GAFA, Geom. funct. anal.
Vol. 13 (2003) 437 – 466
1016-443X/03/020437-30

GAFA Geometric And Functional Analysis

EXTREMAL SUBSPACES AND THEIR SUBMANIFOLDS

D. Kleinbock

Abstract

It was proved in the paper [KM1] that the properties of almost all points of $\mathbb{R}^n$ being not very well (multiplicatively) approximable are inherited by nondegenerate in $\mathbb{R}^n$ (read: not contained in a proper affine subspace) smooth submanifolds. In this paper we consider submanifolds which are contained in proper affine subspaces, and prove that the aforementioned Diophantine properties pass from a subspace to its nondegenerate submanifold. The proofs are based on a correspondence between multidimensional Diophantine approximation and dynamics of lattices in Euclidean spaces.

1 Introduction

We denote by $M_{m,n}$ the space of real matrices with m rows and n columns. $I_k \in M_{k,k}$ stands for the identity matrix. Vectors are named by lowercase boldface letters, such as $\mathbf{x} = (x_i \mid 1 \le i \le k)$. For $\mathbf{x} \in \mathbb{R}^k$ we let $\|\mathbf{x}\| = \max_{1\le i\le k} |x_i|$. 0 stands for a zero vector in any dimension, as well as a zero matrix of any size. The Lebesgue measure in $\mathbb{R}^k$ will be denoted by $|\cdot|$.

We start by recalling several basic facts from the theory of Diophantine approximation. For $v > 0$ and $m, n \in \mathbb{N}$, let us denote by $\mathcal{W}_v(m,n)$ the set of matrices $A \in M_{m,n}$ for which there are infinitely many $\mathbf{q} \in \mathbb{Z}^n$ such that

$$\|A\mathbf{q} + \mathbf{p}\| \le \|\mathbf{q}\|^{-v} \quad \text{for some } \mathbf{p} \in \mathbb{Z}^m\,. \tag{1.1}$$

Clearly $\mathcal{W}_{v_1}(m,n) \supset \mathcal{W}_{v_2}(m,n)$ if $v_1 \le v_2$. We will also use the notation

$$\mathcal{W}_v^+(m,n) \stackrel{\text{def}}{=} \bigcup_{u>v} \mathcal{W}_u(m,n) \quad \text{and} \quad \mathcal{W}_v^-(m,n) \stackrel{\text{def}}{=} \bigcap_{u<v} \mathcal{W}_u(m,n)\,.$$

One knows that $\mathcal{W}_{n/m}(m,n) = M_{m,n}$ by Dirichlet's theorem, and that the Lebesgue measure of $\mathcal{W}_v(m,n)$ is zero whenever $v > n/m$ due to the

Supported in part by NSF Grant DMS-0072565.

Borel–Cantelli lemma. In particular, the set $\mathcal{W}^+_{n/m}(m,n)$ has zero measure. Matrices from the latter set are called *very well approximable*, to be abbreviated as VWA.

It follows from Khintchine's transference principle, see e.g. [C, Chapter V], that the statements $A \in \mathcal{W}^+_{n/m}(m,n)$ and $A^T \in \mathcal{W}^+_{m/n}(n,m)$ are equivalent. In particular, a vector $\mathbf{y} \in \mathbb{R}^n$ interpreted as an $n \times 1$ matrix is VWA iff it is VWA when viewed as a $1 \times n$ matrix. Our goal in the present paper is to look at VWA vectors in $\mathbb{R}^n$, and it will be more convenient for us to use the *row vector* approach, so that, for $\mathbf{y}$ as above and for $\mathbf{q} \in \mathbb{R}^n$, $\mathbf{yq}$ will stand for $y_1q_1 + \cdots + y_nq_n$. In view of the aforementioned duality, this causes no loss of generality, and, hopefully, will cause no confusion.

We now specialize to the case $m = 1$; that is, consider Diophantine properties of vectors (= row matrices) $\mathbf{y} \in \mathbb{R}^n$. Following the terminology introduced by V. Sprindžuk, say that a submanifold $\mathcal{M}$ of $\mathbb{R}^n$ is *extremal* if almost all $\mathbf{y} \in \mathcal{M}$ (with respect to the natural measure class) are not VWA. In other words, if the property of generic $\mathbf{y} \in \mathbb{R}^n$ being not VWA is *inherited* by the submanifold. Pushing this terminology a little further, let us say that a map $\mathbf{f}$ from an open subset U of $\mathbb{R}^d$ to $\mathbb{R}^n$ is *extremal* if $\mathbf{f}(\mathbf{x})$ is not VWA for a.e. $\mathbf{x} \in U$.

Proving extremality of smooth manifolds/maps has been one of the central topics of metric Diophantine approximation for the last 40 years, the major driving force being Sprindžuk's 1964 solution [Sp1] of a long-standing problem of K. Mahler [M], that is, proving the extremality of the so-called *rational normal* or *Veronese* curve

$$\mathcal{M} = \left\{(x, x^2, \ldots, x^n) \mid x \in \mathbb{R}\right\}. \tag{1.2}$$

See [Sp2,3], [BeD] for history and references.

In his 1980 survey of the field [Sp4], Sprindžuk conjectured that a real analytic manifold $\mathcal{M}$ is extremal whenever it is not contained in any proper affine subspace of $\mathbb{R}^n$. The latter condition, loosely put, says that $\mathcal{M}$ 'remembers' the dimension of the space it is imbedded into; and the conjecture asserts that $\mathcal{M}$ must also 'remember' the law of almost all points being not VWA.

This conjecture was proved by G.A. Margulis and the author [KM1] in a stronger form, with the aforementioned geometric condition replaced by an analytic one, and the real analytic class extended to C^k for large enough k. We need the following definitions. Let U be an open subset of $\mathbb{R}^d$, $\mathcal{L}$ an affine subspace of $\mathbb{R}^n$, and let $\mathbf{f} = (f_1, \ldots, f_n)$ be a C^k map $U \to \mathcal{L}$. For

$l \le k$ and $\mathbf{x} \in U$, say that $\mathbf{f}$ is *l-nondegenerate in $\mathcal{L}$ at $\mathbf{x}$* if

$$\begin{array}{c}\text{the linear part of } \mathcal{L} \text{ is spanned by} \\ \text{partial derivatives of } \mathbf{f} \text{ at } \mathbf{x} \text{ of order up to } l\end{array} \tag{1.3}$$

(a linear subspace $\mathcal{L}_0$ of $\mathbb{R}^n$ is called the *linear part* of $\mathcal{L}$ if $\mathcal{L} = \mathcal{L}_0 + \mathbf{y}$ for some $\mathbf{y} \in \mathbb{R}^n$). We will say that $\mathbf{f}$ is *nondegenerate in $\mathcal{L}$ at* $\mathbf{x}$ if (1.3) holds for some l. If $\mathcal{M}$ is a d-dimensional submanifold of $\mathcal{L}$, we will say that $\mathcal{M}$ is *nondegenerate in $\mathcal{L}$ at* $\mathbf{y} \in \mathcal{M}$ if any (equivalently, some) diffeomorphism $\mathbf{f}$ between an open subset U of $\mathbb{R}^d$ and a neighborhood of $\mathbf{y}$ in $\mathcal{M}$ is nondegenerate in $\mathcal{L}$ at $\mathbf{f}^{-1}(\mathbf{y})$. We will say that $\mathbf{f} : U \to \mathcal{L}$ (resp. $\mathcal{M} \subset \mathcal{L}$) is *nondegenerate in $\mathcal{L}$* if it is nondegenerate in $\mathcal{L}$ at almost every point of U (resp. $\mathcal{M}$, in the sense of the natural measure class on $\mathcal{M}$).

One of the main results of [KM1] is the following:

Theorem 1.1. *Let $\mathbf{f} : U \to \mathbb{R}^n$, $U \subset \mathbb{R}^d$, be a smooth map which is nondegenerate in $\mathbb{R}^n$. Then $\mathbf{f}$ is extremal.*

In particular, smooth submanifolds of $\mathbb{R}^n$ which are nondegenerate in $\mathbb{R}^n$ are extremal. Note that many special cases were proved before the general case; see [KM1], [BeD] for a detailed account, and [BeKM], [BeBKM], [KLW] for further developments.

The goal of the present paper is to study manifolds for which the aforementioned non-degeneracy-in-$\mathbb{R}^n$ condition fails. In fact, the simplest ones, namely proper affine subspaces of $\mathbb{R}^n$ themselves, have been the subject of several papers [S2], [Sp3], [BeBDD], and certain conditions have been found sufficient for their extremality. To the best of the author's knowledge, nobody has yet turned attention to proper submanifolds of affine subspaces of $\mathbb{R}^n$. Let us now state one of the main results of the present paper, which addresses this gap.

Theorem 1.2. *Let $\mathcal{L}$ be an affine subspace of $\mathbb{R}^n$. Then:*

(a) *if $\mathcal{L}$ is extremal and $\mathbf{f} : U \to \mathcal{L}$, $U \subset \mathbb{R}^d$, is a smooth map which is nondegenerate in $\mathcal{L}$, then $\mathbf{f}$ is extremal;*

(b) *if $\mathcal{L}$ is not extremal, then all points of $\mathcal{L}$ are VWA (in particular, no subset of $\mathcal{L}$ is extremal).*

This result generalizes Theorem 1.1, showing that the extremality of affine subspaces is inherited by their nondegenerate submanifolds. It also implies that a manifold nondegenerate in some affine subspace of $\mathbb{R}^n$ is extremal if and only if this subspace contains at least one not very well approximable point. (Cf. a similar statement conjectured by B. Weiss in the context of interval exchange transformations and Teichmüller flows [W, Conjecture 2.1].)

The proof is based on the methods of [KM1], that is, on the correspondence between approximation properties of vectors and trajectories of lattices in Euclidean spaces. Necessary background is reviewed in §2. Then in §3 we use the language of lattices to give a necessary and sufficient condition for a map $\mathbf{f} : U \to \mathbb{R}^n$, $U \subset \mathbb{R}^d$, within a certain class of *good* maps to be extremal, and then show that this condition is inherited by nondegenerate submanifolds of affine subspaces.

Dealing with an s-dimensional affine subspace of $\mathbb{R}^n$, one can be more specific and phrase the aforementioned condition in terms of coefficients of an affine map parametrizing the subspace. By permuting variables one can without loss of generality choose a parametrizing map of the form $\mathbf{x} \mapsto (\mathbf{x}, \mathbf{x}A' + \mathbf{a}_0)$, where A' is a matrix of size $s \times (n-s)$ and $\mathbf{a}_0 \in \mathbb{R}^{n-s}$ (here both $\mathbf{x}$ and $\mathbf{a}_0$ are row vectors). In an even more abbreviated way, we will denote the vector $(1, x_1, \dots, x_s)$ by $\tilde{\mathbf{x}}$, and the matrix $\binom{\mathbf{a}_0}{A'}$ by $A \in M_{s+1,n-s}$; then $\mathcal{L}$ is parametrized by

$$\mathbf{x} \mapsto (\mathbf{x}, \tilde{\mathbf{x}}A)\,. \tag{1.4}$$

We show in §4 how the results of §3 allow one to write down a condition on A (see Theorem 4.3) equivalent to the extremality of the map (1.4). On the other hand, it easily follows from the definitions, as explained in §4, that every point of $\mathcal{L}$ parametrized by (1.4) is VWA whenever A belongs to $\mathcal{W}_n^+(s+1, n-s)$. We show that the converse is also true in the following two cases, and by the following two methods:

$$s = n-1 \text{ (that is, } \mathcal{L} \text{ is an affine hyperplane)}\,, \tag{1.5}$$

– as a consequence of Theorem 4.3, and

$$\begin{gathered} s = 1 \text{ and } A \text{ is of the form } \binom{0}{\mathbf{b}} \text{ for a row vector } \mathbf{b} \\ \text{(that is, } \mathcal{L} \text{ is a line passing through the origin)}\,. \end{gathered} \tag{1.6}$$

– using an argument borrowed from [BeBDD]. In other words, the following can be proved:

Theorem 1.3. *In the two special cases* (1.5) *and* (1.6)*, the map* (1.4) *is extremal if and only if*

$$A \notin \mathcal{W}_n^+(s+1, n-s)\,. \tag{1.7}$$

Whether the same is true for an arbitrary affine subspace is not clear. Since matrices A as above provide local coordinate charts to the set of s-dimensional affine subspaces of $\mathbb{R}^n$, and in view of M. Dodson's [Do] formula for the Hausdorff dimension of the sets $\mathcal{W}_v(m,n)$, the affirmative answer to the above question would imply that the dimension of the (null) set of non-extremal s-dimensional affine subspaces of $\mathbb{R}^n$ is equal to

$$\dim\left(\mathcal{W}_n^+(s+1,n-s)\right) = (n-s-1)(s+1)+1\,, \tag{1.8}$$

which is precisely 1 + the Hausdorff dimension of the set of 'rational' s-dimensional subspaces, i.e. of the set

$$\left\{A \in M_{s+1,n-s} \;\middle|\; A\mathbf{q}+\mathbf{p}=0 \text{ for some } \mathbf{p}\in\mathbb{Z}^{s+1}, \mathbf{q}\in\mathbb{Z}^{n-s}\smallsetminus\{0\}\right\}.$$

Other open problems and generalizations are discussed in the last two sections of the paper. This includes the so-called *multiplicative* modification of the standard set-up, which is the subject of §5. Namely, there we define *not very well multiplicatively approximable* (not VWMA, a property stronger than 'not VWA' but still generic in $\mathbb{R}^n$) vectors and *strongly extremal* manifolds (i.e. those for which almost all points are not VWMA). It was proved in [KM1] that smooth nondegenerate submanifolds of $\mathbb{R}^n$ are strongly extremal; we generalize this as follows:

Theorem 1.4. *Let $\mathcal{L}$ be an affine subspace of $\mathbb{R}^n$. Then:*

(a) *if $\mathcal{L}$ is strongly extremal and $\mathbf{f} : U \to \mathcal{L}$ is a smooth map which is nondegenerate in $\mathcal{L}$, then $\mathbf{f}$ is strongly extremal;*

(b) *if $\mathcal{L}$ is not strongly extremal, then all points of $\mathcal{L}$ are VWMA (in particular, no subset of $\mathcal{L}$ is strongly extremal).*

Similarly to Theorem 1.2, this is done by writing down a necessary and sufficient condition (see Theorem 5.3) for a good map to be strongly extremal, and then showing that condition to be inherited by nondegenerate submanifolds of affine subspaces. Following the lines of §4, we are able to simplify that condition in the case (1.5), thus explicitly describing strongly extremal hyperplanes and identifying those which are extremal but not strongly extremal. Whether this can be extended beyond the codimension one case is an open question.

