KHINTCHINE-TYPE THEOREMS ON MANIFOLDS: THE CONVERGENCE CASE FOR STANDARD AND MULTIPLICATIVE VERSIONS.

V. Bernik, D. Kleinbock and G. A. Margulis

September 1999

Abstract. An analogue of the convergence part of the Khintchine-Groshev Theorem, as well as its multiplicative version, is proved for non-degenerate smooth curves in $\mathbb{R}^n$, as well as non-degenerate submanifolds which can be foliated by such curves (e.g. analytic). The proof combines methods from metric number theory with a new approach involving the geometry of lattices in Euclidean spaces.

1. Introduction

Basics on Diophantine approximation. We start by recalling several basic facts from metric theory of Diophantine approximation. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we let $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i$. Unless otherwise indicated, the norm $\|\mathbf{x}\|$ of a vector $\mathbf{x} \in \mathbb{R}^n$ will stand for $\|\mathbf{x}\| = \max_{1 \le i \le n} |x_i|$. In some cases however we will work with the Euclidean norm $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$, keeping the same notation. This distinction will be clearly emphasized to avoid confusion. We will denote by $\mathbb{R}^n_1$ the set of unit vectors in $\mathbb{R}^n$ (with respect to the Euclidean norm).

We will use the notation $\lfloor x \rceil$ for the distance between $x \in \mathbb{R}$ and the closest integer, $\lfloor x \rceil \stackrel{\text{def}}{=} \min_{k \in \mathbb{Z}} |x - k|$. (It is quite customary to use $\|x\|$ instead, but we are not going to do this in order to save the latter notation for norms in vector spaces.) If $B \subset \mathbb{R}$, we let $|B|$ stand for the Lebesgue measure of B.

Let ψ be a positive function defined on the set of positive integers. Following [KM2], say that a vector $\mathbf{y} \in \mathbb{R}^n$ (viewed as a linear form in n variables[1]) is *ψ-approximable*[2], to be abbreviated as ψ-*A*, if there are infinitely many $\mathbf{q} \in \mathbb{Z}^n$ such that

$$\lfloor \mathbf{q} \cdot \mathbf{y} \rceil \le \psi(\|\mathbf{q}\|^n) . \tag{1.1}$$

The work of the second named author was supported in part by NSF Grant DMS-9704489, and that of the third named author by NSF Grant DMS-9800607.

[1]See [KM2] for a general definition of a system of linear forms being ψ-approximable. In particular, a (column) vector $\mathbf{y} \in \mathbb{R}^n$ is called ψ-approximable if there exist infinitely many $q \in \mathbb{Z}$ such that $\|q\mathbf{y} + \mathbf{p}\|^n \le \psi(|q|)$ for some $\mathbf{p} \in \mathbb{Z}^n$. Such a setting is usually referred to as the *simultaneous approximation* of the components of $\mathbf{y}$ by rationals, and the counterparts of the main problems settled in this paper remain open.

[2]We are grateful to M. M. Dodson for a permission to modify his terminology introduced in [Do]. In our opinion, the use of (1.1) instead of the traditional $\lfloor \mathbf{q} \cdot \mathbf{y} \rceil \le \psi(\|\mathbf{q}\|)$ makes the structure more transparent and less dimension-dependent, and simplifies the connection (cf. §5 below) with homogeneous flows. See [KM1, KM2] for justification.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

It is a straightforward consequence of the Borel-Cantelli Lemma that almost all $\mathbf{y} \in \mathbb{R}^n$ are not ψ-approximable if

$$\sum_{k=1}^{\infty} \psi(k) < \infty \,. \tag{1.2}$$

and conversely, by the Khintchine-Groshev Theorem (see [G] or [Sp3, Chapter I, Theorem 12]), almost all $\mathbf{y} \in \mathbb{R}^n$ are ψ-approximable if the series $\sum_{k=1}^{\infty} \psi(k)$ diverges (these two statements will be referred to as the convergence and divergence parts of the theorem). An example: almost all $\mathbf{y} \in \mathbb{R}^n$ are not VWA (*very well approximable*, see [S2, Chapter IV, §5]); the latter is defined to be ψ_ε-approximable for some positive ε, with $\psi_\varepsilon(k) \stackrel{\text{def}}{=} k^{-(1+\varepsilon)}$.

In this paper we also discuss another approach to metric Diophantine approximation in $\mathbb{R}^n$. For an integer vector $\mathbf{q}$, let us consider

$$\Pi_+(\mathbf{q}) \stackrel{\text{def}}{=} \prod_{i=1}^{n} \max(|q_i|, 1) \,,$$

i.e. the absolute value of the product of all the nonzero coordinates of $\mathbf{q}$. Now, for ψ as above, say that $\mathbf{y} \in \mathbb{R}^n$ is *ψ-MA* (*ψ-multiplicatively approximable*) if there are infinitely many $\mathbf{q} \in \mathbb{Z}^n$ such that

$$\lfloor \mathbf{q} \cdot \mathbf{y} \rfloor \leq \psi\big(\Pi_+(\mathbf{q})\big) \,. \tag{1.1M}$$

It is not difficult to show using Borel-Cantelli that almost all $\mathbf{y} \in \mathbb{R}^n$ are not ψ-MA if

$$\sum_{k=1}^{\infty} (\log k)^{n-1} \psi(k) < \infty \,. \tag{1.2M}$$

Clearly $\Pi_+(\mathbf{q})$ is not greater than $\|\mathbf{q}\|^n$, therefore, assuming ψ is non-increasing, any ψ-approximable $\mathbf{y}$ is automatically ψ-MA. For example, one can define *very well multiplicatively approximable* (*VWMA*) points to be ψ_ε-multiplicatively approximable for some positive ε, with ψ_ε as above; it follows that almost all $\mathbf{y} \in \mathbb{R}^n$ are not VWMA.

Diophantine approximation on manifolds. Much more intricate questions arise if one restricts $\mathbf{y}$ to lie on a submanifold M of $\mathbb{R}^n$. In 1932 K. Mahler [M] conjectured that almost all points on the curve

$$\{(x, x^2, \dots, x^n) \mid x \in \mathbb{R}\} \tag{1.3}$$

are not VWA. V. Sprindžuk's proof of this conjecture (see [Sp1, Sp2]) has eventually led to the development of a new branch of metric number theory, usually referred to as "Diophantine approximation with dependent quantities" or "Diophantine approximation on manifolds". In particular, Sprindžuk's result was improved by A. Baker [B1] in 1966: he showed that if ψ is a positive non-increasing function such that

$$\sum_{k=1}^{\infty} \frac{\psi(k)^{1/n}}{k^{1-1/n}} < \infty \,, \tag{1.4}$$

then almost all points on the curve (1.3) are not ψ-approximable. Baker conjectured that (1.4) could be replaced by the optimal condition (1.2); this conjecture was proved later by V. Bernik [Bern]. And recently V. Beresnevich [Be1] proved the complementary divergence case: namely, that almost all points on the curve (1.3) are ψ-A whenever $\sum_{k=1}^{\infty} \psi(k) = \infty$. That is, the complete analogue of the Khintchine-Groshev Theorem for the curve (1.3) has been established. As for the multiplicative approximation, it was conjectured by A. Baker in his book [B2] that almost all points on the curve (1.3) are not VWMA; the validity of this conjecture for $n \le 4$ was verified in 1997 by V. Bernik and V. Borbat [BB].

Since the mid-sixties, a lot of efforts have been directed to obtaining similar results for larger classes of smooth submanifolds of $\mathbb{R}^n$. A new method, based on combinatorics of the space of lattices, was developed in 1998 in the paper [KM1] by Kleinbock and Margulis. Let us introduce the following definition: if V is an open subset of $\mathbb{R}^d$, say that an n-tuple $\mathbf{f} = (f_1, \dots, f_n)$ of C^k functions $V \mapsto \mathbb{R}$ is *non-degenerate at* $x \in V$ if the space $\mathbb{R}^n$ is spanned by partial derivatives of $\mathbf{f}$ at x of order up to k. If $M \subset \mathbb{R}^n$ is a d-dimensional C^k submanifold, we will say that M is *non-degenerate at* $\mathbf{y} \in M$ if any (equivalently, some) diffeomorphism $\mathbf{f}$ between an open subset V of $\mathbb{R}^d$ and a neighborhood of $\mathbf{y}$ in M is non-degenerate at $\mathbf{f}^{-1}(\mathbf{y})$. We will say that $\mathbf{f} : V \to \mathbb{R}^n$ (resp. $M \subset \mathbb{R}^n$) is *non-degenerate* if it is non-degenerate at almost every point of V (resp. M, in the sense of the natural measure class on M).

Theorem A [KM1]. *Let M be a non-degenerate C^k submanifold of $\mathbb{R}^n$. Then almost all points of M are not VWMA (hence not VWA as well).*

In particular, the aforementioned multiplicative conjecture of Baker follows from this theorem. Note also that if the functions $f_1, \dots, f_n$ are analytic and V is connected, the non-degeneracy of $\mathbf{f}$ is equivalent to the linear independence of $1, f_1, \dots, f_n$ over $\mathbb{R}$; in this setting the above statement was conjectured by Sprindžuk [Sp4, Conjectures H_1, H_2]. On the other hand, in the case $d = 1$ the non-degeneracy of a curve parametrized by $\mathbf{f} : V \mapsto \mathbb{R}^n$, where $V \subset \mathbb{R}$ is an interval, is equivalent to non-vanishing of the *Wronskian*

$$\begin{vmatrix} f_1'(x) & f_2'(x) & \dots & f_n'(x) \\ f_1''(x) & f_2''(x) & \dots & f_n''(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n)}(x) & f_2^{(n)}(x) & \dots & f_n^{(n)}(x) \end{vmatrix}$$

for almost all $x \in V$. For example, when $n = 2$ this means that the curve $\big(f_1(x), f_2(x)\big)$ has non-zero curvature for almost all x. W. Schmidt [S1] proved in 1964 that almost all points on such a curve are not VWA; later Sprindžuk [Sp3] obtained the multiplicative strengthening. Recently the complete analogue of the Khintchine-Groshev Theorem for non-degenerate planar curves was obtained in [BDD] (the convergence part) and [BBDD] (the divergence part), and then Beresnevich [Be2–4] extended these results to non-degenerate C^4 curves in $\mathbb{R}^3$. For other special cases see the references in [Sp4] and also [DRV1–3].

Main results and the structure of the paper. The primary goal of the present paper is to obtain a Khintchine-type analogue of Theorem A. More precisely, we specialize to the case $d = 1$ and prove the following

Theorem 1.1. *Let M be a non-degenerate C^n curve in $\mathbb{R}^n$, and let $\psi : \mathbb{N} \mapsto (0, \infty)$ be a non-increasing function. Then:*

(S) *assuming* (1.2), *almost all points of M are not ψ-A;*
(M) *assuming* (1.2M), *almost all points of M are not ψ-MA.*

Observe that the statement (S) gives the convergence part of the Khintchine-Groshev Theorem for non-degenerate C^n curves in $\mathbb{R}^n$, and, in particular, improves the aforementioned results for $n = 2, 3$ by cutting down the number of needed derivatives by one.

Corollary 1.2. *Let ψ be as in Theorem 1.1, and let M be a C^n submanifold of $\mathbb{R}^n$ such that*

$$\text{for a.e. } \mathbf{y} \in M \text{ there exists a curve non-degenerate at } \mathbf{y} \text{ and contained in } M\,. \tag{1.5}$$

Then the statements (S) *and* (M) *hold. In particular, these statements hold for non-degenerate real analytic submanifolds.*

Proof. Let $\mathbf{f} : V \to \mathbb{R}^n$, $V \subset \mathbb{R}^d$, be a diffeomorphic parametrization of M, and suppose that for some $x \in V$ there exists a C^n curve $\varphi : (-\varepsilon, \varepsilon) \mapsto V$, $\varphi(0) = x$, such that the n-tuple $\mathbf{f} \circ \varphi : (-\varepsilon, \varepsilon) \mapsto \mathbb{R}^n$ is non-degenerate at 0. Then by the continuity of the derivatives of $\mathbf{f}$ there exists a neighborhood U of $0 \in \mathbb{R}^d$ and $\delta > 0$ such that for any translation $t \mapsto \varphi(t) + a$, $a \in U$, of this curve, the n-tuple $t \mapsto \mathbf{f}\big(\varphi(t) + a\big)$ is non-degenerate at every t with $|t| < \delta$. This way, assuming (1.5), we obtain a foliation of M at almost every point by non-degenerate curves, which reduces Corollary 1.2 to Theorem 1.1. To finish the proof it remains to mention that the validity of (1.5) for real analytic submanifolds is shown in [Sp4, §3] (see also [P] for a related "slicing" technique). $\square$

It seems plausible to conjecture that (1.5) holds for any non-degenerate C^n submanifold $M \subset \mathbb{R}^n$. See §6 for a discussion of possible further improvements in this direction, and for other related problems, including the complementary divergence case, which is going to be the subject of a forthcoming paper.

Our proof of Theorem 1.1 is based on carefully measuring sets of solutions of certain Diophantine inequalities. Specifically, we will fix an interval $B \subset \mathbb{R}$ and look at the set of all $x \in B$ for which there exists an integer vector $\mathbf{q}$ in a certain range such that the value of the function $F(x) = \mathbf{q} \cdot \mathbf{f}(x)$ is close to an integer. Our estimates will require considering two special cases: when the absolute value of the derivative $F'(x) = \mathbf{q} \cdot \mathbf{f}'(x)$ is big, or, respectively, not very big. More precisely, in §2 we show that, by means of straightforward measure computations, Theorem 1.1 reduces to the following two theorems:

Theorem 1.3. *Let an interval $B \subset \mathbb{R}$ and functions $\mathbf{f} = (f_1, \dots, f_n) \in C^2(B)$ be given. Fix $\delta > 0$ and define*

$$L = \max_{1 \le j \le n,\, x \in B} |f_j''(x)|\,. \tag{1.6a}$$

Then for every $\mathbf{q} \in \mathbb{Z}^n$ such that

$$\|\mathbf{q}\| \ge \frac{1}{4nL|B|^2}\,, \tag{1.6b}$$

the set of solutions $x \in B$ of the inequalities

$$\begin{cases} \lfloor \mathbf{q} \cdot \mathbf{f}(x) \rceil < \delta \\ |\mathbf{q} \cdot \mathbf{f}'(x)| \ge \sqrt{nL\|\mathbf{q}\|} \end{cases} \tag{1.6c}$$

has measure at most $32\delta|B|$.

