Invent. math. (2015) 202:1199–1240
DOI 10.1007/s00222-015-0586-8

Badly approximable points on manifolds

Victor Beresnevich

Received: 26 October 2014 / Accepted: 20 February 2015 / Published online: 5 March 2015

Abstract This paper is motivated by two problems in the theory of Diophantine approximation, namely, Davenport's problem regarding badly approximable points on submanifolds of a Euclidean space and Schmidt's problem regarding the intersections of the sets of weighted badly approximable points. The problems have been recently settled in dimension two but remain open in higher dimensions. In this paper we develop new techniques that allow us to tackle them in full generality. The techniques rest on lattice points counting and a powerful quantitative result of Bernik, Kleinbock and Margulis. The main theorem of this paper implies that any finite intersection of the sets of weighted badly approximable points on any analytic nondegenerate submanifold of $\mathbb{R}^n$ has full dimension. One of the consequences of this result is the existence of transcendental real numbers badly approximable by algebraic numbers of any bounded degree.

Mathematics Subject Classification 11J13 · 11J83

1 Introduction

The notion of badly approximable numbers, as much of the classical and modern theory of Diophantine approximation, is underpinned by Dirichlet's

Supported by EPSRC Grant EP/J018260/1.

V. Beresnevich (✉)
University of York, Heslington, York YO10 5DD, UK
e-mail: victor.beresnevich@york.ac.uk

fundamental result. It states that for every $\alpha \in \mathbb{R}$ and any $Q > 1$ there exists $q \in \mathbb{N}$ and $p \in \mathbb{Z}$ such that $|q\alpha - p| < Q^{-1}$ and $q \leq Q$. In particular, it implies that for every real irrational number α the inequality

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}$$

holds for infinitely many rational numbers p/q written as reduced fractions of integers p and q. A real number α is then called *badly approximable* if there exists a constant $c = c(\alpha) > 0$ such that

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c}{q^2} \tag{1}$$

for all $(q, p) \in \mathbb{N} \times \mathbb{Z}$. In what follows, the set of badly approximable real numbers will be denoted by **Bad**.

It is well known that a real irrational number α is badly approximable if and only if the partial quotients of its continued fraction expansion are uniformly bounded. For instance, any real quadratic irrational number is in **Bad**, since its continued fraction expansion is eventually periodic.[1] Using continued fractions one can easily produce continuum many examples of badly approximable real numbers. Beyond the cardinality, Jarník [27] established that dim **Bad** (the Hausdorff dimension of **Bad**) is 1. However, the Lebesgue measure of **Bad** is known to be zero. This is a trivial consequence of the divergence case of Khintchine's theorem [37], and can also be relatively easily proved using the Lebesgue density theorem, see [18] or [15, Corollary 2].

1.1 Higher dimensions: Schmidt's conjecture

Higher dimensions offer various ways of generalising the notion of badly approximable numbers. For now, we restrict ourselves to considering simultaneous Diophantine approximations by rationals. The point $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ is called *badly approximable* if there exists a constant $c = c(\mathbf{y}) > 0$ such that

$$\max_{1 \leq i \leq n} \|qy_i\| \geq cq^{-1/n} \tag{2}$$

for all $q \in \mathbb{N}$, where $\|x\|$ denotes the distance of x from the nearest integer. The quantities $\|qy_i\|$ are equal to $|qy_i - p_i|$ for some $p_i \in \mathbb{Z}$ and thus give rise to 'approximating' rationals $p_1/q, \ldots, p_n/q$. Once again, the notion of badly approximable points is underpinned by Dirichlet's theorem, this time

[1] It is not known whether there are any real algebraic numbers of degree ≥ 3 that are badly approximable.

for $\mathbb{R}^n$, which implies that the inequality $\max_{1\le i\le n}\|qy_i\| < q^{-1/n}$ holds for infinitely many $q \in \mathbb{N}$. The set of badly approximable points in $\mathbb{R}^n$ will be denoted by $\mathbf{Bad}(n)$. Observe that $\mathbf{Bad}(1) = \mathbf{Bad}$.

The first examples of badly approximable points in $\mathbb{R}^n$ were given by Perron [40] who used an algebraic construction and produced infinitely yet countably many elements of $\mathbf{Bad}(n)$. For instance, $(\alpha, \dots, \alpha^n) \in \mathbf{Bad}(n)$ whenever α is a real algebraic number of degree $n+1$. However, it was not until 1954 when first Davenport [21] for $n = 2$ and then Cassels [19] for $n \ge 2$ showed that $\mathbf{Bad}(n)$ was uncountable. The fact that $\mathbf{Bad}(n)$ has full Hausdorff dimension was proved by Schmidt [43] who introduced powerful ideas based on a specific type of games. The dimension result for $\mathbf{Bad}(n)$ comes about as a consequence of the fact that $\mathbf{Bad}(n)$ is winning for Schmidt's game. Furthermore, Schmidt proved that affine transformations of $\mathbf{Bad}(n)$ are winning and that the collection of winning sets in $\mathbb{R}^n$ is closed under countable intersections.

In his 1983 paper [46] Schmidt formulated a conjecture that later became the catalysis for some remarkable developments. Schmidt's conjecture rests on the modified notion of badly approximable points in which approximations in each coordinate are given some weights, say $r_1, \dots, r_n$. In short, he conjectured that there exist points in $\mathbb{R}^2$ that are simultaneously badly approximable with respect to two different collections of weights. The weights of approximation are required to satisfy the following conditions:

$$r_1 + \cdots + r_n = 1 \quad \text{and} \quad r_i \ge 0 \text{ for all } \ i = 1, \dots, n\,. \tag{3}$$

Throughout this paper the set of all n-tuples $\mathbf{r} = (r_1, \dots, r_n)$ subject to (3) will be denoted by $\mathcal{R}_n$. Formally, given $\mathbf{r} \in \mathcal{R}_n$, the point $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ will be called $\mathbf{r}$-*badly approximable* if there exists $c = c(\mathbf{y}) > 0$ such that

$$\max_{1\le i\le n} \|qy_i\|^{1/r_i} \ge cq^{-1} \tag{4}$$

for all $q \in \mathbb{N}$. Here, by definition, $\|qy_i\|^{1/0} = 0$. Again, a version of Dirichlet's theorem tells us that when $c = 1$ inequality (4) fails infinitely often.

The set of $\mathbf{r}$-badly approximable points in $\mathbb{R}^n$ will be denoted by $\mathbf{Bad}(\mathbf{r})$. As is readily seen, the classical set of badly approximable points $\mathbf{Bad}(n)$ is simply $\mathbf{Bad}(\frac{1}{n}, \dots, \frac{1}{n})$. Using this notation we can now specify the following concrete statement conjectured by Schmidt:

$$\mathbf{Bad}\left(\tfrac{1}{3}, \tfrac{2}{3}\right) \cap \mathbf{Bad}\left(\tfrac{2}{3}, \tfrac{1}{3}\right) \ne \emptyset.$$

It is worth mentioning that the sets $\mathbf{Bad}(\mathbf{r})$ have been studied at length in all dimensions and for arbitrary collections of weights, see [22,34–36,41]. Partly the interest was fueled by natural links with homogeneous dynamics and Little-

wood's conjecture in multiplicative Diophantine approximation, another long standing problem—see [13] for further details. Schmidt's conjecture withstood attacks for nearly 30 years. However, the recent progress has been dramatic.

In 2011 Badziahin et al. [13] made a breakthrough by proving that for any sequence $\mathbf{r}_k = (i_k, j_k) \in \mathcal{R}_2$ such that

$$\liminf_{k\to\infty} \min\{i_k, j_k\} > 0 \tag{5}$$

and any vertical line $L_\theta = \{(\theta, y) : y \in \mathbb{R}\} \subset \mathbb{R}^2$ with $\theta \in \mathbf{Bad}$ one has that

$$\dim \bigcap_k \mathbf{Bad}(\mathbf{r}_k) \cap L_\theta = 1. \tag{6}$$

This readily gives that $\dim \bigcap_k \mathbf{Bad}(\mathbf{r}_k) = 2$ and proves Schmidt's conjecture in a much stronger sense. Shortly thereafter, An [1] proves that for any $\mathbf{r} \in \mathcal{R}_2$ and any $\theta \in \mathbf{Bad}$ the set $\mathbf{Bad}(\mathbf{r}) \cap L_\theta$ is winning for a Schmidt game in L_θ. This immediately leads him to removing condition (5) from the theorem of Badziahin et al., since the collection of Schmidt's winning sets is closed under arbitrary countable intersections. In a related paper An [2] establishes that $\mathbf{Bad}(i, j)$ is winning for the 2-dimensional Schmidt game, thus giving another proof of Schmidt's conjecture. Generalising the techniques of [13] in yet another direction Nesharim [39], independently from An, proves that the set in the left hand side of (6) intersected with naturally occurring fractals embedded in L_θ is uncountable for any sequence $(\mathbf{r}_k)_{k\in\mathbb{N}}$. Subsequently, Nesharim jointly with Weiss establishes the winning property of these intersections—see Appendix B in [39].

As already mentioned, the sets $\mathbf{Bad}(\mathbf{r})$ and even their restrictions to naturally occurring fractals have been investigated in higher dimensions, see [26,34–36]. In particular, the sets $\mathbf{Bad}(\mathbf{r})$ were shown to have full Hausdorff dimension for any $\mathbf{r} \in \mathcal{R}_n$. However, the theory of their mutual intersections is a different story. In an apparent attempt to prove Schmidt's conjecture, Kleinbock and Weiss [36] introduced a modified version of Schmidt's games. As they have shown, winning sets for the same modified Schmidt game inherit the properties of classical winning sets. Namely, they have full Hausdorff dimension and their countable intersections are winning with respect to the same game. Also Kleinbock and Weiss have proved that $\mathbf{Bad}(\mathbf{r})$ is winning for a relevant modified Schmidt game. However, it was not possible to prove that the intersection $\mathbf{Bad}(\mathbf{r}_1) \cap \mathbf{Bad}(\mathbf{r}_2)$ was a winning set for some modified Schmidt game as, with very few exceptions, the corresponding modified Schmidt games were not 'compatible'. As a result the following key problem that generalises Schmidt's original conjecture has remained open in dimensions $n \geq 3$:

Problem 1 Let $n \in \mathbb{N}$. Prove that for any finite or countable subset W of $\mathcal{R}_n$ one has that

$$\dim \bigcap_{\mathbf{r}\in W} \mathbf{Bad}(\mathbf{r}) = n\,. \tag{7}$$

The main result of this paper implies (7) in arbitrary dimensions n and for arbitrary countable subsets W of weights satisfying a condition similar to (5). For instance, the result is applicable to arbitrary finite collections of weights W. The proof will be given by restricting the sets of interest to a suitable family of curves in $\mathbb{R}^n$. Interestingly, this approach, which was innovated in [13] in the case $n = 2$, turns out to face another intricate problem that was first communicated by Davenport.

1.2 $\mathbf{Bad}(\mathbf{r})$ on manifolds and Davenport's problem

In 1964 Davenport [22] established that, given a finite collection $\mathbf{f}_i : \mathbb{R}^m \to \mathbb{R}^{n_i}$ $(1 \le i \le N)$ of C^1 maps, if for some $\mathbf{x}_0 \in \mathbb{R}^m$ and every $i = 1, \dots, N$ the Jacobian of $\mathbf{f}_i$ at $\mathbf{x}_0$ has rank n_i, then the set

$$\bigcap_{i=1}^{N} \mathbf{f}_i^{-1}(\mathbf{Bad}(n_i))$$

has the power of continuum. For instance, taking $f_1(x, y) = x$, $f_2(x, y) = y$ and $\mathbf{f}_3(x, y) = (x, y)$ shows that $\mathbf{Bad}(1, 0) \cap \mathbf{Bad}\left(\frac{1}{2}, \frac{1}{2}\right) \cap \mathbf{Bad}(0, 1)$ has the power of continuum. Another natural example obtained by taking $f_i(x) = x^i$ for $i = 1, \dots, k$ shows that there are continuum many $\alpha \in \mathbb{R}$ such that $\alpha, \alpha^2, \dots, \alpha^k$ are all in $\mathbf{Bad}$.

Clearly, the Jacobian condition above implies that $m \ge n_i$ for every i. Commenting on this, Davenport writes [22, p. 52] "Problems of a much more difficult character arise when the number of independent parameters is less than the dimension of simultaneous approximation. I do not know whether there is a set of α with the cardinal of the continuum such that the pair (α, α^2) is badly approximable for simultaneous approximation". Essentially, if $m < n_i$ then $\mathbf{f}_i(\mathbf{x})$ lies on a submanifold of $\mathbb{R}^{n_i}$. Hence, Davenport's problem boils down to investigating badly approximable points restricted to submanifolds of Euclidean spaces.

In the theory of Diophantine approximation on manifolds, see for instance, [7–9,31,33], there are already well established classes of manifolds of interest. These include non-degenerate manifolds and affine subspaces and should likely be of primary interest when resolving Davenport's problem.

It is worth pointing out that the result of Perron [40] mentioned in § 1.1 implies the existence of algebraic badly approximable points on the Veronese

curves $\mathcal{V}_n = \{(x, \dots, x^n) : x \in \mathbb{R}\}$. However, there are only countably many of them. Khintchine [28] proved that $\mathbf{Bad}(n) \cap \mathcal{V}_n$ had zero 1-dimensional Lebesgue measure. Baker [4] generalised this to arbitrary C^1 submanifold of $\mathbb{R}^n$. Apparently, $\mathbf{Bad}(n)$ can be relatively easily replaced with $\mathbf{Bad}(\mathbf{r})$ in Baker's result, though, to the best of author's knowledge, this has never been formally addressed. To make a long story short, until recently there has been no success in relation to Davenport's problem even for planar curves, let alone manifolds in higher dimension. The aforementioned work of Badziahin et al. [13] was the first step forward. Very recently, assuming (5), Badziahin and Velani [17] have proved (6) with L_θ replaced by any C^2 planar curve which is not a straight line. In particular, this shows that there exist uncountably many real numbers α such that (α, α^2) is in $\mathbf{Bad}(2)$. Also they have dealt with a family of lines in $\mathbb{R}^2$ satisfying a natural Diophantine condition. The most recent results established in [3] by An, Velani and the author of this paper remove condition (5) from the findings of [17] and at the same time settle Davenport's problem for a larger class of lines in $\mathbb{R}^2$ defined by a near optimal condition. As a result, the following general version of Davenport's problem is essentially settled in the case $n = 2$:

Problem 2 Let $n, m \in \mathbb{N}$, B be a ball in $\mathbb{R}^m$, W be a finite or countable subset of $\mathcal{R}_n$ and $\mathcal{F}_n(B)$ be a finite or countable collection of maps $\mathbf{f} : B \to \mathbb{R}^n$. Determine sufficient (and possibly necessary) conditions on W and/or $\mathcal{F}_n(B)$ so that

$$\dim \bigcap_{\mathbf{f}\in\mathcal{F}_n(B)} \bigcap_{\mathbf{r}\in W} \mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r})) = m\,. \tag{8}$$

Despite the success in resolving Problem 2 for planar curves, no progress has been made on Davenport's problem for $n \geq 3$. The results of this paper imply (8) in arbitrary dimensions n and for arbitrary countable subsets W of weights satisfying a condition similar to (5) and arbitrary finite collection $\mathcal{F}_n(B)$ of analytic non-degenerate maps. The proof introduces new ideas based on lattice points counting and a powerful quantitative result of Bernik, Kleinbock and Margulis. Indeed, the arguments presented should be of independent interest even for $n = 2$.