2 Diophantine Approximation and Lattices

In this section we introduce some notation and terminology which will help us work with discrete subgroups Γ of $\mathbb{R}^k$, $k\in\mathbb{N}$. We define the *rank* $\mathrm{rk}(\Gamma)$ of Γ to be the dimension of $\mathbb{R}\Gamma$. Also define $\delta(\Gamma)$ to be the norm of a nonzero element of Γ with the smallest norm, that is,

$$\delta(\Gamma) \stackrel{\text{def}}{=} \inf_{\mathbf{v}\in\Gamma\smallsetminus\{0\}} \|\mathbf{v}\|\,.$$

For $0 \le j \le k$, let us denote by $\mathcal{S}_{k,j}$ the set of all subgroups of $\mathbb{Z}^k$ of rank j, and by $\mathcal{S}_k$ the set of all nonzero subgroups of $\mathbb{Z}^k$ of rank smaller than k, that is, $\mathcal{S}_k \stackrel{\text{def}}{=} \cup_{j=1}^{k-1} \mathcal{S}_{k,j}$.

It will be useful to consider exterior products of vectors generating Γ. Namely, if $\Gamma \in \mathcal{S}_{k,j}$, say that $\mathbf{w} \in \bigwedge^j(\mathbb{R}^k)$ *represents* Γ if

$$\mathbf{w} = \begin{cases} 1 & \text{if } j = 0 \\ \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_j & \text{if } j > 0 \text{ and } \mathbf{v}_1, \dots, \mathbf{v}_j \text{ is a basis of } \Gamma\,. \end{cases}$$

Clearly the element representing Γ is defined up to a sign. With some abuse of notation, we will also denote by $\mathcal{S}_{k,j}$ and $\mathcal{S}_k$ the set of $\mathbf{w} \in \bigwedge(\mathbb{R}^k)$ representing $\Gamma \in \mathcal{S}_{k,j}$ and $\in \mathcal{S}_k$ respectively.

The set of all *lattices* (discrete subgroups of maximal rank) in $\mathbb{R}^k$ of covolume one can be identified with the homogeneous space $\mathrm{SL}_k(\mathbb{R})/\mathrm{SL}_k(\mathbb{Z})$, which we will denote by Ω_k. It is a noncompact space with finite $\mathrm{SL}_k(\mathbb{R})$-invariant measure, and the restriction of the function $\delta(\cdot)$ defined above to this space can be used to describe its geometry at infinity. Namely, Mahler's compactness criterion [R, Corollary 10.9] says that a subset of Ω_k is relatively compact if and only if δ is bounded away from zero on this subset. Further, it follows from the *reduction theory* for $\mathrm{SL}_k(\mathbb{Z})$, see e.g. [Si, Satz 4], that the ratio of $1 + \log(1/\delta(\cdot))$ and $1 + \mathrm{dist}(\,\cdot\,, \mathbb{Z}^k)$ is bounded between two positive constants for any right invariant Riemannian metric 'dist' on the space of lattices. In other words, a lattice $\Lambda \in \Omega_k$ for which $\delta(\Lambda)$ is small is approximately $\log(1/\delta(\Lambda))$ away from the base point $\mathbb{Z}^k$. The reader is referred to [K1] for more details.

This justifies the following definition: for $\gamma \ge 0$ and any one-parameter semigroup $F = \{g_t \mid t \ge 0\}$ acting on Ω_k, say that the F-trajectory of $\Lambda \in \Omega_k$ *grows with exponent* $\ge \gamma$ if there exist arbitrarily large positive t such that

$$\delta(g_t\Lambda) \le e^{-\gamma t}\,.$$

Also define the *growth exponent* $\gamma_F(\Lambda)$ of Λ with respect to F to be the supremum of all γ for which the F-trajectory of Λ grows with exponent $\ge \gamma$. In view of the preceding remark, one has

$$\gamma_F(\Lambda) = \limsup_{t\to\infty} \frac{\mathrm{dist}(g_t\Lambda, \mathbb{Z}^m)}{t}\,.$$

Now let us describe a correspondence, dating back to [S3] and [D], between approximation properties of vectors $\mathbf{y} \in \mathbb{R}^n$ and dynamics of certain trajectories in Ω_{n+1}. Given a row vector $\mathbf{y} \in \mathbb{R}^n$ one considers a lattice $u_{\mathbf{y}}\mathbb{Z}^{n+1}$ in $\mathbb{R}^{n+1}$, where $u_{\mathbf{y}} \stackrel{\text{def}}{=} \left(\begin{smallmatrix} 1 & \mathbf{y} \\ 0 & I_n \end{smallmatrix}\right)$; that is, the collection of vectors of the form $\binom{\mathbf{y}\mathbf{q}+p}{\mathbf{q}}$, where $p \in \mathbb{Z}$ and $\mathbf{q} \in \mathbb{Z}^n$. Then one reads Diophantine properties of $\mathbf{y}$ from the behavior of the trajectory $Fu_{\mathbf{y}}\mathbb{Z}^{n+1}$, where

$$F = \{g_t \mid t \geq 0\}, \quad \text{with } g_t = \operatorname{diag}(e^t, e^{-t/n}, \dots, e^{-t/n}), \tag{2.1}$$

is a one-parameter subsemigroup of $\mathrm{SL}_{n+1}(\mathbb{R})$ which expands the first coordinate and uniformly contracts the last n coordinates of vectors in $\mathbb{R}^{n+1}$.

The passage from Diophantine approximation to growth exponents of trajectories will be based on the following elementary lemma:

Lemma 2.1. *Suppose we are given a set $E \subset \mathbb{R}^2$ which is discrete and homogeneous with respect to positive integers, that is, $kE \subset E$ for any $k \in \mathbb{N}$. Also take $a, b > 0$, $v > a/b$, and define γ by*

$$\gamma = \frac{bv - a}{v + 1} \quad \Leftrightarrow \quad v = \frac{a + \gamma}{b - \gamma}. \tag{2.2}$$

Then the following are equivalent:

[2.1-i] *there exist $(x, z) \in E$ with arbitrarily large $|z|$ such that $|x| \leq |z|^{-v}$;*
[2.1-ii] *there exist arbitrarily large $t > 0$ such that for some $(x, z) \in E \smallsetminus \{0\}$ one has*

$$\max\left(e^{at}|x|, e^{-bt}|z|\right) \leq e^{-\gamma t}. \tag{2.3}$$

Proof. Assume [2.1-i], take $(x, z) \in E$ with $|x| \leq |z|^{-v}$, and define t by $e^{-bt}|z| = e^{-\gamma t}$, that is, $|z| = e^{(b-\gamma)t}$. (Note that it follows from (2.2) that $\gamma < b$.) Then one has

$$e^{at}|x| \leq e^{at}|z|^{-v} = e^{at}(e^{(b-\gamma)t})^{-v} \underset{(2.2)}{=} e^{-\gamma t},$$

that is, (2.3) holds for this choice of x, z and t. Taking $|z|$ arbitrarily large produces arbitrarily large t as well.

Assume now that [2.1-ii] holds. Then one can find a sequence $t_n \to \infty$ and $(x_n, z_n) \in E \smallsetminus \{0\}$ such that

$$e^{at_n}|x_n| \leq e^{-\gamma t_n} \quad \text{and} \quad e^{-bt_n}|z_n| \leq e^{-\gamma t_n}, \tag{2.4}$$

and write

$$|x_n| \underset{(2.4)}{\leq} e^{-(a+\gamma)t_n} \underset{(2.2)}{=} e^{-v(b-\gamma)t_n} \underset{(2.4)}{\leq} |z_n|^{-v}.$$

If the sequence $\{z_n\}$ is unbounded, [2.1-i] is proved. Otherwise, note that $x_n \to 0$ due to (2.4); by the discreteness of E, the sequence $\{(x_n, z_n)\}$ must stabilize, and thus one has $(0, z) \in E$ for some $z > 0$. But then $(0, kz) \in E$ for any $k \in \mathbb{N}$ by the homogeneity, and the proof of [2.1-i] is finished. □

Corollary 2.2. *$\mathbf{y} \in \mathcal{W}_v(1,n)$ iff the growth exponent $\gamma_F(u_{\mathbf{y}}\mathbb{Z}^{n+1})$ of $u_{\mathbf{y}}\mathbb{Z}^{n+1}$ with respect to F as in (2.1) is not less than γ, the latter being defined by*

$$\gamma = \frac{v-n}{n(v+1)} \quad \Leftrightarrow \quad v = \frac{n(1+\gamma)}{1-n\gamma}\,. \tag{2.5}$$

Proof. This corollary is in fact a special case of Theorem 8.5 from [KM2]. However one can easily derive it from the previous lemma by taking $a = 1$, $b = 1/n$ and

$$E = \left\{(\mathbf{y}\mathbf{q} + p, \|\mathbf{q}\|) \;\middle|\; (p, \mathbf{q}) \in \mathbb{Z}^{n+1}\right\},$$

and noticing that the inequality

$$\delta(g_t u_{\mathbf{y}}\mathbb{Z}^{n+1}) \le e^{-\gamma t} \tag{2.6}$$

amounts to the validity of (2.3) for some $(x,z) \in E \smallsetminus \{0\}$. □

Corollary 2.3. *The following are equivalent for $\mathbf{y} \in \mathbb{R}^n$ and F as in (2.1):*

[2.3-i] *$\mathbf{y}$ is VWA;*

[2.3-ii] *$\gamma_F(u_{\mathbf{y}}\mathbb{Z}^{n+1}) > 0$;*

[2.3-iii] *for some $\gamma > 0$ there exist infinitely many $t \in \mathbb{N}$ such that (2.6) holds.*

Proof. The equivalence of [2.3-i] and [2.3-ii] is straightforward from Corollary 2.2 and (2.5), while to derive [2.3-iii] one notices that the ratio of $\delta(g_t\cdot)$ and $\delta(g_{t'}\cdot)$ is uniformly bounded from both sides when $|t - t'| < 1$. □

We return now to the setting of the Diophantine approximation on subsets of $\mathbb{R}^n$. More precisely, we consider a map $\mathbf{f} = (f_1, \dots, f_n) : U \to \mathbb{R}^n$, where U is an open subset of $\mathbb{R}^d$, and study Diophantine properties of vectors $\mathbf{f}(\mathbf{x})$ for a.e. $\mathbf{x} \in U$. This calls for considering the corresponding map from U into Ω_{n+1}, namely $\mathbf{x} \mapsto u_{\mathbf{f}(\mathbf{x})}\mathbb{Z}^{n+1}$, where

$$u_{\mathbf{f}(\mathbf{x})} \stackrel{\text{def}}{=} \begin{pmatrix} 1 & \mathbf{f}(\mathbf{x}) \\ 0 & I_n \end{pmatrix}, \tag{2.7}$$

and then looking at growth of trajectories of lattices $u_{\mathbf{f}(\mathbf{x})}\mathbb{Z}^{n+1}$ under the action of g_t as in (2.1).

In the next section we will describe a method, introduced in [KM1], which is based on keeping track on what happens to every subgroup Γ of $\mathbb{Z}^{n+1}$ under the action by $u_{\mathbf{f}(\mathbf{x})}$ and then by g_t. Fix a basis $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_n$ of $\mathbb{R}^{n+1}$, and for $I = \{i_1, \dots, i_j\} \subset \{0, \dots, n\}$, $i_1 < i_2 < \cdots < i_j$, let

$$\mathbf{e}_I \stackrel{\text{def}}{=} \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_j} \in \bigwedge^j(\mathbb{R}^{n+1}),$$

with the convention $\mathbf{e}_\varnothing = 1$. We extend the norm $\|\cdot\|$ from $\mathbb{R}^{n+1}$ to the exterior algebra $\bigwedge(\mathbb{R}^{n+1})$ by $\|\sum_{I\subset\{1,\dots,j\}} w_I\mathbf{e}_I\| = \max_{I\subset\{0,\dots,n\}} |w_I|$.

