MOSCOW MATHEMATICAL JOURNAL
Volume 2, Number 2, April–June 2002, Pages 203–225

METRIC DIOPHANTINE APPROXIMATION: THE KHINTCHINE–GROSHEV THEOREM FOR NONDEGENERATE MANIFOLDS

V. V. BERESNEVICH, V. I. BERNIK, D. Y. KLEINBOCK, AND G. A. MARGULIS

Dedicated to Yu. I. Manin on the occasion of his 65th birthday

ABSTRACT. The main objective of this paper is to prove a Khintchine type theorem on divergence of linear Diophantine approximation on nondegenerate manifolds, which completes earlier results for convergence.

2000 MATH. SUBJ. CLASS. Primary 11J83; Secondary 11K60.

KEY WORDS AND PHRASES. Diophantine approximation, Khintchine type theorems, metric theory of Diophantine approximation.

1. Background and the main result

1.1. Notation. The Vinogradov symbol $\ll$ (resp., $\gg$) means $\leqslant$ (resp., $\geqslant$) up to a positive constant multiplier; $a \asymp b$ is equivalent to $a \ll b \ll a$. The usual inner product in $\mathbb{R}^n$ of $\boldsymbol{a}$ and $\boldsymbol{b}$ will be denoted by $\boldsymbol{a} \cdot \boldsymbol{b}$; $\|\boldsymbol{a}\| = \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}}$ is the Euclidean norm of $\boldsymbol{a}$. Also, $\|\boldsymbol{a}\|_\infty = \max_{1\leqslant i\leqslant n} |a_i|$ and $\|\boldsymbol{a}\|_1 = \sum_{i=1}^n |a_i|$, where a_i are the coordinates of $\boldsymbol{a}$ in the standard basis of $\mathbb{R}^n$. The Lebesgue measure of $A \subset \mathbb{R}^d$ is denoted by $|A|_d$. We write $|A|$ instead of $|A|_d$ if there is no risk of confusion. Given a subset A of $\mathbb{R}^n$, we define $\operatorname{diam}(A) = \sup_{\boldsymbol{a},\boldsymbol{b}\in A} \|\boldsymbol{a} - \boldsymbol{b}\|$. Given two subsets A and B of $\mathbb{R}^n$, we define $\operatorname{dist}(A, B) = \inf_{\boldsymbol{a}\in A,\boldsymbol{b}\in B} \|\boldsymbol{a}-\boldsymbol{b}\|$; and $\operatorname{dist}(\boldsymbol{a}, A) = \operatorname{dist}(\{\boldsymbol{a}\}, A)$. Given an $\boldsymbol{x} \in \mathbb{R}^n$, there is a unique point $\boldsymbol{a} \in \mathbb{Z}^n$ such that $\boldsymbol{x} - \boldsymbol{a} \in (-1/2, 1/2]^n$. This difference will be denoted by $\langle \boldsymbol{x} \rangle$. Given a set $A \subset \mathbb{R}^d$ and a number $r > 0$, let $\mathcal{B}(A, r) = \{\boldsymbol{x} \in \mathbb{R}^d \colon \operatorname{dist}(\boldsymbol{x}, A) < r\}$. In particular, $\mathcal{B}(\boldsymbol{a}, r) = \mathcal{B}(\{\boldsymbol{a}\}, r)$ is the open ball in $\mathbb{R}^d$ of radius r centered at $\boldsymbol{a}$. Given a ball $\mathcal{B} = \mathcal{B}(\boldsymbol{x}, r)$ and a positive number λ, $\lambda\mathcal{B}$ will denote the ball $\mathcal{B}(\boldsymbol{x}, \lambda r)$. Given a map $\boldsymbol{f} \colon U \to \mathbb{R}^n$, where U is an open subset of $\mathbb{R}^d$, we will denote by $\partial_i \boldsymbol{f} \colon U \to \mathbb{R}^n$, $i = \overline{1, d}$, its partial derivative with respect to x_i. We also define a map $\nabla \boldsymbol{f} \colon U \to M_{n\times d}(\mathbb{R})$, where $M_{n\times d}(\mathbb{R})$ is the space of $n \times d$ matrices over $\mathbb{R}$, by setting $\nabla \boldsymbol{f}(\boldsymbol{x}) = (\partial_j f_i(\boldsymbol{x}))_{1\leqslant i\leqslant n, 1\leqslant j\leqslant d}$. We will also need higher order differentiation: for a *multiindex* $\beta = (i_1, \ldots, i_d)$,

Received January 30, 2002; in revised form April 11, 2002.

The work of the first and second named authors was supported in part by grant F00-249 of the Belarussian Fund of Fundamental Research. The work of the third named author was supported in part by NSF Grant DMS-0072565, and that of the fourth named author by NSF Grant DMS-9800607.

$i_j \in \mathbb{Z}_{\geqslant 0}$, where $\mathbb{Z}_{\geqslant 0} = \{x \in \mathbb{Z} \colon x \geqslant 0\}$, we let $\partial_\beta = \partial_1^{i_1} \circ \cdots \circ \partial_d^{i_d}$. Throughout the paper, $\psi \colon \mathbb{R}_+ \to \mathbb{R}_+$ is a non-increasing function unless a different condition is assumed.

1.2. Metric Diophantine approximation in $\mathbb{R}^n$. Metric Diophantine approximation began with the work of E. Borel and A. J. Khintchine, who considered approximation to real numbers by rational numbers. In 1924 for $n = 1$ Khintchine [Khi24] and in 1938 for $n > 1$ A. V. Groshev [Gro38] established a criterion for the solubility of the inequality

$$|\langle \boldsymbol{a} \cdot \boldsymbol{y} \rangle| < \psi(\|\boldsymbol{a}\|_\infty^n) \tag{1.1}$$

in $\boldsymbol{a} \in \mathbb{Z}^n$ for generic $\boldsymbol{y} \in \mathbb{R}^n$. At this point we need the following

Definition 1.1. The point $\boldsymbol{y} \in \mathbb{R}^n$ is called *ψ-approximable* if (1.1) has infinitely many solutions $\boldsymbol{a} \in \mathbb{Z}^n$. The point $\boldsymbol{y} \in \mathbb{R}^n$ is called *very well approximable* (VWA) if it is ψ_ε-approximable for some positive ε, where $\psi_\varepsilon(h) = h^{-(1+\varepsilon)}$.

In view of this definition, the Khintchine–Groshev theorem [Khi24], [Gro38] asserts that if the sum

$$\sum_{h=1}^{\infty} \psi(h) \tag{1.2}$$

diverges (converges), then almost all (almost no) points $\boldsymbol{y} \in \mathbb{R}^n$ are ψ-approximable.

Remark 1.2. Originally the inequality $|\langle \boldsymbol{a} \cdot \boldsymbol{x} \rangle| < \psi(\|\boldsymbol{a}\|_\infty)$ was considered instead of (1.1). In this setting $\sum_{q=1}^{\infty} q^{n-1}\psi(q)$ should be used instead of (1.2). Khintchine assumed that $h\psi(h)$ was non-increasing, and Groshev's requirement was the monotonicity of $h^{n-1}\psi(h)$. Later W. M. Schmidt succeeded to avoid the monotonicity restriction when $n > 1$ (see Section 6).

Remark 1.3. The Khintchine–Groshev theorem implies that almost all $\boldsymbol{y} \in \mathbb{R}^n$ are not VWA. The convergence case of the theorem can be easily derived from the Borel–Cantelli lemma. The main difficulty is contained in the divergence case.

1.3. The concept of Diophantine approximation on manifolds. This concept emerges if one restricts the point $\boldsymbol{y}$ to lie on a submanifold $\mathcal{M}$ of $\mathbb{R}^n$. Since the manifold $\mathcal{M}$ of dimension $< n$ itself has zero measure, the Khintchine–Groshev theorem does not even guarantee the existence of a single ψ-approximable point on $\mathcal{M}$. To make the theory rich in content, one tries to establish if a given property holds for almost all points of this manifold with respect to the Lebesgue measure induced on the manifold. We will use the following terminology (more details can be found in [BD99] and Section 6).

Definition 1.4. Let $\mathcal{M}$ be a submanifold of $\mathbb{R}^n$. One says that $\mathcal{M}$ is *extremal* if almost all points of $\mathcal{M}$ are not VWA. One says that $\mathcal{M}$ is of *Groshev type for divergence* (*for convergence*) if almost all (almost no) points of $\mathcal{M}$ are ψ-approximable whenever the sum (1.2) diverges (converges).

1.4. Diophantine approximation on the Veronese curves. In 1932 K. Mahler [Mah32] made a conjecture which in the terminology of this paper claimed that for any $n \in \mathbb{N}$ the *Veronese curve*

$$\mathcal{V}_n = \{(x, x^2, \ldots, x^n) : x \in \mathbb{R}\} \tag{1.3}$$

was extremal. It arose in transcendental number theory in connection with a classification of real numbers suggested by Mahler himself. A great deal of work to prove Mahler's conjecture had been undertaken by J. Kubilius, B. Volkmann, W. LeVeque, F. Kash, and W. M. Schmidt. In particular, the problem was solved for $n = 2$ by Kubilius [Kub49] and for $n = 3$ by Volkmann [Vol61]. The complete solution was given by V. G. Sprindẑuk [Spr69] in 1964.

In 1966 A. Baker [Bak66] improved Sprindẑuk's result by replacing the "powering" error function with a general monotonic function ψ by showing that if

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

then almost all points on the curve (1.3) are not ψ-approximable. In the same paper Baker conjectured that (1.4) could be replaced with the convergence of (1.2), i. e., he conjectured that $\mathcal{V}_n$ is of Groshev type for convergence. This conjecture was proved by V.I. Bernik [Ber89] in 1989.

The divergence case was considered by V. V. Beresnevich [Ber99a] in 1999 who proved that the Veronese curves (1.3) are of Groshev type for divergence. The proof is based on a new method involving regular systems, introduced by Baker and Schmidt [BS70] and used for computing the Hausdorff dimension of sets of well approximable points.

1.5. Diophantine approximation on differentiable manifolds. In the sixties of the last century the investigations related to the problem of Mahler eventually led to the development of a new branch of metric number theory, usually referred to as "Diophantine approximation of dependent quantities" or "Diophantine approximation on manifolds". The first result involving manifolds defined by functions satisfying some mild and natural properties was obtained by Schmidt [Sch64b], who proved that any $C^{(3)}$ planar curve with curvature non-vanishing almost everywhere is extremal. Schmidt's theorem was subsequently improved by R. Baker [Bak78], who showed that almost all points on Schmidt's curves are not ψ-approximable whenever $(1.4)_{n=2}$ is satisfied. It has been recently shown that Schmidt's curves are of Groshev type for convergence [BDD98] and for divergence [BBDD99].

Until the mid-nineties most of the results in metric Diophantine approximation dealt with manifolds of a special structure or of high enough dimension. M. M. Dodson, B. P. Rynne and J. A. G. Vickers [DRV90b], [DRV91], [DRV96] investigated a class of manifolds satisfying a geometric condition which for surfaces in $\mathbb{R}^3$ assumed two-convexity (e. g. a cylinder does not satisfy that condition). Schmidt [Sch64b] has investigated certain straight lines in $\mathbb{R}^n$ for extremality, and recently such lines have been shown to be of Groshev type [BBDD00].

A new method, based on combinatorics of the space of lattices, was developed in [KM98] by D. Y. Kleinbock and G. A. Margulis[1], who proved the extremality of the so-called nondegenerate manifolds (they also proved these manifolds to be strongly extremal, see Section 6).

Definition 1.5. Let $\boldsymbol{f}\colon U \to \mathbb{R}^n$ be a map defined on an open set $U \subset \mathbb{R}^d$. We say that $\boldsymbol{f}$ is *l-nondegenerate at $\boldsymbol{x}_0 \in U$* if $\boldsymbol{f}$ is l times continuously differentiable on some sufficiently small ball centered at $\boldsymbol{x}_0$ and partial derivatives of $\boldsymbol{f}$ at $\boldsymbol{x}_0$ of orders up to l span $\mathbb{R}^n$. We say that $\boldsymbol{f}$ is *nondegenerate at $\boldsymbol{x}_0$* if it is l-nondegenerate at $\boldsymbol{x}_0$ for some $l \in \mathbb{N}$. We say that $\boldsymbol{f}$ is *nondegenerate* if it is nondegenerate almost everywhere on U.

