ON FRACTAL MEASURES AND DIOPHANTINE APPROXIMATION

DMITRY KLEINBOCK, ELON LINDENSTRAUSS, AND BARAK WEISS

Abstract. We study diophantine properties of a typical point with respect to measures on $\mathbb{R}^n$. Namely, we identify geometric conditions on a measure μ on $\mathbb{R}^n$ guaranteeing that μ-almost every $\mathbf{y} \in \mathbb{R}^n$ is not very well multiplicatively approximable by rationals. Measures satisfying our conditions are called 'friendly'. Examples include smooth measures on nondegenerate manifolds, thus the present paper generalizes the main result of [KM]. Another class of examples is given by measures supported on self-similar sets satisfying the open set condition, as well as their products and pushforwards by certain smooth maps.

1. Introduction

The metric theory of diophantine approximation is concerned with the following question: if $\mathbf{y} \in \mathbb{R}^n$ is a typical point in the sense of Lebesgue measure, how well can $\mathbf{y}$ be approximated by rational vectors $\mathbf{p}/q$, in terms of the size of q. Definitive answers to this and similar questions were obtained in work of A. Khintchine, A.V. Groshev and others in the 1920s and 1930s. A conjecture of K. Mahler from the 1930s, which was settled three decades later by V. Sprindžuk, led to the theory of diophantine approximation on manifolds, where instead of studying diophantine properties of almost every point in $\mathbb{R}^n$, one considers a more delicate question regarding diophantine properties of almost every point on a proper submanifold of this space.

In order to state more precisely Mahler's conjecture, as well as the results of G.A. Margulis and the first-named author [KM] which were the starting point for this work, we introduce some standard notions from the theory of diophantine approximation.

A point $\mathbf{y} \in \mathbb{R}^n$ is said to be *very well approximable* if for some $\delta > 0$ there are infinitely many solutions $\mathbf{p} \in \mathbb{Z}^n$ and $q \in \mathbb{Z}_+$ to the inequality

$$\|q\mathbf{y} - \mathbf{p}\| < q^{-(\frac{1}{n}+\delta)}\,. \tag{1.1}$$

A slightly less restrictive notion than very well approximable is the following: a point $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ is said to be *very well multiplicatively approximable* if for some $\delta > 0$ there are infinitely many solutions $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{Z}^n$ and $q \in \mathbb{Z}_+$ to

$$\prod_{i=1}^{n} |qy_i - p_i| < q^{-(1+\delta)}\,. \tag{1.2}$$

It is an immediate consequence of the pigeonhole principle that there exists $C > 0$ such that for any $\mathbf{y}$ there are infinitely many solutions to $\|q\mathbf{y} - \mathbf{p}\| < Cq^{-1/n}$. It is also well-known that the set of very well approximable points has Lebesgue measure zero but full Hausdorff dimension, and the same is true for the set of very well multiplicatively approximable points.

Mahler conjectured that for almost every $x \in \mathbb{R}$, the point

$$(x, x^2, \dots, x^n) \tag{1.3}$$

is not very well approximable. More generally, we shall say that a submanifold $\mathcal{M} \subset \mathbb{R}^n$ is *extremal* if almost every point on $\mathcal{M}$ (with respect to the smooth measure class) is not very well approximable. If almost every point on $\mathcal{M}$ is not very well multiplicatively approximable then $\mathcal{M}$ is said to be *strongly extremal*. In these terms, Mahler's conjecture states that the algebraic curve (1.3) is extremal. The line

$$\{(x, x, \dots, x) : x \in \mathbb{R}\}$$

on the other hand is easily seen to be non-extremal.

In [KM] it was shown that *nondegenerate* submanifolds of $\mathbb{R}^n$ (see §2 for a definition; an example is any real analytic submanifold not contained in any proper affine subspace of $\mathbb{R}^n$) are strongly extremal, thereby settling a conjecture of Sprindžuk [Sp2].

In this paper the primary objects of study are not submanifolds but measures. We say that a measure μ on $\mathbb{R}^n$ is *extremal* (resp., *strongly extremal*) if μ-almost every point is not very well approximable (resp., not very well multiplicatively approximable). This viewpoint is more general, and includes the discussion of submanifolds as a special case: by definition, a submanifold $\mathcal{M}$ is extremal (resp., strongly extremal) if and only if the induced Riemannian measure on $\mathcal{M}$ (considered as a measure on $\mathbb{R}^n$) is extremal (resp., strongly extremal).

In [W1], the third-named author treated the one-dimensional case, and showed that any measure μ on $\mathbb{R}$ satisfying a certain geometric decay condition introduced by W. Veech [V] (for example, the Hausdorff measure on the standard Cantor ternary set) is extremal (which in the one-dimensional case is equivalent to being strongly extremal). We refer the reader to [W2] for further discussion. Since rational points in the higher dimensional case are distributed much less regularly than in the one-dimensional case, the multi-dimensional case we address here is harder.

In this paper we identify purely geometric conditions on measures which are sufficient to guarantee strong extremality. Measures that satisfy our conditions are called *friendly* measures (a somewhat fuzzy abbreviation of *Federer, nonplanar and decaying*); these conditions are defined in §2, and our main result is

Theorem 1.1. *Let μ be a friendly measure on $\mathbb{R}^n$. Then μ is strongly extremal.*

The class of friendly measures includes volume measures on smooth manifolds considered in [KM], thus the above theorem generalizes the main result of that paper.

Various measures supported on fractal subsets of $\mathbb{R}^n$ are also friendly, for instance (these are all special cases of more general results stated in §2):

(1) Hausdorff measures on self-similar sets satisfying the open set condition (such as the Cantor set, Koch snowflake or Sierpinski gasket), provided that the set is not contained in the union of finitely many proper hyperplanes.
(2) Pushforwards of such measures by nonsingular nondegenerate maps.

The class of friendly measures is also closed with respect to products of measures, giving rise to further examples.

Overview: We start in §2 by defining and describing the class of friendly measures. We also review some basic facts about self-similar sets, and give precise statements of our results.

In §§3–5 we prove Theorem 1.1. Our proof follows the method developed in [KM]. This involves translating the diophantine properties we are interested in to properties of trajectories for the action of a semigroup of diagonal matrices on the noncompact homogeneous space $\mathrm{SL}_{n+1}(\mathbb{R})/\mathrm{SL}_{n+1}(\mathbb{Z})$. Such a translation is classical for the case $n = 1$ and was used for $n \geq 2$ in the work of W. Schmidt [Sch2], S. G. Dani [Da1], Margulis, the first-named author, and others. We discuss this correspondence and reduce Theorem 1.1 to a quantitative nondivergence result for the aforementioned action (Theorem 4.3) in §§3–4.

In §5 we prove Theorem 4.3, which is an extension of the results of [KM, §§4–5]. We remark that the argument involved played an important role in the study of unipotent flows, and refer the reader to the surveys [KSS, Chapter 3] and [K1] for further description of this method and historical background. Our proof is very close to that of [KM], but is presented from a somewhat different perspective, and incorporates several minor improvements which were useful in our framework.

We give a non-uniform variant of the friendly condition in §6, and outline how the proofs in §§3–5 can be modified to establish that measures satisfying that condition are strongly extremal.

After that we exhibit some friendly measures. In §7 we prove that the pushforward of a measure from a certain class of measures on $\mathbb{R}^d$ (more restrictive than the class of friendly measures) to $\mathbb{R}^n$ via a nonsingular nondegenerate map is friendly. For the special case of Lebesgue measure, this follows easily from [KM, Proposition 3.4], but that proof does not apply to more general measures, and the argument we give is new.

In §8 we discuss Hausdorff measures on self-similar sets satisfying the open set condition, and show that such measures are friendly, and moreover satisfy the stronger property needed for the pushforward result.

In §9 we show that the class of friendly measures is closed with respect to Cartesian products.

We conclude the paper with a discussion of possible extensions and generalizations of the main results and a list of open questions.

Acknowledgements: Part of the work was done during the authors' collaboration at the Newton Institute (Cambridge), ETH (Zurich) and at Brandeis University; the hospitality of these institutions is gratefully acknowledged. We also thank Amnon Besser for his help in producing the figures. This research was supported in part by NSF grant DMS-0196124, NSF grant DMS-0140497, and BSF grant 2000247.

2. Friendly measures

We begin by introducing some notation. For a point x in a metric space and $r > 0$, $B(x, r)$ stands for the open ball of radius r centered at x. The standard inner product of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle$. For an affine subspace $\mathcal{L} \subset \mathbb{R}^n$ we denote by $d_{\mathcal{L}}(\mathbf{x})$ the (Euclidean) distance from $\mathbf{x}$ to $\mathcal{L}$. By $\mathcal{L}^{(\varepsilon)}$ we denote the ε-neighborhood of $\mathcal{L}$, that is, the set

$$\mathcal{L}^{(\varepsilon)} \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbb{R}^n : d_{\mathcal{L}}(\mathbf{x}) < \varepsilon\} . \tag{2.1}$$

If $B \subset X$ and f is a real-valued function on X, let

$$\|f\|_B \stackrel{\text{def}}{=} \sup_{x \in B} |f(x)| .$$

If μ is a measure on X such that $\mu(B) > 0$, 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.

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

Note that for $B \subset \mathbb{R}^n$ and an affine hyperplane (that is, a translate of an $(n-1)$-dimensional linear subspace) $\mathcal{L} \subset \mathbb{R}^n$, the quantity $\|d_{\mathcal{L}}\|_B$ measures the 'width of B with respect to $\mathcal{L}$', i.e., the infimum of ε for which $B \subset \mathcal{L}^{(\varepsilon)}$. Similarly, $\|d_{\mathcal{L}}\|_{\mu,B}$ measures the 'μ-essential width of B with respect to $\mathcal{L}$', which, if B is open, coincides with $\inf\{\varepsilon : \mu(B \smallsetminus \mathcal{L}^{(\varepsilon)}) = 0\}$.

In what follows, μ will stand for a locally finite Borel measure on a σ-compact metric space. Our first definition works in the context of arbitrary metric spaces X. If $D > 0$ and $U \subset X$ is an open subset, let us say that μ is *D-Federer on U* if for all $x \in \operatorname{supp} \mu \cap U$ one has

$$\frac{\mu\big(B(x, 3r)\big)}{\mu\big(B(x, r)\big)} < D \tag{2.2}$$

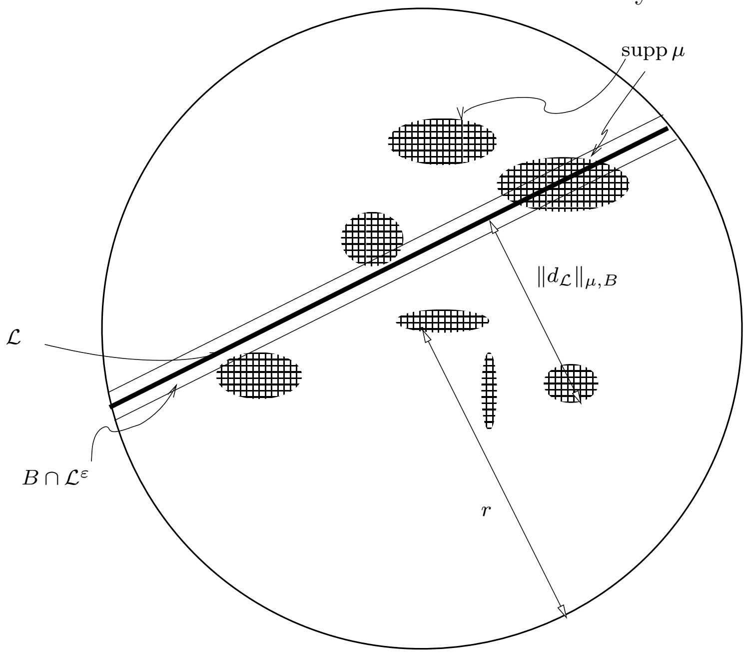
whenever $B(x, 3r) \subset U$.

Equivalently (with a different value of D) one can replace '3' in (2.2) by any number bigger than 1. Another equivalent condition is the existence of c, $\beta > 0$ such that for all $x \in \operatorname{supp} \mu \cap U$ and every $0 < \varepsilon \le r$ with $B(x, r) \subset U$ one has

$$\frac{\mu\big(B(x, \varepsilon)\big)}{\mu\big(B(x, r)\big)} \ge c \left(\frac{\varepsilon}{r}\right)^\beta . \tag{2.3}$$

A version of this definition with $U = X$ is also known as the 'doubling property' and plays an important role in geometric measure theory; we refer the reader to [MU] for references and examples. However the class of measures which are D-Federer on X is too narrow for our purposes, since it is not closed with respect to restrictions to open subsets of X. Thus we propose the following

FIGURE 1. Absolute and relative decay



localized modification: we will say that a measure μ on X is *Federer*[1] if for μ-a.e. $x \in X$ there exist a neighborhood U of x and $D > 0$ such that μ is D-Federer on U.

For the next definitions we take $X = \mathbb{R}^n$. Say that μ on $\mathbb{R}^n$ is *nonplanar* if $\mu(\mathcal{L}) = 0$ for any affine hyperplane $\mathcal{L}$ of $\mathbb{R}^n$. Given C, $\alpha > 0$ and an open subset U of $\mathbb{R}^n$, say that μ is (C, α)-*decaying on* U if for any non-empty open ball $B \subset U$ centered in $\operatorname{supp}\mu$, any affine hyperplane $\mathcal{L} \subset \mathbb{R}^n$, and any $\varepsilon > 0$ one has

$$\mu\left(B \cap \mathcal{L}^{(\varepsilon)}\right) \le C \left(\frac{\varepsilon}{\|d_{\mathcal{L}}\|_{\mu,B}}\right)^{\alpha} \mu(B)\,. \tag{2.4}$$

Also say that μ is *absolutely* (C, α)-*decaying on* U if for any B, $\mathcal{L}$ and ε as above, one has

$$\mu\left(B \cap \mathcal{L}^{(\varepsilon)}\right) \le C \left(\frac{\varepsilon}{r}\right)^{\alpha} \mu(B)\,, \tag{2.5}$$

where r is the radius of B. Equivalently (up to a slight change of C) one can replace r in (2.5) with $\|d_{\mathcal{L}}\|_B$, since the latter is between r and $2r$ whenever the left hand side of (2.5) is positive.

We will say that μ is *decaying* (resp. *absolutely decaying*) if for for μ-a.e. $\mathbf{y}_0 \in \mathbb{R}^n$ there exist a neighborhood U of $\mathbf{y}_0$ and $C, \alpha > 0$ such that μ is (C, α)-decaying (resp. *absolutely* (C, α)-*decaying*) on U. Finally, let us say that μ is *friendly* if it is Federer, nonplanar and decaying.

One can easily see, comparing (2.4) and (2.5), that any absolutely decaying measure is decaying, and it is also clear that any absolutely decaying measure

[1]This term was used in [Sa] to describe a larger class of measures, which we refer to as *non-uniformly Federer*, see §6.

is nonplanar. Thus any absolutely decaying Federer measure is friendly, in particular, so is Lebesgue measure on $\mathbb{R}^n$. Also, the class of friendly measures is clearly closed with respect to restriction to open subsets of $\mathbb{R}^n$, and, further, μ is friendly if and only if μ-almost every point has a neighborhood U such that $\mu|_U$ is friendly.

We now turn to establishing that certain families of measures are friendly. The first such family of nontrivial examples comes from measures on smooth submanifolds of $\mathbb{R}^n$ which are 'curved enough' to deviate from any affine hyperplane. Such manifolds were called *nondegenerate* in [KM]. More precisely, let U be an open subset of $\mathbb{R}^d$, and let $\mathbf{f} = (f_1, \dots, f_n) : U \to \mathbb{R}^n$ be a C^ℓ map, $\ell \in \mathbb{N}$. Say that $\mathbf{f}$ is *ℓ-nondegenerate at $\mathbf{x} \in U$* if the space $\mathbb{R}^n$ is spanned by partial derivatives of $\mathbf{f}$ at $\mathbf{x}$ of order up to ℓ. If $\mathcal{M} \subset \mathbb{R}^n$ is a d-dimensional C^ℓ submanifold, we will say that $\mathcal{M}$ is *ℓ-nondegenerate 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 ℓ-nondegenerate at $\mathbf{f}^{-1}(\mathbf{y})$. We will say that $\mathbf{f}$ or $\mathcal{M}$ are *nondegenerate* if they are ℓ-nondegenerate for some ℓ.

Let us denote Lebesgue measure on $\mathbb{R}^d$ by λ. One has the following

Theorem 2.1. *Let U be an open subset of $\mathbb{R}^d$, and $\mathbf{f}$ a map from U to $\mathbb{R}^n$.*

(a) *Suppose that $\mathbf{f}$ is C^ℓ, nonsingular and ℓ-nondegenerate at λ-almost every point of U; then $\mathbf{f}_*\lambda$ is friendly.*
(b) *Suppose that μ is an absolutely decaying Federer measure on U, and $\mathbf{f}$ is a $C^{\ell+1}$ map which is nonsingular and ℓ-nondegenerate at μ-almost every point; then $\mathbf{f}_*\mu$ is friendly.*

Note that the main result of [KM] is the strong extremality of $\mathbf{f}_*\lambda$ for $\mathbf{f}$ as in (a) above; in view of Theorem 2.1, it is included in Theorem 1.1 as a special case.

Note also that Theorem 2.1(b) highlights the difference between decay and absolute decay: one cannot weaken the hypotheses and replace 'absolutely decaying' with 'decaying'. Indeed, it is not hard to see that volume measures on proper C^1 submanifolds of $\mathbb{R}^n$ are never absolutely decaying, and when $n \geq 2$ one can easily construct an everywhere nondegenerate map from $\mathbb{R}^n$ to $\mathbb{R}^n$ which sends a proper submanifold into a proper affine subspace. However, one can show that any decaying non-atomic ($\Leftrightarrow$ nonplanar) measure on $\mathbb{R}$ is absolutely decaying.