Theorem 1.4. *Let $V \subset \mathbb{R}$ be an interval, $x_0 \in V$, and let $\mathbf{f} = (f_1, \ldots, f_n)$ be an n-tuple of C^n functions on V which is non-degenerate at x_0. Then there exists a subinterval B of V containing x_0 and constants $E > 0$ and $0 < \rho < 1$ such that for any choice of $\varkappa \geq n$, $\delta, K > 0$ and $Q_1, \ldots, Q_n \geq 1$ subject to the constraint*

$$\delta^{\varkappa} \leq \frac{KQ_1 \cdot \ldots \cdot Q_n}{\max_i Q_i} \leq \frac{\rho^{\varkappa+1}}{\delta}, \tag{1.7a}$$

the set

$$\left\{x \in B \;\middle|\; \exists\, \mathbf{q} \in \mathbb{Z}^n \smallsetminus \{0\} \text{ such that } \begin{cases} \lfloor \mathbf{q} \cdot \mathbf{f}(x) \rfloor < \delta \\ |\mathbf{q} \cdot \mathbf{f}'(x)| < K \\ |q_i| < Q_i, \quad i = 1, \ldots, n \end{cases} \right\} \tag{1.7b}$$

has measure at most $E\left(\frac{\delta K Q_1 \cdot \ldots \cdot Q_n}{\max_i Q_i}\right)^{\frac{1}{(\varkappa+1)(2n-1)}} |B|$.

Theorem 1.3, roughly speaking, says that a function with big first derivative and not very big second derivative cannot have values very close to integers on a set of big measure. It is proved in §3 using an argument which is apparently originally due to Bernik, and is implicitly contained in one of the steps of the paper [BDD].

As for Theorem 1.4, it is done by a modification of a method from [KM1] involving the geometry of lattices in Euclidean spaces. This method is exposed in §5, and, to make the paper self-contained, the main technical result from [KM1] is proved in the Appendix. We also rely on the notion of (C, α)-good functions introduced in [KM1]. This concept is reviewed in §4, where we prove that certain functions arising in the proof of Theorem 1.4 are (C, α)-good for suitable C, α. The last section of the paper is devoted to several concluding remarks, applications and some open questions. In particular, there we discuss the complementary divergence case of Theorem 1.1, and present applications involving approximation of zero by values of functions and their derivatives.

2. Reduction of Theorem 1.1 to Theorems 1.3 and 1.4

In this section we prove our main result, Theorem 1.1, modulo the measure estimates of Theorems 1.3 and 1.4.

Let us begin with

Proof of Part (S). Recall that we are given an open interval $V \subset \mathbb{R}$, an n-tuple $\mathbf{f} = (f_1, \ldots, f_n)$ of C^n functions on V, and a non-increasing function ψ satisfying (1.2). Take any $x_0 \in V$ such that $\mathbf{f}$ is non-degenerate at x_0, and consider the subinterval B of V given by Theorem 1.4. We are going to prove that for a.e. $x \in B$ there are at most finitely many $\mathbf{q} \in \mathbb{Z}^n$ with

$$\lfloor \mathbf{q} \cdot \mathbf{f}(x) \rfloor < \psi(\|\mathbf{q}\|^n)\,. \tag{2.1}$$

For a nonnegative integer t, denote by $A(t)$ the set of $x \in V$ such that the inequality (2.1) holds for some $\mathbf{q}$ with

$$2^t \leq \|\mathbf{q}\| < 2^{t+1}\,. \tag{2.2}$$

In view of the Borel-Cantelli Lemma, it suffices to show that $\sum_{t=0}^{\infty} |A(t)| < \infty$. Observe that (2.1), (2.2) and the monotonicity of ψ imply that any solution $(x, \mathbf{q})$ of (2.1) and (2.2) satisfies

$$\lfloor \mathbf{q} \cdot \mathbf{f}(x) \rfloor < \psi(2^{nt}) \,. \tag{2.3}$$

Now let us cover the set $A(t)$ by the union of two sets as follows. Denote by $A_{\geq}(t)$ the set of $x \in B$ such that there exists $\mathbf{q} \in \mathbb{Z}^n$ satisfying (2.2), (2.3) and

$$|\mathbf{q} \cdot \mathbf{f}'(x)| \geq \sqrt{nL2^{t+1}} \,, \tag{2.4$\geq$}$$

where L is defined as in (1.6a). Similarly, denote by $A_{<}(t)$ the set of $x \in B$ such that there exists $\mathbf{q} \in \mathbb{Z}^n$ satisfying (2.2), (2.3) and

$$|\mathbf{q} \cdot \mathbf{f}'(x)| < \sqrt{nL2^{t+1}} \,. \tag{2.4$<$}$$

Obviously $A(t)$ belongs to the union of $A_{\geq}(t)$ and $A_{<}(t)$.

To estimate the measure of $A_{\geq}(t)$, note that if $x \in B$ and $\mathbf{q} \in \mathbb{Z}^n$ satisfy (2.2), (2.3) and (2.4$\geq$), and if δ stands for $\psi(2^{nt})$, then x belongs to the solution set of the system (1.6c). Furthermore, if $2^t \geq 1/4nL|B|^2$, then inequality (1.6b) holds, therefore, by Theorem 1.3, the measure of each of the solution sets (for every value of $\mathbf{q}$) is at most $32|B|\psi(2^{nt})$. But there are at most $(2 \cdot 2^{t+1})^n$ choices of $\mathbf{q}$ one can make to satisfy (2.2). Hence for large enough values of t one gets

$$|A_{\geq}(t)| \leq 32|B|(2 \cdot 2^{t+1})^n \psi(2^{nt}) = 2^{4n+5} 2^{nt} \psi(2^{nt}) \,,$$

and it follows from (1.2) and the monotonicity of ψ that $\sum_{t=0}^{\infty} |A_{\geq}(t)| < \infty$.

As for $A_{<}(t)$, note that it is contained in the set (1.7b) if one makes substitutions $\delta = \psi(2^{nt})$, $K = \sqrt{nL2^{t+1}}$ and $Q_1 = \cdots = Q_n = 2^{t+1}$. Further, the expression in the middle of (1.7a) is then equal to $2^{n-1/2}\sqrt{nL}2^{(n-1/2)t}$. Take E, ρ as in Theorem 1.4 and t large enough to guarantee that $\rho^{n+1}/\delta = \rho^{n+1}/\psi(2^{nt})$ is not less than 2^{nt} (possible due to (1.2)). Then both parts of (1.7a) will hold (with $\varkappa = n$), so, by Theorem 1.4, one has

$$|A_{<}(t)| \leq E\big(2^{n-1/2}\rho^{n+1}\sqrt{nL}2^{-t/2}\big)^{\frac{1}{(n+1)(2n-1)}} |B| \,.$$

This implies that $\sum_{t=0}^{\infty} |A_{<}(t)|$ is also finite, finishing the proof. $\square$

Proof of Part (M). Here we follow the same lines, except that we have to proceed by induction on n. The case $n = 1$ is standard, so let us assume that the claim is proven for any nondegenerate $(n-1)$-tuple of functions, and take an open interval $V \subset \mathbb{R}$, an n-tuple $\mathbf{f} = (f_1, \dots, f_n)$ of C^n functions on V which is non-degenerate at $x_0 \in V$ and a non-increasing function ψ satisfying (1.2M). Again consider the subinterval B of V given by Theorem 1.4. Because of the induction assumption and the fact that projections of a non-degenerate manifold are non-degenerate, we know that for a.e. $x \in B$ there are at most finitely many $\mathbf{q} \in \mathbb{Z}^n$ such that $q_i = 0$ for some $i = 1, \dots, n$, and also

$$\lfloor \mathbf{q} \cdot \mathbf{f}(x) \rfloor < \psi\big(\Pi_+(\mathbf{q})\big) \,. \tag{2.1M}$$

We need to prove that the same is true if one includes integer vectors $\mathbf{q}$ with all coordinates different from zero.

The crucial difference between this setting and the previous one is that we need to take into account the values of each component q_i of $\mathbf{q}$. Therefore instead of just one number t it is natural to use the n-tuple $\mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{Z}_+^n$, and denote by $A(\mathbf{t})$ the set of $x \in B$ such that the inequality (2.1M) holds for some $\mathbf{q}$ with

$$2^{t_i} \le \mathbf{q}_i < 2^{t_i+1}, \ i = 1, \ldots, n. \tag{2.2M}$$

Clearly to prove (M) it now suffices to show that $\sum_{\mathbf{t}\in\mathbb{Z}_+^n} |A(\mathbf{t})| < \infty$.

Denote by $\bar{t}$ the sum of all the components of $\mathbf{t}$ (this notation will be used throughout this subsection, so that whenever $\bar{t}$ and $\mathbf{t}$ appear in the same formula, $\bar{t}$ will stand for $\sum_{i=1}^n t_i$). It immediately follows from (2.2M) that $\Pi_+(\mathbf{q}) \ge 2^{\bar{t}}$, therefore, by the monotonicity of ψ, any solution $(x, \mathbf{q})$ of (2.1M) and (2.2M) satisfies

$$\lfloor \mathbf{q} \cdot \mathbf{f}(x) \rfloor \le \psi(2^{\bar{t}}). \tag{2.3M}$$

As before, we are going to consider two cases, based on the magnitude of the derivative of the function $\mathbf{q}\cdot\mathbf{f}$. That is, we will cover $A(\mathbf{t})$ by the union of $A_\ge(\mathbf{t})$ and $A_<(\mathbf{t})$, where $A_\ge(\mathbf{t})$ (resp. $A_<(\mathbf{t})$) is the set of $x \in B$ such that there exists $\mathbf{q} \in \mathbb{Z}^n$ satisfying (2.2M), (2.3M) and

$$|\mathbf{q} \cdot \mathbf{f}'(x)| \ge \sqrt{nL2^{\|\mathbf{t}\|+1}}, \tag{2.4M$\ge$}$$

or, respectively,

$$|\mathbf{q} \cdot \mathbf{f}'(x)| < \sqrt{nL2^{\|\mathbf{t}\|+1}}, \tag{2.4M$<$}$$

where L is again defined by (1.6a).

To estimate the measure of $A_\ge(\mathbf{t})$, note that if $x \in B$ and $\mathbf{q} \in \mathbb{Z}^n$ satisfy (2.2M), (2.3M) and (2.4M$\ge$), and if δ stands for $\psi(2^{\bar{t}})$, then x belongs to the solution set of the system (1.6c). Furthermore, (1.6b) holds whenever $2^{\|\mathbf{t}\|} \ge 1/4nL|B|^2$, therefore, by Theorem 1.3, for large enough $\|\mathbf{t}\|$ the measure of each of the solution sets is at most $32|B|\psi(2^{\bar{t}})$. But now there are at most $\prod_{i=1}^n 2^{t_i+2} = 2^{\bar{t}+2n}$ choices of $\mathbf{q}$ one can make to satisfy (2.2M). Hence for $\mathbf{t}$ with big enough norm one gets

$$|A_\ge(\mathbf{t})| \le 32|B|2^{\bar{t}+2n}\psi(2^{n\bar{t}}) = 2^{4n+5}2^{\bar{t}}\psi(2^{\bar{t}}).$$

Observe that for any $t \in \mathbb{N}$ the number of possible vectors $\mathbf{t} \in \mathbb{Z}_+^n$ with $\bar{t} = t$ is not greater than $\mathrm{const}\cdot t^{n-1}$. (the values of constants here and below are different and depend on n but not on t). Therefore $\sum_{\mathbf{t}\in\mathbb{Z}_+^n} |A_\ge(\mathbf{t})|$ can be bounded from above by $\mathrm{const}\cdot\sum_{t=1}^\infty t^{n-1}2^t\psi(2^t)$, which is finite due to (1.2M).

In order to use Theorem 1.4 to estimate $|A_<(\mathbf{t})|$, let us take $\delta = \psi(2^{\bar{t}})$, $K = \sqrt{nL2^{\|\mathbf{t}\|+1}}$ and $Q_i = 2^{t_i+1}$, $i = 1, \ldots, n$. Then one has

$$\frac{\delta K Q_1 \cdot \ldots \cdot Q_n}{\max_i Q_i} = 2^{n-1/2}\sqrt{nL}\psi(2^{\bar{t}})2^{\bar{t}}2^{-\|\mathbf{t}\|/2}.$$

If E, ρ are as in Theorem 1.4 and one takes $\mathbf{t}$ far enough from the origin, so that $\psi(2^{\bar{t}}) \le \rho^{n+1}2^{-\bar{t}}$, then the constraint (1.7a) with $\varkappa = n$ will be satisfied, and by Theorem 1.4, one will have

$$|A_<(\mathbf{t})| \le E\big(2^{n-1/2}\rho^{n+1}\sqrt{nL}2^{-\|\mathbf{t}\|/2}\big)^{\frac{1}{(n+1)(2n-1)}}|B|.$$

This clearly implies that $\sum_{\mathbf{t}\in\mathbb{Z}_+^n} |A_<(\mathbf{t})|$ is finite as well. $\square$

3. Proof of Theorem 1.3

We first state a very simple "bounded distortion" lemma.

Lemma 3.1. *Let $f \in C([a,b])$ be such that $c < f'(x) < d$ for some $c, d > 0$ and all $x \in [a,b]$. Then for any r and s such that $f(a) \le r \le s \le f(b)$ one has*

$$\frac{s-r}{d} \le |\{x \in [a,b] \mid r \le f(x) \le s\}| \le \frac{s-r}{c}.$$

Proof. Put $y = f^{-1}(r)$ and $z = f^{-1}(s)$; since f is increasing, the set whose measure we need to estimate is the interval $[y,z]$. Using Lagrange's formula, one writes $z - y = \frac{f(z)-f(y)}{f'(x)}$ for some $x \in [y,z]$, and the lemma follows. $\square$

Lemma 3.2. *Let $B \subset \mathbb{R}$ be an interval, and let the numbers*

$$M \ge 1/4|B|^2 \tag{3.1a}$$

and $\delta > 0$ be given. Take a function $F \in C^2(B)$ such that

$$\sup_{x \in B} |F''(x)| \le M\,, \tag{3.1b}$$

and denote by S the set of all $x \in B$ for which the inequalities

$$\begin{cases} \lfloor F(x) \rfloor < \delta \\ |F'(x)| \ge \sqrt{M} \end{cases} \tag{3.1c}$$

hold. Then $|S| \le 32\delta|B|$.