The nondegeneracy of a manifold is naturally defined via the nondegeneracy of its appropriate parameterization. Geometrically the l-nondegeneracy of a manifold $\mathcal{M} \subset \mathbb{R}^n$ at a point $\boldsymbol{y}_0 \in \mathcal{M}$ means that for any hyperplane Π in $\mathbb{R}^n$, $\limsup_{\boldsymbol{y}\to\boldsymbol{y}_0, \boldsymbol{y}\in\mathcal{M}} \operatorname{dist}(\boldsymbol{y}, \Pi)\cdot\|\boldsymbol{y}-\boldsymbol{y}_0\|^{-l} > 0$; that is, the manifold cannot be approximated by a hyperplane "too well" (see [Ber02], [Ber99b]).

Recently Beresnevich [Ber02] (a short version is published in [Ber00a], [Ber00b]), and independently Bernik, Kleinbock and Margulis [BKM01] using different techniques, have proved that any nondegenerate manifold is of Groshev type for convergence (also there is a multiplicative analogue and a more general version of the result in [BKM01], see Section 6).

Nondegenerate curves have been proved to be of Groshev type for divergence [Ber00d] (also [Ber00a], [Ber00c] contain auxiliary parts of the proof). Moreover, by Pyartli's method [Pya69] one can extend this result to analytic nondegenerate manifolds. The goal of the present paper is to show that any nondegenerate manifold is of Groshev type for divergence. The proof makes use of a new technique, which involves a multidimensional analogue of regular systems and extends the ideas of [Ber99a].

1.6. The main result and the structure of the paper.

Theorem 1.6. *Let U be an open subset of $\mathbb{R}^d$ and let $\boldsymbol{f}\colon U \to \mathbb{R}^n$ be a nondegenerate map. Also let $\psi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a non-increasing function such that the sum (1.2) diverges. Then for almost all $\boldsymbol{x} \in U$ the point $\boldsymbol{f}(\boldsymbol{x})$ is ψ-approximable, i. e., for almost all $\boldsymbol{x} \in U$ there are infinitely many solutions $\boldsymbol{a} \in \mathbb{Z}^n$ to the inequality*

$$|\langle \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{a}\rangle| < \psi(\|\boldsymbol{a}\|_\infty^n). \tag{1.5}$$

The proof of Theorem 1.6 is based on a method of regular systems first suggested in [Ber99a] for dimension one. In particular, we generalize it for any dimension. In Section 3 we construct a regular system of resonant sets corresponding to a given nondegenerate map. In Section 4 we prove a general theorem on approximation by resonant sets. And finally, Section 5 completes the proof of Theorem 1.6.

[1]See also [KM99] and [Kle01] for more on interactions between dynamics on the space of lattices and metric Diophantine approximation.

2. Effective upper bounds

The result of this section will be applied to construct a regular system of resonant sets. We show the following

Theorem 2.1. *Let $\boldsymbol{f}\colon U\to\mathbb{R}^n$ be nondegenerate at $\boldsymbol{x}_0\in U$. Then there exists a sufficiently small ball $\mathcal{B}_0\subset U$ centered at $\boldsymbol{x}_0$ and a constant $C_0>0$ such that for any ball $\mathcal{B}\subset\mathcal{B}_0$ and any $\varepsilon>0$, for all sufficiently large Q, one has*

$$|\mathcal{L}_{\boldsymbol{f}}(\mathcal{B};\,\varepsilon;\,Q)|\leqslant C_0\varepsilon|\mathcal{B}|,\tag{2.1}$$

where

$$\mathcal{L}_{\boldsymbol{f}}(\mathcal{B};\,\varepsilon;\,Q)=\bigcup_{\boldsymbol{a}\in\mathbb{Z}^n\colon\, 0<\|\boldsymbol{a}\|_\infty\leqslant Q}\left\{\boldsymbol{x}\in\mathcal{B}\colon |\langle\boldsymbol{f}(\boldsymbol{x})\cdot\boldsymbol{a}\rangle|<\varepsilon Q^{-n}\right\}.\tag{2.2}$$

To prove Theorem 2.1 we will consider two special cases: when the norm of the gradient $\boldsymbol{a}\nabla\boldsymbol{f}(x)$ is big, or, respectively, not very big. Theorem 2.2 below is essentially due to Bernik and for $d=1$ has appeared earlier [Ber00d]. Its proof relies on the ideas of the method of essential and inessential domains developed by Sprindžuk, when he solved the problem of Mahler. Theorem 2.3 below is due to Kleinbock and Margulis [BKM01] and is proved by means of the method involving lattices, which was first developed in [KM98]. The dichotomy of big/small derivatives has been extensively used; in particular, it is used in [Ber00a], [Ber02], [BKM01] to prove the convergence case.

Theorem 2.2 (Theorem 1.3 in [BKM01])**.** *Let $\mathcal{B}_0\subset\mathbb{R}^d$ be a ball, and let $\boldsymbol{f}\in C^{(2)}(3\mathcal{B}_0)$. Fix $\delta>0$ and define*

$$L_1=\max_{\|\beta\|_1=2}\ \max_{\boldsymbol{x}\in 2\mathcal{B}_0}\|\partial_\beta\boldsymbol{f}(\boldsymbol{x})\|_\infty.\tag{2.3}$$

Then for every ball $\mathcal{B}\subset\mathcal{B}_0$ and any $\boldsymbol{a}\in\mathbb{Z}^n$ such that

$$\|\boldsymbol{a}\|_\infty\geqslant\frac{1}{nL_1(\operatorname{diam}\mathcal{B})^2},\tag{2.4}$$

the set

$$\mathcal{L}^{(1)}_{\boldsymbol{f}}(\mathcal{B};\,\delta;\,\boldsymbol{a})=\left\{\boldsymbol{x}\in\mathcal{B}\colon\ |\langle\boldsymbol{f}(\boldsymbol{x})\cdot\boldsymbol{a}\rangle|<\delta,\ \|\boldsymbol{a}\nabla\boldsymbol{f}(\boldsymbol{x})\|_\infty\geqslant\sqrt{ndL_1\|\boldsymbol{a}\|_\infty}\right\}\tag{2.5}$$

has measure at most $C_1\delta|\mathcal{B}|$, where $C_1>0$ is a constant depending on d only.

Theorem 2.3 (Theorem 1.4 in [BKM01])**.** *Let $U\subset\mathbb{R}^d$ be an open set, $\boldsymbol{x}_0\in U$, and let $\boldsymbol{f}\colon U\to\mathbb{R}^n$ be a map l-nondegenerate at $\boldsymbol{x}_0$. Then there exists a ball $\mathcal{B}_0\subset U$ centered at $\boldsymbol{x}_0$ such that $3\mathcal{B}_0\subset U$ with the following property: there exists a constant $C_2>1$ such that for any ball $\mathcal{B}\subset\mathcal{B}_0$, any ε with $0<\varepsilon<1$ and any $Q\geqslant1$ the set*

$$\mathcal{L}^{(2)}_{\boldsymbol{f}}(\mathcal{B};\,\varepsilon;\,Q)=\bigcup_{\boldsymbol{a}\in\mathbb{Z}^n\colon\, 0<\|\boldsymbol{a}\|_\infty\leqslant Q}\Big\{\boldsymbol{x}\in\mathcal{B}\colon\ |\langle\boldsymbol{f}(\boldsymbol{x})\cdot\boldsymbol{a}\rangle|<\varepsilon Q^{-n},\\ \|\boldsymbol{a}\nabla\boldsymbol{f}(\boldsymbol{x})\|_\infty<\sqrt{ndL_1Q}\Big\}\tag{2.6}$$

satisfies

$$|\mathcal{L}^{(2)}_{\boldsymbol{f}}(\mathcal{B};\,\varepsilon;\,Q)|\leqslant C_2(\varepsilon Q^{-1/2})^{\frac{1}{d(n+1)(2l-1)}}\cdot|\mathcal{B}|,\tag{2.7}$$

where L_1 is defined in (2.3).

Proof of Theorem 2.1. Fix a ball $\mathcal{B}_0$ as in the statement of Theorem 2.3 and fix any ball $\mathcal{B} \subset \mathcal{B}_0$. It is easy to see that the set $\mathcal{L}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon;\, Q)$ is expressed as the following union of three subsets

$$\mathcal{L}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon;\, Q) = \left(\bigcup_{\boldsymbol{a}\in\mathbb{Z}^n:\ Q_1\leqslant\|\boldsymbol{a}\|_\infty\leqslant Q} \mathcal{L}^{(1)}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon Q^{-n};\, \boldsymbol{a}) \right) \cup \\ \cup\, \mathcal{L}^{(2)}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon;\, Q) \cup \left(\bigcup_{\boldsymbol{a}\in\mathbb{Z}^n:\ \|\boldsymbol{a}\|_\infty\leqslant Q_1} \mathcal{L}^{(1)}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon Q^{-n};\, \boldsymbol{a}) \right), \tag{2.8}$$

where $Q_1 = [1/(nL_1(\operatorname{diam}\mathcal{B})^2)] + 1$. The measure of the first subset is estimated by Theorem 2.2:

$$\left| \bigcup_{\boldsymbol{a}\in\mathbb{Z}^n:\ Q_1\leqslant\|\boldsymbol{a}\|_\infty\leqslant Q} \mathcal{L}^{(1)}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon Q^{-n};\, \boldsymbol{a}) \right| \leqslant C_1 \varepsilon Q^{-n} |\mathcal{B}| (2Q+1)^n. \tag{2.9}$$

Next, for every $\boldsymbol{a} \in \mathbb{Z}^n$ such that $0 < \|\boldsymbol{a}\|_\infty < Q_1$ we obviously have

$$\mathcal{L}^{(1)}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon Q^{-n};\, \boldsymbol{a}) \subset \mathcal{L}^{(1)}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon Q^{-n} Q_1;\, \boldsymbol{a}_1), \tag{2.10}$$

where $\boldsymbol{a}_1 = Q_1 \boldsymbol{a}$. It is clear that $\|\boldsymbol{a}_1\|_\infty \geqslant Q_1$. Therefore, we can apply Theorem 2.2 to the set in the right-hand side of (2.10). Thus,

$$|\mathcal{L}^{(1)}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon Q^{-n};\, \boldsymbol{a})| \leqslant |\mathcal{L}^{(1)}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon Q^{-n} Q_1;\, \boldsymbol{a}_1)| \leqslant C_1 \varepsilon Q^{-n} Q_1 |\mathcal{B}|. \tag{2.11}$$

Since the number of points $\boldsymbol{a} \in \mathbb{Z}^n$ with $0 < \|\boldsymbol{a}\|_\infty \leqslant Q_1$ is less than $(2Q_1+1)^n$, we get

$$\left| \bigcup_{\boldsymbol{a}\in\mathbb{Z}^n:\ \|\boldsymbol{a}\|_\infty\leqslant Q_1} \mathcal{L}^{(1)}_{\boldsymbol{f}}(\mathcal{B};\, \varepsilon Q^{-n};\, \boldsymbol{a}) \right| \leqslant (2Q_1+1)^n C_1 \varepsilon Q^{-n} Q_1 |\mathcal{B}|. \tag{2.12}$$

On combining (2.7), (2.9), (2.12) and (2.8) and letting $C_0 > 2^n C_1$, we obtain (2.1) for all sufficiently large Q. This completes the proof of Theorem 2.1. □

We will also use the following

Lemma 2.4 (Lemma 6 in [Ber02])**.** *Let $\alpha, \beta \in \mathbb{R}_+$, $d \in \mathbb{N}$, $\mathcal{B}$ be a ball in $\mathbb{R}^d$, $f\colon \mathcal{B} \to \mathbb{R}$ be a function such that $f \in C^{(k)}$ and for some j with $1 \leqslant j \leqslant d$ one has*

$$\inf_{\boldsymbol{x}\in\mathcal{B}} |\partial_j^k f(\boldsymbol{x})| \geqslant \beta. \tag{2.13}$$

