\dots,x_{k+1})$$

exists, and that $f \in C^k(U)$ if f is C^k at every point of U. The latter is equivalent to $\Phi^k f$ being extendable to a continuous function $\bar{\Phi}^k f\colon U^{k+1} \to F$. Note that $\nabla^{k+1}U$ is dense in U^{k+1} if U has no isolated points, so the extension is unique if it exists. We refer the reader to [Sc, §27–29] for basic facts about C^k functions. For instance, one can show that C^k functions f are k times differentiable, and in fact

$$f^{(k)}(x) = k!\bar{\Phi}^k(x,\dots,x). \tag{4.1}$$

In particular, $f \in C^k$ implies that $f^{(k)}$ is continuous. However the converse is not true, see [Sc, §27, Remark 1] for a counterexample. On the other hand, locally analytic functions are C^k for every k.

The definition of C^k functions of several ultrametric variables is a straighforward generalization of the one for single-variable functions. If f is an F-valued function on $U_1 \times \cdots \times U_d$, where each U_i is a subset of F without isolated points, let us denote by $\Phi_i^k f$ the kth order difference quotient of f with respect to the variable x_i, and, more generally, for a *multiindex* $\beta = (i_1, \dots, i_d)$ let

$$\Phi_\beta f \stackrel{\text{def}}{=} \Phi_1^{i_1} \circ \cdots \circ \Phi_d^{i_d} f,$$

where it is not hard to check that the composition can be taken in any order. The latter "difference quotient of order β" is defined on $\nabla^{i_1}U_1 \times \cdots \times \nabla^{i_d}U_d$, and as before we say that f belongs to $C^k(U_1 \times \cdots \times U_d)$ if for any multiindex β with $|\beta| \stackrel{\text{def}}{=} \sum_{j=1}^d i_j$ at most k, $\Phi_\beta f$ is extendable to a continuous function $\bar{\Phi}_\beta f\colon U_1^{i_1+1} \times \cdots \times U_d^{i_d+1} \to F$. As in the one-variable case, one can show that partial derivatives $\partial_\beta f \stackrel{\text{def}}{=} \partial_1^{i_1} \circ \cdots \circ \partial_d^{i_d} f$ of a C^k function f exist and are continuous as long as $|\beta| \le k$. Moreover, one has

$$\partial_\beta f(x_1,\dots,x_d) = \beta!\bar{\Phi}_\beta(x_1,\dots,x_1,\dots,x_d,\dots,x_d). \tag{4.2}$$

where $\beta! \stackrel{\text{def}}{=} \prod_{j=1}^d i_j!$, and each of the variables x_j on the right-hand side of (4.2) is repeated $i_j + 1$ times.

4.2. An elementary observation, which will be repeatedly used, is that if a function $f\colon U \to F$, where U is an open subset of F^d, is continuous at $\boldsymbol{x}_0 \in U$ and $f(\boldsymbol{x}_0) \neq 0$, then there exists a neighborhood V of $\boldsymbol{x}_0$ such that $|f(\boldsymbol{x})| = |f(\boldsymbol{x}_0)|$ for all $\boldsymbol{x} \in V$. Thus, a natural ultrametric replacement for inequalities of type (3.10) would be assuming that the absolute value of certain difference quotients of f is identically equal to some $A > 0$ on some open set.

With this in mind, let us state an ultrametric analogue of Lemma 3.5.

Theorem 4.1. *Let $V_1, \dots, V_d$ be nonempty open sets in F, and let $A_1, \dots, A_d > 0$, $k \in \mathbb{N}$ and $f \in C^k(V_1 \times \cdots \times V_d)$ be such that*

$$|\Phi_j^k f| \equiv A_j \quad \text{on } \nabla^{k+1} V_j \times \prod_{i \neq j} V_i, \quad j = 1, \dots, d. \tag{4.3}$$

Then f is $(dk^{3-1/k}, 1/dk)$-good on $V_1 \times \cdots \times V_d$.

One can immediately observe that (4.3) amounts to saying that the absolute value of the kth order difference quotient of each of the one-variable functions (3.8), $j = 1, \dots, d$, is equal to A_j on $\nabla^{k+1} V_i$. Therefore one can use Corollary 3.3 to easily derive the above theorem from its one-dimensional case. In other words, it suffices to take an open subset V of F, let $k \in \mathbb{N}$, $A > 0$ and $f \in C^k(V)$ be such that

$$|\Phi^k f(x_1, x_2, \dots, x_{k+1})| = A \quad \text{for all } (x_1, x_2, \dots, x_{k+1}) \in \nabla^{k+1} V, \tag{4.4}$$

and prove that f is $\big(k^{3-1/k}, 1/k\big)$-good on V.

The strategy of the proof will be similar to the one used in [KM] to prove the one-dimensional case of Lemma 3.5. However, we need to pay special attention to the following implication of (4.4) which one gets for free in a similar situation when $F = \mathbb{R}$:

Proposition 4.2. *Let V be a ball in F, and let $k \in \mathbb{N}$, $A > 0$ and $f : V \to F$ be such that* (4.4) *holds. Then for any $\varepsilon > 0$, the set $V^{f,\varepsilon}$ is a disjoint union of at most k balls.*

If in addition one assumes that $f \in C^k(V)$, (4.4), in view of (4.1), would imply that the absolute value of $f^{(k)}(x)$ for $x \in V$ is a nonzero constant. Note that nonvanishing of the kth derivative of a real function f on an interval $V \subset \mathbb{R}$ immediately implies, due to the Mean Value Theorem, that $V^{f,\varepsilon}$ consists of at most k intervals. Unfortunately such a theorem is not present in the ultrametric calculus, so one has to look for alternative approaches.

4.3. To prove the proposition, we will need the following auxiliary lemma:

Lemma 4.3. *Let V be an open subset of F, f a function $V \to F$, $k \geq 2$, and let $x_1, \dots, x_k, y \in V$ be pairwise different. Also assume that*

$$|y - x_k| \leq |x_i - x_k| \quad \textit{for all } i < k, \tag{4.5a}$$

$$|\Phi^{k-i} f(x_i, \dots, x_k)| \geq |\Phi^{k-i} f(x_{i-1}, \dots, x_{k-1})| \quad \textit{for all } i = 2, \dots, k, \tag{4.5b}$$

and

$$|\Phi^{k-1} f(x_1, \dots, x_k)| \geq |\Phi^{k-1} f(x_1, \dots, x_{k-1}, y)|. \tag{4.5c}$$

Then

$$|f(y)| \le \max\big(|f(x_k)|, |f(x_{k-1})|\big). \tag{4.5d}$$

Proof. Note that (4.5a) implies that

$$|y - x_1| = |y - x_k + x_k - x_1| \le |x_k - x_1|,$$

and from (4.5c) one gets

$$\begin{aligned} |\Phi^{k-2} f(x_2, \dots, x_k) - \Phi^{k-2} f(x_1, \dots, x_{k-1})| &= |x_k - x_1| \cdot |\Phi^{k-1} f(x_1, \dots, x_k)| \\ &\ge |x_k - x_1| \cdot |\Phi^{k-1} f(x_1, \dots, x_{k-1}, y)| \ge |y - x_1| \cdot |\Phi^{k-1} f(x_1, \dots, x_{k-1}, y)| \\ &= |\Phi^{k-2} f(x_2, \dots, x_{k-1}, y) - \Phi^{k-2} f(x_1, \dots, x_{k-1})|. \end{aligned} \tag{4.6}$$

Now let us use induction on k. If $k = 2$, (4.6) says that $|f(x_2) - f(x_1)| \ge |f(y) - f(x_1)|$, which readily implies that $|f(y)| \le \max\big(|f(x_1)|, |f(x_2)|\big)$. If $k > 2$ and the claim is true with k replaced by $k - 1$, observe that (4.6) and the "$i = 2$" case of (4.5b) imply that $|\Phi^{k-2} f(x_2, \dots, x_k)| \ge |\Phi^{k-2} f(x_2, \dots, x_{k-1}, y)|$. Therefore the lemma can be applied to $x_2, \dots, x_k, y$, and (4.5d) follows. □

4.4. *Proof of Proposition* 4.3. Replacing f by f/A without loss of generality we may, and will, assume that $A = 1$.

Let $\varepsilon > 0$. Note that it follows from the discreteness of the valuation that $V^{f,\varepsilon}$ is the union of finitely many balls. Assume, by contradiction, that $V^{f,\varepsilon} = \cup_{i=1}^{n} B_i$, where $n \ge k + 1$ and B_i are different components of $V^{f,\varepsilon}$. There exist $x_1, \dots, x_k \in V^{f,\varepsilon}$ such that each x_i belongs to a different component (which after changing the indices we can assume to be B_i) and

$$|\Phi^{k-1} f(x_1, \dots, x_k)| = \sup_{\substack{y_i \in B_{\ell(i)} \\ i \ne j \Rightarrow \ell(i) \ne \ell(j)}} |\Phi^{k-1} f(y_1, \dots, y_k)|.$$

Next we rearrange $x_1, \dots, x_k$ in such a way that for all $\ell = 1, \dots, k - 2$ one has

$$|\Phi^{\ell} f(x_{k-\ell}, \dots, x_k)| \ge |\Phi^{\ell} f(x_{k-\ell-1}, \dots, \check{x}_i, \dots, x_k)| \quad \text{for all } i = k-\ell-1, \dots, k, \tag{4.7a}$$

where $\check{x}_i$ means that the term x_i is missing.

Denote

$$R \stackrel{\text{def}}{=} \min\big(|\Phi^{k-1} f(x_1, \dots, x_k)|, |x_k - x_1|, \dots, |x_k - x_{k-1}|\big), \tag{4.7b}$$

and take $y \in B(x_k, R^+)$. Then, using (4.4) and $A = 1$, one writes

$$|y - x_k| = |\Phi^{k-1} f(x_1, \dots, x_k) - \Phi^{k-1} f(x_1, \dots, x_{k-1}, y)| \le |\Phi^{k-1} f(x_1, \dots, x_k)|. \tag{4.7c}$$

It follows from (4.7a)–(4.7c) that conditions (4.5a)–(4.5c) hold, therefore, by Lemma 4.3,

$$\varepsilon > \max\big(|f(x_k)|, |f(x_{k-1})|\big) \geq |f(y)|.$$

This proves that $B_k \supset B(x_k, R^+)$, and from the fact that balls B_i are disjoint it follows that $|x_k - x_i| \geq |\Phi^{k-1} f(x_1, \dots, x_k)|$ for all $i \neq k$, hence $R = |\Phi^{k-1} f(x_1, \dots, x_k)|$.

Now let $y \in B_{k+1}$. By the choice of $x_1, \dots, x_k$ one has

$$|\Phi^{k-1} f(x_1, \dots, x_k)| \geq |\Phi^{k-1} f(x_1, \dots, x_{k-1}, y)|,$$

hence, again by (4.4),

$$\begin{aligned} |y - x_k| &= |\Phi^{k-1} f(x_1, \dots, x_k) - \Phi^{k-1} f(x_1, \dots, x_{k-1}, y)| \\ &\leq |\Phi^{k-1} f(x_1, \dots, x_k)| = R. \end{aligned}$$

Consequently $x \in B_k$, which is a contradiction. □

4.5. Now we can proceed with the

Proof of Theorem 4.2. We need to show that for any open ball $B \subset V$ one has

$$\lambda\big(B^{f,\varepsilon}\big) \leq k^{3-1/k} \left(\frac{\varepsilon}{\|f\|_B}\right)^{1/k} \lambda(B) \quad \text{for all } \varepsilon > 0$$

whenever V is an open subset of f and $f \in C^k(V)$ satisfies (4.4).

It is clear that the result does not depend on the normalization of λ, and it will be convenient to assume $\lambda(\mathcal{O}) = 1$, so that $\lambda(J) = \operatorname{diam}(J)$ for any ball J. In view of Proposition 4.2, it suffices to show that for any ball $J \subset B$ with $\|f\|_J < \varepsilon$ one has

$$\lambda(J) \leq k^{2-1/k} \left(\frac{\varepsilon}{\|f\|_B}\right)^{1/k} \lambda(B). \tag{4.8}$$

Also, as in the proof of Proposition 4.2, let us replace f by f/A and thus assume $A = 1$. Note however that now we have in addition assumed that $f \in C^k(V)$, therefore (4.4) implies that

$$|\bar{\Phi}^k f(x_1, x_2, \dots, x_{k+1})| = 1 \quad \text{for all } x_1, x_2, \dots, x_{k+1} \in V. \tag{4.9}$$

It is easy to see that one can choose $x_1, \dots, x_{k+1} \in J$ such that

$$|x_i - x_j| \geq \lambda(J)/k \quad \text{for } i \neq j. \tag{4.10}$$

After that let P be the Lagrange polynomial of degree k formed by using values of f at these points, i.e. given by

$$P(x) = \sum_{i=1}^{k+1} f(x_i) \frac{\prod_{j=1,\, j\neq i}^{k+1} (x - x_j)}{\prod_{j=1,\, j\neq i}^{k+1} (x_i - x_j)}.$$

Then we have $\Phi^k(f - P)(x_1, \dots, x_{k+1}) = 0$, that is,

$$\Phi^k f(x_1, \dots, x_{k+1}) = \Phi^k P(x_1, \dots, x_{k+1}) = \text{the leading coefficient of } P$$
$$= \sum_{i=1}^{k+1} f(x_i)\Big(\prod_{j=1,\, j\neq i}^{k+1} (x_i - x_j)\Big)^{-1}.$$

Taking absolute values, one obtains

$$\begin{aligned} 1 &= |\Phi^k f(x_1, \dots, x_{k+1})| = |\Phi^k P(x_1, \dots, x_{k+1})| \\ &\le \max_i \Big| f(x_i)\Big(\prod_{j=1,\, j\neq i}^{k+1} (x_i - x_j)\Big)^{-1} \Big| \underset{(4.10) \text{ and } \|f\|_J < \varepsilon}{<} \varepsilon \frac{k^k}{\lambda(J)^k}. \end{aligned} \tag{4.11}$$

Next, take any $y \in J$ and let Q be the Taylor polynomial of f at y of degree $k-1$. By Taylor's formula [Sc, Theorem 29.4], for any x one has

$$f(x) = Q(x) + (x - y)^k \bar{\Phi}^k f(x, y, \dots, y), \tag{4.12}$$

hence

$$\|f - Q\|_J \le \lambda(J)^k \|\bar{\Phi}^k\|_J \underset{(4.9)}{=} \lambda(J)^k \underset{(4.11)}{\le} k^k \varepsilon.$$

This implies

$$\|Q\|_J \le \max\big(\|f\|_J, \lambda(J)^k\big) < \max\big(\varepsilon, k^k\varepsilon\big) = k^k\varepsilon. \tag{4.13}$$

Now let us apply Lagrange's formula to reconstruct Q on B by its values at $x_1, \dots, x_k$. Namely, for $x \in B$ write

$$\begin{aligned} |Q(x)| &= \Big| \sum_{i=1}^{k} Q(x_i) \frac{\prod_{j=1,\, j\neq i}^{k} (x - x_j)}{\prod_{j=1,\, j\neq i}^{k} (x_i - x_j)} \Big| \\ &\underset{(4.13),\,(4.10)}{<} k^k \varepsilon \lambda(B)^{k-1} \frac{k^{k-1}}{\lambda(J)^{k-1}} \le k^{2k-1} \varepsilon \Big(\frac{\lambda(B)}{\lambda(J)}\Big)^k. \end{aligned} \tag{4.14}$$

Finally, the difference between f and Q on B is, again in view of (4.12) and (4.9), bounded from above by $\lambda(B)^k$, hence

$$\begin{aligned} \|f\|_B &\le \max\big(\|Q\|_B, \lambda(B)^k\big) \underset{(4.14)}{<} \max\left(k^{2k-1}\varepsilon\Big(\frac{\lambda(B)}{\lambda(J)}\Big)^k, \lambda(B)^k \right) \\ &= \lambda(B)^k \max\left(\varepsilon \frac{k^{2k-1}}{\lambda(J)^k}, 1 \right) \underset{(4.11)}{<} k^{2k-1}\varepsilon\Big(\frac{\lambda(B)}{\lambda(J)}\Big)^k, \end{aligned}$$

which is equivalent to (4.8). □

5. Nondegenerate, nonplanar and good maps

5.1. In this section we will consider vector-valued functions of an ultrametric variable. If $\boldsymbol{f} = (f_1, \dots, f_n)$ is a map from an open subset U of F^d into F^n, for any multiindex $\beta = (i_1, \dots, i_d)$ we let

$$\Phi_\beta \boldsymbol{f} \stackrel{\text{def}}{=} (\Phi_\beta f_1, \dots, \Phi_\beta f_n),$$

and say that $\boldsymbol{f}$ is C^k if so is each f_i. In the latter case one denotes by $\bar{\Phi}_\beta \boldsymbol{f}$ the continuous function extending $\Phi_\beta f$ to $U_1^{i_1+1} \times \cdots \times U_d^{i_d+1}$, so that (4.2) holds with f replaced by $\boldsymbol{f}$.

Let us now take F to be either $\mathbb{R}$ or an ultrametric valued field, and say that a map $\boldsymbol{f}\colon U \to F^n$, where U is an open subset of F^d, is *k-nondegenerate at $\boldsymbol{x}_0 \in U$* if it is C^k on a neighborhood of $\boldsymbol{x}_0$, and the space F^n is spanned by all the partial derivatives $\partial_\beta \boldsymbol{f}(\boldsymbol{x}_0)$ of $\boldsymbol{f}$ at $\boldsymbol{x}_0$ with $|\beta| \le k$. We will say that $\boldsymbol{f}$ is *nondegenerate at $\boldsymbol{x}_0$* if it is k-nondegenerate at $\boldsymbol{x}_0$ for some k. Another way of saying this is as follows: $\boldsymbol{f}$ is k-nondegenerate at $\boldsymbol{x}_0$ iff for any function f of the form $f = c_0 + \boldsymbol{c} \cdot \boldsymbol{f}$, where $c_0 \in F$ and $\boldsymbol{c} \in F^n \setminus \{0\}$ there exists a multiindex β with $|\beta| \le k$ such that $\partial_\beta f(\boldsymbol{x}_0) \neq 0$.

