

# METRIC DIOPHANTINE APPROXIMATION FOR SYSTEMS OF LINEAR FORMS VIA DYNAMICS

DMITRY KLEINBOCK, GREGORY MARGULIS, AND JUNBO WANG

ABSTRACT. The goal of this paper is to generalize the main results of [KM1] and subsequent papers on metric Diophantine approximation with dependent quantities to the set-up of systems of linear forms. In particular, we establish ‘joint strong extremality’ of arbitrary finite collection of smooth nondegenerate submanifolds of  $\mathbb{R}^n$ . The proofs are based on generalized quantitative nondivergence estimates for translates of measures on the space of lattices.

## 1. INTRODUCTION

The theory of simultaneous Diophantine approximation is concerned with the following question: if  $Y$  is an  $m \times n$  real matrix (interpreted as a system of  $m$  linear forms in  $n$  variables), how small, in terms of the size of  $\mathbf{q} \in \mathbb{Z}^n$ , can be the distance from  $Y\mathbf{q}$  to  $\mathbb{Z}^m$ . This generalizes the classical theory of approximation of real numbers by rationals, where  $m = n = 1$ .

In the case of a single linear form ( $m = 1$ ), or, dually, a single vector ( $n = 1$ ), significant progress has been made during recent years in showing that some important approximation properties of vectors/forms happen to be generic with respect to certain measures other than Lebesgue measure. This circle of problems dates back to the 1930s, namely, to Mahler’s work on transcendental numbers. In order to describe more precisely Mahler’s original problem, as well as subsequent results and conjectures, let us introduce some standard notions from the theory of Diophantine approximation.

Denote by  $M_{m,n}$  the space of real matrices with  $m$  rows and  $n$  columns. It follows from Dirichlet’s Theorem on simultaneous approximation that for any  $Y \in M_{m,n}$  there are infinitely many  $\mathbf{q} \in \mathbb{Z}^n$  such that  $\|Y\mathbf{q} - \mathbf{p}\| < \|\mathbf{q}\|^{-n/m}$  for some  $\mathbf{p} \in \mathbb{Z}^m$  (here  $\|\cdot\|$  is given by  $\|\mathbf{x}\| = \max_i |x_i|$ .) On the other hand, if  $\delta > 0$ , the set of  $Y \in M_{m,n}$  such that there exist infinitely many  $\mathbf{q} \in \mathbb{Z}^n$  with

$$\|Y\mathbf{q} - \mathbf{p}\| < \|\mathbf{q}\|^{-n/m-\delta} \text{ for some } \mathbf{p} \in \mathbb{Z}^m \quad (1.1)$$

is null with respect to Lebesgue measure  $\lambda$ . One says that  $Y$  is *very well approximable* (abbreviated by VWA) if (1.1) holds for some positive  $\delta$  and infinitely many  $\mathbf{q} \in \mathbb{Z}^n$ . It follows that the set of VWA matrices has zero Lebesgue measure. However its Hausdorff dimension is equal to the dimension

---

*Date:* July 2009.

*1991 Mathematics Subject Classification.* 11J13; 37A17.

*Key words and phrases.* Simultaneous Diophantine approximation, strong extremality, homogeneous flows, quantitative non-divergence.

of  $M_{m,n}$  [Do], so in this sense this set is rather big. Note also that by Khintchine's Transference Principle, see e.g. [C, Chapter V],  $Y$  is VWA iff so is the transpose of  $Y$ .

Let us now turn to a conjecture made by Mahler [M] in 1932 and proved three decades later by Sprindžuk, see [Sp1, Sp2]. It states that for  $\lambda$ -almost every  $x \in \mathbb{R}$ , the row vector  $\mathbf{f}(x) = (x, x^2, \dots, x^n)$  is not VWA. Sprindžuk's proof of the above conjecture has led to the development of a new branch of number theory, the so-called 'Diophantine approximation with dependent quantities'. One of the goals of the theory has been showing that certain smooth maps  $\mathbf{f}$  from open subsets of  $\mathbb{R}^d$  to  $\mathbb{R}^n$  are, in the terminology introduced by Sprindžuk, *extremal*, that is, vectors  $\mathbf{f}(\mathbf{x})$  are not VWA for  $\lambda$ -a.e.  $\mathbf{x}$  (the reader is referred to [BD] for history and references). Thus it seems natural to propose the following general problem: exhibit sufficient conditions on a measure  $\mu$  on  $M_{m,n}$  (for example of the form  $F_*\lambda$  where  $F$  is a smooth map from an open subset of  $\mathbb{R}^d$  to  $M_{m,n}$ ) guaranteeing that  $\mu$  is *extremal*, which by definition means that  $\mu$ -a.e.  $Y \in M_{m,n}$  is not VWA. When  $\mu = F_*\lambda$  for  $F: \mathbb{R}^d \rightarrow M_{m,n}$ , one can interpret this problem as studying  $m$  maps  $\mathbb{R}^d \rightarrow \mathbb{R}^n$  (rows of  $F$ ) simultaneously. Some special cases were done by Kovalevskaya in the 1980s, who used the terminology 'jointly extremal' for the rows (or columns) of  $F$  for which  $F_*\lambda$  is extremal.

The present paper, among other things, suggests possible solutions to this problem. In fact this will be done in a stronger, multiplicative way. For  $\mathbf{x} = (x_i)$  we let

$$\Pi(\mathbf{x}) \stackrel{\text{def}}{=} \prod_i |x_i| \quad \text{and} \quad \Pi_+(\mathbf{x}) \stackrel{\text{def}}{=} \prod_i \max(|x_i|, 1).$$

Then say that  $Y \in M_{m,n}$  is *very well multiplicatively approximable* (VWMA) if for some  $\delta > 0$  there are infinitely many  $\mathbf{q} \in \mathbb{Z}^n$  such that

$$\Pi(Y\mathbf{q} - \mathbf{p}) < \Pi_+(\mathbf{q})^{-(1+\delta)} \tag{1.2}$$

for some  $\mathbf{p} \in \mathbb{Z}^m$ . Since  $\Pi(Y\mathbf{q} - \mathbf{p})$  is always not greater than  $\|Y\mathbf{q} - \mathbf{p}\|^m$  and  $\Pi_+(\mathbf{q}) \leq \|\mathbf{q}\|^n$  for  $\mathbf{q} \in \mathbb{Z}^n \setminus \{0\}$ , VWA implies VWMA. Still it can be easily shown that Lebesgue-a.e.  $Y$  is not VWMA<sup>1</sup>. Therefore one can ask for stronger sufficient conditions on a measure  $\mu$  on  $M_{m,n}$  guaranteeing that it is *strongly extremal*, that is,  $\mu$ -a.e.  $Y \in M_{m,n}$  is not VWMA.

An approach to this class of problems based on homogeneous dynamics was developed in the paper [KM1], which dealt with the case  $m = 1$ . The problem of extending that approach to the matrix set-up was raised in [KM1, §6.2] and then in [Go, §9.1]. To state the main result of [KM1], which verified a conjecture made by Sprindžuk in [Sp3], let us recall the following definitions. A smooth map  $\mathbf{f}$  from  $U \subset \mathbb{R}^d$  to  $\mathbb{R}^n$  is called  *$\ell$ -nondegenerate at  $\mathbf{x} \in U$*  if partial derivatives of  $\mathbf{f}$  at  $\mathbf{x}$  up to order  $\ell$  span  $\mathbb{R}^n$ . We will say that  $\mathbf{f}$  is *nondegenerate at  $\mathbf{x}$*  if it is  $\ell$ -nondegenerate at  $\mathbf{x}$  for some  $\ell$ , and that it is *nondegenerate* if it is nondegenerate at  $\lambda$ -a.e.  $\mathbf{x} \in U$ . Here is the statement of [KM1, Theorem A]:

<sup>1</sup>Also, generalizing Khintchine's Transference Principle one can show that  $Y$  is VWMA iff so is the transpose of  $Y$ , see a remark at the end of §3.

**Theorem 1.1.** *Let  $\mathbf{f}$  be a smooth nondegenerate map from an open subset  $U$  of  $\mathbb{R}^d$  to  $\mathbb{R}^n$ . Then  $\mathbf{f}_*\lambda$  is strongly extremal.*

The goal of this paper is to describe a fairly large class of strongly extremal measures on  $M_{m,n}$ . Here is an important special case of our general results:

**Theorem 1.2.** *For every  $i = 1, \dots, m$ , let  $\mathbf{f}_i$  be a nondegenerate map from an open subset  $U_i$  of  $\mathbb{R}^{d_i}$  to  $\mathbb{R}^n$ , and let*

$$F : U_1 \times \dots \times U_m \rightarrow M_{m,n}, \quad (\mathbf{x}_1, \dots, \mathbf{x}_m) \mapsto \begin{pmatrix} \mathbf{f}_1(\mathbf{x}_1) \\ \vdots \\ \mathbf{f}_m(\mathbf{x}_m) \end{pmatrix}. \quad (1.3)$$

*Then the pushforward of Lebesgue measure on  $U_1 \times \dots \times U_m$  by  $F$  is strongly extremal.*

The case  $d_1 = \dots = d_m = 1$ , i.e. that of  $n$  nondegenerate curves in  $\mathbb{R}^m$ , had been previously studied by Kovalevskaya [Ko1, Ko2, Ko3]. A special case of the above theorem where  $U_1 = \dots = U_m$  and  $\mathbf{f}_1 = \dots = \mathbf{f}_m$  is also of interest: it describes approximation properties of generic  $m$ -tuples of points (viewed as row vectors, or linear forms) on a given nondegenerate manifold. In this form the above statement had been conjectured earlier by Bernik (private communication). We remark that recently V. Beresnevich informed us of an alternative approach allowing to prove Theorem 1.2 when  $\mathbf{f}_1, \dots, \mathbf{f}_m$  are real analytic.

The structure of the paper is as follows. In §2 we introduce the terminology needed to state our general result (Theorem 2.1) of which Theorem 1.2 is a special case. In §3 we discuss a dynamical approach to Diophantine approximation problems and describe Diophantine properties introduced above in the language of flows on the space of lattices. Then in §4 and §5 we present the main ‘quantitative nondivergence’ measure estimate and use it to state and prove a more precise version of Theorem 2.1. §6 is devoted to proving Proposition 2.2, which explains why Theorem 1.2 follows from Theorem 2.1. Then the results of §4 are used in §7 for construction of examples of extremal and strongly extremal measures not covered by Theorem 2.1. Finally in the last section we mention several additional results and further open questions.

**Acknowledgements.** Part of the work was done during the collaboration of D.K. and G.M. at ETH-Zurich, the University of Bielefeld and Yale University; the hospitality of these institutions is gratefully acknowledged. This research was supported in part by NSF grants DMS-0239463, DMS-0244406, DMS-0801064 and DMS-0801195. Several results of this paper were part of Ph.D. Thesis of the third named author [W] defended at Brandeis in 2008. The authors are grateful to Victor Beresnevich and the reviewer for useful remarks.

## 2. THE MAIN THEOREM

We now introduce some terminology needed to state a more general version of Theorem 1.2. Let  $X$  be a metric space. If  $x \in X$  and  $r > 0$ , we denote by  $B(x, r)$  the open ball of radius  $r$  centered at  $x$ . If  $B = B(x, r)$  and  $c > 0$ ,  $cB$

will denote the ball  $B(x, cr)$ . For  $B \subset X$  and a real-valued function  $f$  on  $B$ , let

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

If  $\nu$  is a measure on  $X$  such that  $\nu(B) > 0$ , define  $\|f\|_{\nu, B} \stackrel{\text{def}}{=} \|f\|_{B \cap \text{supp } \nu}$ . All measures on metric spaces will be assumed to be Radon.

If  $D > 0$  and  $U \subset X$  is an open subset, let us say that a measure  $\nu$  on  $X$  is  $D$ -Federer on  $U$  if one has  $\nu(\frac{1}{3}B) > \nu(B)/D$  for any ball  $B \subset U$  centered at  $\text{supp } \nu$ . This condition is often called ‘doubling’ in the literature; see [KLW, MU] for examples and references. A measure  $\nu$  will be called *Federer* if for  $\nu$ -a.e.  $x \in X$  there exist a neighborhood  $U$  of  $x$  and  $D > 0$  such that  $\nu$  is  $D$ -Federer on  $U$ .

Given  $C, \alpha > 0$  and open  $U \subset X$ , say that  $f : U \rightarrow \mathbb{R}$  is  $(C, \alpha)$ -good on  $U$  with respect to a measure  $\nu$  if for any ball  $B \subset U$  centered in  $\text{supp } \nu$  and any  $\varepsilon > 0$  one has

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

This condition was formally introduced in [KM1] for  $\nu$  being Lebesgue measure on  $\mathbb{R}^d$ , and in [KLW] for arbitrary  $\nu$ . If  $\mathbf{f} = (f_1, \dots, f_N)$  is a map from  $U$  to  $\mathbb{R}^N$ , following [K2], we will say that a pair  $(\mathbf{f}, \nu)$  is *good* if for  $\nu$ -a.e.  $x$  there exists a neighborhood  $V$  of  $x$  such that any linear combination of  $1, f_1, \dots, f_N$  is  $(C, \alpha)$ -good on  $V$  with respect to  $\nu$ .

Here is another useful definition:  $(\mathbf{f}, \nu)$  is said to be *nonplanar* if for any ball  $B$  with  $\nu(B) > 0$ , the restrictions of  $1, f_1, \dots, f_N$  to  $B \cap \text{supp } \nu$  are linearly independent over  $\mathbb{R}$ ; in other words,  $\mathbf{f}(B \cap \text{supp } \nu)$  is not contained in any proper affine subspace of  $\mathbb{R}^N$ .

Important examples of good and nonplanar pairs  $(\mathbf{f}, \nu)$  are  $\nu = \lambda$  (Lebesgue measure on  $\mathbb{R}^d$ ) and  $\mathbf{f}$  smooth and nondegenerate. In this case the fact that  $(\mathbf{f}, \lambda)$  is good follows from [KM1, Proposition 3.4], and nonplanarity is immediate. In [KLW] a class of *friendly* measures was introduced: a measure  $\nu$  on  $\mathbb{R}^n$  is friendly if and only if it is Federer and the pair  $(\text{Id}, \nu)$  is good and nonplanar; many examples of those can be found in [KLW, U, SU]. In the paper [KLW] the approach to metric Diophantine approximation developed in [KM1] has been extended to maps and measures satisfying the conditions described above. One of its main results is the following theorem [K3, Theorem 4.2], implicitly contained in [KLW]: let  $\nu$  be a Federer measure on  $\mathbb{R}^d$ ,  $U \subset \mathbb{R}^d$  open, and  $\mathbf{f} : U \rightarrow \mathbb{R}^n$  a continuous map such that  $(\mathbf{f}, \nu)$  is good and nonplanar; then  $\mathbf{f}_* \nu$  is strongly extremal.