Then

$$\big|\{\boldsymbol{x} \in \mathcal{B}\colon |f(\boldsymbol{x})| \leqslant \alpha\}\big| \leqslant 3^{(k+1)/2}(k(k+1)/2+1)(\operatorname{diam}\mathcal{B})^{d-1}\Big(\frac{\alpha}{\beta}\Big)^{1/k}.$$

3. Regular systems of resonant sets

Definition 3.1. Let U be an open subset of $\mathbb{R}^d$, $\mathcal{R}$ be a family of subsets of $\mathbb{R}^d$, $N\colon \mathcal{R} \to \mathbb{R}_+$ be a function and let s be a number satisfying $0 \leqslant s < d$. The triple $(\mathcal{R},\, N,\, s)$ is called a *regular system* in U if there exist constants $K_1,\, K_2,\, K_3 > 0$ and a function $\lambda\colon \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{x\to+\infty}\lambda(x) = +\infty$ such that for any ball $\mathcal{B} \subset U$ and for any $T > T_0$, where $T_0 = T_0(\mathcal{R},\, N,\, s,\, \mathcal{B})$ is a sufficiently large number, there exist sets

$$R_1,\, \dots,\, R_t \in \mathcal{R} \quad \text{with} \quad \lambda(T) \leqslant N(R_i) \leqslant T \quad \text{for} \quad i = \overline{1,\, t} \tag{3.1}$$

and disjoint balls

$$\mathcal{B}_1,\, \dots,\, \mathcal{B}_t \quad \text{with} \quad 2\mathcal{B}_i \subset \mathcal{B} \quad \text{for} \quad i = \overline{1,\, t} \tag{3.2}$$

such that

$$\operatorname{diam}(\mathcal{B}_i) = T^{-1} \quad \text{for} \quad i = \overline{1,\, t}, \tag{3.3}$$

$$t \geqslant K_1 |\mathcal{B}| T^d \tag{3.4}$$

and such that for any $\gamma \in \mathbb{R}$ with $0 < \gamma < T^{-1}$ one has

$$K_2 \gamma^{d-s} T^{-s} \leqslant |\mathcal{B}(R_i,\, \gamma) \cap \mathcal{B}_i|, \tag{3.5}$$

$$|\mathcal{B}(R_i,\, \gamma) \cap 2\mathcal{B}_i| \leqslant K_3 \gamma^{d-s} T^{-s}. \tag{3.6}$$

The elements of $\mathcal{R}$ will be called *resonant sets*.

This definition generalizes the concept of *regular system of points* of Baker and Schmidt. In fact, it is equivalent to the Baker–Schmidt definition when $U = \mathbb{R}$, $\mathcal{R}$ consists of points in the real line, and $s = 0$ [BS70]. In this situation conditions (3.5) and (3.6) hold automatically. Also this definition covers the multidimensional concept of a regular system of points [Ber00c] when $s = 0$. Definition 3.1 is closely related to *ubiquitous systems* [DRV90a].

The goal of this section is to establish the following

Theorem 3.2. *Let $\boldsymbol{f} = (f_1,\, \dots,\, f_n)\colon U \to \mathbb{R}^n$ be a nondegenerate map, where U is an open subset of $\mathbb{R}^d$. Given an $\boldsymbol{a} \in \mathbb{Z}^n$, $\boldsymbol{a} \neq 0$ and an $a_0 \in \mathbb{Z}$, let*

$$R_{\boldsymbol{a},a_0} = \{\boldsymbol{x} \in U\colon \boldsymbol{a} \cdot \boldsymbol{f}(\boldsymbol{x}) + a_0 = 0\}.$$

Define the following set,

$$\mathcal{R}_{\boldsymbol{f}} = \{R_{\boldsymbol{a},a_0}\colon \boldsymbol{a} \in \mathbb{Z}^n,\, \boldsymbol{a} \neq 0,\, a_0 \in \mathbb{Z}\},$$

and the following function,

$$N(R_{\boldsymbol{a},a_0}) = (\|\boldsymbol{a}\|_\infty)^{n+1}.$$

Then for almost every point $\boldsymbol{x}_0 \in U$ there is a ball $\mathcal{B}_0 \subset U$ centered at $\boldsymbol{x}_0$ such that $(\mathcal{R},\, N,\, d-1)$ is a regular system in $\mathcal{B}_0$.

Proof. There is no loss of generality in assuming that $f_1(\boldsymbol{x}) = x_1$. In fact, using the nondegeneracy of $\boldsymbol{f}$, it is possible to show that $\boldsymbol{f}'(\boldsymbol{x}) \neq \boldsymbol{0}$ almost everywhere (see [Ber02, Section 5]). Thus we can take a sufficiently small neighborhood of a point $\boldsymbol{x}_0$ with $\boldsymbol{f}'(\boldsymbol{x}_0) \neq \boldsymbol{0}$ instead of the original domain U, and then make $f_1(\boldsymbol{x})$ equal x_1 by an appropriate change of variables. Also, as $\boldsymbol{f}$ is nondegenerate, we can take U

to be a sufficiently small neighborhood of a point $\boldsymbol{x}_0$ such that $\boldsymbol{f}$ is nondegenerate at this point. Moreover, we can take $\mathcal{B}_0$ satisfying Theorem 2.1. Thus, in view of that theorem, for any ball $\mathcal{B} \subset \mathcal{B}_0$ the set

$$\mathcal{G}(\mathcal{B};\, (4C_0)^{-1};\, Q) = \tfrac{3}{4}\mathcal{B} \setminus \mathcal{L}_{\boldsymbol{f}}(\tfrac{3}{4}\mathcal{B};\, (4C_0)^{-1};\, Q)$$

will satisfy the estimate

$$|\mathcal{G}(\mathcal{B};\, (4C_0)^{-1};\, Q)| \geqslant \frac{1}{2}|\mathcal{B}| \tag{3.7}$$

for all sufficiently large Q.

Note also that there is no loss of generality in assuming that

$$\max_{1\leqslant j\leqslant n} \sup_{\boldsymbol{x}\in\mathcal{B}_0} \|\nabla f_j(\boldsymbol{x})\|_\infty \leqslant L_2, \tag{3.8}$$

for some constant $L_2 > 0$.

The proof of Theorem 3.2 will be completed with the help of

Proposition 3.3. *There is a sufficiently large number Q_0 such that for any $Q \geqslant Q_0$ for any $\boldsymbol{x} \in \mathcal{G}(\mathcal{B};\, (4C_0)^{-1};\, Q)$ there is an integer point $\boldsymbol{a} \in \mathbb{Z}^n$, $\boldsymbol{a} \neq \boldsymbol{0}$ and an integer a_0 with*

$$Q^{n+1} = T/C_3 \leqslant N(R_{\boldsymbol{a},a_0}) \leqslant T = C_3 Q^{n+1}, \tag{3.9}$$

where $C_3 = \big(4C_0(nL_2)^{n-1}\big)^{n+1}$, and a point $\boldsymbol{z} \in R_{\boldsymbol{a},a_0}$ such that

$$\|\boldsymbol{x} - \boldsymbol{z}\| < C_4 T^{-1}, \tag{3.10}$$

where $C_4 = C_3 n/(2C_0)$, and such that for any γ with $0 < \gamma < T^{-1}$ we have

$$K_2\gamma T^{-(d-1)} \leqslant |\mathcal{B}(R_{\boldsymbol{a},a_0},\, \gamma) \cap \mathcal{B}(\boldsymbol{z},\, T^{-1}/2)|, \tag{3.11}$$

$$|\mathcal{B}(R_{\boldsymbol{a},a_0},\, \gamma) \cap \mathcal{B}(\boldsymbol{z},\, T^{-1})| \leqslant K_3\gamma T^{-(d-1)}, \tag{3.12}$$

where K_2, $K_3 > 0$ are some constants which depend on neither $\mathcal{B}$ nor T.

Proof of Proposition 3.3. Let $\boldsymbol{x} \in \mathcal{G}(\mathcal{B};\, (4C_0)^{-1};\, Q)$. By Minkowski's linear forms theorem, there are integers $\boldsymbol{a} \in \mathbb{Z}^n$, $\boldsymbol{a} \neq \boldsymbol{0}$ and $a_0 \in \mathbb{Z}$ such that

$$\begin{cases} |\boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{a} + a_0| \leqslant (4C_0)^{-1} Q^{-n}, \\ |a_1| \leqslant 4C_0(nL_2)^{n-1} Q, \\ |a_i| \leqslant Q/(nL_2), \qquad i = \overline{2,\, n}. \end{cases} \tag{3.13}$$

Define the function $F(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{a} + a_0$. It follows from (3.13) that

$$\|\boldsymbol{a}\|_\infty \leqslant 4C_0(nL_2)^{n-1} Q = T^{1/(n+1)}. \tag{3.14}$$

Since $\boldsymbol{x} \in \mathcal{G}(\mathcal{B};\, (4C_0)^{-1};\, Q)$, $\|\boldsymbol{a}\|_\infty$ must be $> Q$, which, combined with (3.14), gives (3.9).

As $|a_j| < Q$ for $j = \overline{2,\, n}$, we have $|a_1| > Q$. Now, using (3.8) and the condition $f_1(\boldsymbol{x}) = x_1$, we get

$$|\partial_1 F(\boldsymbol{x})| = |a_1| \cdot |\partial_1 f_1(\boldsymbol{x})| - \sum_{i=2}^{n} |a_i| \cdot |\partial_1 f_j(\boldsymbol{x})| > Q - \sum_{i=2}^{n} Q/(nL_2) \cdot L_2 = \frac{Q}{n}. \tag{3.15}$$

Since $\partial_1 \boldsymbol{f}$ is uniformly continuous on $\mathcal{B}_0$, there is a sufficiently small number $r_1 > 0$ such that for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in U$ with $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| < r_1$ we have

$$\|\partial_1 \boldsymbol{f}(\boldsymbol{x}_1) - \partial_1 \boldsymbol{f}(\boldsymbol{x}_2)\|_\infty < \frac{1}{8n^2 C_0 (nL_2)^{n-1}}.$$

It follows that

$$|\partial_1 F(\boldsymbol{x}_1) - \partial_1 F(\boldsymbol{x}_2)| \leqslant n\|\boldsymbol{a}\|_\infty \|\partial_1 \boldsymbol{f}(\boldsymbol{x}_1) - \partial_1 \boldsymbol{f}(\boldsymbol{x}_2)\|_\infty \leqslant \frac{1}{8nC_0(nL_2)^{n-1}} \|\boldsymbol{a}\|_\infty.$$

Applying (3.14) now gives $|\partial_1 F(\boldsymbol{x}_1) - \partial_1 F(\boldsymbol{x}_2)| \leqslant Q/(2n)$ for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in U$ with $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| < r_1$. This and (3.15) imply

$$|\partial_1 F(\boldsymbol{y})| \geqslant |\partial_1 F(\boldsymbol{x})| - |\partial_1 F(\boldsymbol{x}) - \partial_1 F(\boldsymbol{y})| > Q/(2n)$$

for all $\boldsymbol{y} \in U$ with $\|\boldsymbol{x} - \boldsymbol{y}\| < r_1$.