On the other hand, the class of friendly (and even absolutely decaying Federer) measures turns out to be big enough to include measures supported on fractal subsets of $\mathbb{R}^n$. An important family of examples is given by measures supported on self-similar sets satisfying the so-called open set condition, which we now define.

One says that $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^n$ is a *similarity map* if it can be written as

$$\mathbf{h}(\mathbf{y}) = \varrho\Theta(\mathbf{y}) + \mathbf{a}\,, \tag{2.6}$$

where $\varrho \in \mathbb{R}_+$, $\Theta \in \mathrm{O}(n, \mathbb{R})$ and $\mathbf{a} \in \mathbb{R}^n$. It is said to be *contracting* if $\varrho < 1$. It is known, see [H], that for any finite family $\mathbf{h}_1, \dots, \mathbf{h}_m$ of contracting similarity maps there exists a unique set K, called the *attractor* of the family, such that

$$K = \bigcup_{i=1}^{m} \mathbf{h}_i(K). \tag{2.7}$$

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

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

and

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

J. Hutchinson [H] proved that if $\mathbf{h}_i = \varrho_i \Theta_i + \mathbf{a}_i$, $i = 1, \dots, m$, satisfy the open set condition, and if $s > 0$ is the unique solution of

$$\sum_i \varrho_i^s = 1\,, \tag{2.8}$$

which we will call the *similarity dimension* of the family $\{\mathbf{h}_i\}$, then the s-dimensional Hausdorff measure H^s of K is positive and finite. Measures obtained via the above construction have been thoroughly studied during recent decades; perhaps the simplest example is given by the $\frac{\log 2}{\log 3}$-dimensional Hausdorff measure on Cantor's ternary set.

We would like to prove these measures to be friendly. However a natural obstruction arises if there exists a finite collection of proper affine subspaces $\mathcal{L}_1, \dots, \mathcal{L}_k$ which is invariant under the family $\{\mathbf{h}_i\}$. Let us say that a family of maps is *irreducible* if this does not happen. The following is true:

Theorem 2.2. *Let $\{\mathbf{h}_1, \dots, \mathbf{h}_m\}$ be an irreducible family of contracting similarity self-maps of $\mathbb{R}^n$ satisfying the open set condition, s its similarity dimension, μ the restriction of H^s to its attractor. Then μ is absolutely decaying and Federer, and, in particular, friendly.*

Combining the two theorems above, one can already construct a variety of friendly measures. Another source of examples comes from product measures:

Theorem 2.3. *For $i = 1, \dots, k$, let μ_i be a measure on $\mathbb{R}^{n_i}$ and let $\mu = \mu_1 \times \cdots \times \mu_k$ on $\mathbb{R}^n$, $n = \sum n_i$. Then:*

(1) *If each μ_i is absolutely decaying and Federer, then so is μ.*
(2) *If each μ_i is friendly, then μ is friendly.*

The proofs of all the theorems from this section are contained in §§7–9.

Remark. Note that, unlike strong extremality, the Federer and decay conditions are not measure class invariant, see §6 for more detail. Moreover, these conditions can be relaxed somewhat without sacrificing the validity of Theorem 1.1. For clarity of exposition we first present the proof that friendly measures are strongly extremal. We then define non-uniform versions of the conditions for a measure to be friendly, and prove that measures satisfying these non-uniform conditions are also strongly extremal. We do not know whether the latter conditions are measure class invariant, but, in view of a result of H. Sato [Sa] regarding a non-uniform version of the Federer condition, we suspect that they are.

3. Diophantine approximation and flows on homogeneous spaces

In this section we recall the connection between diophantine approximation and flows on homogeneous spaces, and reduce Theorem 1.1 to a quantitative nondivergence result.

Let $G = \mathrm{SL}_{n+1}(\mathbb{R})$, $\Gamma = \mathrm{SL}_{n+1}(\mathbb{Z})$, and denote by $\pi : G \to G/\Gamma$, $g \mapsto g\Gamma$, the natural projection map. G acts on G/Γ by left translations via the rule $g\pi(h) = \pi(gh)$, $g, h \in G$. Equivalently one can describe G/Γ as the space of unimodular lattices in $\mathbb{R}^{n+1}$, with $\pi(g) = g\mathbb{Z}^{n+1}$ and the action of G on G/Γ coming from the linear action of G on $\mathbb{R}^{n+1}$. We will be interested in the action of a certain subsemigroup of G on G/Γ. Namely, we let $\mathbb{R}^n_+$ (resp. $\mathbb{Z}^n_+$) denote the (integer) vectors all of whose entries are non-negative, and for $\mathbf{t} \in \mathbb{R}^n_+$ let

$$g_{\mathbf{t}} = \mathrm{diag}(e^{t_1}, \dots, e^{t_n}, e^{-t}), \quad \mathbf{t} = (t_1, \dots, t_n), \quad t = \sum_{i=1}^{n} t_i \,. \tag{3.1}$$

Then $\{g_{\mathbf{t}}\}$ acts on G/Γ by the rule $g_{\mathbf{t}}\pi(g) = \pi(g_{\mathbf{t}}g)$. Thus the action of $g_{\mathbf{t}}$ on a lattice Λ will contract the last component of every vector of Λ and expand the remaining components. Here and throughout, $\|\cdot\|$ denotes the Euclidean norm and $\|\cdot\|_\infty$ the maximum norm.

For $\varepsilon > 0$ let

$$K_\varepsilon \stackrel{\text{def}}{=} \pi\big(\{g \in G : \|g\mathbf{v}\| \ge \varepsilon \quad \forall\, \mathbf{v} \in \mathbb{Z}^{n+1} \smallsetminus \{0\}\}\big), \tag{3.2}$$

i.e., K_ε is the collection of all unimodular lattices in $\mathbb{R}^{n+1}$ which contain no nonzero vector smaller than ε. Recall that G/Γ is noncompact and has finite G-invariant measure. Each K_ε, however, is compact, and $\{K_\varepsilon\}_{\varepsilon>0}$ is an exhaustion of G/Γ [R, Chapter 10].

We define the following maps from $\mathbb{R}^n$ to G and G/Γ:

$$\tau(\mathbf{y}) \stackrel{\text{def}}{=} \begin{pmatrix} I_n & \mathbf{y} \\ 0 & 1 \end{pmatrix}, \qquad \bar\tau \stackrel{\text{def}}{=} \pi \circ \tau \tag{3.3}$$

(here I_n stands for the $n \times n$ identity matrix). The following proposition relates the orbit of $\bar\tau(\mathbf{y})$ under the semigroup (3.1) with the diophantine properties of $\mathbf{y}$. For an equivalent but slightly different approach, compare with [KM, Corollary 2.2], [K2, Corollary 5.2].

Proposition 3.1. *For $\mathbf{y} \in \mathbb{R}^n$, the following are equivalent:*

(i) $\mathbf{y}$ *is not very well multiplicatively approximable;*

(ii) *for any $\gamma > 0$ one has*

$$g_{\mathbf{t}}\bar\tau(\mathbf{y}) \in K_{e^{-\gamma t}} \tag{3.4}$$

whenever $\mathbf{t} \in \mathbb{R}^n_+$ is large enough;

(iii) *for any $\gamma > 0$ (3.4) holds for all but finitely many $\mathbf{t} \in \mathbb{Z}^n_+$.*

We first prove a lemma which shows that very well multiplicatively approximable points automatically satisfy a seemingly more stringent condition.

Lemma 3.2. *Let $\mathbf{y} \in \mathbb{R}^n$ be very well multiplicatively approximable. Then there exists $\delta > 0$ for which there are infinitely many solutions $\mathbf{p} \in \mathbb{Z}^n$, $q \in \mathbb{Z}_+$*

to (1.2) *in addition satisfying*

$$\|q\mathbf{y} - \mathbf{p}\|_\infty < q^{-\delta/n}\,. \tag{3.5}$$

Proof. Choose $\delta_0 > 0$ for which there are infinitely many solutions $\mathbf{p} \in \mathbb{Z}^n$, $q \in \mathbb{Z}_+$ to the inequality

$$\prod_{i=1}^n |qy_i - p_i| < q^{-(1+\delta_0)}\,, \tag{3.6}$$

and set $\delta \stackrel{\text{def}}{=} \dfrac{\delta_0}{n+2+\delta_0}$.

Let $\mathbf{p}$, q be a solution to (3.6), and let $q_1 \stackrel{\text{def}}{=} [q^{\frac{\delta_0}{n+2}}]$. For every $k \in \{1, \dots, q_1+1\}$ set

$$\mathbf{v}_k \stackrel{\text{def}}{=} kq\,\mathbf{y} \bmod 1$$

(where we take the fractional part in each coordinate). Since $\{\mathbf{v}_1, \dots, \mathbf{v}_{q_1+1}\}$ are q_1+1 points in the unit cube $[0,1)^n$, there must be two points, say $\mathbf{v}_k$, $\mathbf{v}_\ell$ with $1 \le k < \ell \le q_1+1$, such that

$$\|\mathbf{v}_k - \mathbf{v}_\ell\|_\infty \le q_1^{-\frac{1}{n}} \le q^{-\frac{\delta_0}{n(n+2)}}. \tag{3.7}$$

We set $\bar q \stackrel{\text{def}}{=} (\ell - k)q$ and let $\bar{\mathbf{p}} \in \mathbb{Z}^n$ be the integer vector closest to $\bar q\mathbf{y}$. Note that $\bar q \le q^{\frac{n+2+\delta_0}{n+2}}$. Then by (3.7),

$$\|\bar q\mathbf{y} - \bar{\mathbf{p}}\|_\infty \le q^{-\frac{\delta_0}{n(n+2)}} \le \bar q^{-\frac{\delta_0}{n(n+2+\delta_0)}} = \bar q^{-\frac{\delta}{n}}.$$

Furthermore,

$$\begin{aligned}\prod_{i=1}^n |\bar q y_i - \bar{\mathbf{p}}_i| &\le (\ell-k)^n \prod_{i=1}^n |qy_i - p_i| \\ &\le q^{\frac{n\delta_0}{n+2}}\, q^{-(1+\delta_0)} \\ &\le \bar q^{-\frac{n+2+2\delta_0}{n+2+\delta_0}} = \bar q^{-(1+\delta)}.\end{aligned} \tag{3.8}$$

Thus $\bar q$, $\bar{\mathbf{p}}$ is a solution to both (1.2) and (3.5). $\square$

Proof of Proposition 3.1. Let us show that (ii) implies (i). Indeed, suppose by contradiction that for some $\delta > 0$ we have infinitely many solutions to both inequalities (1.2) and (3.5). Fix arbitrary positive $\gamma < \frac{\delta}{n+\delta}$, and let $\mathbf{p}$, q be one of the solutions. Our goal is to find $\mathbf{t} = \mathbf{t}(q)$ so that the norm of the vector

$$\mathbf{v} \stackrel{\text{def}}{=} g_{\mathbf{t}}\tau(\mathbf{y})\begin{pmatrix}-\mathbf{p}\\ q\end{pmatrix} = \begin{pmatrix} e^{t_1}(qy_1 - p_1)\\ \vdots \\ e^{t_n}(qy_n - p_n)\\ e^{-t}q\end{pmatrix} \in g_{\mathbf{t}}\tau(\mathbf{y})\,\mathbb{Z}^{n+1} \smallsetminus \{0\} \tag{3.9}$$

is less than $e^{-\gamma t}$, and so that $\mathbf{t} \to \infty$ as $q \to \infty$.

We consider two cases. If at least one coordinate of $q\mathbf{y} - \mathbf{p}$, say the last one, is equal to zero, one can take $\delta' < \delta$ with $\gamma = \frac{\delta'}{n+\delta'}$, and define $\mathbf{t}$ by

$$e^{t_1} = \dots = e^{t_{n-1}} = q^{\frac{\delta-\delta'}{n}}, \quad e^{t_n} = q^{1-\frac{n-1}{n}\delta+\delta'}\,.$$

Then $e^t = q^{1+\delta'/n}$, therefore

$$\|\mathbf{v}\|_\infty = \max(e^{t_1}\|q\mathbf{y}-\mathbf{p}\|_\infty, e^{-t}q) \overset{(3.5)}{\leq} q^{-\delta'/n} = e^{-\gamma t}.$$

Otherwise, take $1 < s < 1+\delta$, set

$$Q \overset{\text{def}}{=} \left(\prod_{i=1}^n |qy_i - p_i|\right)^{-1} \quad \text{and} \quad \alpha \overset{\text{def}}{=} \frac{\log Q}{s\log q},$$

and then define $\mathbf{t}$ by

$$e^{t_i} = |qy_i - p_i|^{-1/\alpha}, \qquad i = 1, \dots, n.$$

Note that $\alpha > \frac{1+\delta}{s}$ in view of (1.2), or, equivalently, $1 - 1/\alpha > \frac{1+\delta-s}{1+\delta}$. Also

$$e^t = Q^{1/\alpha} = q^s.$$

Therefore one has

$$e^{-t}q = q^{1-s} = e^{-\frac{s-1}{s}t}$$

and

$$e^{t_i}|qy_i - p_i| = |qy_i - p_i|^{1-1/\alpha} \overset{(3.5)}{\leq} q^{-\frac{\delta}{n}(1-1/\alpha)} = e^{-\frac{\delta}{n}\frac{1-1/\alpha}{s}t} < e^{-\frac{\delta}{n}\frac{1+\delta-s}{s(1+\delta)}t}.$$

Choosing the optimal s one gets

$$\|\mathbf{v}\|_\infty \leq e^{-\frac{\delta}{n+\delta}t} \leq e^{-\gamma t},$$

thus finishing the proof of (ii) $\Longrightarrow$ (i).

The proof of the opposite direction is similar; since we will not use it, we leave the details to the reader. Finally, the equivalence of (ii) and (iii) follows in a straightforward manner from the continuity of the G-action on G/Γ. $\square$

We now complete our description of the strategy for proving Theorem 1.1. We will establish the following:

Theorem 3.3. *Suppose that μ is a friendly measure on $\mathbb{R}^n$. Then for μ-almost every $\mathbf{y}_0 \in \mathbb{R}^n$ there is a ball B centered at $\mathbf{y}_0$ and $\widetilde{C}, \alpha > 0$ such that for any $\mathbf{t} \in \mathbb{Z}_+^n$ and any $\varepsilon > 0$,*

$$\mu(\{\mathbf{y} \in B : g_{\mathbf{t}}\bar{\tau}(\mathbf{y}) \notin K_\varepsilon\}) \leq \widetilde{C}\varepsilon^\alpha.$$

This theorem will be deduced from the more general Theorem 4.3 in the next section. Informally speaking, Theorem 3.3 shows that for fixed $\mathbf{t}$, the 'orbit' $\{g_{\mathbf{t}}\bar{\tau}(\mathbf{y}) : \mathbf{y} \in B\}$ does not diverge, that is, a very significant proportion of it (computed in terms of μ and uniform in $\mathbf{t}$) stays inside compact sets K_ε. Theorem 3.3 generalizes [KM, Proposition 2.3], where a similar estimate is proved for volume measures on nondegenerate submanifolds of $\mathbb{R}^n$.

Proof of Theorem 1.1 assuming Theorem 3.3. The Borel-Cantelli Lemma states that if μ is a measure on a space X and $\{A_i\}$ is a countable collection of measurable subsets of X with $\sum_i \mu(A_i) < \infty$, then μ-almost every $x \in X$ is contained in at most finitely many sets A_i. Assuming Theorem 3.3, we see that if μ is friendly, then for μ-almost every $\mathbf{y}_0 \in \mathbb{R}^n$ there is a ball B centered at $\mathbf{y}_0$ such that for any $\gamma > 0$,

$$\sum_{\mathbf{t}\in\mathbb{Z}_+^n} \mu\big(\{\mathbf{y} \in B : g_{\mathbf{t}}\bar{\tau}(\mathbf{y}) \notin K_{e^{-\gamma t}}\}\big) < \infty\,.$$

From Proposition 3.1 we can conclude that μ-a.e. $\mathbf{y} \in B$ is not very well multiplicatively approximable, and since a countable collection of such balls covers μ-a.e. point, it follows that μ is strongly extremal. $\square$

4. Quantitative nondivergence

In order to state the main result of this section we need to introduce some notation and definitions.

Say that a metric space X is *Besicovitch* if there exists $N > 0$ such that for any bounded set $A \subset X$ and any collection of balls $\mathcal{B}$ such that any $x \in A$ is the center of a ball in $\mathcal{B}$, there is a countable subcollection $\Omega \subset \mathcal{B}$ such that

$$A \subset \bigcup_{B\in\Omega} B$$

and

$$\max_{x\in X} \#\{B \in \Omega : x \in B\} \le N.$$

It is well known, see e.g. [Mat], that $\mathbb{R}^n$ equipped with the Euclidean metric, and hence any of its subsets, is Besicovitch.

Given C, $\alpha > 0$, a metric space X, a subset U of X, a measure μ on X, and a real-valued function f on X, say that f is *(C, α)-good on U with respect to μ* if for any open ball $B \subset U$ centered in $\operatorname{supp}\mu$ and any $\varepsilon > 0$ one has

$$\mu\big(\{x \in B : |f(x)| < \varepsilon\}\big) \le C\left(\frac{\varepsilon}{\|f\|_{\mu,B}}\right)^{\alpha} \mu(B)\,. \tag{4.1}$$

The class of (C, α)-good functions with respect to Lebesgue measure on $\mathbb{R}^n$ was introduced in [KM]. We refer the reader to [KM] and [BKM] for various properties and examples.