Thus it makes sense to define the norm of Γ as above by $\|\Gamma\| \stackrel{\text{def}}{=} \|\mathbf{w}\|$, where $\mathbf{w}$ represents Γ. Note that the ratio of $\|\Gamma\|$ and the volume of the quotient space $\mathbb{R}\Gamma/\Gamma$ is uniformly bounded between two positive constants (depending on n and on the choice of the norm). Also note that it follows from Minkowski's theorem that $\delta(\Gamma)$ must be small whenever $\|\Gamma\|$ is small; more precisely, for any $j > 0$ there exists a positive constant $c(j)$ such that

$$\delta(\Gamma) \leq c(j)\|\Gamma\|^{1/j} \tag{2.8}$$

for any Γ of rank j.

As a preparation for the next section, let us write down a formula for $g_t u_{\mathbf{f}(\mathbf{x})}\mathbf{w}$, where $\mathbf{w}$ represents a subgroup Γ of $\mathbb{Z}^{n+1}$. Note that the action of $u_{\mathbf{f}(\mathbf{x})}$ leaves $\mathbf{e}_0$ invariant and sends $\mathbf{e}_i$ to $\mathbf{e}_i + f_i(\mathbf{x})\mathbf{e}_0$, $i = 1, \dots, n$. Therefore

$$u_{\mathbf{f}(\mathbf{x})}\mathbf{e}_I = \begin{cases} \mathbf{e}_I & \text{if } 0 \in I \\ \mathbf{e}_I + \sum_{i \in I} (-1)^{l(I,i)} f_i(\mathbf{x}) \mathbf{e}_{I \cup \{0\} \smallsetminus \{i\}} & \text{otherwise}, \end{cases}$$

where one defines

$$l(I,i) \stackrel{\text{def}}{=} \text{ the number of elements of } I \text{ strictly between } 0 \text{ and } i.$$

Taking $\mathbf{w}$ of the form $\sum_I w_I \mathbf{e}_I$, one gets

$$u_{\mathbf{f}(\mathbf{x})}\mathbf{w} = \sum_{0 \in I} \Big(w_I + \sum_{i \notin I} (-1)^{l(I,i)} w_{I \cup \{i\} \smallsetminus \{0\}} f_i(\mathbf{x}) \Big) \mathbf{e}_I + \sum_{0 \notin I} w_I \mathbf{e}_I \,,$$

and, further, for $\mathbf{w} \in \bigwedge^j(\mathbb{R}^{n+1})$,

$$g_t u_{\mathbf{f}(\mathbf{x})}\mathbf{w} = e^{\frac{n+1-j}{n}t} \sum_{0 \in I} \Big(w_I + \sum_{i \notin I} (-1)^{l(I,i)} w_{I \cup \{i\} \smallsetminus \{0\}} f_i(\mathbf{x}) \Big) \mathbf{e}_I + e^{-\frac{j}{n}t} \sum_{0 \notin I} w_I \mathbf{e}_I . \tag{2.9}$$

What is important here is that each of the coordinates of $g_t u_{\mathbf{f}(\mathbf{x})}\mathbf{w}$ is expressed as a linear combination of functions $1, f_1, \dots, f_n$.

3 Extremality Criteria for Good Maps

Let us recall the definition introduced in [KM1]. If C and α are positive numbers and V a subset of $\mathbb{R}^d$, let us say that a function $f : V \to \mathbb{R}$ is *(C, α)-good on V* if

$$\begin{gathered}\text{for any open ball } B \subset V \text{ and any } \varepsilon > 0, \text{ one has} \\ \Big|\{x \in B \;\big|\; |f(x)| < \varepsilon \cdot \sup_{x \in B} |f(x)|\}\Big| \leq C\varepsilon^\alpha |B| .\end{gathered} \tag{3.1}$$

See [KM1], [BeKM] for various properties and examples of (C, α)-good functions. One property will be particularly useful: it is easy to see that

$$f_i,\ i \in I, \text{ are } (C,\alpha)\text{-good on } V \Rightarrow \text{so is } \sup_{i \in I} |f_i| . \tag{3.2}$$

Now let $\mathbf{f} = (f_1, \dots, f_n)$ be a map from an open subset U of $\mathbb{R}^d$ to $\mathbb{R}^n$. We will say that $\mathbf{f}$ is *good at* $\mathbf{x}_0 \in U$ if there exists a neighborhood $V \subset U$ of $\mathbf{x}_0$ and positive C, α such that any linear combination of $1, f_1, \dots, f_n$ is (C, α)-good on V. We will say that $\mathbf{f} : U \to \mathbb{R}^n$ is *good* if the set of $\mathbf{x}_0 \in U$ such that $\mathbf{f}$ is good at $\mathbf{x}_0$ has full measure. Note that C, α do not have to be uniform in $\mathbf{x}_0 \in U$; however, once $V \ni \mathbf{x}_0$ is chosen, every function of the form $f = c_0 + c_1 f_1 + \dots + c_n f_n$ must satisfy (3.1) for some uniformly chosen C and α.

Recall (see [KM1, Lemma 3.2]) that the basic example of (C, α)-good functions is given by polynomials: any polynomial map $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^n$ is good at every point of $\mathbb{R}^d$. A more general class of examples is given by linear combinations of coordinate functions of nondegenerate maps:

PROPOSITION 3.1 [KM1, Proposition 3.4]. *Let $\mathbf{f} = (f_1, \dots, f_n)$ be a smooth map from an open subset U of $\mathbb{R}^d$ to $\mathbb{R}^n$ which is l-nondegenerate in $\mathbb{R}^n$ at $\mathbf{x}_0 \in U$. Then there exists a neighborhood $V \subset U$ of $\mathbf{x}_0$ and positive C such that any linear combination of $1, f_1, \dots, f_n$ is $(C, 1/dl)$-good on V.*

In other words, $\mathbf{f}$ is good at every point at which it is nondegenerate in $\mathbb{R}^n$. From this one easily derives

COROLLARY 3.2. *Let $\mathcal{L}$ be an affine subspace of $\mathbb{R}^n$ and let $\mathbf{f} = (f_1, \dots, f_n)$ be a smooth map from an open subset U of $\mathbb{R}^d$ to $\mathcal{L}$ which is nondegenerate in $\mathcal{L}$ at $\mathbf{x}_0 \in U$. Then $\mathbf{f}$ is good at $\mathbf{x}_0$.*

Proof. Put $\dim(\mathcal{L}) = s$, choose any affine map $\mathbf{h}$ from $\mathbb{R}^s$ onto $\mathcal{L}$, and define $\mathbf{g} = (g_1, \dots, g_s)$ by $\mathbf{g} = \mathbf{h}^{-1} \circ \mathbf{f}$. It follows from the nondegeneracy of $\mathbf{f}$ in $\mathcal{L}$ that $\mathbf{g}$ is nondegenerate in $\mathbb{R}^s$ at $\mathbf{x}_0$, hence, by Proposition 3.1, it is good at $\mathbf{x}_0$. To finish the proof it suffices to observe that any linear combination of $1, f_1, \dots, f_n$ is a linear combination of $1, g_1, \dots, g_s$. □

COROLLARY 3.3. *Let $\mathbf{f}$ be a real analytic map from a connected open subset U of $\mathbb{R}^d$ to $\mathbb{R}^n$. Then there exists an affine subspace $\mathcal{L}$ of $\mathbb{R}^n$ such that $\mathbf{f}$ is nondegenerate in $\mathcal{L}$ at every point of U; consequently, $\mathbf{f}$ is good at every point of U.*

Proof. For any $\mathbf{x} \in U$, denote by $\mathcal{L}_0(\mathbf{x})$ the linear space spanned by all partial derivatives of $\mathbf{f}$ at $\mathbf{x}$, and put $\mathcal{L}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathcal{L}_0(\mathbf{x})$. Then for any $\mathbf{x}_0, \mathbf{x} \in U$ such that the Taylor series of $\mathbf{f}$ centered at $\mathbf{x}_0$ converges at $\mathbf{x}$, one has $\mathbf{f}(\mathbf{x}) \in \mathcal{L}(\mathbf{x}_0)$ and $\mathcal{L}(\mathbf{x}) \subset \mathcal{L}(\mathbf{x}_0)$. Since U is connected, for any $\mathbf{x}' \in U$ one can find a finite sequence $\mathbf{x}_1, \dots, \mathbf{x}_k = \mathbf{x}'$ such that the Taylor series of $\mathbf{f}$ centered at $\mathbf{x}_{i-1}$ converges at $\mathbf{x}_i$ for all $i = 1, \dots, k$. Therefore

$\mathcal{L}(\mathbf{x}') = \mathcal{L}(\mathbf{x}_0)$, and, by reversing the roles of $\mathbf{x}_0$ and $\mathbf{x}'$, one sees that $\mathcal{L} \stackrel{\text{def}}{=} \mathcal{L}(\mathbf{x})$ is independent of $\mathbf{x} \in U$. It remains to notice that from the construction it follows that $\mathbf{f}$ is nondegenerate in $\mathcal{L}$ at every $\mathbf{x} \in U$, and apply Corollary 3.2. $\square$

Note that it follows from the proof that $\mathcal{L}$ can be defined as the intersection of all the affine subspaces of $\mathbb{R}^n$ containing $\mathbf{f}(U)$, or, equivalently, as

$$\mathbf{f}(\mathbf{x}_0) + \operatorname{Span}\big\{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) \mid \mathbf{x} \in U\big\}$$

for any $\mathbf{x}_0 \in U$.

Example 3.4. It is instructive for better understanding of the class of good maps to remark that the assumption of the analyticity of $\mathbf{f}$ cannot be dropped. Indeed, let us sketch a construction of a C^∞ function from $[0,1]$ to $\mathbb{R}_+$ which is not good on a subset of $[0,1]$ of positive measure. First for every $J = (a,b) \subset [0,1]$ define

$$\psi_J(x) \stackrel{\text{def}}{=} \varphi(x-a)\varphi(b-x)\,, \quad \text{where } \varphi(x) \stackrel{\text{def}}{=} \begin{cases} 0\,, & x \le 0\,, \\ e^{-1/x^2}\,, & x \ge 0\,. \end{cases}$$

One can easily verify that for every neighborhood V of either a or b it is impossible to find $C, \alpha > 0$ such that ψ_J is (C,α)-good on V. Then consider a Cantor set $K \subset [0,1]$ of positive measure, and for $k \in \mathbb{N}$ let $\mathcal{J}_k$ be the collection of disjoint subsegments of $[0,1]$ thrown away at the kth stage of the construction of K. (For example, one can divide every interval left at the kth stage onto 3^{k+1} equal pieces and then throw away the middle interval.) After that define

$$\psi(x) \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} c_k \sum_{J \in \mathcal{J}_k} \psi_J(x)\,,$$

where c_k decays fast enough as $k \to \infty$ to guarantee that ψ is C^∞ (in the aforementioned example, one can take $c_k = 3^{-3^k}$). Since every neighborhood of every point of K contains an endpoint of $J \in \mathcal{J}_k$ for some $k \in \mathbb{N}$, it follows that ψ is not good at any $x \in K$.

We now state an estimate from [KM1], which will be used to derive a criterion for the extremality of $\mathbf{f}$ once the latter is chosen within the class of good maps. It will be convenient to use the following notation: if $B = B(\mathbf{x}, r)$ is a ball in $\mathbb{R}^d$ and $c > 0$, we will denote by cB the ball $B(\mathbf{x}, cr)$.

Theorem 3.5 (cf. [KM1, Theorem 5.2]). *For any $d, k \in \mathbb{N}$ there exists a positive constant C' (explicitly estimated in* [KM1]*) such that the following* *holds. Given $C,\alpha>0$, $0<\rho\le 1/k$, a ball $B\subset\mathbb{R}^d$ and a continuous map $h: 3^kB\to \mathrm{GL}_k(\mathbb{R})$, let us assume that*

[3.5-i] *for any $\Gamma\in\mathcal{S}_k$, the function $x\mapsto \|h(\mathbf{x})\Gamma\|$ is (C,α)-good on 3^kB, and*

[3.5-ii] *for any $\Gamma\in\mathcal{S}_k$, $\sup_{\mathbf{x}\in B}\|h(\mathbf{x})\Gamma\|\ge\rho$.*

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

$$\left|\{\mathbf{x}\in B\mid \delta(h(\mathbf{x})\mathbb{Z}^k)<\varepsilon\}\right|\le CC'\left(\tfrac{\varepsilon}{\rho}\right)^{\alpha}|B|\,.$$

Informally speaking, the conclusion of the above theorem says that the 'orbit' $\{h(\mathbf{x})\mathbb{Z}^k\mid \mathbf{x}\in B\}\subset\Omega_k$ 'does not diverge', that is, its very significant proportion (computed in terms of Lebesgue measure on B) stays inside compact sets $\{\Lambda\in\Omega_k\mid\delta(\Lambda)\ge\varepsilon\}$. We remark that such nondivergence results have a long history, dating back to the work of Margulis [Ma] in the 1970s, and many applications in the theory of dynamics on homogeneous spaces, see e.g. [KSS, Chapter 3] for a historical account.

The next lemma sharpens [KM1, Theorem 5.4], giving a condition sufficient for the extremality of a good map $\mathbf{f}$.