As $\boldsymbol{x} \in \frac{3}{4}\mathcal{B}$, we have $\mathcal{B}(\boldsymbol{x}, \operatorname{diam}\mathcal{B}/8) \subset \mathcal{B}$. Define $r_0 = \min(r_1, \operatorname{diam}\mathcal{B}/8)$. Thus,

$$|\partial_1 F(\boldsymbol{y})| > Q/(2n) \quad \text{for all } \boldsymbol{y} \in \mathcal{B}(\boldsymbol{x}, r_0). \tag{3.16}$$

Let $|\theta| < r_0$. Then $\boldsymbol{x}_\theta = (x_1+\theta, x_2, \ldots, x_d) \in \mathcal{B}(\boldsymbol{x}, r_0)$, where $\boldsymbol{x} = (x_1, \ldots, x_d)$. By the Mean Value Theorem, we have $F(\boldsymbol{x}_\theta) = F(\boldsymbol{x}) + \partial_1 F(\tilde{\boldsymbol{x}}_\theta)\theta$, where $\tilde{\boldsymbol{x}}_\theta \in \mathcal{B}(\boldsymbol{x}, r_0)$. This can equivalently be written as

$$\frac{F(\boldsymbol{x}_\theta)}{\partial_1 F(\tilde{\boldsymbol{x}}_\theta)} = \frac{F(\boldsymbol{x})}{\partial_1 F(\tilde{\boldsymbol{x}}_\theta)} + \theta. \tag{3.17}$$

Assume that $Q > (n/(2r_0C_0))^{1/(n+1)}$. This condition implies that for any

$$\theta \in \left[-n/(2C_0) \cdot Q^{-n-1},\, n/(2C_0) \cdot Q^{-n-1}\right]$$

we have $|\theta| < r_0$, and therefore $\boldsymbol{x}_\theta, \tilde{\boldsymbol{x}}_\theta \in \mathcal{B}(\boldsymbol{x}, r_0)$. Now using (3.13) and (3.16) we get

$$|F(\boldsymbol{x})/\partial_1 F(\tilde{\boldsymbol{x}}_\theta)| < n/(2C_0) \cdot Q^{-n-1}.$$

It follows from this and (3.17) that $F(\boldsymbol{x}_\theta)/\partial_1 F(\tilde{\boldsymbol{x}}_\theta)$ is positive at $\theta = n/(2C_0) \times Q^{-n-1}$ and negative at $\theta = -n/(2C_0) \cdot Q^{-n-1}$. By continuity, there is a number θ_0 with

$$|\theta_0| < n/(2C_0) \cdot Q^{-n-1}$$

such that $F(\boldsymbol{x}_{\theta_0})/\partial_1 F(\tilde{\boldsymbol{x}}_{\theta_0}) = 0$, or, equivalently, $F(\boldsymbol{x}_{\theta_0}) = 0$. Define $\boldsymbol{z}$ to be $\boldsymbol{x}_{\theta_0}$. By construction, $\boldsymbol{z} \in R_{\boldsymbol{a},a_0}$, and $\|\boldsymbol{x} - \boldsymbol{z}\| = |\theta_0| < n/(2C_0) \cdot Q^{-n-1}$. This proves (3.10).

Now we are going to show (3.12). Assume that $T > (C_4+1)/r_0$. This condition and (3.10) imply that

$$\mathcal{B}(\boldsymbol{z}, T^{-1}) \subset \mathcal{B}(\boldsymbol{x}, r_0).$$

Let $0 < \gamma < T^{-1}$. By definition, for any point $\boldsymbol{y} \in \mathcal{B}(R_{\boldsymbol{a},a_0}, \gamma)$ there is a point $\boldsymbol{y}_0 \in R_{\boldsymbol{a},a_0}$ such that $\|\boldsymbol{y} - \boldsymbol{y}_0\| < \gamma$.

Assume that $\boldsymbol{y} \neq \boldsymbol{y}_0$. Then, by the Mean Value Theorem, we have

$$F(\boldsymbol{y}) = F(\boldsymbol{y}_0) + \nabla F(\boldsymbol{y}_1) \cdot (\boldsymbol{y} - \boldsymbol{y}_0) = \nabla F(\boldsymbol{y}_1) \cdot (\boldsymbol{y} - \boldsymbol{y}_0) = (\boldsymbol{a}\nabla \boldsymbol{f}(\boldsymbol{y}_1)) \cdot (\boldsymbol{y} - \boldsymbol{y}_0),$$

where $\boldsymbol{y}_1$ is a point between $\boldsymbol{y}_0$ and $\boldsymbol{y}$. Using (3.14), we find that

$$|F(\boldsymbol{y})| \leqslant d\|\boldsymbol{a}\nabla f(\boldsymbol{y}_1)\|_\infty \cdot \|\boldsymbol{y} - \boldsymbol{y}_0\|_\infty \leqslant dn\|\boldsymbol{a}\|_\infty L_2 \gamma \leqslant C_5 Q\gamma,$$

where $C_5 = dn\,4C_0(nL_2)^{n-1}L_2$. It follows that

$$\mathcal{B}(R_{\boldsymbol{a},a_0},\,\gamma)\cap\mathcal{B}(\boldsymbol{z},\,T^{-1})\subset\{\boldsymbol{y}\in\mathcal{B}(\boldsymbol{z},\,T^{-1})\colon |F(\boldsymbol{y})|\leqslant C_5Q\gamma\}.$$

Now using Lemma 2.4, this inclusion, (3.16), and the fact that $\mathcal{B}(\boldsymbol{z},\,T^{-1})\subset\mathcal{B}(\boldsymbol{x},\,r_0)$, we obtain

$$|\mathcal{B}(R_{\boldsymbol{a},a_0},\,\gamma)\cap\mathcal{B}(\boldsymbol{z},\,T^{-1})|\leqslant 12nC_5\gamma T^{-(d-1)}.$$

This implies inequality (3.12) with $K_3 = 12nC_5$.

It remains to show (3.11). If $d = 1$, then (3.11) holds with $K_2 = 1/2$. Thus we assume that $d > 1$.

Define the constant

$$C_6 = \min\left\{\frac{1}{8},\ \frac{1}{16(d-1)n^2L_2C_3^{1/(n+1)}}\right\}.$$

Let $\boldsymbol{z}' = (z_2,\,\dots,\,z_d)$, where $\boldsymbol{z} = (z_1,\,\dots,\,z_d)$. Fix any point $\boldsymbol{y}' = (y_2,\,\dots,\,y_d)\in\mathbb{R}^d$ such that $\|\boldsymbol{y}'-\boldsymbol{z}'\| < C_6T^{-1}$. Given $y_1\in\mathbb{R}$, we define the point $\boldsymbol{y} = (y_1,\,\boldsymbol{y}') = (y_1,\,y_2,\,\dots,\,y_d)$. If $|y_1 - z_1|\leqslant T^{-1}/8$ then

$$\begin{aligned}\|\boldsymbol{y}-\boldsymbol{z}\| = \sqrt{|y_1-z_1|^2+\|\boldsymbol{y}'-\boldsymbol{z}'\|^2} &\leqslant |y_1-z_1| + \|\boldsymbol{y}'-\boldsymbol{z}'\|\\ &< T^{-1}/8 + C_6T^{-1} = T^{-1}/4. \qquad (3.18)\end{aligned}$$

It follows that $\boldsymbol{y}\in\mathcal{B}(\boldsymbol{z},\,T^{-1}/4)$ whenever $|y_1-z_1|\leqslant T^{-1}/8$. By the Mean Value Theorem,

$$F(\boldsymbol{y}) = F(\boldsymbol{z}) + \nabla F(\tilde{\boldsymbol{y}})\cdot(\boldsymbol{y}-\boldsymbol{z}),$$

where $\tilde{\boldsymbol{y}}\in\mathcal{B}(\boldsymbol{z},\,T^{-1}/4)$. Since $F(\boldsymbol{z}) = 0$, we obtain

$$F(\boldsymbol{y})/\partial_1F(\tilde{\boldsymbol{y}}) = (y_1-z_1) + \sum_{i=2}^{d}\partial_iF(\tilde{\boldsymbol{y}})/\partial_1F(\tilde{\boldsymbol{y}})\cdot(y_i-z_i). \qquad (3.19)$$

Using (3.14), (3.16) and the inequality $\|\boldsymbol{z}'-\boldsymbol{y}'\| < C_6T^{-1}$, we find that

$$\left|\sum_{i=2}^{d}\partial_iF(\tilde{\boldsymbol{y}})/\partial_1F(\tilde{\boldsymbol{y}})\cdot(y_i-z_i)\right| < T^{-1}/8.$$

Therefore, the expression on the right of (3.19) is positive when $y_1 - z_1 = T^{-1}/8$ and is negative when $y_1 - z_1 = -T^{-1}/8$. Thus, the function $f(y_1) = F(\boldsymbol{y})/\partial_1F(\tilde{\boldsymbol{y}})$ has different signs at $\pm T^{-1}/8$. By the continuity, there is a point $y_1\in(-T^{-1}/8,\,T^{-1}/8)$ such that $f(y_1) = 0$, or, equivalently, $F(y_1,\,\dots,\,y_d) = 0$.

Thus, we have proved that for any $\boldsymbol{y}'$ with $\|\boldsymbol{y}'-\boldsymbol{z}'\| < C_6T^{-1}$ there is a point $y_1(\boldsymbol{y}')\in\mathbb{R}$ such that $\boldsymbol{y} = (y_1(\boldsymbol{y}'),\,\boldsymbol{y}')\in R_{\boldsymbol{a},a_0}\cap\mathcal{B}(\boldsymbol{z},\,T^{-1}/4)$. It is now easy to see that for any $\theta\in\mathbb{R}$ with $|\theta|\leqslant T^{-1}/4$ we have $(y_1(\boldsymbol{y}')+\theta,\,\boldsymbol{y}')\in\mathcal{B}(\boldsymbol{z},\,T^{-1}/2)$. Thus, for any positive γ with $\gamma < T^{-1}$ the set

$$A(\gamma) = \left\{(y_1(\boldsymbol{y}')+\theta,\,\boldsymbol{y}')\colon \|\boldsymbol{y}'-\boldsymbol{z}'\| < C_6T^{-1},\ |\theta|\leqslant\gamma/4\right\}$$

satisfies

$$A(\gamma)\subset\mathcal{B}(R_{\boldsymbol{a},a_0},\,\gamma)\cap\mathcal{B}(\boldsymbol{z},\,T^{-1}/2). \qquad (3.20)$$

By the Fubini theorem, it is easy to calculate that

$$|A(\gamma)| = |\mathcal{B}_{d-1}(\boldsymbol{z}', C_6T^{-1})|_{d-1} \cdot \gamma/2 = |\mathcal{B}_{d-1}(\mathbf{0}, C_6)|_{d-1}/2 \cdot \gamma \cdot T^{-(d-1)}.$$

Applying (3.20) now gives inequality (3.11) with $K_2 = |\mathcal{B}_{d-1}(\mathbf{0}, C_6)|_{d-1}/2$. □

Now we proceed to the proof of Theorem 3.2.

Assume that $Q > Q_0$. Choose a collection

$$(\boldsymbol{a}_1, a_{0,1}, \boldsymbol{z}_1), \dots, (\boldsymbol{a}_t, a_{0,t}, \boldsymbol{z}_t) \in (\mathbb{Z}^n \setminus \{\mathbf{0}\}) \times \mathbb{Z} \times \mathcal{B} \quad \text{with} \quad \boldsymbol{z}_i \in R_{\boldsymbol{a}_i, a_{0,i}}$$

such that

$$Q^{n+1} = T/C_3 \leqslant N(R_{\boldsymbol{a}_i, a_{0,i}}) \leqslant T = C_3Q^{n+1} \quad (1 \leqslant i \leqslant t) \tag{3.21}$$

and such that for any γ with $0 < \gamma < T^{-1}$ we have

$$K_2\gamma T^{-(d-1)} \leqslant |\mathcal{B}(R_{\boldsymbol{a}_i, a_{0,i}}, \gamma) \cap \mathcal{B}(\boldsymbol{z}_i, T^{-1}/2)| \quad (1 \leqslant i \leqslant t), \tag{3.22}$$

$$|\mathcal{B}(R_{\boldsymbol{a}_i, a_{0,i}}, \gamma) \cap \mathcal{B}(\boldsymbol{z}_i, T^{-1})| \leqslant K_3\gamma T^{-(d-1)} \quad (1 \leqslant i \leqslant t), \tag{3.23}$$

$$\mathcal{B}(\boldsymbol{z}_i, T^{-1}/2) \cap \mathcal{B}(\boldsymbol{z}_j, T^{-1}/2) = \varnothing \quad \text{for all different } i, j \quad (1 \leqslant i, j \leqslant t), \tag{3.24}$$

and the number t is maximal possible.