In particular, it follows from the nondegeneracy of $\boldsymbol{f}$ at $\boldsymbol{x}_0$ that for any neighborhood B of $\boldsymbol{x}_0$ the restrictions of $1, f_1, \dots, f_n$ to B are linearly independent over F; in other words, $\boldsymbol{f}(B)$ is not contained in any proper affine subspace of F^n. On the other hand, the converse is true under an additional assumption that $\boldsymbol{f}$ is analytic in a neighborhood of $\boldsymbol{x}_0$: indeed, if $\boldsymbol{f}$ can be written as a Taylor series in a neighborhood B of $\boldsymbol{x}_0$, and it is known that *all* partial derivatives of $\boldsymbol{f}$ at $\boldsymbol{x}_0$ belong to a proper subspace L of F^n, then $\boldsymbol{f}(B)$ must be contained in $L + \boldsymbol{x}_0$.

In more general situations it will be convenient to use the following terminology: if X is a metric space, μ a measure on X and $\boldsymbol{f} = (f_1, \dots, f_n)$ a map from X to F^n, the pair $(\boldsymbol{f}, \mu)$ will be called *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 $F \Leftrightarrow \boldsymbol{f}(B \cap \operatorname{supp} \mu)$ is not contained in any proper affine subspace of F^n. We will omit the dependence on the measure when it is taken to be Lebesgue or Haar and will simply say that $\boldsymbol{f}$ is nonplanar at x_0. Thus the above remark translates into saying that for a C^k (resp., analytic) function $\boldsymbol{f}\colon F^d \to F^n$, nondegeneracy implies (resp., is equivalent to) nonplanarity.

5.2. We are now going to discuss another property of $\boldsymbol{f}$ which will also be implied by nondegeneracy. Namely, if $(F, |\cdot|)$ is a valued field, X a metric space, μ a measure on X and $\boldsymbol{f}$ a map from X to F^n, let us say that $(\boldsymbol{f}, \mu)$ is *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 Lemma 3.4 that polynomial maps are good at every point. Similarly, in [KM] Lemma 3.5 was used to show that a smooth map $\boldsymbol{f}: \mathbb{R}^d \to \mathbb{R}^n$ is good at every point where it is nondegenerate.

Our goal in this section is to prove an ultrametric analogue of the aforementioned result, using Theorem 4.1 in place of Lemma 3.4. Namely, we have

Proposition 5.1. *Let F be an ultrametric valued field, and let $\boldsymbol{f} = (f_1, \dots, f_n)$ be a C^ℓ map from an open subset U of F^d to F^n which is ℓ-nondegenerate at $\boldsymbol{x}_0 \in U$. Then there exists a neighborhood $V \subset U$ of $\boldsymbol{x}_0$ such that any linear combination of $1, f_1, \dots, f_n$ is $(d\ell^{3-1/\ell}, 1/d\ell)$-good on V. In particular, the nondegeneracy of $\boldsymbol{f}$ at $\boldsymbol{x}_0$ implies that $\boldsymbol{f}$ is good at $\boldsymbol{x}_0$.*

Proof. Without loss of generality we can put $\boldsymbol{x}_0 = 0$, and consider the family of functions

$$\mathcal{H} \stackrel{\text{def}}{=} \big\{h = c_0 + \textstyle\sum_{i=1}^n c_i f_i \mid \max_{i=0,1,\dots,n} |c_i| = 1\big\}.$$

It is enough (see Lemma 3.1(b)) to find a neighborhood V of 0 in F^d such that any $h \in \mathcal{H}$ is $\big(d\ell^{3-1/k}, 1/d\ell\big)$-good on V.

From the nondegeneracy assumption it follows that for any $h \in \mathcal{H}$ one can find a multiindex β with

$$1 \le |\beta| = k \le \ell \quad \text{and} \quad |\partial_\beta h(0)| = \Big| \sum_{i=1}^n c_i \, \partial_\beta f_i(\boldsymbol{x}_0) \Big| \ne 0. \tag{5.1}$$

Now take $h \in \mathcal{H}$ and consider the functions $h \circ g$, where g runs through the group $\mathrm{GL}(d, \mathcal{O})$ of linear isometries of F^d. (We recall that $\mathcal{O} = \{x \in F \mid |x| \le 1\}$, see Example 2.6.) For any given multiindex γ, $\partial_\gamma(h \circ g)(0)$ is a homogeneous polynomial in matrix elements of g of degree $\le |\gamma|$ with coefficients given by $\partial_{\gamma'} f(0)$ where $|\gamma'| = |\gamma|$. It follows from (5.1) that for any γ with $|\gamma| = k$ this polynomial is nonzero. Hence it is possible to choose g so that $\partial_\gamma(h \circ g)(0) \ne 0$ for *all* multiindices γ with $|\gamma| = k$. In fact, we are only interested in choosing g with

$$\partial_i^k (h \circ g)(0) \ne 0 \quad \text{for each } i = 1, \dots, d. \tag{5.2}$$

Using (5.2) and the compactness of both $\mathcal{H}$ and $\mathrm{GL}(d, \mathcal{O})$, one can find a ball $V = V_1 \times \cdots \times V_d \ni 0$ in F^d (here V_i are balls in F of the same radius) such that for any $h \in \mathcal{H}$ there exist $1 \le k \le \ell$, $g \in \mathrm{GL}(d, \mathcal{O})$ and $A_1, \dots, A_d \in F \setminus \{0\}$ such that (4.3) holds for $f = h \circ g$. Therefore, by Theorem 4.1, for any $h \in \mathcal{H}$ one can find $1 \le k \le \ell$ and $g \in \mathrm{GL}(d, \mathcal{O})$ such that for any ball $B \subset V$ and $\varepsilon > 0$ one has

$$\lambda\big(B^{h\circ g, \varepsilon}\big) \le dk^{3-1/k} \left(\frac{\varepsilon}{\|h \circ g\|_B}\right)^{1/dk} \lambda(B) \le d\ell^{3-1/\ell} \left(\frac{\varepsilon}{\|h \circ g\|_B}\right)^{1/d\ell} \lambda(B).$$

To finish the proof it remains to notice that g leaves V invariant, sends balls to balls, and one clearly has $g(B^{h\circ g, \varepsilon}) = g(B)^{h, \varepsilon}$ and $\|h \circ g\|_B = \|h\|_{g(B)}$. □

5.3. For convenience let us summarize the results of this section as follows:

Theorem 5.2. *Let F be either $\mathbb{R}$ or an ultrametric valued field, and let $\boldsymbol{f}$ be a C^ℓ map from an open subset U of F^d to F^n. Then $\boldsymbol{f}$ is nonplanar and good at every point of U where it is nondegenerate.*

6. Maps of posets into spaces of good functions

6.1. 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. More precisely, we will work with mappings of partially ordered sets (*posets*) $\mathfrak{P}$ 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 6.1) for the measure of the set of "unmarked" points[3].

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 $\mathfrak{P} \setminus \mathfrak{S}$ comparable with any element of $\mathfrak{S}$. Note that one always has

$$l\big(\mathfrak{P}(\mathfrak{S})\big) \le l(\mathfrak{P}) - l(\mathfrak{S}). \tag{6.1}$$

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) $\varepsilon \le |\psi_s(z)| \le \rho$ for all $s \in \mathfrak{S}_z$;

(M2) $|\psi_s(z)| \ge \rho$ for all $s \in \mathfrak{P}(\mathfrak{S}_z)$.

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 6.1. *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 $\tilde{B} \stackrel{\text{def}}{=} B(x, 3^m r)$, such that the following holds:*

(A0) $l(\mathfrak{P}) \le m$;

[3] 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 for the special case of $\mathfrak{P}$ being the poset of nonzero rational subspaces of $\mathbb{R}^m$.

(A1) *ψ_s is (C, α)-good on $\tilde{B}$ with respect to μ for all $s \in \mathfrak{P}$;*

(A2) *$\|\psi_s\|_{\mu,B} \geq \rho$ for all $s \in \mathfrak{P}$;*

(A3) *$\#\{s \in \mathfrak{P} \mid |\psi_s(y)| < \rho\} < \infty$ for all $y \in \tilde{B} \cap \operatorname{supp}\mu$.*

Then one has

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

for all $\varepsilon \leq \rho$

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 $m \geq 1$ and suppose that the claim is proved for all smaller values of m.

Fix $C, \alpha, \rho, \mathfrak{P}, B = B(x, r)$ 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 $|\psi_s(y)| \geq \rho$ 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

$$\begin{aligned} E &\stackrel{\text{def}}{=} \{y \in B \cap \operatorname{supp}\mu \mid H(y) \neq \varnothing\} \\ &= \{y \in B \cap \operatorname{supp}\mu \mid \text{there exists an } s \in \mathfrak{P} \text{ with } |\psi_s(y)| < \rho\}. \end{aligned}$$

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

$$r_{s,y} \stackrel{\text{def}}{=} \sup\{t > 0 \mid \|\psi_s\|_{\mu,B(y,t)} < \rho\}. \tag{6.2}$$

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 (6.2) immediately implies that

$$\|\psi_s\|_{\mu,B_{s,y}} \leq \rho. \tag{6.3}$$

Now for any $y \in E$ choose an element s_y of $H(y)$ such that $r_{s_y,y} \geq r_{s,y}$ 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 $\mathfrak{P}_y$.

6.1.1. The next lemma allows one to show that a point $z \in B_y$ will 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 6.2. *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,*

$$(B_y \cap \operatorname{supp} \mu) \setminus \Phi(\varepsilon, \rho, \mathfrak{P}) \subset \big(B_y \setminus \Phi(\varepsilon, \rho, \mathfrak{P}_y)\big) \cup (B_y)^{\psi_{s_y}, \varepsilon}. \tag{6.4}$$

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

$$\varepsilon \le |\psi_s(z)| \le \rho \quad \text{for all } s \in \mathfrak{S}_{y,z} \tag{6.5}$$

and

$$|\psi_s(z)| \ge \rho \quad \text{for all } s \in \mathfrak{P}_y(\mathfrak{S}_{y,z}). \tag{6.6}$$

Put $\mathfrak{S}_z \stackrel{\text{def}}{=} \mathfrak{S}_{y,z} \cup \{s_y\}$. Then $\mathfrak{P}(\mathfrak{S}_z) = \mathfrak{P}_y(\mathfrak{S}_{y,z})$; therefore (M2) immediately follows from (6.6), and, in view of (6.5), 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 (6.3). □

6.1.2. 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

$$\|\psi_s\|_{\mu, B'_y} \ge \rho \quad \text{for any } y \in E \text{ and } s \in \mathfrak{P}. \tag{6.7}$$

(Indeed, the definition of r_y and (6.2) 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 $\tilde{B}'_y \stackrel{\text{def}}{=} B(y, 3^{m-1}r'_y)$ satisfy properties

- (A0) with m replaced by $m - 1$, in view of (6.1);
- (A2), in view of (6.7);
- (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

$$\begin{aligned} \mu\big(B_y \setminus \Phi(\varepsilon, \rho, \mathfrak{P}_y)\big) &\le \mu\big(B'_y \setminus \Phi(\varepsilon, \rho, \mathfrak{P}_y)\big) \\ &\le (m-1)C\big(N_X D_\mu^2\big)^{m-1} \left(\frac{\varepsilon}{\rho}\right)^{\alpha} \mu(B'_y) \\ &\le D_\mu (m-1)C\big(N_X D_\mu^2\big)^{m-1} \left(\frac{\varepsilon}{\rho}\right)^{\alpha} \mu(B_y) \end{aligned} \tag{6.8}$$

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

$$\begin{aligned}\mu\left((B_y)^{\psi_{s_y},\varepsilon}\right) \le \mu\left((B'_y)^{\psi_{s_y},\varepsilon}\right) &\le C\left(\frac{\varepsilon}{\|\psi_{s_y}\|_{\mu,B'_y}}\right)^{\alpha}\mu(B'_y)\\ &\underset{(6.7)}{\le} C\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\mu(B'_y) \le CD_\mu\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\mu(B_y).\end{aligned} \tag{6.9}$$

Recall that we need to estimate the measure of $E \setminus \Phi(\varepsilon,\rho,\mathfrak{P})$. For any $y \in E$, in view of (6.4), (6.8) and (6.9) one has

$$\begin{aligned}\mu\big(B_y \setminus \Phi(\varepsilon,\rho,\mathfrak{P})\big) &\le C\big((m-1)N_X^{m-1}D_\mu^{2m-1} + D_\mu\big)\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\mu(B_y)\\ &\le mCN_X^{m-1}D_\mu^{2m-1}\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\mu(B_y).\end{aligned} \tag{6.10}$$

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

$$\sum_{y\in Y}\mu(B_y) \le N_X\mu\Big(\bigcup_{y\in Y}B_y\Big) \le N_X\mu\big(B(x,3r)\big) \le N_XD_\mu\mu(B). \tag{6.11}$$

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

$$\begin{aligned}\sum_{y\in Y}\mu\big(B_y \setminus \Phi(\varepsilon,\rho,\mathfrak{P})\big) &\underset{(6.10)}{\to\le} mCN_X^{m-1}D_\mu^{2m-1}\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\sum_{y\in Y}\mu(B_y)\\ &\underset{(6.11)}{\to\le} mC\big(N_XD_\mu^2\big)^m\left(\frac{\varepsilon}{\rho}\right)^{\alpha}\mu(B).\end{aligned}$$

□

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

7.1. 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 7.1. *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 $\beta = 0$. □

7.2. If Δ is a $\mathcal{D}$-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* $\mathrm{rk}\,\Delta$ of Δ by

$$\mathrm{rk}(\Delta) \stackrel{\text{def}}{=} \dim_K(K\Delta). \tag{7.1}$$

For example, one has $\mathrm{rk}(\mathcal{D}^m) = m$ for any $m \in \mathbb{N}$. If Λ is a $\mathcal{D}$-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 $\mathrm{rk}(\Delta)$ 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 $\mathrm{rk}(\Lambda)$.

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

Lemma 7.2. *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 7.1, 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 7.2 implies that for any $\Delta' \subset \mathcal{D}^m$ there exists a unique primitive $\Delta \supset \Delta'$ of the same rank, namely, $\Delta = K\Delta' \cap \mathcal{D}^m$.

7.3. 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:

$$\mathfrak{M}(\mathcal{R}, \mathcal{D}, m) \stackrel{\text{def}}{=} \{g\Delta \mid g \in \mathrm{GL}(m, \mathcal{R}),\ \Delta \text{ is a submodule of } \mathcal{D}^m\}, \tag{7.2}$$

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.

We would like to have a way to measure "size" of submodules from the above collection. Specifically, let us say that a function $\nu\colon \mathfrak{M}(\mathcal{R}, \mathcal{D}, m) \mapsto \mathbb{R}_+$ 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 $\nu(\boldsymbol{v})$ being equal to the Euclidean norm of a vector $\boldsymbol{v} \in \mathbb{R}^m$; 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 6.1 to the poset $\mathfrak{P}(\mathcal{D}, m)$.

Theorem 7.3. *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 $h\colon \tilde{B} \to \mathrm{GL}(m, \mathcal{R})$ 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) $\#\{\Delta \in \mathfrak{P}(\mathcal{D}, m) \mid \psi_\Delta(x) < \rho\} < \infty$ *for all $x \in \tilde{B} \cap \operatorname{supp}\mu$.*

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

$$\mu\Big(\big\{x \in B \mid \nu(h(x)\gamma) < \varepsilon \text{ for some } \gamma \in \mathcal{D}^m \setminus \{0\}\big\}\Big) \leq mC\big(N_X D_\mu^2\big)^m \left(\frac{\varepsilon}{\rho}\right)^\alpha \mu(B). \tag{7.3}$$

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 6.1. Thus it suffices to prove that for any positive $\varepsilon \leq \rho$ one has

$$\Phi(\varepsilon, \rho, \mathfrak{P}) \subset \big\{x \in B \mid \nu(h(x)\gamma) \geq \varepsilon \text{ for all } \gamma \in \mathcal{D}^m \setminus \{0\}\big\}. \tag{7.4}$$

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 $\gamma \in \mathcal{D}^m \setminus \{0\}$. Then there exists i, $1 \le i \le l$, such that $\gamma \in \Delta_i \setminus \Delta_{i-1}$. From the primitivity of Δ_{i-1} and Lemma 7.2 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

$$\nu(h(x)\Delta') \le C_\nu \nu(h(x)\Delta_{i-1})\nu(h(x)\gamma). \tag{7.5}$$

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),

$$\nu(h(x)\Delta) \le \nu(h(x)\Delta'). \tag{7.6}$$

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

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

and then, in view of (7.5) and (7.6), conclude that

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

This shows (7.4) and completes the proof of the theorem. □

8. Discrete submodules of $\mathbb{Q}_S^m$

8.1. 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.