Our goal in this paper is to replace  $\mathbb{R}^n$  with  $M_{m,n}$  in the above statements. For this, given  $Y = (y_{i,j}) \in M_{m,n}$  and subsets  $I = \{i_1, \dots, i_r\} \subset \{1, \dots, m\}$  and  $J = \{j_1, \dots, j_s\} \subset \{1, \dots, n\}$  of equal cardinality and with  $i_1 < \dots < i_r$  and  $j_1 < \dots < j_s$ , we define

$$y_{I,J} \stackrel{\text{def}}{=} \begin{vmatrix} y_{i_1, j_1} & \cdots & y_{i_1, j_s} \\ \cdots & \cdots & \cdots \\ y_{i_r, j_1} & \cdots & y_{i_r, j_s} \end{vmatrix}, \text{ with the convention } y_{\emptyset, \emptyset} = 1. \quad (2.1)$$

Denote by

$$N \stackrel{\text{def}}{=} \binom{m+n}{m} - 1 \tag{2.2}$$

the number of different square submatrices of an  $m \times n$  matrix, and consider the map  $\mathbf{d} : M_{m,n} \rightarrow \mathbb{R}^N$  given by

$$\mathbf{d}(Y) \stackrel{\text{def}}{=} (y_{I,J})_{I \subset \{1, \dots, m\}, J \subset \{1, \dots, n\}, 0 < |I|=|J| \leq \min(m,n)}.$$

In other words,  $\mathbf{d}(Y)$  is a vector whose coordinates are determinants of all possible square submatrices of  $Y$  (the order in which they appear does not matter).

At this point we can state the main result of the paper:

**Theorem 2.1.** *Let  $\nu$  be a Federer measure on  $\mathbb{R}^d$ ,  $U \subset \mathbb{R}^d$  open, and  $F : U \rightarrow M_{m,n}$  a continuous map such that  $(\mathbf{d} \circ F, \nu)$  is (i) good and (ii) nonplanar. Then  $F_*\nu$  is strongly extremal.*

Note that if  $\min(m,n) = 1$ ,  $\mathbf{d} \circ F$  coincides with  $F$ , and  $N$  is equal to  $\max(m,n)$ ; thus [K3, Theorem 4.2] cited above is a special case of Theorem 2.1. If  $\min(m,n) > 1$ , the assumptions (i) and (ii) above can be verified for a wide variety of examples. For instance, when a map  $F : U \rightarrow M_{m,n}$  is real analytic and  $\nu$  is Lebesgue measure, assumption (i) of both theorems is satisfied (this follows from the results of [KM1] and [K1]). And if  $F$  is differentiable and  $\nu = \lambda$ , both (i) and (ii) would follow from an assumption that the map  $\mathbf{d} \circ F : \mathbb{R}^d \rightarrow \mathbb{R}^N$  is nondegenerate. We explain this in more detail in §6, where we also prove

**Proposition 2.2.** *Let  $F : U_1 \times \dots \times U_m \rightarrow \mathbb{R}^n$  be as in Theorem 1.2. Then the pair  $(\mathbf{d} \circ F, \lambda)$  is good and nonplanar.*

In view of the above proposition and since Lebesgue measure is Federer, Theorem 1.2 follows from Theorem 2.1.

We remark that the assumptions (i) and (ii) of Theorem 2.1 are not the most general possible; in particular, assuming (i) one can establish necessary and sufficient conditions for the extremality and strong extremality of  $F_*\nu$ , see Theorem 4.3. Some examples of extremal and strongly extremal measures not covered by Theorem 2.1 are discussed in §7.

### 3. DIOPHANTINE APPROXIMATION AND FLOWS ON HOMOGENEOUS SPACES

From now on we will let  $k = m + n$  and put  $G = \text{SL}_k(\mathbb{R})$ ,  $\Gamma = \text{SL}_k(\mathbb{Z})$  and  $\Omega = G/\Gamma$ . Note that  $\Omega$  is naturally identified with the space of unimodular lattices in  $\mathbb{R}^k$  via the correspondence  $g\Gamma \mapsto g\mathbb{Z}^k$ . Define

$$u_Y \stackrel{\text{def}}{=} \begin{pmatrix} I_m & Y \\ 0 & I_n \end{pmatrix}, \quad \Lambda_Y \stackrel{\text{def}}{=} u_Y \mathbb{Z}^k,$$

where  $I_\ell$  stands for the  $\ell \times \ell$  identity matrix. To highlight the relevance of the objects defined above to the Diophantine problems considered in the

introduction, note that

$$\Lambda_Y = \left\{ \begin{pmatrix} Y\mathbf{q} - \mathbf{p} \\ \mathbf{q} \end{pmatrix} : \mathbf{p} \in \mathbb{Z}^m, \mathbf{q} \in \mathbb{Z}^n \right\}.$$

The main theme of this section is a well known restatement of Diophantine properties of  $Y$  in terms of behavior of certain orbits of  $\Lambda_Y$  on  $\Omega$ . Let us denote by  $\mathcal{A}$  the set of  $k$ -tuples  $\mathbf{t} = (t_1, \dots, t_k) \in \mathbb{R}^k$  such that

$$t_1, \dots, t_k > 0 \quad \text{and} \quad \sum_{i=1}^m t_i = \sum_{j=1}^n t_{m+j}. \quad (3.1)$$

To any  $\mathbf{t} \in \mathcal{A}$  let us associate the diagonal matrix

$$g_{\mathbf{t}} \stackrel{\text{def}}{=} \text{diag}(e^{t_1}, \dots, e^{t_m}, e^{-t_{m+1}}, \dots, e^{-t_k}) \in G.$$

If  $\mathcal{T}$  is a subset of  $\mathcal{A}$ , we let  $g_{\mathcal{T}} \stackrel{\text{def}}{=} \{g_{\mathbf{t}} : \mathbf{t} \in \mathcal{T}\}$ . We are going to consider  $g_{\mathcal{T}}$ -orbits of lattices  $\Lambda_Y$ . The two most important special cases will be  $\mathcal{T} = \mathcal{A}$  and  $\mathcal{T} = \mathcal{R}$ , where

$$\mathcal{R} \stackrel{\text{def}}{=} \left\{ \left( \frac{t}{m}, \dots, \frac{t}{m}, \frac{t}{n}, \dots, \frac{t}{n} \right) : t > 0 \right\} \quad (3.2)$$

is the ‘central ray’ in  $\mathcal{A}$ . Also it will be convenient to use the following notation: for  $\mathbf{t} \in \mathcal{A}$ , we will denote

$$t = \sum_{i=1}^m t_i = \sum_{j=1}^n t_{m+j}, \quad (3.3)$$

so that whenever  $\mathbf{t}$  and  $t$  appear in the same formula, (3.3) will be assumed. Clearly one has  $t \geq \|\mathbf{t}\| \geq t/\min(m, n)$ . Note also that this agrees with the notation of (3.2).

Given  $\varepsilon > 0$ , consider

$$K_{\varepsilon} \stackrel{\text{def}}{=} \left\{ \Lambda \in \Omega \mid \|\mathbf{v}\| \geq \varepsilon \quad \forall \mathbf{v} \in \Lambda \setminus \{0\} \right\},$$

i.e. the collection of all unimodular lattices in  $\mathbb{R}^k$  which contain no nonzero vector of norm smaller than  $\varepsilon$ ). By Mahler’s compactness criterion (see e.g. [R, Chapter 10]), each  $K_{\varepsilon}$  is compact. It has been observed in the past<sup>2</sup> that the existence of infinitely many solutions of inequalities (1.1) and (1.2) corresponds to an unbounded sequence of excursions of certain trajectories outside of the increasing family of compact subsets described above – roughly speaking, to the trajectories growing with certain rate. To make this specific, given  $\mathcal{T} \subset \mathcal{A}$  and a lattice  $\Lambda \in \Omega$ , say that the trajectory  $g_{\mathcal{T}}\Lambda$  has *linear growth* if there exists  $\gamma > 0$  such that

$$g_{\mathbf{t}}\Lambda \notin K_{e^{-\gamma t}} \text{ for an unbounded set of } \mathbf{t} \in \mathcal{T}.$$

(The terminology is justified by the fact that for small  $\varepsilon$ , the diameter of  $K_{\varepsilon}$  is bounded from both sides by  $\text{const} \cdot \log(1/\varepsilon)$ .)

The next proposition gives the desired correspondence between approximation and dynamics:

**Proposition 3.1.** *Let  $Y \in M_{m,n}$ .*

<sup>2</sup>See also [D3] where it is proved that  $Y$  is badly approximable iff  $g_{\mathcal{R}}\Lambda_Y$  is bounded.

- (a)  $Y$  is VWA  $\Leftrightarrow g_{\mathcal{R}}\Lambda_Y$  has linear growth;
- (b)  $Y$  is VWMA  $\Leftrightarrow g_{\mathcal{A}}\Lambda_Y$  has linear growth.

Part (a) is a special case of [KM2, Theorem 8.5]. Part (b), more precisely, its ‘ $\Rightarrow$ ’ direction, has been worked out in [KM1] and [KLW] in the cases  $m = 1$  and  $n = 1$  respectively (converse direction is easier and was not required for applications). See also [KM2, Theorem 9.2] for a related statement. The proof of the general case of (b) combines the argument of the aforementioned papers; to make this paper self-contained we include the proof of both directions.

*Proof of Proposition 3.1(b).* Start with the ‘if’ part. Suppose there exists  $\gamma > 0$  and an unbounded subset  $\mathcal{T}$  of  $\mathcal{A}$  such that whenever  $\mathbf{t} \in \mathcal{T}$ , for some  $(\mathbf{p}, \mathbf{q}) \neq 0$  one has

$$e^{t_i}|Y_i\mathbf{q} - p_i| < e^{-\gamma t}, \quad i = 1, \dots, m, \quad (3.4)$$

and

$$e^{-t_{m+j}}|q_j| < e^{-\gamma t}, \quad j = 1, \dots, n. \quad (3.5)$$

We need to prove that  $Y$  is VWMA. Let  $\ell$  be the number of nonzero components of  $\mathbf{q}$ . (Note that  $\mathbf{q} \neq 0$ , otherwise from (3.4) it would follow that  $\mathbf{p} = 0$ , hence  $(\mathbf{p}, \mathbf{q}) = 0$ .) Multiplying the inequalities in (3.5) corresponding to  $q_i \neq 0$  one gets  $e^{-t}\Pi_+(\mathbf{q}) < e^{-\ell\gamma t}$ , or  $\Pi_+(\mathbf{q}) < e^{(1-\ell\gamma)t}$ . On the other hand, after multiplying inequalities from (3.4) one has  $e^t\Pi(Y\mathbf{q} - \mathbf{p}) < e^{-n\gamma t}$ , or

$$\Pi(Y\mathbf{q} - \mathbf{p}) \leq e^{-(1+n\gamma)t} = (e^{(1-\ell\gamma)t})^{-\frac{1+n\gamma}{1-\ell\gamma}} < \Pi_+(\mathbf{q})^{-\frac{1+n\gamma}{1-\ell\gamma}}. \quad (3.6)$$

Therefore, (1.2) is satisfied with some positive  $\delta = \delta(\gamma)$ . Finally observe that  $Y$  is obviously VWMA if  $Y_i\mathbf{q} \in \mathbb{Z}$  for some  $i$  and  $\mathbf{q} \in \mathbb{Z}^n \setminus \{0\}$ : indeed, it suffices to take integer multiples of  $\mathbf{q}$  to satisfy (1.2). Otherwise, taking  $\mathbf{t} \rightarrow \infty$  in  $\mathcal{T}$  we get infinitely many  $\mathbf{q}$  for which (3.6), and hence (1.2), holds.

For the other direction, let us prove two auxiliary lemmas.

**Lemma 3.2.** *Let  $Y \in M_{m,n}$  be VWMA. Then there exists  $\delta > 0$  for which there are infinitely many solutions  $\mathbf{p} \in \mathbb{Z}^m$ ,  $\mathbf{q} \in \mathbb{Z}^n \setminus \{0\}$  to (1.2) in addition satisfying*

$$\|Y\mathbf{q} - \mathbf{p}\| < \Pi_+(\mathbf{q})^{-\delta/m}. \quad (3.7)$$

*Proof.* We follow the argument of [KLW]. Choose  $\delta_0 > 0$  so that we have

$$\Pi(Y\mathbf{q} - \mathbf{p}) < \Pi_+(\mathbf{q})^{-(1+\delta_0)}, \quad (3.8)$$

for infinitely many  $\mathbf{p} \in \mathbb{Z}^m$ ,  $\mathbf{q} \in \mathbb{Z}^n$ . Let  $\mathbf{p}, \mathbf{q}$  be a solution to (3.8), and let

$$q \stackrel{\text{def}}{=} [\Pi_+(\mathbf{q})^{\frac{\delta_0}{m+n+1}}].$$

We can assume that  $\Pi_+(\mathbf{q})$  is large enough so that  $q \geq \frac{1}{2}\Pi_+(\mathbf{q})^{\frac{\delta_0}{m+n+1}}$ . For every  $\ell \in \{1, \dots, q+1\}$  set

$$\mathbf{v}_\ell \stackrel{\text{def}}{=} \ell Y\mathbf{q} \bmod 1$$

(here the fractional part is taken in each coordinate). Since  $\{\mathbf{v}_1, \dots, \mathbf{v}_{q+1}\}$  are  $q+1$  points in the unit cube  $[0, 1]^m$ , there must be two points, say  $\mathbf{v}_i, \mathbf{v}_j$ , with  $1 \leq i < j \leq q+1$ , such that

$$\|\mathbf{v}_i - \mathbf{v}_j\| \leq q^{-\frac{1}{m}} \leq \left(\frac{1}{2}\Pi_+(\mathbf{q})^{\frac{\delta_0}{m+n+1}}\right)^{-1/m} = 2^{1/m}\Pi_+(\mathbf{q})^{-\frac{\delta_0}{m(m+n+1)}}. \quad (3.9)$$

We set  $\bar{\mathbf{q}} \stackrel{\text{def}}{=} (j-i)\mathbf{q}$  and choose  $\bar{\mathbf{p}} \in \mathbb{Z}^m$  to be an integer vector closest to  $Y\bar{\mathbf{q}}$ . Note that

$$\Pi_+(\bar{\mathbf{q}}) \leq (j-i)^n \Pi_+(\mathbf{q}) \leq \Pi_+(\mathbf{q})^{\frac{n\delta_0}{m+n+1}+1}. \quad (3.10)$$

Then by inequality (3.9),

$$\begin{aligned} \|Y\bar{\mathbf{q}} - \bar{\mathbf{p}}\| &\leq 2^{1/m} \Pi_+(\mathbf{q})^{\frac{\delta_0}{m(m+n+1)}} \stackrel{(3.10)}{\leq} 2^{1/m} \Pi_+(\bar{\mathbf{q}})^{-\frac{m+n+1}{m+n+1+n\delta_0} \frac{\delta_0}{m(m+n+1)}} \\ &= 2^{1/m} \Pi_+(\bar{\mathbf{q}})^{-\frac{\delta_0}{m(m+n+1+n\delta_0)}}. \end{aligned}$$