Lemma 4.1. *Suppose that $f_1, \dots, f_k$ are (C, α)-good on U with respect to μ. Then $(f_1^2 + \cdots + f_k^2)^{1/2}$ is $(k^{\alpha/2}C, \alpha)$-good on U with respect to μ.*

Proof. It is immediate from the definition that if $f_1, \dots, f_k$ are all (C, α)-good then so is $\max(|f_1|, \dots, |f_k|)$. Using this, we see that

$$\begin{aligned}\mu\Big(\Big\{x \in B : \Big(\sum_{i=1}^k f_i^2(x)\Big)^{1/2} < \varepsilon\Big\}\Big) &\le \mu\big(\{x \in B : \max_{1\le i\le k} |f_i(x)| < \varepsilon\}\big)\\ &\le C\left(\frac{\varepsilon}{\|\max_{1\le i\le k} |f_i|\,\|_{\mu,B}}\right)^{\alpha}\mu(B)\\ &\le Ck^{\alpha/2}\left(\frac{\varepsilon}{\|\left(\sum f_i^2\right)^{1/2}\|_{\mu,B}}\right)^{\alpha}\mu(B)\,.\end{aligned}$$

$\square$

One can immediately notice the similarity between the definition of (C,α)-good functions and that of (C,α)-decaying measures. Indeed, one easily verifies:

Lemma 4.2. *A measure μ on $\mathbb{R}^n$ is (C,α)-decaying on $U \subset \mathbb{R}^n$ if and only if all functions of the form $d_{\mathcal{L}}(x)$, where $\mathcal{L}$ is an affine hyperplane in $\mathbb{R}^n$, are (C,α)-good on U with respect to μ.*

The setup of Theorem 4.3 involves a Besicovitch metric space X and a family of functions assumed to be (C,α)-good on some ball in X with respect to some measure μ on X. More precisely, these functions are compositions of continuous maps from X to $G = \mathrm{SL}_{n+1}(\mathbb{R})$ with certain functions on G, which arise from the action of G on

$$\mathcal{W} \stackrel{\text{def}}{=} \text{ the set of nonzero rational subspaces of } \mathbb{R}^{n+1}.$$

Namely, fix a Euclidean structure on $\mathbb{R}^{n+1}$, and for $g \in G$ and $V \in \mathcal{W}$ define $\ell_V(g)$ to be the covolume of $gV \cap g\mathbb{Z}^{n+1}$ in gV. Equivalently, one can extend the Euclidean norm $\|\cdot\|$ from $\mathbb{R}^{n+1}$ to its exterior algebra, and set

$$\ell_V(g) \stackrel{\text{def}}{=} \|g(\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k)\|,$$

where $\{\mathbf{v}_1,\dots,\mathbf{v}_k\}$ is a generating set for $\mathbb{Z}^{n+1} \cap V$; note that the above quantity does not depend on the choice of $\{\mathbf{v}_i\}$.

Here is the main result required for proving Theorem 1.1.

Theorem 4.3. *Given an open subset U of a Besicovitch metric space X, positive constants C, D, α, and a measure μ which is D-Federer on U, there exists $C' > 0$ with the following property. Suppose h is a continuous map $U \to G$, $0 < \varrho \le 1$, $z \in U \cap \operatorname{supp}\mu$, and $B = B(z,r)$ is a ball such that $\widetilde{B} \stackrel{\text{def}}{=} B(z, 3^n r)$ is contained in U, and that for each $V \in \mathcal{W}$*

(a) *the function $\ell_V \circ h$ is (C,α)-good on $\widetilde{B}$ with respect to μ,*

and

(b) $\|\ell_V \circ h\|_{\mu,B} \ge \varrho$.

Then for any $0 < \varepsilon \le \varrho$,

$$\mu\big(\{x \in B : \pi\big(h(x)\big) \notin K_\varepsilon\}\big) \le C' \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \mu(B).$$

Theorem 4.3 is proved in §5 by modifying arguments from [KM, §4–5]. In the remainder of this section we will use it to derive Theorem 3.3.

Proof of Theorem 3.3 assuming Theorem 4.3. The strategy is to take $X = \mathbb{R}^n$ and $h(\mathbf{y}) = g_{\mathbf{t}}\tau(\mathbf{y})$, with $g_{\mathbf{t}}$ and τ as defined in §3, and then relate the decay property of μ to condition (a) of Theorem 4.3, and the nonplanar property of μ to condition (b).

Since μ is decaying and Federer, for almost every $\mathbf{y}_0 \in \mathbb{R}^n$ one can choose a ball $B = B(\mathbf{y}_0, r)$ and $C_0, D, \alpha > 0$ such that μ is (C_0,α)-decaying and D-Federer on the dilated ball $\widetilde{B} = B(\mathbf{y}_0, 3^n r)$. An easy compactness argument using the assumption that μ is nonplanar shows the existence of $0 < \varrho < 1$ such that

$$\|d_{\mathcal{L}}\|_{\mu,B} \geq \varrho \qquad \text{for any affine hyperplane } \mathcal{L} \subset \mathbb{R}^n. \tag{4.2}$$

Suppose that V is a k-dimensional subspace of $\mathbb{R}^{n+1}$ spanned by the integer vectors

$$\mathbf{v}_1 = \begin{pmatrix} -\mathbf{p}_1 \\ q_1 \end{pmatrix}, \ldots, \mathbf{v}_k = \begin{pmatrix} -\mathbf{p}_k \\ q_k \end{pmatrix}.$$

By performing Gaussian elimination over the integers, there is no loss of generality in assuming that

$$q_i = 0, \quad i = 2, \ldots, k.$$

We need to show that $\ell_V \circ h(\mathbf{y})$ satisfies conditions (a) and (b) of Theorem 4.3.

Denote by V_0 the subspace

$$V_0 \stackrel{\text{def}}{=} \{(x_1, \ldots, x_{n+1}) : x_{n+1} = 0\}$$

of $\mathbb{R}^{n+1}$. We separate into two cases: suppose first that q_1 is also 0, that is, $V \subset V_0$. Since every element of V_0 is $\tau(\mathbb{R}^n)$-invariant, we have

$$h(\mathbf{y})\,[\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k] = g_{\mathbf{t}}\,[\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k]\,,$$

so $\mathbf{y} \mapsto \ell_V \circ h(\mathbf{y})$ is a constant function, obviously satisfying (a). Also, since $g_{\mathbf{t}}$ expands V_0, we can write

$$\ell_V \circ h(\mathbf{y}) = \|g_{\mathbf{t}}\,[\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k]\| \geq \|\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k\| \geq 1\,,$$

hence (b) is satisfied as well.

Now suppose $q_1 \neq 0$. Set

$$\tilde{g}_{\mathbf{t}} \stackrel{\text{def}}{=} \begin{pmatrix} e^{t_1} & \ldots & 0 \\ \vdots & \ddots & 0 \\ 0 & \ldots & e^{t_n} \end{pmatrix} \qquad \text{and} \qquad \mathbf{w} \stackrel{\text{def}}{=} g_{\mathbf{t}}\mathbf{v}_2 \wedge \cdots \wedge g_{\mathbf{t}}\mathbf{v}_k\,,$$

and let $\mathcal{P}$ denote the linear subspace of $\mathbb{R}^n$ corresponding to $\tilde{g}_{\mathbf{t}}\mathbf{p}_2 \wedge \cdots \wedge \tilde{g}_{\mathbf{t}}\mathbf{p}_k$. Then

$$\begin{aligned} h(\mathbf{y})\,[\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k] &= [g_{\mathbf{t}}\tau(\mathbf{y})\mathbf{v}_1 \wedge \cdots \wedge g_{\mathbf{t}}\tau(\mathbf{y})\mathbf{v}_k] \\ &= [g_{\mathbf{t}}\tau(\mathbf{y})\mathbf{v}_1 \wedge g_{\mathbf{t}}\mathbf{v}_2 \wedge \cdots \wedge g_{\mathbf{t}}\mathbf{v}_k] \\ &= \left(g_{\mathbf{t}}\begin{pmatrix} q_1\mathbf{y} - \mathbf{p}_1 \\ q_1 \end{pmatrix}\right) \wedge \mathbf{w} \\ &= \begin{pmatrix} \tilde{g}_{\mathbf{t}}(q_1\mathbf{y} - \mathbf{p}_1) \\ e^{-t}q_1 \end{pmatrix} \wedge \mathbf{w} \\ &= \begin{pmatrix} 0 \\ e^{-t}q_1 \end{pmatrix} \wedge \mathbf{w} + \begin{pmatrix} \tilde{g}_{\mathbf{t}}(q_1\mathbf{y} - \mathbf{p}_1) \\ 0 \end{pmatrix} \wedge \mathbf{w}. \end{aligned} \tag{4.3}$$

Since $\mathbf{w}$ is orthogonal to $\begin{pmatrix} 0 \\ e^{-t}q_1 \end{pmatrix}$, we have

$$\left\|\begin{pmatrix} 0 \\ e^{-t}q_1 \end{pmatrix} \wedge \mathbf{w}\right\|^2 = \|\mathbf{w}\|^2\, q_1^2 e^{-2t},$$

and furthermore the two k-forms in (4.3) are orthogonal.

We are left with evaluating the norm of the second expression in (4.3). Since all vectors involved are in V_0 this is simply the volume of the parallelepiped spanned by $\tilde{g}_{\mathbf{t}}(q_1\mathbf{y} - \mathbf{p}_1)$, $\tilde{g}_{\mathbf{t}}\mathbf{p}_2$, $\dots$, $\tilde{g}_{\mathbf{t}}\mathbf{p}_k$, hence

$$\left\| \begin{pmatrix} \tilde{g}_{\mathbf{t}}(q_1\mathbf{y} - \mathbf{p}_1) \\ 0 \end{pmatrix} \wedge \mathbf{w} \right\|^2 = d_{\mathcal{P}}\left(\tilde{g}_{\mathbf{t}}(q_1\mathbf{y} - \mathbf{p}_1)\right)^2 \|\mathbf{w}\|^2.$$

Thus we obtain:

$$\begin{aligned}(\ell_V \circ h(\mathbf{y}))^2 &= \|h(\mathbf{y})\,[\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k]\|^2 \\ &= \|\mathbf{w}\|^2 \left[d_{\mathcal{P}}\big(\tilde{g}_{\mathbf{t}}(q_1\mathbf{y} - \mathbf{p}_1)\big)^2 + q_1^2 e^{-2t} \right].\end{aligned} \tag{4.4}$$

It follows from Lemma 4.1, in conjunction with Lemma 4.2 and the (C_0, α)-decaying property of μ, that $\ell_V \circ h$ is (C, α)-good on $\widetilde{B}$ with respect to μ, with $C = (n+1)^{\alpha/2} C_0$, establishing (a).

As in the previous case, since for $i > 1$ we know that $\mathbf{v}_i$ is an integer vector in V_0, and $g_{\mathbf{t}}$ expands V_0, we have

$$\|\mathbf{w}\| \geq 1\,.$$

Thus (4.4) implies that for every $\mathbf{y}$,

$$\ell_V \circ h(\mathbf{y}) > d_{\mathcal{P}}\big(\tilde{g}_{\mathbf{t}}(q_1\mathbf{y} - \mathbf{p}_1)\big)\,.$$

Since q_1 is a nonzero integer and $\tilde{g}_{\mathbf{t}}$ expands $\mathbb{R}^n$, there exists a $(k-1)$-dimensional affine subspace $\mathcal{P}'$ such that $\ell_V \circ h(\mathbf{y}) > d_{\mathcal{P}'}(\mathbf{y})$. By (4.2) this implies that $\|\ell_V \circ h(\mathbf{y})\|_{\mu,B} > \varrho$.

Using Theorem 4.3 with $U = \widetilde{B}$, and enlarging C' if necessary to account for the case $\varepsilon \geq \varrho$, we obtain the conclusion of Theorem 3.3. $\square$

5. Proof of Theorem 4.3

Throughout this section, we retain the hypotheses and notations of Theorem 4.3, and let N denote the constant involved in the definition of the Besicovitch condition. *With no loss of generality we will replace X with* $\operatorname{supp}\mu$, *inducing the metric from X; this will allow us to use $\|\cdot\|_B$ in place of $\|\cdot\|_{\mu,B}$ throughout the proof.*

We need some additional terminology.

By a *flag* we mean a chain (not necessarily maximal)

$$\mathcal{F} = \big(\{0\} = V_0 \subsetneqq V_1 \subsetneqq \cdots \subsetneqq V_m \subsetneqq V_{m+1} = \mathbb{R}^{n+1}\big)\,,$$

where $V_i \in \mathcal{W}$ for $i = 1, \dots, m$.

We will call m the *length* of $\mathcal{F}$. The subspaces V_i, $i = 1, \dots, m$ will be called the subspaces *belonging to* $\mathcal{F}$. For $V \in \mathcal{W}$ we will say that V *can be added to* $\mathcal{F}$ if, for some $0 \leq i \leq m$, $V_i \subsetneqq V \subsetneqq V_{i+1}$. If this holds we will denote the flag obtained by adding V to $\mathcal{F}$ by $\mathcal{F} + V$.

Let $\mathcal{F}$ be as above, and let $\mathcal{F}'$ be another flag. We will say that $\mathcal{F}$ is *subordinate to* $\mathcal{F}'$ if every subspace belonging to $\mathcal{F}$ can be added to $\mathcal{F}'$. Note that if the length of $\mathcal{F}'$ is m, then the length of any flag subordinate to $\mathcal{F}'$ is at most $n - m$.

If $\mathcal{F}$ is a flag which is subordinate to $\mathcal{F}'$, we will say that a point $y \in X$ is *marked by* $(\mathcal{F}, \varepsilon, \varrho)$ *relative to* $\mathcal{F}'$ (or simply *marked by* $\mathcal{F}$ *relative to* $\mathcal{F}'$), if the following hold:

(M1) for any subspace V belonging to $\mathcal{F}$, $\ell_V\big(h(y)\big) \le \varrho$.
(M2) for any subspace V belonging to $\mathcal{F}$, $\ell_V\big(h(y)\big) \ge \varepsilon$.
(M3) the flag $\mathcal{F}$ has maximal length, relative to all flags which are subordinate to $\mathcal{F}'$ and satisfy (M1).

We will say that y is *marked relative to* $\mathcal{F}'$ if there is some $\mathcal{F}$ subordinate to $\mathcal{F}'$ such that y is marked by $\mathcal{F}$ relative to $\mathcal{F}'$.

Let
$$\mathcal{F}_0 = (\{0\} \subsetneqq \mathbb{R}^n)\,.$$
We say that y is *marked by* $(\mathcal{F}, \varepsilon, \varrho)$ (or simply *marked by* $\mathcal{F}$) if y is marked by $(\mathcal{F}, \varepsilon, \varrho)$ relative to $\mathcal{F}_0$.

The reason to mark points by flags is the following:

Proposition 5.1. *Let the notation be as above, and suppose that* $0 < \varepsilon < \varrho \le 1$*, and that* $y \in X$ *is marked by* $(\mathcal{F}, \varepsilon, \varrho)$*. Then* $\pi\big(h(y)\big) \in K_\varepsilon$.

Proof. Let $g = h(y)$ and suppose by contradiction that there is a nonzero $\mathbf{v} \in \mathbb{Z}^{n+1}$ such that $\|g\mathbf{v}\| < \varepsilon$. Choose i so that $\mathbf{v} \notin V_i$, $\mathbf{v} \in V_{i+1}$, and let V be the span of V_i and $\mathbf{v}$. Let $k = \dim V$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{Z}^{n+1}$ be linearly independent vectors such that $\mathbb{Z}^{n+1} \cap V_i$ is generated by $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ and $\mathbb{Z}^{n+1} \cap V$ is generated by $\mathbf{v}_1, \dots, \mathbf{v}_k$. Clearly
$$\ell_V(g) = \|g\mathbf{v}_1 \wedge \cdots \wedge g\mathbf{v}_k\| \le \|g\mathbf{v}_1 \wedge \cdots \wedge g\mathbf{v}_{k-1}\| \|g\mathbf{v}_k\|.$$
Then one has
$$\ell_V(g) \le \ell_{V_i}(g)\|g\mathbf{v}\| < \varepsilon.$$
In particular, $\ell_V(g) \le \varrho$, so by (M3), $V = V_{i+1}$. Thus $\ell_{V_{i+1}}(g) < \varepsilon$, in contradiction to (M2). □

The following lemma constructs flags by induction.

Lemma 5.2. *For* $d = 0, \dots, n$ *there is a constant* $C' = C'(D, N, C, d)$ *such that the following holds. Let* $\mathcal{F}$ *be a flag, and let* d *be the maximal length of a flag which is subordinate to* $\mathcal{F}$*. Let* $A = B(z_A, r_A)$ *be a ball such that* $\widetilde{A} \overset{\text{def}}{=} B(z_A, 3^d r_A) \subset U$*, and suppose that:*

(a′) *For each* $V \in \mathcal{W}$*, the function* $\ell_V \circ h$ *is* (C, α)*-good on* $\widetilde{A}$ *with respect to* μ.
(b′) *If* V *can be added to* $\mathcal{F}$ *then* $\|\ell_V \circ h\|_A \ge \varrho$.

Let A_0 *be the set of all* $y \in A$ *which are marked relative to* $\mathcal{F}$*.*

Then for every $\varepsilon < \varrho$,
$$\mu(A \smallsetminus A_0) \le C' \left(\frac{\varepsilon}{\varrho}\right)^\alpha \mu(A)\,. \tag{5.1}$$

Proof. For the proof we use the notation $Y \ll Z$ to mean that there is a constant $C' = C'(D, N, C, d, n)$ such that $Y \le C'Z$ (C' may change throughout the proof).