By Proposition 3.3, for any point $\boldsymbol{x} \in \mathcal{G}(\mathcal{B}; (4C_0)^{-1}; Q)$ there is a triple

$$(\boldsymbol{a}, a_0, \boldsymbol{z}) \in (\mathbb{Z}^n \setminus \{\mathbf{0}\}) \times \mathbb{Z} \times \mathcal{B} \quad \text{with} \quad \boldsymbol{z} \in R_{\boldsymbol{a}, a_0}$$

satisfying (3.9)–(3.12). By the maximality of t there is an index $i \in \{1, \dots, t\}$ such that

$$\mathcal{B}(\boldsymbol{z}_i, T^{-1}/2) \cap \mathcal{B}(\boldsymbol{z}, T^{-1}/2) \neq \varnothing.$$

It follows that $\|\boldsymbol{z} - \boldsymbol{z}_i\| < T^{-1}$. This inequality and (3.10) imply that $\|\boldsymbol{x} - \boldsymbol{z}_i\| < (C_4 + 1)T^{-1}$. Therefore,

$$\mathcal{G}(\mathcal{B}; (4C_0)^{-1}; Q) \subset \bigcup_{i=1}^{t} \mathcal{B}(\boldsymbol{z}_i, (C_4 + 1)T^{-1}).$$

By this inclusion and (3.21), we obtain

$$|\mathcal{B}|/2 \leqslant |\mathcal{G}(\mathcal{B}; (4C_0)^{-1}; Q)| \leqslant t \cdot |\mathcal{B}(\mathbf{0}, C_4 + 1)|T^{-d}.$$

Therefore, $t \geqslant K_1|\mathcal{B}|T^d$ with $K_1 = (2|\mathcal{B}(\mathbf{0}, C_4 + 1)|)^{-1}$.

Let $\lambda(x) = x/C_3$. Now, setting $R_i = R_{\boldsymbol{a}_i, a_{0,i}}$ and $\mathcal{B}_i = \mathcal{B}(\boldsymbol{z}_i, T^{-1}/2)$ gives the required collections of resonant sets and balls in the definition of regular system. This completes the proof of Theorem 3.2. □

4. Approximation by resonant sets

In this section we prove the following general result, which is an extension of Theorem 2 in [Ber99a].

Theorem 4.1. *Let U be an open set in $\mathbb{R}^d$, and let $(\mathcal{R}, N, s)$ be a regular system in U. Let $\Psi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a non-increasing function such that the sum*

$$\sum_{h=1}^{\infty} h^{d-s-1}\Psi^{d-s}(h) \tag{4.1}$$

diverges. Then for almost all points $\boldsymbol{x} \in U$ the inequality

$$\operatorname{dist}(\boldsymbol{x}, R) < \Psi(N(R)) \tag{4.2}$$

has infinitely many solutions $R \in \mathcal{R}$.

4.1. Auxiliary lemmas.

Lemma 4.2. *Let $E \subset \mathbb{R}^d$ be a measurable set, and let $U \subset \mathbb{R}^d$ be an open subset. Assume that there is a constant $\delta > 0$ such that for any finite ball $\mathcal{B} \subset U$ we have $|E \cap \mathcal{B}| \geqslant \delta|\mathcal{B}|$. Then E has full measure in U, i. e. $|U \setminus E| = 0$.*

Proof. Let $\tilde{E} = U \setminus E$. As $U \setminus \tilde{E} = U \cap E$, for any ball $\mathcal{B} \subset U$ we have $|\mathcal{B} \setminus \tilde{E}| \geqslant \delta|\mathcal{B}|$. Next, for any $\varepsilon > 0$ there is a cover of $\tilde{E}$ consisting of balls $\mathcal{B}_i$ such that

$$\sum_{i=1}^{\infty} |\mathcal{B}_i| - \varepsilon \leqslant |\tilde{E}| \leqslant \sum_{i=1}^{\infty} |\mathcal{B}_i|.$$

Notice that the sets $\mathcal{B}_i \setminus \tilde{E}$ and $\mathcal{B}_i \cap \tilde{E}$ are disjoint and satisfy $\mathcal{B}_i = (\mathcal{B}_i \setminus \tilde{E}) \cup (\mathcal{B}_i \cap \tilde{E})$. Then we get

$$\begin{aligned} |\tilde{E}| \geqslant \sum_{i=1}^{\infty} |\mathcal{B}_i| - \varepsilon = \sum_{i=1}^{\infty} |\mathcal{B}_i \setminus \tilde{E}| + \sum_{i=1}^{\infty} |\mathcal{B}_i \cap \tilde{E}| - \varepsilon \\ \geqslant \delta \sum_{i=1}^{\infty} |\mathcal{B}_i| + \left| \bigcup_{i=1}^{\infty} \mathcal{B}_i \cap \tilde{E} \right| - \varepsilon \geqslant \delta|\tilde{E}| + |\tilde{E}| - \varepsilon. \end{aligned}$$

Therefore, $|\tilde{E}| \leqslant \varepsilon/\delta \to 0$ as $\varepsilon \to 0$. Hence, $\tilde{E}$ is null and E has full measure in U. □

Lemma 4.3 (Lemma 5, Chapter 1 in [Spr79]). *Let $E_i \subset \mathbb{R}^d$ be a sequence of measurable sets, and let the set E consist of points $\boldsymbol{x}$ belonging to infinitely many E_i. If there is a sufficiently large ball in $\mathbb{R}^d$ which contains all the sets E_i, and the sum $\sum_{i=1}^{\infty} |E_i|$ diverges, then*

$$|E| \geqslant \limsup_{N \to \infty} \frac{\left(\sum_{i=1}^{N} |E_i|\right)^2}{\sum_{i=1}^{N} \sum_{j=1}^{N} |E_i \cap E_j|}. \tag{4.3}$$

Lemma 4.4. *Let Ψ satisfy the conditions of Theorem 4.1, and define $\tilde{\Psi}(h) = \min\{ch^{-1}, \Psi(h)\}$, where $c > 0$ is a constant. Then $\tilde{\Psi}$ is non-increasing and the sum*

$$\sum_{h=1}^{\infty} h^{d-s-1} \tilde{\Psi}^{d-s}(h) \tag{4.4}$$

diverges.

Proof. The monotonicity of $\tilde{\Psi}$ is easily verified. Assume that (4.4) converges. Then, by the monotonicity, we have

$$l^{d-s} \tilde{\Psi}^{d-s}(l) \ll \sum_{l/2 \leqslant h \leqslant l} h^{d-s-1} \tilde{\Psi}^{d-s}(h) \to 0 \quad \text{as} \quad l \to \infty.$$

It follows that $l\tilde{\Psi}(l) = \min\{c, l\Psi(l)\} \to 0$ as $l \to \infty$. This is possible only if $l\Psi(l) \to 0$ as $l \to \infty$. It follows that $\tilde{\Psi}(l) = \Psi(l)$ for all sufficiently large l. Therefore, the sum (4.1) converges, contrary to the conditions of Lemma 4.4. □

Lemma 4.5. *Let $\Psi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be non-increasing. Fix any $d > 0$. Then the sums*

$$\sum_{h=1}^{\infty} h^{d-1}\Psi^d(h) \quad \textit{and} \quad \sum_{k=0}^{\infty} 2^{kd}\Psi^d(2^k)$$

converge or diverge simultaneously.

Proof. Using the monotonicity of Ψ we get the following inequalities:

$$2^{(k+1)d}\Psi^d(2^{k+1}) \ll \sum_{2^k \leqslant h < 2^{k+1}} h^{d-1}\Psi^d(h) \ll 2^{kd}\Psi^d(2^k).$$

Summing these over all $k \in \mathbb{N}$ gives the required property. □

4.2. Proof of Theorem 4.1. By Lemma 4.4, there is no loss of generality in assuming that for all $h > 0$

$$\Psi(h) \leqslant h^{-1}/2. \tag{4.5}$$

Fix any ball $\mathcal{B} \subset U$ and set $T = 2^k$. By Definition 1.5, there are constants K_1, K_2, $K_3 > 0$, which do not depend on $\mathcal{B}$, and there is a sufficiently large number k_0 satisfying the following properties: for any natural number $k \geqslant k_0$ there are resonant sets $R_k^{(i)} \in \mathcal{R}$ $(1 \leqslant i \leqslant t_k)$ and balls $\mathcal{B}_k^{(i)}$ with $2\mathcal{B}_k^{(i)} \subset \mathcal{B}$ $(1 \leqslant i \leqslant t_k)$ such that

$$\lambda(2^k) \leqslant N(R_k^{(i)}) \leqslant 2^k \quad (1 \leqslant i \leqslant t_k), \tag{4.6}$$

$$\operatorname{diam} \mathcal{B}_k^{(i)} = 2^k \quad (1 \leqslant i \leqslant t_k), \tag{4.7}$$

$$\mathcal{B}_k^{(i)} \cap \mathcal{B}_k^{(j)} = \varnothing \quad (1 \leqslant i, j \leqslant t_k, \ \ i \neq j), \tag{4.8}$$

$$K_2\gamma^{d-s}2^{-sk} \leqslant |\mathcal{B}(R_k^{(i)}, \gamma) \cap \mathcal{B}_k^{(i)}|, \tag{4.9}$$

and

$$|\mathcal{B}(R_k^{(i)}, \gamma) \cap 2\mathcal{B}_k^{(i)}| \leqslant K_3\gamma^{d-s}2^{-sk} \tag{4.10}$$

for any γ, $0 < \gamma < 2^{-k}$,

$$K_1 2^{dk}|\mathcal{B}| \leqslant t_k \leqslant 2^{dk}|\mathcal{B}|. \tag{4.11}$$

For every natural number $k \geqslant k_0$ and $i \in \{1, \ldots, t_k\}$ we define the sets

$$E_k^{(i)} = \mathcal{B}(R_k^{(i)}, \Psi(2^k)) \cap \mathcal{B}_k^{(i)}$$

and

$$E_k = \bigcup_{i=1}^{t_k} E_k^{(i)}. \tag{4.12}$$

It follows from (4.9) and (4.10) that

$$K_2\Psi^{d-s}(2^k)2^{-sk} \leqslant |E_k^{(i)}| \leqslant K_3\Psi^{d-s}(2^k)2^{-sk}. \tag{4.13}$$

It follows from (4.8) that

$$E_k^{(i)} \cap E_k^{(j)} = \varnothing \quad \text{if } i \neq j,\ 1 \leqslant i, j \leqslant t_k. \tag{4.14}$$

Therefore, $|E_k| = \sum_{i=1}^{t_k} |E_k^{(i)}|$. Using (4.11) and (4.13), we find that

$$K_1K_2\Psi^{d-s}(2^k)2^{(d-s)k}|\mathcal{B}| \leqslant |E_k| \leqslant 2K_1K_3\Psi^{d-s}(2^k)2^{(d-s)k}|\mathcal{B}|.$$

Let $\phi_k = 2^{(d-s)k}\Psi^{d-s}(2^k)$. Then we have

$$K_1K_2|\mathcal{B}|\phi_k \leqslant |E_k| \leqslant 2K_1K_3|\mathcal{B}|\phi_k. \tag{4.15}$$

Using the divergence of (4.1) and applying Lemma 4.5, we obtain

$$\sum_{k=1}^{\infty} \phi_k = \infty. \tag{4.16}$$

It follows from (4.15) and (4.16) that $\sum_{k=k_0}^{\infty} |E_k| = \infty$. Since $\mathcal{B}$ is bounded and all the sets E_k are contained in $\mathcal{B}$, Lemma 4.3 can be applied to the sequence E_k. We are now going to obtain estimates for the numerator and the denominator in (4.3).

When $K > k_0$, inequalities (4.15) imply that

$$\sum_{k=k_0}^{K} |E_k| \geqslant K_1K_2|\mathcal{B}| \sum_{k=k_0}^{K} \phi_k. \tag{4.17}$$

Now we proceed to the estimate of the measure of $E_k \cap E_l$. Let $k_0 \leqslant k < l \leqslant K$, where $K > k_0$. Using (4.12), we can write

$$E_l \cap E_k^{(i)} = \bigcup_{j=1}^{t_l} E_l^{(j)} \cap E_k^{(i)}.$$

By (4.13), we find that $|E_l^{(j)} \cap E_k^{(i)}| \leqslant K_3\Psi^{d-s}(2^l)2^{-sl}$. Hence,

$$|E_l \cap E_k^{(i)}| \leqslant K_3\Psi^{d-s}(2^l)2^{-sl} \cdot q(l,\, k,\, i), \tag{4.18}$$

where $q(l,\, k,\, i)$ is the number of different indices j such that $E_l^{(j)} \cap E_k^{(i)} \neq \varnothing$.