LEMMA 3.6. *Let B be a ball in $\mathbb{R}^d$, and let $\mathbf{f}=(f_1,\dots,f_n)$ be a continuous map from $3^{n+1}B$ to $\mathbb{R}^n$. Suppose that*

[3.6-i] *$\exists\, C,\alpha>0$ such that any linear combination of $1,f_1,\dots,f_n$ is (C,α)-good on $3^{n+1}B$;*

[3.6-ii] *for any $\beta>0$ there exists $T=T(\beta)>0$ such that for any $t\ge T$ and any $\Gamma\in\mathcal{S}_{n+1}$ one has*

$$\sup_{\mathbf{x}\in B}\|g_tu_{\mathbf{f}(\mathbf{x})}\Gamma\|\ge e^{-\beta t}\,,$$

where $u_{\mathbf{f}(\mathbf{x})}$ is as in (2.7) and $F=\{g_t\}$ is as in (2.1).

Then $\mathbf{f}(\mathbf{x})$ is not VWA for a.e. $\mathbf{x}\in B$.

Proof. We apply Theorem 3.5 with $k=n+1$ and $h(\mathbf{x})=g_tu_{\mathbf{f}(\mathbf{x})}$. Our goal is to show that for any $\gamma>0$, the set $\{\mathbf{x}\in B\mid \gamma_F(u_{\mathbf{f}(\mathbf{x})}\mathbb{Z}^{n+1})>\gamma\}$ has measure zero. As was observed in the preceding section, see (2.9), for every $\mathbf{w}\in\mathcal{S}_{n+1}$ all the coordinates of $h(\mathbf{x})\mathbf{w}$ are linear combinations of $1,f_1,\dots,f_n$, which, by (3.2), implies [3.5-i]. Now choose any $\beta<\gamma$, take $t\ge\max((T(\beta),\log(n+1)/\beta)$ and put $\rho=e^{-\beta t}$. This guarantees $\rho\le 1/k$ and verifies condition [3.5-ii]. Thus, for t as above, the measure of each of the sets

$$\left\{\mathbf{x}\in B\mid \delta(g_tu_{\mathbf{f}(\mathbf{x})}\mathbb{Z}^{n+1})<e^{-\gamma t}\right\}\tag{3.3}$$

is not greater than $CC'e^{-\alpha(\gamma-\beta)t}$. One then applies the Borel–Cantelli lemma to conclude that almost every $\mathbf{x} \in B$ belongs to at most finitely many of the sets (3.3) with $t \in \mathbb{N}$, which completes the proof in view of Corollary 2.3. □

The next lemma shows that assumption [3.6-ii] is in fact necessary for the extremality of $\mathbf{f}$; furthermore, the consequences of [3.6-ii] being not true are much stronger than positive measure of the set $\{\mathbf{x} \in B \mid \mathbf{f}(\mathbf{x}) \text{ is VWA}\}$.

LEMMA 3.7. *Let B be a ball in $\mathbb{R}^d$, and let $\mathbf{f}$ be a map from B to $\mathbb{R}^n$ such that* [3.6-ii] *does not hold. Then $\mathbf{f}(\mathbf{x})$ is VWA for all $\mathbf{x} \in B$.*

Proof. The assumption says that there exists $\beta > 0$ such that one has

$$\sup_{\mathbf{x}\in B} \|g_t u_{\mathbf{f}(\mathbf{x})}\Gamma\| < e^{-\beta t}$$

for arbitrarily large t (and $\Gamma \in \mathcal{S}_{n+1}$ dependent on t). In view of (2.8), for any $\mathbf{x} \in B$ this implies

$$\delta(g_t u_{\mathbf{f}(\mathbf{x})}\mathbb{Z}^{n+1}) \le c(j)\|g_t u_{\mathbf{f}(\mathbf{x})}\Gamma\|^{1/j} < c(j)e^{-\beta t/j}\,,$$

where j is the rank of Γ. Hence $\gamma_F(u_{\mathbf{f}(\mathbf{x})}\mathbb{Z}^{n+1}) \ge \beta/n$, and an application of Corollary 2.3 finishes the proof. □

We now combine the two lemmas above to obtain the desired extremality criterion.

Theorem 3.8. *Let U be an open subset $\mathbb{R}^d$, and let $\mathbf{f}$ be a map from U to $\mathbb{R}^n$ which is continuous and good. Then the following are equivalent:*

[3.8-i] *the set $\{\mathbf{x} \in U \mid \mathbf{f}(\mathbf{x}) \text{ is not VWA}\}$ is dense in U;*

[3.8-ii] $\mathbf{f}$ *is extremal* (*that is, the above set has full measure*);

[3.8-iii] *for a.e. $\mathbf{x}_0 \in U$ and any $r > 0$ there exists a ball $B \subset U$ centered at $\mathbf{x}_0$ of radius less than r satisfying* [3.6-ii];

[3.8-iv] *any ball $B \subset U$ satisfies* [3.6-ii].

Proof. Obviously [3.8-ii]⇒[3.8-i] and [3.8-iv]⇒[3.8-iii]. The implication [3.8-i]⇒[3.8-iv] is immediate from Lemma 3.7. Assuming [3.8-iii] and using the fact that $\mathbf{f}$ is good, for a.e. $\mathbf{x}_0 \in U$ one finds a ball B centered at $\mathbf{x}_0$ such that the dilated ball $3^{n+1}B$ is contained in U, and both [3.6-i] and [3.6-ii] hold. Thus Lemma 3.6 applies, and [3.8-ii] follows. □

We remark that for the equivalence of [3.8-i] and [3.8-ii] it is essential that $\mathbf{f}$ be chosen within the class of good maps. Indeed, one can consider the map $\mathbf{f}(x) = \big(x, \psi(x)\big)$, with ψ from Example 3.4. Clearly $\mathbf{f}(x)$ is VWA for any x from K, which is assumed to have positive measure. On the other hand, the restriction of $\mathbf{f}$ to any $J \in \mathcal{J}_k$, $k \in \mathbb{N}$, is nondegenerate in $\mathbb{R}^2$, hence the set $\{\mathbf{x} \in (0,1) \mid \mathbf{f}(\mathbf{x}) \text{ is not VWA}\}$ has full measure in $(0,1) \smallsetminus K$, and the latter is dense in $(0,1)$.

Our next task is to rephrase [3.6-ii]. For any $B \subset U$ let us denote by $\mathcal{F}_B$ the $\mathbb{R}$-linear span of the restrictions of $1, f_1, \dots, f_n$ to B. Then, for any ball $B \subset U$, let us assume that functions $g_1, \dots, g_s : B \to \mathbb{R}$ are chosen so that $1, g_1, \dots, g_s$ form a basis of $\mathcal{F}_B$ (here the dimension $s+1$ of $\mathcal{F}_B$ may depend on B). Further, the choice of functions $1, g_1, \dots, g_s$ defines a matrix $P \in M_{s+1,n+1}$ formed by coefficients in the expansion of $1, f_1, \dots, f_n$ as linear combinations of $1, g_1, \dots, g_s$. In other words, with the notation $\tilde{\mathbf{f}} \stackrel{\text{def}}{=} (1, f_1, \dots, f_n)$ and $\tilde{\mathbf{g}} \stackrel{\text{def}}{=} (1, g_1, \dots, g_s)$, one has

$$\tilde{\mathbf{f}}(\mathbf{x}) = \tilde{\mathbf{g}}(\mathbf{x})P \quad \forall\, \mathbf{x} \in B\,. \tag{3.4}$$

This way, the restriction of $\mathbf{f}$ to B is described by two pieces of data: the $(s+1)$-tuple $\tilde{\mathbf{g}}$ and the matrix P. We now proceed to show that, assuming the map $\mathbf{f}$ is good (which is an assumption involving $\tilde{\mathbf{g}}$), a criterion for its extremality can be written in terms of P.

Indeed, any $f \in \mathcal{F}_B$ can be written as $f(\mathbf{x}) = \tilde{\mathbf{g}}(\mathbf{x})\mathbf{v}$ for some $\mathbf{v} \in \mathbb{R}^{n+1}$, and because of the linear independence of components of $\tilde{\mathbf{g}}$ over $\mathbb{R}$, the 'supremum-on-B' norm of f, that is, $f \mapsto \sup_{\mathbf{x}\in B} |f(\mathbf{x})|$, is equivalent to $\|\mathbf{v}\|$, the constant in the equivalence depending on B and the choice of $\tilde{\mathbf{g}}$. Now recall that for any $\mathbf{w} = \sum_I w_I \mathbf{e}_I$, the Ith component of $u_{\mathbf{f}(\mathbf{x})}\mathbf{w}$ is equal to w_I if $0 \notin I$, and to

$$w_I + \sum_{i\notin I} (-1)^{l(I,i)} w_{I\cup\{i\}\smallsetminus\{0\}} f_i(\mathbf{x})$$

if I contains 0. It will be convenient to simplify the latter expression by introducing the following notation: given $I \subset \{0,\dots,n\}$ containing 0 with $|I| = j$, and an element $\mathbf{w} = \sum_I w_I \mathbf{e}_I$ of $\bigwedge^j(\mathbb{R}^{n+1})$, let us define a vector $\mathbf{c}_{I,\mathbf{w}} \in \mathbb{R}^{n+1}$ by

$$\mathbf{c}_{I,\mathbf{w}} \stackrel{\text{def}}{=} \sum_{i\notin(I\smallsetminus\{0\})} (-1)^{l(I,i)} w_{I\cup\{i\}\smallsetminus\{0\}}\mathbf{e}_i \;=\; w_I\mathbf{e}_0 + \sum_{i\notin I}(-1)^{l(I,i)} w_{I\cup\{i\}\smallsetminus\{0\}}\mathbf{e}_i\,. \tag{3.5}$$

Then the nonconstant components of $u_{\mathbf{f}(\mathbf{x})}\mathbf{w}$ can be written as $\tilde{\mathbf{f}}(\mathbf{x})\mathbf{c}_{I,\mathbf{w}} = \tilde{\mathbf{g}}(\mathbf{x})P\mathbf{c}_{I,\mathbf{w}}$; therefore the 'supremum-on-B' norm of each of the functions can be replaced by $\|P\mathbf{c}_{I,\mathbf{w}}\|$. Modifying all the norms and using (2.9), one replaces $\sup_{\mathbf{x}\in B}\|g_t u_{\mathbf{f}(\mathbf{x})}\mathbf{w}\|$ by $\max(e^{\frac{n+1-j}{n}t}\max_{0\in I}\|P\mathbf{c}_{I,\mathbf{w}}\|, e^{-\frac{j}{n}t}\max_{0\notin I}|w_I|)$. We summarize the above discussion as follows:

Proposition 3.9. *Let B be a ball in $\mathbb{R}^d$, $\mathbf{f}$ a map from B to $\mathbb{R}^n$, $\{1, g_1,...,g_s\}$ a basis of $\mathcal{F}_B$, and P a matrix satisfying* (3.4). *Then* [3.6-ii] *is equivalent to*

$$\begin{aligned}&\forall\,\beta > 0\ \ \exists\, T > 0 \ \text{ such that } \ \forall\, t \geq T\,,\ > \forall\, j = 1,\dots,n \text{ and } \forall\,\mathbf{w} \in \mathcal{S}_{n+1,j}\\ &\quad \text{one has } \max\Big(e^{\frac{n+1-j}{n}t}\max_{0\in I}\|P\mathbf{c}_{I,\mathbf{w}}\|, e^{-\frac{j}{n}t}\max_{0\notin I}|w_I|\Big) \geq e^{-\beta t}\,. \qquad (3.6)\end{aligned}$$

In other words, we have shown that the extremality of a continuous good map $\mathbf{f} : U \to \mathbb{R}^n$ is equivalent to the validity of certain Diophantine conditions involving matrices P 'coordinatizing' $\mathbf{f}|_B$. These conditions will be made more precise in the next section, and Theorem 1.2 will be obtained as a corollary.

4 Extremality Criteria for Affine Subspaces

Note that in general, as was mentioned above, it may not be possible to choose the same matrix P uniformly for all balls B in U. Let us now consider an important special case when this is possible.

Theorem 4.1. *Let $U \subset \mathbb{R}^d$ be an open subset, and let $\mathbf{g} = (g_1,\dots,g_s) : U \to \mathbb{R}^s$ be a continuous good map such that*

$$\begin{aligned}&\forall\,\mathbf{x} \in U \ \text{(equivalently, } \forall\,\mathbf{x} \text{ from a dense subset of } U\text{)}\\ &\text{the germs of } 1, g_1,\dots,g_s \text{ at } \mathbf{x} \text{ are linearly independent over } \mathbb{R}\,.\end{aligned} \qquad (4.1)$$

Also fix $P \in M_{s+1,n+1}$, and let $\mathbf{f}$ be given by (3.4). *Then* (3.6) *is equivalent to each of the following conditions:*

[4.1-i] *the set $\{\mathbf{x} \in U \mid \mathbf{f}(\mathbf{x})$ is not VWA$\}$ is non-empty;*

[4.1-ii] $\mathbf{f}$ *is extremal* (*that is, the above set has full measure*).

Proof. It is clear from (3.4) that $\mathbf{f}$ is also continuous and good. [4.1-ii] follows from (3.6) in view of Proposition 3.9 and the implication [3.8-iii]⇒[3.8-ii]. On the other hand, if (3.6) is violated, Lemmas 3.7 and Proposition 3.9 imply that every point of U has a neighborhood B such that $\mathbf{f}(\mathbf{x})$ is VWA for all $\mathbf{x} \in B$, contradicting [4.1-i]. □

Informally speaking, the assumption of Theorem 4.1 says that $\mathbf{f}$ is not 'assembled from several pieces unrelated to each other'. Without that assumption, one can easily construct examples of good continuous (and even C^∞) maps $\mathbf{f}$ satisfying [4.1-i] but not [4.1-ii].