2 Main results and corollaries

In what follows, an analytic map $\mathbf{f} : B \to \mathbb{R}^n$ defined on a ball $B \subset \mathbb{R}^m$ will be called *nondegenerate* if the functions $1, f_1, \dots, f_n$ are linearly independent over $\mathbb{R}$. The more general notion of nondegeneracy that does not require analyticity can be found in [33]. Given an integer $n \geq 2$, $\mathcal{F}_n(B)$ will denote a family of maps $\mathbf{f} : B \to \mathbb{R}^n$ with a common domain B. To avoid ambiguity, let us agree

from the beginning that all the intervals and balls mentioned in this paper are of positive and finite diameter. Recall that $\mathcal{R}_n$ denotes the collection of weights of approximation and is defined by (3). Given $\mathbf{r} = (r_1, \ldots, r_n) \in \mathcal{R}_n$, let

$$\tau(\mathbf{r}) \stackrel{\text{def}}{=} \min\{r_i : r_i \neq 0\}, \tag{9}$$

that is $\tau(\mathbf{r})$ is the smallest strictly positive weight within $\mathbf{r}$. The following result regarding Problem 2 represents the main finding of this paper.

Theorem 1 *Let $m, n \in \mathbb{N}$, $1 \leq m \leq n$, B be an open ball in $\mathbb{R}^m$ and $\mathcal{F}_n(B)$ be a finite family of analytic nondegenerate maps. Let W be a finite or countable subset of $\mathcal{R}_n$ such that*

$$\inf\{\tau(\mathbf{r}) : \mathbf{r} \in W\} > 0\,. \tag{10}$$

Then (8) *is satisfied.*

Condition (10) matches (5) and is satisfied whenever W is finite. Now we consider the following basic corollary regarding badly approximable points on manifolds.

Corollary 1 *Let $\mathcal{M}$ be a manifold immersed into $\mathbb{R}^n$ by an analytic nondegenerate map. Let $W \subset \mathcal{R}_n$ be a finite or countable set of weights. Assume that* (10) *is satisfied. Then* $\dim \bigcap_{\mathbf{r}\in W} \mathbf{Bad}(\mathbf{r}) \cap \mathcal{M} = \dim \mathcal{M}$ *. In particular, for any finite collection $\mathbf{r}_1, \ldots, \mathbf{r}_N \in \mathcal{R}_n$ we have that*

$$\dim \bigcap_{k=1}^{N} \mathbf{Bad}(\mathbf{r}_k) \cap \mathcal{M} = \dim \mathcal{M}.$$

Note that the corollary is applicable to $\mathcal{M} = \mathbb{R}^n$, which is clearly analytic and nondegenerate. In this case Corollary 1 establishes an analogue of Schmidt's conjecture in arbitrary dimensions $n \geq 2$ by settling Problem 1 subject to condition (10).

2.1 Reduction to curves

When $m = 1$ the nondegeneracy of an analytic map $\mathbf{f} = (f_1, \ldots, f_n)$ is equivalent to the Wronskian of $f_1', \ldots, f_n'$ being not identically zero. More generally, the map $\mathbf{f}$ (not necessarily analytic) defined on an interval $I \subset \mathbb{R}$ will be called *nondegenerate at $x_0 \in I$* if $\mathbf{f}$ is C^n on a neighborhood of x_0 and the Wronskian of $f_1', \ldots, f_n'$ does not vanish at x_0. This definition of nondegeneracy at a single point is adopted within the following more general result for curves. Note that if $\mathbf{f}$ is nondegenerate at least at one point, then the functions $1, f_1, \ldots, f_n$ are linearly independent over $\mathbb{R}$.

Theorem 2 *Let $n \in \mathbb{N}$, $n \geq 2$, $I \subset \mathbb{R}$ be an open interval and $\mathcal{F}_n(I)$ be a finite family of maps defined on I nondegenerate at the same point $x_0 \in I$. Let W be a finite or countable subset of $\mathcal{R}_n$ satisfying (10). Then*

$$\dim \bigcap_{\mathbf{f}\in\mathcal{F}_n(I)} \bigcap_{\mathbf{r}\in W} \mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r})) = 1. \tag{11}$$

Our immediate goal is to show that Theorem 1 is a consequence of Theorem 2. In metric Diophantine approximation the idea of reducing the case of manifolds to curves is not new. For instance, Badziahin et al. [13] use fibering of $\mathbb{R}^2$ into vertical lines in their proof of Schmidt's conjecture. Underpinning our reduction of Theorem 1 to Theorem 2 is the following version of Marstrand's slicing lemma, see [24, Corollary 7.12] or [25, Theorem 10.11].

Marstrand's slicing Lemma *Let $m > 1$ and S be a subset of $\mathbb{R}^m$. Let $s > 0$ and let U be a subset of $\mathbb{R}^{m-1}$ such that $\dim\{(t, u_2, \dots, u_m) \in S\} \geq s$ for each $(u_2, \dots, u_m) \in U$. Then*

$$\dim S \geq \dim U + s.$$

We will also need the following formal statement which is a slightly modified extract from Sprindžuk's survey [48, pp. 9–10].

The Fibering Lemma *Let $f_0, \dots, f_n$ be analytic functions in m real variables defined on an open neighborhood of $\mathbf{0}$. Assume that $f_0, \dots, f_n$ are linearly independent over $\mathbb{R}$. Then there is a sufficiently large integer $d_0 > 1$ such that for every $d > d_0$ and every $\mathbf{u} = (u_1, u_2, \dots, u_m) \in \mathbb{R}^m$ with $u_1 \dots u_m \neq 0$ the following functions of one real variable*

$$\phi_{\mathbf{u},i} : E_{\mathbf{u}} \to \mathbb{R} \quad (0 \leq i \leq n)$$

given by

$$\phi_{\mathbf{u},i}(t) \stackrel{\text{def}}{=} f_i\left(u_1 t^{1+d^m}, u_2 t^{d+d^m}, \dots, u_m t^{d^{m-1}+d^m}\right),$$

where $E_{\mathbf{u}} \subset \mathbb{R}$ is a neighbourhood of 0, are linearly independent over $\mathbb{R}$.

Although the proof of the Fibering Lemma mostly follows the argument of [48, pp. 9–10], for completeness full details are given in Appendix C. Note that Sprindžuk's version of fibering involves the parametrisation $\widetilde{\phi}_{\mathbf{u},i}(t) = f_i(u_1 t, u_2 t^d, \dots, u_m t^{d^{m-1}})$.

Proof of Theorem 1 modulo Theorem 2 Let $\mathcal{F}_n(B)$ be as in Theorem 1 and let $\mathbf{f} = (f_1, \dots, f_n) \in \mathcal{F}_n(B)$. Without loss of generality we will assume that B

is centred at $\mathbf{0}$. Also assume that $m \geq 2$ as otherwise there is nothing to prove. Let $u_1 = 1$, $t_0 > 0$ and $\delta_2, \ldots, \delta_m > 0$ be sufficiently small numbers such that

$$\left(t^{1+d^m}, u_2 t^{d+d^m}, u_3 t^{d^2+d^m}, \ldots, u_m t^{d^{m-1}+d^m}\right) \in B$$

whenever

$$\tfrac{1}{2}t_0 < t < t_0, \quad \tfrac{1}{2}\delta_i < u_i < \delta_i \quad (2 \leq i \leq m)\,. \tag{12}$$

The existence of $t_0, \delta_2, \ldots, \delta_m$ is guaranteed by the fact that $\mathbf{0}$ is an interior point of B. Let U be the set of $\mathbf{u} = (u_2, \ldots, u_m)$ satisfying the right hand side inequalities of (12) and D be the set of $(t, u_2, \ldots, u_m)$ satisfying (12).

By the nondegeneracy of $\mathbf{f}$, the functions $1, f_1, \ldots, f_n$ are linearly independent. Since they are also analytic, by the Fibering Lemma, there exists $d_0(\mathbf{f}) > 0$ such that for every $d > d_0(\mathbf{f})$ and every $\mathbf{u} \in U$ the coordinate functions of the map

$$\mathbf{f}_{\mathbf{u}}(t) = \mathbf{f}\left(t^{1+d^m}, u_2 t^{d+d^m}, u_3 t^{d^2+d^m}, \ldots, u_m t^{d^{m-1}+d^m}\right) \tag{13}$$

defined on the interval $I = \left(\frac{1}{2}t_0, t_0\right)$ together with 1 are linearly independent over $\mathbb{R}$. Since $\mathcal{F}_n(B)$ is finite,

$$d_0 = \max\{d_0(\mathbf{f}) : \mathbf{f} \in \mathcal{F}_n(B)\}$$

is well defined. Let $d > d_0$. Then for every $\mathbf{f} \in \mathcal{F}_n(B)$ and every $\mathbf{u} \in U$ the coordinate functions of the map (13) together with 1 are linearly independent over $\mathbb{R}$. By the well known criterion of linear independence, their Wronskian is not identically zero. Hence, the Wronskian of $\mathbf{f}'_{\mathbf{u}} = \frac{d}{dt}\mathbf{f}_{\mathbf{u}}$ is not identically zero. As an analytic function, it has isolated zeros. Hence, for a fixed $\mathbf{u}$, there are at most countably many points in I where the Wronskian of $\mathbf{f}'_{\mathbf{u}}$ vanishes for some $\mathbf{f} \in \mathcal{F}_n(B)$. Hence, there exists a point $x_0 \in I$, which may depend on $\mathbf{u}$, such that for every $\mathbf{f} \in \mathcal{F}_n(B)$ the Wronskian of $\mathbf{f}'_{\mathbf{u}}$ is not zero, that is $\mathbf{f}_{\mathbf{u}}$ is non-degenerate at x_0. Thus, Theorem 2 is applicable and we conclude that the following subset of I

$$S_{\mathbf{u}} = \bigcap_{\mathbf{f} \in \mathcal{F}_n(B)} \bigcap_{\mathbf{r} \in W} \mathbf{f}_{\mathbf{u}}^{-1}(\mathbf{Bad}(\mathbf{r}))$$

has Hausdorff dimension 1. Here, by definition, $\mathbf{f}_{\mathbf{u}}^{-1}(\mathbf{Bad}(\mathbf{r}))$ is the set of $t \in I = \left(\frac{1}{2}t_0, t_0\right)$ such that $\mathbf{f}_{\mathbf{u}}(t) \in \mathbf{Bad}(\mathbf{r})$. Then, by Marstrand's slicing lemma, the set

$$S = \{(t, u_2, \ldots, u_m) : t \in S_{\mathbf{u}},\ \mathbf{u} \in U\} \subset D$$

has Hausdorff dimension $\geq \dim U + 1 = m$. Let $S' \subset B$ be the image of S under the map

$$\begin{aligned}(t, u_2, \dots, u_m) &\mapsto (x_1, \dots, x_m) \\ &\stackrel{\text{def}}{=} \left(t^{1+d^m}, u_2 t^{d+d^m}, u_3 t^{d^2+d^m}, \dots, u_m t^{d^{m-1}+d^m}\right). \end{aligned} \tag{14}$$

Then, in view of the definitions of S, $S_{\mathbf{u}}$ and $\mathbf{f}_{\mathbf{u}}$, we have that

$$S' \subset \bigcap_{\mathbf{f} \in \mathcal{F}_n(B)} \bigcap_{\mathbf{r} \in W} \mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r})) \,. \tag{15}$$

Further, note that (14) maps D into B injectively and is bi-Lipschitz on D, since the map itself and its inverse (defined on the image of D) have continuous bounded derivatives. It is well known that bi-Lipschitz maps preserves Hausdorff dimension, see for example [24, Corollary 2.4]. Therefore, $\dim S' = \dim S \geq m$. By (15), and the fact that any subset of $\mathbb{R}^m$ is of dimension $\leq m$, we obtain (8) and thus complete the proof of Theorem 1 modulo Theorem 2. □

2.2 The dual form of approximation

So far we have been dealing with simultaneous rational approximations. Here we introduce the dual definition of badly approximable points—see part (iii) of Lemma 1 below. This has two purposes. Firstly, it is the dual form that will be used in the proof of the results. Secondly, the dual form provides a natural environment for considering Diophantine approximation by algebraic numbers and will allow us to deduce further corollaries of our main results.

Lemma 1 (Equivalent definitions of $\mathbf{Bad}(\mathbf{r})$) *Let $\mathbf{r} = (r_1, \dots, r_n) \in \mathcal{R}_n$ and $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$. Then the following three statements are equivalent:*

(i) $\mathbf{y} \in \mathbf{Bad}(\mathbf{r})$.

(ii) *There exists $c > 0$ such that for any $Q \geq 1$ the only integer solution $(q, p_1, \dots, p_n)$ to the system*

$$|q| < Q, \quad |q y_i - p_i| < \left(c\, Q^{-1}\right)^{r_i} \quad (1 \leq i \leq n) \tag{16}$$

is $q = p_1 = \dots = p_n = 0$.

(iii) *There exists $c > 0$ such that for any $H \geq 1$ the only integer solution $(a_0, a_1, \dots, a_n)$ to the system*

$$|a_0 + a_1 y_1 + \dots + a_n y_n| < c H^{-1}, \quad |a_i| < H^{r_i} \quad (1 \leq i \leq n) \tag{17}$$

is $a_0 = \dots = a_n = 0$.

The equivalence of (i) and (ii) is a straightforward consequence of the definition of $\mathbf{Bad}(\mathbf{r})$. The equivalence of (ii) and (iii) is relatively well known, see Appendix in [13] for a similar statement. Indeed, this equivalence is essentially a special case of Mahler's version of Khintchine's transference Principle appearing in [38]. To make this paper self-contained we provide further details in Appendix A.

2.3 Approximation by algebraic numbers of bounded degree

There are two classical interrelated settings in the theory of approximation by algebraic numbers of bounded degree. One of them boils down to investigating small values of integral polynomials P with $\deg P \le n$ at a given number ξ. The other deals with the proximity of algebraic numbers α of degree $\le n$ to a given number ξ, see [14] for further background. In particular, the long standing Wirsing–Schmidt conjecture [45, p. 258], which was motivated by Wirsing's theorem [49], states that for any $n \in \mathbb{N}$ and any real transcendental number ξ there is a constant $C = C(\xi, n) > 0$ such that

$$|\xi - \alpha| \le C(\xi, n) H(\alpha)^{-n-1}$$

holds for infinitely many algebraic numbers α of degree $\le n$, where $H(\alpha)$ denotes the height of α (to be recalled a few lines below). The $n = 1$ case of the conjecture is a trivial consequence of the theory of continued fractions. For $n = 2$ it was proved by Davenport and Schmidt [23]. However, there are only partial results for $n > 2$. Note, however, that using Dirichlet's theorem it is easily shown that for any $\xi \in \mathbb{R}$ there exists $c_0 = c_0(\xi, n) > 0$ such that $|P(\xi)| < c_0 H(P)^{-n}$ for infinitely many $P \in \mathbb{Z}[x]$ with $\deg P \le n$.