Proof. Note that the desired estimate is trivial when $\delta \ge 1/32$, so without loss of generality we can assume that δ is less than $1/32$. Given $x \in B$, let us denote by $J(x)$ the maximal subinterval of B containing x such that $\lfloor F(y) \rfloor < 1/4$ for all $y \in J(x)$. Because of the maximality, for any two points $x_1, x_2 \in B$ the intervals $J(x_1)$ and $J(x_2)$ either are disjoint or coincide. Therefore it is enough to prove that for any $x \in S$ one has

$$|S \cap J(x)| \le 32\delta|J(x)|\,. \tag{3.2}$$

Fix $x \in S$. It is clear that there exists a unique $p \in \mathbb{Z}$ such that $|F(y) + p| < 1/4$ for all $y \in J(x)$. We claim that $J(x)$ is contained in the interval $[x - \frac{1}{2|F'(x)|}, x + \frac{1}{2|F'(x)|}]$. Indeed, if $x \pm \frac{1}{2|F'(x)|} \in B$, one has

$$F\big(x \pm \tfrac{1}{2|F'(x)|}\big) + p = F(x) + p \pm \tfrac{F'(x)}{2|F'(x)|} + \tfrac{F''(z)}{2}\big(\tfrac{1}{2|F'(x)|}\big)^2$$

for some $z \in B$. Thus

$$|F\big(x \pm \tfrac{1}{2|F'(x)|}\big) + p| \underset{(3.1b)}{\ge} -\delta + 1/2 - \tfrac{M}{8|F'(x)|^2} \underset{(3.1c)}{\ge} 1/2 - 1/8 - \delta \underset{(\delta<1/32)}{>} 1/4\,,$$

hence the claim.

Applying Lagrange's formula to F', for all $y \in J(x)$ and some $z \in B$ one gets

$$|F'(y) - F'(x)| = |F''(z)||y - x| \underset{(3.1\mathrm{b})}{\leq} \tfrac{M}{2|F'(x)|} \underset{(3.1\mathrm{c})}{\leq} \tfrac{|F'(x)|}{2}\,.$$

This implies that the absolute value of $F'(y)$, $y \in J(x)$, is strictly between $\frac{1}{2}|F'(x)|$ and $\frac{3}{2}|F'(x)|$; in particular, F is monotonic on $J(x)$. Denote by $J'(x)$ the interval $\big\{x \in J(x) \mid |F(x) + p| < \delta\big\}$. Clearly $S \cap J(x)$ is contained in $J'(x)$. Now one can apply Lemma 3.1 to the function $\pm\big(F(x) + p\big)$ to conclude that

$$|J'(x)| \leq 2\delta \tfrac{2}{|F'(x)|} = \tfrac{4\delta}{|F'(x)|}\,,$$

and also

$$|J(x)| \geq \left(\tfrac{1}{4} - \delta\right) \tfrac{2}{3|F'(x)|} \underset{(\delta < 1/32)}{\geq} \tfrac{1}{8|F'(x)|}$$

(here we used the fact that, due to (3.1a), at least one of the intervals $[x, x + \frac{1}{2|F'(x)|}]$ and $[x - \frac{1}{2|F'(x)|}, x]$ is contained in B). Therefore $|J'(x)| \leq 32\delta|J(x)|$, and (3.2) follows. $\square$

Now it takes very little to complete the

Proof of Theorem 1.3. Given an interval $B \subset \mathbb{R}$, an n-tuple of C^2 functions $\mathbf{f}$ on B, a positive δ and $\mathbf{q} \in \mathbb{Z}^n$ satisfying (1.6b) with L as in (1.6a), denote $F(x) \stackrel{\text{def}}{=} \mathbf{q} \cdot \mathbf{f}(x)$ and $M \stackrel{\text{def}}{=} nL\|\mathbf{q}\|$. Then inequalities (1.6abc) can be rewritten as (3.1bac), and the theorem follows. $\square$

4. (C, α)-GOOD FUNCTIONS

Let us recall the definition introduced in [KM1]. If C and α are positive numbers and B a subset of $\mathbb{R}$, let us say that a function $f : B \mapsto \mathbb{R}$ is (C, α)-*good on* B if for any open interval $J \subset B$ and any $\varepsilon > 0$ one has

$$\left|\left\{x \in J \;\middle|\; |f(x)| < \varepsilon \cdot \sup_{x \in J} |f(x)|\right\}\right| \leq C\varepsilon^\alpha |J|\,. \tag{4.1}$$

Several elementary facts about (C, α)-good functions are listed below:

Lemma 4.1. (a) *f is (C, α)-good on $B \Rightarrow$ so is λf $\forall\, \lambda \in \mathbb{R}$;*

(b) *f_i, $i \in I$, are (C, α)-good on $B \Rightarrow$ so is $\sup_{i \in I} |f_i|$;*

(c) *If f is (C, α)-good on B and $c_1 \leq \left|\frac{f(x)}{g(x)}\right| \leq c_2$ for all $x \in B$, then g is $\big(C(c_2/c_1)^\alpha, \alpha\big)$-good on B;*

(d) *f is (C, α)-good on $B \Rightarrow$ it is (C', α')-good on B' for every $C' \geq C$, $\alpha' \leq \alpha$ and $B' \subset B$;*

(e) [KM1, Proposition 3.2] *any polynomial $f \in \mathbb{R}[x]$ of degree not greater than k is $(4k, 1/k)$-good on $\mathbb{R}$.*

The following statement is a special case of [KM1, Lemma 3.3]:

Lemma 4.2. *Let B be an open subset of $\mathbb{R}$, and let $f \in C^k(B)$ be such that for some constants $c, \tilde{c} > 0$ one has*

$$|f^{(i)}(x)| \le \tilde{c} \quad \forall\, i = 1, \dots, k \text{ and } \forall\, x \in B\,, \tag{4.2$\le$}$$

and

$$|f^{(k)}(x)| \ge c \quad \forall\, x \in B\,. \tag{4.2$\ge$}$$

Then there exists $C > 0$ (dependent only on k, c and $\tilde{c}$) such that f is (C, α)-good on B.

(See [KM1] for the exact expression for the constant C.)

Corollary 4.3. *Let V be an open subset of $\mathbb{R}$, $x_0 \in V$, and let $\mathbf{f} = (f_1, \dots, f_n) : V \mapsto \mathbb{R}^n$ be an n-tuple of C^n functions which is non-degenerate at x_0. Then there exists a neighborhood $B \subset V$ of x_0 and a positive C such that*

(a) *any linear combination of $1, f_1, \dots, f_n$ is $(C, \frac{1}{n})$-good on B;*

(b) *any linear combination of $f_1', \dots, f_n'$ is $(C, \frac{1}{n-1})$-good on B.*

Proof. Part (a) is a special case of [KM1, Proposition 3.4]. To prove (b), consider $f' = \sum_{i=1}^n v_i f_i'$; in view of Lemma 4.1(a), one can without loss of generality assume that $\mathbf{v}$ is a unit vector. From the nondegeneracy assumption it follows that there exists a constant $c' > 0$ such that for any $\mathbf{v} \in \mathbb{R}^n_1$ one can find $k = 1, \dots, n$ with

$$\Big|\sum_{i=1}^n v_i f_i^{(k)}(x_0)\Big| = |f^{(k)}(x_0)| \ge c'\,. \tag{4.3}$$

Then one can use the continuity of the derivatives of $f_1, \dots, f_n$ to choose a neighborhood $B \subset V$ of x_0 and positive $c, \tilde{c}$ (again independently of $\mathbf{v}$) such that the inequalities (4.2) hold.

Now we have to consider separately the case when k as in (4.3) is equal to 1. Then $c \le |f'(x)| \le \tilde{c}$ for $x \in B$, therefore for any positive ε and any subinterval J of B one has

$$\Big|\{x \in J \;\big|\; |f'(x)| < \varepsilon \cdot \sup_{x \in J} |f(x)|\}\Big| \le \left(\frac{\tilde{c}}{c}\right)^{\frac{1}{n-1}} \varepsilon^{\frac{1}{n-1}} |J|\,.$$

Otherwise (if $k \ge 1$) one can apply Lemma 4.2 to f' and establish that f' is $(C, \frac{1}{n-1})$-good on B for appropriate B and C. $\square$

For the rest of the section we will be studying planar curves parametrized by C^n maps $\mathbf{g} = (g_1, g_2) : V \mapsto \mathbb{R}^2$. For such a curve, let us define

$$A_{\mathbf{g}}(x) \stackrel{\text{def}}{=} \begin{vmatrix} g_1(x) & g_2(x) \\ g_1'(x) & g_2'(x) \end{vmatrix}\,,$$

that is, the signed area of the parallelogram spanned by $\mathbf{g}(x)$ and $\mathbf{g}'(x)$.

We will need the following result:

Proposition 4.4. *Let V be an open subset of $\mathbb{R}$, $x_0 \in V$, and let $\mathcal{G} \subset C^n(V)$ be a family of 2-tuples $\mathbf{g} = (g_1, g_2) : V \mapsto \mathbb{R}^2$ such that*

$$\{\mathbf{g}' \mid \mathbf{g} \in \mathcal{G}\} \text{ is compact in } C^{n-1}(V)\,. \tag{4.4a}$$

Assume also that

$$\inf_{\substack{\mathbf{g}\in\mathcal{G}\\ \mathbf{u}\in\mathbb{R}^2_1}} \sup_{1\le k\le n} |\mathbf{u}\cdot\mathbf{g}^{(k)}(x_0)| > 0 \tag{4.4b}$$

(in other words, the derivatives of $\mathbf{g}$ at x_0 uniformly (in $\mathbf{g} \in \mathcal{G}$) generate the unit ball in $\mathbb{R}^2$). Then

(a) *there exists a neighborhood $B \subset V$ of x_0 such that for any $\mathbf{g} \in \mathcal{G}$ the function $A_{\mathbf{g}}$ is $(8n, \frac{1}{2n-1})$-good on B;*

(b) *for every neighborhood $B \subset V$ of x_0 there exists $\rho = \rho(\mathcal{G}, B)$ such that*

$$\sup_{x\in B} |A_{\mathbf{g}}(x)| \ge \rho \quad \text{for every } \mathbf{g} \in \mathcal{G}\,.$$

To prove this proposition, we will use four lemmas below.

Lemma 4.5. *Let $B \subset \mathbb{R}$ be an interval and $f \in C^k(B)$ be such that (4.2$\ge$) holds. Then*

$$\sup_{x,y\in B} |f(x) - f(y)| \ge \frac{c}{k^k(k+1)!}|B|^k\,.$$

Proof. Denote $\sup_{x,y\in B} |f(x) - f(y)|$ by s. Take any $y \in B$, divide B into k equal segments and let $p(x)$ be the Lagrange polynomial of degree k formed by using values of $f(x) - f(y)$ at the boundary points of these segments. Then there exists $x \in B$ such that $p^{(k)}(x) = f^{(k)}(x)$, hence, by (4.2$\ge$), $|p^{(k)}(x)| \ge c$. On the other hand, after differentiating $p(x)$ k times (see [KM1, (3.3a)]) one gets $|p^{(k)}(x)| \le (k+1)\dfrac{sk!}{(|B|/k)^k}$. Combining the last two inequalities, one easily gets the desired estimate. $\square$

Lemma 4.6. *Let f be a continuous function on an interval $B \subset \mathbb{R}$, and suppose $C, \alpha > 0$ and $0 < \delta < 1$ are such that (4.1) holds for any subinterval J of B and any $\varepsilon \ge \delta$. Then f is (C, α')-good on B whenever $0 < \alpha' < \alpha$ is such that $C\delta^{\alpha-\alpha'} \le 1$.*

Proof. Take $m \in \mathbb{N}$ such that $\delta^{m+1} \le \varepsilon < \delta^m$. It is easy to show by induction on m that

$$\Big|\{x \in J \;\Big|\; |f(x)| < \varepsilon \cdot \sup_{x\in J} |f(x)|\}\Big| \le C^{m+1}\varepsilon^\alpha |J|\,. \tag{4.5}$$

Indeed, the case $m = 0$ follows from the assumption. Assume that (4.5) holds for some m, and observe that, by the continuity of f, the set $\{x \in J \mid |f(x)| < \varepsilon \cdot \sup_{x\in J} |f(x)|\}$, $\varepsilon \le 1$, is a union of intervals I_i with $\sup_{x\in I_i} |f(x)| = \varepsilon \cdot \sup_{x\in J} |f(x)|$ for every i. Therefore

$$\begin{aligned}\Big|\{x \in J \;\Big|\; |f(x)| < \delta\varepsilon \cdot \sup_{x\in J} |f(x)|\}\Big| &\le \sum_i \Big|\{x \in I_i \;\Big|\; |f(x)| < \delta \sup_{x\in I_i} |f(x)|\}\Big| \le \sum_i C\delta^\alpha |I_i| \\ &= C\delta^\alpha \Big|\{x \in J \;\Big|\; |f(x)| < \varepsilon\}\Big| \le C\delta^\alpha \cdot C^{m+1}\varepsilon^\alpha |J|\,,\end{aligned}$$

which proves (4.5) with $\delta\varepsilon$ in place of ε.

It remains to write

$$C^{m+1}\varepsilon^{\alpha} = C \cdot C^{m}(\delta^{m})^{\alpha-\alpha'}(\delta^{m})^{\alpha'}\left(\varepsilon\delta^{-m}\right)^{\alpha} \le C \cdot (C\delta^{\alpha-\alpha'})^{m}\varepsilon^{\alpha'},$$

which implies that f is (C,α')-good on B provided $C\delta^{\alpha-\alpha'} \le 1$. $\square$

From this point and till the end of this section, we are going to employ the Euclidean norm $\|\mathbf{x}\| = \operatorname{dist}(0,\mathbf{x})$ on $\mathbb{R}^2$.