Now suppose that $\mathbf{f}$ is real analytic and U is connected, or, more generally, that $\mathbf{f}$ is nondegenerate in some affine subspace $\mathcal{L}$ of $\mathbb{R}^n$. Then one can easily find $s \le n$ and a good s-tuple $\mathbf{g}$ satisfying (4.1). Specifically, one takes $s = \dim(\mathcal{L})$ and, as in the proof of Corollary 3.2, defines $\mathbf{g}$ to be equal to $\mathbf{h}^{-1} \circ \mathbf{f}$ where $\mathbf{h}$ is any affine map from $\mathbb{R}^s$ onto $\mathcal{L}$. Furthermore, one easily recovers the corresponding matrix P by writing $\mathbf{h}$ in the form

$$\tilde{\mathbf{h}}(\mathbf{x}) = \tilde{\mathbf{x}} P\,, \tag{4.2}$$

where as usually we have $\tilde{\mathbf{h}} \stackrel{\text{def}}{=} (1, h_1, \dots, h_n)$. It follows that for fixed $\mathcal{L}$, any $\mathbf{f}$ which is nondegenerate in $\mathcal{L}$ will satisfy the assumptions of Theorem 4.1 with some uniformly chosen P. In particular, Theorem 4.1 applies to the map $\mathbf{f} = \mathbf{h}$ given by (4.2), that is, to the subspace $\mathcal{L}$ itself. Thus we have proved

Theorem 4.2. *Let $\mathcal{L}$ be an s-dimensional affine subspace of $\mathbb{R}^n$ parametrized as in* (4.2) *with $P \in M_{s+1,n+1}$. Then each of the following conditions below is equivalent to* (3.6):

[4.2-i] *$\mathcal{L}$ contains at least one not very well approximable point;*

[4.2-ii] *$\mathcal{L}$ is extremal;*

[4.2-iii] *any smooth submanifold of $\mathcal{L}$ which is nondegenerate in $\mathcal{L}$ is extremal.*

One then recovers Theorem 1.2 as the implications [4.2-ii]⇒[4.2-iii] and [4.2-i]⇒[4.2-ii] above. Note also that Theorem 1.1 is obtained by taking $s = n$ and $P = I_{n+1}$.

Now recall that, as described in the introduction, one can by permuting variables without loss of generality parametrize $\mathcal{L}$ by (1.4) for some $A \in M_{s+1,n-s}$; that is, take P of the form $P = \begin{pmatrix} I_{s+1} & A \end{pmatrix}$. In order to restate Theorem 4.2 in terms of A, let us denote by $\mathbf{c}^+_{I,\mathbf{w}}$ (resp. $\mathbf{c}^-_{I,\mathbf{w}}$) the column vector consisting of the first $s+1$ (resp. the last $n-s$) coordinates of $\mathbf{c}_{I,\mathbf{w}}$. In other words, we have $\mathbf{c}_{I,\mathbf{w}} = \begin{pmatrix} \mathbf{c}^+_{I,\mathbf{w}} \\ \mathbf{c}^-_{I,\mathbf{w}} \end{pmatrix}$ where

$$\begin{aligned} \mathbf{c}^+_{I,\mathbf{w}} &= \sum_{i \in \{0\} \cup (\{1,\dots,s\} \smallsetminus I)} (-1)^{l(I,i)} w_{I \cup \{i\} \smallsetminus \{0\}} \mathbf{e}_i \\ &= w_I \mathbf{e}_0 + \sum_{i \in \{1,\dots,s\} \smallsetminus I} (-1)^{l(I,i)} w_{I \cup \{i\} \smallsetminus \{0\}} \mathbf{e}_i \end{aligned}$$

and

$$\mathbf{c}^-_{I,\mathbf{w}} = \sum_{i \in \{s+1,\dots,n\} \smallsetminus I} (-1)^{l(I,i)} w_{I \cup \{i\} \smallsetminus \{0\}} \mathbf{e}_i\,.$$

Then one has

Theorem 4.3. *Let $\mathcal{L}$ be an s-dimensional affine subspace of $\mathbb{R}^n$ parametrized as in* (1.4) *with $A \in M_{s+1,n-s}$. Then the following are equivalent:*

[4.3-i] *$\mathcal{L}$ is extremal* ($\Leftrightarrow$ [4.2-i] *and* [4.2-iii] *hold*);

[4.3-ii] *for any $\beta > 0$ there exists $T > 0$ such that for any $t \geq T$, $j = 1, \dots, n$ and $\mathbf{w} \in \mathcal{S}_{n+1,j}$ one has*

$$\max\left(e^{\frac{n+1-j}{n}t} \max_{0\in I} \|\mathbf{c}^+_{I,\mathbf{w}} + A\mathbf{c}^-_{I,\mathbf{w}}\|, e^{-\frac{j}{n}t} \max_{0\notin I} |w_I|\right) \geq e^{-\beta t}\,;$$

[4.3-iii] *$\forall\, j = 1, \dots, n$ and $\forall\, v > \frac{n+1-j}{j}$, $\exists\, N > 0$ such that for any $\mathbf{w} \in \mathcal{S}_{n+1,j}$ with $\max_{0\notin I} |w_I| > N$, one has*

$$\max_{0\in I} \|\mathbf{c}^+_{I,\mathbf{w}} + A\mathbf{c}^-_{I,\mathbf{w}}\| > \left(\max_{0\notin I} |w_I|\right)^{-v}.$$

Proof. The preceding argument shows the equivalence [4.3-i]$\Leftrightarrow$[4.3-ii], while the fact that [4.3-ii] and [4.3-iii] are equivalent is a special case of Lemma 2.1, where for each $j = 1, \dots, n$ one considers

$$E = \left\{\left(\max_{0\in I} \|\mathbf{c}^+_{I,\mathbf{w}} + A\mathbf{c}^-_{I,\mathbf{w}}\|, \max_{0\notin I} |w_I|\right) \,\middle|\, \mathbf{w} \in \mathcal{S}_{n+1,j}\right\}. \qquad \square$$

It is instructive to write down a special case of the above inequality corresponding to $j = 1$. That is, let us take $\mathbf{v} \in \mathbb{Z}^{n+1} \smallsetminus \{0\} = \mathcal{S}_{n+1,1}$ in place of $\mathbf{w}$; one sees that the only one-element subset I of $\{0, \dots, n\}$ for which $\mathbf{c}_{I,\mathbf{v}}$ is defined is $I = \{0\}$, and it easily follows from (3.5) that $\mathbf{c}_{\{0\},\mathbf{v}} = \mathbf{v}$. Writing $\mathbf{v} = \binom{\mathbf{p}}{\mathbf{q}}$, where $\mathbf{p} \in \mathbb{Z}^{s+1}$ and $\mathbf{q} \in \mathbb{Z}^{n-s}$, one gets $\mathbf{c}^+_{\{0\},\mathbf{v}} = \mathbf{p}$ and $\mathbf{c}^-_{\{0\},\mathbf{v}} = \mathbf{q}$. Then denoting by $\mathbf{p}'$ the vector with components $p_1, \dots, p_s$, one writes the $j = 1$ case of [4.3-iii] as follows:

[4.3-iii]$_{j=1}$ *for any $v > n$ there exist at most finitely many $\mathbf{q} \in \mathbb{Z}^{n-s}$ such that for some $\mathbf{p} \in \mathbb{Z}^{s+1}$ one has*

$$\|\mathbf{p} + A\mathbf{q}\| \leq \left\|\begin{pmatrix}\mathbf{p}'\\ \mathbf{q}\end{pmatrix}\right\|^{-v}. \tag{4.3}$$

Now observe that one can safely replace the latter inequality by

$$\|\mathbf{p} + A\mathbf{q}\| \leq \|\mathbf{q}\|^{-v}\,, \tag{4.4}$$

perhaps slightly changing v. Indeed, (4.4) clearly follows from (4.3). On the other hand, (4.4) implies that $\|\mathbf{p}\| \leq C\|\mathbf{q}\|$ for some C dependent only on A; thus, for a slightly smaller v and large enough $\|\mathbf{q}\|$, (4.3) would follow. We arrive to the conclusion that [4.3-iii]$_{j=1}$ is equivalent to (1.7).

However, let us point out that one does not need the full strength of Theorem 4.3 to see that (1.7) is one of the conditions necessary for the extremality of $\mathcal{L}$ as in (1.4). Indeed, as shown above, the assumption

$A \in \mathcal{W}_n^+(s+1, n-s)$ amounts to the existence of $v > n$ such that for infinitely many $\mathbf{q} \in \mathbb{Z}^{n-s}$ one can find $\mathbf{p} \in \mathbb{Z}^{s+1}$ satisfying (4.3). Then one can take any $\mathbf{x} \in \mathbb{R}^s$ and write

$$\left|p_0 + (\mathbf{x}, \tilde{\mathbf{x}}A)\left(\begin{smallmatrix}\mathbf{p}'\\ \mathbf{q}\end{smallmatrix}\right)\right| = |p_0 + \mathbf{x}\mathbf{p}' + \tilde{\mathbf{x}}A\mathbf{q}| = \left|\tilde{\mathbf{x}}(A\mathbf{q}+\mathbf{p})\right| \le (s+1)\|\tilde{\mathbf{x}}\|\|A\mathbf{q}+\mathbf{p}\|\,.$$

Slightly decreasing v if needed, one gets infinitely many solutions of

$$\left|p_0 + (\mathbf{x}, \tilde{\mathbf{x}}A)\begin{pmatrix}\mathbf{p}'\\ \mathbf{q}\end{pmatrix}\right| \le \left\|\begin{pmatrix}\mathbf{p}'\\ \mathbf{q}\end{pmatrix}\right\|^{-v},$$

that is, $(\mathbf{x}, \tilde{\mathbf{x}}\tilde{A})$ is proved to be VWA for all $\mathbf{x}$.

Let us now ask the following question: *could it be the case that the remaining* $n-1$ *conditions of Theorem 4.3 are redundant, that is, follow from* $[4.3\text{-iii}]_{j=1}$? The affirmative answer to this question would provide a very easy to state extremality criterion, i.e. the validity of (1.7), for affine subspaces and their submanifolds.

The answer to this question is (in general) not known to the author. However, the next result shows that the case $j = n$ of [4.3-iii] is indeed redundant.

LEMMA 4.5. *For any* $s = 1, \dots, n-1$, *any* $A \in M_{s+1,n-s}$ *and any* $\mathbf{w} \in \mathcal{S}_{n+1,n}$ *one has*

$$\max_{0\in I} \|\mathbf{c}^+_{I,\mathbf{w}} + A\mathbf{c}^-_{I,\mathbf{w}}\| \ge 1\,. \tag{4.5}$$

Proof. Denote by J_i the set $\{0, \dots, n\} \smallsetminus \{i\}$. Then any $\mathbf{w} \in \mathcal{S}_{n+1,n}$ can be written in the form $\mathbf{w} = \sum_{i=0}^n w_i \mathbf{e}_{J_i}$, and from (3.5) it follows that for $i = 1, \dots, n$ one has

$$\mathbf{c}_{J_i,\mathbf{w}} = w_i\mathbf{e}_0 + (-1)^{i-1}w_0\mathbf{e}_i\,. \tag{4.6}$$

Therefore for any $i = 1, \dots, s$ one has $\mathbf{c}^-_{J_i,\mathbf{w}} = 0$, and hence

$$\mathbf{c}^+_{J_i,\mathbf{w}} + A\mathbf{c}^-_{J_i,\mathbf{w}} = \mathbf{c}^+_{J_i,\mathbf{w}} = w_i\mathbf{e}_0 + (-1)^{i-1}w_0\mathbf{e}_i\,.$$

Consequently, (4.5) is satisfied whenever $w_0 \ne 0$. On the other hand, $w_0 = 0$, in view of (4.6), implies that $\mathbf{c}^-_{J_i,\mathbf{w}} = 0$ for any i. Taking $i > s$ for which $w_i \ne 0$, one gets $\|\mathbf{c}^+_{J_i,\mathbf{w}} + A\mathbf{c}^-_{J_i,\mathbf{w}}\| = |w_i|$. □

This, in particular, gives an affirmative answer to the above question in the case $n = 2$: a line in $\mathbb{R}^2$ given by $y = a_0 + a_1x$ is extremal if and only if $\left(\begin{smallmatrix}a_0\\ a_1\end{smallmatrix}\right) \notin \mathcal{W}_2^+(2,1)$.