Let $\ell \in \mathbb{N}$ and take $S = \{p_1, \dots, p_{\ell-1}, \infty\}$, where $p_1, \dots, p_{\ell-1}$ are primes. The (possibly empty) subset $\{p_1, \dots, p_{\ell-1}\}$ of S will be denoted by S_f. To every element of S we associate the *normalized* valuation $|\cdot|_v$ of $\mathbb{Q}$; in other words, $|\cdot|_v$ is the usual absolute value if $v = \infty$, and is defined as in Example 2.6 if v is p-adic. We let $\mathbb{Q}_S$ be the direct product of all the completions $\mathbb{Q}_v$, $v \in S$, in which $\mathbb{Q}$ is diagonally imbedded (here we use the notation $\mathbb{Q}_\infty = \mathbb{R}$), and let

$$\mathbb{Z}_S \stackrel{\text{def}}{=} \mathbb{Z}\big[\tfrac{1}{p_1}, \dots, \tfrac{1}{p_{\ell-1}}\big] = \big\{x \in \mathbb{Q} \mid x \in \mathbb{Z}_p \text{ for all primes } p \notin S_f\big\}$$

stand for the ring of S-integers of $\mathbb{Q}$. We also let

$$\mathbb{Z}_{S,f} \stackrel{\text{def}}{=} \prod_{i=1}^{\ell-1} \mathbb{Z}_{p_i} \quad \text{and} \quad \mathbb{Q}_{S,f} \stackrel{\text{def}}{=} \mathbb{R} \times \mathbb{Z}_{S,f}.$$

Denote by λ_v the *normalized* Haar measure on $\mathbb{Q}_v$ (that is, λ_∞ is the usual Lebesgue measure on $\mathbb{R}$ and λ_{p_i} is normalized by $\lambda_{p_i}(\mathbb{Z}_{p_i}) = 1$), and by $\lambda_S = \prod_{v\in S} \lambda_v$ the product measure on $\mathbb{Q}_S$. Elements of $\mathbb{Q}_S$ will be denoted as $x = (x^{(v)})_{v\in S}$ or simply $x = (x^{(v)})$, where $x^{(v)} \in \mathbb{Q}_v$. For x of this form, we define the S-adic *absolute value* $|\boldsymbol{x}|$ and the *content* $c(x)$ of x to be the maximum (resp. the product) of all $|x^{(v)}|_v$, $v \in S$. Since all the valuations are normalized, one has

$$\lambda_S(xM) = c(x)\lambda_S(M), \tag{8.1}$$

where $x \in \mathbb{Q}_S$ and M is a measurable subset of $\mathbb{Q}_S$.

If m is a natural number, we preserve the same notation λ_v and λ_S to denote the product measure on $\mathbb{Q}_v^m$ and $\mathbb{Q}_S^m$, respectively. Elements $\boldsymbol{x} = (x_1, \dots, x_m)$ of $\mathbb{Q}_S^m$ will be denoted as $\boldsymbol{x} = (\boldsymbol{x}^{(v)})$, where $\boldsymbol{x}^{(v)} = (x_1^{(v)}, \dots, x_m^{(v)}) \in \mathbb{Q}_v^m$. We denote by $\|\cdot\|_v$ the usual (Euclidean) norm on $\mathbb{R}^m$ if $v = \infty$, and the sup-norm defined by

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

if v is non-Archimedean. For $\boldsymbol{x} = (\boldsymbol{x}^{(v)})$ in $\mathbb{Q}_S^m$ we define the *norm* $\|\boldsymbol{x}\|$ and the *content* $c(\boldsymbol{x})$ of $\boldsymbol{x}$ to be the maximum (resp., the product) of all the numbers $\|(\boldsymbol{x}^{(v)})\|_v$, $v \in S$. The group $\mathrm{GL}(m, \mathbb{Q}_S) = \prod_{v\in S} \mathrm{GL}(m, \mathbb{Q}_v)$ acts naturally on $\mathbb{Q}_S^m$, and one has

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

where $M \subset \mathbb{Q}_S^m$ is any measurable subset of $\mathbb{Q}_S^m$, $g = (g^{(v)}) \in \mathrm{GL}(m, \mathbb{Q}_S)$, and $\det(g) \stackrel{\text{def}}{=} \big(\det(g^{(v)})\big)$ is an invertible element of $\mathbb{Q}_S$.

8.2. Our goal now is to consider discrete $\mathbb{Z}_S$-submodules Δ of $\mathbb{Q}_S^m$. It turns out that any such Δ is a finitely generated free $\mathbb{Z}_S$-module:

Proposition 8.1. *Let Δ be a discrete $\mathbb{Z}_S$-submodule of $\mathbb{Q}_S^m$. Then $\Delta = \mathbb{Z}_S\boldsymbol{a}_1 \oplus \cdots \oplus \mathbb{Z}_S\boldsymbol{a}_r$ for some $\boldsymbol{a}_1, \dots, \boldsymbol{a}_r \in \mathbb{Q}_S^m$ such that*

$$v\boldsymbol{a}_1^{(v)}, \dots, \boldsymbol{a}_r^{(v)} \text{ are linearly independent over } \mathbb{Q}_v \text{ for any } v \in S. \tag{8.2}$$

Furthermore, there exists $g \in \mathrm{GL}(m, \mathbb{Q}_S)$ such that Δ is contained in $g\mathbb{Z}_S^m$.

Proof. The proposition is trivial if $S = \infty$. So assume that $S \supsetneq \infty$ and denote by Δ_0 the intersection of Δ with $\mathbb{Q}^m_{S,f}$. Let $\pi : \mathbb{Q}^m_{S,f} \to \mathbb{Q}^m_\infty$ be the natural projection. Since $\ker(\pi) = \mathbb{Z}^m_{S,f}$ is compact and $\mathbb{Z}^m_{S,f}$ does not contain nontrivial discrete subgroups, $\pi(\Delta_0)$ is a free abelian group of rank $r \le m$ and $\pi(\Delta_0)$ is isomorphic to Δ_0. If $\boldsymbol{x}$ is any element in Δ, then there exists an element ξ of $\mathbb{Z}^*_S$ (the group of S-adic units) such that $\xi\boldsymbol{x} \in \Delta_0$. This implies that Δ is a free $\mathbb{Z}_S$-module of rank r.

Let $\Delta = \mathbb{Z}_S\boldsymbol{a}_1 \oplus \cdots \oplus \mathbb{Z}_S\boldsymbol{a}_r$. Suppose that $\boldsymbol{a}_1^{(v)} = 0$ for some $v \in S$, and let $\{\xi_i\}$ be a sequence of S-adic units such that $\lim_{i\to\infty} |\xi_i|_w = 0$ for all $w \ne v$. Then $\lim_{i\to\infty} \xi_i\boldsymbol{a}_1 = 0$, which contradicts the discreteness of $\mathbb{Z}_S\boldsymbol{a}_1$. Therefore $\boldsymbol{a}_1^{(v)} \ne 0$ for all $v \in S$, which proves (8.2) for $r = 1$. To complete the proof we use induction on r. Assume that $r > 1$ and (8.2) is true for free modules of rank $< r$. Shifting Δ by an appropriate automorphism from $\mathrm{GL}(m, \mathbb{Q}_S)$, without loss of generality we may and will assume that $\boldsymbol{a}_1 = \boldsymbol{e}_1$, the first vector of the standard basis of $\mathbb{Q}^m_S$. Let $\varphi\colon \mathbb{Q}^m_S \to \mathbb{Q}^m_S/\mathbb{Q}_S\boldsymbol{e}_1$ be the natural homomorphism. Since $\mathbb{Z}_S\boldsymbol{e}_1$ is a cocompact lattice in $\mathbb{Q}_S\boldsymbol{e}_1$, we get that $\varphi(\Delta)$ is discrete in $\varphi(\mathbb{Q}^m_S) \cong \mathbb{Q}^{m-1}_S$. By the induction hypothesis $\varphi(\boldsymbol{a}_2)^{(v)}, \dots, \varphi(\boldsymbol{a}_r)^{(v)}$ are linearly independent over $\mathbb{Q}_v$ for all v, which completes the proof of (8.2). It remains to observe that the last part of the proposition immediately follows from (8.2). □

Note that it follows from Proposition 8.1 that the rank of Δ as a free module is equal to $\mathrm{rk}(\Delta)$ as defined in (7.1) with $K = \mathbb{Q}$.

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

$$\text{if } v = \infty, \quad \text{then } (\boldsymbol{b}_1^{(v)}, \dots, \boldsymbol{b}_r^{(v)}) \text{ is an orthonormal basis of } L_v, \tag{8.3a}$$

and

$$\text{if } v \in S_f, \text{ then } L_v \cap \mathbb{Z}^m_v = \mathbb{Z}_v\boldsymbol{b}_1^{(v)} + \cdots + \mathbb{Z}_v\boldsymbol{b}_r^{(v)}. \tag{8.3b}$$

Then consider the $\mathbb{Q}_S$-linear map sending the standard basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_r)$ of $\mathbb{Q}^r_S$ to $(\boldsymbol{b}_1, \dots, \boldsymbol{b}_r)$, and define the *volume* on L as the pushforward of λ_S (the normalized Haar measure on $\mathbb{Q}^r_S$) 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 $\boldsymbol{b}_1^{(v)}, \dots, \boldsymbol{b}_r^{(v)}$ in the case (8.3b) easily follows from the fact that $\mathbb{Z}_v$ is a principal ideal domain. Note also that the above definition does not depend on the choice of the basis $(\boldsymbol{b}_1, \dots, \boldsymbol{b}_r)$ satisfying (8.3a, b), because if $(\boldsymbol{b}'_1, \dots, \boldsymbol{b}'_r)$ is another such basis and $h \in \mathrm{GL}(r, \mathbb{Q}_S)$ represents the isomorphism of $L \cong \mathbb{Q}^r_S$ sending the first basis to the second one, then $c(\det(h)) = 1$, which implies that h is measure-preserving.

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

$$\bigwedge^r \mathbb{Q}_S^m \simeq \bigoplus_{v \in S} \bigwedge^r \mathbb{Q}_v^m$$

of $\mathbb{Q}_S^m$, which is a free $\mathbb{Q}_S$-module with the standard basis

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

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

8.4. Recall (cf. [W]) that $\mathbb{Z}_S$ is a lattice in $\mathbb{Q}_S$ with covolume 1, that is, it is discrete in $\mathbb{Q}_S$ and the Haar measure (induced by λ_S) of the quotient space is equal to one. Likewise, $\mathbb{Z}_S^m$ is a lattice in $\mathbb{Q}_S^m$ with covolume 1. It follows from Proposition 8.1 that the set of discrete $\mathbb{Z}_S$-submodules of $\mathbb{Q}_S^m$ can be identified with $\mathfrak{M}(\mathbb{Q}_S, \mathbb{Z}_S, m)$ as defined in (7.2). Furthermore, it also follows that any $\Delta \in \mathfrak{M}(\mathbb{Q}_S, \mathbb{Z}_S, m)$ is a lattice in $\mathbb{Q}_S\Delta$. The following lemma shows how one can explicitly compute covolumes:

Lemma 8.2. *Let $\Delta = \mathbb{Z}_S\boldsymbol{a}_1 \oplus \cdots \oplus \mathbb{Z}_S\boldsymbol{a}_r \in \mathfrak{M}(\mathbb{Q}_S, \mathbb{Z}_S, m)$, where $\boldsymbol{a}_1, \dots, \boldsymbol{a}_r \in \mathbb{Q}_S^m$. Then the covolume* $\operatorname{cov}(\Delta)$ *of Δ in $\mathbb{Q}_S\Delta$ with respect to the volume on $L = \mathbb{Q}_S\Delta$ normalized as in* §8.3 *is equal to*

$$c(\boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_r).$$

Proof. Put $L = \mathbb{Q}_S\Delta$, define a basis $(\boldsymbol{b}_1, \dots, \boldsymbol{b}_r)$ of L as in (8.3a, b), and then complete it to a basis $(\boldsymbol{b}_1, \dots, \boldsymbol{b}_m)$ of the whole space $\mathbb{Q}_S^m$. Also let $h_v \in \mathrm{GL}(m, \mathbb{Q}_v)$ be such that $h_v\boldsymbol{b}_i^{(v)} = \boldsymbol{e}_i^{(v)}$ for all $1 \le i \le m$, where $\boldsymbol{e}_1^{(v)}, \dots, \boldsymbol{e}_m^{(v)}$ is the standard basis of $\mathbb{Q}_v^m$. It follows from the definition of the measure on L that the map $h\colon L \to h(L)$, where $h = (h_v) \in \mathrm{GL}(m, \mathbb{Q}_S)$, is measure preserving. Since the map $\bigwedge^r h\colon \bigwedge^r \mathbb{Q}_S^m \to \bigwedge^r \mathbb{Q}_S^m$ preserves the content c on $\bigwedge^r \mathbb{Q}_S^m$, we may reduce the problem to the case $\boldsymbol{b}_i = \boldsymbol{e}_i$ for all $1 \le i \le m$. Let $\varphi \in \mathrm{GL}(r, \mathbb{Q}_S)$ be such that $\varphi(\boldsymbol{e}_i) = \boldsymbol{a}_i$ for all $1 \le i \le r$. Then since $\boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_r = \det(\varphi)\boldsymbol{e}_1 \wedge \cdots \wedge \boldsymbol{e}_r$, we get

$$\operatorname{cov}(\Delta) = \operatorname{cov}\big(\varphi(\mathbb{Z}_S^r)\big) = c(\det\varphi)\operatorname{cov}(\mathbb{Z}_S^r) = c(\boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_r). \qquad \square$$

8.5. Lemma 8.2 immediately implies

Corollary 8.3. *For every $\Delta \in \mathfrak{M}(\mathbb{Q}_S, \mathbb{Z}_S, m)$, the function $\mathrm{GL}(m, \mathbb{Q}_S) \to \mathbb{R}_+$, $g \mapsto \operatorname{cov}(g\Delta)$, is continuous.*

Corollary 8.4. *If $\Delta,\ \Delta' \in \mathfrak{M}(\mathbb{Q}_S, \mathbb{Z}_S, m)$ are such that $\mathbb{Q}_S\Delta \cap \mathbb{Q}_S\Delta' = \{0\}$, then*

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

Proof. Using Proposition 8.1, write $\Delta = \mathbb{Z}_S \boldsymbol{a}_1 \oplus \cdots \oplus \mathbb{Z}_S \boldsymbol{a}_r$ and $\Delta' = \mathbb{Z}_S \boldsymbol{b}_1 \oplus \cdots \oplus \mathbb{Z}_S \boldsymbol{b}_r$. Since

$$\mathbb{Q}_S\Delta + \mathbb{Q}_S\Delta' = \mathbb{Q}_S\Delta \oplus \mathbb{Q}_S\Delta',$$

in view of Lemma 8.2, it is enough to prove that

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

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

8.6. In the remaining part of the section we investigate metric properties of discrete submodules of $\mathbb{Q}_S^m$. Let us state the following S-arithmetic version of the classical Minkowski's Lemma:

Lemma 8.5. *Let $\Delta \in \mathfrak{M}(\mathbb{Q}_S, \mathbb{Z}_S, m)$ be of rank r, and let B be a closed ball in $\mathbb{Q}_S\Delta$ (with respect to the norm $\|\cdot\|$) centered at 0 such that $\lambda_S(B) \ge 2^r \operatorname{cov}(\Delta)$. Then $\Delta \cap B \ne \{0\}$.*

Proof. Since the volume λ_S on $L = \mathbb{Q}_S\Delta$ was defined by identifying L with $\mathbb{Q}_S^r$ via the basis (8.3a, b), without loss of generality we can assume that Δ is a lattice in $\mathbb{Q}_S^r$. Then we can write $B = B_\infty \times B_f$, where $B_\infty \stackrel{\text{def}}{=} B \cap \mathbb{Q}_\infty^r$ and $B_f \stackrel{\text{def}}{=} B \cap \big(\prod_{v \in S_f} \mathbb{Q}_v\big)^r$. Note that

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

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

8.7. We will also need the following result:

Lemma 8.6. *There exists a constant $A > 1$ depending only on S such that if $\boldsymbol{x} \in \mathbb{Q}_S^m$ and $c(\boldsymbol{x}) \ne 0$, then there exists $\xi \in \mathbb{Z}_S^*$ such that*

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

where ℓ is the cardinality of S.

Proof. Let

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

Write $S = \{v_1, \dots, v_\ell\}$; it is easy to see that the group

$$\{(|\xi|_{v_1}, \dots, |\xi|_{v_\ell}) \mid \xi \in \mathbb{Z}_S^*\}$$

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 an S-adic unit ξ with

$$A^{-1} \le |\xi|_{v_i} a_i \le A \tag{8.4}$$

for all i.

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

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

and

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

for all $\xi \in \mathbb{Z}_S^*$. This, in view of (8.4), implies the claim. □

8.8. Let us denote by $\Omega_{S,m}$ the space of all lattices in $\mathbb{Q}_S^m$. It follows from Lemmas 8.2 and 8.4 that it can also be defined as

$$\Omega_{S,m} = \{g\mathbb{Z}_S^m \mid g \in \mathrm{GL}(m, \mathbb{Q}_S)\} \cong \mathrm{GL}(m, \mathbb{Q}_S)/\mathrm{GL}(m, \mathbb{Z}_S).$$

Corollary 8.7. *For any $\Lambda \in \Omega_{S,m}$ and any $\rho > 0$, the number of submodules of Λ with covolume $\le \rho$ is finite.*

Proof. If Δ is a $\mathbb{Z}_S$-submodule with $\mathrm{rk}(\Delta) = r$, then $\bigwedge^r \Delta$ is a $\mathbb{Z}_S$-submodule of rank 1 of the lattice $\bigwedge^r \Lambda$ in $\bigwedge^r \mathbb{Q}_S^m$, and, in view of Lemma 8.2,

$$\mathrm{cov}(\Delta) = \mathrm{cov}\left(\bigwedge^r \Delta\right).$$

Therefore it is enough to prove that the number of rank-one $\mathbb{Z}_S$-submodules of Λ is finite. If $\Delta = \mathbb{Z}_S \boldsymbol{a}$ is such a submodule, then in view of Lemma 8.6 the generator $\boldsymbol{a}$ can be chosen in such a way that

$$\|\boldsymbol{a}\| \le A\rho^{1/\ell}.$$

Since Λ is discrete in $\mathbb{Q}_S^m$, the set of all $\boldsymbol{a} \in \Lambda$ satisfying the above inequality is finite, which proves the corollary. □

8.9. Because of lack of an appropriate reference we will prove an S-adic version of Mahler's Compactness Criterion. Consider the group

$$\mathrm{GL}^1(m, \mathbb{Q}_S) \stackrel{\mathrm{def}}{=} \{g \in \mathrm{GL}(m, \mathbb{Q}_S) \mid c(\det(g)) = 1\},$$

consisting of λ_S-preserving linear automorphisms of $\mathbb{Q}_S^m$, and let

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

Note that $\mathbb{Z}_S^m$ is an element of $\Omega^1_{S,m}$, and its stabilizer in $\mathrm{GL}(m,\mathbb{Q}_S)$ coincides with $\mathrm{GL}(m,\mathbb{Z}_S)$ which is understood to be diagonally imbedded in $\mathrm{GL}(m,\mathbb{Q}_S)$. Since $c(\xi)=1$ for any $\xi\in\mathbb{Z}_S^*$, it follows that $\mathrm{GL}(m,\mathbb{Z}_S)$ is contained in $\mathrm{GL}^1(m,\mathbb{Q}_S)$. Thus $\Omega^1_{S,m}$ is naturally identified with the homogeneous space $\mathrm{GL}^1(m,\mathbb{Q}_S)/\mathrm{GL}(m,\mathbb{Z}_S)$. Since $\mathrm{SL}(m,\mathbb{Z}_S)$ is a lattice in $\mathrm{SL}(m,\mathbb{Q}_S)$ and $\mathbb{Z}_S^*$ is a lattice in $\mathrm{GL}(1,\mathbb{Q}_S)$, $\mathrm{GL}(m,\mathbb{Z}_S)$ is a lattice in $\mathrm{GL}^1(m,\mathbb{Q}_S)$.