In this section we will deal with real numbers badly approximable by algebraic numbers. Given a polynomial P with integer coefficients, $H(P)$ will denote the height of P, which, by definition, is the maximum of the absolute values of the coefficients of P. Given an algebraic number $\alpha \in \mathbb{C}$, $H(\alpha)$ will denote the (naive) height of α, which, by definition, is the height of the minimal defining polynomial P of α over $\mathbb{Z}$. It is also convenient to introduce the following three sets:

$$\mathcal{B}_n = \left\{ \xi \in \mathbb{R} : \begin{array}{l} \exists c_1 = c_1(\xi, n) > 0 \text{ such that } |P(\xi)| \ge c_1 H(P)^{-n} \\ \text{for all non-zero } P \in \mathbb{Z}[x],\ \deg P \le n \end{array} \right\},$$

$$\mathcal{W}_n^* = \left\{ \xi \in \mathbb{R} : \begin{array}{l} \exists c_2 = c_2(\xi, n) > 0 \text{ such that } |\xi - \alpha| < c_2 H(\alpha)^{-n-1} \\ \text{for infinitely many real algebraic } \alpha \text{ with } \deg \alpha \le n \end{array} \right\},$$

$$\mathcal{B}_n^* = \left\{ \xi \in \mathbb{R} : \begin{array}{l} \exists c_3 = c_3(\xi, n) > 0 \text{ such that } |\xi - \alpha| \ge c_3 H(\alpha)^{-n-1} \\ \text{for all real algebraic } \alpha \text{ with } \deg \alpha \le n \end{array} \right\}.$$

The sets $\mathcal{B}_n$ and $\mathcal{B}_n^*$ are the natural generalisations of badly approximable numbers to the context of approximation by algebraic numbers. They are known to have Lebesgue measure zero, e.g., by a Khintchine type theorem proved in [10]. Within this paper we will deal with the following two conjectures that Bugeaud formulated as Problems 24 and 25 in his Cambridge Tract [14, §10.2]:

Conjecture B1 *$\mathcal{B}_n$ contains a real transcendental number.*

Conjecture B2 *$\mathcal{W}_n^* \cap \mathcal{B}_n^*$ contains a real transcendental number.*

Note that Conjecture B1 is stronger than Conjecture B2 since we have that

$$\mathcal{B}_n \subset \mathcal{W}_n^* \cap \mathcal{B}_n^*. \tag{18}$$

The proof of (18) is rather standard. Indeed, it rests on the Mean Value Theorem and Minkowski's theorem for convex bodies, see Appendix B for details. Here we establish the following Hausdorff dimension result that easily settles the above conjectures.

Theorem 3 *For any natural number n and any interval I in $\mathbb{R}$*

$$\dim \bigcap_{k=1}^{n} \mathcal{B}_k \cap I \;=\; \dim \bigcap_{k=1}^{n} (\mathcal{W}_k^* \cap \mathcal{B}_k^* \cap I) \;=\; 1.$$

Proof Without loss of generality we will assume that $n \geq 2$. Let

$$\mathbf{f} : \mathbb{R} \to \mathbb{R}^n \quad \text{such that} \quad \mathbf{f}(x) = (x, x^2, \ldots, x^n),$$

$1 \leq k \leq n$ be an integer and $\mathbf{r}_k = \left(\frac{1}{k}, \ldots, \frac{1}{k}, 0, \ldots, 0\right) \in \mathcal{R}_n$, where the number of zeros is $n-k$. Let $\xi \in \mathbb{R}$ be such that $\mathbf{f}(\xi) \in \mathbf{Bad}(\mathbf{r}_k)$. By Property (iii) of Lemma 1, there exists $c(\xi, n, k) > 0$ such that for any $H \geq 1$ the only integer solution $(a_0, a_1, \ldots, a_n)$ to the system

$$\begin{aligned} &|a_0 + a_1 x + \cdots + a_n x^n| < c(\xi, n, k) H^{-1}, \\ &|a_i| < H^{1/k} \quad (1 \leq i \leq k), \\ &|a_i| < H^{0} \quad (k+1 \leq i \leq n) \end{aligned}$$

is $a_0 = \cdots = a_n = 0$. Hence, for any non-zero polynomial $P(x) = a_k x^k + \cdots + a_0 \in \mathbb{Z}[x]$ with $H(P) < H^{1/k}$ we must have that $|P(\xi)| \geq c(\xi, n, k) H^{-1} > c(\xi, n, k) H(P)^{-k}$. By definition, this means that $\xi \in \mathcal{B}_k$. To sum up, we have just shown that $\mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r}_k)) \subset \mathcal{B}_k$. Hence

$$\bigcap_{k=1}^{n} \mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r}_k)) \subset \bigcap_{k=1}^{n} \mathcal{B}_k \overset{(18)}{\subset} \bigcap_{k=1}^{n} \mathcal{W}_k^* \cap \mathcal{B}_k^*.$$

By Theorem 2, for any interval $I \subset \mathbb{R}$ we have that $\dim \bigcap_{k=1}^{n} \mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r}_k)) \cap I = 1$. In view of the above inclusions the statement of Theorem 3 now readily follows. □

Remark An interesting problem is to show that Theorem 3 holds when $n = \infty$.

3 Lattice points counting

The rest of the paper will be concerned with the proof of Theorem 2, which will rely heavily on efficient counting of lattice points in convex bodies. The lattices will arise upon reformulating $\mathbf{Bad}(\mathbf{r})$ in the spirit of Dani [20] and Kleinbock [30]. This will require the following notation. Given a subset Λ of $\mathbb{R}^{n+1}$, let

$$\delta(\Lambda) = \inf_{\mathbf{a} \in \Lambda \setminus \{\mathbf{0}\}} \|\mathbf{a}\|_\infty, \tag{19}$$

where $\|\mathbf{a}\|_\infty = \max\{|a_0|, \dots, |a_n|\}$ for $\mathbf{a} = (a_0, \dots, a_n)$. Given $0 < \kappa < 1$, let

$$G(\kappa; \mathbf{y}) = \begin{pmatrix} \kappa^{-1} & \kappa^{-1}\mathbf{y} \\ 0 & I_n \end{pmatrix}, \tag{20}$$

where $\mathbf{y} \in \mathbb{R}^n$ is regarded as a row and I_n is the $n \times n$ identity matrix. Finally, given $\mathbf{r} \in \mathcal{R}_n$, $b > 1$ and $t \in \mathbb{R}$, define the $(n+1) \times (n+1)$ unimodular diagonal matrix

$$g_{\mathbf{r},b}^t = \operatorname{diag}\left\{b^t, b^{-r_1 t}, \dots, b^{-r_n t}\right\}. \tag{21}$$

Lemma 2 *Let $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{r} \in \mathcal{R}_n$. Then $\mathbf{y} \in \mathbf{Bad}(\mathbf{r})$ if and only if there exists $\kappa \in (0, 1)$ and $b > 1$ such that for all $t \in \mathbb{N}$*

$$\delta\left(g_{\mathbf{r},b}^t G(\kappa; \mathbf{y}) \mathbb{Z}^{n+1}\right) \geq 1. \tag{22}$$

Proof The necessity is straightforward as all one has to do is to take $H = b^t$ and divide each inequality in (17) by its right hand side. Then, assuming that $\mathbf{y} \in \mathbf{Bad}(\mathbf{r})$, the non-existence of integer solutions to (17) would imply (22) with $\kappa = c$. The sufficiency is only slightly harder. Assume that for some κ and b inequality (22) holds for all $t \in \mathbb{N}$, while $\mathbf{y} \notin \mathbf{Bad}(\mathbf{r})$. Take $c = \kappa/b$. By definition, there is an $H > 1$ such that (17) has a non-zero integer solution $(a_0, \dots, a_n)$. Take $t = [\log H / \log b] + 1$, where $[\cdot]$ denotes

the integer part. Note that $Hb^{-t} < 1$ and $H^{-1}b^t \leq b$. Then (17) implies that $\delta(g^t_{\mathbf{r},b}G(\kappa;\mathbf{y})\mathbb{Z}^{n+1}) < 1$, contrary to (22). The proof is thus complete. □

Remark Lemma 2 can be regarded as a variation of the Dani–Kleinbock correspondence between badly approximable points in $\mathbb{R}^n$ and bounded orbits of certain lattices under the actions by the diagonal semigroup $\left\{g^t_{\mathbf{r},b} : t > 0\right\}$, where $b > 1$. It is easily seen that this semigroup is independent of the choice of $b > 1$, which is usually taken to be $e = \exp(1)$. The correspondence was first established by Dani [20] in the case $\mathbf{r} = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$ and then extended by Kleinbock [30] to the case of arbitrary positive weights and can be stated as follows. The point $\mathbf{y} \in \mathbb{R}^n$ is $\mathbf{r}$-badly approximable if and only if the orbit of the lattice $G(1;\mathbf{y})\mathbb{Z}^{n+1}$ under the action by $\left\{g^t_{\mathbf{r},e} : t > 0\right\}$ is bounded.

We proceed by recalling two classical results from the geometry of numbers. In what follows, $\mathrm{vol}_\ell(X)$ denotes the ℓ-dimensional volume of $X \subset \mathbb{R}^\ell$ and $\#X$ denotes the cardinality of X. Also $\det\Lambda$ will denote the *determinant or covolume* of a lattice Λ.

Minkowski's Convex Body Theorem (see [45, Theorem 2B]) *Let $K \subset \mathbb{R}^\ell$ be a convex body symmetric about the origin and let Λ be a lattice in $\mathbb{R}^\ell$. Suppose that $\mathrm{vol}_\ell(K) > 2^\ell \det\Lambda$. Then K contains a non-zero point of Λ.*

Theorem (Blichfeldt [12]) *Let $K \subset \mathbb{R}^\ell$ be a convex bounded body and let Λ be a lattice in $\mathbb{R}^\ell$ such that* $\mathrm{rank}\,(K \cap \Lambda) = \ell$. *Then*

$$\#(K \cap \Lambda) \;\leq\; \ell!\,\frac{\mathrm{vol}_\ell(K)}{\det\Lambda} + \ell.$$

The following lemma is a straightforward consequence of Blichfeldt's theorem.

Lemma 3 (cf. Lemma 4 in [34]) *Let K be a convex bounded body in $\mathbb{R}^\ell$ with $\mathbf{0} \in K$ and $\mathrm{vol}_\ell(K) < 1/\ell!$. Then* $\mathrm{rank}\,(K \cap \mathbb{Z}^\ell) \leq \ell - 1$.

Proof Assume the contrary, that is assume that $\mathrm{rank}\,(K \cap \mathbb{Z}^\ell) = \ell$ (note that the rank cannot be bigger than ℓ). It means that K contains at least ℓ non-zero integer points. Since $\mathbf{0} \in K$, we then have that $\#\left(K \cap \mathbb{Z}^\ell\right) \geq \ell + 1$. However, since $\det\mathbb{Z}^\ell = 1$ and $\mathrm{vol}_\ell(K) < 1/\ell!$, by Blichfeldt's theorem, we conclude that

$$\#\left(K \cap \mathbb{Z}^\ell\right) \;\leq\; \ell!\,\frac{\mathrm{vol}_\ell(K)}{\det\Lambda} + \ell < \ell!\,\frac{1/\ell!}{1} + \ell < 1 + \ell,$$

contrary to the above lower bound. □

The bodies K of interest will arise as the intersection of parallelepipeds

$$\Pi_{\boldsymbol{\theta}} = \left\{\mathbf{x} = (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} : |x_i| < \theta_i, \quad i = 0, \ldots, n\right\} \tag{23}$$

with ℓ-dimensional subspaces of $\mathbb{R}^{n+1}$, where $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_n)$ is an $(n+1)$-tuple of positive numbers. In view of this, we now obtain an estimate for the volume of the bodies that arise this way (Lemma 4 below) and then verify what Blichfeldt's theorem means for such bodies (Lemma 5 below).

Lemma 4 *Let $\ell \in \mathbb{N}$, $\ell \le n+1$, $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_n)$ with $\theta_0, \ldots, \theta_n > 0$. Then for any linear subspace V of $\mathbb{R}^{n+1}$ of dimension ℓ we have that*

$$\mathrm{vol}_\ell(\Pi_{\boldsymbol{\theta}} \cap V) \le 2^\ell (n+1)^{\ell/2} \Theta_\ell, \quad \textit{where} \quad \Theta_\ell = \max_{\substack{I \subset \{0,\ldots,n\} \\ \#I = \ell}} \prod_{i \in I} \theta_i.$$

Proof Since V is a linear subspace of $\mathbb{R}^{n+1}$ of dimension ℓ, it is given by $n+1-\ell$ linear equations. Using Gaussian elimination, we can rewrite these equations to parametrise V with a linear map $\mathbf{f} : \mathbb{R}^\ell \to \mathbb{R}^{n+1}$ of $x_{i_1}, \ldots, x_{i_\ell}$ such that

$$\mathbf{f}(x_{i_1}, \ldots, x_{i_\ell}) = (x_{i_1}, \ldots, x_{i_\ell}) M,$$

where $M = (m_{i,j})$ is an $\ell \times (n+1)$ matrix with $|m_{i,j}| \le 1$ for all i and j. Then note that $\mathrm{vol}_\ell(\Pi_{\boldsymbol{\theta}} \cap V)$ is bounded by the area of the intersection of V with the cylinder $|x_{i_j}| \le \theta_{i_j}$ for $j = 1, \ldots, \ell$. This area is equal to

$$\int_{-\theta_{i_1}}^{\theta_{i_1}} \cdots \int_{-\theta_{i_\ell}}^{\theta_{i_\ell}} \left\| \frac{\partial \mathbf{f}}{\partial x_{i_1}} \wedge \cdots \wedge \frac{\partial \mathbf{f}}{\partial x_{i_\ell}} \right\|_e dx_{i_1} \ldots dx_{i_\ell}, \tag{24}$$

where $\|\cdot\|_e$ is the Euclidean norm on $\bigwedge^\ell \left(\mathbb{R}^{n+1}\right)$. Since $|m_{i,j}| \le 1$, every coordinate of every partial derivative of $\mathbf{f}$ is bounded by 1 in absolute value. Hence $\|\partial \mathbf{f}/\partial x_{i_j}\|_e \le \sqrt{n+1}$ and the integrand in (24) is bounded above by $(\sqrt{n+1})^\ell$. This readily implies that the area given by (24) is bounded above by $2^\ell (n+1)^{\ell/2} \theta_{i_1} \ldots \theta_{i_\ell} \le 2^\ell (n+1)^{\ell/2} \Theta_\ell$, whence the result follows. □

Lemma 5 *Let $c(n) = 4^{n+1}(n+1)^{(n+1)/2}(n+1)!$ and let $\boldsymbol{\theta}$ and Θ_ℓ be as in Lemma 4. Then for any discrete subgroup Γ of $\mathbb{R}^{n+1}$ with $\ell = \mathrm{rank}\left(\Gamma \cap \Pi_{\boldsymbol{\theta}}\right) > 0$ we have that*

$$\#\left(\Gamma \cap \Pi_{\boldsymbol{\theta}}\right) \;\le\; c(n) \frac{\Theta_\ell}{\delta(\Gamma)^\ell} + n + 1. \tag{25}$$

Proof Let $V = \mathrm{span}(\Gamma \cap \Pi_{\boldsymbol{\theta}})$ and $\Lambda = V \cap \Gamma$. Clearly, $\mathrm{rank}\,(\Lambda) = \ell$ and furthermore Λ is a lattice in V. Also note that $\Gamma \cap \Pi_{\boldsymbol{\theta}} = \Lambda \cap \Pi_{\boldsymbol{\theta}}$. Since

$\Lambda \subseteq \Gamma$, we have that $\delta(\Gamma) \le \delta(\Lambda)$. Let $B(r)$ denote the open ball in V of radius r centred at the origin. Note that the length of any non-zero point in Λ is bigger than or equal to $\delta(\Lambda) \ge \delta(\Gamma)$. Hence, by Minkowski's convex bodies theorem, we must have that $\mathrm{vol}_\ell\big(B(\delta(\Gamma))\big) \le 2^\ell \det\Lambda$, whence we obtain $\det\Lambda \ge \mathrm{vol}_\ell(B(\delta(\Lambda)))2^{-\ell} \ge (\delta(\Lambda)/2)^\ell$. Now using this inequality, Blichfeldt's theorem, Lemma 4 and the fact that $\ell \le n+1$ readily gives (25).