It turns out that an argument similar to the proof of Lemma 4.5 produces an analogous extremality criterion for $(n-1)$-dimensional affine subspaces of $\mathbb{R}^n$ for arbitrary n:

LEMMA 4.6. *Let $A \in M_{n,1}$ be given by a column vector $\mathbf{a} \in \mathbb{R}^n$ (this corresponds to $s = n-1$). Then (4.5) holds for any $\mathbf{w} \in \mathcal{S}_{n+1,j}$ with $j > 1$.*

Proof. Our choice of s implies that for any $\mathbf{w} \in \mathcal{S}_{n+1,j}$ and any $I \ni 0$ of size j, the vector $\mathbf{c}^-_{I,\mathbf{w}}$ consists of a single component, namely

$$c^-_{I,\mathbf{w}} = \begin{cases} (-1)^{l(I,n)} w_{I\cup\{n\}\smallsetminus\{0\}} & \text{if } n \notin I \\ 0 & \text{otherwise.} \end{cases} \tag{4.7}$$

Therefore for $0, n \in I$ one can write

$$\mathbf{c}^+_{I,\mathbf{w}} + \mathbf{a}c^-_{I,\mathbf{w}} = \mathbf{c}^+_{I,\mathbf{w}} = w_I\mathbf{e}_0 + \sum_{i\in\{1,\dots,n-1\}\smallsetminus I} (-1)^{l(I,i)} w_{I\cup\{i\}\smallsetminus\{0\}}\mathbf{e}_i\,.$$

Consequently, (4.5) is satisfied whenever $w_I \neq 0$ for some I containing n. (Here we use the fact that $j > 1$: indeed, such an I must also contain some $i = 0,\dots,n-1$, and thus the ith coordinate of $\mathbf{c}^+_{gI\cup\{0\}\smallsetminus\{i\},\mathbf{w}}$ is equal to w_{gI}.) On the other hand, the assumption $w_{gI} = 0$ for all $I \ni n$, in view of (4.7), implies that $c^-_{I,\mathbf{w}} = 0$ for any I; thus (4.5) is satisfied again. Hence one can take an arbitrary I for which $w_{gI} \neq 0$ and observe that for any $i \in I$, the absolute value of the ith coordinate of $\mathbf{c}^+_{I\cup\{0\}\smallsetminus\{i\},\mathbf{w}}$ is equal to $|w_{gI}|$. □

Combining the above lemma with Theorem 4.3 and the equivalence of $[4.3\text{-iii}]_{j=1}$ and (1.7), one easily obtains Theorem 1.3 under assumption (1.5). In particular, in view of (1.8), one sees that the set of $\mathbf{a} \in \mathbb{R}^n$ for which the $(n-1)$-dimensional affine subspace of $\mathbb{R}^n$ parametrized by

$$\mathbf{x} = (x_1,\dots,x_{n-1}) \mapsto (\mathbf{x}, \tilde{\mathbf{x}}\mathbf{a}) = (x_1,\dots,x_{n-1}, a_0 + a_1x_1 + \dots + a_{n-1}x_{n-1}) \tag{4.8}$$

is not extremal, has Hausdorff dimension 1.

As was mentioned above, we are unable to show the equivalence of (1.7) and [4.3-iii] in the general case. However, let us now turn to assumption (1.6), under which that equivalence can be demonstrated by a direct proof (that is, without a reference to lattices).

Until the end of this section, let us assume $s = 1$, take A of the form $\binom{0}{\mathbf{b}}$ for a row vector $\mathbf{b} = (b_1,\dots,b_{n-1}) \in \mathbb{R}^{n-1}$, and let $\mathcal{L}$ be parametrized by (1.4); that is, $\mathcal{L}$ is a line passing through the origin given by

$$x \mapsto (x, b_1x, \dots, b_{n-1}x)\,. \tag{4.9}$$

It is clear that $A \in \mathcal{W}^+_n(2, n-1)$ if and only if $\mathbf{b} \in \mathcal{W}^+_n(1, n-1)$. Therefore in order to prove Theorem 1.3 assuming (1.6), it suffices to prove the following

PROPOSITION 4.7. *A line given by* (4.9) *is extremal whenever*

$$\mathbf{b} \notin \mathcal{W}_n^+(1, n-1)\,. \tag{4.10}$$

Proof. We follow an argument from the paper [BeBDD], where a stronger (than the extremality) property was proved for $\mathcal{L}$ as in (4.9) under the stronger (than (4.10)) assumption $\mathbf{b} \notin \mathcal{W}_n^-(1, n-1)$.

The goal is to prove that for any $v > n$, the set

$$\left\{ x \in \mathbb{R} \;\middle|\; \begin{array}{c} |p + q_0x + q_1b_1x + \cdots + q_{n-1}b_{n-1}x| \le \|\mathbf{q}\|^{-v} \\ \text{for infinitely many } \mathbf{q} = (q_0, q_1, \ldots, q_{n-1})^T \in \mathbb{Z}^n,\ p \in \mathbb{Z} \end{array} \right\} \tag{4.11}$$

has measure zero. Clearly without loss of generality one can restrict x to lie in the unit interval. Also, our usual notation $\tilde{\mathbf{b}} = (1, \mathbf{b})$ will be helpful, since the left-hand side of the inequality in (4.11) will be then written as $|p + (\tilde{\mathbf{b}}\mathbf{q})x|$.

Let us now state a lemma from which the desired result will easily follow. For $\mathbf{b} \in \mathbb{R}^{n-1}$ and $Q, v > 0$, define the set $\mathcal{A}(\mathbf{b}, v, Q)$ to be the set of $x \in [0, 1)$ for which the inequality

$$\bigl|p + (\tilde{\mathbf{b}}\mathbf{q})x\bigr| < Q^{-v} \tag{4.12}$$

holds for some $p \in \mathbb{Z}, \mathbf{q} \in \mathbb{Z}^n$ with $Q \le \|\mathbf{q}\| < 2Q$.

LEMMA 4.8. *For any* $\mathbf{b}$ *satisfying* (4.10) *and any* $v > n$ *there exists a positive constant* $C = C(\mathbf{b}, v)$ *such that for any* $Q > 1$ *one has*

$$\bigl|\mathcal{A}(\mathbf{b}, v, Q)\bigr| < CQ^{\frac{n-v}{2}}\,.$$

It is easy to see that the intersection of the set (4.11) with $[0, 1)$ is contained in

$$\bigl\{x \bigm| x \in \mathcal{A}(\mathbf{b}, v, 2^k) \text{ for infinitely many } k \in \mathbb{N}\bigr\}\,. \tag{4.13}$$

Assuming Lemma 4.8, one has $|\mathcal{A}(\mathbf{b}, v, 2^k)| < C2^{-\frac{v-n}{2}k}\ \forall\, k$, and the fact that the set (4.13) has measure zero is then immediate from the Borel–Cantelli lemma. □

It remains to write down the

Proof of Lemma 4.8. Define $\mathcal{A}^0(\mathbf{b}, v, 2^k)$ to be the set of $x \in [0, 1)$ for which (4.12) holds for some $\mathbf{q} \in \mathbb{Z}^n$ with $Q \le \|\mathbf{q}\| < 2Q$ and with $p = 0$. It is contained in a union of intervals of the form $[0, Q^{-v}/|\tilde{\mathbf{b}}\mathbf{q}|]$. Due to (4.10), there exists $c = c(v) > 0$ such that the denominator of the above fraction is not less than

$$c \cdot \max(q_1, \ldots, q_{n-1})^{-\frac{n+v}{2}} \ge c \cdot \|\mathbf{q}\|^{-\frac{n+v}{2}} < c \cdot (2Q)^{-\frac{n+v}{2}}\,.$$

Therefore one has

$$\left|\mathcal{A}^0(\mathbf{b}, v, Q)\right| < c^{-1}Q^{-v}(2Q)^{\frac{n+v}{2}} = c^{-1}2^{\frac{n+v}{2}}Q^{\frac{n-v}{2}} .$$

Now let us estimate the measure of $\mathcal{A}(\mathbf{b}, v, Q) \smallsetminus \mathcal{A}^0(\mathbf{b}, v, Q)$. Note that, assuming $p \neq 0$, inequality (4.12) can be solvable in $x \in [0,1)$ only if $|\tilde{\mathbf{b}}\mathbf{q}| > 1-Q^{-v} > 1-1/Q \geq 1/2$. For fixed p and $\mathbf{q}$, (4.12) defines an interval of length at most $2\frac{Q^{-v}}{|\tilde{\mathbf{b}}\mathbf{q}|}$, and, for fixed $\mathbf{q}$, the number of different centers of those intervals, that is, points $p/|\tilde{\mathbf{b}}\mathbf{q}|$, is at most $1 + |\tilde{\mathbf{b}}\mathbf{q}|$. Therefore one can write

$$\begin{aligned}\left|\mathcal{A}(\mathbf{b}, v, Q) \smallsetminus \mathcal{A}^0(\mathbf{b}, v, Q)\right| &\leq \sum_{\substack{\|\mathbf{q}\|<2Q,\\ |\tilde{\mathbf{b}}\mathbf{q}|>1/2}} \tfrac{2Q^{-v}}{|\tilde{\mathbf{b}}\mathbf{q}|}\big(1+|\tilde{\mathbf{b}}\mathbf{q}|\big) = 2Q^{-v} \sum_{\substack{\|\mathbf{q}\|<2Q,\\ |\tilde{\mathbf{b}}\mathbf{q}|>1/2}} \Big(1+\tfrac{1}{|\tilde{\mathbf{b}}\mathbf{q}|}\Big)\\ &\leq 2Q^{-v}(4Q)^n + 2Q^{-v} \sum_{\substack{\|\mathbf{q}\|<2Q,\\ |\tilde{\mathbf{b}}\mathbf{q}|>1/2}} \tfrac{1}{|\tilde{\mathbf{b}}\mathbf{q}|} .\end{aligned}$$

To estimate the sum in the right-hand side of the above formula, note that for fixed $q_1, \dots, q_{n-1}$ and variable q_0, the values of $\tilde{\mathbf{b}}\mathbf{q}$ form an arithmetic progression. Thus, fixing $q_1, \dots, q_{n-1}$, one gets

$$\begin{aligned}\sum_{\substack{\|q_0\|<2Q\,,\\ |\tilde{\mathbf{b}}\mathbf{q}|>1/2}} \frac{1}{|\tilde{\mathbf{b}}\mathbf{q}|} &\leq 2\left(\frac{1}{1/2} + \frac{1}{1/2+1} + \cdots + \frac{1}{1/2+2Q-1}\right)\\ &= 4\left(1 + \frac{1}{3} + \cdots + \frac{1}{4Q-1}\right) < 4\big(1+\log(4Q-1)\big) .\end{aligned}$$

Summing the above estimate over all $q_1, \dots, q_{n-1}$, one obtains

$$\left|\mathcal{A}(\mathbf{b}, v, Q) \smallsetminus \mathcal{A}^0(\mathbf{b}, v, Q)\right| \leq 2^{2n+1}Q^{n-v} + 8Q^{-v}(4Q)^{n-1}\big(1+\log(4Q-1)\big) ,$$

which is not greater than the right-hand side of the desired inequality for an appropriate value of C . □

5 Multiplicative Approximation

The dynamical approach to Diophantine problems described above has an advantage of being quite general to allow various modifications of the set-up. In particular, most of the ideas described in this paper work for the so-called multiplicative approximation. Let us briefly list all the relevant definitions. For $\mathbf{x} \in \mathbb{R}^n$ we let

$$\Pi_+(\mathbf{x}) = \prod_{i=1}^{n} |x_i|_+ \,, \quad \text{where} \quad |x|_+ = \max\bigl(|x|, 1\bigr)\,.$$

For $v > 0$ let us denote by $\mathcal{WM}_v(1,n)$ the set of row vectors $\mathbf{y} \in \mathbb{R}^n$ for which there are infinitely many $\mathbf{q} \in \mathbb{Z}^n$ such that

$$|\mathbf{yq} + p| \leq \Pi_+(\mathbf{q})^{-v/n} \quad \text{for some } p \in \mathbb{Z}\,. \tag{5.1}$$

Since the right-hand side of (5.1) is not less than that of (1.1), one clearly has $\mathcal{WM}_v(1,n) \supset \mathcal{W}_v(1,n)$; in particular, $\mathcal{WM}_n(1,n) = \mathbb{R}^n$ by Dirichlet's theorem. Also it can be shown using the Borel–Cantelli lemma that the Lebesgue measure of $\mathcal{WM}_v(1,n)$ is zero whenever $v > n$. Therefore, with the definition of *very well multiplicatively approximable* (VWMA) vectors as those $\mathbf{y} \in \mathbb{R}^n$ which are in $\mathcal{WM}_v(1,n)$ for some $v > n$, one has that almost all $\mathbf{y} \in \mathbb{R}^n$ are not VWMA.

Let us now say, following the terminology of [Sp4], that a submanifold $\mathcal{M}$ of $\mathbb{R}^n$ (resp. a smooth map $\mathbf{f}$ from an open subset U of $\mathbb{R}^d$ to $\mathbb{R}^n$) is *strongly extremal* if almost all $\mathbf{y} \in \mathcal{M}$ (resp. $\mathbf{f}(\mathbf{x})$ for a.e. $\mathbf{x} \in U$) are not VWMA. It is clear that strong extremality implies extremality, and to prove a manifold to be strongly extremal is usually a harder task than just to prove extremality. For example, the strong extremality of the curve (1.2) (that is, the multiplicative analogue of Mahler's problem) was conjectured by A. Baker in 1975 [B], and the only proof that exists as of now is based on the dynamical approach of [KM1]. In fact, the main result of the latter paper ([KM1, Theorem A], of which Theorem 1.1 is a special case) is the strong extremality of manifolds nondegenerate in $\mathbb{R}^n$ (in the analytic case this was conjectured in [Sp4]).