Furthermore,

$$\begin{aligned} \Pi(Y\bar{\mathbf{q}} - \bar{\mathbf{p}}) &\leq (j-i)^m \Pi(Y\mathbf{q} - \mathbf{p}) \leq \Pi_+(\mathbf{q})^{\frac{m\delta_0}{m+n+1}} \Pi_+(\mathbf{q})^{-(1+\delta_0)} \\ &\stackrel{(3.10)}{\leq} \Pi_+(\bar{\mathbf{q}})^{-\frac{m+n+1}{m+n+1+n\delta_0} \frac{m\delta_0 - (1+\delta_0)(m+n+1)}{m+n+1}} \\ &= \Pi_+(\bar{\mathbf{q}})^{-(1+\frac{\delta_0}{m+n+1+n\delta_0})}. \end{aligned}$$

This, if we choose a positive  $\delta$  not greater than  $\frac{\delta_0}{m+n+1+\delta_0}$  and assume, as we may, that  $\Pi_+(\bar{\mathbf{q}})^{\frac{\delta_0}{m+n+1+\delta_0}-\delta}$  is not less than 2, we obtain a solution  $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$  to both (1.2) and (3.7).  $\square$

**Lemma 3.3.** *Suppose we are given  $z_1, \dots, z_m \geq 0$ ,  $r \geq 0$  and  $C > 1$  such that*

$$z_i < r \quad \text{for each } i = 1, \dots, m, \quad (3.11)$$

and

$$\prod_{i=1}^m z_i < r^m / C. \quad (3.12)$$

Then there exist  $C_1, \dots, C_m \geq 1$  such that

$$C = \prod_{i=1}^m C_i, \quad (3.13)$$

and

$$C_i z_i \leq r \quad \text{for each } i = 1, \dots, m. \quad (3.14)$$

*Proof.* Without loss of generality assume that  $z_m \leq \dots \leq z_1$ . Then define  $C_0 = 1$  and inductively

$$C_i = \min\left(\frac{r}{z_i}, \frac{C}{\prod_{j=0}^{i-1} C_j}\right) \quad (3.15)$$

(here we use the convention  $r/0 = \infty$ ). The validity of (3.14) is clear, and it follows from (3.11) that if for some  $i$  the first term in the right hand side of (3.15) is not less than the second one, the same will happen for all the subsequent values of  $i$ . Also it follows from (3.12) that a scenario under which  $r/z_i < C/\prod_{j=0}^{i-1} C_j$  for all  $i = 1, \dots, m$  is impossible. Therefore for  $i = m$  the minimum in (3.15) is equal to the second term, implying (3.13).  $\square$

Now let us get back to the proof of the remaining part of Proposition 3.1(b). Suppose that  $Y$  is VWMA; in view of Lemma 3.2 we can assume that for some  $\delta > 0$  there are infinitely many solutions to both (1.2) and (3.7). Take an arbitrary positive  $s < \frac{1}{m+n}$ , and for each solution  $(\mathbf{p}, \mathbf{q})$ , let  $r = \Pi_+(\mathbf{q})^{-\delta s}$  and define  $t_{m+1}, \dots, t_n$  by

$$|q_j|_+ = r e^{t_{m+j}}.$$

Then  $e^{-t_{m+j}}|q_j| \leq e^{-t_{m+j}}|q_j|_+ = r$  and  $\Pi_+(\mathbf{q}) = r^n e^t = \Pi_+(\mathbf{q})^{-\delta n s} e^t$ , hence  $r = e^{-\frac{\delta s}{1+\delta n s} t}$  and  $\Pi_+(\mathbf{q}) = e^{\frac{1}{1+\delta n s} t}$ . Thus, denoting  $\gamma = \frac{\delta s}{1+\delta n s}$ , we have  $e^{-t_{m+j}}|q_j| \leq e^{-\gamma t}$  for  $j = 1, \dots, n$ . To finish the proof we need to find  $t_1, \dots, t_m \geq 0$  with  $t = t_1 + \dots + t_m$  such that  $e^{t_i}|Y_i \mathbf{q} - p_i| \leq e^{-\gamma t}$  for each  $i$ ; this would clearly imply the linear growth of  $g_{\mathcal{A}} \Lambda_Y$ .

For that, let us denote  $z_i = |Y_i \mathbf{q} - p_i|$  and  $C = e^t$ , and check (3.11) and (3.12): in view of (3.7), we have

$$z_i \leq \Pi_+(\mathbf{q})^{-\delta/m} = r^{1/ms} < r$$

since  $s < 1/m$ , and also, in view of (1.2),

$$\prod_{i=1}^m z_i = \Pi(Y \mathbf{q} - \mathbf{p}) \leq \Pi_+(\mathbf{q})^{-(1+\delta)} = e^{-\frac{1+\delta}{1+\delta n s} t} = e^{-t} e^{-\frac{\delta(1-n s)}{1+\delta n s} t} = e^{-t} r^{\frac{1-n s}{s}},$$

and the latter is not greater than  $r^m/C$  since  $s < \frac{1}{m+n}$ . Taking  $e^{t_i} = C_i$  where  $C_1, \dots, C_m \geq 1$  are as in Lemma 3.3 finishes the proof.  $\square$

**Remark.** It easily follows from the continuity of the  $G$ -action on  $\Omega$  that whenever  $\mathcal{S}$  is a subset of  $\mathcal{T}$  of bounded Hausdorff distance from  $\mathcal{T}$  (that is,  $\mathcal{T}$  is contained in the  $r$ -neighborhood of  $\mathcal{S}$  for some  $r > 0$ ),  $g_{\mathcal{T}} \Lambda$  has linear growth if and only if so does  $g_{\mathcal{S}} \Lambda$ . In particular, without loss of generality we can take sets  $\mathcal{T}$  to be countable, e.g. replace  $\mathcal{A}$  with the set of vectors in  $\mathcal{A}$  with integer coordinates. See some more explanations in the proof of [KM1, Corollary 2.2].

The correspondence of Proposition 3.1 will be instrumental in our deduction of the main results of this paper from measure estimates on the space of lattices, following the method first introduced in [KM1]. Indeed, in view of the proposition, proving the extremality or strong extremality of  $F_* \nu$  is equivalent to showing that for arbitrary positive  $\gamma$ ,  $\nu$ -almost every  $\mathbf{x}$  is contained in at most finitely many sets  $\{\mathbf{x} : g_{\mathbf{t}} \Lambda_{F(\mathbf{x})} \notin K_{e^{-\gamma t}}\}$ , where  $\mathcal{T}$  is either  $\mathcal{R}$  or  $\mathcal{A}$  and  $\mathbf{t} \in \mathcal{T}$  has integer coordinates. The latter will follow from the Borel-Cantelli Lemma and estimates of type

$$\nu(\{\mathbf{x} \in B : g_{\mathbf{t}} \Lambda_{F(\mathbf{x})} \notin K_{\varepsilon}\}) \leq \text{const} \cdot \varepsilon^{\alpha} \nu(B), \quad (3.16)$$

where  $B \subset U$  is a ball and  $\alpha > 0$ .

Note that so far whenever the norm  $\|\cdot\|$  on a finite-dimensional vector space was used, in particular in the definition of the sets  $K_{\varepsilon}$ , it was meant to be the ‘maximum’ norm. However replacing it by another norm would result only in changes up to fixed multiplicative constants, and therefore Proposition 3.1 will remain true regardless of the norm used to define  $K_{\varepsilon}$ . In what follows, for geometric reasons it will be convenient to describe sets  $K_{\varepsilon}$  using Euclidean norm  $\|\cdot\|$  on  $\mathbb{R}^k$  induced by the standard inner product  $\langle \cdot, \cdot \rangle$ .

Note also that the geometry of  $\Omega$  at infinity can be similarly described using other representations of  $G$ , for example on higher exterior powers of  $\mathbb{R}^k$ . It will be convenient to denote by  $\mathcal{W}_\ell$ , where  $1 \leq \ell \leq k$ , the set of elements  $\mathbf{w} = \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_\ell$  of  $\bigwedge^\ell(\mathbb{Z}^k)$  where  $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell \in \mathbb{Z}^k\}$  can be completed to a basis of  $\mathbb{Z}^k$  (those are called *primitive  $\ell$ -tuples*). In fact, up to a sign elements of  $\mathcal{W}_\ell$  can be identified with rational  $\ell$ -dimensional subspaces of  $\mathbb{R}^k$ , or, equivalently, with primitive subgroups of  $\mathbb{Z}^k$  of rank  $\ell$ . We also let

$$\mathcal{W} \stackrel{\text{def}}{=} \cup_{1 \leq \ell \leq k} \mathcal{W}_\ell \subset \bigwedge(\mathbb{Z}^k).$$

The Euclidean norm and the inner product will be extended from  $\mathbb{R}^k$  to its exterior algebra; this way  $\|\mathbf{w}\|$  is equal to the covolume of the subgroup corresponding to  $\mathbf{w}$ . Then for  $\varepsilon > 0$  define

$$\tilde{K}_\varepsilon \stackrel{\text{def}}{=} \{g\mathbb{Z}^k \in \Omega \mid \|g\mathbf{w}\| \geq \varepsilon \quad \forall \mathbf{w} \in \mathcal{W}\}.$$

Clearly  $\tilde{K}_\varepsilon \subset K_\varepsilon$ ; on the other hand it easily follows from Minkowski's Lemma that for any positive  $\varepsilon$  one has  $K_\varepsilon \subset \tilde{K}_{c\varepsilon^{1/k}}$  where  $c > 0$  depends only on  $k$ . Therefore the following holds:

**Lemma 3.4.** *Given  $\mathcal{T} \subset \mathcal{A}$  and  $\Lambda \in \Omega$ ,  $g_{\mathbf{T}}\Lambda$  has linear growth if and only if there exists  $\gamma > 0$  such that*

$$g_{\mathbf{t}}\Lambda \notin \tilde{K}_{e^{-\gamma t}} \quad \text{for an unbounded set of } \mathbf{t} \in \mathcal{T}. \quad (3.17)$$

**Remark.** One can also use Proposition 3.1 for an alternative proof of the multiplicative version of Khintchine's Transference Principle [SW], that is, the equivalence of  $Y$  and  $Y^T$  being VWMA. Indeed, let  $\sigma$  be the linear transformation of  $\mathbb{R}^k$  induced by the permutation on the  $k$  coordinates which exchanges the group of the first  $m$  of them with that of the last  $n$ , without reordering within groups, and denote by  $\varphi$  the automorphism of  $G$  given by  $\varphi(g) = \sigma((g^T)^{-1})\sigma^{-1}$  for all  $g \in G$ . Then it is easy to see that  $\varphi(g_{\mathbf{t}}) = g_{\sigma(\mathbf{t})}$  and  $\varphi(u_Y) = u_{-Y^T}$ . Since  $\varphi(\Gamma) = \Gamma$ , the automorphism  $\varphi$  induces a self-map of  $\Omega$  which we can also denote by  $\sigma$ ; geometrically it can be interpreted as  $\varphi(\Lambda) = \sigma(\Lambda^*)$  where  $\Lambda^*$  is the lattice dual to  $\Lambda$ . Now the desired equivalence follows from an observation that  $\varphi(K_\varepsilon) \subset K_{c\varepsilon^{k-1}}$  for all  $\varepsilon > 0$ , where  $c$  is a constant dependent only on  $k$ .

#### 4. QUANTITATIVE NONDIVERGENCE AND ITS APPLICATIONS

During the last decade, starting from the paper [KM1], quantitative nondivergence estimates for unipotent trajectories on the space of lattices evolved into a powerful method yielding measure estimates as in (3.16) for a certain broad class of measures  $\nu$  and maps  $F$ . Recall that the sets in the left hand side of (3.16) consist of those  $\mathbf{x}$  for which the lattice  $g_{\mathbf{t}}\Lambda_{F(\mathbf{x})}$  has a vector of length less than  $\varepsilon$ . The crucial ingredient of the method is a way to keep track not just of length of vectors in that lattice, but of covolumes of subgroups of arbitrary dimension. The following is our main estimate:

**Theorem 4.1** ([KLW], Theorem 4.3). *Given  $d, k \in \mathbb{N}$  and positive constants  $\tilde{C}, D, \alpha$ , there exists  $C' = C'(d, k, \tilde{C}, \alpha, D) > 0$  with the following property.*

Suppose a measure  $\nu$  on  $\mathbb{R}^d$  is  $D$ -Federer on a ball  $\tilde{B}$  centered at  $\text{supp } \nu$ ,  $0 < \rho \leq 1$ , and  $h$  is a continuous map  $\tilde{B} \rightarrow G$  such that for each  $\mathbf{w} \in \mathcal{W}$ ,

(i) the function  $\mathbf{x} \mapsto \|h(\mathbf{x})\mathbf{w}\|$  is  $(\tilde{C}, \alpha)$ -good on  $\tilde{B}$  with respect to  $\nu$ ,

and

(ii)  $\|h(\mathbf{x})\mathbf{w}\| \geq \rho$  for some  $\mathbf{x} \in \text{supp } \nu \cap B$ , where  $B = 3^{-(k-1)}\tilde{B}$ .

Then for any  $0 < \varepsilon \leq \rho$ ,

$$\nu(\{\mathbf{x} \in B : h(\mathbf{x})\mathbb{Z}^k \notin K_\varepsilon\}) \leq C'(\varepsilon/\rho)^\alpha \nu(B).$$

This theorem has a long history, starting from Margulis' proof of non-divergence of unipotent flows [Mar], and continuing with a series of papers by Dani [D1, D2, D4]. The way it appeared in [KLW] is essentially the same as in [KM1] but slightly generalized. The crucial step made in [KM1] was the introduction of the requirement (condition labeled by (i) in the above theorem) that covolumes of subgroups should give rise to  $(C, \alpha)$ -good functions; this made it possible to significantly expand the applicability of the estimates. In particular, when

$$h(\mathbf{x}) = g_{\mathbf{t}}u_{F(\mathbf{x})}, \quad (4.1)$$

where  $\mathbf{t} \in \mathcal{A}$  and  $F$  is a map from  $U$  to  $M_{m,n}$ , condition (i) will hold for balls  $\tilde{B}$  centered at  $\nu$ -generic points as long as  $F$  and  $\nu$  satisfy assumption (i) of Theorem 2.1. To show this, it will be helpful to have explicit expressions for the coordinate functions of  $g_{\mathbf{t}}u_{F(\mathbf{x})}\mathbf{w}$ . Let us denote by  $\{\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{v}_1, \dots, \mathbf{v}_n\}$  the standard basis of  $\mathbb{R}^k$ . Then one has

$$u_Y \mathbf{e}_i = \mathbf{e}_i \quad \text{and} \quad u_Y \mathbf{v}_j = \mathbf{v}_j + \sum_{i=1}^m y_{i,j} \mathbf{e}_i = \mathbf{v}_j + \mathbf{y}_j, \quad (4.2)$$

where in the latter equality we have identified the columns  $\mathbf{y}_1, \dots, \mathbf{y}_n$  of  $Y$  with elements of  $E \stackrel{\text{def}}{=} \text{Span}(\mathbf{e}_1, \dots, \mathbf{e}_m)$  via the correspondence  $\mathbf{y}_j \leftrightarrow \sum_{i=1}^m y_{i,j} \mathbf{e}_i$ .