Let us say that a set Q of lattices in $\mathcal{R}^m$, where $\mathcal{R}$ is a locally compact 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 8.8 (Mahler's Compactness Criterion). *A subset $Q\subset\Omega^1_{S,m}$ is bounded if and only if it is separated from* 0.

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

$$\mathrm{GL}^1(m,\mathbb{Q}_S)=\mathbb{Q}_S^1\ltimes\mathrm{SL}(m,\mathbb{Q}_S)\quad\text{and}\quad\mathrm{GL}(m,\mathbb{Z}_S)=\mathbb{Z}_S^*\ltimes\mathrm{SL}(m,\mathbb{Z}_S),$$

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

$$\{g\mathbb{Z}_S^m\mid g\in\mathrm{SL}(m,\mathbb{Z}_S)\}.$$

It follows from the strong approximation theorem [PR, Theorem 7.12] that

$$\mathrm{SL}(m,\mathbb{Q}_S)=\mathrm{SL}(m,\mathbb{Q}_{S,f})\,\mathrm{SL}(m,\mathbb{Z}_S).$$

Thus every $g\in\mathrm{SL}(m,\mathbb{Q}_S)$ can be represented as $g=g_fg_l$, where $g_f\in\mathrm{SL}(m,\mathbb{Q}_{S,f})$ and $g_l\in\mathrm{SL}(m,\mathbb{Z}_S)$. One has

$$g\mathbb{Z}_S^m\cap\mathbb{Q}_{S,f}^m=g_f\big(\mathbb{Z}_S^m\cap\mathbb{Q}_{S,f}^m\big)=g_f\mathbb{Z}^m.\tag{8.5}$$

Let $\tilde{Q}\subset\mathrm{SL}(m,\mathbb{Q}_S)$ be such that the set of lattices $\{g\mathbb{Z}_S^m\mid g\in\tilde{Q}\}$ is separated from 0. It follows from (8.5) that

$$\{g_f\mathbb{Z}^m\mid g\in\tilde{Q}\}\text{ is separated from 0 in }\mathbb{Q}_{S,f}^m.\tag{8.6}$$

Note that $\mathrm{SL}(m,\mathbb{Q}_{S,f})=\mathrm{SL}(m,\mathbb{R})\times\mathrm{SL}(m,\mathbb{Z}_{S,f})$, and, therefore, every $g_f\in\mathrm{SL}(m,\mathbb{Q}_{S,f})$ can be written as $g_f=g_\infty g_c$, where $g_\infty\in\mathrm{SL}(m,\mathbb{R})$ and g_c belongs to the compact group $\mathrm{SL}(m,\mathbb{Z}_{S,f})$. It follows from (8.6) that $\{g_\infty\mathbb{Z}^m\mid g\in\tilde{Q}\}$ is separated from 0. This reduces the proof to the case $S=\infty$, that is, to the original Mahler's Criterion [PR, Proposition 4.18]. □

In particular, it follows from Lemma 8.6 and Theorem 8.8 that for all positive ε, the sets Q_ε defined as in (1.2) are compact.

9. S-arithmetic quantitative nondivergence

9.1. In this section we apply Theorem 7.3 to the triple $(\mathcal{D}, K, \mathcal{R})$ of a particular type. Namely, as in §8, we let $K = \mathbb{Q}$, choose a finite set S of valuations $|\cdot|_v$ of $\mathbb{Q}$ containing the Archimedean one, and for the rest of this section take $\mathcal{D} = \mathbb{Z}_S$ and $\mathcal{R} = \mathbb{Q}_S$. Then for any $m \in \mathbb{N}$, the set $\mathfrak{M}(\mathbb{Q}_S, \mathbb{Z}_S, m)$, as defined in (7.2), is equal to the set of all submodules of all lattices $\Lambda \in \Omega_{S,m}$, where $\Omega_{S,m}$ is as defined in (8.5). Let us now state the following

Lemma 9.1. *The function $\nu\colon \mathfrak{M}(\mathbb{Q}_S, \mathbb{Z}_S, m) \to \mathbb{R}_+$ given by $\nu(\Delta) = \operatorname{cov}(\Delta)$, with $\operatorname{cov}(\cdot)$ as in §8.4, is norm-like, with $C_\nu = 1$.*

Proof. Property (N1) is straightforward since $\Delta' \subset \Delta$ and $\mathbb{Q}_S\Delta' = \mathbb{Q}_S\Delta$ implies that Δ is a subgroup of Δ' of finite index and $\operatorname{cov}(\Delta') = [\Delta : \Delta']\operatorname{cov}(\Delta)$. Property (N2) with $C_\nu = 1$ follows from Corollary 8.4 with $\Delta' = \mathbb{Z}_S\gamma$. Finally, (N3) has already been mentioned as Corollary 8.3. □

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

$$\delta(\Lambda) \stackrel{\text{def}}{=} \min\big\{c(\boldsymbol{x}) \mid \boldsymbol{x} \in \Lambda \setminus \{0\}\big\}.$$

Note that the minimum is well defined due to Lemma 8.6 and every $\Lambda \in \Omega_{S,m}$ being discrete in $\mathbb{Q}_S^m$. We will use the following

Lemma 9.2. *There exists a constant $A > 0$ depending only on 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 $\mathbb{Q}_S\Delta$ centered at 0 of radius ε (with respect to the norm $\|\cdot\|$ introduced in §8.1). Then one has $\lambda_S(B) \le \text{const}\cdot\varepsilon^{r\ell}$, where $r = \operatorname{rk}(\Delta)$, ℓ is the cardinality of S, and the constant depends only on S and m. By Lemma 8.5, Δ has a nontrivial intersection with B whenever $\text{const}\cdot\varepsilon^{r\ell} \ge 2^r\rho$. This shows how one can choose A such that Δ (and hence Λ) is guaranteed to contain a nonzero vector $\boldsymbol{x}$ with $\|\boldsymbol{x}\| \le A^{1/\ell}\rho^{1/r\ell}$, which clearly implies $c(\boldsymbol{x}) \le A\rho^{1/r} \le A\rho^{1/m}$. □

9.3. As in §7, let us use the notation $\mathfrak{P}(\mathbb{Z}_S, m)$ for the set of all nonzero primitive submodules of $\mathbb{Z}_S^m$.

Theorem 9.3. *Let X be a Besicovitch metric space, μ a uniformly Federer measure on X, and let S be as above. For $m \in \mathbb{N}$, let a ball $B = B(x_0, r_0) \subset X$ and a continuous map $h\colon \tilde{B} \to \mathrm{GL}(m, \mathbb{Q}_S)$ 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* $\Delta \in \mathfrak{P}(\mathbb{Z}_S, m)$*, the function* $\operatorname{cov}\big(h(\cdot)\Delta\big)$ *is* (C, α)*-good on* $\tilde{B}$ *with respect to* μ*;*

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

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

$$\mu\left(\{x \in B \mid \delta(h(x)\mathbb{Z}_S^m) < \varepsilon\}\right) \leq mC\big(N_X D_\mu^2\big)^m \left(\frac{\varepsilon}{\rho}\right)^\alpha \mu(B). \tag{9.1}$$

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

Proof. To apply Theorem 7.3, one uses Lemma 9.1 which guarantees the norm-like property of $\operatorname{cov}(\cdot)$, and Corollary 8.7 which implies condition (iii) of Theorem 7.3. To derive (9.1) from (7.3) it remains to observe that $\delta\big(h(x)\mathbb{Z}_S^m\big) < \varepsilon$ amounts to the existence of a vector $\boldsymbol{x} \in h(x)\mathbb{Z}_S^m \setminus \{0\}$ with

$$\operatorname{cov}(\mathbb{Z}_S\boldsymbol{x}) = c(\boldsymbol{x}) < \varepsilon. \qquad \square$$

9.4. In order to interpret the above result, let us assume, as it will be the case in many applications, that the function h takes values in the group $\mathrm{GL}^1(m, \mathbb{Q}_S)$. Then $h(x)\mathbb{Z}_S^m$ belongs to $\Omega^1_{S,m}$ for any x, and the inequality $\delta(h(x)\mathbb{Z}_S^m) < \varepsilon$ can be equivalently written as $h(x)\mathbb{Z}_S^m \notin Q_\varepsilon$. This way, Theorem 9.3 estimates, in terms of ε, the relative measure of points $x \in B$ which are mapped, by $x \mapsto h(x)\mathbb{Z}_S^m$, 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 $X = \prod_{v\in S} \mathbb{Q}_v^{d_v}$, $\mu = \lambda$ as defined in (1.3), and a map $h = (h_v)_{v\in S} \colon X \to \mathrm{GL}^1(m, \mathbb{Q}_S)$, where each h_v is a map from $\mathbb{Q}_v^{d_v}$ to $\mathrm{GL}(m, \mathbb{Q}_v)$. 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 $\bar{\mathbb{Q}}_v^{d_v} \to \mathrm{GL}(m, \bar{\mathbb{Q}}_v)$, where $\bar{\mathbb{Q}}_v$ is the algebraic closure of $\mathbb{Q}_v$).

Theorem 9.4. *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*

$$\lambda\left(\{x \in B \mid h(x)\mathbb{Z}_S^m \notin Q_\varepsilon\}\right) \leq C_0\varepsilon^\alpha \lambda(B) \tag{9.2}$$

whenever $h(B)\mathbb{Z}_S^m \cap Q_\tau \neq \varnothing$*. Furthermore, if* $h(X)\mathbb{Z}_S^m \cap Q_\tau = \varnothing$*, then there exists a proper (not of maximal rank)* $\Delta \in \mathfrak{P}(\mathbb{Z}_S, m)$ *such that* $h(x)(\mathbb{Q}_S\Delta) = h(0)(\mathbb{Q}_S\Delta)$ *for all* $x \in X$*.*

Proof. It was mentioned in Example 2.6 that X is Besicovitch, and in Example 2.7 that λ is uniformly Federer. Using the exterior power representation of Lemma 8.2, one can easily show that for every $\Delta \in \mathfrak{P}(\mathbb{Z}_S, m)$ the function $\operatorname{cov}\big(h(\cdot)\Delta\big)$ has the form $\prod_{v\in S} \|f_v\|_v$, where each f_v is a polynomial map from $\mathbb{Q}_v^{d_v}$ to another vector space over $\mathbb{Q}_v$. Every such function is (C', α')-good with uniform C' and α' due to Lemma 3.4, Lemma 3.1 (c), (d) and Corollary 3.3. Thus condition (i) of Theorem 9.3 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 9.2 and Theorem 8.8 that there exists $\tau = \tau(L) > 0$ such that for any submodule Δ of any $\Lambda \in L$ one has $\operatorname{cov}\big(\Delta\big) \geq \tau$. Without loss of generality we can assume that $\tau < \frac{1}{(2\tilde{C})^{1/\alpha}}$. Note that L is contained in Q_τ, since by definition of τ one has $c(\boldsymbol{x}) = \operatorname{cov}(\mathbb{Z}_S \boldsymbol{x}) \geq \tau$ for any nonzero element $\boldsymbol{x}$ of any $\Lambda \in L$.

If B is such that $h(B)\mathbb{Z}_S^m \cap Q_\tau \neq \varnothing$, it follows from Lemma 9.2 that condition (ii) of Theorem 9.3 is satisfied with $(\tau/A)^m$ in place of ρ (where A is as in Lemma 9.2). Therefore one has

$$\lambda\left(\{x \in B \mid h(x)\mathbb{Z}_S^m \notin Q_\varepsilon\}\right) \leq \tilde{C}\left(\frac{\varepsilon A^m}{\tau^m}\right)^\alpha \lambda(B)$$

for all $\varepsilon \leq \tau^m / A^m$. Replacing, if necessary, $\tilde{C}$ by a larger number C, we conclude that (9.2) is valid for all positive ε.

Now assume that $h(x)\mathbb{Z}_S^m \notin Q_\tau$ for all $x \in X$. Take $\rho \stackrel{\text{def}}{=} (2\tilde{C})^{1/\alpha}\tau < 1$, 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

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

Then clearly $\mathfrak{P}_{i+1} \subset \mathfrak{P}_i$ for all i, and all these sets are finite due to Corollary 8.7. We claim that $\bigcap_{i=1}^\infty \mathfrak{P}_i$ must be nonempty. Indeed, otherwise one obtains a nonempty ball B such that $\|\operatorname{cov}\big(h(\cdot)\Delta\big)\|_B \geq \rho$ for every $\Delta \in \mathfrak{P}(\mathbb{Z}_S, m)$. Thus Theorem 9.3 can be applied, and one can conclude that

$$\lambda(B) = \lambda\left(\{x \in B \mid h(x)\mathbb{Z}_S^m \notin Q_\tau\}\right) \leq \tilde{C}\left(\frac{\tau}{\rho}\right)^\alpha \lambda(B) = \frac{1}{2}\lambda(B),$$

a contradiction.

Consequently, there exists a proper $\Delta \in \mathfrak{P}(\mathbb{Z}_S, m)$ 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 $h(x)(\mathbb{Q}_S\Delta)$ does not depend on x. □

We note that to derive Theorem 1.2 from the above theorem, one needs to take $L = \{\Lambda\}$ and h of the form $x \mapsto h(x)g$, where $g \in \mathrm{GL}^1(m, \mathbb{Q}_S)$ is such that $\Lambda = g\mathbb{Z}_S^m$, and observe that $h(B)\mathbb{Z}_S^m \cap Q_\tau$ is nonempty whenever B contains 0.

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

10.1. Theorem 9.4 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 [MT3], [Rt4] and [To]).

Let us recall some definitions from [MT3]. As in the previous sections, S is a finite set of normalized valuations of $\mathbb{Q}$ containing the Archimedean one, and $\mathbb{Q}_S$ is the direct product of all $\mathbb{Q}_v$, $v \in S$. By a *$\mathbb{Q}_S$-algebraic group* $\boldsymbol{G}$ we mean a (formal) direct product $\prod_{v\in S} \boldsymbol{G}_v$ of $\mathbb{Q}_v$-algebraic groups $\boldsymbol{G}_v$. The group $\prod_{v\in S} \boldsymbol{G}_v(\mathbb{Q}_v)$ will be denoted by $\boldsymbol{G}(\mathbb{Q}_S)$ and called *the group of $\mathbb{Q}_S$-rational points* of $\boldsymbol{G}$. We will also use the simpler notations G for $\boldsymbol{G}(\mathbb{Q}_S)$ and G_v for $\boldsymbol{G}_v(\mathbb{Q}_v)$. If $\boldsymbol{H}$ is another $\mathbb{Q}_S$-algebraic group, then a homomorphism $\varphi\colon \boldsymbol{G} \to \boldsymbol{H}$ is called *$\mathbb{Q}_S$-homomorphism* if φ is a product of $\mathbb{Q}_v$-homomorphisms of algebraic groups $\varphi_v\colon \boldsymbol{G}_v \to \boldsymbol{H}_v$, $v \in S$. We preserve the same terminology for the restriction map $\varphi\colon G \to H$. By the *Zariski topology* on $\boldsymbol{G}$ (respectively, G) we mean the formal product of the Zariski topologies on $\boldsymbol{G}_v$ (respectively, G_v). By a *$\mathbb{Q}_S$-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*.

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.

Let U be a unipotent algebraic subgroup of G. Then $U = \prod_{v\in S} U_v$, where U_v are unipotent algebraic subgroups of G_v. Given $v \in S$, we denote by $\exp_v\colon \mathrm{Lie}(U_v) \to U_v$ 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, $v \in S$, and by $\exp\colon \mathrm{Lie}(U) \to U$ the direct product of the maps $\exp_v$, $v \in S$. By a (rational) *parametrization* of U we mean a product $\phi = (\phi_v)_{v\in S}$ of surjective maps $\phi_v\colon \mathbb{Q}_v^{d_v} \to U_v$, $v \in S$, such that for every $v \in S$ the map $\log_v \circ \phi_v\colon \mathbb{Q}_v^{d_v} \to \mathrm{Lie}(U_v)$ is polynomial and $\phi_v(0) = e$. If d_v is the degree of $\log \circ \phi_v$ then $d = \max\{d_v | v \in S\}$ is called the *degree* of the parametrization ϕ. Clearly exp is a parametrization of U which we call trivial. We let $X = \prod_{v\in S} \mathbb{Q}_v^{d_v}$ and endow it with a product metric as in (2.6) and Haar measure λ as in (1.3).