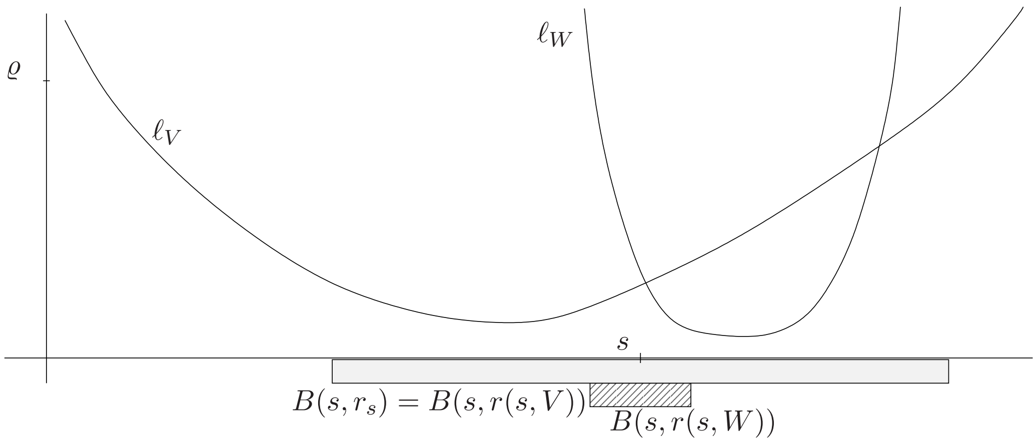


FIGURE 2. Choosing V_0

The proof is by induction on d (equivalently, by a decreasing induction on the length of $\mathcal{F}$). If $d = 0$, then (b′) implies that (M1), (M2), (M3) hold trivially for all $x \in A$, that is $A = A_0$ and (5.1) holds.

Suppose $d \geq 1$. For each $x \in A$, let $\mathcal{W}(x)$ consist of all $V \in \mathcal{W}$ which can be added to $\mathcal{F}$, and for which $\ell_V\big(h(x)\big) < \varrho$. Let

$$E = \{x \in A : \mathcal{W}(x) \neq \varnothing\}.$$

If $x \notin E$ then $\mathcal{W}(x) = \varnothing$ and hence x is marked by $\mathcal{F}_0$ relative to $\mathcal{F}$; that is

$$A \smallsetminus A_0 = E \smallsetminus A_0. \tag{5.2}$$

Now for each $x \in E$ and each $V \in \mathcal{W}(x)$ let

$$r(x, V) \stackrel{\text{def}}{=} \sup\{r : \|\ell_V \circ h\|_{B(x,r)} < \varrho\}.$$

By (b′), $A \not\subset B\big(x, r(x, V)\big)$ and hence the supremum in the above definition is finite. By continuity of $\ell_V \circ h$, $r(x, V) > 0$.

By discreteness, $\mathcal{W}(x)$ is finite. Take $V_0 = V_0(x) \in \mathcal{W}(x)$ such that

$$r_x \stackrel{\text{def}}{=} r(x, V_0) = \max_{V \in \mathcal{W}(x)} r(x, V),$$

and let

$$\mathcal{F}'(x) = \mathcal{F} + V_0.$$

We have a cover of E by the collection of balls $\{B(x, r_x) : x \in E\}$. Using the Besicovitch assumption we take a countable subset $\Omega \subset E$ such that $\{B(x, r_x) : x \in \Omega\}$ is a cover of E and $\max_{y \in X} \#\{x \in \Omega : y \in B(x, r_x)\} \ll 1$, therefore

$$\sum_{x \in \Omega} \mu\big(B(x, r_x)\big) \ll \mu\Big(\bigcup_{x \in \Omega} B(x, r_x)\Big). \tag{5.3}$$

By (b′), $B(x, r)$ does not contain A for $r < r_x$, and hence

$$r_x \leq r_A + \operatorname{dist}(z_A, x) < 2r_A. \tag{5.4}$$

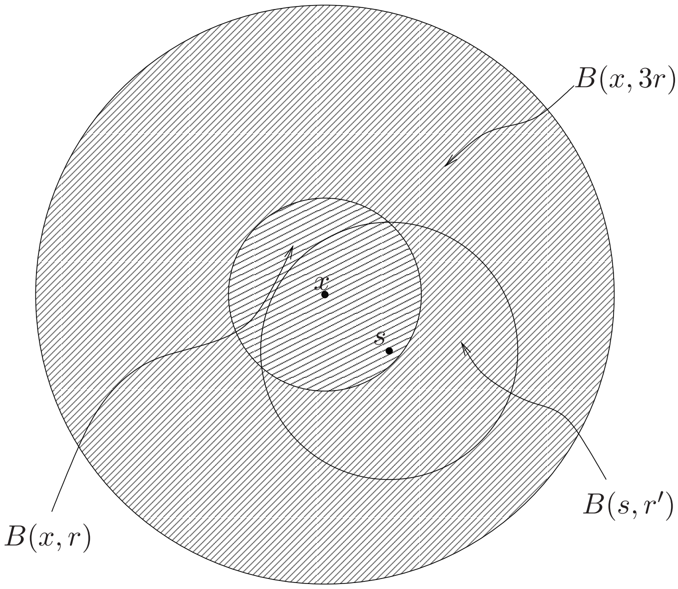


FIGURE 3. New balls cover old balls and aren't much larger

Thus there exist $r_x < r'_x < 2r_A$ such that

$$B(x,r'_x) \subset B(z_A,3r_A)\,.$$

In particular, using Federer,

$$\mu\big(\bigcup_{x\in\Omega} B(x,r_x)\big) \le \mu\big(B(z_A,3r_A)\big) \ll \mu(A). \tag{5.5}$$

It follows from (5.4) that

$$B(x,3^{d-1}r'_x) \subset \widetilde{A}. \tag{5.6}$$

By definition of r_x, $\|\ell_{V_0}\circ h\|_{B(x,r'_x)} \ge \varrho$. By assumption (a′) and Federer we obtain:

(5.7)
$$\begin{aligned}\mu(\{y\in B(x,r_x) : \ell_{V0(x)}\big(h(y)\big) < \varepsilon\}) &\le \mu(\{y\in B(x,r'_x) : \ell_{V0(x)}\big(h(y)\big) < \varepsilon\}) \\ &\ll \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \mu\big(B(x,r'_x)\big) \\ &\ll \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \mu\big(B(x,r_x)\big)\,.\end{aligned}$$

Let

$$Y = \bigcup_{x\in\Omega}\{y\in B(x,r_x) : \ell_{V0(x)}(y) < \varepsilon\}.$$

It follows from (5.3), (5.5) and (5.7) that:

$$\mu(Y) \ll \left(\frac{\varepsilon}{\varrho}\right)^{\alpha}\mu(A). \tag{5.8}$$

Write $V = V_0(x)$, and denote by $E_0(x)$ the set of points of E which are marked relative to $\mathcal{F}'(x)$. If $y \in B(x, r_x) \smallsetminus Y$, then $\varepsilon \le \ell_V \circ h(y) \le \varrho$; thus if y is marked by $\mathcal{F}''$ relative to $\mathcal{F}'(x)$, it is also marked by $\mathcal{F}'' + V$ relative to $\mathcal{F}$. Hence

$$E_0(x) \cap B(x, r_x) \subset A_0 \cup Y.$$

Let $y \in E \smallsetminus (A_0 \cup Y)$, then there is some $x \in \Omega$ such that $y \in B(x, r_x)$, and by the above, $y \notin E_0(x)$. That is,

$$E \smallsetminus (A_0 \cup Y) \subset \bigcup_{x \in \Omega} \big(B(x, r_x) \smallsetminus E_0(x)\big). \tag{5.9}$$

The maximal length of a flag subordinate to $\mathcal{F}'(x)$ is at most $d-1$, and by the induction hypothesis, the lemma is valid for $d-1$ in place of d. We apply the lemma to $B(x, r'_x)$ and $\mathcal{F}'(x)$. Assumption (a$'$) follows from (5.6) and assumption (b$'$) follows from the definitions of r_x and r'_x. We obtain (using Federer again):

$$\begin{aligned} \mu\big(B(x, r_x) \smallsetminus E_0(x)\big) &\le \mu\big(B(x, r'_x) \smallsetminus E_0(x)\big) \\ &\ll \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \mu\big(B(x, r'_x)\big) \\ &\ll \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \mu\big(B(x, r_x)\big). \end{aligned} \tag{5.10}$$

Therefore one has:

$$\begin{aligned} \mu(A \smallsetminus A_0) &\overset{(5.2)}{=} \mu(E \smallsetminus A_0) \le \mu(Y) + \mu\big(E \smallsetminus (Y \cup A_0)\big) \\ &\overset{(5.8)\text{ and }(5.9)}{\ll} \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \mu(A) + \mu\left(\bigcup_{x \in \Omega} \big(B(x, r_x) \smallsetminus E_0(x)\big)\right) \\ &\overset{(5.10)}{\ll} \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \left[\mu(A) + \sum_{x \in \Omega} \mu\big(B(x, r_x)\big)\right] \\ &\overset{(5.3)\text{ and }(5.5)}{\ll} \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \mu(A)\,, \end{aligned}$$

which finishes the proof of the lemma. □

To complete the proof of Theorem 4.3, apply Lemma 5.2 to $A = B$ and to $\mathcal{F} = \mathcal{F}_0$. Then apply Proposition 5.1. □

6. NON-UNIFORMLY FRIENDLY MEASURES

In this section we define non-uniform versions of the Federer and decaying conditions, and prove that every nonplanar, non-uniformly Federer and non-uniformly decaying measure is strongly extremal.

Let D, C, α, s be positive constants. Say that a measure μ on a metric space X is (D, s)-*Federer at* $x \in X$ if (2.2) holds whenever $3r < s$. Similarly, say that a measure μ on $\mathbb{R}^n$ is (C, α, s)-*decaying at* $\mathbf{y} \in \mathbb{R}^n$ if (2.4) holds for any $\varepsilon > 0$, any affine hyperplane $\mathcal{L}$, and any $B = B(\mathbf{y}, r)$ with $r < s$. We will say that μ is *Federer* (resp., *decaying*) *at a point* if it is (D,s)-Federer for some D,s (resp., (C,α,s)-decaying for some C,α,s) at this point, and that μ is *non-uniformly Federer* (resp., *non-uniformly decaying*) if it is Federer (resp., decaying) at μ-almost every point.

The following theorem strengthens Theorem 1.1:

Theorem 6.1. *If μ is a measure on $\mathbb{R}^n$ which is nonplanar, non-uniformly Federer and non-uniformly decaying, then μ is strongly extremal.*

We remark that the above non-uniform versions seem to be more natural than their uniform analogues: the non-uniform Federer condition was proved in [Sa, Example 5] to be measure class invariant, and it is plausible that the same holds for the non-uniform decaying condition. On the other hand it is not hard to construct a positive Lebesgue integrable function φ on $\mathbb{R}^n$ such that the measure μ defined by $d\mu/d\lambda = \varphi$ is both non-uniformly Federer and non-uniformly decaying, but is neither Federer nor decaying. The authors do not know, however, if there exist measures satisfying the assumptions of Theorem 6.1 and at the same time singular to any friendly measure.

Closely related to decaying measures is the class of (C,α,s)-good functions, to which we also give a non-uniform version: a function f is (C,α,s)-*good at x with respect to μ* if (4.1) holds for any $\varepsilon > 0$ and any $B = B(x,r)$ with $r < s$.

The proof of Theorem 6.1 requires the following refined version of Theorem 4.3:

Theorem 6.2. *Given a Besicovitch metric space X, positive constants C, D, α, and a measure μ on X, there exists $C' > 0$ with the following property. Suppose that h is a continuous map $U \to G$, ϱ, r_0 are positive constants, $z \in \operatorname{supp}\mu$, and Ξ is a subset of U such that for every $x \in \Xi$ and every $V \in \mathcal{W}$:*

(1) *μ is (D,s)-Federer at x with $s > 3^n r_0$;*
(2) *the function $\ell_V \circ h$ is (C,α,s)-good at x with respect to μ, with $s > 3^n r_0$;*
(3) *$\|\ell_V \circ h\|_{\mu,B(x,r_0)} \ge \varrho$.*

Then for any $0 < \varepsilon \le \varrho$,

$$\mu\left(\left\{x \in \Xi \cap B : \pi\big(h(x)\big) \notin K_\varepsilon\right\}\right) \le C' \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \mu(\widetilde{B}),$$

where $B \stackrel{\text{def}}{=} B(z,r_0)$ and $\widetilde{B} \stackrel{\text{def}}{=} B(z,3^n r_0)$.

Proof. Note that in the proof of Lemma 5.2, only balls centered in E were used. Therefore, repeating the proof but changing the definition of E to

$$E \stackrel{\text{def}}{=} \{x \in \Xi \cap A : \mathcal{W}(x) \ne \varnothing\},$$

one obtains the following statement:

For $d = 0,\dots,n$ there is a constant $C' = C'(D,N,C,d)$ such that the following holds. Let $\mathcal{F}$ be a flag, and let d be the maximal length of a flag which is subordinate to $\mathcal{F}$. Let $A \stackrel{\text{def}}{=} B(z_A,r_A)$, $\widetilde{A} \stackrel{\text{def}}{=} B(z_A,3^d r_A)$, $\Xi \subset X$, and suppose that:

(a$'$) *For each $x \in A \cap \Xi$, μ is (D,s)-Federer at x with $s > 3^d r_A$.*

(b′) *For each $V \in \mathcal{W}$ and each $x \in A \cap \Xi$, the function $\ell_V \circ h$ is (C, α, s)-good at x with respect to μ, with $s > 3^d r_A$.*

(c′) *If V can be added to $\mathcal{F}$ then $\|\ell_V \circ h\|_A \geq \varrho$.*

Let A_0 be the set of all $y \in A$ which are marked relative to $\mathcal{F}$.
Then for every $\varepsilon < \varrho$,

$$\mu(A \cap \Xi \smallsetminus A_0) \leq C' \left(\frac{\varepsilon}{\varrho}\right)^{\alpha} \mu(\widetilde{A}) .$$

From this statement the theorem follows. □

Proof of Theorem 6.1. By Proposition 3.1 we need to show that for any $\gamma > 0$, for μ-a.e. $\mathbf{y} \in \mathbb{R}^n$,

$$g_{\mathbf{t}} \bar{\tau}(\mathbf{y}) \notin K_{e^{-\gamma t}}$$

has only finitely many solutions $\mathbf{t} \in \mathbb{Z}_+^n$. Let $\delta > 0$ be arbitrary. Then there exists a subset $\Xi \subset \operatorname{supp} \mu$ and positive constants D, C, α, r' so that $\mu(\mathbb{R}^n \smallsetminus \Xi) < \delta$ and μ is (D, r')-Federer and (C, α, r')-decaying at every $\mathbf{y} \in \Xi$. Taking $0 < r_0 < 3^{-n} r'$ we see that conditions (1) and (2) of Theorem 6.2 are satisfied.

By compactness and nonplanarity we see that there exists $\varrho > 0$ so that for every $\mathbf{x} \in \operatorname{supp} \mu$ and any affine hyperplane $\mathcal{L}$ one has

$$\|d_{\mathcal{L}}\|_{\mu, B(\mathbf{x}, r_0)} > \varrho ,$$

and, as in the derivation of Theorem 3.3 from Theorem 4.3, this implies that for any $V \in \mathcal{W}$ and $\mathbf{x} \in \operatorname{supp} \mu$

$$\|\ell_V \circ h\|_{\mu, B(\mathbf{x}, r_0)} > \varrho .$$

In particular, we see that condition (3) of Theorem 6.2 also holds.

Applying that theorem, we see that

$$\sum_{\mathbf{t} \in \mathbb{Z}_+^n} \mu\left(\{\mathbf{y} \in \Xi : g_{\mathbf{t}} \tau(\mathbf{y}) \notin K_{e^{-\gamma t}}\}\right) \leq C' \sum_{\mathbf{t} \in \mathbb{Z}_+^n} e^{-\gamma \alpha t} < \infty ,$$

so by Borel-Cantelli and since γ was arbitrary, μ-almost every $\mathbf{y} \in \Xi$ is not very well multiplicatively approximable. Since δ was also arbitrary, this establishes Theorem 6.1. □

7. Pushforwards

In this section, μ is a measure on $\mathbb{R}^d$, U is an open subset of $\mathbb{R}^d$, and $\mathbf{f} : U \to \mathbb{R}^n$ is a map which pushes μ forward. Let us denote by $\mathcal{S}_{\mathbf{f}}$ the $\mathbb{R}$-linear span of $1, f_1, \ldots, f_n$, that is, the space of functions of the form

$$f = c_0 + c_1 f_1 + \cdots + c_n f_n , \tag{7.1}$$

where $c_i \in \mathbb{R}$. It will be convenient to introduce the following definition. Given C, $\alpha > 0$, a subset U of $\mathbb{R}^d$, a measure μ on U, and a real-valued function f on U, say that f is *absolutely (C, α)-good on U with respect to* μ if for any open ball $B \subset U$ centered in $\operatorname{supp} \mu$ and any $\varepsilon > 0$ one has

$$\mu\left(\{\mathbf{x} \in B : |f(\mathbf{x})| < \varepsilon\}\right) \leq C \left(\frac{\varepsilon}{\|f\|_B}\right)^{\alpha} \mu(B) . \tag{7.2}$$

Clearly being absolutely (C,α)-good implies being (C,α)-good, and the converse is true for measures having full support. Now we can state sufficient conditions, written in terms of μ, for $\mathbf{f}_*\mu$ to be friendly.