Now we will estimate $q(l,\, k,\, i)$. Using (4.7) and (4.8), we get

$$\begin{aligned} \left| \bigcup_{j=1,\dots,t_l:\ E_l^{(j)} \cap E_k^{(i)} \neq \varnothing} \mathcal{B}_l^{(j)} \right| &= |\mathcal{B}(\mathbf{0},\, 2^{-l}/2)| \cdot q(l,\, k,\, i) \\ &= |\mathcal{B}(\mathbf{0},\, 1/2)| \cdot 2^{-dl} q(l,\, k,\, i). \end{aligned} \tag{4.19}$$

Consider any ball $\mathcal{B}_l^{(j)}$ such that $E_l^{(j)} \cap E_k^{(i)} \neq \varnothing$. Fix a point $\boldsymbol{x} \in E_l^{(j)} \cap E_k^{(i)}$. By the definition of $E_k^{(i)}$, there is a point $\boldsymbol{z} \in R_k^{(i)}$ such that

$$\|\boldsymbol{x} - \boldsymbol{z}\| < \Psi(2^k). \tag{4.20}$$

Next, since $\boldsymbol{x} \in E_l^{(j)} \subset \mathcal{B}_l^{(j)}$, for any point $\boldsymbol{y} \in \mathcal{B}_l^{(j)}$ we have

$$\|\boldsymbol{y} - \boldsymbol{x}\| < \operatorname{diam} \mathcal{B}_l^{(j)} = 2^{-l}. \tag{4.21}$$

Then, using (4.20) and (4.21), we obtain

$$\|\boldsymbol{y} - \boldsymbol{z}\| < \|\boldsymbol{y} - \boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{z}\| < 2^{-l} + \Psi(2^k).$$

Therefore,

$$\operatorname{dist}(\boldsymbol{y},\, R_k^{(i)}) < 2^{-l} + \Psi(2^k),$$

whence

$$\mathcal{B}_l^{(j)} \subset \mathcal{B}(R_k^{(i)},\, 2^{-l} + \Psi(2^k)). \tag{4.22}$$

Let $\boldsymbol{x}_0$ denote the center of $\mathcal{B}_k^{(i)}$. For $\boldsymbol{x} \in E_k^{(i)} \subset \mathcal{B}_k^{(i)}$, we have

$$\|\boldsymbol{x} - \boldsymbol{x}_0\| < \frac{1}{2} \operatorname{diam} \mathcal{B}_k^{(i)}. \tag{4.23}$$

Using the inequality $l > k$ and (4.21), we obtain

$$\|\boldsymbol{x} - \boldsymbol{y}\| < \frac{1}{2} \operatorname{diam} \mathcal{B}_k^{(i)}.$$

On combining the last inequality with (4.23), we get

$$\|\boldsymbol{y} - \boldsymbol{x}_0\| < \|\boldsymbol{y} - \boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{x}_0\| < \operatorname{diam} \mathcal{B}_k^{(i)}.$$

Thus, $\mathcal{B}_l^{(j)} \subset 2\mathcal{B}_k^{(i)}$. Using this inclusion and (4.22) gives

$$\bigcup_{j=1,\dots,t_l:\ E_l^{(j)} \cap E_k^{(i)} \neq \varnothing} \mathcal{B}_l^{(j)} \subset \mathcal{B}(R_k^{(i)},\, 2^{-l} + \Psi(2^k)) \cap 2\mathcal{B}_k^{(i)}. \tag{4.24}$$

Now, applying (4.10), (4.24), and the monotonicity of the measure, we derive

$$\begin{aligned}\Bigg|\bigcup_{j=1,\dots,t_l:\ E_l^{(j)} \cap E_k^{(i)} \neq \varnothing} \mathcal{B}_l^{(j)}\Bigg| &\leqslant K_3(2^{-l} + \Psi(2^k))^{d-s} 2^{-sk} \\ &\leqslant K_3 2^{d-s}(2^{-l(d-s)} + \Psi^{d-s}(2^k))2^{-sk}.\end{aligned}$$

On combining this inequality and (4.19), we obtain

$$q(l,\, k,\, i) \ll 2^{s(l-k)} + 2^{dl} 2^{-sk} \Psi^{d-s}(2^k). \tag{4.25}$$

It follows from (4.18) and (4.25) that

$$|E_l \cap E_k^{(i)}| \ll \Psi^{d-s}(2^l) 2^{-sk} + 2^{(d-s)l} 2^{-sk} \Psi^{d-s}(2^k) \Psi^{d-s}(2^l). \tag{4.26}$$

Since the number of different sets $E_k^{(i)}$ does not exceed t_k, we have

$$|E_l \cap E_k| \leqslant t_k \cdot \max_{1 \leqslant i \leqslant t_k} |E_l \cap E_k^{(i)}|.$$

Using this inequality, (4.11), and (4.26), we get

$$\begin{aligned}|E_l \cap E_k| &\ll |\mathcal{B}| \Psi^{d-s}(2^l) 2^{(d-s)k} \big(1 + 2^{(d-s)l} \Psi^{d-s}(2^k)\big) \\ &= |\mathcal{B}|(2^{-(d-s)(l-k)} \phi_l + \phi_k \phi_l).\end{aligned} \tag{4.27}$$

For arbitrary l, k with $k_0 \leqslant l, k \leqslant K$, we have

$$|E_l \cap E_k| \ll |\mathcal{B}|(2^{-(d-s)|l-k|} \phi_l + \phi_k \phi_l). \tag{4.28}$$

By (4.16), there is a sufficiently large number K' such that for all $K > K'$

$$\sum_{k=k_0}^{K} \phi_k > 1. \tag{4.29}$$

Let $K > K'$. Now using (4.15), (4.27), and (4.29), we calculate

$$\begin{aligned}\sum_{l=k_0}^{K}\sum_{k=k_0}^{K}|E_l\cap E_k| &\ll |\mathcal{B}|\sum_{l=k_0}^{K}\sum_{k=k_0}^{K}\phi_k\phi_l+|\mathcal{B}|\sum_{l=k_0}^{K}\sum_{k=k_0}^{K}2^{-(d-s)|l-k|}\phi_l\\ &\leqslant |\mathcal{B}|\sum_{l=k_0}^{K}\sum_{k=k_0}^{K}\phi_k\phi_l+|\mathcal{B}|\sum_{l=k_0}^{K}\phi_l\sum_{k\in\mathbb{Z}}2^{-(d-s)|l-k|}\\ &= |\mathcal{B}|\left(\sum_{l=k_0}^{K}\phi_l\right)^2+(1+2^{s-d})/(1-2^{s-d})|\mathcal{B}|\sum_{l=k_0}^{K}\phi_l \ll |\mathcal{B}|\left(\sum_{k=k_0}^{K}\phi_k\right)^2\end{aligned}$$

where the implicit constant in this estimate does not depend on either $\mathcal{B}$ or K. Using (4.17) now gives

$$\frac{\left(\sum_{k=k_0}^{K}|E_k|\right)^2}{\sum_{l=k_0}^{K}\sum_{k=k_0}^{K}|E_l\cap E_k|} \gg |\mathcal{B}| \tag{4.30}$$

when $K > K'$. By Lemma 4.3, the measure of the set E, consisting of points $\boldsymbol{x}$ which belong to infinitely many sets E_k, is at least $(K_1K_2)^2/C_{10}\cdot|\mathcal{B}|$.

Using the monotonicity of Ψ and inequalities (4.6), it is easy to see that for any point $\boldsymbol{x}\in E$ inequality (4.2) has infinitely many solutions. Let $\mathcal{R}(\Psi)$ denote the set of points $\boldsymbol{x}\in U$ such that inequality (4.2) has infinitely many solutions. Then $E\subset\mathcal{R}(\Psi)\cap\mathcal{B}$. It follows that $|\mathcal{R}(\Psi)\cap\mathcal{B}|\geqslant|E|\gg|\mathcal{B}|$. By Lemma 4.2, the set $\mathcal{R}(\Psi)$ has full measure in U. The proof of Theorem 4.1 is completed.

5. Proof of the main theorem

It is obvious that we can restrict ourselves to a sufficiently small ball $\mathcal{B}_0$ centered at a point belonging to a set with full measure in U. By Theorem 3.2 we can take $\mathcal{B}_0$ to be such that $(\mathcal{R},\, N,\, s)$ is a regular system in $\mathcal{B}_0$, where $s = d-1$, N and $\mathcal{R}$ are defined in the statement of Theorem 3.2. Define the sequence Ψ by setting

$$dnL_2h\Psi(h^{n+1}) = \psi(h^n).$$

Thus $\Psi(k) = k^{-1/(n+1)}\psi(k^{n/(n+1)})/dnL_2$. Since ψ is non-increasing, Ψ is non-increasing as well. Next, we calculate

$$\begin{aligned}\sum_{h=1}^{\infty}h^{d-s-1}\Psi^{d-s}(h) &= \sum_{h=1}^{\infty}\Psi(h)\\ &= \frac{1}{dnL_2}\sum_{k=1}^{\infty}\sum_{(k-1)^{(n+1)/n}<h\leqslant k^{(n+1)/n}}h^{-1/(n+1)}\psi(h^{n/(n+1)})\\ &\gg \sum_{k=1}^{\infty}\sum_{(k-1)^{(n+1)/n}<h\leqslant k^{(n+1)/n}}k^{-1/n}\psi(k)\geqslant\sum_{k=1}^{\infty}\psi(k)=\infty.\end{aligned}$$

By Theorem 4.1, for almost all $\boldsymbol{x}\in U$ there are infinitely many $(\boldsymbol{a},\, a_0)\in\mathbb{Z}^n\times\mathbb{Z}$ satisfying

$$\operatorname{dist}(\boldsymbol{x},\, R_{\boldsymbol{a},a_0}) < \Psi(\|\boldsymbol{a}\|_\infty^{n+1}). \tag{5.1}$$

It follows from (5.1) that there is a point $\boldsymbol{z} \in R_{\boldsymbol{a},a_0}$ such that

$$\|\boldsymbol{x} - \boldsymbol{z}\| < \Psi(\|\boldsymbol{a}\|_\infty^{n+1}). \tag{5.2}$$

By the definition of $R_{\boldsymbol{a},a_0}$, we have $F(\boldsymbol{z}) = \boldsymbol{a} \cdot \boldsymbol{f}(\boldsymbol{z}) + a_0 = 0$. Using the Mean Value Theorem, we obtain

$$F(\boldsymbol{x}) = F(\boldsymbol{z}) + \nabla F(\tilde{\boldsymbol{x}}) \cdot (\boldsymbol{x} - \boldsymbol{z}) = \nabla F(\tilde{\boldsymbol{x}}) \cdot (\boldsymbol{x} - \boldsymbol{z}) = (\boldsymbol{a} \nabla \boldsymbol{f}(\tilde{\boldsymbol{x}})) \cdot (\boldsymbol{x} - \boldsymbol{z}), \tag{5.3}$$

where $\tilde{\boldsymbol{x}}$ is a point between $\boldsymbol{x}$ and $\boldsymbol{z}$. Using (3.8), we find that

$$\begin{aligned} |\langle \boldsymbol{a} \cdot \boldsymbol{f}(\boldsymbol{x})\rangle| = |F(\boldsymbol{x})| &\leqslant d \|\boldsymbol{a} \nabla f(\tilde{\boldsymbol{x}})\|_\infty \cdot \|\boldsymbol{x} - \boldsymbol{z}\|_\infty \\ &< dn \|\boldsymbol{a}\|_\infty L_2 \Psi(\|\boldsymbol{a}\|_\infty^{n+1}) = \psi(\|\boldsymbol{a}\|_\infty^n). \end{aligned} \tag{5.4}$$

As we have shown above, for almost all $\boldsymbol{x} \in U$ there are infinitely many $(\boldsymbol{a},\, a_0) \in \mathbb{Z}^n \times \mathbb{Z}$ satisfying (5.1). Therefore, for almost all $\boldsymbol{x} \in U$ there are infinitely many $\boldsymbol{a}$ satisfying (5.4). This completes the proof of Theorem 1.6.