Now take  $I = \{i_1, \dots, i_r\} \subset \{1, \dots, m\}$  and  $J = \{j_1, \dots, j_s\} \subset \{1, \dots, n\}$ , where  $i_1 < \dots < i_r$  and  $j_1 < \dots < j_s$ , and consider  $\mathbf{e}_I \stackrel{\text{def}}{=} \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_r}$  and  $\mathbf{v}_J \stackrel{\text{def}}{=} \mathbf{v}_{j_1} \wedge \dots \wedge \mathbf{v}_{j_s}$ , with the convention  $\mathbf{e}_\emptyset = \mathbf{v}_\emptyset = 1$ . For any  $1 \leq \ell \leq k$ , elements

$$\mathbf{e}_I \wedge \mathbf{v}_J, \quad \text{where } I \subset \{1, \dots, m\}, J \subset \{1, \dots, n\}, |I| + |J| = \ell, \quad (4.3)$$

form a basis of  $\bigwedge^\ell(\mathbb{R}^k)$ . Then one can write

$$\begin{aligned} u_Y(\mathbf{e}_I \wedge \mathbf{v}_J) &= \mathbf{e}_I \wedge (\mathbf{v}_{j_1} + \sum_{i=1}^m y_{i,j_1} \mathbf{e}_i) \wedge \dots \wedge (\mathbf{v}_{j_s} + \sum_{i=1}^m y_{i,j_s} \mathbf{e}_i) \\ &= \sum_{L \subset J} \sum_{\substack{K \subset \{1, \dots, m\} \setminus I, \\ |K|=|L|}} \pm y_{K,L} \mathbf{e}_{I \cup K} \wedge \mathbf{v}_{J \setminus L}, \end{aligned} \quad (4.4)$$

where  $y_{K,L}$  is defined as in (2.1), and the choice of sign in  $\pm$  depends on  $K$  and  $L$ .

Now we can easily establish

**Lemma 4.2.** *Let  $d, k \in \mathbb{N}$  and  $C, \alpha > 0$ , and suppose  $\tilde{B}$  is a ball in  $\mathbb{R}^d$ ,  $\nu$  is a measure on  $\tilde{B}$ , and  $F : \tilde{B} \rightarrow M_{m,n}$  is a continuous map such that  $(\mathbf{d} \circ F, \nu)$  is  $(C, \alpha)$ -good on  $\tilde{B}$ . Then functions  $\mathbf{x} \mapsto \|g_{\mathbf{t}} u_{F(\mathbf{x})} \mathbf{w}\|$  are  $(N^{\alpha/2} C, \alpha)$ -good on  $\tilde{B}$  with respect to  $\nu$  for any  $\mathbf{t} \in \mathcal{A}$  and  $\mathbf{w} \in \mathcal{W}$ , where  $N$  is as in (2.2).*

*Proof.* Take  $\mathbf{w} \in \bigwedge^\ell(\mathbb{R}^k)$ . In view of (4.4), each coordinate of  $u_{F(\cdot)} \mathbf{w}$  with respect to the basis (4.3) is a linear combination of functions  $F(\cdot)_{K,L}$  for various  $K \subset \{1, \dots, m\}$  and  $L \subset \{1, \dots, n\}$  with  $|K| = |L|$ , that is, of 1 and components of  $\mathbf{d} \circ F$ . The same can be said about coordinates of  $g_{\mathbf{t}} u_{F(\cdot)} \mathbf{w}$ ; in fact, the basis (4.3) consists of eigenvectors for  $g_{\mathbf{t}}$ . It remains to apply a well-known and elementary property, see e.g. [KLW, Lemma 4.1], that whenever  $f_1, \dots, f_N$  are  $(C, \alpha)$ -good on a set  $U$  with respect to a measure  $\nu$ , the function  $(f_1^2 + \dots + f_N^2)^{1/2}$  is  $(N^{\alpha/2} C, \alpha)$ -good on  $U$  with respect to  $\nu$ .  $\square$

Consequently, whenever  $F$  and  $\nu$  satisfy assumption (i) of Theorem 2.1 (in particular, if  $F$  is real analytic and  $\nu$  is Lebesgue measure), for  $\nu$ -almost all  $\mathbf{x}$  it is possible to choose a ball  $\tilde{B}$  centered at  $\mathbf{x}$  and  $\tilde{C}, \alpha > 0$  such that  $h(\cdot)$  as in (4.1) satisfies condition (i) of Theorem 4.1. Our attention will be thus centered on lower bounds for  $\|g_{\mathbf{t}} u_{F(\cdot)} \mathbf{w}\|_{\nu, B}$ ; indeed, a bound uniform in  $\mathbf{w}$  and  $\mathbf{t}$  would make it possible to apply Theorem 4.1 and establish (3.16). Moreover, generalizing a result from [K1] it is possible to write down a condition equivalent to the statement

$$g_{\mathcal{T}} \Lambda_{F(\mathbf{x})} \text{ has linear growth for } \nu\text{-almost no } \mathbf{x} \quad (4.5)$$

within the class of Federer measures and good pairs.

**Theorem 4.3.** *Let an open subset  $U$  of  $\mathbb{R}^d$ , a continuous map  $F : U \rightarrow M_{m,n}$  and a Federer measure  $\nu$  on  $U$  be such that the pair  $(\mathbf{d} \circ F, \nu)$  is good. Also let  $\mathcal{T}$  be an unbounded subset of  $\mathcal{A}$ . Then (4.5) holds if and only if for any ball  $B \subset U$  with  $\nu(B) > 0$  and any  $\beta > 0$  there exists  $T > 0$  such that*

$$\|g_{\mathbf{t}} u_{F(\cdot)} \mathbf{w}\|_{\nu, B} \geq e^{-\beta t} \quad \forall \mathbf{w} \in \mathcal{W} \text{ and any } \mathbf{t} \in \mathcal{T} \text{ with } t \geq T. \quad (4.6)$$

*Proof.* Let us start with the ‘if’ part. Take an arbitrary positive  $\gamma$ . Since  $\nu$  is Federer,  $(\mathbf{d} \circ F, \nu)$  is good and in view of Lemma 4.2, for  $\nu$ -almost every  $\mathbf{x}_0 \in U$  there exists a ball  $\tilde{B}$  centered at  $\mathbf{x}_0$  and constants  $\tilde{C}, \alpha, D$  such that all the functions  $\mathbf{x} \mapsto \|g_{\mathbf{t}} u_{F(\mathbf{x})} \mathbf{w}\|$  are  $(\tilde{C}, \alpha)$ -good on  $\tilde{B}$  with respect to  $\nu$ , and  $\nu$  is  $D$ -Federer on  $\tilde{B}$ . Then take  $B = 3^{-(k-1)} \tilde{B}$ , choose an arbitrary  $0 < \beta < \gamma$  and  $T$  such that (4.6) holds. This will enforce condition (ii) of Theorem 4.1 with  $\rho = e^{-\beta t}$  and  $h$  as in (4.1) with  $t \geq T$ . Applying Theorem 4.1 with  $\varepsilon = e^{-\gamma t}$  will yield

$$\nu(\{\mathbf{x} \in B : g_{\mathbf{t}} \Lambda_{F(\mathbf{x})} \notin K_{e^{-\gamma t}}\}) \leq C'(e^{-(\gamma-\beta)t})^\alpha \nu(B).$$

Now choose a countable subset  $\mathcal{S}$  of  $\mathcal{T}$  with finite Hausdorff distance from  $\mathcal{T}$  and such that  $\inf_{\mathbf{t}_1, \mathbf{t}_2 \in \mathcal{S}, \mathbf{t}_1 \neq \mathbf{t}_2} \|\mathbf{t}_1 - \mathbf{t}_2\| > 0$ . The above estimate implies that

$$\sum_{\mathbf{t} \in \mathcal{S}} \nu(\{\mathbf{x} \in B : g_{\mathbf{t}} \Lambda_{F(\mathbf{x})} \notin K_{e^{-\gamma t}}\}) < \infty.$$

Applying the Borel-Cantelli Lemma, one concludes that for  $\nu$ -a.e.  $\mathbf{x} \in B$  one has

$$g_{\mathbf{t}}\Lambda_{F(\mathbf{x})} \in K_{e^{-\gamma t}}$$

for all but finitely many  $\mathbf{t} \in \mathcal{S}$ , which, in view of the remark before Proposition 3.1 and since  $\gamma$  could be chosen arbitrary small, implies (4.5).

As for the converse, suppose that there exists a ball  $B \subset U$  with  $\nu(B) > 0$  and  $\beta > 0$  such that for an unbounded set of  $\mathbf{t} \in \mathcal{T}$  one has

$$\|g_{\mathbf{t}}u_{F(\cdot)}\mathbf{w}\|_{\nu,B} < e^{-\beta t}$$

for some  $\mathbf{w} \in \mathcal{W}$  (dependent on  $\mathbf{t}$ ). This means that for any  $\mathbf{x} \in B \cap \text{supp } \nu$  and for each  $\mathbf{t}$  as above,  $g_{\mathcal{T}}\Lambda_{F(\mathbf{x})}$  is not in  $\tilde{K}_{e^{-\beta t}}$ . This, in view of Lemma 3.4, implies that  $g_{\mathcal{T}}\Lambda_{F(\mathbf{x})}$  has linear growth for all  $\mathbf{x}$  in  $B \cap \text{supp } \nu$ .  $\square$

In particular, in view of Proposition 3.1, for  $\mathcal{T} = \mathcal{R}$  or  $\mathcal{A}$  we get criteria for extremality and strong extremality of  $F_*\nu$  within the class of good pairs. Note that here we see a dichotomy between a certain property happening either for almost no points or for all points in some nonempty open ball. This is typical for this class of problems, see [K1, K2, K4, Zh].

## 5. PROOF OF THEOREM 2.1

In general, checking a conditions like (4.6) seems to be a complicated task; the full strength of the vector case ( $n = 1$ ) of Theorem 4.3 has been utilized in [K1], see also [K2, Zh]. However, we will show that the nonplanarity assumption of Theorems 2.1 implies a stronger property, namely  $e^{-\beta t}$  in the right hand side of (4.6) can be replaced by a positive constant dependent only on  $B$ . To establish such lower bounds, we are going to look closely at projections of ‘curves’  $\{u_{F(\mathbf{x})}\mathbf{w}\}$  in  $\bigwedge(\mathbb{R}^k)$  onto subspaces expanded by the  $g_{\mathbf{t}}$ -action. Namely, for a fixed  $\mathbf{t}$  let us denote by  $E_{\mathbf{t}}^+$  the span of all the eigenvectors of  $g_{\mathbf{t}}$  in  $\bigwedge(\mathbb{R}^k)$  with eigenvalues greater or equal to one (in other words, those which are not contracted by the  $g_{\mathbf{t}}$ -action). It is easy to see that  $E_{\mathbf{t}}^+$  is spanned by elements  $\mathbf{e}_I \wedge \mathbf{v}_J$  where  $I \subset \{1, \dots, m\}$  and  $J \subset \{1, \dots, n\}$  are such that

$$\sum_{i \in I} t_i \geq \sum_{j \in J} t_{m+j}.$$

Also let  $\pi_{\mathbf{t}}^+$  be the orthogonal projection onto  $E_{\mathbf{t}}^+$ . As a straightforward application of Theorem 4.3, we have

**Corollary 5.1.** *Let  $F : U \rightarrow M_{m,n}$ ,  $\nu$  and  $\mathcal{T}$  be as in Theorem 4.3. Suppose that for any ball  $B \subset U$  with  $\nu(B) > 0$  one has*

$$\inf_{\mathbf{w} \in \mathcal{W}, \mathbf{t} \in \mathcal{T}} \|\pi_{\mathbf{t}}^+ u_{F(\cdot)}\mathbf{w}\|_{\nu,B} > 0. \tag{5.1}$$

Then (4.5) holds.

*Proof.* By the definition of the map  $\pi_{\mathbf{t}}^+$ , for any  $\mathbf{w} \in \bigwedge(\mathbb{R}^k)$  one has

$$\|g_{\mathbf{t}}\mathbf{w}\| \geq \|\pi_{\mathbf{t}}^+ g_{\mathbf{t}}\mathbf{w}\| = \|g_{\mathbf{t}}\pi_{\mathbf{t}}^+\mathbf{w}\| \geq \|\pi_{\mathbf{t}}^+\mathbf{w}\|,$$

hence a uniform lower bound, say  $c$ , on  $\|\pi_{\mathbf{t}}^+ u_{F(\mathbf{x})}\mathbf{w}\|$  implies a similar bound on  $\|g_{\mathbf{t}}u_{F(\mathbf{x})}\mathbf{w}\|$ . Thus (4.6) will hold as long as  $e^{-\beta T} \leq c$ .  $\square$

Our strategy for checking extremality or strong extremality will be to derive estimates of type (5.1) from the nonplanarity assumptions, in particular from those of Theorem 2.1. However before proceeding let us exhibit a partial converse to the above corollary:

**Corollary 5.2.** *Let  $Y \in M_{m,n}$  and let  $\mathcal{T}$  be an unbounded subset of  $\mathcal{A}$ . Suppose that there exist  $\mathbf{t}_0 \in \mathcal{T}$  and  $\mathbf{w} \in \mathcal{W}$  such that:*

- (a)  $\mathcal{S} \stackrel{\text{def}}{=} \{c\mathbf{t}_0 : c > 0\} \cap \mathcal{T}$  is unbounded; and
- (b)  $\pi_{\mathbf{t}_0}^+ u_Y \mathbf{w} = 0$ .

Then  $g_{\mathcal{T}}\Lambda_Y$  has linear growth.

*Proof.* From (a) and (b) it follows that  $u_Y \mathbf{w}$  belongs to the orthogonal complement of  $E_{\mathbf{t}}^+$  whenever  $\mathbf{t} \in \mathcal{S}$  (clearly the spaces  $E_{\mathbf{t}}^+$  do not change if  $\mathbf{t}$  is replaced by a proportional vector). Hence it is exponentially contracted by the  $g_{\mathbf{t}}$ -action, that is, for some  $\beta > 0$  and all  $\mathbf{t} \in \mathcal{S}$  one can write

$$\|g_{\mathbf{t}} u_Y \mathbf{w}\| \leq e^{-\beta t} \|u_Y \mathbf{w}\| \leq C e^{-\beta t},$$

where  $C$  is a constant depending on  $\mathbf{w}$  and  $Y$ . Consequently (3.17) is satisfied with  $\Lambda = \Lambda_Y$ , and Lemma 3.4 readily implies the linear growth of  $g_{\mathcal{T}}\Lambda_Y$ .  $\square$

In particular, whenever conditions (a) and (b) above are satisfied for some  $\mathbf{t}_0 \in \mathcal{T}$  and  $Y$  of the form  $F(\mathbf{x})$  for all  $\mathbf{x} \in \text{supp } \nu \cap B$ , where  $B \subset U$  is a ball of positive measure (that is, the infimum in the left hand side of (5.1) is equal to zero and is attained), it follows that (4.5) does not hold, and, moreover,  $g_{\mathcal{T}}\Lambda_{F(\mathbf{x})}$  has linear growth for all  $\mathbf{x} \in B \cap \text{supp } \nu$ . We will explore this when it comes to discussing specific examples at the end of the paper.