Lemma 4.7. *Let $B \subset \mathbb{R}$ be an interval and let $\mathbf{g} = (g_1, g_2)$ be a C^1 map $B \mapsto \mathbb{R}^2$. Take $y \in B$, denote by $\mathbf{u}$ the unit vector orthogonal to $\mathbf{g}(y)$, and let $d = \sup_{x\in B}\|\mathbf{g}(x) - \mathbf{g}(y)\|$ and $w = \sup_{x\in B}|\mathbf{u}\cdot\mathbf{g}(x)|$. Then*

$$\sup_{x\in B}|A_{\mathbf{g}}(x)| \ge \frac{w\big(\|\mathbf{g}(x)\| - 2d\big)}{|B|}.$$

Proof. Denote the function $\mathbf{u}\cdot\mathbf{g}(x)$ by f, and take $z \in B$ such that $|f(z)| = w$. Then, using Lagrange's Theorem, find t between y and z such that $\mathbf{g}(y) - \mathbf{g}(z) = \mathbf{g}'(t)(y-z)$. Our goal is to show that $|A_{\mathbf{g}}(t)|$ is big enough. Let $\mathcal{L}$ be the line connecting $\mathbf{g}(y)$ and the origin, $\mathcal{M}$ the tangent line to the curve at $\mathbf{g}(t)$ and θ the angle between $\mathcal{L}$ and $\mathcal{M}$. Note that $\sin\theta = |f(z)|/\|\mathbf{g}(y) - \mathbf{g}(z)\| = \frac{|f(z)|}{|y-z|\cdot\|\mathbf{g}'(t)\|}$. From the construction it follows that $\mathcal{L}$ and $\mathcal{M}$ intersect at some point $\mathbf{x}$, and, since $|f(z)| \ge |f(t)|$, the distance between $\mathbf{g}(t)$ and $\mathbf{x}$ is not greater than $\|\mathbf{g}(y) - \mathbf{g}(z)\|$. Therefore

$$\|\mathbf{x}\| \ge \|\mathbf{g}(t)\| - \|\mathbf{x} - \mathbf{g}(t)\| \ge \|\mathbf{g}(t)\| - \|\mathbf{g}(z) - \mathbf{g}(y)\| \ge \|\mathbf{g}(y)\| - 2d.$$

Now, $|A_{\mathbf{g}}(t)|$ is equal to the product of $\|\mathbf{g}'(t)\|$ and the distance from 0 to $\mathcal{M}$, and the latter is equal to $\|\mathbf{x}\|\sin\theta$. Therefore we get

$$|A_{\mathbf{g}}(t)| = \|\mathbf{g}'(t)\|\|\mathbf{x}\|\frac{|f(z)|}{|y-z|\cdot\|\mathbf{g}'(t)\|} = \frac{|f(z)|\|\mathbf{x}\|}{|y-z|} \ge \frac{w(\|\mathbf{g}(x)\| - 2d)}{|B|},$$

completing the proof. $\square$

Lemma 4.8. *Put $J = [0,1]$ and let $\mathbf{p} : J \mapsto \mathbb{R}^2$ be a polynomial map of degree $\le n$ such that*

(4.6a) $$\sup_{x,y\in J}\|\mathbf{p}(x) - \mathbf{p}(y)\| \le 2$$

(the diameter of the image of $\mathbf{p}$ is bounded from above), and

(4.6b) $$\sup_{x\in J}\operatorname{dist}\big(\mathcal{L},\mathbf{p}(x)\big) \ge 1/8 \text{ for any straight line } \mathcal{L} \subset \mathbb{R}^2$$

(that is, the "width" of $\mathbf{p}\big([0,1]\big)$ in any direction is bounded from below). Then:

(a) *there exists a constant $0 < \gamma < 1$ (dependent only on n) such that*

(4.7a) $$\sup_{x\in J}|A_{\mathbf{p}}(x)| \ge \gamma\big(1 + \sup_{x\in J}\|\mathbf{p}(x)\|\big);$$

(b) *there exists $M \geq 1$ (dependent only on n) such that*

$$\sup_{x \in J} \|\mathbf{p}'(x)\| \leq M\,. \tag{4.7b}$$

Note that it is possible to explicitly estimate γ and M using Lagrange's interpolation formula.

Proof. Let $\mathcal{P}$ be the set of polynomial maps $\mathbf{p} : [0,1] \mapsto \mathbb{R}^2$ of degree $\leq n$ satisfying (4.6ab) and such that $\sup_{x\in J} \|\mathbf{p}(x)\| \leq 3$. We first prove that there exists $\gamma > 0$ such that (4.7a) holds for any $\mathbf{p} \in \mathcal{P}$. Indeed, otherwise from the compactness of $\mathcal{P}$ it follows that there exists $\mathbf{p} \in \mathcal{P}$ such that $A_{\mathbf{p}}(x)$ is identically equal to zero. Clearly this can only happen when all coefficients of $\mathbf{p}$ are proportional to each other, which contradicts (4.6b).

Now assume that $\mathbf{p}$ satisfies (4.6ab) and $\|\mathbf{p}(y)\| > 3$ for some $y \in J$. Then one can apply the previous lemma to the map $\mathbf{p} : J \mapsto \mathbb{R}^2$ to get $\sup_{x\in J} |A_{\mathbf{p}}(x)| \geq \frac{\frac{1}{8}(\|\mathbf{p}(y)\|-2)}{|J|} \geq \frac{1}{32}\big(1+\|\mathbf{p}(y)\|\big)$, which finishes the proof of part (a). It remains to observe that part (b) trivially follows from the compactness of the set of polynomials of the form $\mathbf{p}'(x)$ where $\mathbf{p}(x)$ satisfies (4.6a) and has degree $\leq n$. $\square$

Proof of Proposition 4.4. Choose $0 < \delta < 1/8$ such that

$$8n \cdot \delta^{\frac{1}{(2n-1)(2n-2)}} \leq 1\,. \tag{4.8}$$

From the compactness of the family $\{\mathbf{g}' \mid \mathbf{g} \in \mathcal{G}\}$ it follows that one can choose a neighborhood B of x_0 and a positive c such that for every $\mathbf{g} \in \mathcal{G}$ and any $x, y \in B$ one has

$$\forall\, \mathbf{u} \in \mathbb{R}^2_1 \quad \exists\, k = 1,\ldots,n \text{ such that } |\mathbf{v}\cdot \mathbf{g}^{(k)}(x)| \geq c\,, \tag{4.9a}$$

and

$$\|\mathbf{g}^{(n)}(x) - \mathbf{g}^{(n)}(y)\| \leq \frac{\delta c \gamma}{8Mn^n(n+1)!}\,, \tag{4.9b}$$

where γ and M are as in Lemma 4.7. Take a subinterval J of B. In view of Lemma 4.6 and (4.8), in order to demonstrate part (a), it suffices to prove the following: given a C^n map $\mathbf{g} : J \to \mathbb{R}^2$ such that the inequalities (4.9ab) hold for all $x, y \in J$, one has

$$\Big|\{x \in J \;\Big|\; |A_{\mathbf{g}}(x)| < \varepsilon \cdot \sup_{x\in J} |A_{\mathbf{g}}(x)|\}\Big| \leq 8n\varepsilon^{\frac{1}{2n-2}}|J| \quad \text{whenever } \varepsilon \geq \delta\,. \tag{4.10}$$

We will do this in several steps.

Step 1. Note that conditions (4.9ab), as well as the function $A_{\mathbf{g}}$, will not change if one replaces $\mathbf{g}$ by $L\mathbf{g}$ where L is any rotation of the plane (g_1, g_2). Thus one can choose the "g_1-axis" in such a way that it is parallel to the line connecting two most distant points of $\mathbf{g}(B)$. By (4.9a) there exist $1 \leq k_1, k_2 \leq n$ such that $|g_i^{(k_i)}(x)| \geq c$ for $i = 1,2$ and all $x \in J$. Denote $J = [a,b]$ and $s_i = \sup_{x,y\in J} |g_i(x) - g_i(y)|$, $i = 1,2$. Then it follows from Lemma 4.5 that

$$s_i \geq \frac{c}{{k_i}^{k_i}(k_i+1)!}(b-a)^{k_i} \geq \frac{c}{n^n(n+1)!}(b-a)^n\,. \tag{4.11}$$

Step 2. Here we replace the functions $g_i(x)$ by $\frac{1}{s_i} g_i\big(a+(b-a)x\big)$, and the interval J by $[0,1]$. This way the function $A_{\mathbf{g}}$ will be multiplied by a constant, and the statement (4.10) that we need to prove will be left unchanged. However the n-th derivatives of the functions $g_i(x)$ will be multiplied by factors $\frac{(b-a)^n}{s_i}$. In view of (4.11), the inequality (4.9b) will then imply

$$\sup_{x,y\in J} \|\mathbf{g}^{(n)}(x) - \mathbf{g}^{(n)}(y)\| \le \delta\gamma/8M \tag{4.12}$$

(here and until the end of the proof, J stands for [0,1]). Note also that it follows from the construction that $\mathbf{g}(J)$ is contained in a translate of the square $[-\frac{1}{2},\frac{1}{2}]^2$, and that $\sup_{x\in J} \operatorname{dist}\big(\mathcal{L}, \mathbf{g}(x)\big) \ge 1/2\sqrt{2}$ for any straight line $\mathcal{L}\subset\mathbb{R}^2$.

Step 3. Here we introduce the Taylor polynomial $\mathbf{p}(x)$ of $\mathbf{g}(x)$ at 0. Using (4.12) one can show that $\mathbf{p}$ is $\frac{\delta\gamma}{8M}$-close to $\mathbf{g}$ in the C^1 topology, that is,

$$\sup_{x\in J} \|\mathbf{g}(x)-\mathbf{p}(x)\| \le \delta\gamma/8M \quad \text{and} \quad \sup_{x\in J} \|\mathbf{g}'(x)-\mathbf{p}'(x)\| \le \delta\gamma/8M\,.$$

It follows that conditions (4.6ab) are satisfied, and therefore, by Lemma 4.8, the inequalities (4.7ab) hold.

Step 4. Now let us compare the functions $A_{\mathbf{g}}$ and $A_{\mathbf{p}}$: one has

$$A_{\mathbf{g}}(x) - A_{\mathbf{p}}(x) = \begin{vmatrix} g_1(x) & g_2(x)-p_2(x) \\ g_1'(x) & g_2'(x)-p_2'(x) \end{vmatrix} - \begin{vmatrix} g_1(x)-p_1(x) & g_2(x) \\ g_1'(x)-p_1'(x) & g_2'(x) \end{vmatrix}\,,$$

therefore

$$\begin{aligned} |A_{\mathbf{g}}(x) - A_{\mathbf{p}}(x)| &\le \frac{\delta\gamma}{8M}\big(\sup_{x\in J}|g_1(x)| + \sup_{x\in J}|g_1'(x)| + \sup_{x\in J}|p_1(x)| + \sup_{x\in J}|p_1'(x)|\big) \\ &\le \frac{\delta\gamma}{8M}\big(2+\frac{\delta\gamma}{8M}\big)\big(\sup_{x\in J}\|\mathbf{p}(x)\| + \sup_{x\in J}\|\mathbf{p}'(x)\|\big) \\ &\underset{(4.7\text{b})}{\le} \frac{3\delta\gamma}{8}\big(1+\sup_{x\in J}\|\mathbf{p}(x)\|\big) \underset{(4.7\text{a})}{\le} \frac{3\delta}{8}\sup_{x\in J}|A_{\mathbf{p}}(x)|\,. \end{aligned}$$

Step 5. Finally we are ready to prove (4.10): take ε between δ and 1, put $s \overset{\text{def}}{=} \sup_{x\in J}|A_{\mathbf{p}}(x)|$ and observe that, in view of Step 4, the set in the left hand side of (4.10) is contained in

$$\big\{x\in J \bigm| |A_{\mathbf{p}}(x)| - \tfrac{3\delta}{8}s < \varepsilon(1+\tfrac{3\delta}{8})s\big\} = \big\{x\in J \bigm| |A_{\mathbf{p}}(x)| < \big(\varepsilon + \tfrac{3\delta}{8}(1+\varepsilon)\big)\,s\big\}\,.$$

Since $\varepsilon+\frac{3\delta}{8}(1+\varepsilon) \le \varepsilon + \frac{3\delta}{4} \le 2\varepsilon$, and since $A_{\mathbf{p}}$ is a polynomial of degree not greater than $2n-2$, one can apply Lemma 4.1(e) and write

$$\Big|\big\{x\in J \bigm| |A_{\mathbf{g}}(x)| < \varepsilon\cdot\sup_{x\in J}|A_{\mathbf{g}}(x)|\big\}\Big| \le \Big|\big\{x\in J \bigm| |A_{\mathbf{p}}(x)| < 2\varepsilon s\big\}\Big| \le 4(2n-2)(2\varepsilon)^{\frac{1}{2n-2}} \le 8n\varepsilon^{\frac{1}{2n-2}}\,,$$

which shows (4.10) and finishes the proof of part (a).

As for part (b), fix $B \ni x_0$; it is clear that there exists $r > 0$ such that for any $\mathbf{g} \in \mathcal{G}$ one has $\sup_{x\in B} \|\mathbf{g}(x)\| \geq r$ (otherwise, by a compactness argument similar to that of Lemma 4.8, one would get that $0 \in \mathcal{G}$, contradicting (4.4b)). Also take $C \geq 1$ such that

$$\sup_{\mathbf{g}\in\mathcal{G},\, x\in B} \|\mathbf{g}'(x)\| \leq C\,. \tag{4.13}$$

It follows from (4.4ab) that there exists $0 < b < \frac{r}{4C}$ and $c > 0$ with the following property: for any $\mathbf{g} \in \mathcal{G}$, choose $y \in B$ with $\|\mathbf{g}(y)\| \geq r$ and let $\mathbf{u} \in \mathbb{R}^2_1$ be orthogonal to $\mathbf{g}(y)$; then there exists $1 \leq k \leq n$ such that $|\mathbf{u}\cdot\mathbf{g}^{(k)}(x)| > c$ for all x from a b-neighborhood $B' \subset B$ of y.

Applying Lemma 4.5 to the function $\mathbf{u}\cdot\mathbf{g}(x)$ one gets $\sup_{x\in B'} |\mathbf{u}\cdot\mathbf{g}(x)| \geq \frac{c}{k^k(k+1)!}b^k \geq \frac{c}{n^n(n+1)!}b^n$. On the other hand (4.13) shows that $\sup_{x\in B'} \|\mathbf{g}(x) - \mathbf{g}(y)\|$ is not greater than Cb. Now one can apply Lemma 4.7 to the map $\mathbf{g} : B' \mapsto \mathbb{R}^2$ to get

$$\sup_{x\in B'} |A_{\mathbf{g}}(x)| \geq \frac{c}{n^n(n+1)!}b^n(r - 2Cb)/b \geq \frac{cr}{2n^n(n+1)!}b^{n-1}\,,$$

and the proof is completed. $\square$

5. Proof of Theorem 1.4

The main idea of our approach to Theorem 1.4 is to translate the system of inequalities (1.7b) into the language of lattices in $\mathbb{R}^{n+2}$. We need to introduce some terminology from [KM1]. Fix $l \in \mathbb{N}$; in order to describe discrete subgroups of $\mathbb{R}^l$, we fix an orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_l$ of $\mathbb{R}^l$, and extend the norm $\|\cdot\|$ from $\mathbb{R}^l$ to the exterior algebra $\bigwedge(\mathbb{R}^l)$ by $\|\sum_{1\leq i_1<\dots<i_k\leq l} w_{(i_1,\dots,i_k)}\mathbf{e}_{i_1}\wedge\dots\wedge\mathbf{e}_{i_k}\| = \max_{1\leq i_1<\dots<i_k\leq l}|w_{(i_1,\dots,i_k)}|$.