6. Concluding remarks

In this section we give a brief account of other results in metric Diophantine approximation and state the most important problems in this field. We also discuss possible developments of the theory of regular systems and difficulties that prevent us from proving multiplicative divergence Khintchine type results.

6.1. Simultaneous approximation. The point $\boldsymbol{y} \in \mathbb{R}^n$ is called *simultaneously ψ-approximable* if

$$\|\langle q\boldsymbol{y}\rangle\|_\infty^n < \psi(q) \tag{6.1}$$

has infinitely many solutions $q \in \mathbb{Z}$. By the Khintchine transference principle, a point $\boldsymbol{y} \in \mathbb{R}^n$ is very well approximable if and only if it is simultaneously ψ_ε-approximable for some positive ε, where $\psi_\varepsilon(h) = h^{-(1+\varepsilon)}$. Unfortunately there is no such connection between simultaneous and dual approximation for general approximation functions ψ that would make it possible to derive a Khintchine type theorem for the simultaneous case from the dual and vice versa. However, it has been known since the 1926 paper of Khintchine that almost all (almost no) points of $\mathbb{R}^n$ are simultaneously ψ-approximable if the sum (1.2) diverges (converges).

Let $\mathcal{M}$ be a submanifold of $\mathbb{R}^n$. One says that $\mathcal{M}$ is of *Khintchine type for divergence* (*for convergence*) if almost all (almost no) points of $\mathcal{M}$ are simultaneously ψ-approximable whenever the sum (1.2) diverges (converges).

We mostly deal with monotonic approximation errors. However, it is worth saying that for $n > 1$ an analogue of Khintchine's theorem for nonmonotonic error function has been obtained by A. Pollington and R. Vaughan [PV90], who proved a multidimensional analogue of the Duffin–Schaeffer conjecture.

Only special manifolds have been proved to be of Khintchine type. Bernik [Ber79] has shown that the parabola $\{(x,\, x^2) : x \in \mathbb{R}\}$ is of Khintchine type for convergence. He has also proved by a method of trigonometric sums that any manifold given as a topological product of at least 4 planar curves with curvatures non-vanishing almost everywhere is of Khintchine type for both convergence and divergence [Ber73]. A class of manifolds in $\mathbb{R}^n$ with a special geometrical property, which substantially restricts the dimension of the manifolds, has been proved to be of Khintchine type for both convergence and divergence [DRV91], [DRV96].

In the Khintchine type theory for simultaneous Diophantine approximation the following is regarded as the main problem.

Problem 1. Prove that a nondegenerate manifold $\mathcal{M}$ in $\mathbb{R}^n$ is of Khintchine type for convergence and for divergence.

It is of interest to consider some special cases of Problem 1 such as the circle, the sphere and others. There remain two classical special cases of Problem 1: to prove that for $n \geqslant 3$ the curve $\mathcal{V}_n$ is of Khintchine type for convergence and to prove that for $n \geqslant 2$ the curve $\mathcal{V}_n$ is of Khintchine type for divergence.

One difficulty in the simultaneous Diophantine approximation is that there is no longer the dichotomy of big/small derivative (the derivative is always big) but the investigated sets are quite rare. Thus one needs a considerably new technique to break through the problem.

A much deeper problem is to prove asymptotic formulae for the number of solutions of Diophantine inequalities under consideration. This remains unsettled for both linear and simultaneous approximation.

6.2. Multiplicative results. The point $\boldsymbol{y} \in \mathbb{R}^n$ is said to be *ψ-multiplicatively approximable* if the inequality

$$|\langle \boldsymbol{a} \cdot \boldsymbol{x} \rangle| < \psi(\Pi_+(\boldsymbol{a})) \tag{6.2}$$

has infinitely many solutions $\boldsymbol{a} \in \mathbb{Z}^n$, where $\Pi_+(\boldsymbol{a}) = \prod_{i=1}^n \max(|a_i|, 1)$. One can define *very well multiplicatively approximable* points to be ψ_ε-multiplicatively approximable for some positive ε, with $\psi_\varepsilon(h) = h^{-1-\varepsilon}$.

By the Borel–Cantelli lemma, almost all points of $\mathbb{R}^n$ are not ψ-multiplicatively approximable whenever the sum

$$\sum_{h=1}^{\infty} (\log h)^{n-1} \psi(h) \tag{6.3}$$

converges. Since $\Pi_+(\boldsymbol{a})$ is not greater than $\|\boldsymbol{a}\|_\infty^n$, any ψ-approximable point is automatically ψ-multiplicatively approximable. Therefore, a very well approximable point is also very well multiplicatively approximable.

A manifold $\mathcal{M}$ is said to be of *multiplicative Groshev type for divergence* (*convergence*) if almost all (almost no) points of $\mathcal{M}$ are multiplicatively ψ-approximable whenever the sum (6.3) diverges (converges). A manifold $\mathcal{M}$ is said to be *strongly extremal* if almost all points of $\mathcal{M}$ are not very well multiplicatively approximable.

The problem of proving strong extremality in connection with multiplicative approximation was first raised by Baker in [Bak90, Ch. 9, p. 96]. The question, as initially proposed, related to the Veronese curve and it was later generalized to any nondegenerate manifold by Sprindzuk. Baker was motivated in part by the non-metrical instances of specific points known to have the property of strong extremality, *i. e.*, the algebraic numbers and powers of e [Bak90, Ch. 7 and Ch. 10].

Kleinbock and Margulis [KM98] proved that any nondegenerate manifold is strongly extremal, and later jointly with Bernik [BKM99] they proved a stronger result that these manifolds are of multiplicative Groshev type for convergence. They even established a more general result, to be stated in Section 6.3. No manifold (except $\mathbb{R}^n$ itself) has ever been shown to be of multiplicative Groshev type for divergence.

The difficulty of proving multiplicative Groshev type theorems for divergence with the method of this paper is that Minkowski's theorem on convex bodies cannot be efficiently extended to nonconvex bodies, e. g. star bodies, which appear in the context of multiplicative approximation. One might try to relax the definition of regular system used in this paper by taking a multi-valued function N to control any possible difference in the magnitude of integer coefficients. In this way however one would loose a sufficient estimate for denominators in (4.3). Thus more investigation is required to prove a multiplicative Groshev type theorem for divergence.

Problem 2. Prove that any nondegenerate manifold is of multiplicative Groshev type for divergence.

One can also consider a multiplicative version of simultaneous Diophantine approximation when one replaces the right-hand side of (6.1) with $\prod_{i=1}^{n} |\langle qy_i\rangle|$. Khintchine type theorems for this type of approximation have never been proved for convergence or for divergence.

6.3. A general approximation function. Let $\Psi\colon \mathbb{Z} \to \mathbb{R}_+$, n, $m \in \mathbb{N}$. The point $\boldsymbol{y} \in \mathbb{R}^{nm}$ is said to be (Ψ, n, m)-*approximable* if the inequality

$$\|\langle \boldsymbol{a}\boldsymbol{y}\rangle\|_\infty^m < \Psi(\boldsymbol{a}) \tag{6.4}$$

has infinitely many solutions $\boldsymbol{a} \in \mathbb{Z}^n$. The point $\boldsymbol{y}$ is considered to be a matrix with n rows and m columns.

Due to Schmidt [Sch60], [Sch64a] one knows the following most general result on Diophantine approximation of independent quantities.

Let m, $n \in \mathbb{N}$, $n \geqslant 2$, $\Psi\colon \mathbb{Z}^n \to \mathbb{R}_+$. Almost all (almost no) points $\boldsymbol{y} \in \mathbb{R}^{nm}$ are (Ψ, n, m)-approximable whenever the sum

$$\sum_{\boldsymbol{a}\in\mathbb{Z}^n} \Psi(\boldsymbol{a}) \tag{6.5}$$

diverges (converges).

For the case of $m = 1$ and under some monotonicity restrictions on Ψ, Bernik, Kleinbock, and Margulis extended the convergence part of this result to nondegenerate manifolds. More precisely, assuming that for every $i = \overline{1,\, n}$, whenever $|q_i| \leqslant |q_i'|$ and $q_i q_i' > 0$,

$$\Psi(q_1, \ldots, q_i, \ldots, q_n) \geqslant \Psi(q_1, \ldots, q_i', \ldots, q_n), \tag{6.6}$$

they proved that almost no point $\boldsymbol{y} \in \mathcal{M}$ is $(\Psi, n, 1)$-approximable whenever the sum $(6.5)_{m=1}$ converges, where $\mathcal{M}$ is a given nondegenerate manifold.

Problem 3. Assuming (6.6), prove that almost all points $\boldsymbol{y} \in \mathcal{M}$ are $(\Psi, n, 1)$-approximable whenever the sum (6.5) diverges, where $\mathcal{M}$ is a given nondegenerate manifold.

It is also of interest to investigate Diophantine approximation (of any type) with nonmonotonic error function (right-hand side of inequalities).

Another interesting problem is to find reasonable conditions on the entries of the matrix $\boldsymbol{y}$ in (6.4) when they are dependent, so that one would have an extremality type or Khintchine–Groshev (or Schmidt) type theorem.

6.4. Hausdorff dimension. The first results on the Hausdorff dimension of sets arising in Diophantine approximation are due to V. Jarnik and A. S. Besicovitch. They found the exact value of the Hausdorff dimension of the set of w-approximable points (i. e. $\psi_{w/n-1}$-approximable points with $\psi_\varepsilon(h) = h^{-1-\varepsilon}$) in the real line.

The first general method for obtaining lower bounds for the Hausdorff dimension was suggested by Baker and Schmidt. They introduced the concept of regular systems, which made it possible to efficiently describe the distribution of objects that were used for approximation. Baker and Schmidt have proved with their method that the set of w-approximable points on $\mathcal{V}_n$ has dimension at least $\frac{n+1}{w+1}$, and conjectured that this number is the right upper bound as well. The Baker–Schmidt conjecture was proved by Bernik [Ber83] in 1983. Extending the ideas of Baker and Schmidt, Dodson and H. Dickinson [DD00] have shown that for any extremal manifold $\mathcal{M}$ in $\mathbb{R}^n$ the set of w-approximable points on the manifold has Hausdorff dimension at least $\frac{n+1}{w+1} + \dim \mathcal{M} - 1$. Thus we've got a very natural

Problem 4. Let $w > n$ and $\mathcal{M}$ be a nondegenerate manifold in $\mathbb{R}^n$. Prove that the Hausdorff dimension of w-approximable points on $\mathcal{M}$ is exactly $\frac{n+1}{w+1} + \dim \mathcal{M} - 1$.

Also, Dodson [Dod92], [Dod93] has investigated the Hausdorff dimension of the set of (Ψ, n, m)-approximable points when $\Psi(\boldsymbol{a}) = \psi(\|\boldsymbol{a}\|_\infty)$, $\psi\colon \mathbb{R}_+ \to \mathbb{R}_+$ and Dickinson and S. Velani [DV97] answered a very general question on the Hausdorff measure (with respect to a general dimension function) of this set and proved a Khintchine–Groshev type theorem.

The problem of calculating the Hausdorff dimension in the case of simultaneous Diophantine approximation seems to be even more complicated (see [BD99, pp. 92–98]). The Hausdorff dimension of simultaneously v-approximable points (i. e. simultaneously $\psi_{(nv-1)}$-approximable with $\psi_\varepsilon(h) = h^{-1-\varepsilon}$) with v big enough seems to depend on arithmetic and other properties of the manifold one would like to approximate (see [BD99, pp. 90–98]). However there might be a general formula for v close to the extremal exponent $1/n$.

6.5. Beyond the nondegeneracy condition. Looking for new classes of extremal, Khintchine or Groshev type manifolds is a challenging task. The simplest ones for which the nondegeneracy condition fails are proper affine subspaces of $\mathbb{R}^n$. They have been studied in several papers in the past, and some conditions (written in terms of Diophantine properties of coefficients of parameterizing equations) have been found sufficient for their extremality [Sch64b], [Spr79] and, in the case of straight lines passing through the origin, for being of Groshev type for both convergence and divergence [BBDD00].