Lemma 7.1. *Let μ be a D-Federer measure on an open subset U of $\mathbb{R}^d$, and let $\mathbf{f} : U \to \mathbb{R}^n$ and $C, K, \alpha > 0$ be such that*

(i) *$\mathbf{f}$ is K-bi-Lipschitz, that is, for any $\mathbf{x}_1, \mathbf{x}_2 \in U$ one has*

$$\frac{1}{K}\operatorname{dist}(\mathbf{x}_1,\mathbf{x}_2) \le \operatorname{dist}\big(\mathbf{f}(\mathbf{x}_1),\mathbf{f}(\mathbf{x}_2)\big) \le K \operatorname{dist}(\mathbf{x}_1,\mathbf{x}_2)\,;$$

(ii) *any $f \in \mathcal{S}_{\mathbf{f}}$ is absolutely (C,α)-good on U with respect to μ.*

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

Proof. To show that $\mathbf{f}_*\mu$ is Federer, for any $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ with $\mathbf{x}_0 \in U$ take

$$\widetilde{U} \overset{\text{def}}{=} B(\mathbf{y}_0, r_0) \quad \text{where} \quad r_0 \le \frac{1}{2K}\operatorname{dist}(\mathbf{x}_0, \partial U)\,. \tag{7.3}$$

Then for any $\mathbf{y} \in \widetilde{U} \cap \operatorname{supp}\mathbf{f}_*\mu$ one has $\mathbf{y} = \mathbf{f}(\mathbf{x})$ for some $\mathbf{x} \in B(\mathbf{x}_0, Kr_0) \cap \operatorname{supp}\mu$. Now, for $B(\mathbf{y}, 3r) \subset \widetilde{U}$ one can write

$$\begin{aligned}\mathbf{f}_*\mu\big(B(\mathbf{y},3r)\big) &= \mu\left(\mathbf{f}^{-1}\big(B(\mathbf{y},3r)\big)\right) \le \mu\big(B(\mathbf{x},3Kr)\big)\\ &\overset{(2.3)}{\le} \frac{(3K^2)^\beta}{c}\mu\big(B(\mathbf{x},r/K)\big) \le \frac{(3K^2)^\beta}{c}\mathbf{f}_*\mu\big(B(\mathbf{y},r)\big)\,.\end{aligned}$$

To prove that $\mathbf{f}_*\mu$ is decaying, for any $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ with $\mathbf{x}_0 \in U$ take $\widetilde{U}$ as above, and for an affine hyperplane $\mathcal{L}$ put $d_{\mathcal{L}}(\mathbf{y}) = |c_0 + c_1y_1 + \cdots + c_ny_n|$ for some $c_0, c_1, \ldots, c_n \in \mathbb{R}$. Then for any $\mathbf{y}' = \mathbf{f}(\mathbf{x}') \in \widetilde{U} \cap \operatorname{supp}\mathbf{f}_*\mu$ and $B = B(\mathbf{y}', r) \subset \widetilde{U}$ write

$$\begin{aligned}\mathbf{f}_*\mu\left(B \cap \mathcal{L}^{(\varepsilon)}\right) &= \mathbf{f}_*\mu\left(\{\mathbf{y} \in B : |c_0 + c_1y_1 + \cdots + c_ny_n| < \varepsilon\}\right)\\
&\overset{\text{with } f \text{ as in (7.1)}}{\le} \mu\left(\{\mathbf{x} \in \mathbf{f}^{-1}(B) : |f(\mathbf{x})| < \varepsilon\}\right) \le \mu\left(\{\mathbf{x} \in B(\mathbf{x}, Kr) : |f(\mathbf{x})| < \varepsilon\}\right)\\
&\overset{(7.2)}{\le} C\left(\frac{\varepsilon}{\|f\|_{B(\mathbf{x},Kr)}}\right)^\alpha \mu\big(B(\mathbf{x},Kr)\big)\\
&\overset{\text{Federer}}{\le} C\frac{(K^2)^\beta}{c}\left(\frac{\varepsilon}{\|f\|_{B(\mathbf{x},Kr)}}\right)^\alpha \mu\big(B(\mathbf{x},r/K)\big)\\
&= C\frac{(K^2)^\beta}{c}\left(\frac{\varepsilon}{\|d_{\mathcal{L}}\|_{\mathbf{f}(B(\mathbf{x},Kr))}}\right)^\alpha \mathbf{f}_*\mu\big(\mathbf{f}(B(\mathbf{x},r/K))\big)\\
&\le C\frac{(K^2)^\beta}{c}\left(\frac{\varepsilon}{\|d_{\mathcal{L}}\|_{B,\mathbf{f}_*\mu}}\right)^\alpha \mathbf{f}_*\mu(B)\,.\end{aligned}$$

Finally, observe that (7.2) clearly implies that for any ball $B \subset U$ centered at $\operatorname{supp}\mu$ and for $\mathcal{L}, f$ as above, one has

$$\mathbf{f}_*\mu\big(\mathbf{f}(B) \cap \mathcal{L}\big) = \mu\big(\{\mathbf{x} \in B : f(\mathbf{x}) = 0\}\big) = 0\,. \qquad \square$$

Note that if $\mathbf{f}$ is smooth and nonsingular at $\mathbf{x}_0 \in U$, one can find a neighborhood V of $\mathbf{x}_0$ and $K > 0$ such that $\mathbf{f}|_V$ is K-bi-Lipschitz. Thus to prove Theorem 2.1, it suffices to show that for any nondegeneracy point $\mathbf{x}_0 \in U \cap \operatorname{supp}\mu$ of $\mathbf{f}$ one can find a neighborhood V of $\mathbf{x}_0$ and $C, \alpha > 0$ such that any $f \in \mathcal{S}_{\mathbf{f}}$ is absolutely (C, α)-good on V with respect to μ.

The case in which $\mu = \lambda$ is Lebesgue measure (for which the notions of 'good' and 'absolutely good' coincide) was the subject of [KM], where the following statement was proved:

Proposition 7.2 ([KM, Proposition 3.4])**.** *Let $\mathbf{f} = (f_1, \dots, f_n)$ be a C^ℓ map from an open subset U of $\mathbb{R}^d$ to $\mathbb{R}^n$, and let $\mathbf{x}_0 \in U$ be such that $\mathbf{f}$ is ℓ-nondegenerate at $\mathbf{x}_0$. Then there exists a neighborhood $V \subset U$ of $\mathbf{x}_0$ and positive C such that any $f \in \mathcal{S}_{\mathbf{f}}$ is $(C, 1/d\ell)$-good on V with respect to λ.*

The proof of the above proposition involves induction on d, and it is not clear to the authors how to adapt it to measures μ other than Lebesgue. Below we develop an alternative approach which yields a similar result for any absolutely decaying Federer measure, thus finishing the proof of Theorem 2.1, and, in many cases, improves the conclusion of Proposition 7.2 (see a discussion below). We obtain:

Proposition 7.3. *Let $U \subset \mathbb{R}^d$ be open and let $\mathbf{f} : U \to \mathbb{R}^n$ be a $C^{\ell+1}$ map which is ℓ-nondegenerate at $\mathbf{x}_0 \in U$. Let μ be a measure which is D-Federer and absolutely (C, α)-decaying on U for some $D, C, \alpha > 0$. Then there exists a neighborhood $V \subset U$ of $\mathbf{x}_0$ and positive $\widetilde{C}$ such that any $f \in \mathcal{S}_{\mathbf{f}}$ is absolutely $(\widetilde{C}, \frac{\alpha}{2^{\ell+1}-2})$-good on V with respect to μ.*

It is interesting to compare the two propositions above. For $d = 1$, Proposition 7.2 gives the optimal exponent $\zeta(\ell) = \ell^{-1}$, which is much better than $(2^{\ell+1} - 2)^{-1}$ provided by Proposition 7.3. However, as d grows, the exponent in Proposition 7.2 tends to zero, whereas in Proposition 7.3 the exponent does not depend on d. Another difference is that our proof of Proposition 7.3 requires an extra derivative. It seems to be a challenging open problem to find the optimal constant $\eta = \eta(\ell, d)$ with the following property: *let $\mathbf{f} : U \to \mathbb{R}^n$ be a map which is ℓ-nondegenerate at $\mathbf{x}_0 \in U \subset \mathbb{R}^d$, let $\alpha > 0$, and suppose a measure μ is D-Federer and absolutely (C, α)-decaying on U for some positive C, D; then there exist a neighborhood $V \subset U$ of $\mathbf{x}_0$ and $\widetilde{C} > 0$ such that any $f \in \mathcal{S}_{\mathbf{f}}$ is absolutely $(\widetilde{C}, \eta\alpha)$-good on V with respect to μ.*

The rest of the section is devoted to the proof of Proposition 7.3. We need to introduce some notation. Denote by ∂_i the operator of partial differentiation of functions on $\mathbb{R}^d$ with respect to x_i, $i = 1, \dots, d$. For a *multiindex* $\beta = (j_1, \dots, j_d)$, denote $\partial_\beta = \partial_1^{j_1} \circ \cdots \circ \partial_d^{j_d}$, and define the *order* $|\beta|$ of β by $|\beta| \stackrel{\text{def}}{=} j_1 + \cdots + j_d$. If β and γ are two multiindices, then $\beta + \gamma$ denotes the multiindex determined by $\partial_{\beta+\gamma} = \partial_\beta \circ \partial_\gamma = \partial_\gamma \circ \partial_\beta$.

Lemma 7.4. *Let f be a C^k function on a ball B of radius $r > 0$. Then for any multiindex β with $|\beta| = k$,*

$$\|f\|_B \geq \frac{r^k}{2^k(k+1)^k} \inf_B |\partial_\beta f| \,.$$

Proof. The Lemma follows from the following standard generalization of the mean value theorem: let $f \in C^k(\mathbb{R}^d)$ and, for $i = 1, \dots, d$, define

$$D_i f(\mathbf{x}) = f(x_1, \dots, x_i + 1, \dots, x_d) - f(\mathbf{x}) .$$

For $\beta = (j_1, \dots, j_d)$, set

$$D_\beta = D_1^{j_1} \circ \cdots \circ D_d^{j_d} .$$

Thus $D_\beta f(0)$ is a linear combination of the values $f(\mathbf{x}_1), \dots, f(\mathbf{x}_m)$, where $\mathbf{x}_i \in B(0, k+1)$, $i = 1, \dots, m \le 2^k$, and the sum of the absolute value of the coefficients in this combination equal to 2^k. The generalized mean value theorem says that there is some point $\mathbf{x}$ in the convex hull of $\mathbf{x}_1, \dots, \mathbf{x}_m$ such that $D_\beta f(0) = \partial_\beta f(\mathbf{x})$. Thus

$$\inf_{B(0,k+1)} |\partial_\beta f| \le |\partial_\beta f(\mathbf{x})| = |D_\beta f(0)| \le 2^k \, \|f\|_{B(0,k+1)} \ ;$$

rescaling, one gets the desired inequality. □

Lemma 7.5. *Let positive C, α, c, t, ε be given, and let μ be absolutely (C, α)-decaying on $U \subset \mathbb{R}^d$. Suppose we are given a ball $B(\mathbf{y}, t) \subset U$ with $\mathbf{y} \in \operatorname{supp} \mu$ and a C^2 function $f : B(\mathbf{y}, t) \to \mathbb{R}$ such that*

(i) $\|\nabla f(\mathbf{y})\| \ge c$;

(ii) $\|\partial_\beta f\|_{B(\mathbf{y},t)} \le \frac{\varepsilon}{dt^2}$ *for any multiindex β of order* 2.

Then

$$\mu(\{\mathbf{x} \in B(\mathbf{y}, t) : |f(\mathbf{x})| < \varepsilon\}) \le C \left(\frac{3\varepsilon}{ct}\right)^\alpha \mu\big(B(\mathbf{y}, t)\big) .$$

Proof. By the multidimensional Taylor expansion, for any $\mathbf{x} \in B(\mathbf{y}, t)$ one has

$$|f(\mathbf{x}) - f(\mathbf{y}) - \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle| \le \frac{1}{2} d \frac{\varepsilon}{dt^2} \|\mathbf{x} - \mathbf{y}\|^2 < \frac{\varepsilon}{2},$$

and hence for any $\mathbf{x}, \mathbf{x}_0 \in B(\mathbf{y}, t)$,

$$|f(\mathbf{x}) - f(\mathbf{x}_0) - \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{x}_0 \rangle| < \varepsilon .$$

Now suppose there is some $\mathbf{x}_0 \in B(\mathbf{y}, t)$ for which $|f(\mathbf{x}_0)| < \varepsilon$ (otherwise there is nothing to prove). Then for any $\mathbf{x} \in B(\mathbf{y}, t)$ with $|f(\mathbf{x})| < \varepsilon$ one has

$$|\langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{x}_0 \rangle| < 3\varepsilon,$$

i.e. $\mathbf{x}$ is in a strip of width at most $3\varepsilon/c$. The Lemma now follows immediately from the decay condition. □

The next statement is an analogue of [KM, Lemma 3.3].

Theorem 7.6. *Let $B \subset \mathbb{R}^d$ be a ball of radius $r > 0$, and let $k \in \mathbb{Z}_+$, $C, \alpha > 0$ and $0 < s \le 1$, $S \ge 4$ be given. Suppose $f : B \to \mathbb{R}$ is a C^{k+1} function such that for some multiindex β of order k,*

$$\inf_B |\partial_\beta f| > \frac{s}{r^k} \|f\|_B \ , \tag{7.4}$$

and such that for any multiindex γ with $|\gamma| \le k+1$ one has

$$\|\partial_\gamma f\|_B \le \frac{S}{r^{|\gamma|}} \|f\|_B \ . \tag{7.5}$$

Suppose also that μ is absolutely (C,α)-decaying on B, and let $\hat{B}$ be a ball concentric with B of radius $r/2$. Then for any $\varepsilon < s$ one has

$$\mu\left(\left\{\mathbf{x} \in \hat{B} : |f(\mathbf{x})| < \varepsilon \|f\|_B\right\}\right) \le C_k C \left(\frac{S}{s}\right)^{2^k\alpha} \varepsilon^{\zeta_k \alpha} \mu(B)\,,$$

where

$$C_k = 2^k N_d \left(3\sqrt{d} \prod_{i=1}^{k} (2i)^{(i-1)2^{i-1}}\right)^{\alpha}, \tag{7.6}$$

$$\zeta_k = (2^{k+1}-2)^{-1}\,, \tag{7.7}$$

and N_d is a Besicovitch constant for $\mathbb{R}^d$.

Proof. The proof is by induction on k. For $k=0$, the theorem clearly holds for $C_0 = 1$ and $\zeta_0 = \infty$. Assume now that $k \ge 1$, write $\beta = \beta_1 + \beta'$, where $|\beta_1| = 1$, and denote $\partial_{\beta_1} f$ by g.

Define

$$E \stackrel{\text{def}}{=} \{\mathbf{x} \in B : |f(\mathbf{x})| < \varepsilon \|f\|_B\}\,,$$

and, for $\delta > 0$ to be determined later, define

$$E_\delta \stackrel{\text{def}}{=} \left\{\mathbf{x} \in \hat{B} : |g(\mathbf{x})| < \delta \|g\|_B\right\}.$$

First let us estimate the measure of E_δ. Applying Lemma 7.4 to g and β' and using (7.4), one has

$$\frac{(2k)^{k-1}}{r^{k-1}} \|g\|_B \ge \inf_B |\partial_{\beta'} g| = \inf_B |\partial_\beta f| \ge \frac{s}{r^k} \|f\|_B\,,$$

therefore

$$\frac{s}{r(2k)^{k-1}} \|f\|_B \le \|g\|_B = \|\partial_{\beta_1} f\|_B \stackrel{(7.5)}{\le} \frac{S}{r} \|f\|_B\,. \tag{7.8}$$

This calculation shows that g satisfies all the conditions of the theorem with k, β, s and S replaced by $k-1$, β', s/S and $(2k)^{k-1}S/s$ respectively. Applying the induction hypothesis, one gets

(7.9)

$$\mu(E_\delta) \le C_{k-1} C \left(\frac{(2k)^{k-1} S^2}{s^2}\right)^{2^{k-1}\alpha} \delta^{\zeta_{k-1}\alpha} \mu(B) \stackrel{(7.6)}{=} \frac{1}{2} C_k C \left(\frac{S}{s}\right)^{2^k\alpha} \delta^{\zeta_{k-1}\alpha} \mu(B)\,.$$

Now define

$$t = r\sqrt{\frac{\varepsilon}{dS}} \tag{7.10}$$

and consider some ball $B' = B(\mathbf{y}, t)$ with $\mathbf{y} \in (\hat{B} \cap \operatorname{supp}\mu) \smallsetminus E_\delta$. Note that $B' \subset B$ since $\varepsilon < s < 4 \le S$, and one can check, using (7.5) and

$$\|\nabla f(\mathbf{y})\| \ge |\partial_{\beta_1} f(\mathbf{y})| = |g(\mathbf{y})| \ge \delta \|g\|_B\,,$$

that this ball satisfies all the assumptions of Lemma 7.5, with ε and c replaced by $\varepsilon\,\|f\|_B$ and $\delta\,\|g\|_B$, respectively. Therefore

$$\mu(B'\cap E) \le C\left(\frac{3\varepsilon\,\|f\|_B}{\delta\,\|g\|_B\,t}\right)^{\alpha}\mu(B') \overset{(7.8)}{\le} C\left(\frac{3\varepsilon r(2k)^{k-1}}{\delta s t}\right)^{\alpha}\mu(B')$$
$$\overset{(7.10)}{=} C\left((dS/\varepsilon)^{1/2}\,\frac{3\varepsilon(2k)^{k-1}}{\delta s}\right)^{\alpha}\mu(B') \le C\left(3\sqrt{d}(2k)^{k-1}\frac{S}{s}\right)^{\alpha}\frac{\varepsilon^{\alpha/2}}{\delta^{\alpha}}\mu(B').$$

Clearly one can cover $(\hat{B}\cap\operatorname{supp}\mu)\smallsetminus E_\delta$ by balls as above with multiplicity $\le N_d$. Let $\mathcal{F}$ denote such a cover; then

$$\mu\big((\hat{B}\smallsetminus E_\delta)\cap E\big) \le \sum_{B'\in\mathcal{F}}\mu(B'\cap E) \le CN_d\left(3\sqrt{d}(2k)^{k-1}\frac{S}{s}\right)^{\alpha}\frac{\varepsilon^{\alpha/2}}{\delta^{\alpha}}\mu(B)\,. \tag{7.11}$$

Using (7.11) together with (7.9) and (7.6), one gets that

$$\mu(\hat{B}\cap E) \le \frac{1}{2}C_kC\left(\frac{S}{s}\right)^{2^k\alpha}\left[\delta^{\zeta_{k-1}\alpha} + \frac{\varepsilon^{\alpha/2}}{\delta^{\alpha}}\right]\mu(B)\,.$$

Choosing the optimal $\delta = \varepsilon^{\frac{1}{2(\zeta_{k-1}+1)}}$ one obtains

$$\mu(\hat{B}\cap E) \le C_kC\left(\frac{S}{s}\right)^{2^k\alpha}\varepsilon^{\frac{\zeta_{k-1}\alpha}{2(\zeta_{k-1}+1)}}\mu(B).$$

Thus the theorem is proved with $\zeta_k = \frac{\zeta_{k-1}}{2\zeta_{k-1}+2}$, and solving for ζ_k yields (7.7). $\square$

In order to apply Theorem 7.6 to nondegenerate maps, we will need three additional lemmas. Below we fix $d,\,\ell\in\mathbb{N}$, let $\mathcal{P}_{d,\ell}$ denote the space of all polynomial maps $\mathbb{R}^d\to\mathbb{R}$ of degree at most ℓ, and let B_1 stand for the unit ball $B(0,1)\subset\mathbb{R}^d$.