The following theorem generalizes earlier results, which for one-parameter real groups were proved in [D1], [D2] and for one-parametric ultrametric groups were announced, with indications of the proof, in [MT3, Theorem 11.4] and [Rt4, Theorem 9.1]:

Theorem 10.1. *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 $M \subset G/\Gamma$ such that for any $y \in G/\Gamma$*

and any parametrization $\phi\colon X \to U$ of degree $\le d$ 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*

$$\frac{\lambda\left(\{x \in B \mid \phi(x)y \in M\}\right)}{\lambda(B)} \ge 1 - \beta. \tag{10.1}$$

In particular, (10.1) *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.*

10.2. Before proving the theorem we will establish the following result which is well known in the real case.

Proposition 10.2. *Let G and Γ be as in Theorem* 10.1*, and let $R(G)$ be the solvable radical of G (i.e. $R(G)$ is the direct product of the solvable radicals $R(G_v)$ of G_v, $v \in S$). Then $R(G) \cap \Gamma$ is a cocompact lattice in $R(G)$.*

Proof. Let $\widetilde{\Gamma}$ be the Zariski closure of Γ in G. In view of the Borel Density Theorem for $\mathbb{Q}_S$-algebraic groups (see [MT3, Lemma 3.1]) $\widetilde{\Gamma}$ contains all unipotent algebraic subgroups and all S-split tori of G. Therefore $G/\widetilde{\Gamma}$ is compact, and without loss of generality we may (and will) assume that Γ is Zariski dense in G. The proposition will be proved in two steps.

Step 1. First we will prove that $R(G) \cap \Gamma$ is a lattice in $R(G)$. Let $R_u(G)$ be the unipotent radical of G (i.e. $R_u(G)$ is the group of all unipotent elements in $R(G)$). Denote by $\varphi\colon G \to G/R_u(G)$, $\psi\colon G \to G/R(G)$ and $\chi\colon G/R_u(G) \to G/R(G)$ the natural S-rational homomorphisms. (We have $\chi \circ \varphi = \psi$.) Using verbatim an argument from the proof of a theorem of Zassenhaus (see [R, Section 8.14], one proves the following

Claim: There exists a neighborhood Ω of e in $G/R_u(G)$ such that if K is a bounded subset in $\varphi^{-1}(\Omega)$, then $K^{(n)} \to e$, where $K^{(0)} = K$ and $K^{(n)} = [K, K^{(n-1)}]$ for all $n \ge 1$.

Now let Δ be the Hausdorff closure in $\varphi(G)$ of the subgroup generated by $\varphi(\Gamma) \cap \Omega$. Since $\varphi(\Gamma) \cap \Omega$ is dense in $\overline{\varphi(\Gamma)} \cap \Omega$, the group Δ is open in $\overline{\varphi(\Gamma)}$. (As usually in this paper, here and hereafter $\overline{Y}$ stands for the closure of $Y \subset G$ with respect to the Hausdorff topology.) Let $K \subset \Gamma$ be a finite set such that $\varphi(K) \subset \Omega$. In view of the above claim there exists $n_0 > 0$ such that $K^{(n_0)} = \{e\}$. Therefore the group generated by K is nilpotent [R, Lemma 8.17], which implies that Δ is solvable [R, Lemma 8.4]. Since $\varphi(\Gamma)$ is Zariski dense in $\varphi(G)$, the Lie algebra of $\overline{\varphi(\Gamma)}$ is solvable and $\mathrm{Ad}\big(\varphi(G)\big)$-invariant. Therefore $Z\big(\varphi(G)\big) \cap \overline{\varphi(\Gamma)}$ is open in $\overline{\varphi(\Gamma)}$, where $Z\big(\varphi(G)\big)$ denotes the center of $\varphi(G)$. Let H be a maximal semisimple subgroup of $\varphi(G)$. Then $\varphi(G)$ is an almost direct product of H and $Z\big(\varphi(G)\big)$. Since

$Z\big(\varphi(G)\big) \cap \overline{\varphi(\Gamma)}$ is open in $\overline{\varphi(\Gamma)}$, $H \cap \overline{\varphi(\Gamma)}$ is a discrete normal subgroup of $\overline{\varphi(\Gamma)}$. Also, $H \cap \overline{\varphi(\Gamma)}$ is Zariski dense in H because the commutator of $\overline{\varphi(\Gamma)}$ is a Zariski dense subgroup of H. Therefore $\chi\big(H \cap \overline{\varphi(\Gamma)}\big)$ is discrete in $\psi(G)$ and normalized by $\psi(\Gamma)$.

Assume by contradiction that $\psi(\Gamma)$ is not discrete. Then there exists a neighborhood W of e in $\overline{\psi(\Gamma)}$ and an infinite Zariski closed subgroup L of $\psi(G)$ such that if W' is any neighborhood of e contained in W, then the Zariski closure of the subgroup generated by W' coincides with L. Now for every $g \in \chi\big(H \cap \overline{\varphi(\Gamma)}\big)$ there exists a neighborhood W_g of e in W which centralizes g. Therefore L centralizes $\chi\big(H \cap \overline{\varphi(\Gamma)}\big)$, which implies that L is central (because $\chi\big(H \cap \overline{\varphi(\Gamma)}\big)$ is Zariski dense in $\psi(G)$). This contradicts the fact that $\psi(G)$ has finite center. Therefore $\psi(\Gamma)$ is discrete and, in view of [R, Theorem 1.13], $R(G) \cap \Gamma$ is a lattice in $R(G)$.

Step 2. In order to complete the proof of the proposition, it is enough to prove that if G is solvable and Γ is a closed subgroup of G such that G/Γ admits finite G-invariant measure, then G/Γ is compact. (Recall that the cocompactness of lattices in real solvable Lie groups is a theorem of G. Mostow, see [R, ch. 3].)

For G and Γ as above, denote $G = G_\infty \times G_f$, where $G_f = \prod_{v \in S_f} G_v$. Let $\alpha_\infty \colon G \to G_\infty$ and $\alpha_f \colon G \to G_f$ be the natural projections. As in Step 1, using the Borel Density Theorem we reduce the proof to the case when Γ is Zariski dense in G. Let G_f^* be an open compact subgroup of G_f. (The group G_f^* exists because G_f is a direct product of p-adic Lie groups.) Then $(G_\infty \times G_f^*)/\big(\Gamma \cap (G_\infty \times G_f^*)\big)$ has finite $(G_\infty \times G_f^*)$-invariant measure, which, in view of the compactness of G_f^* and the cocompactness of lattices in real solvable Lie groups, implies that $G_\infty/\alpha_\infty\big(\Gamma \cap (G_\infty \times G_f^*)\big)$ is compact and, therefore, $(G_\infty \times G_f^*)/\big(\Gamma \cap (G_\infty \times G_f^*)\big)$ is compact. Assume for a moment that $G_f/\overline{\alpha_f(\Gamma)}$ is compact. Then G/G_1 is compact, where $G_1 \stackrel{\text{def}}{=} \alpha_f^{-1}\big(\overline{\alpha_f(\Gamma)}\big)$. Since $G_1 = \big(G_1 \cap (G_\infty \times G_f^*)\big)\Gamma$, we get that G_1/Γ is compact and, therefore, G/Γ is compact.

So, it remains to prove that if $G = G_f$ and Γ is Zariski dense in G, then G/Γ is compact. Let Γ_u denote the group generated by all unipotent algebraic subgroups contained in Γ. Then Γ_u is normal in G, and replacing G by G/Γ_u we reduce the proof to the case when $\Gamma_u = \{e\}$. Let P be an open subgroup of G containing $R_u(G)$ and such that $P/R_u(G)$ is compact. Then $P/(P \cap \Gamma)$ admits a finite P-invariant measure and, in view of [MT4, Lemma 1.10], $P \cap \Gamma \supset R_u(G)$. Therefore $R_u(G) = \{e\}$ and G is an abelian group. This proves that the quotient G/Γ is a locally compact topological group with finite Haar measure. Therefore G/Γ is compact. □

10.3. *Proof of Theorem* 10.1. By Proposition 10.2, $R(G) \cap \Gamma$ is a cocompact lattice in $R(G)$. Let N be a maximal subgroup in the class of all normal algebraic subgroups of G such that $N \cap \Gamma$ is a cocompact subgroup of N. Let $\mathbf{N}$ be the Zariski closure of N in $\boldsymbol{G}$. By the general structure theory of algebraic groups [Bo], $\boldsymbol{H} = \boldsymbol{G}/\mathbf{N}$ is a $\mathbb{Q}_S$-

algebraic group, $\mathbf{H} = \prod_{v\in S} \boldsymbol{H}_v$, where each $\boldsymbol{H}_v$ is a semisimple group and $\boldsymbol{H}_v(\mathbb{Q}_v)$ has no compact factors, and there exists a $\mathbb{Q}_S$-homomorphism $\varphi\colon G \to H$ (where $H = \boldsymbol{H}(\mathbb{Q}_S)$) 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 $\widetilde{\varphi}\colon G/\Gamma \to H/\Sigma$ is proper. With ϕ as in the formulation of the theorem, 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.

With the above assumptions, let $\operatorname{rank}_S G = \sum_{p\in S} \operatorname{rank}_{\mathbb{Q}_v} G_v$ be the S-rank of G. If $\operatorname{rank}_S G = 1$, 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 in G (see [T]) 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 $\boldsymbol{G}$ is a $\mathbb{Q}$-algebraic subgroup of $\mathbf{SL}(m)$, $\Gamma = G \cap \mathrm{SL}(m, \mathbb{Z}_S)$, and, after eventually replacing $\boldsymbol{G}$ by its image under a $\mathbb{Q}$-irreducible representation, we may also assume that $\boldsymbol{G}(\mathbb{Q})$ acts irreducibly on $\mathbb{Q}^m$.

Writing y in the form $g\Gamma$ for some $g \in G$ and applying Theorem 9.4 to 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

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

whenever $\phi(B)y \cap L_0 \neq \varnothing$. Choosing ε such that $\beta \geq C\varepsilon^\alpha$, we get a compact $M \stackrel{\text{def}}{=} Q_\varepsilon$ satisfying (10.1).

If $Uy \cap L_0 = \varnothing$, then in view of Theorem 9.4 there exists a proper nonzero vector subspace $V \subset \mathbb{Q}^m$ such that $\mathbb{Q}_S V$ 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 $\mathbb{Q}_S V$.) Let $\boldsymbol{P}$ be the Zariski closure of the stabilizer of V in $\boldsymbol{G}(\mathbb{Q})$ under the natural action of $\boldsymbol{G}(\mathbb{Q})$ on $\mathbb{Q}^m$. Since $\boldsymbol{G}(\mathbb{Q})$ acts irreducibly on $\mathbb{Q}^m$, $\boldsymbol{P}$ is a proper $\mathbb{Q}$-algebraic subgroup of $\boldsymbol{G}$. Therefore so is the subgroup $\boldsymbol{P}_u$ of $\boldsymbol{P}$ spanned by all unipotent elements of $\boldsymbol{P}$. It is easy to see that $g^{-1}Ug \subset \boldsymbol{P}_u(\mathbb{Q}_S)$. By the S-adic version of Borel–Harish-Chandra theorem [PR, Theorem 5.7], $\boldsymbol{P}_u(\mathcal{O}_S)$ is a lattice in $\boldsymbol{P}_u(\mathbb{Q}_S)$. Therefore the group $H = g\boldsymbol{P}_u(\mathbb{Q}_S)g^{-1}$ satisfies the requirements of the formulation of the theorem. □

We remark that to derive Theorem 1.2 from the above theorem it suffices to take $K = \mathbb{Q}$, $L = \{y\}$ 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.

10.4. Let σ be a Haar measure on U. Given a bijective parametrization $\phi\colon X \to U$, let $\phi_*\lambda$ be the pushforward of λ to U via ϕ. Note that ϕ can be chosen in such a way

that $\sigma = \phi_*\lambda$. Indeed, 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 $U_1, \dots, U_s$ of U such that the map $\psi\colon U_1 \times \cdots \times U_s \to U$, $\psi(u_1, \dots, u_s) = u_1 \cdots u_s$, is bijective, and for any i, $1 \le i < s$, the product $U_1 \cdots U_i$ 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 $\phi = \psi \circ (\phi_1, \dots, \phi_s)$ is such that $\sigma = \phi_*\lambda$, proving the above claim. Using the bijective map ϕ we translate the metric from X to U. In view of the preceding discussions, Theorem 10.1 immediately implies the following

Theorem 10.3. *Let G and Γ be as in Theorem* 10.1. *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 of G/Γ such that for any $y \in G/\Gamma$ and any unipotent algebraic subgroup U of G 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*

$$\frac{\sigma\left(\{u \in B \mid uy \in M\}\right)}{\sigma(B)} \ge 1 - \beta, \tag{10.2}$$

where σ is a Haar measure on U. In particular, (10.2) *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.*

10.5. In order to prove Theorem 1.3 we will need a version of the Birkhoff ergodic theorem. Let G, Γ, H and μ be the same as in the formulation of Theorem 1.3, and let U be a unipotent algebraic subgroup of G contained in H. 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.

Fix an imbedding $G \subset \mathrm{GL}(m, K_{\mathcal{T}})$ and a maximal unipotent subgroup $W = \prod_{v\in\mathcal{T}} W_v \subset \mathrm{GL}(m, K_{\mathcal{T}})$ which contains U. There exists an element $g = (g_v) \in \mathrm{GL}(m, K_{\mathcal{T}})$ such that $gWg^{-1} = W$, $W = \{x \in \mathrm{GL}(m, K_{\mathcal{T}}) \mid \lim_{n\to\infty} g^{-n}xg^n = e\}$, and for every $v \in \mathcal{T}$ there exists $\rho_v \in K_v$ such that $|\rho_v|_v > 1$ and the eigenvalues of $\mathrm{Ad}(g_v)_{|\,\mathrm{Lie}(W_v)}$ are pairwise different positive powers of ρ_v. It is easy to see [MT1, Proposition 2.2] that the sequence $\mathrm{Lie}(g^{-i}Ug^i)$ has a limit $\mathrm{Lie}(U_0)$ in the Grassmannian variety $\mathrm{Gr}(\mathrm{Lie}\, W) = \prod_{v\in\mathcal{T}} \mathrm{Gr}(\mathrm{Lie}\, W_v)$ of $\mathrm{Lie}\, W$, where U_0 is a $\mathrm{Int}(g)$-invariant $K_{\mathcal{T}}$-algebraic subgroup of W. Also it is known [MT1, Proposition 2.8] that there exists an $\mathrm{Int}(g)$-invariant Zariski closed subset $V \subset W$ such that the maps $U \times V \to W$, $(u, v) \to uv$, and $U_0 \times V \to W$, $(u_0, v) \to uv$, are bijective. Using the latter map, we let $\pi\colon W \to U_0$ (respectively, $\pi'\colon W \to V$) be the projection

of W onto U_0 (respectively, the projection of W onto V). Also put $p = \pi|_U$. Note that p is a homeomorphism between U and U_0. Let $\psi = p^{-1} \circ \mathrm{Int}(g) \circ p$. Then $\psi : U \to U$ acts as an expansion on U and $\lim_{i\to\infty} \psi^{-i}(u) = e$ for all $u \in U$. In particular, if B is a relatively compact neighborhood of e in U, then

$$U = \bigcup_{i\geq 0} B_i, \tag{10.3}$$

where $B_0 = B$ and $B_{i+1} = \psi(B_i)$ for all $i \geq 0$. Further on we fix a set B as above.

Let $u \in U, \varepsilon > 0$, and let $\rho(\cdot,\cdot)$ be a metric on U. Since ψ^{-1} is a contracting map, there exists $i_0 > 0$ such that if $i > i_0$ and $x \in B_i$, then $\rho\big(\psi^{-i}(ux), \psi^{-i}(x)\big) < \varepsilon$. Therefore $\lim_{i\to\infty} \sigma\big(B \triangle \psi^{-i}(uB_i)\big) = 0$, where $C \triangle D$ stands for the symmetric difference between two sets C and D. Since the Jacobian of ψ is constant, we get

$$\lim_{i\to\infty} \frac{\sigma(B_i \triangle uB_i)}{\sigma(B_i)} = \lim_{i\to\infty} \frac{\sigma\big(B \triangle \psi^{-i}(uB_i)\big)}{\sigma(B)} = 0. \tag{10.4}$$

Let us show that there exists a compact $L \subset W$ such that $\mathrm{Int}(g^{-i})(B_i) \subset L$ for all i. Otherwise, it would follow from the definition of B_i and the $\mathrm{Int}(g)$-invariancy of U_0 and V that there exists a sequence $u_{k_i} \in B_{k_i}$ such that $\mathrm{Int}(g^{-k_i})\big(\pi(u_{k_i})\big) \in \pi(B)$ for all i and the sequence $\mathrm{Int}(g^{-k_i})\big(\pi(u_{k_i})\big)$ is unbounded. Since $\mathrm{Lie}(U_0) = \lim_{i\to\infty} \mathrm{Lie}(g^{-i} U g^i)$, the latter implies that $U_0 \cap V \neq \{e\}$, a contradiction. We can choose L such that $L = L^{-1}$. So, if $x, y \in B_i$ then $\mathrm{Int}(g^{-i})(x^{-1}y) \in L \cdot L$. We have

$$\pi\big(\mathrm{Int}(g^{-i})(x^{-1}y)\big) = \mathrm{Int}(g^{-i})\big(p(x^{-1}y)\big) \in \pi(L \cdot L).$$

Therefore

$$\psi^{-i}(B_i^{-1}B_i) \subset p^{-1}\big(\pi(L \cdot L)\big).$$

Since $p^{-1}\big(\pi(L \cdot L)\big)$ is compact, we obtain

$$\sup_i \frac{\sigma(B_i^{-1}B_i)}{\sigma(B_i)} = \sup_i \frac{\sigma\big(\psi^{-i}(B_i^{-1}B_i)\big)}{\sigma(B)} < \infty. \tag{10.5}$$

In view of (10.3), (10.4) and (10.5), the following result directly follows from [Te, Corollary 3.2, Ch. 6] (see also [MT3, Proposition 7.1]):

Proposition 10.4. *With the above notation and assumptions, let f be a continuous μ-integrable function on G/Γ. Then*

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

for μ-almost all $y \in G/\Gamma$. Furthermore, the limit function

$$f^*(y) \stackrel{\text{def}}{=} \int_{G/\Gamma} f(z)\, d\mu_{a(y)}(z) \tag{10.6}$$

is μ-integrable and U-invariant. □

10.6. We now proceed with the

Proof of Theorem 1.3. 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 H-invariant and positive μ-almost everywhere. 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 of G contained 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]). Therefore, in view of the Mautner phenomenon for products of real and p-adic Lie groups [MT4, Proposition 2.1], if $f \in L^1(G/\Gamma, \mu)$ is U-invariant, then there exists an H-invariant μ-integrable function on G/Γ which coincides with f almost everywhere. So, it is enough to prove the theorem for $H = U$.