For a discrete subgroup Γ of $\mathbb{R}^l$, we denote by $\Gamma_{\mathbb{R}}$ the minimal linear subspace of $\mathbb{R}^l$ containing Γ. Let $k = \dim(\Gamma_{\mathbb{R}})$; say that $\mathbf{w} \in \bigwedge^k(\mathbb{R}^l)$ *represents* Γ if

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

Clearly the element representing Γ is defined up to a sign. Therefore it makes sense to define the norm of Γ by $\|\Gamma\| \stackrel{\text{def}}{=} \|\mathbf{w}\|$, where $\mathbf{w}$ represents Γ. Note that if Γ is a lattice (that is, $\Gamma_{\mathbb{R}} = \mathbb{R}^l$), the ratio of $\|\Gamma\|$ and the volume of the quotient space $\mathbb{R}^l/\Gamma$ is uniformly bounded between two positive constants.

Let Λ be a discrete subgroup of $\mathbb{R}^l$. We say that a subgroup Γ of Λ is *primitive* (in Λ) if $\Gamma = \Gamma_{\mathbb{R}} \cap \Lambda$, and denote by $\mathcal{L}(\Lambda)$ the set of all nonzero primitive subgroups of Λ. Example: a cyclic subgroup of $\mathbb{Z}^l$ is primitive in $\mathbb{Z}^l$ iff it is generated by a *primitive* vector (that is, a vector which is not equal to a nontrivial multiple of another element of $\mathbb{Z}^l$).

We need the following theorem, which is a slight generalization of the $d = 1$ case of Theorem 5.2 from [KM1]. Its proof, essentially a verbatim repetition of the proof given in [KM1], is presented in the Appendix for the sake of making this paper self-contained (and also because in the one-dimensional case the proof is simpler and produces a better constant).

Theorem 5.1. *Let $l \in \mathbb{N}$, $C, \alpha > 0$, $0 < \rho \leq 1/l$, an interval $B \subset \mathbb{R}$ and a continuous map $H : B \to GL_l(\mathbb{R})$ be given. Take $\Lambda \in \mathcal{L}(\mathbb{Z}^l)$ and, for $\varepsilon > 0$, denote by $B_{H,\Lambda,\varepsilon}$ the set*

$$B_{H,\Lambda,\varepsilon} \stackrel{\text{def}}{=} \left\{x \in B \mid \|H(x)\mathbf{v}\| < \varepsilon \text{ for some } \mathbf{v} \in \Lambda \smallsetminus \{0\}\right\}.$$

Assume that for any $\Gamma \in \mathcal{L}(\Lambda)$,

(i) *the function* $x \mapsto \|H(x)\Gamma\|$ *is* (C,α)*-good on* B, *and*

(ii) $\exists\, x \in B$ *such that* $\|H(x)\Gamma\| \ge \rho$.

Then for any positive $\varepsilon \le \rho$ *one has* $|B_{H,\Lambda,\varepsilon}| \le C\dim(\Lambda_{\mathbb{R}})2^{\dim(\Lambda_{\mathbb{R}})}\left(\dfrac{\varepsilon}{\rho}\right)^{\alpha}|B|$.

Let us now explain how to connect the above theorem with Theorem 1.4. Denote by Λ the subgroup of $\mathbb{Z}^{n+2}$ consisting of integer vectors with zero second coordinate; that is,

$$\Lambda = \left\{ \begin{pmatrix} p \\ 0 \\ \mathbf{q} \end{pmatrix} \middle|\, p \in \mathbb{Z},\ \mathbf{q} \in \mathbb{Z}^n \right\}. \tag{5.1}$$

Also let U_x stand for the matrix

$$U_x \overset{\text{def}}{=} \begin{pmatrix} 1 & 0 & \mathbf{f}(x)^T \\ 0 & 1 & \mathbf{f}'(x)^T \\ 0 & 0 & I_n \end{pmatrix} \in SL_{n+2}(\mathbb{R}), \tag{5.2}$$

where $\mathbf{f}$ is as in Theorem 1.1. Note that $U_x \begin{pmatrix} p \\ 0 \\ \mathbf{q} \end{pmatrix} = \begin{pmatrix} \mathbf{q}\cdot\mathbf{f}(x)+p \\ \mathbf{q}\cdot\mathbf{f}'(x) \\ \mathbf{q} \end{pmatrix}$ is the vector whose components appear in the right hand sides of the inequalities (1.7b). Therefore the fact that there exists $\mathbf{q} \in \mathbb{Z}^n \setminus \{0\}$ satisfying (1.7b) implies the existence of a nonzero element of $U_x\Lambda$ which belongs to a certain parallelepiped in $\mathbb{R}^{n+2}$. Our strategy will be as follows: we will find a diagonal matrix $D \in GL_{n+2}(\mathbb{R})$ which transforms the above parallelepiped into a small cube; then the solvability of the above system of inequalities will force the lattice $DU_x\Lambda$ to have a small nonzero vector, and we will use Theorem 5.1 to estimate the measure of the set of $x \in B$ for which it can happen.

More precisely, take $\delta, K, Q_1, \ldots, Q_n$ as in Theorem 1.4, fix $\varepsilon > 0$ and denote

$$D = \operatorname{diag}(d_0^{-1}, d_*^{-1}, d_1^{-1}, \ldots, d_n^{-1}), \tag{5.3}$$

where

$$d_0 = \frac{\delta}{\varepsilon},\ d_* = \frac{K}{\varepsilon},\ d_i = \frac{Q_i}{\varepsilon},\ i = 1, \ldots, n. \tag{5.4}$$

It can be easily seen that the set (1.7b) will be exactly equal to $B_{h,\Lambda,\varepsilon}$ if we define $H : B \to GL_{n+2}(\mathbb{R})$ by

$$H(x) = DU_x. \tag{5.5}$$

Therefore one can hope to derive Theorem 1.4 from Theorem 5.1 by finding a neighborhood B of x_0 and constants C, α, ρ such that conditions (i) and (ii) hold. This hope is realized by the two lemmas below. Throughout this section, we will fix $V \subset \mathbb{R}$, $x_0 \in V$ and $\mathbf{f}$ as in Theorem 1.4, and will use the notation Λ, U_x, D and $H(x)$ as defined in (5.1), (5.2), (5.3) and (5.5).

Lemma 5.2. *There exist a subinterval B of V containing x_0 and a constant $C > 0$ such that for any $d_0, d_*, d_1, \dots, d_n$ and any $\Gamma \in \mathcal{L}(\Lambda)$,*

$$\text{the function } x \mapsto \|H(x)\Gamma\| \text{ is } (C, \tfrac{1}{2n-1})\text{-good on } B\,. \tag{5.6}$$

Proof. First note that the norm that is used to define the functions considered depends on the choice of an orthonormal basis of $\mathbb{R}^{n+2}$. However it is clear that passing to a different orthonormal basis would result in nothing more than a slight change of the constant C, see Lemma 4.1(c). Therefore it will suffice to take a discrete subgroup $\Gamma \in \mathcal{L}(\Lambda)$ and then choose an orthonormal basis such that (5.6) will hold with the norm defined with respect to this basis and with B and C independent of this choice and of Γ.

Let k be the dimension of $\Gamma_{\mathbb{R}}$. It will be convenient to denote the standard basis of $\mathbb{R}^{n+2}$ by $\{\mathbf{e}_0, \mathbf{e}_*\mathbf{e}_1, \dots, \mathbf{e}_n\}$. Since $D\Gamma_{\mathbb{R}}$ is a k-dimensional subspace of $\mathbb{R}\mathbf{e}_0 \oplus \mathbb{R}\mathbf{e}_1 \oplus \cdots \oplus \mathbb{R}\mathbf{e}_n$, it is possible to choose an orthonormal set $\mathbf{v}_1, \dots, \mathbf{v}_{k-1} \in D\Gamma_{\mathbb{R}}$ such that each $\mathbf{v}_i$, $i = 1, \dots, k-1$, is orthogonal to $\mathbf{e}_0$. Now let us consider two cases:

Case 1. $D\Gamma_{\mathbb{R}}$ contains $\mathbf{e}_0$; then $\{\mathbf{e}_0, \mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}$ is a basis of $\mathbb{R}\mathbf{e}_0 \oplus D\Gamma_{\mathbb{R}}$. Thus one can find $\mathbf{w} \in \bigwedge^k(\mathbb{R}^l)$ representing Γ such that $D\mathbf{w}$ can be written as $a\mathbf{e}_0 \wedge \mathbf{v}_1 \cdots \wedge \mathbf{v}_{k-1}$ for some $a \geq 1$.

Case 2. $D\Gamma_{\mathbb{R}}$ does not contain $\mathbf{e}_0$; then it is possible to choose $\mathbf{v}_k \in \mathbb{R}\mathbf{e}_0 \oplus D\Gamma_{\mathbb{R}}$ such that $\{\mathbf{e}_0, \mathbf{v}_1, \dots, \mathbf{v}_k\}$ is an orthonormal basis of $\mathbb{R}\mathbf{e}_0 \oplus D\Gamma_{\mathbb{R}}$. In this case, one can represent Γ by $\mathbf{w}$ such that

$$D\mathbf{w} = (a\mathbf{e}_0 + b\mathbf{v}_k) \wedge \mathbf{v}_1 \cdots \wedge \mathbf{v}_{k-1} = a\mathbf{e}_0 \wedge \mathbf{v}_1 \cdots \wedge \mathbf{v}_{k-1} + b\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k \tag{5.7}$$

for some $a, b \in \mathbb{R}$ with $a^2 + b^2 \geq 1$. In fact (5.7) is valid in Case 1 as well; one simply has to put b equal to zero and bear in mind that the vector $\mathbf{v}_k$ is not defined.

Now write

$$H(x)\mathbf{w} = DU_xD^{-1}(D\mathbf{w})\,, \tag{5.8}$$

and introduce the $(n+2)$-tuple of functions

$$\tilde{\mathbf{f}}(x) = \big(1, 0, \frac{d_1}{d_0}f_1(x), \dots, \frac{d_n}{d_0}f_n(x)\big)\,.$$

It is straightforward to verify that $DU_xD^{-1}\mathbf{e}_0 = \mathbf{e}_0$ and

$$DU_xD^{-1}\mathbf{v} = \mathbf{v} + \big(\tilde{\mathbf{f}}(x) \cdot \mathbf{v}\big)\mathbf{e}_0 + \frac{d_0}{d_*}\big(\tilde{\mathbf{f}}'(x) \cdot \mathbf{v}\big)\mathbf{e}_*$$

whenever $\mathbf{v}$ is orthogonal to $\mathbf{e}_0$ and $\mathbf{e}_*$. Therefore the space $\mathbb{R}\mathbf{e}_0 \oplus \mathbb{R}\mathbf{e}_* \oplus D\Gamma_{\mathbb{R}}$ is invariant under DU_xD^{-1}; hence we can restrict ourselves to the coordinates of DU_xD^{-1}-image of $D\mathbf{w}$ with respect to the basis chosen above, and write

$$DU_xD^{-1}(\mathbf{e}_0 \wedge \mathbf{v}_1 \cdots \wedge \mathbf{v}_{k-1}) = \mathbf{e}_0 \wedge \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_{k-1} + \frac{d_0}{d_*}\sum_{i=1}^{k-1} \pm\big(\mathbf{v}_i \cdot \tilde{\mathbf{f}}'(x)\big)\,\mathbf{e}_0 \wedge \mathbf{e}_* \wedge \bigwedge_{s \neq i} \mathbf{v}_s$$

and

$$DU_xD^{-1}(\mathbf{v}_1\wedge\cdots\wedge\mathbf{v}_k) \;=\; \mathbf{v}_1\wedge\cdots\wedge\mathbf{v}_k \;+\; \sum_{i=1}^{k}\pm\big(\mathbf{v}_i\cdot\tilde{\mathbf{f}}(x)\big)\,\mathbf{e}_0\wedge\bigwedge_{s\neq i}\mathbf{v}_s$$
$$+\;\frac{d_0}{d_*}\sum_{i=1}^{k}\pm\big(\mathbf{v}_i\cdot\tilde{\mathbf{f}}'(x)\big)\,\mathbf{e}_*\wedge\bigwedge_{s\neq i}\mathbf{v}_s \;+\; \frac{d_0}{d_*}\sum_{\substack{i=1\\ j>i}}^{k}\pm\begin{vmatrix}\mathbf{v}_i\cdot\tilde{\mathbf{f}}(x) & \mathbf{v}_j\cdot\tilde{\mathbf{f}}(x)\\ \mathbf{v}_i\cdot\tilde{\mathbf{f}}'(x) & \mathbf{v}_j\cdot\tilde{\mathbf{f}}'(x)\end{vmatrix}\,\mathbf{e}_0\wedge\mathbf{e}_*\wedge\bigwedge_{s\neq i,j}\mathbf{v}_s$$

Collecting terms and using (5.7) and (5.8), one finds that

$$\begin{aligned}H(x)\mathbf{w} \;=\;& \Big(a+b\big(\mathbf{v}_k\cdot\tilde{\mathbf{f}}(x)\big)\Big)\,\mathbf{e}_0\wedge\mathbf{v}_1\wedge\cdots\wedge\mathbf{v}_{k-1} \;+\; b\,\mathbf{v}_1\wedge\cdots\wedge\mathbf{v}_k\\ &+\; b\sum_{i=1}^{k-1}\pm\big(\mathbf{v}_i\cdot\tilde{\mathbf{f}}(x)\big)\,\mathbf{e}_0\wedge\bigwedge_{s\neq i}\mathbf{v}_s \;+\; b\frac{d_0}{d_*}\sum_{i=1}^{k}\pm\big(\mathbf{v}_i\cdot\tilde{\mathbf{f}}'(x)\big)\,\mathbf{e}_*\wedge\bigwedge_{s\neq i}\mathbf{v}_s\\ &+\;\frac{d_0}{d_*}\sum_{i=1}^{k-1}\pm\begin{vmatrix}\mathbf{v}_i\cdot\tilde{\mathbf{f}}(x) & b\mathbf{v}_k\cdot\tilde{\mathbf{f}}(x)\pm a\\ \mathbf{v}_i\cdot\tilde{\mathbf{f}}'(x) & b\mathbf{v}_k\cdot\tilde{\mathbf{f}}'(x)\end{vmatrix}\,\mathbf{e}_0\wedge\mathbf{e}_*\wedge\bigwedge_{s\neq i,k}\mathbf{v}_s\\ &+\; b\frac{d_0}{d_*}\sum_{\substack{i,j=1\\ j>i}}^{k-1}\pm\begin{vmatrix}\mathbf{v}_i\cdot\tilde{\mathbf{f}}(x) & \mathbf{v}_j\cdot\tilde{\mathbf{f}}(x)\\ \mathbf{v}_i\cdot\tilde{\mathbf{f}}'(x) & \mathbf{v}_j\cdot\tilde{\mathbf{f}}'(x)\end{vmatrix}\,\mathbf{e}_0\wedge\mathbf{e}_*\wedge\bigwedge_{s\neq i,j}\mathbf{v}_s\end{aligned}\tag{5.9}$$