With the help of the approach developed in [KM1], let us now try to investigate multiplicative approximation properties of generic points of proper affine subspaces and their submanifolds by first describing the set of VWMA vectors in a dynamical language. It turns out that the actions that are relevant for this case are multi-parameter. Namely, one replaces (2.1) by

$$\begin{gathered} g_{\mathbf{t}} = \operatorname{diag}(e^t, e^{-t_1}, \dots, e^{-t_n})\,, \\ \text{where} \quad \mathbf{t} = (t_1, \dots, t_n), t_i \geq 0, \text{ and } t = \sum_{i=1}^{n} t_i\,. \end{gathered} \tag{5.2}$$

The latter notation is used throughout the section, so that whenever t and $\mathbf{t}$ appear in the same context, t stands for $\sum_{i=1}^{n} t_i$.

For the rest of this section, we mostly sketch our argument, as it is very similar to what is done in §§2–4, and only highlight important modifications. The following is a multi-parameter version of Lemma 2.1:

LEMMA 5.1. *Suppose we are given a set E of pairs $(x, \mathbf{z}) \in \mathbb{R}^{n+1}$, which is discrete and homogeneous with respect to positive integers. Then the following are equivalent:*

[5.1-i] *for any $v > n$ there exist $(x, \mathbf{z}) \in E$ with $\mathbf{z}$ arbitrarily far from 0 such that*
$$|x| \le \Pi_+(\mathbf{z})^{-v/n}; \tag{5.3}$$

[5.2-ii] *for any $\gamma > 0$ there exists an unbounded set of $\mathbf{t} \in \mathbb{R}^n_+$ for which one has*
$$e^t|x| \le e^{-\gamma t} \quad \text{and} \quad e^{-t_i}|z_i| \le e^{-\gamma t}, \quad i = 1, \dots, n\,. \tag{5.4}$$
for some $(x, \mathbf{z}) \in E \smallsetminus \{0\}$.

Proof. We follow the lines of the proof of Lemma 2.1. Assuming [5.1-i], take $v > n$ and $(x, \mathbf{z}) \in E$ satisfying (5.3), and define t by
$$e^{(1-n\gamma)t} = \Pi_+(\mathbf{z})\,, \tag{5.5a}$$
where $\gamma < 1/n$ is as in (2.5). Then for every i define t_i by
$$e^{t_i} = e^{\gamma t}|z_i|_+\,. \tag{5.5b}$$
Note that, since $|z_i| \le |z_i|_+$, this implies $e^{-t_i}|z_i| \le e^{-\gamma t}$; and note also that multiplying all the equalities (5.5b) and comparing the result with (5.5a) one can verify that $t = \sum_{i=1}^n t_i$. Then one has
$$e^t|x| \le e^t\Pi_+(\mathbf{z})^{-v/n} = e^t(e^{(1-n\gamma)t})^{-v/n} \underset{(2.5)}{=} e^{-\gamma t}\,,$$
that is, (5.4) is satisfied for this choice of x, $\mathbf{z}$ and $\mathbf{t}$; taking $\mathbf{z}$ with arbitrarily large $\Pi_+(\mathbf{z})$ produces arbitrarily large values of t.

Assume now that [5.1-ii] holds, and take $\gamma < 1/n$. Then one can find an unbounded sequence of vectors $\mathbf{t}$ and a sequence of points $(x, \mathbf{z}) \in E \smallsetminus \{0\}$ satisfying (5.4). Since for any $\mathbf{t}$ one has $t_i \ge \gamma t$ for at least one i, passing to a subsequence and reshuffling the coordinates of $\mathbf{t}$ and $\mathbf{z}$ if necessary, one can assume that for some $k = 1, \dots, n$ and all entries $\mathbf{t}$ of that sequence, one has
$$t_i \ge \gamma t \quad \text{for} \quad i \le k, \qquad \text{and} \quad t_i < \gamma t \quad \text{for} \quad i > k\,. \tag{5.6}$$
It follows from (5.6) and (5.4) that $|z_i|_+ \le e^{t_i - \gamma t}$ for $i \le k$, and $|z_i| < 1$ for $i > k$, hence
$$\Pi_+(\mathbf{z}) \le \prod_{i=1}^k |z_i|_+ \le e^{t_1 + \cdots + t_k - k\gamma t} \le e^{(1-k\gamma)t}\,. \tag{5.7}$$

Now it is time to find an appropriate v. However, because of an extra parameter k, we have to modify (2.5), namely define $v > n$ by

$$\gamma = \frac{v-n}{kv+n} \quad \Leftrightarrow \quad v = \frac{n(1+\gamma)}{1-k\gamma}\,. \tag{5.8}$$

Then the right-hand side of (5.7) is equal to $e^{\frac{n}{v}(1+\gamma)t}$, and (5.4) implies

$$|x| \le e^{-(1+\gamma)t} = \left(e^{(1-k\gamma)t}\right)^{-v/n} \underset{(5.7)}{\le} \Pi_+(\mathbf{z})^{-v/n}\,.$$

which is exactly what was needed. After that, as in the proof of Lemma 2.1, one notices that a uniform bound on $\|\mathbf{z}\|$, by the discreteness of E, implies that $(0,\mathbf{z}_0) \in E$ for some $\mathbf{z}_0$, and integral multiples of $(0,\mathbf{z}_0)$ give infinitely many $(x,\mathbf{z}) \in E$ satisfying (5.3). To finish the proof, it remains to observe that (5.8) forces v to tend to n uniformly in k as $\gamma \to 0$. □

COROLLARY 5.2. *For* $\mathbf{y} \in \mathbb{R}^n$ *and* $g_\mathbf{t}$ *as in* (5.2), *the following are equivalent:*

[5.2-i] $\mathbf{y}$ *is VWMA;*

[5.2-ii] *for some* $\gamma > 0$ *there exists an unbounded set of* $\mathbf{t} \in \mathbb{R}^n_+$ *such that*

$$\delta(g_\mathbf{t} u_\mathbf{y} \mathbb{Z}^{n+1}) \le e^{-\gamma t}\,; \tag{5.9}$$

[5.2-iii] *for some* $\gamma > 0$ *there exist infinitely many* $\mathbf{t} \in \mathbb{Z}^n_+$ *such that* (5.9) *holds.*

Note that Corollary 2.2 in [KM1] provides the implication [5.2-i]⇒ [5.2-iii].

Proof. Taking $E = u_\mathbf{y}\mathbb{Z}^{n+1}$, one sees that (5.9) amounts to the validity of (5.4) for some $(x,\mathbf{z}) \in E \setminus \{0\}$. The rest of the argument mimics the proof of Corollary 2.3. □

From the above corollary and Theorem 3.5 it is not hard to derive multiplicative analogues of extremality criteria of §3 and §4. The crucial condition to consider is an analogue of [3.6-ii]: if $B \subset \mathbb{R}^d$ is a ball and $\mathbf{f}$ a map from B to $\mathbb{R}^n$, it is important to check whether or not

$$\begin{gathered}\forall\,\beta > 0 \;\; \exists\, T = T(\beta) > 0 \text{ such that for any } \mathbf{t} \in \mathbb{R}^n_+ \text{ with } t \ge T \\ \text{and any } \Gamma \in \mathcal{S}_{n+1}, \text{ one has } \sup_{\mathbf{x}\in B} \|g_t u_{\mathbf{f}(\mathbf{x})}\Gamma\| \ge e^{-\beta t}\,.\end{gathered} \tag{5.10}$$

The following can be proved by a straightforward repetition of the argument of §3:

- (cf. Lemma 3.7) if (5.10) does not hold, then $\mathbf{f}(\mathbf{x})$ is VWMA for all $\mathbf{x} \in B$;
- (cf. Lemma 3.6) if $\mathbf{f}$ is continuous, defined on $3^{n+1}B$ and satisfies [3.6-i] and (5.10), then $\mathbf{f}(\mathbf{x})$ is not VWMA for a.e. $\mathbf{x} \in B$.

Therefore one has

Theorem 5.3. *For U and $\mathbf{f}$ as in Theorem 3.8, the following are equivalent:*

[5.3-i] *the set $\{\mathbf{x} \in U \mid \mathbf{f}(\mathbf{x})$ is not VWMA$\}$ is dense in U;*

[5.3-ii] *$\mathbf{f}$ is strongly extremal (that is, the above set has full measure);*

[5.3-iii] *for a.e. $\mathbf{x}_0 \in U$ and any $r > 0$ there exists a ball $B \subset U$ centered at $\mathbf{x}_0$ of radius less than r satisfying (5.10);*

[5.3-iv] *any ball $B \subset U$ satisfies (5.10).*

In order to express (5.10) in Diophantine language, we need some more notation: for $\mathbf{t} \in \mathbb{R}^n_+$ and $I \subset \{0, \dots, n\}$, let

$$t_I \stackrel{\text{def}}{=} \sum_{i \in I \setminus \{0\}} t_i \,.$$

Then, taking $\mathbf{f}$, B, $\{1, g_1, \dots, g_s\}$ and P as in Proposition 3.9, one can observe that (5.10) can be written in the form

$$\begin{aligned} &\forall\, \beta > 0 \;\; \exists\, T > 0 \;\text{ such that }\; \forall\, \mathbf{t} \in \mathbb{R}^n_+ \text{ with } t \geq T \text{ and } \forall\, \mathbf{w} \in \mathcal{S}_{n+1} \\ &\qquad \text{one has} \quad \max\left(e^{t - t_I} \max_{0 \in I} \|P\mathbf{c}_{I,\mathbf{w}}\|, e^{-t_I} \max_{0 \notin I} |w_I|\right) \geq e^{-\beta t} \,. \end{aligned} \tag{5.11}$$

Thus Theorem 5.3 implies

Theorem 5.4. *Let U, $\mathbf{f}$, $\mathbf{g}$ and P be as in Theorem 4.1. Then (5.11) is equivalent to each of the following conditions:*

[5.4-i] *the set $\{\mathbf{x} \in U \mid \mathbf{f}(\mathbf{x})$ is not VWMA$\}$ is non-empty;*

[5.4-ii] *$\mathbf{f}$ is strongly extremal (that is, the above set has full measure).*

Taking P of the form $P = \begin{pmatrix} I_{s+1} & A \end{pmatrix}$, one deduces

Theorem 5.5. *The following are equivalent for an s-dimensional affine subspace $\mathcal{L}$ of $\mathbb{R}^n$ parametrized as in (1.4) with $A \in M_{s+1,n-s}$:*

[5.5-i] *$\mathcal{L}$ contains at least one not VWMA point;*

[5.5-ii] *$\mathcal{L}$ is strongly extremal;*

[5.5-iii] *any smooth submanifold of $\mathcal{L}$ which is nondegenerate in $\mathcal{L}$ is strongly extremal;*

[5.5-iv] *for any $\beta > 0$ there exists $T > 0$ such that for any $\mathbf{t} \in \mathbb{R}^n_+$ with $t \geq T$, and any $\mathbf{w} \in \mathcal{S}_{n+1}$ one has*

$$\max\left(\max_{0 \in I} e^{t - t_I} \|\mathbf{c}^+_{I,\mathbf{w}} + A\mathbf{c}^-_{I,\mathbf{w}}\|, \max_{0 \notin I} e^{-t_I} |w_I|\right) \geq e^{-\beta t} \,. \tag{5.12}$$

This, in particular, proves Theorem 1.4, as well as gives a criterion for the strong extremality of $\mathcal{L}$ written in terms of Diophantine properties of the parametrizing matrix A.

In general, it seems to be a hard problem to simplify condition [5.5-iv]. However, in view of computations made in §4, this can be easily done in the case $s = n-1$, that is, when $\mathcal{L}$ is a codimension one subspace. Namely, due to Lemma 4.6, one knows that (5.12) always holds when $s = n-1$ and $\mathbf{w} \in \mathcal{S}_{n+1,j}$ with $j > 1$; thus it suffices to handle the case $j = 1$. As in §4, one then uses $\mathbf{v} = \binom{\mathbf{p}}{q} \in \mathbb{Z}^{n+1} \smallsetminus \{0\}$ in place of $\mathbf{w}$ (here $\mathbf{p} \in \mathbb{Z}^n$ and $q \in \mathbb{Z}$) and notices that $\mathbf{c}_{I,\mathbf{v}}$ is defined only for $I = \{0\}$, with $\mathbf{c}_{\{0\},\mathbf{v}} = \mathbf{v}$, $\mathbf{c}^+_{\{0\},\mathbf{v}} = \mathbf{p}$ and $\mathbf{c}^-_{\{0\},\mathbf{v}} = q$. Replacing A by a column vector $\mathbf{a} = (a_0, a_1, \dots, a_{n-1})^T \in \mathbb{R}^{n-1}$ and letting $\mathbf{p} = (p_0, p_1, \dots, p_{n-1})^T$ and $\mathbf{p}' = (p_1, \dots, p_{n-1})^T$, one gets

Corollary 5.6. *The following are equivalent for an* $(n-1)$*-dimensional affine subspace* $\mathcal{L}$ *of* $\mathbb{R}^n$ *parametrized by* (4.8)*:*

[5.6-i] $\mathcal{L}$ *is strongly extremal;*

[5.6-ii] *for any* $\beta > 0$ *there exists* $T > 0$ *such that for any* $\mathbf{t} \in \mathbb{R}^n_+$ *with* $t \geq T$ *and any* $\binom{\mathbf{p}}{q} \in \mathbb{Z}^{n+1} \smallsetminus \{0\}$ *one has*

$$\max\big(e^t\|\mathbf{p} + \mathbf{a}q\|, e^{-t_1}|p_1|, \dots, e^{-t_{n-1}}|p_{n-1}|, e^{-t_n}|q|\big) \geq e^{-\beta t}\,.$$

[5.6-iii] *for any* $v > n$ *there exists* $K > 0$ *such that for any* $\mathbf{p} \in \mathbb{Z}^n$ *and* $q \in \mathbb{Z}$ *with* $\max(\|\mathbf{p}'\|, |q|) > K$*, one has*

$$\|\mathbf{p} + \mathbf{a}q\| > \Pi_+(\mathbf{p}', q)^{-v/n}\,.$$

Proof. The only part that requires a comment is the equivalence [5.6-ii]⇔[5.6-iii], which is a special case of Lemma 5.1 with

$$E = \big\{(\|\mathbf{p} + \mathbf{a}q\|, \mathbf{p}', q) \mid \mathbf{p} \in \mathbb{Z}^n, q \in \mathbb{Z}\big\}\,. \qquad \square$$

Corollary 5.7. *Let* $\mathcal{L}$ *be parametrized by* (4.8), *and let*

$$k = \#\{1 \leq i \leq n-1 \mid a_i \neq 0\}\,.$$

Then $\mathcal{L}$ *is strongly extremal iff* $\mathbf{a} \notin \mathcal{W}^+_{k+1}(n,1)$.