Let f be a positive continuous μ-integrable function on G/Γ. Applying Proposition 10.4, we get a U-invariant function f^* defined by formula (10.6). It is enough to prove that $f^* > 0$ μ-a.e. Note that

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

where ψ is as in §10.5. From Theorem 10.1 and the facts that $\sigma = \phi_*\lambda$ for some rational parametrization ϕ of U (see §10.4) and that all $\psi^n \circ \phi$ have the same degree (because ψ is linear), it follows that there exists a compact $M \subset G/\Gamma$ such that for any positive n

$$\frac{\sigma\big(\{g \in B \mid \psi^n(g)y \in M\}\big)}{\sigma(B)} > \frac{1}{2}.$$

Since f is positive, the above formula implies that $f^*(y) > 0$ μ-a.e., which completes the proof of the theorem. □

Remark 10.5. Using the methods from [MT4] it is easy to see that Theorems 10.1, 10.3, 1.3 and Proposition 10.4 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.

11. S-arithmetic Diophantine approximation

11.1. In this section we present a motivation for the definitions of VWA and VWMA vectors given in the introduction, and state the main Diophantine result of the present paper, of which Theorem 1.4 is a special case. Here and for the rest of the paper we fix a set S of cardinality ℓ consisting of distinct normalized valuations of $\mathbb{Q}$ (and not necessarily containing ∞) , and let

We will interpret elements

$$\boldsymbol{y} = (\boldsymbol{y}^{(v)})_{v\in S} = (y_1, \dots, y_n)$$

of $\mathbb{Q}_S^n$, where $\boldsymbol{y}^{(v)} = (y_1^{(v)}, \dots, y_n^{(v)}) \in \mathbb{Q}_v^n$ and $y_i = (y_i^{(v)}) \in \mathbb{Q}_S$, as linear forms on $\mathbb{Q}_S^n$, and will study their values $\boldsymbol{y}\cdot\boldsymbol{q} = y_1q_1 + \cdots + y_nq_n$ at integer points $\boldsymbol{q} = (q_1, \dots, q_n)$. 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) $\boldsymbol{y} \in \mathbb{Q}_S^n$ by rational vectors. See [KLW] and [KM, **?**] for a discussion of this set-up in the real case.

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

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

so that $q_0 + y_1q_1 + \cdots + y_nq_n = q_0 + \boldsymbol{y}\cdot\boldsymbol{q}$ is written as $\tilde{\boldsymbol{y}}\cdot\tilde{\boldsymbol{q}}$. 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.

11.2. 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 $\boldsymbol{y} \in \mathbb{Q}_S^n$ and any $N > 0$ one is guaranteed to have two different integer vectors $\tilde{\boldsymbol{q}}_1, \tilde{\boldsymbol{q}}_2$ of norm $\le N$ such that $\tilde{\boldsymbol{y}}\cdot\tilde{\boldsymbol{q}}_1$ and $\tilde{\boldsymbol{y}}\cdot\tilde{\boldsymbol{q}}_2$ are at most $\text{const}(\boldsymbol{y})N^{-\beta}$ apart.

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

$$|\tilde{\boldsymbol{y}}^{(v)}\cdot\tilde{\boldsymbol{q}}|_v \le \max(\|\boldsymbol{y}^{(v)}\|_v, 1) \tag{11.1}$$

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

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

$$|\tilde{\boldsymbol{y}}\cdot\tilde{\boldsymbol{q}}| \le \text{const}(\boldsymbol{y})\cdot N^{-\frac{n+1}{\ell}}.$$

- If $v = \infty$, there is clearly no universal upper bound similar to (11.1). However one can for any given $\boldsymbol{q} \in \mathbb{Z}^n$ choose $q_0 \in \mathbb{Z}$ such that

$$|\tilde{\boldsymbol{y}}^{(\infty)} \cdot \tilde{\boldsymbol{q}}|_\infty = |q_0 + \boldsymbol{y}^{(\infty)} \cdot \boldsymbol{q}|_\infty \le 1.$$

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

$$|\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}}| = |q_0 + \boldsymbol{y} \cdot \boldsymbol{q}| \le \text{const}(\boldsymbol{y}) N^{-\frac{n}{\ell}} \quad \text{for some } q_0 \in \mathbb{Z}.$$

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

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

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

$n + i_S$. 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 $\boldsymbol{y} \in \mathbb{Q}_S^n$ (with respect to Haar measure λ_S on $\mathbb{Q}_S^n$). Thus it is natural to say that $\boldsymbol{y} \in \mathbb{Q}_S^n$ 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 $\tilde{\boldsymbol{q}} \in \mathbb{Z}^{n+1}$ to

$$|\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}}|^\ell \le \|\tilde{\boldsymbol{q}}\|^{-(n+i_S)(1+\varepsilon)}. \tag{11.2}$$

Note that in the case when $\infty \in S$, any solution $\tilde{\boldsymbol{q}}$ of (11.2) automatically satisfies

$$|q_0| \le 1 + n\|\boldsymbol{y}^{(\infty)}\|_\infty \|\boldsymbol{q}\|, \tag{11.3}$$

and so $\|\tilde{\boldsymbol{q}}\|$ on the right-hand side of (11.2) can be replaced by $\|\tilde{\boldsymbol{q}}\|$, agreeing with (1.4).

11.3. The next step is to define *very well multiplicatively approximable*, or *VWMA*, vectors $\boldsymbol{y} \in \mathbb{Q}_S^n$. To do this, one would like to replace the left-hand side of (11.2) by the product of norms of all the components of $\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}}$, and the norm of $\tilde{\boldsymbol{q}}$ in (11.2) with the geometric mean of its coordinates. However one needs to be careful and keep in mind the dichotomy in the Dirichlet-principle argument.

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

$$c(\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}}) \le \Pi_+(\tilde{\boldsymbol{q}})^{-(1+\varepsilon)}. \tag{11.4}$$

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

$$c(\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}}) = c(q_0 + \boldsymbol{y} \cdot \boldsymbol{q}) \le \Pi_+(\boldsymbol{q})^{-(1+\varepsilon)} \tag{11.4$'$}$$

holds for some $q_0 \in \mathbb{Z}$. Indeed, this coincides with the standard definition when $S = \{\infty\}$, cf. [KM]. However it is not hard to see that, whenever S contains both finite and infinite valuations, for any $\varepsilon > 0$ the set of $\boldsymbol{y} \in \mathbb{Q}_S^n$ for which (11.4$'$) admits infinitely many solutions has full measure. Indeed, the trouble here comes from the fact that an upper estimate for $c(\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}})$ does not imply a bound similar to (11.3), that is, a bound on $|q_0|$ in terms of $\boldsymbol{q}$. And one can easily show that for any fixed $\boldsymbol{q}$ the set of $\boldsymbol{y} \in \mathbb{Q}_S^n$ for which there exists q_0 satisfying (11.4$'$) has full measure.

It follows that in order to achieve a multiplicative analogue of (11.4) 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 $\boldsymbol{q}$. Namely, in the case $\infty \in S$ we will define $\boldsymbol{y} \in \mathbb{Q}_S^n$ to be *VWMA* if for some $\varepsilon > 0$ there are infinitely many solutions $\boldsymbol{q}$ to

$$c(\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}}) \le \Pi_+(\boldsymbol{q})^{-(1+\varepsilon)} |q_0|_+^{-\varepsilon}. \tag{11.4$_\infty$}$$

Put together with (11.4), the latter inequality can be written in the form (1.5), or, equivalently, in the unified form as

$$c(\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}}) \le \Pi_+(\boldsymbol{q})^{-(1+\varepsilon)} |q_0|_+^{-(i_S+\varepsilon)}. \tag{11.5}$$

Several remarks are in order. First, note that in the case $S = \{\infty\}$ (11.4$_\infty$) can be replaced by (11.4$'$), perhaps with a slightly different value of ε: indeed, any solution of (11.5) will satisfy (11.3), hence $|q_0|_+$ is bounded from above by some power of $\Pi_+(\boldsymbol{q})$. Similarly, it can be easily seen that infinitely many solutions to (11.2) imply infinitely many solutions to (11.5) (with the same ε if $\infty \notin S$, and, in view of (11.3) and $\Pi_+(\boldsymbol{q}) \ge \|\boldsymbol{q}\|$, perhaps with a different ε if $\infty \in S$). 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 11.1 below.

11.4. Recall that a measure μ on $\mathbb{Q}_S^n$ is called 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 11.1. *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 $\boldsymbol{f} = (\boldsymbol{f}^{(v)})_{v\in S}$, where $\boldsymbol{f}^{(v)}$ are continuous maps from X_v to $\mathbb{Q}_v^n$ such that pairs $(\boldsymbol{f}^{(v)}, \mu_v)$ are good and nonplanar at μ_v-almost every point of X_v. Then $\boldsymbol{f}_*\mu$ is strongly extremal.*

It is clear from Theorem 4.3, as well as from Examples 2.6 and 2.7, that Theorem 1.4 is a special case of the above result.

11.5. Let us also remark that Theorem 11.1 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 μ-a.e. $\boldsymbol{y} \in F^n$ there exist a neighborhood V of $\boldsymbol{y}$ 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 §5.2, one easily observes that μ is decaying if and only if (Id, μ), where Id is the identity map $F^n \to F^n$, is good at μ-almost every point. It is also not hard to see that the nonplanarity of μ forces (Id, μ) to be nonplanar at μ-almost every point (the converse is true under the additional assumption that μ is decaying).

It is now clear that Theorem 11.1 immediately implies

Corollary 11.2. *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 11.1. (As was mentioned before, our proof is also a generalization of the argument from [KLW], which, in turn, generalizes the one from [KM].)

11.6. 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 inducing the metric "dist" on F^n, and say that a map $\boldsymbol{h}\colon F^n \to F^n$ is a *contracting similitude* with *contraction rate* ρ if $0 < \rho < 1$ and

$$\operatorname{dist}\big(\boldsymbol{h}(\boldsymbol{x}), \boldsymbol{h}(\boldsymbol{y})\big) = \rho \operatorname{dist}(\boldsymbol{x}, \boldsymbol{y}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in F^n.$$

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

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

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

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

and

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

J. Hutchinson [H, §5.3] proved[4] that if $\boldsymbol{h}_i$, $i = 1, \dots, m$, are contracting similitudes with contraction rates ρ_i satisfying the open set condition, and if $s > 0$ is the unique solution of $\sum_i \rho_i^s = 1$, called the *similarity dimension* of the family $\{\boldsymbol{h}_i\}$, then the s-dimensional Hausdorff measure $\mathcal{H}^s$ of Q is positive and finite. Let us also say that the family $\{\boldsymbol{h}_i\}$ is *irreducible* if there does not exist a finite $\{\boldsymbol{h}_i\}$-invariant collection of proper affine subspaces of F^n. The proof of [KLW, Theorem 2.3] applies verbatim and yields

Proposition 11.3. *For any completion $\mathbb{Q}_v$ of $\mathbb{Q}$, let $\{\boldsymbol{h}_1, \dots, \boldsymbol{h}_m\}$ 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 limit set. 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

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

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.

11.7. We conclude this section with the following modification of Theorem 11.1:

Theorem 11.4. *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 $\boldsymbol{f} = (\boldsymbol{f}^{(v)})_{v \in S}$, where $\boldsymbol{f}^{(v)}$ 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:*

$$\begin{gathered} \text{for some } C_v, \alpha_v > 0, \text{ any linear combination of } 1, f_1^{(v)}, \dots, f_n^{(v)} \\ \text{is } (C_v, \alpha_v)\text{-good on } B(x_v, 3^{n+1}r) \text{ with respect to } \mu_v, \end{gathered} \tag{11.6}$$

and

$$\begin{gathered} \text{the restrictions of } 1, f_1^{(v)}, \dots, f_n^{(v)} \text{ to } B_v \cap \operatorname{supp} \mu_v \\ \text{are linearly independent over } \mathbb{Q}_v. \end{gathered} \tag{11.7}$$

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

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

Reduction of Theorem 11.1 *to Theorem* 11.4. Let X and $\boldsymbol{f}$ be as in Theorem 11.1. First note that, replacing X_v by appropriate neighborhoods of its μ_v-generic points for each v, one can without loss of generality assume that μ is uniformly Federer. Then for any v, since $(\boldsymbol{f}^{(v)}, \mu_v)$ is 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 $(\boldsymbol{f}^{(v)}, \mu_v)$ is 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 (11.7) holds, and the conclusion follows. □

In the next two sections we present the proof of Theorem 11.4, 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.

12. Proof of Theorem 11.4 for $\infty \notin S$

12.1. In order to prove Theorem 11.4, we are going to dynamically interpret 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 the section we will work with

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

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

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

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

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

$$\tilde{t} = \sum_{i=0}^{n} t_i \quad \text{and} \quad t = \sum_{i=1}^{n} t_i \tag{12.1}$$

(this convention will be used throughout the next two sections, so that whenever t and $\boldsymbol{t}$, or $\tilde{t}$ and $\tilde{\boldsymbol{t}}$, appear in the same context, (12.1) will be assumed). Then, given $\tilde{\boldsymbol{t}}$

as above and another vector $\boldsymbol{s} = (s_1, \dots, s_\ell) \in \mathbb{Z}_+^\ell$, define $g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} \in \mathrm{GL}(n+1, \mathcal{R})$ by

$$(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}})^{(p_j)} = \begin{pmatrix} p_j^{-s_j} & 0 \\ 0 & I_n \end{pmatrix},\ j = 1, \dots, \ell, \quad (g_{\boldsymbol{s},\tilde{\boldsymbol{t}}})^{(\infty)} = \mathrm{diag}(e^{-t_0}, e^{-t_1}, \dots, e^{-t_n}). \tag{12.2}$$

The next lemma shows how a good approximation for $\boldsymbol{y}$ in the sense of (11.1) gives rise to a translation of $u_{\boldsymbol{y}}\mathcal{D}^{n+1}$ by $g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}$ for some $\boldsymbol{s}, \tilde{\boldsymbol{t}}$, so that $\delta(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{y}}\mathcal{D}^{n+1})$ is small. This allows one to use Theorem 9.3 to derive the needed measure estimate.

Lemma 12.1. *Let $\varepsilon > 0$, $\boldsymbol{y} \in \mathbb{Q}_S^n$ and $\tilde{\boldsymbol{q}} \in \mathbb{Z}^{n+1}$ be such that* (11.4) *holds. For $i = 0, 1, \dots, n$ define $t_i > 0$ by*

$$|q_i|_+ = \Pi_+(\tilde{\boldsymbol{q}})^{-\frac{\varepsilon}{n+1}} e^{t_i}, \tag{12.3a}$$

and let

$$\gamma = \frac{\varepsilon}{(n+1)(1+\varepsilon)}. \tag{12.3b}$$

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

$$\delta(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{y}}\mathcal{D}^{n+1}) \le \sqrt{n+1}\, e^{-\gamma\tilde{t}} \tag{12.3c}$$

and

$$\prod_{j=1}^{\ell} p_j^{s_j} \le e^{\tilde{t}} < \prod_{j=1}^{\ell} p_j^{s_j+1}. \tag{12.3d}$$

Proof. Multiplying equalities (12.3a), we get

$$e^{\tilde{t}} = \Pi_+(\tilde{\boldsymbol{q}})^{1+\varepsilon} \tag{12.3e}$$

and

$$e^{-t_i}|q_i| \le e^{-t_i}|q_i|_+ \underset{(12.3a)}{\le} \Pi_+(\tilde{\boldsymbol{q}})^{-\frac{\varepsilon}{n+1}} \underset{(12.3be)}{=} e^{-\gamma\tilde{t}},\ i = 0, 1, \dots, n,$$

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

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

$$p_j^{s_j} \le \min\left(\frac{e^{\tilde{t}}}{\prod_{i=1}^{j-1} p_i^{s_i}}, \frac{1}{|\tilde{\boldsymbol{y}}^{(p_j)} \cdot \tilde{\boldsymbol{q}}|_{p_j}} \right) < p_j^{s_j+1} \tag{12.4}$$

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

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

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

It remains to check that inequalities (12.3d) are satisfied. Taking $j = \ell$ in (12.4) immediately implies the lower estimate. To prove the upper estimate, let us consider two cases:

- If for some j the minimum in (12.4) 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 $|\tilde{\boldsymbol{y}}^{(p_j)} \cdot \tilde{\boldsymbol{q}}|_{p_j} > p_j^{-(s_j+1)}$ for all j, and to derive the desired estimate it remains to notice that (11.4), in view of (12.3e), can be rewritten as $c(\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}}) = \prod_{j=1}^{\ell} |\tilde{\boldsymbol{y}}^{(p_j)} \cdot \tilde{\boldsymbol{q}}|_{p_j} \le e^{-\tilde{t}}$. □

Corollary 12.2. *Assume that $\boldsymbol{y} \in \mathbb{Q}_S^n$ is VWMA. Then for some $c, \gamma > 0$ there are infinitely many $\tilde{\boldsymbol{t}} \in \mathbb{Z}_+^{n+1}$ and $\boldsymbol{s} \in \mathbb{Z}_+^{\ell}$ such that*

$$e^{-(n+1)} \prod_{j=1}^{\ell} p_j^{s_j} \le e^{\tilde{t}} < \prod_{j=1}^{\ell} p_j^{s_j+1} \tag{12.5a}$$

and

$$\delta(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{y}} \mathcal{D}^{n+1}) \le c e^{-\gamma \tilde{t}}. \tag{12.5b}$$