Recently, in a preprint [Kle02] by Kleinbock, using the dynamical approach of [KM98], necessary and sufficient conditions for extremality and strong extremality of any affine subspace of $\mathbb{R}^n$ have been written down. Also it has been shown there that a smooth submanifold $\mathcal{M}$ of an affine subspace $\mathcal{L}$ of $\mathbb{R}^n$ is extremal (resp. strongly extremal) whenever $\mathcal{L}$ is such, provided $\mathcal{M}$ is *nondegenerate in* $\mathcal{L}$.

The latter notion is a straightforward generalization of Definition 1.5, so that a submanifold $\mathcal{M}$ of $\mathcal{L}$ is nondegenerate in $\mathcal{L}$ if it cannot be "too well" approximated by hyperplanes contained in $\mathcal{L}$.

This naturally leads to the following

Problem 5. Find criteria for an affine subspace $\mathcal{L}$ of $\mathbb{R}^n$ to be of Groshev type for convergence or divergence; or, given a specific function ψ such that the sum (1.2) diverges (converges), find necessary and sufficient conditions for almost all (almost no) points of $\mathcal{L}$ to be ψ-approximable. Also, prove that the aforementioned properties of $\mathcal{L}$ are inherited by its submanifolds which are nondegenerate in $\mathcal{L}$.

It is also worthwhile to mention that one can investigate Diophantine properties of almost all (almost no) points with respect to measures other than Lebesgue measures on smooth manifolds. The latter can be supported on fractal subsets of $\mathbb{R}$ (see [Wei01]) or $\mathbb{R}^n$ ([KLW02], the work currently in progress).

Acknowledgments. Bernik, Kleinbock and Margulis are grateful to SFB-343 and Humboldt Foundation for the support of their 1999 stay at the University of Bielefeld, where a preliminary form of this paper was discussed. Beresnevich is grateful to EPSRC for the support of his stay at the University of York in 2000, where a part of the paper was written, and to Prof. Maurice Dodson for his hospitality and making arrangements for this stay.

References

[Bak66] A. Baker, *On a theorem of Sprindžuk*, Proc. Royal Soc. **292** (1966), 92–104.

[Bak78] R. C. Baker, *Dirichlet's theorem on Diophantine approximation*, Math. Proc. Cambridge Philos. Soc. **83** (1978), no. 1, 37–59. MR 58 #5535

[Bak90] A. Baker, *Transcendental number theory*, 3rd ed., Cambridge Math. Library, Cambridge University Press, Cambridge, 1990. MR **91f:**11049

[BBDD99] V. V. Beresnevich, V. I. Bernik, H. Dickinson, M. M. Dodson, *The Khintchine–Groshev theorem for planar curves*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **455** (1999), no. 1988, 3053–3063. MR **2001j:**11069

[BBDD00] V. V. Beresnevich, V. I. Bernik, H. Dikinson, M. Dodson, *On linear manifolds for which the Khinchin approximation theorem holds*, Vestsī Nats. Akad. Navuk Belarusī Ser. Fīz.-Mat. Navuk (2000), no. 2, 14–17, 139 (Russian). MR **2001j:**11068

[BD99] V. I. Bernik, M. M. Dodson, *Metric Diophantine approximation on manifolds*, Cambridge Tracts in Mathematics, 137, Cambridge University Press, Cambridge, 1999. MR **2001h:**11091

[BDD98] V. I. Bernik, H. Dickinson, M. M. Dodson, *A Khintchine-type version of Schmidt's theorem for planar curves*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **454** (1998), no. 1968, 179–185. MR **99e:**11099

[Ber73] V. I. Bernik, *Asymptotic behavior of the number of solutions of certain systems of inequalities in the theory of Diophantine approximations of dependent variables*, Vescī Akad. Navuk BSSR Ser. Fīz.-Mat. Navuk (1973), no. 1, 10–17, 134 (Russian). MR 49 #2632

[Ber79] V. I. Bernik, *The exact order of approximation of almost all points of a parabola*, Mat. Zametki **26** (1979), no. 5, 657–665, 813 (Russian). MR **81c:**10067

[Ber83] V. I. Bernik, *Application of the Hausdorff dimension in the theory of Diophantine approximations*, Acta Arith. **42** (1983), no. 3, 219–253 (Russian). MR **85k:**11035

[Ber89] V. I. Bernik, *The exact order of approximating zero by values of integral polynomials*, Acta Arith. **53** (1989), no. 1, 17–28. MR **91d:**11079

[Ber99a] V. Beresnevich, *On approximation of real numbers by real algebraic numbers*, Acta Arith. **90** (1999), no. 2, 97–112. MR **2000f:**11087

[Ber99b] V. V. Beresnevich, *A Khintchine–Groshev type theorem on manifolds: the case of convergence*, Math. Dept., The University of York, York (England), April 1999, Preprint.

[Ber00a] V. V. Beresnevich, *An analogue of the Khinchin–Groshev theorem for curves in* $\mathbb{R}^n$, Dokl. Nats. Akad. Nauk Belarusi **44** (2000), no. 3, 29–32 (Russian). MR **2001m:**11123

[Ber00b] V. V. Beresnevich, *An analogue of the Khinchin–Groshev theorem for nondegenerate manifolds in* $\mathbb{R}^n$, Dokl. Nats. Akad. Nauk Belarusi **44** (2000), no. 4, 42–45, 124 (Russian). MR **2002f:**11091

[Ber00c] V. V. Beresnevich, *Application of the concept of regular systems of points in metric number theory*, Vestsī Nats. Akad. Navuk Belarusī Ser. Fīz.-Mat. Navuk (2000), no. 1, 35–39, 140 (Russian). MR **2001e:**11082

[Ber00d] V. V. Beresnevich, *On a proof of analogues of the Khinchin theorem for curves*, Vestsī Nats. Akad. Navuk Belarusī Ser. Fīz.-Mat. Navuk (2000), no. 3, 35–40, 140. MR **2002d:**11090

[Ber02] V. V. Beresnevich, *A Groshev type theorem for convergence on manifolds*, Acta Math. Hungar. **94** (2002), no. 1–2, 99–130.

[BKM99] V. Bernik, D. Kleinbock, G. A. Margulis, *Khintchine-type theorems on manifolds: convergence case for standard and multiplicative versions*, Universität Bielefeld, Bielefeld, 1999, Preprint 99–092.

[BKM01] V. Bernik, D. Kleinbock, G. A. Margulis, *Khintchine-type theorems on manifolds: the convergence case for standard and multiplicative versions*, Internat. Math. Res. Notices (2001), no. 9, 453–486. MR 1 829 381

[BS70] A. Baker, W. M. Schmidt, *Diophantine approximation and Hausdorff dimension*, Proc. London Math. Soc. (3) **21** (1970), 1–11. MR 42 #5916

[DD00] H. Dickinson, M. M. Dodson, *Extremal manifolds and Hausdorff dimension*, Duke Math. J. **101** (2000), no. 2, 271–281. MR **2001f:**11116

[Dod92] M. M. Dodson, *Hausdorff dimension, lower order and Khintchine's theorem in metric Diophantine approximation*, J. Reine Angew. Math. **432** (1992), 69–76. MR **94a:**11125

[Dod93] M. M. Dodson, *Geometric and probabilistic ideas in the metric theory of Diophantine approximations*, Uspekhi Mat. Nauk **48** (1993), no. 5(293), 77–106 (Russian), Russian Math. Surveys **48** (1993), no. 5, 73–102. MR **95d:**11092

[DRV90a] M. M. Dodson, B. P. Rynne, J. A. G. Vickers, *Diophantine approximation and a lower bound for Hausdorff dimension*, Mathematika **37** (1990), no. 1, 59–73. MR **91i:**11098

[DRV90b] M. M. Dodson, B. P. Rynne, J. A. G. Vickers, *Dirichlet's theorem and Diophantine approximation on manifolds*, J. Number Theory **36** (1990), no. 1, 85–88. MR **91k:**11063

[DRV91] M. M. Dodson, B. P. Rynne, J. A. G. Vickers, *Khintchine-type theorems on manifolds*, Acta Arith. **57** (1991), no. 2, 115–130. MR **92m:**11070

[DRV96] M. M. Dodson, B. P. Rynne, J. A. G. Vickers, *Simultaneous Diophantine approximation and asymptotic formulae on manifolds*, J. Number Theory **58** (1996), no. 2, 298–316. MR **97f:**11055

[DV97] D. Dickinson, S. L. Velani, *Hausdorff measure and linear forms*, J. Reine Angew. Math. **490** (1997), 1–36. MR **98e:**11094

[Gro38] A. Groshev, *A theorem on a system of linear forms*, Dokl. Akad. Nauk SSSR **19** (1938), 151–152 (Russian).

[Khi24] A. J. Khintchine, *Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen*, Math. Ann. **92** (1924), 115–125.

[Khi26] A. J. Khintchine, *Zur metrischen Theorie der diophantischen Approximationen*, Math. Zeitschr **24** (1926), 706–714.

[Kle01] D. Kleinbock, *Some applications of homogeneous dynamics to number theory*, Smooth ergodic theory and its applications (Seattle, WA, 1999), Amer. Math. Soc., Providence, RI, 2001, pp. 639–660. MR 1 858 548

[Kle02] D. Kleinbock, *Extremal subspaces and their submanifolds*, Preprint, 2002.

[KLW02] D. Y. Kleinbock, E. Lindenstrauss, B. Weiss, *On fractal sets and Diophantine approximation*, Preprint, 2002.

[KM98] D. Y. Kleinbock, G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. of Math. (2) **148** (1998), no. 1, 339–360. MR **99j:**11083

[KM99] D. Y. Kleinbock, G. A. Margulis, *Logarithm laws for flows on homogeneous spaces*, Invent. Math. **138** (1999), no. 3, 451–494. MR **2001i:**37046

[Kub49] J. P. Kubilius, *On an application of Vinogradov's method to the solving of a problem in metrical number theory*, Dokl. Akad. Nauk SSSR **67** (1949), 783–786 (Russian).

[Mah32] K. Mahler, *Über das Maßder Menge aller S-Zahlen*, Math. Ann. **106** (1932), 131–139.

[PV90] A. D. Pollington, R. C. Vaughan, *The k-dimensional Duffin and Schaeffer conjecture*, Mathematika **37** (1990), no. 2, 190–200. MR **92d:**11084

[Pya69] A. S. Pyartli, *Diophantine approximations of submanifolds of a Euclidean space*, Funktsional. Anal. i Prilozhen. **3** (1969), no. 4, 59–62. MR 40 #4209

[Sch60] W. Schmidt, *A metrical theorem in Diophantine approximation*, Canad. J. Math. **12** (1960), 619–631. MR 22 #9482

[Sch64a] W. M. Schmidt, *Metrical theorems on fractional parts of sequences*, Trans. Amer. Math. Soc. **110** (1964), 493–518. MR 28 #3018

[Sch64b] W. M. Schmidt, *Metrische Sätze über simultane Approximation abhängiger Grössen*, Monatsh. Math. **68** (1964), 154–166. MR 30 #1980

[Spr69] V. G. Sprindzhuk, *Mahler's problem in metric number theory*, Translations of Mathematical Monographs, vol. 25, American Mathematical Society, Providence, R.I., 1969. MR 39 #6833

[Spr79] V. G. Sprindzhuk, *Metric theory of Diophantine approximations*, V. H. Winston & Sons, Washington, D.C., 1979. MR **80k:**10048

[Vol61] B. Volkmann, *The real cubic case of Mahler's conjecture*, Mathematika **8** (1961), 55–57. MR 24 #A1254

[Wei01] B. Weiss, *Almost no points on a Cantor set are very well approximable*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **457** (2001), 949–952. MR 1 875 309

Victor Beresnevich: Department of Number Theory, Institute of Mathematics, the National Belarus Academy of Sciences, 220072, Surganova 11, Minsk, Belarus

E-mail address: beresnevich@im.bas-net.by

Vasily Bernik: Department of Number Theory, Institute of Mathematics, the National Belarus Academy of Sciences, 220072, Surganova 11, Minsk, Belarus

E-mail address: bernik@im.bas-net.by

Dmitry Kleinbock: Department of Mathematics, Brandeis University, Waltham, MA 02454, USA

E-mail address: kleinboc@brandeis.edu

Gregory Margulis: Department of Mathematics, Yale University, New Haven, Connecticut 06520, USA

E-mail address: margulis@math.yale.edu