Now let us get back to Corollary 5.1 and its applications. The next observation immediately follows from the compactness of spheres in finite-dimensional spaces:

**Lemma 5.3.** *Let  $\nu$  be a measure on  $\mathbb{R}^d$  and  $\mathbf{f} = (f_1, \dots, f_N)$  a map  $U \rightarrow \mathbb{R}^N$ , where  $U \subset \mathbb{R}^d$  is open with  $\nu(U) > 0$ . Then  $(\mathbf{f}, \nu)$  is nonplanar if and only if for any ball  $B \subset U$  with  $\nu(B) > 0$  there exists  $c > 0$  such that*

$$\|a_0 + \sum_{i=1}^N a_i f_i\|_{\nu, B} \geq c \text{ for any } a_0, a_1, \dots, a_N \text{ with } \max |a_i| \geq 1, .$$

Since it is assumed in Theorem 2.1 that  $(\mathbf{d} \circ F, \nu)$  is nonplanar, in view of the above lemma and corollary to check (5.1) it would suffice to bound  $\|\pi_{\mathbf{t}}^+ u_{F(\cdot)} \mathbf{w}\|$  from below by the absolute value of a linear combination of 1 coordinates of  $\mathbf{d} \circ F$  with big enough coefficients.

Note that the spaces  $E_{\mathbf{t}}^+$  may be different for different  $\mathbf{t}$  (although, as was mentioned above,  $E_{\mathbf{t}}^+ = E_{\mathbf{t}'}^+$  if  $\mathbf{t}$  and  $\mathbf{t}'$  are proportional). However, it turns out that in the set-up of Theorem 2.1 one can work with the intersection of all those spaces:

$$E^+ \stackrel{\text{def}}{=} \bigcap_{\mathbf{t} \in \mathcal{A}} E_{\mathbf{t}}^+$$

consisting of elements which are not contracted by  $g_{\mathbf{t}}$  for all  $\mathbf{t} \in \mathcal{A}$ . It is easy to see that  $E^+$  is spanned by

$$\{\mathbf{e}_I, \mathbf{e}_{\{1, \dots, m\}} \wedge \mathbf{v}_J : I \subset \{1, \dots, m\}, J \subset \{1, \dots, n\}\}. \quad (5.2)$$

The next proposition explains that for any  $\mathbf{w} \in \mathcal{W}$  it is possible to find an element of  $E^+$  on which the ‘curves’  $\{u_{F(\mathbf{x})}\mathbf{w}\}$  project nontrivially.

**Proposition 5.4.** *For any  $\mathbf{w} \in \mathcal{W}$  it is possible to choose an element  $\mathbf{w}_0$  of the basis (5.2) of  $E^+$  such that the function  $Y \mapsto \langle u_Y \mathbf{w}, \mathbf{w}_0 \rangle$  is a nontrivial integer linear combination of 1 and components of  $\mathbf{d}(Y)$ .*

*Proof.* Denote by  $\pi^+$  the orthogonal projection onto  $E^+$ . We are going to use (4.4) to explicitly write down the coordinates  $\pi^+ u_Y \mathbf{w}$  with respect to the basis (5.2) for any  $\mathbf{w} \in \mathcal{W}_\ell$ , that is,

$$\mathbf{w} = \sum_{I, J, |I|+|J|=\ell} a_{I, J} \mathbf{e}_I \wedge \mathbf{v}_J. \quad (5.3)$$

Consider two cases.

**Case 1.** If  $\ell \leq m$ , using (4.4) one can see that

$$\pi^+ u_Y (\mathbf{e}_I \wedge \mathbf{v}_J) = \mathbf{e}_I \wedge \sum_{K \subset \{1, \dots, m\} \setminus I, |K|=|J|} \pm y_{K, J} \mathbf{e}_K.$$

Note that  $|I|$  can take values between  $\max(0, \ell - n)$  and  $\ell$ ; equivalently,  $|J| = \ell - |I|$  ranges between 0 and  $\ell - \max(0, \ell - n) = \min(\ell, n)$ . Thus

$$\pi^+ u_Y \mathbf{w} = \sum_{\substack{I \subset \{1, \dots, m\} \\ \max(0, \ell - n) \leq |I| \leq \ell}} \mathbf{e}_I \wedge \sum_{\substack{J \subset \{1, \dots, n\} \\ |J| = \ell - |I|}} a_{I, J} \sum_{\substack{K \subset \{1, \dots, m\} \setminus I \\ |K| = |J|}} \pm y_{K, J} \mathbf{e}_K.$$

Rearranging terms and substituting  $L = I \cup K$ , we get

$$\pi^+ u_Y \mathbf{w} = \sum_{\substack{L \subset \{1, \dots, m\} \\ |L| = \ell}} \left( \sum_{\substack{K \subset L \\ 0 \leq |K| \leq \min(\ell, n) \leq |I| \leq \ell}} \sum_{\substack{J \subset \{1, \dots, n\} \\ |J| = |K|}} \pm a_{L \setminus K, J} y_{K, J} \right) \mathbf{e}_L.$$

Recall that the coefficients in the expansion (5.3) are integer and at least one of them, say  $a_{I, J}$ , is nonzero. Take any  $K \subset \{1, \dots, m\} \setminus I$  with  $|K| = |J|$  and denote  $L \stackrel{\text{def}}{=} I \cup K$ . Then

$$\langle u_Y \mathbf{w}, \mathbf{e}_L \rangle = \sum_{\substack{K \subset L \\ 0 \leq |K| \leq \min(\ell, n) \leq |I| \leq \ell}} \sum_{\substack{J \subset \{1, \dots, n\} \\ |J| = |K|}} \pm a_{L \setminus K, J} y_{K, J}$$

will be a nontrivial (since  $a_{I, J}$  is one of the coefficients) integer linear combination of 1 and components of  $\mathbf{d}(Y)$ .

**Case 2.** If  $\ell \geq m$ , we get

$$\begin{aligned} \pi^+ u_Y(\mathbf{e}_I \wedge \mathbf{v}_J) &= \mathbf{e}_I \wedge \left( \sum_{K \subset J, |K|=m-|I|} \pm y_{\{1, \dots, m\} \setminus I, K} \mathbf{e}_{\{1, \dots, m\} \setminus I} \wedge \mathbf{v}_{J \setminus K} \right) \\ &= \mathbf{e}_{\{1, \dots, m\}} \wedge \left( \sum_{K \subset J, |K|=m-|I|} \pm y_{\{1, \dots, m\} \setminus I, K} \mathbf{v}_{J \setminus K} \right). \end{aligned}$$

Note that this time we must have  $\max(0, \ell - n) \leq |I| \leq m$ , or, equivalently,  $\{1, \dots, m\} \setminus I \leq m - \max(0, \ell - n) = \min(m, k - \ell)$ . Therefore:

$$\pi^+ u_Y \mathbf{w} = \mathbf{e}_{\{1, \dots, m\}} \wedge \sum_{\substack{I \subset \{1, \dots, m\} \\ |I| \geq \max(0, \ell - n)}} \sum_{\substack{J \subset \{1, \dots, n\} \\ |J| = \ell - |I|}} a_{I, J} \sum_{\substack{K \subset J \\ |K| = m - |I|}} \pm y_{\{1, \dots, m\} \setminus I, K} \mathbf{v}_{J \setminus K}.$$

Rearranging terms, substituting  $L = J \setminus K$  and replacing  $I$  with  $\{1, \dots, m\} \setminus I$ , we get

$$\pi^+ u_Y \mathbf{w} = \sum_{\substack{L \subset \{1, \dots, n\} \\ |L| = \ell - m}} \left( \sum_{\substack{I \subset \{1, \dots, m\} \\ |I| \leq \min(m, k - \ell)}} \sum_{\substack{K \subset \{1, \dots, n\} \setminus L \\ |K| = |I|}} \pm a_{\{1, \dots, m\} \setminus I, K \cup L} y_{I, K} \right) \mathbf{e}_{\{1, \dots, m\}} \wedge \mathbf{v}_L.$$

Now let  $a_{\{1, \dots, m\} \setminus I, J}$  be a nonzero coefficient. Then one can take any  $K \subset J$  with  $|K| = |I|$  and conclude that

$$\langle u_Y \mathbf{w}, \mathbf{e}_{\{1, \dots, m\}} \wedge \mathbf{v}_L \rangle = \sum_{\substack{I \subset \{1, \dots, m\} \\ |I| \leq \min(m, k - \ell)}} \sum_{\substack{K \subset \{1, \dots, n\} \setminus L \\ |K| = |I|}} \pm a_{\{1, \dots, m\} \setminus I, K \cup L} y_{I, K}$$

is a nontrivial (since  $a_{\{1, \dots, m\} \setminus I, J}$  is one of the coefficients) integer linear combination of 1 and the components of  $\mathbf{d}(Y)$ . This finishes the proof of the proposition.  $\square$

We remark that in the case  $m = 1$  or  $n = 1$  all the spaces  $E_{\mathbf{t}}^+$ ,  $\mathbf{t} \in \mathcal{A}$ , coincide with each other and with  $E^+$ ; in that case in [KM1] and [KLW] a simplified form of the above computation was used to prove [KM1, Theorem 5.4] and [KLW, Theorem 3.3] respectively.

Finally we can complete the

*Proof of Theorem 2.1.* Recall that it suffices to check that the assumption of Corollary 5.1 are satisfied. Take  $B \subset U$  with  $\nu(B) > 0$ , and write

$$\|\pi_{\mathbf{t}}^+ u_{F(\cdot)} \mathbf{w}\| \geq |\langle \pi_{\mathbf{t}}^+ u_{F(\cdot)} \mathbf{w}, \mathbf{w}_0 \rangle| = |\langle u_{F(\cdot)} \mathbf{w}, \mathbf{w}_0 \rangle| \cdot \|\pi_{\mathbf{t}}^+ \mathbf{w}_0\| \geq |\langle u_{F(\cdot)} \mathbf{w}, \mathbf{w}_0 \rangle|,$$

where  $\mathbf{w}_0$  is as in Proposition 5.4, so that  $\langle u_{F(\cdot)} \mathbf{w}, \mathbf{w}_0 \rangle$  is a nontrivial integer linear combination of 1 and the components of  $\mathbf{d} \circ F$ . Therefore Lemma 5.3 and the nonplanarity of  $(\mathbf{d} \circ F, \nu)$  imply (5.1). In view of Corollary 5.1, Theorem 4.3 and Proposition 3.1(b),  $F_* \nu$  is strongly extremal.  $\square$

6. CONSEQUENCES OF THEOREM 2.1

Our goal in this section is to construct examples of pairs  $(F, \nu)$  such that the assumptions of Theorem 2.1 are satisfied. As mentioned in §2, whenever  $\mathbf{f} : U \rightarrow \mathbb{R}^N$  is a nondegenerate smooth map, the pair  $(\mathbf{f}, \lambda)$  is good and nonplanar (see [KM1, Proposition 3.4]). Since Lebesgue measure is Federer, Theorem 2.1 as a special case implies

**Corollary 6.1.** *Let  $F : U \rightarrow M_{m,n}$  be a differentiable map such that  $\mathbf{d} \circ F$  is nondegenerate. Then  $F_*\lambda$  is strongly extremal.*

Specific examples include

$$x \mapsto \begin{pmatrix} x & x^2 \\ x^3 & x^5 \end{pmatrix}, \quad \text{or} \quad x \mapsto \begin{pmatrix} x & x^2 & x^3 \\ x^4 & x^6 & x^8 \end{pmatrix}, \quad (6.1)$$

where  $x \in \mathbb{R}$ . More generally, here is a definition introduced in [KLW]: given  $C$ ,  $\alpha > 0$  and an open subset  $U$  of  $\mathbb{R}^d$ , say that  $\nu$  is *absolutely decaying* if for  $\nu$ -a.e.  $\mathbf{x} \in \mathbb{R}^n$  there exist a neighborhood  $U$  of  $\mathbf{x}$  and  $C, \alpha > 0$  such that for any non-empty open ball  $B \subset U$  centered at  $\text{supp } \nu$ , any affine hyperplane  $\mathcal{L} \subset \mathbb{R}^n$  and any  $\varepsilon > 0$  one has

$$\nu(B \cap \mathcal{L}^{(\varepsilon)}) \leq C \left(\frac{\varepsilon}{r}\right)^\alpha \nu(B),$$

where  $r$  is the radius of  $B$  and  $\mathcal{L}^{(\varepsilon)}$  is the  $\varepsilon$ -neighborhood of  $\mathcal{L}$ . Following a terminology suggested in [PV], say that  $\nu$  is *absolutely friendly* if it is Federer and absolutely decaying. The following was essentially proved in [KLW] (see [KLW, Theorem 2.1(b) and §7]): suppose that  $\nu$  is an absolutely friendly measure,  $\ell \in \mathbb{N}$ , and  $\mathbf{f}$  is a  $C^{\ell+1}$  map which is  $\ell$ -nondegenerate at  $\nu$ -a.e. point; then  $(\mathbf{f}, \nu)$  is good. Since the nonplanarity of  $(\mathbf{f}, \nu)$  is immediate from the nondegeneracy condition, the following is also a special case of Theorem 2.1:

**Corollary 6.2.** *Let  $\nu$  be an absolutely friendly measure on  $\mathbb{R}^d$ ,  $U$  an open subset of  $\mathbb{R}^d$ ,  $\ell \in \mathbb{N}$ , and  $F : U \rightarrow M_{m,n}$  a  $C^{\ell+1}$  map such that  $\mathbf{d} \circ F$  is  $\ell$ -nondegenerate at  $\nu$ -almost every point. Then  $F_*\nu$  is strongly extremal.*

Numerous examples of absolutely friendly measures have been constructed in [KLW, KW1, U, SU]. In particular, limit measures of finite irreducible systems of contracting similarities [KLW, §8] (or, more generally, self-conformal contractions, [U]) satisfying the open set condition are absolutely friendly. Thus, if  $\nu$  is, say, the natural measure on the Cantor set in  $\mathbb{R}$  and  $F$  is one of the maps of the form (6.1), the pushforward of  $\nu$  by  $F$  is strongly extremal.

In general checking the nondegeneracy of  $\mathbf{d} \circ F$  may be a complicated task. However, in the important special case when the rows (or columns) of  $F$  are functions of independent variables, the assumptions of Theorem 2.1 turn out to be easier to check. Namely, the following is true:

**Theorem 6.3.** *For every  $i = 1, \dots, m$ , let  $\mathbf{f}_i$  be a continuous map from an open subset  $U_i$  of  $\mathbb{R}^{d_i}$  to  $\mathbb{R}^n$ , and let  $\nu_i$  be a Federer measure on  $\mathbb{R}^{d_i}$  such that for each  $i$ , the pair  $(\mathbf{f}_i, \nu_i)$  is good and nonplanar. Define  $F$  by (1.3) and let  $\nu = \nu_1 \times \dots \times \nu_m$ . Then (a)  $\nu$  is Federer, and  $(\mathbf{d} \circ F, \nu)$  is (b) good and (c) nonplanar.*

In view of the discussion preceding Corollary 6.1, the above theorem includes Proposition 2.2 as a special case; hence its proof also establishes Theorem 1.1.