We are now approaching the key counting result of this section. Let

$$\Pi(b,u) \stackrel{\text{def}}{=} \Pi_{\boldsymbol{\theta}} \quad \text{with} \quad \boldsymbol{\theta} = (b^u, 1, \dots, 1), \tag{26}$$

where $u > 0$, $b > 1$ and $\Pi_{\boldsymbol{\theta}}$ is given by (23). Given $\mathbf{r} \in \mathcal{R}_n$, let

$$z(\mathbf{r}) \stackrel{\text{def}}{=} \#\{\, i \,:\, r_i = 0\,\} \quad \text{and} \quad \lambda(\mathbf{r}) \stackrel{\text{def}}{=} \big(1+\tau(\mathbf{r})\big)^{-1}. \tag{27}$$

Recall that $\tau(\mathbf{r})$, $\delta(\cdot)$, $g^t_{\mathbf{r},b}$ and $\Pi(b,u)$ are given by (9), (19), (21) and (26) respectively, and $[x]$ denotes the integer part of x.

Lemma 6 *Let $b > 1$, $\mathbf{r} \in \mathcal{R}_n$, $\lambda = \lambda(\mathbf{r})$, $z = z(\mathbf{r})$, $t \in \mathbb{N}$, $u \in \mathbb{R}$, $1 \le \lambda u \le t$ and $c(n)$ be as in Lemma 5. Let $g^t = g^t_{\mathbf{r},b}$. Let Λ be a discrete subgroup of $\mathbb{R}^{n+1}$ such that* rank $\Lambda \le n - z$ *and*

$$\delta\big(g^{t-[\lambda u]}\Lambda\big) \ge 1. \tag{28}$$

Then

$$\#(g^t\Lambda) \cap \Pi(b,u) \le 2c(n)b^\tau b^{\lambda u}. \tag{29}$$

Proof Let $\mathbf{x} = (x_0, \dots, x_n) \in \Lambda$ be such that $g^t\mathbf{x} \in \Pi(b,u)$. By the definitions of $g^t = g^t_{\mathbf{r},b}$ and $\Pi(b,u)$, we have that $b^t|x_0| < b^u$ and $b^{-r_i t}|x_i| < 1$ for $i = 1, \dots, n$. Equivalently, for $s \in \mathbb{Z}$, $1 \le s \le u-1$, we have that

$$b^{t-s}|x_0| < b^{u-s} \quad \text{and} \quad b^{-r_i(t-s)}|x_i| < b^{r_i s} \quad (1 \le i \le n).$$

This can be written as $g^{t-s}\mathbf{x} \in \Pi_{\boldsymbol{\theta}}$, where $\boldsymbol{\theta} = (b^{u-s}, b^{r_1 s}, \dots, b^{r_n s})$. Therefore,

$$(g^t\Lambda) \cap \Pi(b,u) \;=\; \Gamma \cap \Pi_{\boldsymbol{\theta}}, \tag{30}$$

where $\Gamma = g^{t-s}\Lambda$. Now take $s = [\lambda u]$. Recall that $\lambda < 1$ and that, by the conditions of Lemma 6, $[\lambda u] \le t$. Then, by the left hand side of (28), we have that $\delta(\Gamma) \ge 1$. Hence, by Lemma 5 and (30), we get

$$\#(g^t\Lambda) \cap \Pi(b,u) \;=\; \#\big(\Gamma \cap \Pi_{\boldsymbol{\theta}}\big) \;\le\; c(n)\Theta_\ell + n + 1, \tag{31}$$

where $\ell = \operatorname{rank}\Gamma = \operatorname{rank}\Lambda \le n - z$. Note that all the components of $\boldsymbol{\theta}$ are ≥ 1 and exactly z of them equal 1. Then, since $\ell \le n - z$ and $s = [\lambda u]$, we get that

$$\Theta_\ell \le \frac{\theta_0 \dots \theta_n}{\min\{\theta_i : \theta_i > 1\}} = \frac{b^u}{\min\{b^{u-s}, b^{\tau s}\}} \le \max\{b^{\lambda u}, b^{u-\tau(\lambda u-1)}\} = b^\tau b^{\lambda u}.$$

Combining this estimate with (31) and the obvious fact that $n+1 < c(n)b^\tau b^{\lambda u}$ gives (29). □

4 'Dangerous' intervals

In view of Lemma 2, when proving Theorem 2 we will aim to avoid the solutions of the inequalities $\delta\left(g^t_{\mathbf{r},b}G_x\mathbb{Z}^{n+1}\right) < 1$, where $G_x = G(\kappa;\mathbf{y})$ with $\mathbf{y} = \mathbf{f}(x)$ and κ is a sufficiently small constant. For fixed $\mathbf{r}, b, t, \mathbf{f}$ and κ the above inequality is equivalent to the existence of $(a_0, \mathbf{a}) \in \mathbb{Z}^{n+1}$ with $\mathbf{a} \ne \mathbf{0}$ satisfying

$$\begin{cases} |a_0 + \mathbf{a}.\mathbf{f}(x)| < \kappa b^{-t}, \\ |a_i| < b^{r_i t} \quad (1 \le i \le n). \end{cases} \tag{32}$$

Here the dot means the usual inner product. That is $\mathbf{a}.\mathbf{b} = a_1b_1 + \dots + a_nb_n$ for any given $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$. In this section we study intervals arising from (32) that, for obvious reasons, are referred to as *dangerous* (see [45] for similar terminology). We will consider several cases that are tied up with the magnitude of $\mathbf{a}.\mathbf{f}'(x)$; *i.e.*, the derivative of $a_0 + \mathbf{a}.\mathbf{f}(x)$—see Propositions 1 and 2 below.

Throughout $\mathcal{F}_n(I)$ and x_0 are as in Theorem 2. First we discuss some conditions that arise from the nondegeneracy assumption on maps in $\mathcal{F}_n(I)$. Let $\mathbf{f} = (f_1, \dots, f_n) \in \mathcal{F}_n(I)$. Since $\mathbf{f}$ is nondegenerate at $x_0 \in I$, there is a sufficiently small neighborhood $I_\mathbf{f}$ of x_0 such that the Wronskian of $f_1', \dots, f_n'$, which, by definition, is the determinant $\det\left(f_j^{(i)}\right)_{1\le i,j\le n}$, is non zero everywhere in $I_\mathbf{f}$. Then every coordinate function f_j is non-vanishing at all but countably many points of $I_\mathbf{f} \subset I$—see, e.g., [5, Lemma 3]. Since $\mathbf{f} \in C^n$ and $\mathcal{F}_n(I)$ is finite, we can choose a compact interval $I_0 \subset \bigcap_{\mathbf{f}\in\mathcal{F}_n(I)} I_\mathbf{f} \subset I$ satisfying

Property F There are constants $0 < c_0 < 1 < c_1$ such that for every map $\mathbf{f} = (f_1, \dots, f_n) \in \mathcal{F}_n(I)$, for all $x \in I_0$, $1 \le i \le n$ and $0 \le j \le n$ one has that

$$\left|\det\left(f_j^{(i)}(x)\right)_{1\le i,j\le n}\right| > c_0, \quad |f_j'(x)| > c_0 \quad \text{and} \quad |f_j^{(i)}(x)| < c_1. \tag{33}$$

Next, we prove two auxiliary lemmas that are well known in a related context.

Lemma 7 (cf. Lemma 5 in [5]) *Let $I_0 \subset I$ be a compact interval satisfying Property F. Let $2c_2 = c_0 c_1^{-n+1} n!^{-1}$, where c_0 and c_1 arise from (33). Then for any $\mathbf{f} \in \mathcal{F}_n(I)$, any $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{Z}^n \setminus \{0\}$ and any $x \in I_0$ there exists $i \in \{1, \dots, n\}$ such that $|\mathbf{a}.\mathbf{f}^{(i)}(x)| \geq 2c_2 \max_{1 \leq j \leq n} |a_j|$.*

Proof Solving the system $a_1 f_1^{(i)}(x) + \dots + a_n f_n^{(i)}(x) = \mathbf{a}.\mathbf{f}^{(i)}(x)$, where $1 \leq i \leq n$, by Cramer's rule with respect to a_i and using (33) to estimate the determinants involved in the rule we obtain

$$|a_j| \ \leq \ c_1^{n-1} \cdot n! c_0^{-1} \max_{1 \leq i \leq n} \left| \mathbf{a}.\mathbf{f}^{(i)}(x) \right|$$

for each $j = 1, \dots, n$, whence the statement of lemma readily follows. □

Lemma 8 (cf. Lemma 6 in [5]) *Let $I_0 \subset I$ and c_2 be as in Lemma 7. Then there is $\delta_0 > 0$ such that for any interval $J \subset I_0$ of length $|J| \leq \delta_0$, any $\mathbf{f} \in \mathcal{F}_n(I)$ and $\mathbf{a} = (a_1, \cdots, a_n) \in \mathbb{Z}^n \setminus \{0\}$, there is an $i \in \{1, \dots, n\}$ satisfying*

$$\inf_{x \in J} \left| \mathbf{a}.\mathbf{f}^{(i)}(x) \right| \geq c_2 \max_{1 \leq j \leq n} |a_j|. \tag{34}$$

Proof Since I_0 is compact, for each $\mathbf{f} \in \mathcal{F}_n(I)$ and $1 \leq i \leq n$, the map $\mathbf{f}^{(i)}$ is uniformly continuous on I_0. Hence, there is a $\delta_{i,\mathbf{f}} > 0$ such that for any $x, y \in I_0$ with $|x - y| \leq \delta_{i,\mathbf{f}}$ we have $\left| \mathbf{f}^{(i)}(x) - \mathbf{f}^{(i)}(y) \right| < c_2/n$. Let $J \subset I_0$ be an interval of length $|J| \leq \delta_{i,\mathbf{f}}$ and $x, y \in J$. By Lemma 7, there is $i \in \{1, \dots, n\}$ such that $|\mathbf{a}.\mathbf{f}^{(i)}(x)| \geq 2c_2 h$, where $h = \max_{1 \leq j \leq n} |a_j|$. Then

$$|\mathbf{a}.\mathbf{f}^{(i)}(y)| \geq |\mathbf{a}.\mathbf{f}^{(i)}(x)| - |\mathbf{a}.\mathbf{f}^{(i)}(y) - \mathbf{a}.\mathbf{f}^{(i)}(x)| \geq 2c_2 h - nhc_2/n = c_2 h. \tag{35}$$

Since $\mathcal{F}_n(I)$ is finite, $\delta_0 = \inf_{i,\mathbf{f}} \delta_{i,\mathbf{f}} > 0$. Hence (35) implies (34) provided that $|J| \leq \delta_0$. □

Proposition 1 *Let $I_0 \subset I$ be a compact interval satisfying Property F and $\mathbf{f} \in \mathcal{F}_n(I)$. Further, let δ_0 be as in Lemma 8, $\mathbf{r} \in \mathcal{R}_n$ and*

$$\gamma = \gamma(\mathbf{r}) \stackrel{\text{def}}{=} \max\{r_1, \dots, r_n\}. \tag{36}$$

Finally, let $t \in \mathbb{N}$, $\ell \in \mathbb{Z}_{\geq 0}$, $b > 1$, $\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$, $a_0 \in \mathbb{Z}$, $0 < \kappa < 1$ and

$$D^1_{t,\ell,\mathbf{r},b,\kappa,\mathbf{f}}(a_0, \mathbf{a}) = \left\{ x \in I_0 \ : \ \begin{array}{r} |a_0 + \mathbf{a}.\mathbf{f}(x)| < \kappa b^{-t} \\ b^{\gamma t - (1+\gamma)\ell} \leq |\mathbf{a}.\mathbf{f}'(x)| < b^{\gamma t - (1+\gamma)(\ell-1)} \\ |a_i| < b^{r_i t} \end{array} \right\}.$$

Then, there is a constant $c_3 > 0$ depending on n, $|I_0|$, c_1, c_2 and δ_0 only such that the set $D^1_{t,\ell,\mathbf{r},b,\kappa,\mathbf{f}}(a_0, \mathbf{a})$ can be covered by a collection $\mathcal{D}^1_{t,\ell,\mathbf{r},b,\kappa,\mathbf{f}}(a_0, \mathbf{a})$ of at most c_3 intervals Δ of length $|\Delta| \le \kappa b^{-(1+\gamma)(t-\ell)}$.

Proof We will abbreviate $D^1_{t,\ell,\mathbf{r},b,\kappa,\mathbf{f}}(a_0, \mathbf{a})$ as D^1 and naturally assume that $D^1 \neq \emptyset$ as otherwise there is nothing to prove. Since I_0 can be covered by at most $\left[\delta_0^{-1}|I_0|\right] + 1$ intervals J of length $|J| \le \delta_0$, it suffices to prove the proposition under the assumption that $|I_0| \le \delta_0$. Let $f(x) = a_0 + \mathbf{a}.\mathbf{f}(x)$. Then, by Lemma 8, we have that $|f^{(i)}(x)| > 0$ for a fixed $i \in \{1, \cdots, n\}$ and all $x \in I_0$. First consider the case $i > 1$. Then, using Rolle's theorem, one finds that the function $f^{(j)}(x)$ vanishes on I_0 at $\le i - j$ points $(0 \le j \le i-1)$. Assuming that $I_0 = [a, b]$, let $x_0 = a < x_1 < \cdots < x_{s-1} < x_s = b$ be the collection consisting of the points a and b and all the zeros of $\prod_{j=0}^{i-1} f^{(j)}(x)$. Then, as we have just seen $s \le 1 + \sum_{j=0}^{i-1}(i - j) = i(i + 1)/2 + 1$. By the choice of the points x_i, we have that for $1 \le q \le s$ and $0 \le j \le i - 1$ the function $f^{(j)}(x)$ is monotonic and does not change sign on the interval $[x_{q-1}, x_q]$. Therefore, in view of the definition of D^1 we must have that $\Delta_q = D^1 \cap [x_{q-1}, x_q]$ is an interval. Hence, $D^1 = \bigcup_{q=1}^{s} \Delta_q$, a union of at most $(i + 1)i/2 + 1 \le (n + 1)n/2 + 1$ intervals.

It remains to estimate the length of each Δ_q. To this end, take any $x_1, x_2 \in \Delta_q$. By the construction of Δ_q, the numbers $f(x_1)$ and $f(x_2)$ have the same sign and satisfy the inequality $|f(x_i)| < \kappa b^{-t}$. Hence, $|f(x_1) - f(x_2)| < \kappa b^{-t}$. By the Mean Value Theorem, $|f(x_1) - f(x_2)| = |f'(\theta)(x_1 - x_2)|$. Hence $|x_1 - x_2| \le \kappa b^{-t}/|f'(\theta)|$. Since $\Delta_q \subset D^1$ is an interval, $\theta \in D^1$. Hence, $|f'(\theta)| \ge b^{\gamma t-(1+\gamma)\ell}$ and we obtain that $|x_1 - x_2| \le \kappa b^{-t} b^{-\gamma t+(1+\gamma)\ell} = \kappa b^{-(1+\gamma)(t-\ell)}$. This estimate together with the obvious equality $|\Delta_q| = \sup_{x_1,x_2\in\Delta_q} |x_1 - x_2|$ implies that $|\Delta_q| \le \kappa b^{-(1+\gamma)(t-\ell)}$. Thus, if $i > 1$, the set D^1 can be covered by at most $n(n + 1)/2 + 1$ intervals of length $\kappa b^{-(1+\gamma)(t-\ell)}$.