It is clear that the coefficients in the first two lines are linear combinations of either $1, f_1, \dots, f_n$ or $f_1', \dots, f_n'$, hence one can, using Corollary 4.3, find $B_1 \ni x_0$ and $C_1 > 0$ such that all these functions are $(C_1, 1/n)$-good on B_1. The same can be said about the rest of the coefficients when $b = 0$. Otherwise one can observe that the coefficients in the last two lines of (5.9) are of the form $\pm A_{L\mathbf{g}}$ where L is some linear transformation of $\mathbb{R}^2$ and $\mathbf{g}$ belongs to

$$\mathcal{G} \stackrel{\text{def}}{=} \big\{\big(\mathbf{u}_1\cdot\mathbf{f}(x), \mathbf{u}_2\cdot\mathbf{f}(x)+u_0\big) \mid u_0\in\mathbb{R}, \mathbf{u}_1\perp\mathbf{u}_2\in\mathbb{R}^n_1\big\}\,. \tag{5.10}$$

Since $A_{L\mathbf{g}}$ is proportional to $A_{\mathbf{g}}$, it suffices, in view of Lemma 4.1(a), to prove the existence of $C \geq C_1$ and a neighborhood $B \subset B_1$ of x_0 such that the functions $A_{\mathbf{g}}$ are $(C, \frac{1}{2n-1})$-good on B for any $\mathbf{g} \in \mathcal{G}$. The latter is a direct consequence of Proposition 4.4(a), since the family $\{\mathbf{g}' \mid \mathbf{g} \in \mathcal{G}\}$ is compact in $C^n(V)$, and it follows from the non-degeneracy of $\mathbf{f}$ that condition (4.4b) is satisfied. It remains to apply Lemma 4.1(b) and conclude that the function $x \mapsto \|H(x)\Gamma\|$ (which is equal to the supremum of all the coefficients in (5.9) provided the norm is defined with respect to the basis chosen above) is $(C, \frac{1}{2n-1})$-good on B. $\square$

Lemma 5.3. *For any subinterval B of V containing x_0 there exists $\rho > 0$ such that for any $0 < d_0 \leq 1$, $d_* > 0$, $d_n \geq \cdots \geq d_1 \geq 1$ such that*

$$d_0 d_* d_1 \dots d_{n-1} \leq 1 \tag{5.11}$$

and any $\Gamma \in \mathcal{L}(\Lambda)$, one has

$$\sup_{x\in B}\|H(x)\Gamma\| \geq \rho\,. \tag{5.12}$$

Proof. Take any neighborhood $B \subset V$ of x_0. It follows from the linear independence of the functions $1, f_1, \ldots, f_n$ and from the linear independence of their derivatives that there exists $\rho_1 > 0$ such that

$$\forall\, \mathbf{v} \in \mathbb{R}^n_1 \ \forall\, v_0 \in \mathbb{R} \text{ one has } \sup_{x\in B} |\mathbf{v}\cdot\mathbf{f}(x) + v_0| \geq \rho_1 \text{ and } \sup_{x\in B} |\mathbf{v}\cdot\mathbf{f}'(x)| \geq \rho_1\,. \tag{5.13}$$

Also let $\rho_2 = \rho(\mathcal{G}, B)$ where $\mathcal{G}$ is the class of 2-tuples of functions defined in (5.10), and $\rho(\mathcal{G}, B)$ is as in Proposition 4.4(b). Consider

$$M \stackrel{\text{def}}{=} \max\big(\sup_{x\in B}\|\mathbf{f}(x)\|, \sup_{x\in B}\|\mathbf{f}'(x)\|\big)\,;$$

we will show that (5.12) will hold for any nonzero subgroup Γ of Λ if we choose

$$\rho = \frac{\rho_1\rho_2}{\sqrt{\rho_1^2 + (\rho_2 + 2M^2)^2}}\,.$$

First let us consider the case $k = \dim(\Gamma_{\mathbb{R}}) = 1$. Then Γ can be represented by a vector $\mathbf{v} = (v_0, 0, v_1, \ldots, v_n)$ with integer coordinates, and it is straightforward to verify that the first coordinate of $H(x)\mathbf{v}$ will be equal to $\frac{1}{d_0}(v_0 + v_1 f_1(x) + \cdots + v_n f_n(x))$, which will deviate from zero by not less than ρ at some point of B due to (5.13) and since $\rho \leq \rho_1$ and $d_0 \leq 1$.

Now let k be greater than 1. Our method will be similar to that of the proof of the previous lemma: given Γ, we will choose a suitable orthonormal subset of $\bigwedge^k(\mathbb{R}^{n+2})$ and then show that the absolute value of the projection of $H(x)\mathbf{w}$ to at least one of the components will be not less than ρ for some $x \in B$.

In order to prove the desired estimate, it is important to pay special attention to the vector $\mathbf{e}_n$, which is the eigenvector of D with the smallest eigenvalue. We do it by first choosing an orthonormal set $\mathbf{v}_1, \ldots, \mathbf{v}_{k-2} \in \Gamma_{\mathbb{R}}$ such that each $\mathbf{v}_i$, $i = 1, \ldots, k-2$, is orthogonal to both $\mathbf{e}_0$ and $\mathbf{e}_n$. Then choose $\mathbf{v}_{k-1}$ orthogonal to $\mathbf{v}_i$, $i = 1, \ldots, k-2$, and to $\mathbf{e}_0$ (but in general not to $\mathbf{e}_n$). Now, if necessary (see the discussion in the proof of Lemma 5.2), choose a vector $\mathbf{v}_k$ to complete $\{\mathbf{e}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\}$ to an orthonormal basis of $\mathbb{R}\mathbf{e}_0 \oplus \Gamma_{\mathbb{R}}$. This way, similarly to (5.7), we will represent Γ by $\mathbf{w}$ of the form

$$\mathbf{w} = (a\mathbf{e}_0 + b\mathbf{v}_k) \wedge \mathbf{v}_1 \cdots \wedge \mathbf{v}_{k-1} = a\mathbf{e}_0 \wedge \mathbf{v}_1 \cdots \wedge \mathbf{v}_{k-1} + b\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k \tag{5.14}$$

for some $a, b \in \mathbb{R}$ with $a^2 + b^2 \geq 1$. As before, we will use (5.14) even when $\mathbf{v}_k$ is not defined, in this case the coefficient b will vanish.

Now, similarly to the proof of Lemma 5.2, introduce the $(n+2)$-tuple of functions

$$\hat{\mathbf{f}}(x) = \big(1, 0, f_1(x), \ldots, f_n(x)\big)\,,$$

and observe that $U_x\mathbf{e}_0 = \mathbf{e}_0$ and

$$U_x\mathbf{v} = \mathbf{v} + \big(\hat{\mathbf{f}}(x)\cdot\mathbf{v}\big)\mathbf{e}_0 + \big(\hat{\mathbf{f}}'(x)\cdot\mathbf{v}\big)\mathbf{e}_*$$

whenever $\mathbf{v}$ is orthogonal to $\mathbf{e}_0$ and $\mathbf{e}_*$. Using this and (5.14), one can obtain an expression analogous to (5.9). This time however we are interested only in the terms of the form

$\mathbf{e}_0 \wedge \mathbf{e}_* \wedge \bigwedge^{k-2}\big((\mathbb{R}\mathbf{e}^0 \oplus \mathbb{R}\mathbf{e}_*)^\perp\big)$ (note that these terms are present only if $k \geq 2$). Namely, let us write

$$U_x\mathbf{w} \;=\; \Big(a + b\big(\mathbf{v}_k \cdot \hat{\mathbf{f}}(x)\big)\Big)\; \mathbf{e}_0 \wedge \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_{k-1} \;+\; b\,\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k$$
$$+\, b\sum_{i=1}^{k-1} \pm\big(\mathbf{v}_i \cdot \hat{\mathbf{f}}(x)\big)\,\mathbf{e}_0 \wedge \bigwedge_{s\neq i} \mathbf{v}_s \;+\; b\sum_{i=1}^{k} \pm\big(\mathbf{v}_i \cdot \hat{\mathbf{f}}'(x)\big)\,\mathbf{e}_* \wedge \bigwedge_{s\neq i} \mathbf{v}_s \;+\; \mathbf{e}_0 \wedge \mathbf{e}_* \wedge \hat{\mathbf{w}}(x)\,,$$

where

(5.15)
$$\hat{\mathbf{w}}(x) \stackrel{\text{def}}{=} \sum_{i=1}^{k-1} \pm \begin{vmatrix} \mathbf{v}_i \cdot \hat{\mathbf{f}}(x) & b\mathbf{v}_k \cdot \hat{\mathbf{f}}(x) \pm a \\ \mathbf{v}_i \cdot \hat{\mathbf{f}}'(x) & b\mathbf{v}_k \cdot \hat{\mathbf{f}}'(x) \end{vmatrix} \bigwedge_{s\neq i,k} \mathbf{v}_s \;+\; b \sum_{\substack{i,j=1\\ j>i}}^{k-1} \pm \begin{vmatrix} \mathbf{v}_i \cdot \hat{\mathbf{f}}(x) & \mathbf{v}_j \cdot \hat{\mathbf{f}}(x) \\ \mathbf{v}_i \cdot \hat{\mathbf{f}}'(x) & \mathbf{v}_j \cdot \hat{\mathbf{f}}'(x) \end{vmatrix} \bigwedge_{s\neq i,j} \mathbf{v}_s\,.$$

Note that $\hat{\mathbf{w}}(x)$ lies in the space $\bigwedge^{k-2}\big((\mathbb{R}\mathbf{e}^0 \oplus \mathbb{R}\mathbf{e}_*)^\perp\big)$; since $\mathbf{e}_0$ and $\mathbf{e}_*$ are eigenvectors of D (with eigenvalues d_0^{-1} and d_*^{-1} respectively), both the space $\mathbf{e}_0 \wedge \mathbf{e}_* \wedge \bigwedge^{k-2}\big((\mathbb{R}\mathbf{e}^0 \oplus \mathbb{R}\mathbf{e}_*)^\perp\big)$ and its orthogonal complement are D-invariant, and clearly $U_x\mathbf{w} - \mathbf{e}_0 \wedge \mathbf{e}_* \wedge \hat{\mathbf{w}}(x)$ lies in this orthogonal complement. Therefore it will suffice to prove that $\sup_{x\in B}\|D\hat{\mathbf{w}}(x)\|$ is not less than $d_0 d_* \rho$.

In the argument below, we will switch to the Euclidean norm $\|\cdot\|$ on $\bigwedge(\mathbb{R}^{n+2})$. Consider the product $\mathbf{e}_n \wedge \hat{\mathbf{w}}(x)$. We claim that it is enough to show that

$$\|\mathbf{e}_n \wedge \hat{\mathbf{w}}(x)\| \geq \rho \text{ for some } x \in B\,. \tag{5.16}$$

Indeed, since $\mathbf{e}_n$ is an eigenvector of D with eigenvalue d_n^{-1}, for any $x \in B$ the norm of $D\big(\mathbf{e}_n \wedge \hat{\mathbf{w}}(x)\big)$ is not greater than $d_n^{-1}\|D\hat{\mathbf{w}}(x)\|$. Therefore, since the smallest eigenvalue of D on $\bigwedge^{k-1}(\Lambda_\mathbb{R})$ is equal to $d_{n-k+1}\cdot\ldots\cdot d_n$, the norm of $D\hat{\mathbf{w}}(x)$ is not less than

$$d_n\|D\big(\mathbf{e}_n \wedge \hat{\mathbf{w}}(x)\big)\| \geq \frac{d_n}{d_{n-k+1}\cdot\ldots\cdot d_n}\|\mathbf{e}_n \wedge \hat{\mathbf{w}}(x)\|$$
$$\underset{\text{for some } x\in B,\text{ by (5.16)}}{\geq} \frac{\rho}{d_{n-k+1}\cdot\ldots\cdot d_{n-1}} \underset{\text{since } d_i\geq 1}{\geq} \frac{d_0 d_*\rho}{d_0 d_* d_1\cdot\ldots\cdot d_{n-1}} \underset{\text{by (5.11)}}{\geq} d_0 d_*\rho\,.$$

Thus it remains to prove (5.16). For this let us select the term containing $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_{k-2}$, and multiply (5.15) by $\mathbf{e}_n$ as follows:

$$\begin{gathered}\mathbf{e}_n \wedge \hat{\mathbf{w}}(x) = \pm A(x)\mathbf{e}_n \wedge \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_{k-2}\\ +\text{ other terms where one or two of } \mathbf{v}_i,\ i = 1,\ldots,k-2, \text{ are missing}\,,\end{gathered} \tag{5.17}$$

where

$$A(x) \stackrel{\text{def}}{=} \begin{vmatrix} \mathbf{v}_{k-1} \cdot \hat{\mathbf{f}}(x) & b\mathbf{v}_k \cdot \hat{\mathbf{f}}(x) \pm a \\ \mathbf{v}_{k-1} \cdot \hat{\mathbf{f}}'(x) & b\mathbf{v}_k \cdot \hat{\mathbf{f}}'(x) \end{vmatrix} = b \begin{vmatrix} \mathbf{v}_{k-1} \cdot \hat{\mathbf{f}}(x) & \mathbf{v}_k \cdot \hat{\mathbf{f}}(x) \\ \mathbf{v}_{k-1} \cdot \hat{\mathbf{f}}'(x) & \mathbf{v}_k \cdot \hat{\mathbf{f}}'(x) \end{vmatrix} \pm a\mathbf{v}_{k-1} \cdot \hat{\mathbf{f}}'(x)$$

Because of the orthogonality of the two summands in (5.17), and also because $\mathbf{e}_n$ is orthogonal to $\mathbf{v}_i$, $i = 1,\ldots,k-2$, it follows that $\|\mathbf{e}_n \wedge \hat{\mathbf{w}}(x)\|$ is not less than $|A(x)|$. It follows from the first expression for $A(x)$ that $\sup_{x\in B}|A(x)| \geq \rho_2 b$, and from the second one that $\sup_{x\in B}|A(x)| \geq \rho_1 a - 2M^2 b$. An elementary computation shows that ρ as defined by (5.13) is not greater than $\min_{a^2+b^2\geq 1}\max(\rho_2 b, \rho_1 a - 2M^2 b)$. This completes the proof of (5.16), and hence of Lemma 5.3. $\square$

Proof of Theorem 1.4. First note that without loss of generality one can assume that $Q_1 \leq \cdots \leq Q_n$. (Otherwise one can replace $Q_1,\ldots,Q_n$ by a permutation $Q_{i_1},\ldots,Q_{i_n}$ with $Q_{i_1} \leq \cdots \leq Q_{i_n}$, and consider the n-tuple $(f_{i_1},\ldots,f_{i_n})$ instead of the original one, which will still be non-degenerate at x_0.) This way, the middle expression in (1.7a) becomes equal to $\delta K Q_1 \cdot \ldots \cdot Q_{n-1}$, and one has to find $B \subset V$ and $E, \rho > 0$ such that the measure of the set (1.7b) is at most $E(\delta K Q_1 \cdot \ldots \cdot Q_{n-1})^{\frac{1}{(\varkappa+1)(2n-1)}}|B|$ for any choice of $\varkappa \geq n$, $\delta, K > 0$ and $Q_1 \geq \ldots Q_1 \geq 1$ satisfying $\delta^{\varkappa} \leq K Q_1 \cdot \ldots \cdot Q_{n-1} \leq \frac{\rho^{\varkappa+1}}{\delta}$.