*Proof.* The fact that the product of Federer measures is Federer is straightforward, see e.g. [KLW, Theorem 2.4]. For parts (b) and (c) we will use induction on  $m$ . The case  $m = 1$  is obvious since in that case  $\mathbf{d} \circ F$  is the same as  $F$ . The induction step is based on the following elementary observation: given  $Y \in M_{m,n}$  with  $m > 1$ , any linear combination of components of  $\mathbf{d}(Y)$  and 1, that is,

$$\sum_{\substack{I \subset \{1, \dots, m\}, J \subset \{1, \dots, n\} \\ 0 \leq |I| = |J| \leq \min(m, n)}} a_{I,J} y_{I,J} \quad (6.2)$$

can be rewritten as

$$\sum_{\substack{I \subset \{2, \dots, m\}, J \subset \{1, \dots, n\} \\ 0 \leq |I| = |J| \leq \min(m-1, n)}} \left( a_{I,J} + \sum_{j \notin J} a_{I \cup \{1\}, J \cup \{j\}} y_{1,j} \right) y_{I,J}, \quad (6.3)$$

where the choice of signs in  $\pm$  depends on  $j$  and  $J$ .

Let us first establish the nonplanarity of  $(\mathbf{d} \circ F, \nu)$ . Denote  $\mathbf{x}' = (\mathbf{x}_2, \dots, \mathbf{x}_m)$ ,  $\nu' = \nu_2 \times \dots \times \nu_m$ , and let

$$F' : U_2 \times \dots \times U_m \rightarrow M_{m,n}, \quad \mathbf{x}' \mapsto \begin{pmatrix} \mathbf{f}_2(\mathbf{x}_2) \\ \vdots \\ \mathbf{f}_m(\mathbf{x}_m) \end{pmatrix}$$

Assume that the statement is true for  $m - 1$  in place of  $m$ , which in particular implies that the pair  $(\mathbf{d} \circ F', \nu')$  is nonplanar. Take a ball  $B \subset \mathbb{R}^{d_1 + \dots + d_m}$  with  $\nu(B) > 0$ . Choose coefficients  $a_{I,J} \in \mathbb{R}$ , where  $I \subset \{1, \dots, m\}$ ,  $J \subset \{1, \dots, n\}$ ,  $0 \leq |I| = |J| \leq \min(m, n)$ , such that one of them has absolute value at least 1, and denote

$$\varphi(\mathbf{x}) = \sum_{I,J} a_{I,J} f(\mathbf{x})_{I,J}$$

Then, using the equivalence of (6.2) and (6.3), one can write

$$\varphi(\mathbf{x}) = \sum_{\substack{I \subset \{2, \dots, m\}, J \subset \{1, \dots, n\} \\ 0 \leq |I| = |J| \leq \min(m-1, n)}} \left( a_{I,J} + \sum_{j \notin J} a_{I \cup \{1\}, J \cup \{j\}} f_{1,j}(\mathbf{x}_1) \right) f(\mathbf{x}')_{I,J}. \quad (6.4)$$

Since  $\max |a_{I,J}| \geq 1$ , one can choose  $I \subset \{2, \dots, m\}$  and  $J \subset \{1, \dots, n\}$  such that the absolute value of some coefficient in the expression

$$a_{I,J} + \sum_{j \notin J} a_{I \cup \{1\}, J \cup \{j\}} f_{1,j}$$

is at least 1. Since  $(\mathbf{f}_1, \nu_1)$  is nonplanar, Lemma 5.3 implies that there exists  $\tilde{\mathbf{x}}_1 \in B_1 \cap \text{supp } \nu_1$  and  $c_1 > 0$  such that  $|a_{I,J} + \sum_{j \notin J} a_{I \cup \{1\}, J \cup \{j\}} f_{1,j}(\tilde{\mathbf{x}}_1)| \geq c_1$ . Fixing  $\mathbf{x}_1 = \tilde{\mathbf{x}}_1$ , we infer that at least one of the functions  $f(\cdot)_{I,J}$  in the linear combination (6.4) has a coefficient of absolute value at least  $c_1$ . From the nonplanarity of  $(\mathbf{d} \circ F', \nu')$  we can then deduce the existence of  $c > 0$  and  $\mathbf{x}'$

such that  $\mathbf{x} = (\tilde{\mathbf{x}}_1, \mathbf{x}') \in B \cap \text{supp } \nu$  and  $|\varphi(\mathbf{x})| > c$ . This, again in view of Lemma 5.3, shows the nonplanarity of  $(\mathbf{d} \circ F, \nu)$ .

The proof of part (c) goes along similar lines and is based on the following

**Lemma 6.4** ([KT], Lemma 2.2). *Let metric spaces  $X, Y$  with measures  $\mu, \nu$  be given. Suppose  $\varphi$  is a continuous function on  $U \times V$ , where  $U \subset X$  and  $V \subset Y$  are open subsets, and suppose  $C, D, \alpha, \beta$  are positive constants such that*

$$\begin{aligned} & \text{for all } y \in V \cap \text{supp } \nu, \text{ the function } x \mapsto \varphi(x, y) \\ & \text{is } (C, \alpha)\text{-good on } U \text{ with respect to } \mu, \end{aligned}$$

and

$$\begin{aligned} & \text{for all } x \in U \cap \text{supp } \mu, \text{ the function } y \mapsto \varphi(x, y) \\ & \text{is } (D, \beta)\text{-good on } V \text{ with respect to } \nu. \end{aligned}$$

Then  $\varphi$  is  $(E, \gamma)$ -good on  $U \times V$  with respect to  $\mu \times \nu$ , where  $E$  and  $\gamma$  can be explicitly expressed in terms of  $\alpha, \beta, C, D$ .

It is given that  $\nu_1$ -a.e. point of  $\mathbb{R}^{d_1}$  has a neighborhood  $U_1$  such that  $(\mathbf{f}_1, \nu_1)$  is  $(C_1, \alpha_1)$ -good on  $U_1$  for some  $C_1, \alpha_1 > 0$ . From the induction assumption it follows that  $\nu'$ -a.e. point of  $\mathbb{R}^{d_2+\dots+d_m}$  has a neighborhood  $U'$  such that  $(\mathbf{d} \circ F', \nu')$  is  $(C', \alpha')$ -good on  $U'$  for some  $C', \alpha' > 0$ . Taking  $U = U_1, V = U'$  and  $\varphi$  as in (6.4), one sees that the assumptions of the above lemma are satisfied, and therefore for  $\nu$ -a.e.  $(\mathbf{x}_1, \mathbf{x}')$  there exists a neighborhood  $U$  of  $(\mathbf{x}_1, \mathbf{x}')$  and  $C, \alpha > 0$  such that  $(\mathbf{d} \circ F, \nu)$  is  $(C, \alpha)$ -good on  $U$ . This finishes the proof of Theorem 6.3.  $\square$

## 7. LOW-DIMENSIONAL EXAMPLES

It is not hard to guess, looking at the information used in Corollary 5.1, that it might be possible to weaken the nonplanarity assumption of Theorem 2.1 by requiring only some, and not all, linear combinations of components of  $\mathbf{d} \circ F$  to be nonzero. In this section we consider some low-dimensional special cases and exhibit conditions sufficient for strong extremality and extremality of  $F_*\nu$  which are weaker than the ones required by Theorem 2.1, thus generating new examples of extremal and strongly extremal measures.

With some abuse of notation, let us introduce the following definition: say that a pair  $(F, \nu)$ , where  $F : U \rightarrow M_{m,n}$  and  $\nu$  is a measure on  $U$ , is *nonplanar* if for any ball  $B \subset U$  with  $\nu(B) > 0$  and any nonzero  $\mathbf{v} \in \mathbb{R}^n$ , the restriction of the map  $\mathbf{x} \mapsto F(\mathbf{x})\mathbf{v}$  to  $B \cap \text{supp } \nu$  is nonconstant. Clearly it coincides with the definition of nonplanarity if  $m = 1$ , and clearly  $(F, \nu)$  is row-nonplanar if  $F$  has a row  $\mathbf{f}$  such that  $(\mathbf{f}, \nu)$  is nonplanar (but converse is not true).

For the first result of this section, let us take  $m = n = 2$ .

**Theorem 7.1.** *Let  $\nu$  be a Federer measure on  $\mathbb{R}^d$ ,  $U \subset \mathbb{R}^d$  open, and  $F : U \rightarrow M_{2,2}$  a continuous map such that  $(\mathbf{d} \circ F, \nu)$  is good.*

- (a) *Suppose that  $(\mathbf{f}, \nu)$  is nonplanar for any row or column  $\mathbf{f}$  of  $F$ ; then  $F_*\nu$  is strongly extremal.*
- (b) *Suppose that both  $(F, \nu)$  and  $(F^T, \nu)$  are row-nonplanar; then  $F_*\nu$  is extremal.*

As was mentioned before,  $(\mathbf{d} \circ F, \nu)$  happens to be good when  $\nu$  is Lebesgue and functions  $f_{ij}$  are real analytic. Thus, in particular, the pushforwards of Lebesgue measure by

$$x \mapsto \begin{pmatrix} x & x^2 \\ x^3 & x^4 \end{pmatrix} \quad \text{or} \quad x \mapsto \begin{pmatrix} x & x^2 \\ x^2 & x \end{pmatrix}$$

are strongly extremal, even though the determinant of the first map is identically zero, and the image of the second one is contained in a two-dimensional subspace of  $M_{2,2}$  – and therefore the nonplanarity condition of Theorem 2.1 is violated in both cases. Likewise, the pushforwards of Lebesgue measure by  $x \mapsto \begin{pmatrix} x & x \\ x & x^2 \end{pmatrix}$  or even  $x \mapsto \begin{pmatrix} x & x \\ x & 2x \end{pmatrix}$  are extremal (it is clear that strong extremality fails in the latter cases). The proof will be an illustration of techniques described in §5: we will estimate from below the norms of projections of  $u_{F(\cdot)}\mathbf{w}$  onto  $E_{\mathbf{t}}^+$  uniformly in  $\mathbf{w}$  and  $\mathbf{t}$ .

*Proof of Theorem 7.1.* We will be proving both parts simultaneously, since they are based on the same computation. We need to look through elements of  $\mathcal{W}_\ell$  where  $\ell = 1, 2$  or  $3$ . Since the assumptions on  $F$  are obviously invariant under transposition, the computations for  $\ell = 1$  and  $\ell = 3$  are identical, i.e. dual to each other (see the remark at the end of §3). Thus the two cases to consider correspond to vectors and bi-vectors  $\mathbf{w}$  respectively. As in §4, we will denote the standard basis of  $\mathbb{R}^4$  by  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{v}_1, \mathbf{v}_2\}$ , so that the  $u_Y$ -action is described via (4.2) and (4.4).

First consider  $\ell = 1$  and take

$$\mathbf{w} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + b_1\mathbf{v}_1 + b_2\mathbf{v}_2 \in \mathcal{W}_1.$$

Note that for any  $\mathbf{t} \in \mathcal{A}$ ,  $E_{\mathbf{t}}^+ \cap \bigwedge^1(\mathbb{R}^4)$  is spanned by  $\mathbf{e}_1$  and  $\mathbf{e}_2$ , and therefore

$$\pi_{\mathbf{t}}^+(u_{F(\mathbf{x})}\mathbf{w}) = (a_1 + b_1f_{11}(\mathbf{x}) + b_2f_{12}(\mathbf{x}))\mathbf{e}_1 + (a_2 + b_1f_{21}(\mathbf{x}) + b_2f_{22}(\mathbf{x}))\mathbf{e}_2.$$

Identifying  $\mathbf{e}_1$  with  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $\mathbf{e}_2$  with  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  one can write

$$\pi_{\mathbf{t}}^+u_{F(\mathbf{x})}\mathbf{w} = F(\mathbf{x}) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Since at least one of  $a_i, b_j$  is a nonzero integer, the row-nonplanarity of  $(F, \nu)$  implies that for any ball  $B \subset U$  with  $\nu(B) > 0$  there exists  $c > 0$  such that norms of the above vectors, uniformly in  $\mathbf{w} \in \mathcal{W}_1$  and  $\mathbf{t} \in \mathcal{A}$ , are not less than  $c$  for some  $\mathbf{x} \in \text{supp } \nu \cap B$ . Hence Corollary 5.1 applies. Note that in this case the weaker assumption of part (b) was sufficient to draw the required conclusion.

For the case  $\ell = 2$ , take

$$\mathbf{w} = a\mathbf{e}_1 \wedge \mathbf{e}_2 + b_{11}\mathbf{e}_1 \wedge \mathbf{v}_1 + b_{12}\mathbf{e}_1 \wedge \mathbf{v}_2 + b_{21}\mathbf{e}_2 \wedge \mathbf{v}_1 + b_{22}\mathbf{e}_2 \wedge \mathbf{v}_2 + c\mathbf{v}_1 \wedge \mathbf{v}_2 \in \mathcal{W}_2$$

and write

$$\begin{aligned} u_{F(\cdot)}\mathbf{w} &= (a + b_{11}f_{21} + b_{12}f_{22} - b_{21}f_{11} - b_{22}f_{21} + c \det(F))\mathbf{e}_{\{1,2\}} \\ &\quad + (b_{11} - cf_{12})\mathbf{e}_1 \wedge \mathbf{v}_1 + (b_{12} + cf_{11})\mathbf{e}_1 \wedge \mathbf{v}_2 \\ &\quad + (b_{21} - cf_{22})\mathbf{e}_2 \wedge \mathbf{v}_1 + (b_{22} + cf_{21})\mathbf{e}_2 \wedge \mathbf{v}_2 + c\mathbf{v}_{\{1,2\}} \end{aligned} \tag{7.1}$$

First let us describe the argument in case (b). When  $\mathbf{t} \in \mathcal{R}$ , that is,  $t_1 = t_2 = t_3 = t_4$ , it is easy to see that all the elements  $\mathbf{e}_i \wedge \mathbf{v}_j$  are in  $E_{\mathbf{t}}^+$ . Let  $\pi$  be the orthogonal projection of  $\bigwedge^2(\mathbb{R}^4)$  onto the span of  $\mathbf{e}_1 \wedge \mathbf{v}_1$  and  $\mathbf{e}_2 \wedge \mathbf{v}_1$ . Identifying these with  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , one can write

$$\pi(u_{F(\cdot)}\mathbf{w}) = F(\cdot) \begin{pmatrix} 0 \\ -c \end{pmatrix} + \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}.$$

Thus the desired estimate holds whenever at least one of  $b_{11}, b_{21}, c$  is nonzero. Otherwise, either  $b_{12} = \langle u_{F(\cdot)}\mathbf{w}, \mathbf{e}_1 \wedge \mathbf{v}_2 \rangle$  or  $b_{22} = \langle u_{F(\cdot)}\mathbf{w}, \mathbf{e}_2 \wedge \mathbf{v}_2 \rangle$  or  $a = \langle u_{F(\cdot)}\mathbf{w}, \mathbf{e}_{\{1,2\}} \rangle$  is a nonzero integer, and therefore the estimate of Corollary 5.1 holds in this case as well.