Now consider the case $i = 1$. Recall that $f(x) = a_0 + \mathbf{a}.\mathbf{f}(x)$. Then, by the definition of D^1 and (33), for $x \in D^1$ we get

$$b^{\gamma t-(1+\gamma)\ell} \le |f'(x)| = |\mathbf{a}.\mathbf{f}'(x)| \le c_1 n \max_{1\le j\le n} |a_j|. \tag{37}$$

Further, $(34)_{i=1}$ implies that $\inf_{x\in I_0} |f'(x)| \ge c_2 \max_{1\le j\le n} |a_j|$. Therefore, f is monotonic on I_0 and D^1 is covered by a single interval Δ defined by the inequality $|f(x)| < \kappa b^{-t}$. Arguing as above and using (37) we get

$$|\Delta| \le \frac{2\kappa b^{-t}}{\inf_{x\in I_0} |f'(x)|} \le \frac{2\kappa b^{-t}}{c_2 \max_{1\le j\le n} |a_j|}$$

$$\le \frac{2c_1 n\kappa b^{-t}}{c_2 b^{\gamma t-(1+\gamma)\ell}} = \frac{2c_1 n}{c_2} \times \kappa b^{-(1+\gamma)(t-\ell)}.$$

Thus, by splitting Δ into smaller intervals if necessary, D^1 can be covered by at most $\left[\frac{2c_1 n}{c_2}\right] + 1$ intervals of length $\kappa b^{-(1+\gamma)(t-\ell)}$. □

Proposition 2 *Let $I_0 \subset I$ be a compact interval satisfying Property F and $\gamma = \gamma(\mathbf{r})$ be given by (36). Then there are constants $K_0 > 0$ and $0 < \kappa_0 < 1$ such that for any $\mathbf{f} \in \mathcal{F}_n(I)$, any $\mathbf{r} \in \mathcal{R}_n$, $t \in \mathbb{N}$, $0 \le \varepsilon < \gamma$, $b > 1$ and $0 < \kappa < \kappa_0$ the set*

$$D^2_{t,\varepsilon,\mathbf{r},b,\kappa,\mathbf{f}} = \left\{ x \in I_0 : \exists\, \mathbf{a} \in \mathbb{Z}^n \setminus \{0\} \text{ and } a_0 \in \mathbb{Z} \text{ such that } \begin{array}{l} |a_0 + \mathbf{a}.\mathbf{f}(x)| < \kappa\, b^{-t} \\ |\mathbf{a}.\mathbf{f}'(x)| < nc_1 b^{(\gamma-\varepsilon)t} \\ |a_i| < b^{r_i t} \end{array} \right\}$$

can be covered by a collection $\mathcal{D}^2_{t,\varepsilon,\mathbf{r},b,\kappa,\mathbf{f}}$ of intervals such that

$$|\Delta| \le \delta_t \quad \text{for all} \quad \Delta \in \mathcal{D}^2_{t,\varepsilon,\mathbf{r},b,\kappa,\mathbf{f}} \tag{38}$$

and

$$\#\mathcal{D}^2_{t,\varepsilon,\mathbf{r},b,\kappa,\mathbf{f}} \le \frac{K_0\,(\kappa\, b^{-\varepsilon t})^\alpha}{\delta_t}, \tag{39}$$

where $\delta_t = \kappa\, b^{-t(1+\gamma-\varepsilon)}$ and $\alpha = \frac{1}{(n+1)(2n-1)}$.

Proposition 2 will be derived from a theorem due to Bernik, Kleinbock and Margulis using the ideas of [6]. In what follows $|X|$ denotes the Lebesgue measure of a set $X \subset \mathbb{R}$. The following is a simplified version of Theorem 1.4 from [11] that refines the results of [33].

Theorem 4 (Theorem 1.4 in [11]) *Let $I \subset \mathbb{R}$ be an open interval, $x_0 \in I$ and $\mathbf{f} : I \to \mathbb{R}^n$ be nondegenerate at x_0. Then there is an open interval $J \subset I$ centred at x_0 and $E_J > 0$ such that for any real $\omega, K, T_1, \dots, T_n$ satisfying*

$$0 < \omega \le 1, \quad T_1, \dots, T_n \ge 1, \quad K > 0 \quad \text{and} \quad \omega K T_1 \cdots T_n \le \max_i T_i$$

the set

$$S(\omega, K, T_1, \dots, T_n) \stackrel{\text{def}}{=} \left\{ x \in I : \exists \mathbf{a} \in \mathbb{Z}^n \setminus \{0\} \quad \begin{array}{l} \|\mathbf{a}.\mathbf{f}(x)\| < \omega \\ |\mathbf{a}.\mathbf{f}'(x)| < K \\ |a_i| < T_i \quad (1 \le i \le n) \end{array} \right\}$$

satisfies

$$|S(\omega, K, T_1, \dots, T_n) \cap J| \;\le\; E_J \cdot \max\left(\omega, \left(\frac{\omega K T_1 \dots T_n}{\max_i T_i}\right)^{\frac{1}{n+1}}\right)^{\frac{1}{2n-1}}. \quad (40)$$

We will also use the following elementary consequence of Taylor's formula.

Lemma 9 *Let $f : J \to \mathbb{R}$ be a C^2 function on an interval J. Let $\omega, K > 0$ and $y \in J$ be such that $|f''(x)| < K^2/\omega$ for all $x \in J$ and*

$$|f(y)| < \omega/2 \quad \text{and} \quad |f'(y)| < K/2\,. \quad (41)$$

Then $|f(x)| < \omega$ and $|f'(x)| < K$ for all $x \in J$ with $|x - y| < \omega/2K$.

Proof of Proposition 2 Fix any $\mathbf{f} \in \mathcal{F}_n(I)$. We will abbreviate $D^2_{t,\varepsilon,\mathbf{r},b,\kappa,\mathbf{f}}$ as D^2 and naturally assume that it is non-empty as otherwise there is nothing to prove. By (33), $\mathbf{f}$ is nondegenerate at any $x \in I_0$ and therefore Theorem 4 is applicable. Let $J = J(x)$ be the interval centred at x that arises from Theorem 4. Since I_0 is compact there is a finite cover of I_0 by intervals $J(x_1), \dots, J(x_s)$, where $s = s_{\mathbf{f}}$ depends on $\mathbf{f}$. Let $0 < \kappa_0 < 1$ and $\kappa_0 \le \min_{1\le i\le s} |I_0 \cap J(x_i)|$. The existence of κ_0 is obvious because $|I_0 \cap J(x)| > 0$ for each $x \in I_0$.

Let $0 < \kappa < \kappa_0$, $\mathbf{r} \in \mathcal{R}_n$, $t \in \mathbb{N}$, $0 \le \varepsilon < \gamma$, $b > 1$ and let

$$\omega = 2\kappa\, b^{-t}, \qquad K = 2nc_1 b^{(\gamma-\varepsilon)t} \quad \text{and} \quad T_i = b^{r_i t}\ (1 \le i \le n). \quad (42)$$

Note that since $\varepsilon < \gamma$ and $c_1 > 1$ we have that $K > 2$. Also note that $\omega < 2\kappa$. For each $i \in \{1, \dots, s\}$ define the interval $J_i = (a_i + \omega/2K, b_i - \omega/2K)$, where $[a_i, b_i]$ is the intersection of I_0 and the closure of $J(x_i)$. Since $\kappa < \kappa_0 \le |I_0 \cap J(x_i)|$, $\omega < 2\kappa$ and $K > 2$, we have that $J_i \ne \emptyset$ for each i. Let

$$\tilde{D}^2 = \bigcup_{1\le i\le s} \bigcup_{y\in D^2\cap J_i} \big(y - \omega/2K,\, y + \omega/2K\big). \quad (43)$$

Our goal now is to use Lemma 9 with $f(x) = a_0 + \mathbf{a}.\mathbf{f}(x)$ in order to show that

$$\tilde{D}^2 \subset \bigcup_{1\le i\le s} S(\omega, K, T_1, \dots, T_n) \cap J(x_i). \quad (44)$$

In view of the definitions of D^2 and $S(\omega, K, T_1, \dots, T_n)$ and the choice of parameters (42), inequalities (41) hold for every $y \in D^2$. Further, by (33), the inequalities $|a_i| < b^{r_i t}$ and the fact that $r_i \le \gamma$ for all i implied by (36), we get that

$$|f''(x)| \le nc_1 \max_{1\le j\le n} |a_j| \le nc_1 \max_{1\le j\le n} b^{r_i t} \le nc_1 b^{\gamma t}. \quad (45)$$

Next, $K^2/\omega = \frac{1}{2}n^2c_1^2\kappa^{-1}b^{2(\gamma-\varepsilon)t}b^t > nc_1b^{\gamma t}$ because $\varepsilon \le \gamma \le 1$, $c_1 > 1$ and $\kappa < 1$. Therefore, by (45), we have that $|f''(x)| \le K^2/\omega$ for all $x \in I_0$. Thus, Lemma 9 is applicable and for $1 \le i \le s$ we have that $\{x : |x - y| < \omega/2K\} \subset S(\omega, K, T_1, \dots, T_n) \cap J(x_i)$ each $y \in D^2 \cap J_i$. This proves (44).

Next, by Theorem 4, condition $r_1 + \cdots + r_n = 1$ and (44) we conclude that

$$|\tilde{D}^2| \le E_{\mathbf{f}} \cdot \left(4nc_1\kappa b^{-\varepsilon t}\right)^{\alpha}, \tag{46}$$

where $E_{\mathbf{f}} = s \max_{1\le i\le s} E_{J(x_i)}$. By (43), $\tilde{D}^2$ can be written as a union of disjoint intervals of length $\ge \omega/K = (nc_1)^{-1}\kappa\, b^{-t(1+\gamma-\varepsilon)} = (nc_1)^{-1}\delta_t$. By splitting some of these intervals if necessary, we get a collection $\tilde{\mathcal{D}}^2$ of disjoint intervals Δ such that $\frac{1}{c_1 n}\delta_t \le |\Delta| \le \delta_t$. Let

$$K_0 = \max_{\mathbf{f}\in\mathcal{F}_n(I)} \max\{4s_{\mathbf{f}}, (4nc_1)^{1+\alpha}E_{\mathbf{f}}\}.$$

Then, by (46) and the above inequality, we get

$$\#\tilde{\mathcal{D}}^2 \le \frac{E_{\mathbf{f}} \cdot \left(4nc_1\kappa b^{-\varepsilon t}\right)^{\alpha}}{\frac{1}{c_1 n}\delta_t} \le \frac{K_0\,(\kappa\, b^{-\varepsilon t})^{\alpha}}{2\delta_t}. \tag{47}$$

Let $\mathcal{D}^2$ be the collection of all the intervals in $\tilde{\mathcal{D}}^2$ together with the $2s$ intervals $[a_i, a_i + \omega/2K]$ and $[b_i - \omega/2K, b_i]$ $(1 \le i \le s)$. It is easily seen that $2s$ is less than or equal to the right hand side of (47). Then, by (47) and the definition of $\mathcal{D}^2$, we get (38) and (39). Also, by construction, we see that $\mathcal{D}^2$ is a cover of D^2. The proof is thus complete. □

5 A Cantor sets framework

Let $R \ge 2$ be an integer. Given a collection $\mathcal{I}$ of compact intervals in $\mathbb{R}$, let $\frac{1}{R}\mathcal{I}$ denote the collection of intervals obtained by dividing each interval in $\mathcal{I}$ into R equal closed subintervals. For example, for $R = 3$ and $\mathcal{I} = \{[0, 1]\}$ we have that $\frac{1}{R}\mathcal{I} = \left\{[0, \frac{1}{3}], [\frac{1}{3}, \frac{2}{3}], [\frac{2}{3}, 1]\right\}$. Let $I_0 \subset \mathbb{R}$ be a compact interval. The sequence $(\mathcal{I}_q)_{q\ge 0}$ will be called *an R-sequence in I_0* if

$$\mathcal{I}_0 = \{I_0\} \quad \text{and} \quad \mathcal{I}_q \subset \tfrac{1}{R}\mathcal{I}_{q-1} \quad \text{for} \quad q \ge 1. \tag{48}$$

The intervals lying in $\mathcal{I}_q$ will be called to be of *level q*. Thus, the intervals of level q are obtained from intervals of level $q - 1$ by, firstly, splitting the intervals of $\mathcal{I}_{q-1}$ into R equal parts to form $\frac{1}{R}\mathcal{I}_{q-1}$, and, secondly, removing some of the intervals from $\frac{1}{R}\mathcal{I}_{q-1}$ to form $\mathcal{I}_q$. Given $q \in \mathbb{N}$, the intervals that are being removed in this procedure will be denoted by

$$\widehat{\mathcal{I}}_q \stackrel{\text{def}}{=} \left(\tfrac{1}{R}\mathcal{I}_{q-1}\right)\setminus\mathcal{I}_q.$$

Naturally, I_q will denote any interval from the collection $\mathcal{I}_q$, that is any interval of level q. Observe that

$$|I_q| = R^{-q}|I_0| \quad \text{for} \quad q \geq 0. \tag{49}$$

By definition, given $I_q \in \mathcal{I}_q$ with $q \geq 1$, there is a unique interval $I_{q-1} \in \mathcal{I}_{q-1}$ such that $I_q \subset I_{q-1}$; this interval I_{q-1} will be called the *precursor* of I_q. Obviously it is independent of the choice of the R-sequence $(\mathcal{I}_q)_{q\geq 0}$ with $I_q \in \mathcal{I}_q$.

We also define the limit set of $(\mathcal{I}_q)_{q\geq 0}$ as

$$\mathcal{K}((\mathcal{I}_q)_{q\geq 0}) \stackrel{\text{def}}{=} \bigcap_{q\geq 0} \bigcup_{I_q\in\mathcal{I}_q} I_q. \tag{50}$$

This is *a Cantor type set*. The classical middle third Cantor set can be constructed this way in an obvious manner with $R = 3$ and $I_0 = [0, 1]$. Theorem 2 will be proved by finding suitable Cantor type sets $\mathcal{K}((\mathcal{I}_q)_{q\geq 0})$. The construction of the corresponding R-sequences will be based on removing the intervals that intersect dangerous intervals—see Sect. 4.

Note that if $\mathcal{I}_q \neq \emptyset$ for all q so that $(\mathcal{I}_q)_{q\geq 0}$ is genuinely an infinite sequence, then $\mathcal{K}((\mathcal{I}_q)_{q\geq 0}) \neq \emptyset$. However, ensuring that $\mathcal{K}((\mathcal{I}_q)_{q\geq 0})$ is large requires better understanding of the sets $\mathcal{I}_q$. There are various techniques in fractal geometry that are geared towards this task—see [24]. We shall use a recent powerful result of Badziahin and Velani [16] restated below using our notation. Naturally, if we expect that the Cantor set $\mathcal{K}((\mathcal{I}_q)_{q\geq 0})$ is large, then the number of removed intervals at level q, that is the cardinality of $\widehat{\mathcal{I}}_q$, should be relatively small. In what follows, given $q \in \mathbb{N}$ and an interval J, let

$$\widehat{\mathcal{I}}_q \sqcap J \stackrel{\text{def}}{=} \{I_q \in \widehat{\mathcal{I}}_q : I_q \subset J\}.$$

This denotes the subcollection of removed intervals (when going from level $q-1$ to level q) that lie over a given interval J. The key characteristic that is 'assessing' the proportion of removed intervals at a particular level is given by

$$d_q(\mathcal{I}_q) = \min_{\{\widehat{\mathcal{I}}_{q,p}\}} \sum_{p=0}^{q-1} \left(\frac{4}{R}\right)^{q-p} \max_{I_p\in\mathcal{I}_p} \#\left(\widehat{\mathcal{I}}_{q,p} \sqcap I_p\right), \tag{51}$$

where the minimum is taken over all partitions $\{\widehat{\mathcal{I}}_{q,p}\}_{p=0}^{q-1}$ of $\widehat{\mathcal{I}}_q$, that is $\widehat{\mathcal{I}}_q = \bigcup_{p=0}^{q-1} \widehat{\mathcal{I}}_{q,p}$. Also define the corresponding global characteristic as

$$d\left((\mathcal{I}_q)_{q\ge 0}\right) = \sup_{q>0} d_q(\mathcal{I}_q).$$

The goal is to ensure that $d((\mathcal{I}_q)_{q\ge 0})$ is small. Then as we shall shortly see the corresponding Cantor set is large. Note that when estimating $d_q(\mathcal{I}_q)$ the key is to arrange the removed intervals into a partition $\bigcup_{p=0}^{q-1} \widehat{\mathcal{I}}_{q,p}$ which makes the sum on the right of (51) small.