Proof. By definition, for some $\varepsilon > 0$ there are infinitely many solutions $\tilde{\boldsymbol{q}} \in \mathbb{Z}^{n+1}$ of (11.4). Therefore, by the above lemma and with γ as in (12.3b), there exists an unbounded set of $\tilde{\boldsymbol{t}} \in \mathbb{R}_+^{n+1}$ such that (12.3c) holds for some $\boldsymbol{s} \in \mathbb{Z}_+^{\ell}$ satisfying (12.3d). Denote by $[\tilde{\boldsymbol{t}}]$ the vector consisting of integer parts of t_i, then clearly the ratio of $\delta(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{y}} \mathcal{D}^{n+1})$ and $\delta(g_{\boldsymbol{s},[\tilde{\boldsymbol{t}}]} u_{\boldsymbol{y}} \mathcal{D}^{n+1})$ is bounded from above by some uniform constant. Thus, replacing $\tilde{\boldsymbol{t}}$ by $[\tilde{\boldsymbol{t}}]$, for some $c > 0$ one gets infinitely many solutions $\tilde{\boldsymbol{t}} \in \mathbb{Z}_+^{n+1}$ of (12.5b), with $e^{\tilde{t}}$ being smaller than before by at most a factor of e^{n+1}, hence (12.5a). □

Corollary 12.3. *Let X be a Besicovitch metric space and μ a uniformly Federer measure on X. Suppose we are given a continuous map $\boldsymbol{f}\colon X \to \mathbb{Q}_S^n$ 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 $\boldsymbol{s} \in \mathbb{Z}_+^{\ell}$, $\tilde{\boldsymbol{t}} \in \mathbb{Z}_+^{n+1}$ satisfying* (12.5a)*, one has*

$$\text{the function } x \mapsto \operatorname{cov}(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{f}(x)} \Delta) \text{ is } (C,\alpha)\text{-good on } B(x_0, 3^{n+1} r) \text{ w.r.t. } \mu, \tag{12.6a}$$

and

$$\sup_{x\in B\cap \operatorname{supp}\mu} \operatorname{cov}\left(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}u_{\boldsymbol{f}(x)}\Delta\right) \geq \rho. \tag{12.6b}$$

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

Proof. Applying Theorem 9.3, with $h(x) = g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}u_{\boldsymbol{f}(x)}$ and $m = n+1$, we conclude that

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

whenever $ce^{-\gamma\tilde{t}} \leq \rho$ and (12.5a) holds. Note that for fixed $\tilde{\boldsymbol{t}}$, the number of different $\boldsymbol{s}\in\mathbb{Z}_+^\ell$ satisfying (12.5a) is at most const $\cdot\tilde{t}^{\,\ell-1}$. Therefore the sum (over all integer $\boldsymbol{s},\tilde{\boldsymbol{t}}$ for which (12.5a) holds) of measures of sets $\{x\in B \mid \delta(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}u_{\boldsymbol{f}(x)}\mathcal{D}^{n+1}) < ce^{-\gamma\tilde{t}}\}$ 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 $\boldsymbol{s},\tilde{\boldsymbol{t}}$ to (12.5a), (12.5b). Corollary 12.2 then implies that $\boldsymbol{f}(x)$ is not VWMA for μ-a.e. $x\in B$. □

12.2. We are now ready for the

Proof of Theorem 11.4*, 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 12.3. Thus we need to have explicit expressions for functions $x\mapsto \operatorname{cov}\left(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}u_{\boldsymbol{f}(x)}\Delta\right)$.

Using Proposition 8.1 and Lemma 8.2, one can associate to any nonzero submodule $\Delta\subset\mathcal{D}^{n+1}$ of rank r an element $\boldsymbol{w}$ of $\bigwedge^r(\mathcal{D}^{n+1})$ such that $\operatorname{cov}(\Delta) = c(\boldsymbol{w})$ and $\operatorname{cov}(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}u_{\boldsymbol{f}(x)}\Delta) = c(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}u_{\boldsymbol{f}(x)}\boldsymbol{w})$. It will be convenient to use the standard basis $\boldsymbol{e}_0,\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$ of $\mathcal{R}^{n+1}$, where

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

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

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

- $u_{\boldsymbol{f}(x)}^{(\infty)}$ is trivial, and each $\boldsymbol{e}_I^{(\infty)}$ is an eigenvector of $(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}})^{(\infty)}$ with eigenvalue e^{-t_I}, where $t_I\stackrel{\text{def}}{=}\sum_{i\in I}t_i$;

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

$$(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{f}(x)} \boldsymbol{e}_I)^{(p_j)} = \begin{cases} p_j^{-s_j} \boldsymbol{e}_I^{(p_j)} & \text{if } 0 \in I, \\ \boldsymbol{e}_I^{(p_j)} + p_j^{-s_j} \sum_{i \in I} \pm f_i^{(p_j)}(x) \boldsymbol{e}_{I \cup \{0\} \setminus \{i\}}^{(p_j)} & \text{otherwise.} \end{cases} \tag{12.7}$$

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

$$(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{f}(x)} \boldsymbol{w})^{(p_j)} = \sum_{0 \notin I} w_I \boldsymbol{e}_I^{(p_j)} + p_j^{-s_j} \sum_{0 \in I} \Big(w_I + \sum_{i \notin I} \pm w_{I \cup \{i\} \setminus \{0\}} f_i^{(p_j)}(x) \Big) \boldsymbol{e}_I^{(p_j)} \tag{12.8}$$

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

In particular, real components of all the coordinates of $g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{f}(x)} \boldsymbol{w}$ are constant, and p_j-adic components are linear combinations of $1, f_1^{(p_j)}, \dots, f_n^{(p_j)}$. Condition (12.6a) then immediately follows from Lemma 3.1 (b), (c), (11.6) and Corollary 3.3. On the other hand, for any $j = 1, \dots, \ell$ one can use (11.7) and the compactness of the unit sphere in $\mathbb{Q}_{p_j}^{n+1}$ to find $\rho_j > 0$ such that for any $\boldsymbol{a} = (a_0, a_1, \dots, a_n) \in \mathbb{Q}_{p_j}^{n+1}$ one has

$$\sup_{x \in B_j \cap \operatorname{supp} \mu_j} |a_0 + a_1 f_1^{(p_j)}(x) + \dots + a_n f_n^{(p_j)}(x)|_{p_j} \geq \rho_j \|\boldsymbol{a}\|_{p_j}. \tag{12.9}$$

It remains to notice that all the components of $\boldsymbol{w}$ necessarily appear in the second sum in (12.8) (that is, the sum of terms with $0 \in I$). Therefore (12.8) and (12.9) imply

$$\sup_{x \in B_j \cap \operatorname{supp} \mu_j} \|(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{f}(x)} \boldsymbol{w})^{(p_j)}\|_{p_j} \geq \rho_j {p_j}^{s_j} \max_I |w_I|_{p_j}, \tag{12.10}$$

and hence

$$\begin{aligned} \sup_{x \in B \cap \operatorname{supp} \mu} c(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{f}(x)} \boldsymbol{w}) &\geq \Big(\prod_{j=1}^{\ell} \rho_j {p_j}^{s_j} \max_I |w_I|_{p_j} \Big) \max_I e^{-t_I} |w_I|_\infty \\ &\geq \Big(\prod_{j=1}^{\ell} \rho_j \Big) e^{-\tilde{t}} \Big(\prod_{j=1}^{\ell} {p_j}^{s_j} \Big) \max_I c(w_I) \underset{(12.5a)}{\geq} \prod_{j=1}^{\ell} \frac{\rho_j}{p_j}. \end{aligned}$$

Condition (12.6b) is thus established, and the theorem follows. □

13. Proof of Theorem 11.4 for $\infty \in S$

13.1. 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 $u_{\boldsymbol{y}} \in \mathrm{GL}^1(n+1, \mathcal{R})$ can be simply defined by $u_{\boldsymbol{y}} \stackrel{\text{def}}{=} \left(\begin{smallmatrix} 1 & \boldsymbol{y} \\ 0 & I_n \end{smallmatrix}\right)$, so that

$$u_{\boldsymbol{y}} \mathcal{D}^{n+1} = \left\{ \begin{pmatrix} \tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}} \\ \boldsymbol{q} \end{pmatrix} \mid \tilde{\boldsymbol{q}} \in \mathcal{D}^{n+1} \right\}.$$

The definition (12.2) of the diagonal element $g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} \in \mathrm{GL}(n+1, \mathcal{R})$ given in the previous section will still be valid, except for $\boldsymbol{s}$ now having $\ell - 1$ components.

Now let us split the set of VWMA vectors into two parts: say that a very well multiplicatively approximable $\boldsymbol{y} \in \mathbb{Q}_S^n$ is *VWMA*$_\le$ if for some positive ε there are infinitely many solutions $\tilde{\boldsymbol{q}}$ to (11.4_∞) satisfying

$$|q_0| \le \big(1 + n\|\boldsymbol{y}^{(\infty)}\|_\infty\big)\|\boldsymbol{q}\|, \tag{13.1}$$

and that it is *VWMA*$_>$ otherwise. 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 11.4 to take care of the second case.

Here is a replacement for Lemma 12.1.

Lemma 13.1. *Let $\varepsilon > 0$, $\boldsymbol{y} \in \mathbb{Q}_S^n$ and $\tilde{\boldsymbol{q}} \in \mathbb{Z}^{n+1}$ be such that (11.4_∞) and (13.1) hold. For $i = 1, \dots, n$ define $t_i > 0$ by*

$$|q_i|_+ = \Pi_+(\boldsymbol{q})^{-\frac{\varepsilon}{n+1}} e^{t_i}, \tag{13.2a}$$

and let

$$\gamma = \frac{\varepsilon}{n + 1 + n\varepsilon}. \tag{13.2b}$$

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

$$\delta(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{y}} \mathcal{D}^{n+1}) \le \sqrt{n+1} e^{-\gamma t}, \tag{13.2c}$$

$$-t \le t_0 \le t + \ln\big(1 + 2n\|\boldsymbol{y}^{(\infty)}\|_\infty\big), \tag{13.2d}$$

and

$$\prod_{j=1}^{\ell-1} p_j^{s_j} \le e^{\tilde{t}} < \prod_{j=1}^{\ell-1} p_j^{s_j+1}. \tag{13.2e}$$

Proof. As in the proof of Lemma 12.1, we consider the product of equalities (13.2a), namely

$$e^t = \Pi_+(\boldsymbol{q})^{1+\frac{n}{n+1}\varepsilon}, \tag{13.2f}$$

and then write

$$e^{-t_i}|q_i| \le e^{-t_i}|q_i|_+ \underset{(13.2a)}{\le} \Pi_+(\boldsymbol{q})^{-\frac{\varepsilon}{n+1}} \underset{(13.2bf)}{=} e^{-\gamma t},\ i = 1, \dots, n. \tag{13.2g}$$

After that define t_0 by

$$e^{-t_0} \stackrel{\text{def}}{=} \min\left(e^t, \frac{e^{-\gamma t}}{|\tilde{\boldsymbol{y}}^{(\infty)} \cdot \tilde{\boldsymbol{q}}|_\infty}\right). \tag{13.3}$$

It follows that $|e^{-t_0}\tilde{\boldsymbol{y}}^{(\infty)} \cdot \tilde{\boldsymbol{q}}|_\infty \le e^{-\gamma t}$, hence $\|(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}})^{(\infty)} u_{\boldsymbol{y}}^{(\infty)} \tilde{\boldsymbol{q}}\|_\infty \le \sqrt{n+1}e^{-\gamma t}$. The lower estimate in (13.2d) is immediate from (13.3), while the upper estimate clearly holds if the minimum in (13.3) is equal to e^t, and otherwise one has

$$e^{-t_0} = \frac{e^{-\gamma t}}{|q_0 + \boldsymbol{y}^{(\infty)} \cdot \boldsymbol{q}|_\infty} \underset{(13.1)}{\ge} \frac{e^{-\gamma t}}{\left(1 + 2n\|\boldsymbol{y}^{(\infty)}\|_\infty\right)\|\boldsymbol{q}\|} \underset{(13.2g)}{\ge} \frac{e^{-t}}{1 + 2n\|\boldsymbol{y}^{(\infty)}\|_\infty}.$$

Now that all the components of $\tilde{\boldsymbol{t}}$ are chosen, we can define s_j, $j = 1, \dots, \ell - 1$, as in (12.4). After that one can verify, following the lines of the proof of Lemma 12.1, that

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

for each j, so that $c(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{y}} \tilde{\boldsymbol{q}}) \le \sqrt{n+1}e^{-\gamma t}$, and that inequalities (13.2e) are satisfied. □

Corollary 13.2. *Assume that $\boldsymbol{y} \in \mathbb{Q}_S^n$ is VWMA$_\le$. Then for some $c_0, c, \gamma > 0$ there are infinitely many $\tilde{\boldsymbol{t}} \in \mathbb{Z} \times \mathbb{Z}_+^n$ and $\boldsymbol{s} \in \mathbb{Z}_+^{\ell-1}$ satisfying*

$$-t - 1 \le t_0 \le t + c_0, \tag{13.4a}$$

$$e^{-(n+1)} \prod_{j=1}^{\ell-1} p_j^{s_j} \le e^{\tilde{t}} < \prod_{j=1}^{\ell-1} p_j^{s_j+1}, \tag{13.4b}$$

and

$$\delta(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{y}} \mathcal{D}^{n+1}) \le ce^{-\gamma t}. \tag{13.4c}$$

Proof. By definition, for some $\varepsilon > 0$ there are infinitely many solutions $\tilde{\boldsymbol{q}}$ to (11.4$_\infty$) and (13.1). Therefore, by the above lemma and with γ as in (13.2b), there exists an unbounded set of $\boldsymbol{t} \in \mathbb{R}_+^n$ such that inequalities (13.2c)–(13.2e) hold for some $t_0 \in \mathbb{R}$ and $\boldsymbol{s} \in \mathbb{Z}_+^{\ell-1}$. The rest of the proof of Corollary 12.2 applies verbatim. □

Corollary 13.3. *Let X be a Besicovitch metric space and μ a uniformly Federer measure on X. Suppose we are given a continuous map $\boldsymbol{f}\colon X \to \mathbb{Q}_S^n$ with the following 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* (12.6a), (12.6b) *hold for any $\Delta \in \mathfrak{P}(\mathcal{D}, n+1)$ and any $\boldsymbol{s} \in \mathbb{Z}_+^{\ell-1}$, $\boldsymbol{t} \in \mathbb{Z} \times \mathbb{Z}_+^n$ satisfying* (13.4a), (13.4b). *Then $\boldsymbol{f}(x)$ is not VWMA$_<$ for μ-a.e. $x \in X$.*

Proof. An application of Theorem 9.3, again with $h(x) = g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{f}(x)}$ and $m = n+1$, yields

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

whenever $ce^{-\gamma t} \le \rho$ and (13.4a, b) hold. Now observe that for fixed $\boldsymbol{t}$, the number of different $t_0 \in \mathbb{Z}$ and $\boldsymbol{s} \in \mathbb{Z}_+^{\ell-1}$ satisfying (13.4a, b) is at most const $\cdot t^{\ell-1}$. Therefore for any $c, \gamma > 0$ the sum (over all integers $\boldsymbol{s}, \tilde{\boldsymbol{t}}$ for which inequalities (13.4a, b) hold) of measures of sets $\{x \in B \mid \delta(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}} u_{\boldsymbol{f}(x)} \mathcal{D}^{n+1}) < ce^{-\gamma t}\}$ 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 $\boldsymbol{s}, \tilde{\boldsymbol{t}}$ to (13.4a)–(13.4c) for any $c, \gamma > 0$. Corollary 13.2 then implies that $\boldsymbol{f}(x)$ is not VWMA$_<$ for μ-a.e. $x \in B$. □

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

Lemma 13.4. *Assume that $\boldsymbol{y} = \left(\boldsymbol{y}^{(v)}\right)_{v \in S}$ is VWMA$_>$; then $\left(\boldsymbol{y}^{(v)}\right)_{v \in S_f}$ is VWMA.*

Proof. By assumption, there exist infinitely many solutions $\tilde{\boldsymbol{q}}$ of (11.4$_\infty$) for which (13.1) fails. For each of them one can write

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

Therefore one has

$$\begin{aligned} \prod_{v \in S_f} |\tilde{\boldsymbol{y}}^{(v)} \cdot \tilde{\boldsymbol{q}}|_v = \frac{c(\tilde{\boldsymbol{y}} \cdot \tilde{\boldsymbol{q}})}{|\tilde{\boldsymbol{y}}^{(\infty)} \cdot \tilde{\boldsymbol{q}}|_\infty} &\le \Pi_+(\boldsymbol{q})^{-(1+\varepsilon)} |q_0|_+^{-\varepsilon} \cdot \left(1 + n\|\boldsymbol{y}^{(\infty)}\|_\infty\right) |q_0|_+^{-1} \\ &\le \left(1 + n\|\boldsymbol{y}^{(\infty)}\|_\infty\right) \Pi_+(\tilde{\boldsymbol{q}})^{-(1+\varepsilon)}, \end{aligned}$$

which finishes the proof modulo a slight change of ε. □

13.3. Finally we are ready for the

Proof of Theorem 11.4*, the case* $\infty \in S$. Applying the case "$\infty \notin S$" of Theorem 11.4 to the map $(\boldsymbol{f}^{(v)})_{v\in S_f}$, we obtain that the pushforward of $\prod_{v\in S_f}\mu_v$ by $(\boldsymbol{f}^{(v)})_{v\in S_f}$ must be strongly extremal, which, in view of Lemma 13.4, implies that $\boldsymbol{f}(x)$ is not $\mathrm{VWMA}_>$ 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 13.3 are satisfied by means of writing down explicit expressions for functions $x \mapsto \mathrm{cov}\left(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}\boldsymbol{u}_{\boldsymbol{f}(x)}\Delta\right)$. This again boils down to the computation of components of $g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}\boldsymbol{u}_{\boldsymbol{f}(x)}\boldsymbol{w}$, where $\boldsymbol{w} = \sum_{I\subset\{0,\dots,n\}} w_I \boldsymbol{e}_I \in \bigwedge^r(\mathcal{D}^{n+1})$.