Now turn to part (a). It is not hard to see that for any  $\mathbf{t} \in \mathcal{A}$ , the dimension of  $E_{\mathbf{t}}^+ \cap \bigwedge^2(\mathbb{R}^4)$  is at least three. Specifically, let  $i_0$  be such that  $t_{i_0} = \max_{i=1,\dots,4} t_i$ . If  $i_0 \leq 2$ , then clearly  $\mathbf{e}_{i_0} \wedge \mathbf{v}_i \in E_{\mathbf{t}}^+$ ,  $i = 1, 2$ , and otherwise  $\mathbf{e}_i \wedge \mathbf{v}_{i_0-2} \in E_{\mathbf{t}}^+$ ,  $i = 1, 2$ . In addition,  $\mathbf{e}_{\{1,2\}}$  is clearly also always in  $E_{\mathbf{t}}^+$ . Without loss of generality let us assume that  $i_0 = 1$  (the other cases are treated similarly). Then both  $\mathbf{e}_1 \wedge \mathbf{v}_1$  and  $\mathbf{e}_1 \wedge \mathbf{v}_2$  belong to  $E_{\mathbf{t}}^+$ , so whenever at least one of  $b_{11}, b_{12}, c$  is nonzero, the nonplanarity of  $((f_{11}, f_{12}), \nu)$  implies the desired estimate. Otherwise, the projection of  $u_{F(\cdot)}\mathbf{w}$  onto  $\mathbf{e}_{\{1,2\}}$  is equal to  $a - b_{21}f_{11} - b_{22}f_{21}$ , and one of  $b_{21}, b_{22}, a$  is definitely nonzero; therefore the nonplanarity of  $((f_{11}, f_{21}), \nu)$  applies and finishes the proof.  $\square$

We would like to point out that the nonplanarity conditions of the above theorem are as close to being optimal as the standard nonplanarity assumption on the pair  $(\mathbf{f}, \nu)$  in the case  $\min(m, n) = 1$ . Indeed, if the nonplanarity of  $(\mathbf{f}, \nu)$  is violated by the existence of a nontrivial *integer* linear combination of  $1, f_1, \dots, f_n$  vanishing on  $B \cap \text{supp } \nu$ , then clearly every point of  $\mathbf{f}(B \cap \text{supp } \nu)$  is very well approximable. Likewise, if a nontrivial integer linear combination of 1 and the components of some row or column of  $F$  vanishes on  $B \cap \text{supp } \nu$ , then  $\mathbf{f}(B \cap \text{supp } \nu)$  consists of VWMA matrices. Indeed, if the above is the case for one of the rows of  $F$ , then  $\Pi(F(\cdot)\mathbf{q} + \mathbf{p}) \equiv 0$  on  $B \cap \text{supp } \nu$ ; the same conclusion for one of the columns follows from the transference principle. Similarly, if the row-nonplanarity assumption is violated by the existence of a nonzero integer vector  $\mathbf{q} \in \mathbb{Z}^n$  such that the restriction of the map  $\mathbf{x} \mapsto F(\mathbf{x})\mathbf{q}$  to  $B \cap \text{supp } \nu$  is a constant integer  $\mathbf{p} \in \mathbb{Z}^m$ , then  $F(\cdot)\mathbf{q} - \mathbf{p} \equiv 0$  on  $B \cap \text{supp } \nu$ , and hence obviously  $F(B \cap \text{supp } \nu)$  consists of VWA matrices.

Of course there is a gap between vanishing of *all* linear combinations and non-vanishing of a non-trivial *integer* linear combination; a precise criterion (in the class of Federer measures and good pairs) is likely to involve some Diophantine conditions on the parameterizing coefficients of the smallest affine subspace containing the image of  $F$ , similarly to the results of [K1, K2, Zh],

Looking at Theorem 7.1 one may wonder whether or not it is possible in general to derive strong extremality or at least extremality of  $F_*\nu$  from conditions involving just linear combinations of rows/columns of  $F$ . This turns out not to be the case when  $\max(m, n) > 2$ . Indeed, as we have seen in Corollary 5.2, linear growth of  $g_{\mathcal{R}}\Lambda_Y$  is implied by vanishing of the projection of  $u_Y\mathbf{w}$

for some  $\mathbf{w} \in \mathcal{W}$  onto the space  $E_{\mathbf{t}}^+$ , where  $\mathbf{t} \in \mathcal{R}$  is arbitrary. Next we are going to show that such vanishing conditions can boil down to higher degree polynomial relations between the columns (or rows) of  $Y$ .

For simplicity consider the case  $n = 2$  (similarly one can treat the general case). Fix  $t > 0$  and  $\mathbf{t} = (\frac{t}{m}, \dots, \frac{t}{m}, \frac{t}{2}, \frac{t}{2})$ , and observe that

$$\bigwedge^2(\mathbb{R}^{m+2}) \cap E_{\mathbf{t}}^+ \text{ is spanned by } \mathbf{e}_i \wedge \mathbf{e}_j, \quad 1 \leq j < i \leq m. \quad (7.2)$$

Indeed, unlike the case  $m = 2$  considered in Theorem 7.1, elements  $\mathbf{e}_i \wedge \mathbf{v}_j$  are contracted by  $g_{\mathbf{t}}$ , namely one has  $g_{\mathbf{t}}(\mathbf{e}_i \wedge \mathbf{v}_j) = e^{t(\frac{1}{n} - \frac{1}{2})} \mathbf{e}_i \wedge \mathbf{v}_j$ . Denote by  $\mathbf{y}_1, \mathbf{y}_2$  the columns of  $Y$ . Now take an arbitrary  $\mathbf{w} \in \mathcal{W}_2$ , and denote by  $W$  the plane in  $\mathbb{R}^{m+2}$  corresponding to  $\mathbf{w}$ . Also denote by  $V$  the plane spanned by  $\mathbf{v}_1, \mathbf{v}_2$  and by  $E$  the span of  $\{\mathbf{e}_i : i = 1, \dots, m\}$ . Clearly the following three cases can occur: the orthogonal projection of  $W$  onto  $V$  can have dimension 1, 2 or 0. In the latter case  $\mathbf{w}$  belongs to  $E_{\mathbf{t}}^+$  and is  $u_Y$ -invariant, therefore  $\pi_{\mathbf{t}}^+ u_Y \mathbf{w}$  does not vanish. The other two cases are more interesting.

**Case 1.** If  $W$  projects onto a one-dimensional subspace of  $V$ , one can write  $\mathbf{w} = \mathbf{v} \wedge \mathbf{u}$  where  $\mathbf{v}, \mathbf{u}$  are nonzero integer vectors in  $V$  and  $E$  respectively. In other words (identifying  $E$  with  $\mathbb{R}^m$  as before),

$$\mathbf{w} = \mathbf{u} \wedge (a\mathbf{v}_1 + b\mathbf{v}_2), \quad \text{where } \mathbf{u} \in \mathbb{Z}^m \setminus \{0\}, \quad (a, b) \in \mathbb{Z}^2 \setminus \{0\}.$$

From (4.2) and (7.2) it then follows that

$$\pi_{\mathbf{t}}^+ u_Y \mathbf{w} = \pi_{\mathbf{t}}^+ \mathbf{u} \wedge (a(\mathbf{v}_1 + \mathbf{y}_1) + b(\mathbf{v}_2 + \mathbf{y}_2)) = \mathbf{u} \wedge (a\mathbf{y}_1 + b\mathbf{y}_2).$$

**Conclusion 1:** if a nontrivial integer linear combination of columns of  $Y$  is proportional to an integer vector, then  $Y$  is very well approximable. In particular, if this happens for  $Y = F(\mathbf{x})$  with the coefficient of proportionality being a function of  $\mathbf{x}$ , then  $F(\mathbf{x})$  is VWA for every  $\mathbf{x}$ . Consider for

example  $F(x) = \begin{pmatrix} x & x^2 + x^3 \\ x^2 & x + x^3 \\ x^3 & x + x^2 \end{pmatrix}$ . Each row (resp., column) of  $F$  is a non-

degenerate polynomial map  $\mathbb{R} \rightarrow \mathbb{R}^2$  (resp.,  $\mathbb{R} \rightarrow \mathbb{R}^3$ ). However  $F(x)$  is very well approximable for every  $x$ , since the sum of its columns is equal to  $(x + x^2 + x^3)(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3)$ .

**Case 2.** In the generic situation, when the plane  $W$  projects surjectively onto  $V$ , using Gaussian reduction over integers, one can express  $\mathbf{w}$  as

$$\mathbf{w} = (\mathbf{u}_1 + a\mathbf{v}_1) \wedge (\mathbf{u}_2 + b\mathbf{v}_2), \quad \text{where } \mathbf{u}_1, \mathbf{u}_2 \in \mathbb{Z}^m \setminus \{0\}, \quad a, b \in \mathbb{Z} \setminus \{0\}.$$

Then

$$\pi_{\mathbf{t}}^+ u_Y \mathbf{w} = \pi_{\mathbf{t}}^+ (\mathbf{u}_1 + a(\mathbf{v}_1 + \mathbf{y}_1)) \wedge (\mathbf{u}_2 + b(\mathbf{v}_2 + \mathbf{y}_2)) = (\mathbf{u}_1 + a\mathbf{y}_1) \wedge (\mathbf{u}_2 + b\mathbf{y}_2).$$

**Conclusion 2:** if a integer translate of an integer multiple of a column of  $Y$  is proportional to an integer translate of an integer multiple of the other column, then  $Y$  is very well approximable. For example matrices

$$F_1(x) = \begin{pmatrix} x & x^4 \\ x^2 & x^5 \\ x^3 & x^6 \end{pmatrix} \quad \text{and} \quad F_2(x) = \begin{pmatrix} x & 2x^2 + 3x \\ x^2 & 2x^3 + 2x^2 - x \\ x^3 & 2x^4 + 2x^3 + x \end{pmatrix}$$

are VWA for every  $x$  (even though, as in the previous example, their rows and columns are nondegenerate polynomial maps). This is completely clear as far as  $F_1$  is concerned – its columns are proportional. However it is far less obvious to understand the reason for the non-extremality of  $(F_2)_*\lambda$ , namely, that

$$2 \begin{pmatrix} x \\ x^2 \\ x^3 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 2x^2 + 3x \\ 2x^3 + 2x^2 - x \\ 2x^4 + 2x^3 + x \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \text{ are proportional.}$$

It appears to be a challenging task to devise an algorithm which detects all the aforementioned obstructions to extremality, say for matrices whose elements are integer polynomials in one real variable<sup>3</sup>. This is part of a vague general problem, asked in [Go, §9.1], to describe general conditions which are sufficient for extremality or strong extremality and are ‘close to being optimal’, in the sense of the discussion after Theorem 7.1, within certain class of maps. (The latter theorem, incidentally, settles the problem for  $m = n = 2$  in the class of Federer measures and good pairs.) This circle of problems will be addressed in a forthcoming paper [BKM].

### 8. CONCLUDING REMARKS AND OPEN QUESTIONS

**8.1. Improving Dirichlet’s Theorem.** Another application of techniques developed in this paper yields a generalization of a theorem from [KW2], which in its turn has generalized many earlier results. The starting point for the general set-up of the problem is a multi-parameter form of Dirichlet’s Theorem<sup>4</sup>: for any system of linear forms  $Y_1, \dots, Y_m$  (rows of  $Y \in M_{m,n}$ ) and for any  $\mathbf{t} \in \mathcal{A}$  there exist solutions  $\mathbf{q} = (q_1, \dots, q_n) \in \mathbb{Z}^n \setminus \{0\}$  and  $\mathbf{p} = (p_1, \dots, p_m) \in \mathbb{Z}^m$  of

$$\begin{cases} |Y_i \mathbf{q} - p_i| < e^{-t_i}, & i = 1, \dots, m \\ |q_j| \leq e^{t_{m+j}}, & j = 1, \dots, n. \end{cases} \tag{8.1}$$

Then, given an unbounded subset  $\mathcal{T}$  of  $\mathcal{A}$  and positive  $\varepsilon < 1$ , one says that *Dirichlet’s Theorem can be  $\varepsilon$ -improved for  $Y$  along  $\mathcal{T}$* , or  $Y \in \text{DI}_\varepsilon(\mathcal{T})$ , if there is  $T$  such that for every  $\mathbf{t} = (t_1, \dots, t_{m+n}) \in \mathcal{T}$  with  $t > T$ , the inequalities

$$\begin{cases} |Y_i \mathbf{q} - p_i| < \varepsilon e^{-t_i}, & i = 1, \dots, m \\ |q_j| < \varepsilon e^{t_{m+j}}, & j = 1, \dots, n, \end{cases} \tag{8.2}$$

i.e., (8.1) with the right hand side terms multiplied by  $\varepsilon$ , have nontrivial integer solutions. Using an elementary argument dating back to Khintchine, one can show that for any  $m, n$  and any unbounded  $\mathcal{T} \subset \mathcal{A}$ ,  $\text{DI}_\varepsilon(\mathcal{T})$  has Lebesgue measure zero as long as  $\varepsilon < 1/2$ . In [KW2] a similar statement was proved for pushforwards of Federer measures to  $\mathbb{R}^n \cong M_{1,n}$  by continuous maps  $\mathbf{f}$ .

---

<sup>3</sup>Arguing similarly to the proof of Theorem 7.1, it is possible to show that the obstructions listed in Cases 1 and 2 above, together with the linear ones taken care of by assuming the row-nonplanarity of  $F$  and  $F^T$ , can be used to generate a complete list of obstructions within the class of Federer measures and good pairs when  $n = 2$  and  $m = 3$ . For higher dimensions the situation is more complicated, that is, one can produce non-trivial obstructions by considering  $u_Y$ -action on  $\mathcal{W}_p$  for  $p \geq 3$ .

<sup>4</sup>In [Sh2] it was referred to as Dirichlet-Minkowski Theorem.

Namely, let  $\nu$  be a  $D$ -Federer measure on  $\mathbb{R}^d$ ,  $U \subset \mathbb{R}^d$  open, and  $\mathbf{f} : U \rightarrow \mathbb{R}^n$  continuous such that the pair  $(\mathbf{f}, \nu)$  is  $(C, \alpha)$ -good and nonplanar. Then it was proved in [KW2, Theorem 1.5] that  $\mathbf{f}_*\nu(\text{DI}_\varepsilon(\mathcal{T})) = 0$  for any unbounded  $\mathcal{T} \subset \mathcal{A}$  and any  $\varepsilon < \varepsilon_0$ , where  $\varepsilon_0$  depends only on  $d, n, C, \alpha, D$ . Note that here one needs a uniform version of the definition of a good pair:  $(\mathbf{f}, \nu)$  is said to be  $(C, \alpha)$ -good if for  $\nu$ -a.e.  $x$  there exists a neighborhood  $U$  of  $x$  such that  $(\mathbf{f}, \nu)$  is  $(C, \alpha)$ -good on  $U$ .