Lemma 7.7. *There exist $s,S>0$ such that for any nonzero $P\in\mathcal{P}_{d,\ell}$ one has*

$$\max_{\beta}\inf_{\mathbf{x}\in B_1}|\partial_\beta P(\mathbf{x})| > s\|P\|_{B_1}\,,$$

and

$$\max_{\gamma}\|\partial_\gamma P\|_{B_1} < S\|P\|_{B_1}.$$

Proof. For the first statement, let $s=\frac{1}{2(2\sqrt{d}+1)^{\ell+1}}$. Given $P\in\mathcal{P}_{d,\ell}$, let k be the largest integer j for which there is a multiindex $\beta=\beta_j$ of order j and $\mathbf{y}_j\in B_1$ such that

$$|\partial_\beta P(\mathbf{y}_j)| > \frac{\|P\|_{B_1}}{2(2\sqrt{d}+1)^j}\,.$$

Clearly $0\le k\le\deg P\le\ell$.

We claim that $\inf_{\mathbf{x}\in B_1}|\partial_\beta P(\mathbf{x})| > s\|P\|_{B_1}$ for $\beta=\beta_k$. Replacing P with a scalar multiple, we may assume with no loss of generality that $\|P\|_{B_1}=1$. By definition of k, $|\partial_\beta P(\mathbf{y}_k)| > \frac{1}{2(2\sqrt{d}+1)^k}$. Suppose by contradiction that there is

$\mathbf{y}' \in B_1$ with $|\partial_\beta P(\mathbf{y}')| \le s = \frac{1}{2(2\sqrt{d}+1)^{k+1}}$. Then by the mean value theorem, there is a point $\mathbf{y}''$ on the segment between $\mathbf{y}_k$ and $\mathbf{y}'$ for which

$$\begin{aligned}\frac{\sqrt{d}}{(2\sqrt{d}+1)^{k+1}} &= \frac{1}{2(2\sqrt{d}+1)^k} - \frac{1}{2(2\sqrt{d}+1)^{k+1}} \\ &< |\partial_\beta P(\mathbf{y}_k) - \partial_\beta P(\mathbf{y}')| = |\langle \nabla \partial_\beta P(\mathbf{y}''), \mathbf{y}_k - \mathbf{y}'\rangle| \\ &\le \|\nabla \partial_\beta P(\mathbf{y}'')\| \|\mathbf{y}_k - \mathbf{y}'\| \le 2\sqrt{d} \max_{|\beta_1|=1} |\partial_{\beta_1+\beta} P(\mathbf{y}'')| \,.\end{aligned}$$

Hence there is β_1, of order 1, such that for $\beta' = \beta_1 + \beta$ one has $|\partial_{\beta'} P(\mathbf{y}'')| > \frac{1}{2(2\sqrt{d}+1)^{k+1}}$, contradicting the definition of k.

The second statement follows from the fact that both $P \mapsto \|P\|_{B_1}$ and $P \mapsto \max_\gamma \|\partial_\gamma P\|_{B_1}$ are norms on $\mathcal{P}_{d,\ell}$. $\square$

We now define

$$\mathcal{S}^1_{\mathbf{f}} \stackrel{\text{def}}{=} \{c_0 + c_1 f_1 + \cdots + c_n f_n : \sum_{i=0}^n c_i^2 = 1\} \,,$$

and for each C^ℓ function $g : U \to \mathbb{R}$, where U is an open subset of $\mathbb{R}^d$ and $\mathbf{y} \in U$, let $P^\ell_{g,\mathbf{y}}$ be the ℓ-th order Taylor polynomial of g at $\mathbf{y}$.

Lemma 7.8. *Suppose $U \subset \mathbb{R}^d$ is open and $\mathbf{f} : U \to \mathbb{R}^n$ is a C^ℓ map. Then:*

(i) *For any $\eta > 0$ there is a neighborhood $V \subset U$ of $\mathbf{x}_0$ such that for any ball $B(\mathbf{y}, r) \subset V$, any multiindex β of order at most ℓ, and any $f \in \mathcal{S}^1_{\mathbf{f}}$,*

$$\|\partial_\beta f - \partial_\beta P^\ell_{f,\mathbf{y}}\|_{B(\mathbf{y},r)} < \eta r^{\ell - |\beta|} \,.$$

(ii) *If $\mathbf{f}$ is ℓ-nondegenerate at $\mathbf{x}_0$, then there is an open set $V_0 \subset U$ containing $\mathbf{x}_0$ and $\varkappa > 0$ such that for any $B(\mathbf{y}, r) \subset V$, and any $f \in \mathcal{S}^1_{\mathbf{f}}$,*

$$\|f\|_{B(\mathbf{y},r)} \ge \varkappa r^\ell$$

and

$$\frac{1}{2} \le \frac{\|P^\ell_{f,\mathbf{y}}\|_{B(\mathbf{y},r)}}{\|f\|_{B(\mathbf{y},r)}} \le \frac{3}{2} \,.$$

Proof. Recall the following error estimate, which is an easy consequence of the Lagrange form of the error term in Taylor's formula of order $l-1$:

$$\|g - P^\ell_{g,\mathbf{y}}\|_{B(\mathbf{y},r)} \le \frac{d^\ell}{\ell!} \max_{|\gamma|=\ell,\, \mathbf{x} \in B(\mathbf{y},r)} \|\partial_\gamma g(\mathbf{x}) - \partial_\gamma g(\mathbf{y})\| \, r^\ell. \tag{7.12}$$

Note that for any multiindex β of order $k \le \ell$ one has $\partial_\beta P^\ell_{f,\mathbf{y}} = P^{l-k}_{\partial_\beta f, \mathbf{y}}$. By the chain rule, the derivative of any $f \in \mathcal{S}^1_{\mathbf{f}}$ is a composition of the derivative of $\mathbf{f}$ and a linear functional of uniformly bounded norm. Therefore the collection

$$\{\partial_\gamma f : \; f \in \mathcal{S}^1_{\mathbf{f}}, \, |\gamma| = \ell\}$$

is equicontinuous, and hence, by making V sufficiently small, one can ensure that for any multiindex γ of order ℓ, any $f \in \mathcal{S}^1_{\mathbf{f}}$, and any $\mathbf{x}_1, \mathbf{x}_2 \in V$ one has

$$|\partial_\gamma f(\mathbf{x}_1) - \partial_\gamma f(\mathbf{x}_2)| < \eta \min_{0 \le k \le \ell} \frac{k!}{d^k}.$$

Now one obtains (i) by applying (7.12) to $g = \partial_\beta f$ with l replaced by $l - k$.

For (ii), let $V \subset U$ be a bounded open subset containing $\mathbf{x}_0$ which is small enough so that $\mathbf{f}$ is ℓ-nondegenerate at $\mathbf{y}$ for any $\mathbf{y} \in \overline{V}$. Equip $\mathcal{P}_{d,\ell}$ with the norm assigning to a polynomial the maximum of the absolute values of its coefficients. For each $\mathbf{y} \in \overline{V}$, $f \in \mathcal{S}^1_{\mathbf{f}}$ let

$$Q_{f,\mathbf{y}}(\mathbf{x}) = P^\ell_{f,\mathbf{y}}(\mathbf{x} - \mathbf{y}).$$

By nondegeneracy, $Q_{f,\mathbf{y}} \neq 0$ for any $\mathbf{y} \in \overline{V}$ and any $f \in \mathcal{S}^1_{\mathbf{f}}$, and by compactness, the set $\{Q_{f,\mathbf{y}} : f \in \mathcal{S}^1_{\mathbf{f}},\ \mathbf{y} \in \overline{V}\} \subset \mathcal{P}_{d,\ell}$ is bounded away from zero. For each r and $\mathbf{y}$ for which $B(\mathbf{y}, r) \subset V$, let

$$Q_{r,f,\mathbf{y}}(\mathbf{x}) \stackrel{\text{def}}{=} r^{-\ell} Q_{f,\mathbf{y}}(r\mathbf{x}).$$

By comparing the coefficients of $Q_{f,\mathbf{y}}$ and $Q_{r,f,\mathbf{y}}$ one sees that

$$\{Q_{r,f,\mathbf{y}} :\ B(\mathbf{y}, r) \subset V,\ f \in \mathcal{S}^1_{\mathbf{f}}\} \subset \mathcal{P}_{d,\ell}$$

is also bounded away from zero. Now comparing with the norm $P \mapsto \|P\|_{B(0,1)}$ one finds that there is $\varkappa$ such that for all $r, f, \mathbf{y}$ as above,

$$\|P^\ell_{f,\mathbf{y}}\|_{B(\mathbf{y},r)} = r^\ell \|Q_{r,f,\mathbf{y}}\|_{B(0,1)} \geq 2\varkappa r^\ell. \tag{7.13}$$

Also, from the case $|\beta| = 0$ in (i) it follows that if V is sufficiently small,

$$\|f - P^\ell_{f,\mathbf{y}}\|_{B(\mathbf{y},r)} \leq \varkappa r^\ell. \tag{7.14}$$

Putting together (7.13) and (7.14) yields (ii). □

From now until the end of the section, we let U, $\mathbf{f}$, $\mathbf{x}_0$ and ℓ be as in Proposition 7.3. It will be useful to consider restrictions of functions to small balls and renormalize them as follows: if V is a small enough neighborhood of $\mathbf{x}_0$, $B(\mathbf{y}, r) \subset V$ and $f \in \mathcal{S}^1_{\mathbf{f}}$, we let $g_{r,f,\mathbf{y}}$ be a function on B_1 given by

$$g_{r,f,\mathbf{y}}(\mathbf{x}) \stackrel{\text{def}}{=} \frac{f(\mathbf{y} + r\mathbf{x})}{\|f\|_{B(\mathbf{y},r)}},$$

and define

$$\mathcal{G}(\mathbf{f}, V) \stackrel{\text{def}}{=} \{g_{r,f,\mathbf{y}} : f \in \mathcal{S}^1_{\mathbf{f}},\ B(\mathbf{y}, r) \subset V\}.$$

To simplify notation, for $g \in \mathcal{G}(\mathbf{f}, V)$ we will write P_g instead of $P^\ell_{g,0}$; note that for $g = g_{r,f,\mathbf{y}}$ one has $P_g(\mathbf{x}) = \frac{P^\ell_{f,\mathbf{y}}(\mathbf{y}+r\mathbf{x})}{\|f\|_{B(\mathbf{y},r)}}$.

Lemma 7.9. *For any $\eta > 0$ there exists a neighborhood $V \subset U$ of $\mathbf{x}_0$ such that for any $g \in \mathcal{G}(\mathbf{f}, V)$ one has*

$$\max_{|\beta| \leq \ell+1} \|\partial_\beta g - \partial_\beta P_g\|_{B_1} < \eta.$$

Proof. When $|\beta| \leq \ell$ we apply Lemma 7.8, replacing η with $\eta' = \eta\varkappa$, and obtain for small enough V:

$$\begin{aligned}\|\partial_\beta g - \partial_\beta P_g\|_{B_1} &= \frac{1}{\|f\|_{B(\mathbf{y},r)}} \left\|\partial_\beta f(\mathbf{y} + r\mathbf{x}) - \partial_\beta P^\ell_{f,\mathbf{y}}(\mathbf{y} + r\mathbf{x})\right\|_{B_1} \\ &= \frac{1}{\|f\|_{B(\mathbf{y},r)}} \left\|r^{|\beta|}\partial_\beta f - r^{|\beta|}\partial_\beta P^\ell_{f,\mathbf{y}}\right\|_{B(\mathbf{y},r)} < \frac{r^{|\beta|}}{\varkappa r^\ell}\, \eta' r^{\ell - |\beta|} = \eta.\end{aligned}$$

When $|\beta| = \ell + 1$ we argue as in the proof of Lemma 7.8 to find a uniform upper bound K on $\|\partial_\beta f\|_V$ for all $f \in \mathcal{S}^1_{\mathbf{f}}$, and then write

$$\|\partial_\beta g - \partial_\beta P_{g,0}\|_{B_1} = \|\partial_\beta g\|_{B_1} = \frac{\|\partial_\beta f(\mathbf{y} + r\mathbf{x})\|_{B_1}}{\|f\|_{B(\mathbf{y},r)}} \le \frac{r^{\ell+1}}{\varkappa r^\ell}\|\partial_\beta f\|_V \le \frac{rK}{\varkappa} < \eta$$

whenever $\operatorname{diam}(V) < \eta\varkappa/K$. □

We are now ready for the

Proof of Proposition 7.3. By Theorem 7.6, it is enough to find positive s, S and a neighborhood $V \subset U$ of $\mathbf{x}_0$ such that for any $B = B(\mathbf{y}, r) \subset V$ and any $f \in \mathcal{S}_{\mathbf{f}}$ there is a multiindex β with $k = |\beta| \le \ell$ such that (7.4) holds, and such that for all multiindices γ with $|\gamma| \le |\beta| + 1$, (7.5) holds. Since the validity of (7.4) and (7.5) is unaffected by replacing f with a scalar multiple, we may assume that $f \in \mathcal{S}^1_{\mathbf{f}}$. Thus it suffices to find s, S and V such that for any $g \in \mathcal{G}(\mathbf{f}, V)$ there is a multiindex β of order $\le \ell$ such that

$$\inf_{\mathbf{x} \in B_1} |\partial_\beta g(\mathbf{x})| > s\,, \tag{7.15}$$

and for each multiindex γ of order at most $\ell + 1$,

$$\|\partial_\gamma g\|_{B_1} < S\,. \tag{7.16}$$

Note that Lemma 7.8(ii) allows one to choose V such that for any $g \in \mathcal{G}(\mathbf{f}, V)$ one has

$$\frac{1}{2} \le \|P_g\|_{B_1} \le \frac{3}{2}\,;$$

therefore one can use Lemma 7.7 to find s and S such that for any $g \in \mathcal{G}(\mathbf{f}, V)$ (7.15) and (7.16) are satisfied with P_g in place of g. Now to finish the proof it remains to take $\eta = s/2$ and apply Lemma 7.9. □

8. The open set condition

In this section we will discuss measures supported on attractors of finite systems of contracting similarities. The open set condition, defined in §2, was shown by Hutchinson to be very useful. The following summarizes some of the results of [H]:

Theorem 8.1. *Let $\{\mathbf{h}_i = \varrho_i\Theta_i + \mathbf{a}_i : i = 1, \dots, m\}$, be a family of contracting similarity self-maps of $\mathbb{R}^n$ satisfying the open set condition, s its similarity dimension, μ the restriction of H^s to its attractor K. Then:*

(i) $0 < \mu(K) < \infty$*;*

(ii) *for any measurable $A \subset \mathbb{R}^n$,*

$$\mu(A) = \sum_i \varrho_i^s \mu\big(\mathbf{h}_i^{-1}(A)\big)\,; \tag{8.1}$$

(iii) *there exist constants $0 < \lambda_1 \le \lambda_2 < \infty$ such that for every $\mathbf{y} \in K$*

$$\lambda_1 \le \inf_{0<r<\operatorname{diam}(K)} \frac{\mu\big(B(\mathbf{y}, r)\big)}{r^s} \le \sup_{r>0} \frac{\mu\big(B(\mathbf{y}, r)\big)}{r^s} \le \lambda_2\,. \tag{8.2}$$

Remark. In [H], assertion (iii) involves $\limsup, \liminf$ in place of $\sup, \inf$ respectively, but the proof given there, see [H, 5.3(1)(b)], in fact yields (8.2).