Now take B and C as in Lemma 5.2, denote by ρ the minimum of $\frac{1}{n+2}$ and ρ as in Lemma 5.3, choose $\varepsilon = (\delta K Q_1 \cdot \ldots \cdot Q_{n-1})^{\frac{1}{\varkappa+1}}$, and define $d_0, d_*, d_1, \ldots, d_n$ as in (5.4). Then

$$d_0 = \frac{\delta}{\varepsilon} = \frac{\delta^{\frac{\varkappa}{\varkappa+1}}}{(KQ_1 \cdot \ldots \cdot Q_{n-1})^{\frac{1}{\varkappa+1}}} = \left(\frac{\delta^{\varkappa}}{KQ_1 \cdot \ldots \cdot Q_{n-1}}\right)^{\frac{1}{\varkappa+1}} \leq 1\,,$$

and also

$$d_0 d_* d_1 \ldots d_{n-1} = \frac{\delta K Q_1 \cdot \ldots \cdot Q_{n-1}}{\varepsilon^{n+1}} = \big(\delta K Q_1 \cdot \ldots \cdot Q_{n-1}\big)^{1-\frac{n+1}{\varkappa+1}} \leq \rho^{\varkappa-n} < 1\,.$$

Thus Lemma 5.3 applies and we get $\sup_{x\in B}\|H(x)\Gamma\| \geq \rho$ for all $\Gamma \in \mathcal{L}(\Lambda)$. Also Lemma 5.2 says that the functions $x \mapsto \|H(x)\Gamma\|$ are $(C, \frac{1}{2n-1})$-good on B for all $\Gamma \in \mathcal{L}(\Lambda)$. Therefore conditions (i) and (ii) of Theorem 5.1, with $l = n+2$ and $\alpha = \frac{1}{2n-1}$, are satisfied, so whenever $\varepsilon \leq \rho$ (which reduces to the second inequality in (1.7a)), one has

$$|B_{H,\Lambda,\varepsilon}| \leq (n+1)2^{n+1}C\left(\frac{\varepsilon}{\rho}\right)^{\frac{1}{2n-1}}|B| = (n+1)2^{n+1}C\rho^{\frac{1}{2n-1}}(KQ_1 \cdot \ldots \cdot Q_{n-1})^{\frac{1}{(\varkappa+1)(2n-1)}}|B|\,.$$

It remains to observe that, as it was mentioned in the discussion after Theorem 5.1, for any $\varepsilon > 0$ the set $B_{H,\Lambda,\varepsilon}$, with Λ as in (5.1) and H as in (5.5), coincides with the set (1.7b) if one uses the values of $d_0, d_*, d_1, \ldots, d_n$ as in (5.4). $\square$

6. Concluding remarks

6.1. The main result of the paper (Theorem 1.1) was proved already in the summer of 1998, but only in the case when the functions $f_1,\ldots,f_n$ are analytic. More precisely, in addition to the non-degeneracy of $\mathbf{f}$, we had to assume that there exist positive constants C and α such that for almost all $x \in V$ one can find a subinterval B of V containing x such that

$$\text{Span}\left(f_1',\ldots,f_n';\begin{vmatrix} f_i & f_j \\ f_i' & f_j'\end{vmatrix}, 1 \leq i < j \leq n\right) \text{consists of functions } (C,\alpha)\text{-good on B}\,.$$

For analytic functions $f_1,\ldots,f_n$ this condition can be easily verified by applying Corollary 4.3(a) to the basis of the above function space.

In our original approach we considered sets more general than (1.7b), namely the sets

$$\left\{x \in B \;\middle|\; \exists\, \mathbf{q} \in \mathbb{Z}^n \smallsetminus \{0\} \text{ such that } \begin{cases} \lfloor \mathbf{q}\cdot\mathbf{f}(x) \rfloor < \delta \\ |\mathbf{q}\cdot\mathbf{g}(x)| < K \\ |q_i| < Q_i, \quad i = 1,\dots,n \end{cases} \right\}, \tag{6.1}$$

where $\mathbf{f}$ is as in Theorem 1.4 and $\mathbf{g}$ is another non-degenerate n-tuple of functions on V. We were able to prove an analogue of Theorem 1.4 for sets (6.1) but only for $K \geq 1$, and for n-tuples $\mathbf{f}$ and $\mathbf{g}$ with an additional assumption that

$$\operatorname{Span}\left(g_1,\dots,g_n; \begin{vmatrix} f_i & f_j \\ g_i & g_j \end{vmatrix}, 1 \leq i < j \leq n\right) \text{ consists of } (C,\alpha)\text{-good functions}. \tag{6.2}$$

Instead of U_x as in (5.2) we considered more general matrices $U_x^{\mathbf{f},\mathbf{g}} \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 0 & \mathbf{f}(x)^T \\ 0 & 1 & \mathbf{g}(x)^T \\ 0 & 0 & I_n \end{pmatrix}$. To prove an analogue of Lemma 5.2, or, more precisely, the statement that for some positive C, α and any $\Gamma \in \mathcal{L}(\Lambda)$ the function $x \mapsto \|DU_x^{\mathbf{f},\mathbf{g}}\Gamma\|$ is (C,α)-good on some neighborhood B of x_0, it was enough to consider the standard basis $\{\mathbf{e}_0, \mathbf{e}_*\mathbf{e}_1,\dots,\mathbf{e}_n\}$ of $\mathbb{R}^{n+2}$ and the corresponding basis

$$\left\{\mathbf{e}_I \stackrel{\text{def}}{=} \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k} \mid I = \{i_1,\dots,i_k\} \subset \{0,*,1,\dots,n\}\right\}$$

of $\bigwedge^k(\mathbb{R}^{n+2})$, decompose an element $\mathbf{w}$ representing Γ as $\mathbf{w} = \sum_I w_I \mathbf{e}_I$, write an expansion similar to (5.9) and use (6.2).

To prove an analogue of Lemma 5.3, that is ths statement that for any neighborhood B of x_0 there exists $\rho > 0$ such that $\sup_{x\in B} \|DU_x^{\mathbf{f},\mathbf{g}}\Gamma\| \geq \rho$ for every $\Gamma \in \mathcal{L}(\Lambda)$, we used the fact that the coefficients w_I are integers and considered the following two cases: 1) $w_I \neq 0$ for some $I \subset \{1,\dots,n\}$, and 2) $w_I = 0$ for all $I \subset \{1,\dots,n\}$. In the just described approach it was important that $K \geq 1$. This was enough for the proof of Theorem 1.1; however we will see in §6.4 that the stronger version, allowing arbitrarily small positive values of K, is important for other applications.

6.2. As it was mentioned in the introduction, it is possible to prove the complementary divergence counterpart of part (S) of Theorem 1.1: that is, assuming the divergence of the series $\sum_{k=1}^\infty \psi(k)$ and with a curve M as in Theorem 1.1, almost all points of M are ψ-A. The proof is based on the method of *regular systems*, which dates back to [BS] and has been extensively used in the existing proofs of divergence Khintchine-type results for special classes of manifolds [DRV3, BBDD, Be1, Be3]. We plan a detailed exposition of the method and its possible generalizations and applications in a forthcoming paper.

6.3. It seems natural to conjecture that in Corollary 1.2, the assumption (1.5) can be replaced by the non-degeneracy of M, and, furthermore, that it is not essential for the functions $f_1,\dots,f_n$ parametrizing M to be n times differentiable[3]; the only thing that should matter

[3]Recall that in [DRV2, DRV3] it was proved that statement (S), as well as in some cases its divergence counterpart, holds for a certain class of C^3 manifolds of big enough dimension, with some curvature conditions imposed.

is, as in the setting of Theorem A, having enough derivatives to span $\mathbb{R}^n$. Note that the necessity to foliate the manifold by non-degenerate curves was avoided in the paper [KM1], where the main measure estimate was proved for a subset V of $\mathbb{R}^d$, $d \ge 1$. It seems plausible that one can prove d-dimensional analogues of Theorems 1.3 and 1.4, thus demonstrating (S) and (D) for any non-degenerate C^k submanifold $M \subset \mathbb{R}^n$, possibly with $k < n$. It is worthwhile to mention that after the present work was completed the authors received a preprint [Be5] of V. Beresnevich where the statement (S) (but not (M)) was proved for a non-degenerate submanifold $\mathbf{f}(V)$ of $\mathbb{R}^n$ without assuming existence of derivatives of $\mathbf{f}$ up to the nth order.

6.4. The idea to study the set of points x such that $F(x) \stackrel{\text{def}}{=} \mathbf{q}\cdot\mathbf{f}(x)$ is close to an integer by looking at the values of the derivative $F'(x) = \mathbf{q}\cdot\mathbf{f}'(x)$ of F has a long history. It was extensively used by Sprindžuk in his proof of Mahler's Conjecture [Sp2, Sp4], that is, when $f_i(x) = x^i$. Also from a paper of A. Baker and W. Schmidt [BS] it follows that for some $\gamma, \varepsilon > 0$, on a set of positive measure the system

$$\begin{cases} |P(x)| < \|\mathbf{q}\|^{-n+\gamma} \\ |P'(x)| < \|\mathbf{q}\|^{1-\gamma-\varepsilon} \end{cases} \tag{6.3}$$

(here $P(x)$ is the polynomial $p + q_1x + \cdots + q_nx^n$) has at most finitely many solutions $p \in \mathbb{Z}$, $\mathbf{q} \in \mathbb{Z}^n$. This was used to construct a certain regular system of real numbers and obtain the sharp lower estimate for the Hausdorff dimension of the set

$$\left\{x \in \mathbb{R} \;\middle|\; \lfloor q_1x + \cdots + q_nx^n \rfloor < \|\mathbf{q}\|^{-\lambda} \text{ for infinitely many } \mathbf{q} \in \mathbb{Z}^n\right\}$$

for $\lambda > n$. Note also that the system (6.3) is related to the distribution of values of discriminants of integer polynomials, see [D, Sp2, Bern].

In 1995 V. Borbat [Bo] proved that given any $\varepsilon > 0$ and $0 < \gamma < 1$, for almost all x there are at most finitely many solutions of (6.3). Now we can use Theorem 1.4 to relax the restriction $\gamma < 1$ (here it becomes important that K in Theorem 1.4 can be taken to be arbitrarily small). More precisely, we derive the following generalization and strengthening of the aforementioned result of Borbat:

Theorem. *Let $V \subset \mathbb{R}$ be an open subset, and let $\mathbf{f} = (f_1, \dots, f_n)$ be a non-degenerate n-tuple of C^n functions on V. Take $\varepsilon > 0$ and $0 < \gamma < n$. Then for almost all $x \in V$ there exist at most finitely many solutions $p \in \mathbb{Z}$, $\mathbf{q} \in \mathbb{Z}^n$, of the system*

$$\begin{cases} \lfloor \mathbf{q}\cdot\mathbf{f}(x) \rfloor < \Pi_+(\mathbf{q})^{-1+\gamma/n} \\ |\mathbf{q}\cdot\mathbf{f}'(x)| < \|\mathbf{q}\|^{1-\gamma-\varepsilon} \end{cases} \tag{6.4}$$

The reduction of this theorem to Theorem 1.4 is similar to the proof of Part (M) of Theorem 1.1 (see §2) and is left to the reader. The above theorem naturally invites one to think about a possibility of Khintchine-type results involving derivative estimates; that is, replacing (6.4) by, say,

$$\begin{cases} \lfloor \mathbf{q}\cdot\mathbf{f}(x) \rfloor < \psi_1(\|\mathbf{q}\|) \text{ or } \psi_1(\Pi_+(\mathbf{q})) \\ |\mathbf{q}\cdot\mathbf{f}'(x)| < \psi_2(\|\mathbf{q}\|) \end{cases} \tag{6.5}$$

and finding optimal conditions on ψ_i implying at most finitely many (or, conversely, infinitely many) solutions of (6.5) for almost all x.

7. Appendix: Proof of Theorem 5.1

Our goal here is to describe a simplified version of a combinatorial construction which was used in [KM1] to prove the main measure estimate. For a partially ordered set (*poset*) S, we will denote by $l(S)$ the *length* of S (i.e. the number of elements in a maximal linearly ordered subset of S). If T is a subset of S, we let $S(T)$ be the poset of elements of $S \smallsetminus T$ comparable with any element of T. Note that one always has

$$l\big(S(T)\big) \le l(S) - l(T)\,. \tag{7.1}$$

If f is a function on B, we will use the notation $\|f\|_B \stackrel{\text{def}}{=} \sup_{x\in B} |f(x)|$.

For $k \in \mathbb{Z}_+$ and $C, \alpha, \rho > 0$, define $\mathcal{A}(k, C, \alpha, \rho)$ to be the set of triples (S, φ, B) where S is a poset, $B \subset \mathbb{R}$ is an interval, and φ is a mapping from S to the space of continuous functions on B (this mapping will be denoted by $s \to \varphi_s$) such that the following holds:

(A0) $l(S) \le k$;
(A1) $\forall\, s \in S\,,\quad \varphi_s$ is (C, α)-good on B;
(A2) $\forall\, s \in S\,,\quad \|\varphi_s\|_B \ge \rho$;
(A3) $\forall\, x \in B,\quad \#\{s \in S \mid |\varphi_s(x)| < \rho\} < \infty$.