Theorem 5 (Theorem 4 in [16]) *Let $R \ge 4$ be an integer, I_0 be a compact interval in $\mathbb{R}$ and $(\mathcal{I}_q)_{q\ge 0}$ be an R-sequence in I_0. If $d((\mathcal{I}_q)_{q\ge 0}) \le 1$ then*

$$\dim \mathcal{K}((\mathcal{I}_q)_{q\ge 0}) \ \ge \ \left(1 - \frac{\log 2}{\log R}\right). \tag{52}$$

In order to facilitate the comparison of Theorem 5 to [16, Theorem 4] we summarise the correspondence between the notation and objects used in this paper and in [16]:

Our notation/object	Corresponding notation/object in [16]
q	$n+1$
R	R_n (allowed to vary with n)
$\frac{1}{R}\mathcal{I}_{q-1}$	$\mathcal{I}_{n+1}$
$\mathcal{I}_q$	$\mathcal{J}_{n+1}$
p	$n-k$ $(0 \le k \le n)$
$\max_{I_p \in \mathcal{I}_p} \#\big(\widehat{\mathcal{I}}_{q,p} \sqcap I_p\big)$	$r_{n-k,n}$

Given the above correspondence table, it is readily verified that our condition $d((\mathcal{I}_q)_{q\ge 0}) \le 1$ corresponds to condition (16) within [16, Theorem 4]. Hence Theorem 5 above is an immediate consequence of Theorem 4 from [16].

Let $M > 1$, $X \subset \mathbb{R}$ and I_0 be a compact interval. We will say that X is *M-Cantor rich in I_0* if for any $\varepsilon > 0$ and any integer $R \ge M$ there exists an R-sequence $(\mathcal{I}_q)_{q\ge 0}$ in I_0 such that $\mathcal{K}((\mathcal{I}_q)_{q\ge 0}) \subset X$ and $d((\mathcal{I}_q)_{q\ge 0}) \le \varepsilon$. We will say that X is *Cantor rich in I_0* if it is *M-Cantor rich in I_0* for some M. We will say that X is *Cantor rich* if it is *Cantor rich in I_0* for some compact interval I_0. The following statement readily follows from Theorem 5 and our definitions.

Theorem 6 *Any Cantor rich set X satisfies* $\dim X = 1$.

We now proceed with a discussion of the intersections of Cantor rich sets. To some extent this already appears in [16, Theorem 5] and in [13]. First we prove the following auxiliary statement.

Lemma 10 *Let* $\left(\mathcal{I}_q^j\right)_{q\geq 0}$ *be a family of* R*-sequences in* I_0 *indexed by* j*. Given* $q \in \mathbb{Z}_{\geq 0}$*, let* $\mathcal{J}_q = \bigcap_j \mathcal{I}_q^j$*. Then* $(\mathcal{J}_q)_{q\geq 0}$ *is an* R*-sequence in* I_0 *such that*

$$\widehat{\mathcal{J}}_q \subset \bigcup_j \widehat{\mathcal{I}}_q^j \quad \text{for all} \quad q \geq 0 \tag{53}$$

and

$$\mathcal{K}\left((\mathcal{J}_q)_{q\geq 0}\right) \subset \bigcap_j \mathcal{K}\left(\left(\mathcal{I}_q^j\right)_{q\geq 0}\right). \tag{54}$$

Proof The validity of (48) for $(\mathcal{J}_q)_{q\geq 0}$ follows from the uniqueness of the precursor of an interval in any R-sequence from that sequence and the fact that $\mathcal{I}_0^j = \{I_0\}$ for all j, which means that $\mathcal{J}_0 = \bigcap_j \mathcal{I}_0^j = \{I_0\}$. Thus, $(\mathcal{J}_q)_{q\geq 0}$ is truly an R-sequence. The inclusion (53) is obvious for $q = 0$ for both sides of the inclusion are empty sets in this case. To see (53) for $q > 0$, observe that $\mathcal{J}_{q-1} \subset \mathcal{I}_{q-1}^j$ and this implies that $\frac{1}{R}\mathcal{J}_{q-1} \subset \frac{1}{R}\mathcal{I}_{q-1}^j$ for each j. Then we have

$$\begin{aligned}\widehat{\mathcal{J}}_q \;=\; \tfrac{1}{R}\mathcal{J}_{q-1}\setminus\mathcal{J}_q = \tfrac{1}{R}\mathcal{J}_{q-1}\setminus\textstyle\bigcap_j \mathcal{I}_q^j \;=\; \bigcup_j \left(\tfrac{1}{R}\mathcal{J}_{q-1}\setminus\mathcal{I}_q^j\right)\\ \subset \bigcup_j \left(\tfrac{1}{R}\mathcal{I}_{q-1}^j\setminus\mathcal{I}_q^j\right) \;=\; \bigcup_j \widehat{\mathcal{I}}_q^j.\end{aligned}$$

Finally, by the inclusion $\mathcal{J}_q \subset \mathcal{I}_q^j$, we have that $\bigcup J_q \subset \bigcup I_q^j$ for each pair of j and q, where the union is taken over $J_q \in \mathcal{J}_q$ and $I_q^j \in \mathcal{I}_q^j$ respectively. Hence, by (50), we have that $\mathcal{K}((\mathcal{J}_q)_{q\geq 0}) \subset \mathcal{K}((\mathcal{I}_q^j)_{q\geq 0})$ for all j, whence (54) now follows. □

Theorem 7 *Let* I_0 *be a compact interval. Then any countable intersection of* M*-Cantor rich sets in* I_0 *is* M*-Cantor rich in* I_0*. In particular, any finite intersection of Cantor rich sets in* I_0 *is Cantor rich in* I_0*.*

Proof Let $\{X_j\}_{j\in\mathbb{N}}$ be a collection of M-Cantor rich sets in I_0. Let $\varepsilon > 0$. Then, by definition, for each $j \in \mathbb{N}$ and $R \geq M$ there is an R-sequence $(\mathcal{I}_q^j)_{q\geq 0}$ in I_0 such that $\mathcal{K}((\mathcal{I}_q^j)_{q\geq 0}) \subset X_j$ and $d_q(\mathcal{I}_q^j) \leq \varepsilon 2^{-j}$ for all $q > 0$. By (51), for each j and $q > 0$ there exists a partition $\{\widehat{\mathcal{I}}_{q,p}^j\}_{p=0}^{q-1}$ of $\widehat{\mathcal{I}}_q^j$ such that

$$\sum_{p=0}^{q-1} \left(\frac{4}{R}\right)^{q-p} \max_{I_p\in\mathcal{I}_p^j} \#\big(\widehat{\mathcal{I}}_{q,p}^j \sqcap I_p\big) \;\leq \varepsilon 2^{-j}. \tag{55}$$

For $q \in \mathbb{Z}_{\geq 0}$ define $\mathcal{J}_q = \bigcap_{j\in\mathbb{N}} \mathcal{I}_q^j$ and $\widehat{\mathcal{J}}_{q,p} = \widehat{\mathcal{J}}_q \cap \bigcup_{j\in\mathbb{N}} \widehat{\mathcal{I}}_{q,p}^j$. Since $\widehat{\mathcal{I}}_q^j = \bigcup_{p=0}^{q-1} \widehat{\mathcal{I}}_{q,p}^j$ for each j, by (53), we have that $\widehat{\mathcal{J}}_q = \bigcup_{p=0}^{q-1} \widehat{\mathcal{J}}_{q,p}$, where $q > 0$. Then, for each $q > 0$ we get that

$$\sum_{p=0}^{q-1}\left(\frac{4}{R}\right)^{q-p}\max_{J_p\in\mathcal{J}_p}\#\left(\widehat{\mathcal{J}}_{q,p}\sqcap J_p\right)\le\sum_{j=1}^{\infty}\sum_{p=0}^{q-1}\left(\frac{4}{R}\right)^{q-p}\max_{I_p\in\mathcal{I}_p^j}\#\left(\widehat{\mathcal{I}}_{q,p}^{\,j}\sqcap I_p\right).$$

This inequality together with (55) and the definition of $d((\mathcal{J}_q)_{q\ge0})$ implies that $d((\mathcal{J}_q)_{q\ge0})\le\varepsilon$. By (54) and the fact that $\mathcal{K}\left(\left(\mathcal{I}_q^j\right)_{q\ge0}\right)\subset X_j$ for each j, we have that $\mathcal{K}((\mathcal{J}_q)_{q\ge0})\subset\bigcap_j X_j$. Thus the intersection $\bigcap_j X_j$ meets the definition of M-Cantor rich sets and the proof is complete. □

The winning sets in the sense of Schmidt have been used a lot to investigate various sets of badly approximable points. Hence we suggest the following

Problem 3 Verify if an α-winning set in $\mathbb{R}$ as defined by Schmidt [45] is M-Cantor rich for some M and, if this so, find an explicit relation between M and α.

6 Proof of Theorem 2

The following proposition is a key step to establishing Theorem 2. We will use the Vinogradov symbol $\ll$ to simplify the calculations. The expression $X\ll Y$ will mean that $X\le CY$ for some $C>0$, which only depends on n, the family of maps $\mathcal{F}_n(I)$ from Theorem 2 and the interval I_0 occurring in Property F.

Proposition 3 *Let $\mathcal{F}_n(I)$ be as in Theorem 2, $I_0\subset I$ be a compact interval satisfying Property F, c_0, c_1 be the same as in (33), $\sigma=1-(2n)^{-4}$ and κ_0 be as in Proposition 2. Further, let*

$$\varrho_1=nc_1|I_0|+1\quad\text{and}\quad\varrho_0=\varrho_1|I_0|+1 \tag{56}$$

and let

$$R_0=\max\left\{\varrho_0,\ nc_1,\ \frac{2^{n+1}\varrho_0\varrho_1(n+1)!}{c_0}\right\} \tag{57}$$

and

$$m_0=\max\left\{4,\ \frac{-\log\kappa_0}{\log R_0}+1\right\}. \tag{58}$$

Then for any $\mathbf{f}\in\mathcal{F}_n(I)$, $\mathbf{r}\in\mathcal{R}_n$ and any integers $m\ge m_0$ and $R\ge R_0$, there exists an R-sequence $(\mathcal{I}_q)_{q\ge0}$ in I_0 such that

(i) *for any $t\in\mathbb{N}$ and any $I_{t+m}\in\mathcal{I}_{t+m}$ we have that*

$$\delta\left(g^tG_x\mathbb{Z}^{n+1}\right)\ge1\quad\text{for all}\quad x\in I_{t+m}; \tag{59}$$

where $g^t = g^t_{\mathbf{r},b}$ *is given by* (21) *with* $b^{1+\gamma} = R$, $\gamma = \gamma(\mathbf{r})$ *and* $G_x = G(\kappa; \mathbf{f}(x))$ *is given by* (20) *with* $\kappa = R^{-m}$;

(ii) *if* $q \le m$ *then* $\#\widehat{\mathcal{I}}_q = 0$;

(iii) *if* $q = t + m$ *for some* $t \in \mathbb{N}$ *then* $\widehat{\mathcal{I}}_q$ *can be written as the union* $\widehat{\mathcal{I}}_q = \bigcup_{p=0}^{q-1} \widehat{\mathcal{I}}_{q,p}$ *such that for integers* $p = t + 3 - 2\ell$ *with* $0 \le \ell \le \ell_t = [t/2n] + 1$ *and* $I_p \in \mathcal{I}_p$ *we have that*

$$\#(\widehat{\mathcal{I}}_{q,p} \sqcap I_p) \ll R^{\frac{1+\lambda}{2}(q-p)-\frac{1-\lambda}{2}m+3}, \tag{60}$$

$$\#\widehat{\mathcal{I}}_{q,0} \ll R^{\sigma q} \tag{61}$$

and $\widehat{\mathcal{I}}_{q,p} = \emptyset$ *for all other* $p < q$, *where* $\lambda = \lambda(\mathbf{r})$ *is given by* (27).

Proof Note that since $\varrho_0, \varrho_1 > 1$ and $c_0 < 1$, we have that $R_0 > 4$. Let $m \ge m_0$ and $R \ge R_0$ be any integers. Define $\mathcal{I}_0 = \{I_0\}$ and then for $q = 1, \dots, m$ let $\mathcal{I}_q = \frac{1}{R}\mathcal{I}_{q-1}$. In this case conditions (i) and (iii) are irrelevant, while (ii) is obvious. Continuing by induction, let $q = t + m$ with $t \ge 1$ and let us assume that $\mathcal{I}_{q'}$ with $q' < q$ are given and satisfy conditions (i)–(iii). Define $\mathcal{I}_q$ to be the collection of intervals from $\frac{1}{R}\mathcal{I}_{q-1}$ that satisfy (59). By construction, (i) holds, (ii) is irrelevant and we only need to verify condition (iii). We shall assume that $\widehat{\mathcal{I}}_q \neq \emptyset$ as otherwise (iii) is obvious. By construction, $\widehat{\mathcal{I}}_q$ consists of intervals I_q such that $\delta\left(g^t G_x \mathbb{Z}^{n+1}\right) < 1$ for some $x \in I_q$. Recall that this is equivalent to the existence of $(a_0, \mathbf{a}) \in \mathbb{Z}^{n+1}$ with $\mathbf{a} \neq \mathbf{0}$ satisfying the system (32). We shall use Propositions 1 and 2 and Lemma 6 to estimate the number of these intervals I_q. Before we proceed with the estimates note that, by (33) and (36), the validity of (32) implies that $|\mathbf{a}.\mathbf{f}'(x)| \le nc_1 \max_{1\le j\le n} |a_j| \le nc_1 \max_{1\le j\le n} b^{r_j t} = nc_1 b^{\gamma t}$. Thus,

$$\forall x \in I_0 \quad \delta\left(g^t G_x \mathbb{Z}^{n+1}\right) < 1 \quad \Rightarrow \quad |\mathbf{a}.\mathbf{f}'(x)| \le nc_1 b^{\gamma t}. \tag{62}$$

The arguments split into two cases depending on the size of t as follows. Note that in view of our choice of m_0 we have that

$$\kappa = R^{-m} < \kappa_0$$

and so Proposition 2 is applicable as appropriate.