Since there was no change in the ultrametric components of $g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}$ and $\boldsymbol{u}_{\boldsymbol{f}(x)}$, formula (12.7) is still valid. Furthermore, an expression for the Archimedean components turns out to be similar to (12.7):

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

Therefore one has (12.8) for $j = 1,\dots,\ell-1$, and, in addition,

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

We see that real (resp. p_j-adic) components of all the coordinates of $g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}\boldsymbol{u}_{\boldsymbol{f}(x)}\boldsymbol{w}$ are linear combinations of $1, f_1^{(\infty)},\dots,f_n^{(\infty)}$ (resp. $1, f_1^{(p_j)},\dots,f_n^{(p_j)}$). Condition (12.6a) then immediately follows from Lemma 3.1(b), (c), (11.6) and Corollary 3.3. 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 (12.10) holds, and also that there exists $\rho_\infty > 0$ such that

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

Therefore

$$\begin{aligned}\sup_{x\in B\cap\mathrm{supp}\,\mu} c(g_{\boldsymbol{s},\tilde{\boldsymbol{t}}}\boldsymbol{u}_{\boldsymbol{f}(x)}\boldsymbol{w}) &\ge \Big(\prod_{j=1}^{\ell-1}\rho_j p_j^{s_j}\max_I |w_I|_{p_j}\Big)\rho_\infty e^{-\tilde{t}}\max_I |w_I|_\infty\\ &\ge \rho_\infty\Big(\prod_{j=1}^{\ell-1}\rho_j\Big)e^{-\tilde{t}}\Big(\prod_{j=1}^{\ell-1}p_j^{s_j}\Big)\max_I c(w_I)\end{aligned}$$

$$\underset{(13.4b)}{\geq} \rho_\infty \prod_{j=1}^{\ell-1} \frac{\rho_j}{p_j}.$$

This implies (12.6b) and shows that $\boldsymbol{f}(x)$ is not $\mathrm{VWMA}_{\leq}$ for μ-a.e. $x \in X$, thus finishing the proof of the theorem. □

14. More on S-arithmetic Diophantine approximation

14.1. Extensions of $\mathbb{Q}$. It seems to be a natural task to extend the metric Diophantine approximation results proved in this paper to the framework of an arbitrary number field K. Indeed, the main quantitative nondivergence estimate of the paper (Theorem 9.3) can be rather straightforwardly generalized to the setting of maps from Besicovitch metric spaces into $\mathrm{GL}(m, K_S)$, where K is a finite extension of $\mathbb{Q}$, S is a finite set of its normalized valuations containing all the Archimedean ones, and K_S is the direct product of completions K_v of K over $v \in S$. See the earlier version [KT] of the present paper for more detail. Similarly one can mimic the presentation of §§11–13 to define *very well approximable* elements of K_S^n, and prove that those form a null set with respect to pushforwards of Haar measure by products of nondegenerate maps.

However, understanding multiplicative approximation over an arbitrary number field turns out to be more complicated[5]. Indeed, if K has more than one infinite valuation, the group of units of the ring of integers of K is infinite, which complicates the definition of VWMA vectors and makes proofs of the corresponding results more delicate. The case of an arbitrary number field will be treated in a forthcoming paper.

14.2. Khintchine-type theorems. Another way to generalize the Diophantine setup of this paper would be to replace the right-hand side of (1.4) by an arbitrary function of $\|\boldsymbol{q}\|$ or $\|\tilde{\boldsymbol{q}}\|$. With the notation of §1.4, let us introduce the following definition: for a non-increasing function $\psi : \mathbb{N} \to (0, \infty)$, say that $\boldsymbol{y} \in \mathbb{Q}_S^n$ is *ψ-approximable* if there are infinitely many solutions $\tilde{\boldsymbol{q}} = (q_0, q_1, \dots, q_n) \in \mathbb{Z}^{n+1}$ to

$$|q_0 + \boldsymbol{q} \cdot \boldsymbol{y}|^\ell \leq \begin{cases} \psi(\|\tilde{\boldsymbol{q}}\|) & \text{if } \infty \notin S, \\ \psi(\|\boldsymbol{q}\|) & \text{if } \infty \in S. \end{cases}$$

As in the case of (1.4), it is easy to check using the Borel–Cantelli Lemma that λ-a.e. $\boldsymbol{y} \in \mathbb{Q}_S^n$ is not ψ-approximable whenever the series

$$\begin{cases} \sum_{k=1}^\infty k^n \psi(k) & \text{if } \infty \notin S, \\ \sum_{k=1}^\infty k^{n-1} \psi(k) & \text{if } \infty \in S \end{cases} \tag{14.1}$$

[5] except when K is an imaginary quadratic extension of $\mathbb{Q}$ (cf. [DK]), in which case the proof of the analogs of our results can be carried out without major changes

converges, and using the methods of [L] it should be possible to prove that λ-a.e. $\boldsymbol{y} \in \mathbb{Q}_S^n$ is ψ-approximable if the above series diverges. Similar questions then arise regarding measures other than λ, in particular, volume measures on nondegenerate smooth manifolds or, in the case $\ell > 1$, their products.

In recent years the case $S = \{\infty\}$ has been completely understood, see [BKM, Be] for the convergence case and [BBKM] for the divergence case. That is, the convergence (divergence) of (14.1) was shown to imply that almost no (almost all) points on nondegenerate submanifolds of $\mathbb{R}^n$ are ψ-approximable. Combining the approach of the present paper with the methods of [BKM] and [BBKM] respectively, it seems plausible that both convergence and divergence cases can be proved for $\boldsymbol{f}_*\lambda$ as in Theorem 1.4. Note that when $S = \{p\}$, both cases were recently established for the curve (1.6) [BBK] and for λ-a.e. nondegenerate $\boldsymbol{f} : \mathbb{Z}_p \to \mathbb{Z}_p^2$ which is normal in the sense of Mahler [BK]. The convergence case for nondegenerate curves in $\mathbb{Z}_p^3$ was treated by E. Kovalevskaya in [Ko1], [Ko2], and in another paper [Ko3] she extended the method of [BK] to obtain a result involving both p-adic and infinite valuations. Note also that the paper [BKM] contains a more general (in particular, multiplicative) version of the convergence case for nondegenerate submanifolds of $\mathbb{R}^n$, and it would be interesting to see whether the S-arithmetic set-up can be treated in a similar way.

14.3. Analogues of other results over $\mathbb{R}$. Since the introduction [KM] of the dynamical approach to Diophantine approximation on manifolds, various extensions and generalizations of the method have been found. We expect that many of the ideas developed recently for Diophantine approximation over $\mathbb{R}$ can be applied in the non-Archimedean setting. Specifically we would like to propose two conjectures, in which $\mathbb{Q}_v$ stands for an arbitrary completion of $\mathbb{Q}$.

Conjecture IS (Inheritance for subspaces). Let $\mathcal{L}$ be an affine subspace of $\mathbb{Q}_v^n$ and let $\boldsymbol{f} : \mathbb{Q}_v^d \to \mathcal{L}$ be a C^k map which is *nondegenerate in*[6] $\mathcal{L}$ at λ-a.e. point. Suppose that the volume measure on $\mathcal{L}$ is extremal (resp. strongly extremal); then so is $\boldsymbol{f}_*\lambda$.

This was proved in [K2] for $v = \infty$, and in addition explicit necessary and sufficient conditions, involving coefficients of linear functions parametrizing $\mathcal{L}$, were found for the volume measure on $\mathcal{L}$ to be extremal (strongly extremal). There should be no major difficulties in extending these results to the non-Archimedean case.

Conjecture FP (Friendliness of pushforwards). Let μ be a self-similar measure on the limit set of an irreducible family of contracting similitudes of $\mathbb{Q}_v^d$ satisfying the open set condition (see §11.6), and let $\boldsymbol{f} : \mathbb{Q}_v^d \to \mathbb{Q}_v^n$ be a smooth enough map which is nondegenerate at μ-a.e. point. Then $\boldsymbol{f}_*\mu$ is strongly extremal.

[6] that is, the linear part of $\mathcal{L}$ is spanned by partial derivatives of $\boldsymbol{f}$

The case $v = \infty$ of the above conjecture is one of the main results of [KLW]. Note that a key step of the proof, see [KLW, Proposition 7.3], crucially involves the Mean Value Theorem, and for its non-Archimedean analogue one would need to come up with a replacement, perhaps similarly to our approach to Proposition 4.2.

References

[B] A. Baker, *Transcendental number theory*. Cambridge University Press, Cambridge 1975. Zbl 0297.10013 MR 0422171

[Be] V. Beresnevich, A Groshev type theorem for convergence on manifolds. *Acta Math. Hungar.* **94** (2002), 99–130. Zbl 0997.11053 MR 1905790

[BBKM] V. Beresnevich, V. Bernik, D. Kleinbock, and G. A. Margulis, Metric Diophantine approximation: the Khintchine–Groshev theorem for non-degenerate manifolds. *Moscow Math. J.* **2** (2002), 203–225. Zbl 1013.11039 MR 1944505

[BBK] V. Beresnevich, V. Bernik and E. Kovalevskaya, On approximation of p-adic numbers by p-adic algebraic numbers. *J. Number Theory* **111** (2005), 33–56 Zbl 1078.11050 MR 2124042

[BK] V. Beresnevich and E. Kovalevskaya, On Diophantine approximations of dependent quantities in the p-adic case. *Math. Notes* **73** (2003), 21–35. Zbl 1037.11047 MR 1993537

[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** (9) (2001), 453–486. Zbl 0986.11053 MR 1829381

[Bo] A. Borel, *Linear Algebraic Groups*. Grad. Texts in Math. 126, Springer-Verlag, New York 1991. Zbl 0726.20030 MR 1102012

[BP] A. Borel and G. Prasad, Values of isotropic quadratic forms at S-integral points. *Compositio Math.* **83** (1992), 347–372. Zbl 0777.11008 MR 1175945

[D1] S. G. Dani, On invariant measures, minimal sets, and a lemma of Margulis. *Invent. Math.* **51** (1979), 239–260. Zbl 0415.28008 MR 0530631

[D2] ——, On orbits of unipotent flows on homogeneous spaces. *Ergodic Theory Dynam. Systems* **4** (1984), 25–34; Corrigendum *ibid.* **6** (1986), 321. Zbl 0557.22008 MR 0758891

[D3] ——, On orbits of unipotent flows on homogeneous spaces, II. *Ergodic Theory Dynam. Systems* bf6 (1986), 167–182. Zbl 0601.22003 MR 0857195

[DM] S. G. Dani and G. A. Margulis, Limit distributions of orbits of unipotent flows and values of quadratic forms. In *I. M. Gel'fand Seminar*, Adv. Soviet Math. 16, Part 1, Amer. Math. Soc., Providence, RI, 1993, 91–137. Zbl 0814.22003 MR 1237827

[DK] M. M. Dodson and S. Kristensen, Hausdorff dimension and Diophantine approximation. In *Fractal Geometry and Applications: a Jubilee of Benoît Mandelbrot*, Part 1, Proc. Sympos. Pure Math. 72, Amer. Math. Soc., Providence, RI, 2004, 305–347. Zbl 02154225 MR 2112110

[H] J. E. Hutchinson, Fractals and self-similarity. *Indiana Univ. Math. J.* **30** (1981), 713–747. Zbl 0598.28011 MR 0625600

[K1] D. Kleinbock, Some applications of homogeneous dynamics to number theory. In *Smooth Ergodic Theory and Its Applications* (Seattle, WA, 1999), Proc. Sympos. Pure Math. 68, Amer. Math. Soc. Providence, RI, 2001, 639–660 Zbl 1067.11045 MR 1858548

[K2] ——, Extremal subspaces and their submanifolds. *Geom. Funct. Anal.* **13** (2003), 437–466. Zbl 05056113 MR 1982150

[KLW] D. Kleinbock, E. Lindenstrauss, and B. Weiss, On fractal measures and Diophantine approximation. *Selecta Math.* **10** (2004), 479–523. Zbl 02168146 MR 2134453

[KM] D. Kleinbock and G. A. Margulis, Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. of Math.* **148** (1998), 339–360. Zbl 0922.11061 MR 1652916

[KT] D. Kleinbock and G. Tomanov, Flows on S-arithmetic homogeneous spaces and applications to metric Diophantine approximation. Max Planck Institute, Preprint 2003–65, 2003.

[Ko1] E. Kovalevskaya, p-adic variant of the convergence Khintchine theorem for curves over $\mathbb{Z}_p$. *Acta Math. Inform. Univ. Ostraviensis* **10** (2002), 71–78. Zbl 1069.11027 MR 1943025

[Ko2] ——, Diophantine approximations of points of a class of space curves over Z_p *Dokl. Nats. Akad. Nauk Belarusi* **46** (2) (2002), 53–56 (in Russian). MR 2011857

[Ko3] ——, Simultaneous approximation of zero by values of integral polynomials with respect to different valuations. *Math. Slovaca* **54** (2004), 479–486. Zbl 1108.11050 MR 2114619

[L] E. Lutz, *Sur les approximations diophantiennes linéaires P-adiques.* Actualités Sci. Ind. 1224, Hermann & Cie, Paris 1955. Zbl 0064.28401 MR 0069224

[M] G. A. Margulis, On the action 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, 365–370 Zbl 0305.22014 MR 0470140

[MT1] G. A. Margulis and G. Tomanov, Measure rigidity for algebraic groups over local fields. *C. R. Acad. Sci. Paris Sér. I Math.* **315** (1992), 1221–1226. Zbl 0789.22023 MR 1194522

[MT3] ——, Invariant measures for actions of unipotent groups over local fields on homogeneous spaces. *Invent. Math.* **116** (1994), 347–392. Zbl 0816.22004 MR 1253197

[MT4] ——, Measure rigidity for almost linear groups and its applications. *J. Anal. Math.* **69** (1996), 25–54. Zbl 0864.22005 MR 1428093

[Mt] P. Mattila, *Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability.* Cambridge Stud. in Adv. Math. 44, Cambridge University Press, Cambridge 1995. Zbl 0819.28004 MR 1333890

[R] M. S. Raghunathan, *Discrete subgroups of Lie groups.* Ergeb. Math. Grenzgeb. 68, Springer-Verlag, New York 1972. Zbl 0254.22005 MR 0507234

[PR] V. Platonov and A. Rapinchuk, *Algebraic Groups and Number Theory.* Academic Press, Boston, New York 1994 Zbl 0841.20046 MR 1278263

[Rt1] M. Ratner, Raghunathan's topological conjecture and distributions of unipotent flows. *Duke Math. J.* **63** (1991), 235–280. Zbl 0733.22007 MR 1106945

[Rt2] ——, Invariant measures and orbit closures for unipotent actions on homogeneous spaces. *Geom. Funct. Anal.* **4** (1994), 236–257. Zbl 0801.22008 MR 1262705

[Rt3] ——, Raghunathan's conjectures for p-adic Lie groups. *Internat. Math. Res. Notices* **1993** (5) (1993), 141–146. Zbl 0801.22007 MR 1219864

[Rt4] ——, Raghunathan's conjectures for cartesian products of real and p-adic Lie groups. *Duke Math. J.* **77** (1995), 275–382. Zbl 0914.22016 MR 1321062

[S] H. Sato, Global density theorem for a Federer measure. *Tohoku Math. J.* **44** (1992), 581–595. Zbl 0785.28003 MR 1190914

[Sc] W. H. Schikhof, *Ultrametric calculus*. An introduction to p-adic analysis, Cambridge Stud. in Adv. Math. 4, Cambridge University Press, Cambridge 1984. Zbl 0553.26006 MR 0791759

[Sp1] V. G. Sprindžuk, Mahler's problem in metric number theory. Transl. Math. Monogr., Amer. Math. Soc. Providence, R.I., 1969. Zbl 0181.05502 MR 0245527

[Sp2] ——, Achievements and problems in Diophantine approximation theory. *Russian Math. Surveys* **35** (1980), 1–80. Zbl 0463.10020 MR 0586190

[T] T. Tamagawa, On discrete subgroups of p-adic algebraic groups. In *Arithmetic Algebraic Geometry* (ed. by O. F. G. Schilling), Harper and Row, New York 1965, 11–17. Zbl 0158.27801 MR

[Te] A. Tempelman, *Ergodic Theorems for Group Actions*. Math. Appl. 78, Kluwer Academic Publishers, Dordrecht 1992. Zbl 0753.28014 MR 1172319

[To] G. Tomanov, Orbits on homogeneous spaces of arithmetic origin and approximations. In *Analysis on homogeneous spaces and representation theory of Lie groups*, Adv. Stud. Pure Math. 26, Math. Soc. Japan, Tokyo 2000, 265–297. Zbl 0960.22006 MR 1770724

[W] A. Weil, *Basic Number Theory*. Springer-Verlag, New York 1995. Zbl 0823.11001 MR 1344916

[Z] F. F. Želudevič, Simultane diophantische Approximationen abhängiger Größen in mehreren Metriken. *Acta Arith.* **46** (1986), 285–296. Zbl 0545.10040 MR 0864263

[Zi] R. Zimmer, *Ergodic theory and semisimple groups*. Monogr. Math. 81, Birkhäuser, Basel 1984. Zbl 0571.58015 MR 0776417

Received March 29, 2005

Dmitry Kleinbock, Department of Mathematics, Brandeis University, Waltham, MA 02454-9110, U.S.A.

E-mail: kleinboc@brandeis.edu

George Tomanov, Université Claude Bernard – Lyon 1, Institut Camille Jordan, U.M.R. 5208 CNRS, 69622 Villeurbanne cedex, France

E-mail: tomanov@math.univ-lyon1.fr