Let Σ denote the set of all finite words in the alphabet $\{1, \dots, m\}$. Define for any word $w = i_1 \cdots i_\ell \in \Sigma$

$$\mathbf{h}_w = \mathbf{h}_{i_\ell} \circ \cdots \circ \mathbf{h}_{i_1}$$

and define $a_w \in \mathbb{R}^n,\ \varrho_w \in (0,1),\ \Theta_w \in \mathrm{O}(n, \mathbb{R})$ by the equation

$$\mathbf{h}_w : \mathbf{y} \mapsto a_w + \varrho_w \Theta_w(\mathbf{y})$$

Set

$$\varrho_{\min} = \min_{1 \le i \le m} \varrho_i \,.$$

We say that a finite subset P of Σ is a *complete prefix set* if the corresponding cylindrical sets form a partition of $\{1, \dots, m\}^{\mathbb{N}}$; equivalently, for every $w \in \{1, \dots, m\}^{\mathbb{N}}$ there is a unique $w' \in P$ that coincides with some initial segment of w. It is clear that for any $0 < r < 1$, we can find a complete prefix set P so that for any $w \in P$ we have

$$\varrho_{\min} r \le \varrho_w \le r. \tag{8.3}$$

Proof of Theorem 2.2. Assume using Theorem 8.1(i) that $\mu(K) = 1$. It follows from Theorem 8.1(iii) that μ is D-Federer on $\mathbb{R}^n$, with $D = 3^s \lambda_2 / \lambda_1$. Irreducibility of $\{\mathbf{h}_1, \dots, \mathbf{h}_m\}$ will now be used to show that $\mu(\mathcal{L}) = 0$ for all proper affine subspaces $\mathcal{L}$. Let k be the smallest number for which

$$M \stackrel{\mathrm{def}}{=} \sup_{\dim \mathcal{L} = k} \mu(\mathcal{L}) > 0.$$

Then for any two distinct k-dimensional affine subspaces $\mathcal{L}_1,\ \mathcal{L}_2$, we have $\mu(\mathcal{L}_1 \cap \mathcal{L}_2) = 0$, and hence for any sequence $\mathcal{L}_1, \mathcal{L}_2, \dots$ of distinct k-dimensional affine subspaces, $\mu(\bigcup_j \mathcal{L}_j) = \sum_j \mu(\mathcal{L}_j)$. This implies that

$$\#\{\mathcal{L} : \dim \mathcal{L} = k,\ \mu(\mathcal{L}) \ge M/2\} < \infty,$$

hence there exists $\mathcal{L}_0$ satisfying

$$\mu(\mathcal{L}_0) = M.$$

Applying (8.1) and (2.8) we obtain that $\mu(\mathcal{L}_0)$ is a weighted average of the numbers $\mu\big(\mathbf{h}_i^{-1}(\mathcal{L}_0)\big) \le M$ for $i = 1, \dots, m$, and hence $\mu\big(\mathbf{h}_i^{-1}(\mathcal{L}_0)\big) = M$ for all i. By induction

$$w \in \Sigma \implies \mu\big(\mathbf{h}_w^{-1}(\mathcal{L}_0)\big) = M.$$

Therefore

$$\mathcal{B} = \{\mathbf{h}_w^{-1}(\mathcal{L}_0) : w \in \Sigma\}$$

is finite. For each i, $\mathbf{h}_i(\mathcal{B}) \subset \mathcal{B}$ and by finiteness of $\mathcal{B}$, $\mathbf{h}_i(\mathcal{B}) = \mathcal{B}$. This implies that $\mathcal{B}$ is invariant under $\mathbf{h}_i$, $i = 1, \dots, m$, so by the irreducibility hypothesis, $k = n$.

We now define

$$t(\varepsilon) = \sup_{\mathcal{L}} \mu(\mathcal{L}^{(\varepsilon)}),$$

where the supremum is taken over all affine hyperplanes $\mathcal{L}$.

Lemmas 8.2 and 8.3 below, together with the Federer condition, imply the existence of $\widetilde{C}$ and α such that μ is absolutely $(\widetilde{C}, \alpha)$-decaying on any ball of radius 1, and hence will complete the proof.

Lemma 8.2. *There are positive constants C, α such that for any $\varepsilon > 0$,*

$$t(\varepsilon) \leq C\varepsilon^{\alpha}$$

Lemma 8.3. *There are constants M, c such that for any $\mathbf{y} \in K$, $0 < r < 1$, $\varepsilon > 0$, and any affine hyperplane $\mathcal{L}$, we have*

$$\mu\big(\mathcal{L}^{(r\varepsilon)} \cap B(\mathbf{y}, r)\big) \leq Mt\left(\frac{\varepsilon}{\varrho_{\min}}\right) \mu\big(B(\mathbf{y}, cr)\big).$$

The following technical result, which will be crucial for proving the lemmas, shows why the open set condition is so useful. This idea for putting the open set condition to use is due to Hutchinson.

Lemma 8.4. *There exist a constant M_1 and a ball B_1 containing K so that for any $r > 0$ and any complete prefix set P satisfying* (8.3)*, the collection*

$$\mathcal{F}_P = \{\mathbf{h}_w(B_1) : w \in P\}$$

is an open cover of K, and no point in $\mathbb{R}^n$ is in more than M_1 elements of $\mathcal{F}_P$.

Proof. Let $U \subset \mathbb{R}^n$ be an open set as in the open set condition, $\mathbf{y} \in U$ and $B_0 = B(\mathbf{y}, r_0) \subset U$. Take $B_1 = B(\mathbf{y}, r_1)$ with $r_1 > r_0$ large enough so that $K \subset B_1$.

It follows easily from (2.7) that for any complete prefix set P,

$$K = \bigcup_{w \in P} \mathbf{h}_w(K).$$

Since $K \subset B_1$, $\mathcal{F}_P$ is an open cover of K.

Suppose now that for some $P_0 \subset P$ and some $\mathbf{y} \in \mathbb{R}^n$ we have

$$\mathbf{y} \in \bigcap_{w \in P_0} \mathbf{h}_w(B_1). \tag{8.4}$$

By (8.3), for any $w \in P$, the radius of the ball $\mathbf{h}_w(B_1)$ is at most rr_1 and the radius of $\mathbf{h}_w(B_0)$ is at least $\varrho_{\min} r r_0$. It is also clear from the open set condition that for any complete prefix set, $\{\mathbf{h}_w(B_0) : w \in P\}$ are disjoint.

In other words, if (8.4) holds, the ball of radius $2rr_1$ around $\mathbf{y}$ contains $\#P_0$ disjoint balls of radius $\varrho_{\min} r r_0$. By considering the volumes of these balls,

$$M_1 \stackrel{\text{def}}{=} \left(\frac{2r_1}{\varrho_{\min} r_0}\right)^n \geq \#P_0.$$

$\square$

Proof of Lemma 8.2. Using an easy compactness argument, from the fact that μ is nonplanar we obtain that $t(\varepsilon) \to 0$ as $\varepsilon \to 0$.

We will show below that for M_1 as in Lemma 8.4,

$$\exists\, c > 1 \text{ such that } \quad \forall\, \varepsilon > 0,\ t(\varepsilon^2) \leq M_1 t(c\varepsilon)^2\,. \tag{8.5}$$

Assuming this we obtain by induction for $\ell = 1, 2, \dots,$

$$t(\varepsilon^{2^\ell}) \le M_1^{2^\ell - 1} t(c^{\frac{2^\ell-1}{2^{\ell-1}}}\varepsilon)^{2^\ell}\,. \tag{8.6}$$

Choose δ so that $t(c^2\delta) < \frac{1}{2M_1}$, let $\alpha = \frac{\log 2}{2|\log\delta|}$, and assume with no loss of generality that $M_1 > 1 > \delta$. Suppose first that $\varepsilon \le \delta$, choose ℓ so that $\delta^{2^{\ell+1}} < \varepsilon \le \delta^{2^\ell}$ and apply (8.6) with δ instead of ε. Then

$$t(\varepsilon) \le 2^{-\frac{\log\varepsilon}{2\log\delta}} = \varepsilon^\alpha\,.$$

Now choosing $C > 1$ big enough so that $C\delta^\alpha > 1$ we obtain the conclusion of the lemma for all $\varepsilon > 0$.

It remains to prove (8.5). Recall that for any measurable $A \subset \mathbb{R}^n$, the measure H^s satisfies $H^s\big(\mathbf{h}(A)\big) = \varrho^s H^s(A)$ for $\mathbf{h}$ as in (2.6), and hence

$$w \in \Sigma \implies \mu\big(\mathbf{h}_w(A \cap K)\big) = \varrho_w^s \mu(A).$$

Let P be a complete prefix set for $r = \varepsilon$ as in Lemma 8.4, so

$$w \in P \implies \varepsilon\varrho_{\min} \le \varrho_w \le \varepsilon.$$

Also, take B_1 to be a ball of radius r_1 and M_1 as in that lemma. Let

$$P' \stackrel{\text{def}}{=} \{w \in P : \mathbf{h}_w(K) \cap \mathcal{L}^{(\varepsilon^2)} \ne \varnothing\}.$$

Since P is a complete prefix set,

$$K \cap \mathcal{L}^{(\varepsilon^2)} = \bigcup_{w\in P'} [\mathbf{h}_w(K) \cap \mathcal{L}^{(\varepsilon^2)}].$$

It is also clear that

$$\bigcup_{w\in P'} \mathbf{h}_w(K) \subset \mathcal{L}^{(c'\varepsilon)}$$

for $c' \stackrel{\text{def}}{=} r_1 + 1$. Since no point is in more than M_1 of the $\mathbf{h}_w(K)$, we have that

$$\begin{aligned} t(c'\varepsilon) &\ge \mu(\mathcal{L}^{(c'\varepsilon)}) \ge \mu\Big(\bigcup_{w\in P'} \mathbf{h}_w(K)\Big) \\ &\ge \frac{1}{M_1}\sum_{w\in P'} \mu\big(\mathbf{h}_w(K)\big) = \frac{1}{M_1}\sum_{w\in P'} \varrho_w^s\,. \end{aligned} \tag{8.7}$$

Hence

$$\begin{aligned} \mu(\mathcal{L}^{(\varepsilon^2)}) &\le \sum_{w\in P'} \mu(\mathbf{h}_w(K) \cap \mathcal{L}^{(\varepsilon^2)}) = \sum_{w\in P'} \varrho_w^s \mu\big(\mathbf{h}_w^{-1}(\mathcal{L}^{(\varepsilon^2)})\big) \\ &\le t\left(\frac{\varepsilon}{\varrho_{\min}}\right)\sum_{w\in P'} \varrho_w^s \\ &\stackrel{(8.7)}{\le} M_1 t\left(\frac{\varepsilon}{\varrho_{\min}}\right) t(c'\varepsilon)\,. \end{aligned}$$

Taking the supremum over $\mathcal{L}$ we obtain (8.5), with $c \stackrel{\text{def}}{=} \max(c', \frac{1}{\varrho_{\min}})$. $\square$

Proof of Lemma 8.3. We employ a strategy similar to the one used to prove (8.5). Let $B(\mathbf{y}, r)$ be arbitrary with $r < 1$. Let P be a complete prefix set as in Lemma 8.4, so (8.3) holds for every $w \in P$, and let

$$P' \stackrel{\text{def}}{=} \{w \in P : \mathbf{h}_w(K) \cap B(\mathbf{y}, r) \neq \varnothing\}.$$

As before,

$$K \cap B(\mathbf{y}, r) \subset \bigcup_{w \in P'} \mathbf{h}_w(K)$$

and arguing as in the proof of (8.7) one has

$$\sum_{w \in P'} \varrho_w^s \le M_1 \mu\big(B(\mathbf{y}, c'r)\big). \tag{8.8}$$

Thus for any affine hyperplane $\mathcal{L}$,

$$\begin{aligned} \mu\big(\mathcal{L}^{(r\varepsilon)} \cap B(\mathbf{y}, r)\big) &\le \sum_{w \in P'} \mu\big(\mathbf{h}_w(K) \cap \mathcal{L}^{(r\varepsilon)}\big) = \sum_{w \in P'} \varrho_w^s \mu\big(\mathbf{h}_w^{-1}(\mathcal{L}^{(r\varepsilon)})\big) \\ &\stackrel{(8.8)}{\le} t\left(\frac{\varepsilon}{\varrho_{\min}}\right) M_1 \mu\big(B(\mathbf{y}, c'r)\big)\,. \end{aligned}$$

□

9. Products of friendly measures

In this section we prove Theorem 2.3, that is, show that the classes of friendly (resp., absolutely decaying and Federer) measures are stable under Cartesian products.

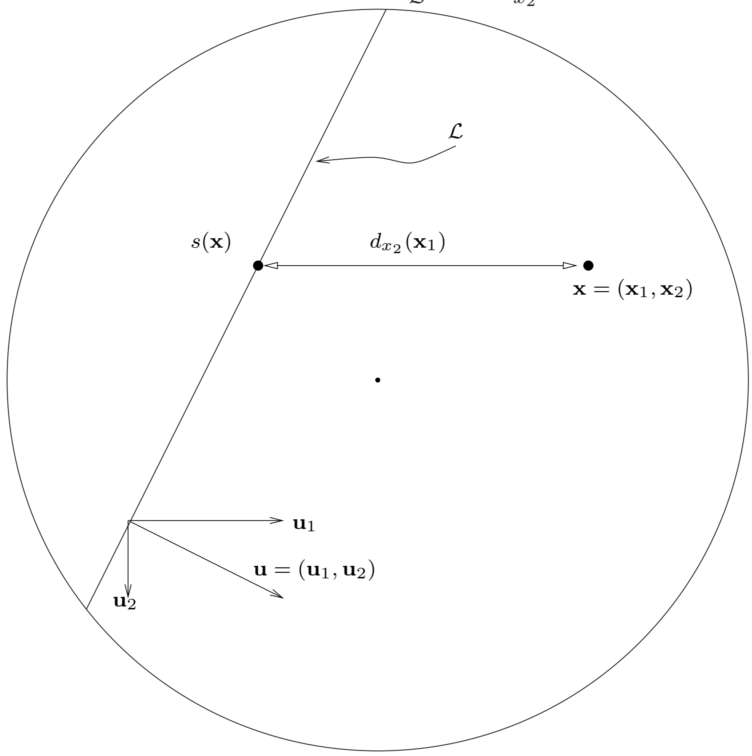
Lemma 9.1. *For $i = 1, 2$, let μ_i be D_i-Federer measures on open $U_i \subset \mathbb{R}^{n_i}$, and consider $\mu = \mu_1 \times \mu_2$ on $U = U_1 \times U_2 \subset \mathbb{R}^n$, $n = n_1 + n_2$. Then:*

(i) *If each μ_i is absolutely (C_i, α_i)-decaying on U_i, then μ is absolutely (C, α)-decaying on U for some $C > 0$ and $\alpha = \min(\alpha_1, \alpha_2)$.*

(ii) *If each μ_i is (C_i, α_i)-decaying on U_i, then μ is (C, α')-decaying on U for some $C > 0$ and $\alpha' = \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2}$.*

Proof. For $\mathbf{y} \in \mathbb{R}^n$ we write $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$ with $\mathbf{y}_i \in \mathbb{R}^{n_i}$. Let $B = B(\mathbf{y}, r)$ be a ball contained in U, where $\mathbf{y} \in \operatorname{supp} \mu$, let $\mathcal{L}$ be an affine hyperplane, and let $\varepsilon > 0$. For both parts of the Lemma, our goal is to find an upper estimate for $\mu(B \cap \mathcal{L}^{(\varepsilon)})$. First let us describe the argument informally. Call a set of the form $\mathbb{R}^{n_1} \times \{\mathbf{x}_2\}$ a 'slice'. By exchanging the factors if necessary, the intersection of $\mathcal{L}^{(\varepsilon)}$ with each slice is a set of the form $\mathcal{L}_{\mathbf{x}_2}^{(c\varepsilon)} \times \{\mathbf{x}_2\}$ where $\mathcal{L}_{\mathbf{x}_2}$ is an appropriate affine hyperplane in $\mathbb{R}^{n_1}$ (depending on the slice) and c is at most $\sqrt{2}$. Thus, if μ_1 is absolutely decaying, the measure of each such intersection is small, and (i) follows. If one assumes only that μ_1 is decaying, then the measures of such intersections will be small provided the slice contains points in the support of μ_1 which are far away from $\mathcal{L}_1$. That this holds for most slices follows from the assumption that μ_2 is decaying and the fact that $\operatorname{supp} \mu = \operatorname{supp} \mu_1 \times \operatorname{supp} \mu_2$.

FIGURE 4. $d_{\mathcal{L}}$ and d_{x_2}



We now proceed to the details. Assume that $B \cap \mathcal{L}^{(\varepsilon)} \neq \varnothing$, otherwise there is nothing to prove. Replacing μ with a translate if necessary, let us assume that $0 \in \mathcal{L}$. Then we can write

$$\mathcal{L} = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{u} \rangle = 0\}, \tag{9.1}$$

where $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2)$ is a unit vector. Assume with no loss of generality that $\|\mathbf{u}_2\| \le \|\mathbf{u}_1\|$, so that

$$\|\mathbf{u}_2\| \le \frac{1}{\sqrt{2}} \le \|\mathbf{u}_1\|. \tag{9.2}$$

For $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$, we define $s(\mathbf{x})$ to be the point in the intersection of $\mathcal{L}$ with the slice $\mathbb{R}^{n_1} \times \{\mathbf{x}_2\}$ closest to $\mathbf{x}$. This point is given by the formula

$$s(\mathbf{x}) \stackrel{\text{def}}{=} (s_1(\mathbf{x}_2), \mathbf{x}_2), \quad \text{where } \; s_1(\mathbf{x}_2) \stackrel{\text{def}}{=} -\frac{\langle \mathbf{x}_2, \mathbf{u}_2 \rangle}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 \in \mathbb{R}^{n_1}, \tag{9.3}$$

i.e. it is the intersection of $\mathcal{L}$ with the line passing through $\mathbf{x}$ in the direction of $\mathbf{u}_1$.

For fixed $\mathbf{x}_2$, let us also denote by $\mathcal{L}_{\mathbf{x}_2}$ the affine hyperplane of $\mathbb{R}^{n_1}$ passing through $s_1(\mathbf{x}_2)$ and orthogonal to $\mathbf{u}_1$, or, equivalently, the affine hyperplane in $\mathbb{R}^{n_1}$ obtained by intersecting $\mathcal{L}$ with the slice $\mathbb{R}^{n_1} + \mathbf{x}_2$. Also let $d_{\mathbf{x}_2} : \mathbb{R}^{n_1} \to \mathbb{R}_+$ stand for the distance from $\mathcal{L}_{\mathbf{x}_2}$ in $\mathbb{R}^{n_1}$. See Figure 4.