Then, given $(S, \varphi, B) \in \mathcal{A}(k, C, \alpha, \rho)$ and $\varepsilon > 0$, say that a point $z \in B$ is $(\varepsilon, S, \varphi)$-*marked* if there exists a linearly ordered subset Σ_z of S such that

(M1) $\varepsilon \le |\varphi_s(z)| \le \rho \quad \forall\, s \in \Sigma_z$;
(M2) $|\varphi_s(z)| \ge \rho \quad \forall\, s \in S(\Sigma_z)$.

We will denote by $\Phi(\varepsilon, S, \varphi, B)$ the set of all the $(\varepsilon, S, \varphi)$-marked points $z \in B$.

Theorem 7.1 (cf. [KM1, Theorem 4.1]). *Let $k \in \mathbb{Z}_+$ and $C, \alpha, \rho > 0$ be given. Then for all $(S, \varphi, B) \in \mathcal{A}(k, C, \alpha, \rho)$ and $\varepsilon > 0$ one has*

$$\big|\big(B \smallsetminus \Phi(\varepsilon, S, \varphi, B)\big)\big| \le kC2^k \left(\frac{\varepsilon}{\rho}\right)^{\alpha} |B|\,.$$

Proof. Our argument here is almost identical to the proof of Theorem 4.1 from [KM1], with minor simplifications. We proceed by induction on k. First take $(S, \varphi, B) \in \mathcal{A}(0, C, \alpha, \rho)$; since the poset S is empty, all points of B are $(\varepsilon, S, \varphi)$-marked for any $\varepsilon > 0$, which means that in the case $k = 0$ the claim is trivial. Now take $j \ge 1$ and suppose that the theorem is proved for $k = j - 1$; put $k = j$ and take $(S, \varphi, B) \in \mathcal{A}(d, k, C, \alpha, \rho)$ and a positive ε.

For any $x \in B$ we define

$$H(x) \stackrel{\text{def}}{=} \{s \in S \mid |\varphi_s(x)| < \rho\}$$

(this is a finite subset of S in view of (A3)), and let

$$E \stackrel{\text{def}}{=} \{x \in B \mid H(x) \ne \varnothing\} = \{x \in B \mid \exists\, s \in S \text{ with } |\varphi_s(x)| < \rho\}\,.$$

Observe that any point z of B which does not belong to E is $(\varepsilon, S, \varphi)$-marked for any positive ε, that is, $B \smallsetminus E \subset \Phi(\varepsilon, S, \varphi, B)$. (Indeed, by the definition of E, one has $|\varphi_s(z)| \ge \rho$

for all $s \in S$. Then one can take Σ_z to be the empty set and check that (M1) and (M2) are satisfied.) Therefore

$$B \smallsetminus \Phi(\varepsilon, S, \varphi, B) \subset E \smallsetminus \Phi(\varepsilon, S, \varphi, B)\,,$$

and to prove the theorem it remains to estimate the measure of the set in the right hand side.

We will do it by constructing a special covering $\{B_i\}$ of the set E. Take $x \in E$ and $s \in H(x)$, and define $B_{s,x}$ to be the maximal interval of the form $B \cap (x-r, x+r)$ on which the absolute value of φ_s is not greater than ρ. From the definition of $H(x)$ and the continuity of functions φ_s it follows that $B_{s,x}$ contains some neighborhood of x. Further, the maximality property of $B_{s,x}$ implies that $|\varphi_s(y)| = \rho$ for some $y \in B_{s,x}$, that is,

$$\|\varphi_s\|_{B_{s,x}} = \rho\,. \tag{7.2}$$

Indeed, either $B_{s,x} = B$ (then the claim follows from (A2)) or at one of the endpoints of $B_{s,x}$ $|\varphi_s|$ must attain the value ρ (otherwise one can enlarge the interval and still have $|\varphi_s|$ not greater than ρ).

Now for any $x \in E$ let us choose an element s_x of $H(x)$ such that $B_{s_x,x} = \bigcup_{s\in H(x)} B_{s,x}$ (this can be done since $H(x)$ is finite). For brevity we will denote $B_{s_x,x}$ by B_x. We now claim that

$$\|\varphi_s\|_{B_x} \ge \rho \text{ for any } x \in E \text{ and } s \in S\,. \tag{7.3}$$

Indeed, if $\|\varphi_s\|_{B_x} < \rho$, then one necessarily has $|\varphi_s(x)| < \rho$, hence $s \in H(x)$ and $B_{s,x}$ is defined. But $B_{s,x}$ is contained in B_x, so (7.3) follows from (7.2).

Next for any $x \in E$ we let $S_x \stackrel{\text{def}}{=} S(\{s_x\})$. From the fact that $(S, \varphi, B) \in \mathcal{A}(k, C, \alpha, \rho)$ it immediately follows that

$$(S_x, \varphi|_{S_x}, B_x) \in \mathcal{A}(d, k-1, C, \alpha, \rho) \text{ for any } x \in E\,. \tag{7.4}$$

Indeed, the properties (A1) and (A3) are straightforward, while (A0) and (A2) follow from (7.1) and (7.3) respectively.

Further, one can prove exactly as in [KM1, Lemma 4.6] that any point $z \in B_x$ which is $(\varepsilon, S_x, \varphi|_{S_x})$-marked, is $(\varepsilon, S, \varphi)$-marked provided $|\varphi_{s_x}(z)| \ge \varepsilon$. Equivalently,

$$B_x \smallsetminus \Phi(\varepsilon, S, \varphi, B) \subset B_x \smallsetminus \Phi(\varepsilon, S_x, \varphi|_{S_x}, B_x) \cup \{z \in B_x \bigm| |\varphi_{s_x}(z)| < \varepsilon\}\,,$$

Note that the measure of the first set in the right hand side is not greater than $(k-1)C2^{k-1}\left(\dfrac{\varepsilon}{\rho}\right)^{\alpha}|B_x|$ by (7.4) and the induction assumption, while that of the second one, in view of φ_{s_x} being (C,α)-good on B, is bounded by

$$C\left(\frac{\varepsilon}{\|\varphi_{s_x}\|_{B_x}}\right)^{\alpha}|B_x| \underset{(7.3)}{\le} C\left(\frac{\varepsilon}{\rho}\right)^{\alpha}|B_x|\,.$$

Consequently, one gets

$$|B_x \smallsetminus \Phi(\varepsilon, S, \varphi, B)| \le C\left((k-1)2^{k-1}+1\right)\left(\frac{\varepsilon}{\rho}\right)^{\alpha}|B_x| \le kC2^{k-1}\left(\frac{\varepsilon}{\rho}\right)^{\alpha}|B_x|\,.$$

Finally, consider the covering $\{B_x \mid x \in E\}$ of E and choose a subcovering $\{B_i\}$ of multiplicity at most 2. Then the measure of $E \smallsetminus \Phi(\varepsilon, S, \varphi, B)$ is not greater than

$$\sum_i |B_i \smallsetminus \Phi(\varepsilon, S, \varphi, B)| \le kC2^{k-1}\left(\frac{\varepsilon}{\rho}\right)^{\alpha} \sum_i |B_i| \le kC2^k \left(\frac{\varepsilon}{\rho}\right)^{\alpha} |B|$$

and the theorem is proven. $\square$

7.2. *Proof of Theorem 5.1.* We will apply Theorem 7.1 to the triple (S, φ, B), where $S = \mathcal{L}(\Lambda)$ is the partially ordered set of nonzero primitive subgroups of Λ and φ is defined by $\varphi_\Gamma(x) \overset{\text{def}}{=} \|h(x)\Gamma\|$. It is easy to verify that $(S, \varphi, B) \in \mathcal{A}\big(\dim(\Lambda_{\mathbb{R}}), C, \alpha, \rho\big)$. Indeed, the functions φ_Γ are continuous since so is h, the property (A0) is clear, (A1) is given by (i), (A2) by (ii), and (A3) follows from the discreteness of $\bigwedge(\Lambda)$ in $\bigwedge(\mathbb{R}^l)$.

In view of Theorem 7.1, it remains to prove that

$$\Phi(\varepsilon, S, \varphi, B) \subset \big\{x \in B \mid \|h(x)\mathbf{v}\| \ge \varepsilon \text{ for all } \mathbf{v} \in \Lambda \smallsetminus \{0\}\big\}. \tag{7.5}$$

Take an $(\varepsilon, S, \varphi)$-marked point $x \in B$, and let $\{0\} = \Gamma_0 \subsetneq \Gamma_1 \subsetneq \cdots \subsetneq \Gamma_m = \Lambda$ be all the elements of $\Sigma_x \cup \big\{\{0\}, \Lambda\big\}$. Take any $\mathbf{v} \in \Lambda \smallsetminus \{0\}$. Then there exists i, $1 \le i \le m$, such that $\mathbf{v} \in \Gamma_i \smallsetminus \Gamma_{i-1}$. Denote $(\Gamma_{i-1} + \mathbb{R}\mathbf{v}) \cap \Lambda$ by Δ. Clearly Δ is a primitive subgroup of Λ contained in Γ_i, therefore $\Delta \in \Sigma_x \cup S(\Sigma_x)$. Further, it is easy to show (see [KM1, Lemma 5.1]) that the norm of Δ is not greater than $l\|\Gamma_{i-1}\| \cdot \|\mathbf{v}\|$. Now one can use properties (M1) and (M2) to deduce that $|\varphi_\Delta(x)| = \|h(x)\Delta\| \ge \min(\varepsilon, \rho) = \varepsilon$, and then conclude that

$$\|h(x)\mathbf{v}\| \ge \|h(x)\Delta\|/l\|h(x)\Gamma_{i-1}\| \ge \varepsilon/l\rho \ge \varepsilon\,.$$

This shows (7.5) and completes the proof of the theorem. $\square$

Acknowledgements

A substantial part of this work was done during the authors' stays at the University of Bielefeld in 1998 and 1999. Thess stays were supported by SFB-343 and Humboldt Foundation. Thanks are also due to V. Beresnevich for useful comments.

References

[B1] A. Baker, *On a theorem of Sprindzhuk*, Proc. Roy. Soc. London **A 292** (1966), 92–104.

[B2] ______, *Transcendental number theory*, Cambridge Univ. Press, Cambridge, 1975.

[Be1] V. Beresnevich, *On approximation of real numbers by real algebraic numbers*, Acta Arith. (to appear).

[Be2] ______, *The best approximation of points of smooth curves in 3-dimensional euclidean space*, Dokladi NAN Belarusi (to appear).

[Be3] ______, *Application of the concept of a regular system of points in metric number theory*, Vestsī Akad. Navuk Belarusī Ser. Fīz.-Mat. Navuk (to appear).

[Be4] ______, *On proof of Khintchine type theorem for curves*, Vestsī Akad. Navuk Belarusī Ser. Fīz.-Mat. Navuk (to appear).

[Be5] ______, *A Khintchine-Groshev type theorem on manifolds: the case of convergence*, the University of York, Preprint (1999).

[Bern] V. Bernik, *A proof of Baker's conjecture in the metric theory of transcendental numbers*, Doklady Akad. Nauk SSSR **277** (1984), 1036–1039. (Russian)

[Bo] V. Borbat, *Joint zero approximation by the values of integral polynomials and their derivatives*, Vestsī Akad. Navuk Belarusī Ser. Fīz.-Mat. Navuk **123** (1995,), no. 1, 9–16.
[BB] V. Bernik and V. Borbat, *Polynomials with differences in values of coefficients and a conjecture of A. Baker*, Vestsī Akad. Navuk Belarusī Ser. Fīz.-Mat. Navuk (1997), no. 3, 5–8. (Russian)
[BBDD] V. Beresnevich, V. Bernik, H. Dickinson and M. Dodson, *A divergent Khintchine theorem for planar curves*, Preprint (1998).
[BDD] V. Bernik, H. Dickinson and M. Dodson, *A Khintchine-type version of Schmidt's theorem for planar curves*, Proc. Roy. Soc. London **A 454** (1998), 179–185.
[BS] A. Baker and W. Schmidt, *Diophantine approximation and Hausdorff dimension*, Proc. Lond. Math. Soc. **21** (1970), 1–11.
[D] H. Davenport, *A note on binary cubic forms*, Mathematika **8** (1961), 58–62.
[Do] M. M. Dodson, *Geometric and probabilistic ideas in metric Diophantine approximation*, Russian Math. Surveys **48** (1993), 73–102.
[DRV1] M. M. Dodson, B. P. Rynne and J. A. G. Vickers, *Metric Diophantine approximation and Hausdorff dimension on manifolds*, Math. Proc. Cambridge Philos. Soc. **105** (1989), 547–558.
[DRV2] ———, *Khintchine-type theorems on manifolds*, Acta Arith. **57** (1991), 115–130.
[DRV3] ———, *Simultaneous Diophantine approximation and asymptotic formulae on manifolds*, J. Number Theory **58** (1996), 298–316.
[G] A. V. Groshev, *Une théorème sur les systèmes des formes linéaires*, Dokl. Akad. Nauk SSSR **9** (1938), 151–152.
[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*, Inv. Math. (to appear).
[M] K. Mahler, *Über das Mass der Menge aller S-Zahlen*, Math. Ann. **106** (1932), 131–139.
[P] A. S. Pyartli, *Diophantine approximations on submanifolds of Euclidean space*, Functional Anal. Appl. **3** (1969), 303–306.
[S1] W. Schmidt, *Metrische Sätze über simultane Approximation abhänginger Grössen*, Monatsch. Math. **68** (1964), 154–166.
[S2] ———, *Diophantine approximation*, Springer-Verlag, Berlin and New York, 1980.
[Sp1] V. Sprindžuk, *More on Mahler's conjecture*, Doklady Akad. Nauk SSSR **155** (1964), 54–56 (Russian); English transl. in Soviet Math. Dokl **5** (1964), 361–363.
[Sp2] ———, *Mahler's problem in metric number theory*, Translations of Mathematical Monographs, vol. 25, Amer. Math. Soc., Providence, RI, 1969.
[Sp3] ———, *Metric theory of Diophantine approximations*, John Wiley & Sons, New York-Toronto-London, 1979.
[Sp4] ———, *Achievements and problems in Diophantine approximation theory*, Russian Math. Surveys **35** (1980), 1–80.

V. I. Bernik, Institute of Mathematics, Belarus Academy of Sciences, Minsk, 220072 Belarus
E-mail address: `bernik@im.bas-net.by`

D. Y. Kleinbock, Department of Mathematics, Rutgers University, New Brunswick, NJ 08903
E-mail address: `kleinboc@math.rutgers.edu`

G. A. Margulis, Department of Mathematics, Yale University, New Haven, CT 06520
E-mail address: `margulis@math.yale.edu`