Case 1: $t \le 2nm$. In this case let $\widehat{\mathcal{I}}_{q,0} = \widehat{\mathcal{I}}_q$ and $\widehat{\mathcal{I}}_{q,p} = \emptyset$ for $0 < p < q$. Then, the only thing we need to verify is (61). Let $\varepsilon = 0$. Then, by (62), we have that

$$\left\{x \in I_0 : \delta\left(g^t G_x \mathbb{Z}^{n+1}\right) < 1\right\} = D_0^2, \tag{63}$$

where $D_0^2 = D^2_{t,\varepsilon,\mathbf{r},b,\kappa,\mathbf{f}}$ (with $\varepsilon = 0$) as defined in Proposition 2. Hence, $\#\widehat{\mathcal{I}}_{q,0}$ is bounded by the number of intervals in $\frac{1}{R}\mathcal{I}_{q-1}$ that intersect an interval from the corresponding collection $\mathcal{D}_0^2$ of intervals arising from Proposition 2. By (49), the intervals in $\frac{1}{R}\mathcal{I}_{q-1}$ are of length $R^{-q}|I_0|$. By (38), the intervals from $\mathcal{D}_0^2$ have length $\le \delta_t = \kappa\, b^{-t(1+\gamma)}$. Hence, each interval from $\mathcal{D}_0^2$ can intersect at most $\delta_t/(R^{-q}|I_0|) + 2 = \kappa\, b^{-t(1+\gamma)} R^q |I_0|^{-1} + 2$ intervals from $\frac{1}{R}\mathcal{I}_{q-1}$. Since $b^{1+\gamma} = R$, $\kappa = R^{-m}$ and $q = t + m$, we have that $\delta_t/(R^{-q}|I_0|) = |I_0|^{-1}$. Hence each interval from $\mathcal{D}_0^2$ can intersect $\ll \delta_t R^q$ intervals from $\frac{1}{R}\mathcal{I}_{q-1}$. Then, by (39), we get

$$\#\widehat{\mathcal{I}}_{q,0} \ \ll \delta_t R^q \times \frac{K_0\,\kappa^\alpha}{\delta_t} \ll R^q \times \kappa^\alpha. \tag{64}$$

Using $q = t + m$, $\kappa = R^{-m}$ and $t \le 2nm$ we obtain from (64) that

$$\#\widehat{\mathcal{I}}_{q,0} \ \ll R^{t+m} \times (R^{-m})^\alpha \le R^{\left(1-\frac{\alpha}{2n+1}\right)(t+m)} = R^{\left(1-\frac{\alpha}{2n+1}\right)q}. \tag{65}$$

Recall from Proposition 2 that $\alpha = \frac{1}{(n+1)(2n-1)}$. Consequently, $\sigma \ge 1 - \frac{\alpha}{2n+1}$ and (65) implies (61).

Case 2: $t > 2nm$. Let $\varepsilon = (2n)^{-1}$. Since $\sum_i r_i = 1$ and $\gamma = \max\{r_1, \dots, r_n\}$, we have that $\gamma \ge 1/n$. Hence $\varepsilon < \gamma$. Recall that $R > nc_1$. Then, by (62) and the choice of ε, for any $x \in I_0$ such that $\delta(g^t G_x \mathbb{Z}^{n+1}) < 1$ we have that either

$$|\mathbf{a}.\mathbf{f}'(x)| < nc_1 b^{(\gamma-\varepsilon)t}$$

or for some $\ell \in \mathbb{Z}$ with $0 \le \ell \le \ell_t = [t/2n] + 1$

$$b^{\gamma t-(1+\gamma)\ell} \le |\mathbf{a}.\mathbf{f}'(x)| < b^{\gamma t-(1+\gamma)(\ell-1)}.$$

Then, once again using the equivalence of $\delta\left(g^t G_x \mathbb{Z}^{n+1}\right) < 1$ to the existence of $(a_0, \mathbf{a}) \in \mathbb{Z}^{n+1}$ with $\mathbf{a} \ne \mathbf{0}$ satisfying (32), we write that

$$\left\{x \in I_0 : \delta\left(g^t G_x \mathbb{Z}^{n+1}\right) < 1\right\} = \bigcup_{\ell=0}^{\ell_t} \bigcup_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \bigcup_{a_0\in\mathbb{Z}} D^1_\ell(a_0,\mathbf{a}) \cup D^2, \tag{66}$$

where

$$D^1_\ell(a_0,\mathbf{a}) = D^1_{t,\ell,\mathbf{r},b,\kappa,\mathbf{f}}(a_0,\mathbf{a}) \quad \text{and} \quad D^2 = D^2_{t,\varepsilon,\mathbf{r},b,\kappa,\mathbf{f}}$$

as defined in Propositions 1 and 2 respectively.

By definition, intervals in $\widehat{\mathcal{I}}_q$ are characterised by having a non-empty intersection with the left hand side of (66). We now use the right hand side of (66) to define the subcollections $\widehat{\mathcal{I}}_{q,p}$ of $\widehat{\mathcal{I}}_q$. More precisely, for $p = t+3-2\ell$ with $0 \le \ell \le \ell_t$ let $\widehat{\mathcal{I}}_{q,p}$ consist of the intervals $I_q \in \widehat{\mathcal{I}}_q$ that intersect $D^1_\ell(a_0, \mathbf{a})$ for some $\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ and $a_0 \in \mathbb{Z}$. Next, let $\widehat{\mathcal{I}}_{q,0}$ consist of the intervals $I_q \in \widehat{\mathcal{I}}_q$ that intersect D^2. Finally, define $\widehat{\mathcal{I}}_{q,p} = \emptyset$ for all other $p < q$. By (66), it is easily seen that $\widehat{\mathcal{I}}_q = \bigcup_{p=0}^{q-1} \widehat{\mathcal{I}}_{q,p}$. It remains to verify (60) and (61).

• *Verifying* (61) This is very much in line with Case 1. The goal is to count the number intervals in $\frac{1}{R}\mathcal{I}_{q-1}$ that intersect some interval from the collection $\mathcal{D}^2$ arising from Proposition 2. By (49), the intervals in $\frac{1}{R}\mathcal{I}_{q-1}$ are of length $R^{-q}|I_0|$. By (38), the intervals from $\mathcal{D}^2$ have length $\le \delta_t = \kappa\, b^{-t(1+\gamma-\varepsilon)}$. Hence, each interval from $\mathcal{D}^2$ can intersect at most $\delta_t/(R^{-q}|I_0|) + 2 \ll \delta_t R^q$ intervals from $\frac{1}{R}\mathcal{I}_{q-1}$. Then, by (39), we get

$$\#\widehat{\mathcal{I}}_{q,0} \ll \delta_t R^q \times \frac{K_0 \left(\kappa\, b^{-\varepsilon t}\right)^\alpha}{\delta_t} \ll R^q \times \left(\kappa\, b^{-\varepsilon t}\right)^\alpha.$$

Using $\kappa = R^{-m}$, $b^{1+\gamma} = R$, $q = t+m$ and $0 < \gamma \le 1$, we obtain that

$$\#\widehat{\mathcal{I}}_{q,0} \ll R^q \times \left(R^{-m}\, R^{-\varepsilon t/(1+\gamma)}\right)^\alpha \le R^q R^{-\frac{\varepsilon\alpha}{2}(t+m)} = R^{(1-\varepsilon\alpha/2)q}. \quad (67)$$

Once again using the value of α from Proposition 2 we verify that $\sigma \ge 1 - \frac{1}{2}\varepsilon\alpha$ and so (67) implies (61) as required.

• *Verifying* (60). Let $p = t+3-2\ell$ with $0 \le \ell \le \ell_t$ and $I_p \in \mathcal{I}_p$. Let $S(I_p)$ be the set of points $(a_0, \mathbf{a}) \in \mathbb{Z}^{n+1}$ with $\mathbf{a} \ne \mathbf{0}$ such that $D^1_\ell(a_0, \mathbf{a}) \cap I_p \ne \emptyset$. By Proposition 1, for every $(a_0, \mathbf{a}) \in S(I_p)$ any interval in $\mathcal{D}^1_\ell(a_0, \mathbf{a})$ is of length

$$\le \kappa b^{-(1+\gamma)(t-\ell)} = R^{-m} R^{-(t-\ell)} = R^{-(t+m-\ell)} = R^{\ell-q}.$$

as $\kappa = R^{-m}$, $b^{1+\gamma} = R$ and $q = t+m$. Then, by (49), any interval from $\mathcal{D}^1_\ell(a_0, \mathbf{a})$ intersects $\ll R^\ell$ intervals from $\frac{1}{R}\mathcal{I}_{q-1}$. By Proposition 1, $\#\mathcal{D}^1_\ell(a_0, \mathbf{a}) \ll 1$. Hence,

$$\#(\widehat{\mathcal{I}}_{q,p} \sqcap I_p) \ll \#S(I_p) \times R^\ell \quad (68)$$

and our main concern becomes to obtain a bound for $\#S(I_p)$. We shall prove that

$$\#S(I_p) \ll R^{\frac{\tau}{1+\gamma} + \lambda(m+\ell-1)}. \quad (69)$$

Armed with this estimate establishing (60) and thus completing our task becomes simple. Indeed, using (68) and (69) gives

$$\#(\widehat{\mathcal{I}}_{q,p} \sqcap I_p) \ll R^{\frac{1+\lambda}{2}(2\ell+m-3)-\frac{1-\lambda}{2}m+\frac{3+\lambda}{2}+\frac{\tau}{1+\gamma}},$$

which implies (60) upon observing that $2\ell+m-3 = q-p$ and $\frac{3+\lambda}{2}+\frac{\tau}{1+\gamma} < 3$.

Proof of (69) We assume that $S(I_p) \neq \emptyset$ as otherwise (69) is trivial. The proof will be split into several relatively simple steps.

Step 1: We show that for any $(a_0, \mathbf{a}) \in S(I_p)$ and any $x \in I_p$ we have

$$|a_0 + \mathbf{a}.\mathbf{f}(x)| < \varrho_0 b^{-t+(1+\gamma)(\ell-2)} \quad \text{and} \quad |\mathbf{a}.\mathbf{f}'(x)| < \varrho_1 b^{\gamma t-(1+\gamma)(\ell-1)}, \tag{70}$$

where ϱ_0 and ϱ_1 are given by (56).

First we prove the right hand side of (70). To this end, fix any $(a_0, \mathbf{a}) \in S(I_p)$ and let $x_0 \in D^1_\ell(a_0, \mathbf{a}) \cap I_p$. To simplify notation define $f(x) = a_0 + \mathbf{a}.\mathbf{f}(x)$. By the Mean Value Theorem, for any $x \in I_p$ we have

$$|f'(x)| = |f'(x_0) + f''(\tilde{x}_0)(x - x_0)| \leq |f'(x_0)| + |f''(\tilde{x}_0)(x - x_0)|, \tag{71}$$

where $\tilde{x}_0$ is a point between x and x_0. By the definition of $D^1_\ell(a_0, \mathbf{a})$, we have that $|f'(x_0)| \leq b^{\gamma t-(1+\gamma)(\ell-1)}$. Proceeding as in (45), we get that $|f''(\tilde{x}_0)| < nc_1 b^{\gamma t}$. Substituting the estimates for $|f'(x_0)|$ and $|f''(\tilde{x}_0)|$ into (71) and using the inequity

$$|x - x_0| \leq |I_p| = R^{-p}|I_0| = R^{-(t+3-2\ell)}|I_0| \tag{72}$$

implied by (49), we get

$$|f'(x)| < b^{\gamma t-(1+\gamma)(\ell-1)} + nc_1 b^{\gamma t} \times R^{-(t+3-2\ell)}|I_0|.$$

Since $b^{1+\gamma} = R$, we have that

$$|f'(x)| < b^{\gamma t-(1+\gamma)(\ell-1)} + nc_1|I_0|b^{\gamma t-(1+\gamma)(t+3-2\ell)}. \tag{73}$$

Since $\ell \leq \ell_t < t/4 + 1$, one easily verifies that $(t + 3 - 2\ell) > (\ell - 1)$. Therefore (73) implies the right hand side of (70).

Now we prove the left hand side of (70). Again fix any $(a_0, \mathbf{a}) \in S(I_p)$, $x_0 \in D^1_\ell(a_0, \mathbf{a}) \cap I_p$ and let $f(x) = a_0 + \mathbf{a}.\mathbf{f}(x)$. By the Mean Value Theorem, for any $x \in I_p$ we have that

$$|f(x)| = |f(x_0) + f'(\widehat{x}_0)(x - x_0)| \leq |f(x_0)| + |f'(\widehat{x}_0)(x - x_0)|, \tag{74}$$

where $\widehat{x}_0$ is a point between x and x_0. In particular, $\widehat{x}_0 \in I_p$ and therefore, by the right hand side of (70), which we have already established, $|f'(\widehat{x}_0)| < \varrho_1 b^{\gamma t-(1+\gamma)(\ell-1)}$. By the definition of $D^1_\ell(a_0, \mathbf{a})$, we have that $|f(x_0)| \leq$

$\kappa b^{-t} = b^{-t-m(1+\gamma)}$. Hence, using these estimates together with inequality (72) and equation $b^{1+\gamma} = R$, we get from (74) that

$$\begin{aligned} |f(x)| &< b^{-t-m(1+\gamma)} + \varrho_1 b^{\gamma t-(1+\gamma)(\ell-1)} \times b^{-(1+\gamma)(t+3-2\ell)}|I_0| \\ &= b^{-t-m(1+\gamma)} + \varrho_1|I_0|b^{-t+(1+\gamma)(\ell-2)}\,. \end{aligned} \tag{75}$$

Since $m \geq m_0 \geq 4$ we have that $-m(1+\gamma) \leq (1+\gamma)(\ell-2)$ for all $\ell \geq 0$. Therefore (75) implies the left hand side of (70).

Step 2: Now we utilize (70) to show that rank $S(I_p) \leq n - z$. First of all, observe that if $(a_0, \mathbf{a}) \in S(I_p)$, where $\mathbf{a} = (a_1, \ldots, a_n)$, then $|a_j| < b^{r_j t} = 1$ whenever $r_j = 0$. Since $a_j \in \mathbb{Z}$ in this case, we have that

$$\forall\, (a_0, \mathbf{a}) \in S(I_p) \quad a_j = 0 \quad \text{whenever} \quad r_j = 0. \tag{76}$$

Let $J = \{j : r_j \neq 0\}$ and $\overline{J} = \{1, \ldots, n\} \setminus J$. Note that J contains exactly $n - z > 0$ elements, where $z = z(\mathbf{r})$ is the number of zeros in $\mathbf{r}$. Let J_0 be the subset of J obtained by removing the smallest index j_0 such that $r_{j_0} = \gamma(\mathbf{r})$. Note that if $\mathbf{r}$ has only one non-zero component then $J_0 = \emptyset$. Let $x \in I_p$. Then, using (76) and (70) we obtain that every $(a_0, \mathbf{a}) \in S(I_p)$ satisfies the system

$$\begin{cases} |a_0 + \sum_{j\in J} a_j f_j(x)| < \varrho_0 b^{-t+(1+\gamma)(\ell-2)}, \\ |\sum_{j\in J} a_j f_j'(x)| < \varrho_1 b^{\gamma t-(1+\gamma)(\ell-1)}, \\ |a_j| < b^{r_j t} \qquad (j \in J_0), \\ a_j = 0 \qquad (j \in \overline{J})\,, \end{cases} \tag{77}$$

where ϱ_0 and ϱ_1 ar given by (56). Let $\mathbf{B}_{p,x}$ denote the set of $(a_0, a_1, \ldots, a_n) \in \mathbb{R}^{n+1}$ satisfying (77). Then, $S(I_p) \subset \mathbf{B}_{p,x}$. Clearly, $\mathbf{B}_{p,x}$ is a convex body lying over the $n - z + 1$ dimensional linear subspace of $\mathbb{R}^{n+1}$ given by the equations $a_j = 0$ for $j \in \overline{J}$. As is well known the $n - z + 1$-dimensional volume of $\mathbf{B}_{p,x}$ is equal to

$$\frac{2\varrho_0 b^{-t+(1+\gamma)(\ell-2)} \times 2\varrho_1 b^{\gamma t-(1+\gamma)(\ell-1)} \times \prod_{j\in J_0} 2b^{r_j t}}{|\Omega|} = \frac{2^{n+1-z}\varrho_0\varrho_1 b^{-(1+\gamma)}}{|\Omega|},$$

where Ω is the determinant of the system of linear forms in the variables a_j, $j \in J \cup \{0\}$, staying in the first three lines of (77). Note that $|\Omega| = |f_{j_0}'(x)|$. Hence, using (33) and the fact that $b^{1+\gamma} = R \geq R_0$, we conclude that the volume of $\mathbf{B}_{p,x}$ is

$$\frac{2^{n+1-z}\varrho_0\varrho_1 b^{-(1+\gamma)}}{|f_{j_0}'(x)|} < \frac{2^{n+1-z}\varrho_0\varrho_1}{c_0 R} \leq \frac{2^{n+1}\varrho_0\varrho_1}{c_0 R_0} \leq \frac{1}{(n+1)!} \leq \frac{1}{(n-z+1)!}\,.$$

In this case, Lemma 3 is applicable and we have that rank $S(I_p) \le n - z$ as claimed at the start of Step 2.