We refer the reader to [KW2] and [Sh1] for a history of the subject, which had been initiated in [DS1, DS2] for the case  $\mathcal{T} = \mathcal{R}$ , that is, dealing with Dirichlet's Theorem in its classical form. Also note that recent results of Shah [Sh1, Sh2] show that in many cases, with  $\nu = \lambda$  and  $\mathbf{f}$  real analytic, a similar result holds with  $\varepsilon_0 = 1$ .

It turns out that a combination of methods of [KW2] and the present paper can produce the following generalization to the case  $\min(m, n) > 1$ :

**Theorem 8.1.** *For any  $d, m, n \in \mathbb{N}$  and  $C, \alpha, D > 0$  there exists  $\varepsilon_0$  with the following property. Let  $U$  be an open subset of  $\mathbb{R}^d$ ,  $F : U \rightarrow M_{m,n}$  continuous and  $\nu$  a measure on  $U$ . Assume that  $\nu$  is  $D$ -Federer, and  $(\mathbf{d} \circ F, \nu)$  is  $(C, \alpha)$ -good and nonplanar. Then  $F_*\nu(\text{DI}_\varepsilon(\mathcal{T})) = 0$  for any unbounded  $\mathcal{T} \subset \mathcal{A}$  and any  $\varepsilon < \varepsilon_0$ .*

It can be shown, by combining the argument of §6 with [KW2, Proposition 4.4], that a uniform version of Proposition 2.2 holds; that is, for  $\nu = \lambda$  and  $F$  as in (1.3) one can choose  $C, \alpha$  such that  $(\mathbf{d} \circ F, \nu)$  is  $(C, \alpha)$ -good; thus the conclusion of the above theorem holds for  $F_*\lambda$  as in Theorem 1.1, with some positive  $\varepsilon_0$ . Details and further results along these lines will appear in a forthcoming paper. It seems natural to conjecture that for  $\nu = \lambda$  and real analytic  $F$  such that  $\mathbf{d} \circ F$  is nonplanar an analogue of Shah's result holds, that is, sets  $\text{DI}_\varepsilon(\mathcal{T})$  are  $F_*\lambda$ -null for any  $\varepsilon < 1$ .

**8.2. Inhomogeneous Diophantine problems.** A method allowing to transfer results on extremality and strong extremality of measures on  $M_{m,n}$  to inhomogeneous Diophantine approximation has been recently developed by Beresnevich and Velani in [BV]. In the inhomogeneous set-up, instead of systems of linear forms given by  $Y \in M_{m,n}$ , one considers systems of *affine* forms  $(Y, \mathbf{z})$ , that is, maps  $\mathbf{q} \mapsto Y\mathbf{q} + \mathbf{z}$  where  $Y \in M_{m,n}$  and  $\mathbf{z} \in \mathbb{R}^m$ . Generalizing the homogeneous setting by identifying  $Y$  with  $(Y, 0)$ , let us say that  $(Y, \mathbf{z})$  is VWA if for some  $\delta > 0$  there are infinitely many  $\mathbf{q} \in \mathbb{Z}^n$  such that

$$\|Y\mathbf{q} + \mathbf{z} - \mathbf{p}\| < \|\mathbf{q}\|^{-n/m-\delta} \text{ for some } \mathbf{p} \in \mathbb{Z}^m,$$

and that it is VWMA if for some  $\delta > 0$  there are infinitely many  $\mathbf{q} \in \mathbb{Z}^n$  such that

$$\Pi(Y\mathbf{q} + \mathbf{z} - \mathbf{p}) < \Pi_+(\mathbf{q})^{-(1+\delta)} \text{ for some } \mathbf{p} \in \mathbb{Z}^m.$$

From the Borel-Cantelli Lemma it is clear that for any  $\mathbf{z} \in \mathbb{R}^m$  the set

$$\text{VWMA}_{\mathbf{z}} \stackrel{\text{def}}{=} \{Y \in M_{m,n} : (Y, \mathbf{z}) \text{ is VWMA}\}$$

has zero Lebesgue measure; and, since VWA obviously implies VWMA, the same is true for

$$\text{VWA}_{\mathbf{z}} \stackrel{\text{def}}{=} \{Y \in M_{m,n} : (Y, \mathbf{z}) \text{ is VWA}\}.$$

Following [BV], let us say that a measure  $\mu$  on  $M_{m,n}$  is *inhomogeneously extremal* (resp., *inhomogeneously strongly extremal*) if  $\mu(\text{VWA}_{\mathbf{z}}) = 0$  for any  $\mathbf{z} \in \mathbb{R}^m$  (resp.,  $\mu(\text{VWMA}_{\mathbf{z}}) = 0 \forall \mathbf{z} \in \mathbb{R}^m$ ).

One of the main results of [BV] is the following transference phenomenon: under some regularity conditions on  $\mu$ , the inhomogeneous properties defined above are equivalent to their (a priori weaker) homogeneous analogues. Specifically, Beresnevich and Velani define the class of measures on  $M_{m,n}$  which they call *contracting almost everywhere* and a subclass of measures *strongly contracting almost everywhere* (we refer the reader to [BV] for precise definitions). According to [BV, Theorem 1], a (strongly) contracting almost everywhere measure on  $M_{m,n}$  is (strongly) extremal if and only if it is inhomogeneously (strongly) extremal. Using this, [BV] establishes inhomogeneous strong extremality of many measures proved earlier to be strongly extremal, such as  $\mathbf{f}_* \lambda$  where  $\mathbf{f}$  is as in Theorem 1.1, or, more generally, arbitrary friendly measures on  $\mathbb{R}^n$ .

As remarked at the end of [BV], ‘any progress on the homogeneous extremality problem can be transferred over to the inhomogeneous setting’. Indeed, many measures on  $M_{m,n}$  discussed in the present paper can be shown to be strongly contracting almost everywhere. Here is an example: suppose that for any  $i = 1, \dots, m$  we are given a contracting measure  $\mu_i$  on  $\mathbb{R}^n$ , where the latter space is identified with the space of  $i$ th rows of  $Y \in M_{m,n}$ . Then it is clear from the definitions that  $\mu_1 \times \dots \times \mu_m$  is strongly contracting. Therefore Theorem 1.2 and the results of [BV] imply

**Theorem 8.2.** *Let  $F$  be as in Theorem 1.2. Then  $F_* \lambda$  is inhomogeneously strongly extremal.*

This motivates a problem of checking contracting and strongly contracting properties of other measures on  $M_{m,n}$  proved in the present paper to be extremal or strongly extremal. For example, it would be interesting to understand under what conditions on a smooth submanifold of  $M_{m,n}$  its Riemannian volume measure is (strongly) contracting (Theorem 4 of [BV] deals with the case  $n = 1$ ).

**8.3. What is next?** Here is an incomplete list of other possible directions for further research:

8.3.1. Can one characterize extremal or strongly extremal affine subspaces of  $M_{m,n}$  in the spirit of [K1] which settled the problem for  $\min(m, n) = 1$ ? Or, more generally, subspaces with a given Diophantine exponent following [K2]?

8.3.2. Is it possible to obtain some Khintchine-type results for smooth submanifolds of  $M_{m,n}$  with  $\min(m, n) > 1$ ? That is, study inequalities of type (1.1) with a power of the norm of  $\mathbf{q}$  in the right hand side replaced by a general non-increasing function of  $\|\mathbf{q}\|$  satisfying the convergence or divergence conditions of the Khintchine-Groshev Theorem. Note that, as of now,

the convergence case of the problem is not settled even for nondegenerate submanifolds of  $M_{m,1}$  where  $m > 2$ ; however recent divergence theorems [B, BDV] gives a hope of possible extensions to curves in the space of matrices. Likewise, convergence-type results of [VV] for planar curves give a hope for a complete Khintchine-type theorem for smooth ‘sufficiently nondegenerate’ (in the spirit of Theorem 7.1) smooth curves in  $M_{2,2}$ .

8.3.3. Following [KT] and [G], it should not be difficult to extend the results of the present paper to metric Diophantine problems over non-Archimedean local fields and their products.

## REFERENCES

- [B] V. Beresnevich, *Rational points near manifolds and metric Diophantine approximation*, [arXiv:0904.0474](#).
- [BD] V. I. Bernik and M. M. Dodson, *Metric Diophantine approximation on manifolds*, Cambridge University Press, Cambridge, 1999.
- [BDV] V. Beresnevich, H. Dickinson and S. Velani, *Diophantine approximation on planar curves and the distribution of rational points*, *Ann. of Math.* **166** (2007), 367–426.
- [BKM] V. Beresnevich, D. Kleinbock and G. A. Margulis, in preparation.
- [BV] V. Beresnevich and S. Velani, *An Inhomogeneous Transference Principle and Diophantine Approximation*, [arXiv:0802.1837](#).
- [C] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge University Press, New York, 1957.
- [D1] S. G. Dani, *On invariant measures, minimal sets, and a lemma of Margulis*, *Invent. Math.* **51** (1979), 239–260.
- [D2] ———, *On orbits of unipotent flows on homogeneous spaces*, *Ergodic Theory Dynam. Systems* **4** (1984), 25–34.
- [D3] ———, *Divergent trajectories of flows on homogeneous spaces and diophantine approximation*, *J. Reine Angew. Math.* **359** (1985), 55–89.
- [D4] ———, *On orbits of unipotent flows on homogeneous spaces, II*, *Ergodic Theory Dynam. Systems* **6** (1986), 167–182.
- [Do] M. M. Dodson, *Hausdorff dimension, lower order and Khintchine’s theorem in metric Diophantine approximation*, *J. Reine Angew. Math.* **432** (1992), 69–76.
- [DS1] H. Davenport and W. M. Schmidt, *Dirichlet’s theorem on diophantine approximation*, in: *Symposia Mathematica*, Vol. IV (INDAM, Rome, 1968/69), pp. 113–132, 1970.
- [DS2] ———, *Dirichlet’s theorem on diophantine approximation. II*, *Acta Arith.* **16** (1969/1970) 413–424.
- [G] A. Ghosh, *Metric Diophantine approximation over a local field of positive characteristic*, *J. Number Theory*, **124** (2007), 454–469.
- [Go] A. Gorodnik, *Open problems in dynamics and related fields*, *J. Mod. Dyn.* **1** (2007), no. 1, 1–35.
- [K1] D. Kleinbock, *Extremal subspaces and their submanifolds*, *Geom. Funct. Anal.* **13** (2003), no. 2, 437–466.
- [K2] ———, *An extension of quantitative nondivergence and applications to Diophantine exponents*, *Trans. Amer. Math. Soc.* **360** (2008), 6497–6523.
- [K3] ———, *Diophantine exponents of measures and homogeneous dynamics*, *Pure Appl. Math. Q.* **4** (2008), 81–97.
- [K4] ———, *An ‘almost all versus no’ dichotomy in homogeneous dynamics and Diophantine approximation*, [arXiv:0904.1614](#).
- [KLW] D. Kleinbock, E. Lindenstrauss and B. Weiss, *On fractal measures and diophantine approximation*, *Selecta Math.* **10** (2004), 479–523.

- [KM1] D. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. Math. **148** (1998), 339–360.
- [KM2] ———, *Logarithm laws for flows on homogeneous spaces*, Invent. Math. **138** (1999), 451–494.
- [Ko1] E. Kovalevskaya, *Simultaneously extremal manifolds*, Dokl. Akad. Nauk BSSR **31** (1987), 405–408 (Russian).
- [Ko2] ———, *Jointly extremal manifolds*, Mat. Zametki **41** (1987), 38.
- [Ko3] ———, *Strongly jointly extremal manifolds*, Vestsi Akad. Navuk BSSR Ser. Fiz.-Mat. Navuk **6** (1987), 16–19 (Russian).
- [KT] D. Kleinbock and G. Tomanov, *Flows on  $S$ -arithmetic homogeneous spaces and applications to metric Diophantine approximation*, Comm. Math. Helv. **82** (2007), 519–58.
- [KW1] D. Kleinbock and B. Weiss, *Badly approximable vectors on fractals*, Israel J. Math. **149** (2005), 137–170.
- [KW2] ———, *Dirichlet’s theorem on diophantine approximation and homogeneous flows*, J. Mod. Dyn. **2** (2008), 43–62.
- [M] K. Mahler, *Über das Mass der Menge aller  $S$ -Zahlen*, Math. Ann. **106**, (1932), 131–139.
- [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.
- [MU] D. Mauldin and M. Urbanski, *The doubling property of conformal measures of infinite iterated function systems*, J. Number Th. **102** (2003), 23–40.
- [PV] A. Pollington and S. Velani, *Metric Diophantine approximation and ‘absolutely friendly’ measures*, Selecta Math. **11** (2005), 297–307.
- [R] M. S. Raghunathan, *Discrete Subgroups of Lie Groups*, Springer, Berlin, 1972.
- [Sh1] N. A. Shah, *Equidistribution of expanding translates of curves and Dirichlet’s theorem on diophantine approximation*, arXiv:0802.3278.
- [Sh2] ———, *Expanding translates of curves and Dirichlet-Minkowski theorem on linear forms*, arXiv:0804.1424.
- [Sp1] V. G. Sprindžuk, *More on Mahler’s conjecture*, Dokl. Akad. Nauk SSSR, 155:54–56, 1964.
- [Sp2] ———, *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.
- [Sp3] ———, *Achievements and problems in Diophantine approximation theory*, Russian Math. Surveys **35** (1980), 1–80.
- [SU] B. Stratmann and M. Urbanski, *Diophantine extremality of the Patterson measure*, Math. Proc. Cambridge Phil. Soc. **140** (2006), 297–304.
- [SW] W.M. Schmidt and Y. Wang, *A note on a transference theorem of linear forms*, Sci. Sinica **22** (1979), 276–280.
- [U] M. Urbanski, *Diophantine approximation of self-conformal measures*, J. Number Th. **110** (2005), no. 2, 219–235.
- [VV] R. Vaughan and S. Velani, *Diophantine approximation on planar curves: the convergence theory*, Invent. Math. **166** (2006), no. 1, 103–124.
- [W] J. Wang, *Diophantine approximation of linear forms*, Ph.D. Thesis, Brandeis University, 2008.
- [Zh] Y. Zhang, *Diophantine Exponents of Affine Subspaces: The Simultaneous Approximation Case*, arXiv:0809.0386, J. Number Theory, to appear.

BRANDEIS UNIVERSITY, WALTHAM MA 02454-9110 kleinboc@brandeis.edu

YALE UNIVERSITY, NEW HAVEN, CT 06520 margulismath.yale.edu

BRANDEIS UNIVERSITY, WALTHAM MA 02454-9110 junbo@brandeis.edu