Proof. By the previous corollary, the fact that $\mathcal{L}$ is not strongly extremal is equivalent to saying that for some $v > n$ there exist $(\mathbf{p}, q) \in \mathbb{Z}^n$ with arbitrarily large $\max(\|\mathbf{p}'\|, |q|)$ such that

$$\|\mathbf{p} + \mathbf{a}q\| \leq \Pi_+(\mathbf{p}', q)^{-v/n}. \tag{5.13}$$

Equivalently, there exists a sequence of solutions of (5.13) with p_i arbitrarily close to $a_i q$ for all $i = 1, \dots, n-1$, which in particular happens if and only if p_i is equal to zero for any i with $a_i = 0$. Hence the ratio of $\Pi_+(\mathbf{p}', q)$ and $|q|^{k+1}$ is bounded from both sides, and, slightly changing v, one gets infinitely many solutions of $\|\mathbf{p} + \mathbf{a}q\| \leq |q|^{-\frac{v}{n}(k+1)}$. $\square$

In particular, one can see that the set of vectors $\mathbf{a} \in \mathbb{R}^n$ for which the map (4.8) is not strongly extremal is slightly bigger than the one corresponding to non-extremality, agrees with the latter outside of all the coordinate planes, and still has Hausdorff dimension 1.

6 Further Generalizations and Open Questions

6.1 As was mentioned before, it would be very interesting to find out whether Theorem 1.3 can be extended to the cases when the rank of A is greater than one. If it cannot, it would be nice to find a reasonable description of the set of non-extremal subspaces, e.g. such that would allow to compute its Hausdorff dimension. Similar questions are open in the case of multiplicative approximation; in particular, it is not clear, except for the case of hyperplanes, how much smaller than the set of extremal subspaces is the set of strongly extremal ones.

6.2 Given that one of the main results of this paper is that the extremality of an affine subspace is inherited by its nondegenerate submanifolds, one can ask whether any 'passage of information' takes place when the subspace is not extremal. Let us introduce the following definition: for $A \in M_{m,n}(\mathbb{R})$, define the *Diophantine exponent* $\omega(A)$ of A by

$$\omega(A) \stackrel{\text{def}}{=} \sup\left\{v \mid A \in \mathcal{W}_v(m,n)\right\}.$$

Clearly $n/m \le \omega(A) \le \infty$ for all A, and A is VWA iff $\omega(A) > n/m$. Now, for a map $\mathbf{f} : U \to \mathbb{R}^n$ define $\omega(\mathbf{f})$ to be the essential infimum of $\omega(\mathbf{f}(\cdot))$, i.e.

$$\omega(\mathbf{f}) \stackrel{\text{def}}{=} \sup\left\{v \mid |\{\mathbf{x} \in U \mid \mathbf{f}(\mathbf{x}) \in \mathcal{W}_v(1,n)\}| > 0\right\}.$$

Naturally, if $\mathcal{M}$ is a smooth manifold, we let the Diophantine exponent $\omega(\mathcal{M})$ of $\mathcal{M}$ to be the Diophantine exponent of its parametrizing map.

Clearly a manifold (map) is extremal iff its Diophantine exponent is the smallest possible, i.e. is equal to $n = \omega(\mathbb{R}^n)$. Now the following two questions arise:

- Is it true that the Diophantine exponent of an affine subspace $\mathcal{L}$ of $\mathbb{R}^n$ is inherited by manifolds nondegenerate in $\mathcal{L}$?
- how to efficiently describe the class of affine subspaces with a given Diophantine exponent?

The answer to the first question is 'yes', which can be proved by a refinement of the 'dynamical' approach developed in [KM1] and the present paper. Furthermore, similarly to Theorems 3.8 and 4.2–4.3, for any $v \ge n$ one can write down necessary and sufficient conditions for a good (resp. affine) map $\mathbf{f}$ to have $\omega(\mathbf{f}) \le v$.

The answer to the second question is as obscure as the problem of extending Theorem 1.3 beyond the cases (1.5) and (1.6). Indeed, the conclusion of the latter theorem amounts to saying that, for $\mathcal{L}$ parametrized as in (1.4), $\omega(A) \le n$ implies $\omega(\mathcal{L}) = n$. In general, one can easily prove that $\omega(\mathcal{L}) \ge \omega(A)$; thus one is left to ask whether $\omega(\mathcal{L})$ is always equal to $\max\big(n, \omega(A)\big)$ (this can be verified in the 'rank-one' cases (1.5) and (1.6)). All this is going to be the topic of a forthcoming paper [K2].

6.3 Even more generally, one can replace the right-hand side of (1.1) by an arbitrary function of $\|\mathbf{q}\|$. Let us specialize to the case of row vectors and introduce the following definition: for a non-increasing function $\psi : \mathbb{N} \to (0, \infty)$, define $\mathcal{W}_\psi(1, n)$ to be the set of $\mathbf{y} \in \mathbb{R}^n$ for which there are infinitely many $\mathbf{q} \in \mathbb{Z}^n$ such that

$$\|\mathbf{y}\mathbf{q} + p\| \le \psi(\|\mathbf{q}\|) \quad \text{for some } p \in \mathbb{Z}\,.$$

It is a theorem of A.V. Groshev ([G], see also [S1]) that almost no (resp. almost all) $\mathbf{y} \in \mathbb{R}^n$ belong to $\mathcal{W}_\psi(1, n)$ if the series

$$\sum_{k=1}^{\infty} k^{n-1}\psi(k) \tag{6.1}$$

converges (resp. diverges). (The case of convergence easily follows from the Borel–Cantelli lemma.) It has been recently proved, see [BeKM] for the convergence part and [BeBKM] for the divergence part, that the same dichotomy takes place for any nondegenerate submanifold of $\mathbb{R}^n$; in other words, for a smooth map $\mathbf{f} : U \to \mathbb{R}^n$ which is nondegenerate in $\mathbb{R}^n$, one has

$$\begin{gathered}\mathbf{f}(\mathbf{x}) \in \mathcal{W}_\psi(1, n) \text{ for almost no (resp. almost all) } \mathbf{x} \in U \\ \text{if the series (6.1) converges (resp. diverges).}\end{gathered} \tag{6.2}$$

It is natural to expect a similar dichotomy for any smooth submanifold $\mathcal{M}$ of $\mathbb{R}^n$, with the convergence/divergence of (6.1) replaced by another 'dividing line' condition, possibly involving the Diophantine exponent of $\mathcal{M}$. The following problems remain open:

- Is it true that the aforementioned 'dividing line' condition of an affine subspace $\mathcal{L}$ of $\mathbb{R}^n$ is inherited by manifolds nondegenerate in $\mathcal{L}$?
- Given an affine subspace, find its 'dividing line' condition; or, vice versa, describe the class of subspaces with a given 'dividing line'.

The only result along these lines known to the author is the paper [BeBDD], where it is shown that the convergence/divergence of (6.1) serves as the 'dividing line' condition for one-dimensional subspaces of $\mathbb{R}^n$ of the form (1.9) with $\mathbf{b} \notin \mathcal{W}_n^-(1, n-1)$.

Finally, let us note that the paper [BeKM] also contains a more general (in particular, multiplicative) version of the convergence case of (6.2), and it is of considerable interest to see if the argument from that paper can be applied to the set-up of proper affine subspaces and their nondegenerate submanifolds.

Acknowledgements. The author is grateful to Professor Gregory Margulis for conjecturing some of the results of this paper and for many stimulating and enlightening discussions, and to Victor Beresnevich, Barak Weiss and the reviewer for useful remarks.

References

[B] A. Baker, Transcendental Number Theory, Cambridge Univ. Press, Cambridge, 1975.

[BeBDD] V. Beresnevich, V. Bernik, H. Dickinson, M.M. Dodson, On linear manifolds for which the Khinchin approximation theorem holds (in Belorussian), Vestsi Nats. Acad. Navuk Belarusi. Ser. Fiz.-Mat. Navuk (2000), 14–17.

[BeBKM] V. Beresnevich, V. Bernik, D. Kleinbock, G.A. Margulis, Metric diophantine approximation: the Khintchine–Groshev theorem for nondegenerate manifolds, Moscow Math. J. 2:2 (2002), 203–225.

[BeD] V. Bernik, M.M. Dodson, Metric Diophantine approximation on Manifolds, Cambridge Univ. Press, Cambridge, 1999.

[BeKM] V. Bernik, D. Kleinbock, G.A. Margulis, Khintchine-type theorems on manifolds: the convergence case for standard and multiplicative versions, Internat. Math. Res. Notices 9 (2001), 453–486.

[C] J.W.S. Cassels, An Introduction to Diophantine Approximation, Cambridge Tracts in Math. 45, Cambridge Univ. Press, Cambridge, 1957.

[D] S.G. Dani, Divergent trajectories of flows on homogeneous spaces and Diophantine approximation, J. reine angew. Math. 359 (1985), 55–89.

[Do] M.M. Dodson, Hausdorff dimension, lower order and Khintchine's theorem in metric Diophantine approximation, J. reine angew. Math. 432 (1992), 69–76.

[G] A.V. Groshev, Une théorème sur les systèmes des formes linéaires, Dokl. Akad. Nauk SSSR 9 (1938), 151–152 .

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

[K2] D. Kleinbock, An extension of quantitative nondivergence and applications to Diophantine exponents, in preparation.

[KLW] D. Kleinbock, E. Lindenstrauss, B. Weiss, On fractal measures and Diophantine approximation, preprint, 2003.

[KM1] D. Kleinbock, G.A. Margulis, Flows on homogeneous spaces and Diophantine approximation on manifolds, Ann. Math. 148 (1998), 339–360.

[KM2] D. Kleinbock, G.A. Margulis, Logarithm laws for flows on homogeneous spaces, Invent. Math. 138 (1999), 451–494.

[KSS] D. Kleinbock, N. Shah, A. Starkov, Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory, in "Handbook on Dynamical Systems, Volume 1A", Elsevier Science, North Holland (2002), 813–930.

[M] K. Mahler, Über das Mass der Menge aller S-Zahlen, Math. Ann. 106 (1932), 131–139.

[Ma] G.A. Margulis, On the action of unipotent group in the space of lattices, Proceedings of the Summer School on Group Representations, (Budapest 1971), Académiai Kiado, Budapest (1975), 365–370.

[R] M.S. Raghunathan, Discrete Subgroups of Lie Groups, Springer-Verlag, Berlin and New York, 1972 .

[S1] W. Schmidt, A metrical theorem in Diophantine approximation, Canadian J. Math. 12 (1960), 619–631.

[S2] W. Schmidt, Metrische Sätze über simultane Approximation abhänginger Grössen, Monatsch. Math. 68 (1964), 154–166.

[S3] W. Schmidt, Diophantine approximation and certain sequences of lattices, Acta Arith. 18 (1971), 195–178.

[Si] C.L. Siegel, Zur Reduktionstheorie quadratischer Formen, The Mathematical Society of Japan, Tokyo, 1959.

[Sp1] V. Sprindžuk, More on Mahler's conjecture (in Russian), Doklady Akad. Nauk SSSR 155 (1964), 54–56; English transl. Soviet Math. Dokl 5 (1964), 361–363.

[Sp2] V. Sprindžuk, Mahler's Problem in Metric Number Theory, Translations of Mathematical Monographs 25, Amer. Math. Soc., Providence, RI, 1969.

[Sp3] V. Sprindžuk, Metric Theory of Diophantine Approximations, John Wiley & Sons, New York-Toronto-London, 1979.

[Sp4] V. Sprindžuk, Achievements and problems in Diophantine approximation theory, Russian Math. Surveys 35 (1980), 1–80.

[W] B. Weiss, Dynamics on parameter spaces: submanifold and fractal subset questions, in "Rigidity in Dynamics and Geometry (Cambridge 2000)" (M. Burger, A. Iozzi, eds.), Springer, Berlin (2002), 425–440.

Dmitry Kleinbock, Department of Mathematics, Brandeis University, Waltham, MA 02454-9110, USA kleinboc@brandeis.edu

Submitted: March 2002
Revision: August 2002