Clearly one has

$$d_{\mathbf{x}_2}(\mathbf{x}_1) = \frac{1}{\|\mathbf{u}_1\|} d_{\mathcal{L}}(\mathbf{x}) = |\langle \mathbf{x}_1 - s_1(\mathbf{x}_2), \bar{\mathbf{u}}_1 \rangle|\,, \quad \text{where } \bar{\mathbf{u}}_1 \stackrel{\text{def}}{=} \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}\,.$$

Since $s(\mathbf{x}) \in \mathcal{L}$, one has

$$\begin{aligned} B \cap \mathcal{L}^{(\varepsilon)} &= \{\mathbf{x} \in B : |\langle \mathbf{x} - s(\mathbf{x}), \mathbf{u}\rangle| < \varepsilon\} \\ &= \{(\mathbf{x}_1, \mathbf{x}_2) \in B : |\langle \mathbf{x}_1 - s_1(\mathbf{x}_2), \mathbf{u}_1\rangle| < \varepsilon\} \\ &\subset X \stackrel{\text{def}}{=} \{(\mathbf{x}_1, \mathbf{x}_2) \in B_1 \times B_2 : d_{\mathbf{x}_2}(\mathbf{x}_1) < \sqrt{2}\varepsilon\}, \end{aligned} \tag{9.4}$$

where $B_i = B(\mathbf{y}_i, r) \subset U_i$ is the corresponding ball in $\mathbb{R}^{n_i}$.

Applying the absolute decay property to μ_1, we obtain that for any $\mathbf{x}_2$,

$$\frac{\mu_1\left(\{\mathbf{x}_1 \in B_1 : d_{\mathbf{x}_2}(\mathbf{x}_1) < \sqrt{2}\varepsilon\}\right)}{\mu_1(B_1)} \le C_1 \left(\frac{\sqrt{2}\varepsilon}{r}\right)^{\alpha_1} \le C_3 \left(\frac{\varepsilon}{r}\right)^{\alpha}, \tag{9.5}$$

where $C_3 = 2^{\alpha/2} C_1$.

We can now estimate $\mu(X)$ by disintegrating into slices parallel to $\mathbb{R}^{n_1}$ to obtain:

$$\mu(X) \le C_3 \left(\frac{\varepsilon}{r}\right)^{\alpha} \mu_1(B_1)\mu_2(B_2) = C_3 \left(\frac{\varepsilon}{r}\right)^{\alpha} \mu(B_1 \times B_2).$$

Since $B_1 \times B_2 \subset B(\mathbf{y}, \sqrt{2}r)$, we can use Federer to find a constant C such that

$$\mu(B \cap \mathcal{L}^{(\varepsilon)}) \le \mu(X) \le C \left(\frac{\varepsilon}{r}\right)^{\alpha} \mu(B),$$

whence (2.5).

To prove (ii), note that by (9.1) we have $d_{\mathcal{L}}(\mathbf{x}) = |\langle \mathbf{x}, \mathbf{u}\rangle|$. Let $\mathbf{x}' = (\mathbf{x}'_1, \mathbf{x}'_2) \in B \cap \operatorname{supp} \mu$ such that

$$|\langle \mathbf{x}', \mathbf{u}\rangle| = d_{\mathcal{L}}(\mathbf{x}') > \|d_{\mathcal{L}}\|_{\mu,B}/2\,; \tag{9.6}$$

since $\mathbf{x}'_1 \in \operatorname{supp} \mu_1 \cap B_1$, by the decaying property for μ_1 we have

$$\frac{\mu_1\left(\{\mathbf{x}_1 \in B_1 : |d_{\mathbf{x}_2}(\mathbf{x}_1)| < \sqrt{2}\varepsilon\}\right)}{\mu_1(B_1)} \le C_1 \left(\frac{\varepsilon}{\|d_{\mathbf{x}_2}\|_{\mu_1,B_1}}\right)^{\alpha_1} \le C_1 \left(\frac{\varepsilon}{|d_{\mathbf{x}_2}(\mathbf{x}'_1)|}\right)^{\alpha_1}. \tag{9.7}$$

On the other hand, let us consider the function $D : \mathbb{R}^{n_2} \to \mathbb{R}$ defined by

$$D(\mathbf{x}_2) \stackrel{\text{def}}{=} d_{\mathbf{x}_2}(\mathbf{x}'_1) = \frac{1}{\|\mathbf{u}_1\|} d_{\mathcal{L}}(\mathbf{x}'_1, \mathbf{x}_2)\,.$$

By (9.6) and since $\mathbf{x}'_2 \in \operatorname{supp} \mu_2 \cap B_2$, one has

$$\|D\|_{\mu_2,B_2} \ge D(\mathbf{x}'_2) = \frac{1}{\|\mathbf{u}_1\|} d_{\mathcal{L}}(\mathbf{x}') \ge d_{\mathcal{L}}(\mathbf{x}') > \|d_{\mathcal{L}}\|_{\mu,B}/2\,.$$

Thus, by the decaying property of μ_2, for every $\eta > 0$,

$$\frac{\mu_2\left(\{\mathbf{x}_2 \in B_2 : D(\mathbf{x}_2) < \eta\}\right)}{\mu(B_2)} \le C_4 \left(\frac{\eta}{\|d_{\mathcal{L}}\|_{\mu,B}}\right)^{\alpha_2}, \tag{9.8}$$

where $C_4 \stackrel{\text{def}}{=} 2^{\alpha_2} C_2$. Set

$$\eta \stackrel{\text{def}}{=} \varepsilon^{\frac{\alpha_1}{\alpha_1+\alpha_2}} \left(\|d_{\mathcal{L}}\|_{\mu,B}\right)^{\frac{\alpha_2}{\alpha_1+\alpha_2}}\,.$$

We have

$$X \subset \{(\mathbf{x}_1, \mathbf{x}_2) \in X : d_{\mathbf{x}_2}(\mathbf{x}'_1) \ge \eta\} \cup \{(\mathbf{x}_1, \mathbf{x}_2) \in B_1 \times B_2 : D(\mathbf{x}_2) < \eta\}\,.$$

Putting together (9.7) and (9.8) we obtain:

$$\begin{aligned}\mu\left(B \cap \mathcal{L}^{(\varepsilon)}\right) \le \mu(X) &\le \left[C_3\left(\frac{\varepsilon}{\eta}\right)^{\alpha_1} + C_4\left(\frac{\eta}{\|d_{\mathcal{L}}\|_{\mu,B}}\right)^{\alpha_2}\right]\mu_1(B_1)\mu_2(B_2)\\ &\le C''\left(\frac{\varepsilon}{\|d_{\mathcal{L}}\|_{\mu,B}}\right)^{\alpha'}\mu(B_1\times B_2)\\ &\le C'\left(\frac{\varepsilon}{\|d_{\mathcal{L}}\|_{\mu,B}}\right)^{\alpha'}\mu(B),\end{aligned}$$

proving (ii). □

Proof of Theorem 2.3. By induction, it suffices to consider the case $k = 2$, and it is easy to verify that the product of two nonplanar (resp. Federer) measures is also nonplanar (resp. Federer). The rest follows from Lemma 9.1. □

10. Concluding remarks and open questions

There are many intriguing questions about diophantine properties of measures; concerning some of them we have partial results or ideas how to proceed, about others we know very little.

10.1. **Khintchine's theorem.** For a decreasing function $\psi : \mathbb{Z}_+ \to \mathbb{R}$, say that $\mathbf{y} \in \mathbb{R}^n$ is *ψ-approximable* if there are infinitely many $\mathbf{p} \in \mathbb{Z}^n$, $q \in \mathbb{Z}_+$ satisfying

$$\|q\mathbf{y} - \mathbf{p}\| < \psi(q)\,.$$

A classical theorem of Khintchine (see [Kh1] or [C, Chapter VII]) characterizes those ψ for which Lebesgue-a.e. $\mathbf{y} \in \mathbb{R}^n$ is ψ-approximable, the condition being that

$$\sum_{q=1}^{\infty}\psi(q)^n = \infty\,.$$

If this criterion fails, then Lebesgue a.e. $\mathbf{y} \in \mathbb{R}^n$ is not ψ-approximable. For example, by taking $\psi(q) = q^{-(\frac{1}{n}+\delta)}$ one gets that Lebesgue-a.e. $\mathbf{y} \in \mathbb{R}^n$ is not very well approximable.

Question 10.1. *Let $\psi : \mathbb{Z}_+ \to \mathbb{R}$ be a decreasing function, and let μ be a friendly measure on $\mathbb{R}^n$.*

1. *Is it always true that either the set of ψ-approximable points or its complement has measure 0?*
2. *Is there an explicit necessary and sufficient criterion so that μ-a.e. point is ψ-approximable?*

For example, this condition might be phrased in terms of a series involving ψ and μ.

Note that even for volume measures on proper nondegenerate submanifolds this problem is still open (see [BD, Chapter 2] for history and references), although there exist definite results [BKM, BBKM] for a dual (linear form) setting. Some partial results for fractal measures are contained in [W1], [W2]. For example, it is proved there that μ-a.e. $y \in \mathbb{R}$ is not ψ-approximable when μ is the Cantor measure on $\mathbb{R}$ and $\sum_q \psi(q)^\alpha q^{\alpha-1} < \infty$, $\alpha = \frac{\log 2}{\log 3}$. However it is not known whether μ-a.e. y is ψ-approximable in case the above series diverges.

10.2. **Badly approximable vectors.** Our next question is the analogue of results of V. Jarnik and W. Schmidt. Say that $\mathbf{y} \in \mathbb{R}^n$ is *badly approximable* if there exists $\varepsilon > 0$ such that for every $\mathbf{p} \in \mathbb{Z}^n$, $q \in \mathbb{Z}_+$ one has

$$\|q\mathbf{y} - \mathbf{p}\| \geq \frac{\varepsilon}{q^{1/n}}\,.$$

It was proved by Schmidt [Sch1] (and earlier by Jarnik [Ja] for $n = 1$) that badly approximable vectors form a set of full Hausdorff dimension in $\mathbb{R}^n$. One can ask whether a similar statement is true for the support of a friendly measure. Even the following more modest question is completely open:

Question 10.2. *Is there at least one badly approximable vector in the support of every friendly measure?*

Note that even for volume measures on nondegenerate manifolds it seems that nothing is known about Question 10.2.

In a forthcoming paper [KW] the first and third named authors show:

Theorem 10.3. *Let K be the attractor of a finite irreducible family of contracting similarity self-maps of $\mathbb{R}^n$ satisfying the open set condition. Then the Hausdorff dimension of the intersection of K with the set of badly approximable vectors is the same as the Hausdorff dimension of K.*

10.3. **Singular vectors.** Recall that $\mathbf{y} \in \mathbb{R}^n$ is called *singular* if for every $\varepsilon > 0$ there exists Q_0 such that

$$\forall\, Q \geq Q_0 \ \exists\, \mathbf{p} \in \mathbb{Z}^n \text{ and } q \in \mathbb{Z}_+ \text{ with } q \leq Q \text{ and } \|q\mathbf{y} - \mathbf{p}\| \leq \frac{\varepsilon}{Q^{1/n}}\,.$$

It is easy to see that Lebesgue-a.e. $\mathbf{y} \in \mathbb{R}^n$ is not singular, and that $y \in \mathbb{R}$ is singular if and only if $y \in \mathbb{Q}$; however for $n > 1$ one can construct plenty of nontrivial singular vectors, see [Kh2] or [C, Chapter V].

Note that it was observed by Dani [Da1] that $\mathbf{y}$ is singular if and only if a certain one-parameter trajectory of $\bar{\tau}(\mathbf{y}) \in G/\Gamma$ is divergent. This makes it possible to deduce the following result from Theorem 4.3 (details to appear in [KW]):

Theorem 10.4. *If μ is a friendly measure on $\mathbb{R}^n$, then μ-a.e. $\mathbf{y}$ is not singular.*

10.4. **Other natural measures and strong extremality.** We suspect that many more natural examples of nonplanar measures are strongly extremal. In this spirit, we propose the following two concrete conjectures:

Conjecture 10.5. *Suppose $\{\mathbf{h}_i : \mathbb{R}^n \to \mathbb{R}^n,\ i = 1, \dots, m\}$ is an irreducible system of real analytic contractions (that is, there is $\varrho < 1$ such that for every i and every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\|\mathbf{h}_i(\mathbf{x}) - \mathbf{h}_i(\mathbf{y})\| \leq \varrho\|\mathbf{x} - \mathbf{y}\|$), and suppose that $p_1, \dots, p_m$ are positive with $\sum p_i = 1$. Let μ be a measure on $\mathbb{R}^n$ satisfying*

$$\mu = \sum_i p_i (\mathbf{h}_i)_* \mu.$$

Then μ is strongly extremal.

The following special case is already interesting:

Conjecture 10.6. *Suppose* $\{\mathbf{h}_i : \mathbb{R}^n \to \mathbb{R}^n,\ i = 1, \dots, m\}$ *is an irreducible system of conformal contractions satisfying the open set condition, let* K *be their attractor, let* s *be the Hausdorff dimension of* K*, and let* μ *be the restriction of* H^s *to* K*. Then* μ *is strongly extremal.*

10.5. **Not-so-friendly measures.** A measure μ is friendly if it is Federer, nonplanar, and decaying. It is interesting to ask to what extent these conditions can be relaxed without sacrificing strong extremality of μ. One direction is to take the non-uniform version of the above conditions as in §6.

By considering Lebesgue measures on proper rational affine subspaces of $\mathbb{R}^n$ it is clear that in general the nonplanarity condition cannot be dropped. On the other hand, given an affine subspace $\mathcal{L}$ of $\mathbb{R}^n$, the paper [K2] exhibits necessary and sufficient conditions, involving coefficients of affine functions parametrizing $\mathcal{L}$, for the (strong) extremality of $\mathcal{L}$. In fact, it is proved there that $\mathcal{L}$ is (strongly) extremal if and only if it contains at least one not very well (multiplicatively) approximable point. It is also proved in that paper that (strong) extremality of $\mathcal{L}$ implies the same for any smooth submanifold of $\mathcal{L}$ which is nondegenerate in $\mathcal{L}$.

It seems plausible that decaying and Federer measures supported on proper affine subspaces of $\mathbb{R}^n$ exhibit similar behavior. Namely, we believe that it is possible to show the following: *suppose* $\mathcal{L}$ *is an extremal (resp., strongly extremal) affine subspace of* $\mathbb{R}^n$*, and* μ *is a Federer and decaying measure on* $\mathcal{L}$ *with the property that* $\mu(\mathcal{L}') = 0$ *for any proper affine subspace* $\mathcal{L}'$ *of* $\mathcal{L}$*; then* μ *is also extremal (resp., strongly extremal).*

We do not know whether the Federer assumption can be lifted in general. If μ is assumed to be absolutely decaying and nonplanar, we believe that it is possible to deduce that μ is strongly extremal without any additional assumptions, by proving a variant of Theorem 4.3. In the statement of this modified theorem, the assumption that μ is Federer is replaced by an additional assumption on the function h, namely that there exists C such that for every $V \in \mathcal{W}$, μ-a.e. $x \in X$ and every $r > 0$,

$$\frac{\|\ell_V \circ h\|_{\mu, B(x,3r)}}{\|\ell_V \circ h\|_{\mu, B(x,r)}} \le C\,.$$

References

[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**, no. 2 (2002), 203–225.

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

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

[C] J.W.S. Cassels, **An introduction to Diophantine approximation**, Cambridge University Press, New York, 1957.

[Da1] S.G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*, J. Reine Angew. Math., **359** (1985), 55–89.

[Da2] S.G. Dani, *On orbits of unipotent flows on homogeneous spaces. II*, Ergodic Theory Dynam. Systems, **6** (1986), 167–182.

[Ja] V. Jarnik, *Über die simulatanen diophantischen Approximationen*, Math. Zeit. **33** (1931), 505–543.

[H] J.E. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. **30** (1981), no. 5, 713–747.

[K1] D. Kleinbock, *Some applications of homogeneous dynamics to number theory*, in: **Smooth ergodic theory and its applications** (Seattle, WA, 1999), 639–660, Proc. Sympos. Pure Math., 69, Amer. Math. Soc., Providence, RI, 2001.

[K2] D. Kleinbock, *Extremal subspaces and their submanifolds*, to appear in Geom. Funct. Anal.

[Kh1] A. Khintchine, *Zur metrischen Theorie der diophantischen Approximationen*, Math. Zeit. **24** (1926), 706–714.

[Kh2] A. Khintchine, *Uber eine Klasse linearer Diophantischer Approximationen*, Rend. Circ. Mat. Palermo **50** (1926), 175–195.

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

[KSS] D. Kleinbock, N. Shah, and 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.

[KW] D. Kleinbock and B. Weiss, in preparation.

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

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

[MU] D. Mauldin and M. Urbanski, *The doubling property of conformal measures of infinite iterated function systems*, (2001), preprint.

[R] M. S. Raghunathan, **Discrete subgroups of Lie groups**, Springer, New York, 1972.

[Sa] H. Sato, *Global density theorem for a Federer measure*, Tohoku Math. J. (2) **44** (1992), no. 4, 581–595.

[Sch1] W.M. Schmidt, *Badly approximable systems of linear forms*, J. Number Theory **1** (1969), 139–154.

[Sch2] W.M. Schmidt, *Diophantine approximation and certain sequences of lattices*, Acta Arith. **18** (1971), 165–178.

[Sp1] V.G. Sprindžuk, **Mahler's problem in metric number theory**, Translated from the Russian by B. Volkmann. Translations of Mathematical Monographs, Amer. Math. Soc., Providence, R.I., 1969.

[Sp2] V.G. Sprindžuk, *Achievements and problems in Diophantine approximation theory,* Russian Math. Surveys **35** (1980), 1–80.

[V] W.A. Veech, *Measures supported on the set of uniquely ergodic directions of an arbitrary holomorphic 1-form*, Ergodic Theory Dynam. Systems **19** (1999), 1093–1109.

[W1] B. Weiss, *Almost no points on a Cantor set are very well approximable*, Proc. R. Soc. Lond. **457** (2001), 949–952.

[W2] B. Weiss, *Dynamics on parameter spaces: submanifold and fractal subset questions*, in: **Rigidity in Dynamics and Geometry** M. Burger and A. Iozzi (eds.), Springer, 2002.

Brandeis University, Waltham MA 02454-9110 `kleinboc@brandeis.edu`

Stanford University, Stanford CA 94305 `elonl@math.stanford.edu`

Ben Gurion University, Be'er Sheva, Israel 84105 `barakw@math.bgu.ac.il`