Step 3: Finally, we obtain (69). To this end, let Γ denote the $\mathbb{Z}$-span of $S(I_p)$. Since rank $S(I_p) \le n-z$, we have that rank $\Gamma \le n-z$. Discarding the second inequality from (77) and using the fact that $\varrho_0 \le R = b^{1+\gamma}$, which is implied by (57), we obtain that the points $(a_0, \mathbf{a}) \in S(I_p)$ satisfy the system

$$\begin{cases} |a_0 + \sum_{j=1}^n a_j f_j(x)| & < & b^{-t+(1+\gamma)(\ell-1)}, \\ |a_j| & < & b^{r_j t} \quad (1 \le j \le n)\,. \end{cases} \tag{78}$$

On applying g^t to both sides of the system and dividing its first inequality by $\kappa = R^{-m} = b^{-m(1+\gamma)}$, (78) becomes

$$\begin{cases} b^t \kappa^{-1} |a_0 + \sum_{j=1}^n a_j f_j(x)| & < & \kappa^{-1} b^{(1+\gamma)(\ell-1)} = b^{(1+\gamma)(m+\ell-1)}, \\ b^{-r_j t}|a_j| & < & 1 \quad (1 \le j \le n). \end{cases}$$

Hence, in view of the definitions of $\Pi(b,u)$, $g^t = g^t_{\mathbf{r},b}$ and $G_x = G(\kappa; \mathbf{f}(x))$ given in Sect. 3, namely (20), (21) and (26), we obtain that

$$g^t G_x S(I_p) \subset g^t G_x \Gamma \cap \Pi(b,u) \quad \text{with} \quad u = (1+\gamma)(m+\ell-1). \tag{79}$$

Note that $0 \le \tau(\mathbf{r}) \le 1/n \le \gamma(\mathbf{r}) \le 1$. Therefore $\lambda(1+\gamma) = (1+\gamma)/(1+\tau) \in [1,2]$. Hence $(m+\ell-1) \le \lambda u \le 2(m+\ell-1)$. Since $t > 2nm$, $m \ge 4$ and $\ell - 1 \le t/2n$, one can easily see that $1 < \lambda u < t$. Hence $1 \le t - [\lambda u] < t$. Take $x \in I_{t-[\lambda u]} \cap I_p$ for an appropriate interval $I_{t-[\lambda u]}$. By induction, (59) holds when t is replaced by $t - [\lambda u]$. This verifies (28) with $\Lambda = G_x \Gamma$. Clearly, rank Λ = rank $\Gamma \le n - z$. Hence, by Lemma 6 and (79), we obtain that $\#S(I_p) = \#g^t G_x S(I_p) \ll b^\tau b^{\lambda u}$. Now (69) readily follows upon substituting $b = R^{1/(1+\gamma)}$ and $u = (1+\gamma)(m+\ell-1)$. □

The following key statement is essentially a corollary of Proposition 3.

Theorem 8 *Let $\mathcal{F}_n(I)$ be as in Theorem 2, $I_0 \subset I$ be a compact interval satisfying Property F. Then there is a constant $M_0 \ge 4$ such that for any $\mathbf{r} \in \mathcal{R}_n$ and any $\mathbf{f} \in \mathcal{F}_n(I)$ the set $\mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r}))$ is M-Cantor rich in I_0 for any $M > \max\left\{M_0, 16^{1+1/\tau}\right\}$, where $\tau = \tau(\mathbf{r})$ is defined by (9).*

Proof Let R_0 and m_0 be as in Proposition 3 and $M_0 = \max\left\{R_0, 4^{(2n)^4}\right\}$. Let $M > \max\left\{M_0, 16^{1+1/\tau}\right\}$, $R \ge M$ and $m \ge m_0$. Take any $\mathbf{f} \in \mathcal{F}_n(I)$ and $\mathbf{r} \in \mathcal{R}_n$. Let $(\mathcal{I}_q)_{q\ge 0}$ denote the R-sequence in I_0 that arises from Proposition 3. By (59) and Lemma 2, we have that $\mathcal{K}((\mathcal{I}_q)_{q\ge 0}) \subset \mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r}))$. Thus, by definition, the fact that $\mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r}))$ is M-Cantor rich in I_0 will follow on showing that $d((\mathcal{I}_q)_{q\ge 0})$ can be made $\le \varepsilon$ for any $\varepsilon > 0$.

Observe that $(1-\lambda)/2 = \tau/(2+2\tau)$. Then, since $R \geq M > 16^{1+1/\tau}$, we have that $4R^{-\frac{1-\lambda}{2}} < 1$. By conditions (ii) and (iii) of Proposition 3, for $q > 0$

$$\sum_{p=1}^{q-1} \left(\frac{4}{R}\right)^{q-p} \max_{I_p \in \mathcal{I}_p} \#\big(\widehat{\mathcal{I}}_{q,p} \sqcap I_p\big) \ll \sum_{q-p \geq m-3} \left(\frac{4}{R}\right)^{q-p} R^{\frac{1+\lambda}{2}(q-p)-\frac{1-\lambda}{2}m+3}$$

$$< R^{-\frac{1-\lambda}{2}m+3} \sum_{\ell \geq m-3} \left(4R^{-\frac{1-\lambda}{2}}\right)^{\ell} = R^{-\frac{1-\lambda}{2}m+3} \frac{\left(4R^{-\frac{1-\lambda}{2}}\right)^{m-3}}{1-4R^{-\frac{1-\lambda}{2}}} \to 0 \quad (80)$$

as $m \to \infty$. Further, since $R \geq M > M_0 \geq 4^{(2n)^4}$, we have that $4R^{-(1-\sigma)} = 4R^{-(2n)^{-4}} < 1$. Once again, by conditions (ii) and (iii) of Proposition 3, for $q \leq m$ we have that $\#\left(\widehat{\mathcal{I}}_{q,0} \sqcap I_0\right) = 0$, while for $q > m$

$$\left(\frac{4}{R}\right)^{q} \#\left(\widehat{\mathcal{I}}_{q,0} \sqcap I_0\right) \ll \left(\frac{4}{R}\right)^{q} R^{\sigma q} = \left(4R^{-(1-\sigma)}\right)^{q} < \left(4R^{-(2n)^{-4}}\right)^{m} \to 0 \quad (81)$$

as $m \to \infty$. By (51), combining (80) and (81) gives $d_q(\mathcal{I}_q) \leq \varepsilon$ for all $q > 0$ provided that m is sufficiently large. This completes the proof. □

Proof of Theorem 2 Let I_0 and M_0 be the same as in Theorem 8 and $M = \max\left\{M_0, 16^{1+1/\tau_0}\right\} + 1$, where $\tau_0 = \inf\{\tau(\mathbf{r}) : \mathbf{r} \in W\}$. By (10), $\tau_0 > 0$ and so $M < \infty$. By Theorem 8, $\mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r}))$ is M-Cantor rich in I_0 for each $\mathbf{f} \in \mathcal{F}_n(I)$ and each $\mathbf{r} \in W$. By Theorem 7, so is $S = \bigcap_{\mathbf{f} \in \mathcal{F}_n(I)} \bigcap_{\mathbf{r} \in W} \mathbf{f}^{-1}(\mathbf{Bad}(\mathbf{r}))$. By Theorem 6, $\dim S = 1$. The proof is thus complete. □

7 Final remarks

In this section we discuss possible generalisations of our main results and further problems. First of all, the analyticity assumption within Theorem 1 can be relaxed by making use of more general fibering techniques such as that of [42]. This however leaves the question of whether Theorem 1 holds for arbitrary nondegenerate submanifold of $\mathbb{R}^n$ as defined in [33] open. Beyond nondegenerate manifolds, it would be interesting to obtain generalisations of Theorems 1 and 2 for friendly measures as defined in [32] as well as for affine subspaces of $\mathbb{R}^n$ and their submanifolds—see [31] for a related context. In another direction, it would be interesting to develop the theory of badly approximable systems of linear forms. Removing condition (10) is another appealing problem that would be settled if the sets of interest were shown to be winning in the sense of Schmidt (see [1], [3], [17, §1.3] and [45]). However,

the techniques of this paper could also help accomplishing this task: the key is to make the lower bound on M appearing in Theorem 8 independent of $\tau(\mathbf{r})$. Finally, all of the above questions make sense and are of course interesting in the case of Diophantine approximation over $\mathbb{Q}_p$ and in positive characteristic.

Acknowledgments The author is grateful to Maurice Dodson, Sanju Velani and Dmitry Kleinbock for their valuable comments on an earlier version of this paper and to Evgeniy Zorin whose suggestion of the modification of Sprindžuk's fibering technique (specifically Lemma 12) helped filling a gap in Sprindžuk's argument. The author is also really grateful to the anonymous reviewer of this paper for the very detailed report providing many helpful suggestions.

Appendix A: Proof of Lemma 1

As was mentioned, the equivalence of (i) and (ii) is straightforward and thus left to the reader. The proof of the equivalence of (ii) and (iii) will make use of the following

Lemma 11 (Mahler [38]) *Let $L_0, \dots, L_n$ be a system of linear forms in variables $u_0, \dots, u_n$ with real coefficients and determinant $d \neq 0$, and let $L_0', \dots, L_n'$ be the transposed system of linear forms in variables $v_0, \dots, v_n$, so that $\sum_{i=0}^n L_i L_i' = \sum_{i=0}^n u_i v_i$. Let $\lambda = T_0 \cdots T_n/|d|$. Suppose there exists an integer point $(u_0, \dots, u_n) \neq (0, \dots, 0)$ such that*

$$|L_i(u_0, \dots, u_n)| \leq T_i \quad (0 \leq i \leq n). \tag{82}$$

Then there exists an integer point $(v_0, \dots, v_n) \neq (0, \dots, 0)$ such that

$$|L_0'(v_0, \dots, v_n)| \leq n\lambda/T_0 \quad \text{and} \quad |L_i'(v_0, \dots, v_n)| \leq \lambda/T_i \ \ (1 \leq i \leq n). \tag{83}$$

Proof of the equivalence of (ii) *and* (iii) *in Lemma 1* First consider the case when $r_i > 0$ for all i. If $n = 1$ then there is nothing to prove because (16) and (17) coincide when $c = c'$ and $Q = H$. Thus we will assume that $n \geq 2$. Define the linear forms $L_0 = u_0$ and $L_i = u_0 y_i - u_i$ $(1 \leq i \leq n)$. Then the transposed forms are $L_0' = v_0 + v_1 y_1 + \cdots + v_n y_n$ and $L_i' = -v_i$ $(1 \leq i \leq n)$. It is easily verified that Mahler's lemma is applicable with $d = 1$. Let $0 < c < 1$. Then, the existence of a non-zero integer solution $(q, p_1, \dots, p_n)$ to (16) implies the existence of a non-zero integer solution $(u_0, \dots, u_n)$ to (82) with $T_0 = Q$ and $T_i = \delta Q^{-r_i}$ $(1 \leq i \leq n)$, where $\delta = c^\tau < 1$ and $\tau = \min r_i > 0$. By Mahler's lemma, there is a non-zero integer solution $(v_0, \dots, v_n)$ to (83),

where $\lambda = \delta^n$. This implies (17) with $H = Q$ and $c' = n\delta$. Note that $c' \to 0$ as $c \to 0$. Thus if there is $c' > 0$ such that the only integer solution to (17) is $a_0 = a_1 = \cdots = a_n = 0$, then there must exist a $c > 0$ such that the only integer solution to (16) is $q = p_1 = \cdots = p_n = 0$. The converse is proved in exactly the same way by swapping the roles of L_i and L_i' and taking $T_0 = c'H^{-1}$, $T_i = H^{r_i}$ and $Q = (n+1)H$.

The case when $\mathbf{r}$ contains a zero is treated by induction. The case $n = 1$ meaning $\mathbf{r} = (r_1)$ with $r_1 \neq 0$ has already been done. Assume that $n > 1$ and our desired statement holds for smaller dimensions. Assume that $\mathbf{r}$ contains a zero component. Without loss of generality assume that $r_n = 0$. Since $\|x\|^{1/0} = 0$, we have that $\max_{1\le i\le n} \|qy_i\|^{1/r_i} = \max_{1\le i\le n-1} \|qy_i\|^{1/r_i}$. Therefore, $\mathbf{y} \in \mathbf{Bad}(\mathbf{r})$ if and only if $\mathbf{y}' = (y_1, \dots, y_{n-1}) \in \mathbf{Bad}(\mathbf{r}')$, where $\mathbf{r}' = (r_1, \dots, r_{n-1})$. By induction, this is equivalent to the existence of $c > 0$ such that for any $H \ge 1$ the only integer solution $(a_0, a_1, \dots, a_{n-1})$ to the system

$$|a_0 + a_1 y_1 + \cdots + a_{n-1} y_{n-1}| < cH^{-1}, \quad |a_i| < H^{r_i} \ (1 \le i \le n-1)$$

is $a_0 = \cdots = a_{n-1} = 0$. In turn, the latter statement is equivalent to (iii), since, by $r_n = 0$, the inequality $|a_n| < H^{r_n}$ implies that $a_n = 0$ whenever $a_n \in \mathbb{Z}$. □

Appendix B: Proof of (18)

Recall that (18) is the following inclusion

$$\mathcal{B}_n \subset \mathcal{W}_n^* \cap \mathcal{B}_n^*.$$

Since for $n = 1$ (18) becomes trivial, we will assume that $n \ge 2$. Fix any $\xi \in \mathcal{B}_n$. Define

$$c_1(\xi, n) \stackrel{\text{def}}{=} \inf_{P\in\mathbb{Z}[x],\ 1\le \deg P\le n} H(P)^n |P(\xi)|.$$

Note that, by Dirichlet's theorem, $c_1(\xi, n) \le 1$ and, by the assumption that $\xi \in \mathcal{B}_n$, we have that

$$c_1(\xi, n) > 0. \tag{84}$$

Also note that ξ is not algebraic of degree $\le n$, since otherwise $c_1(\xi, n) = 0$.

Assume for a moment that $\xi \notin \mathcal{B}_n^*$. Then there exists a sequence $(\alpha_i)_{i\in\mathbb{N}}$ of algebraic numbers of degree $\le n$ such that $H(\alpha_i)^{n+1}|\xi - \alpha_i| \to 0$ as $i \to \infty$. Let $P_i \in \mathbb{Z}[x]$ be the minimal polynomial of α_i over $\mathbb{Z}$. In particular, $P_i(\alpha_i) = 0$, $1 \le \deg P_i = \deg \alpha_i \le n$ and $H(P_i) = H(\alpha_i)$. Using Taylor's

Theorem, we get that

$$H(P_i)^n|P_i(\xi)| = H(P_i)^n \left| \sum_{j=1}^{n} \tfrac{1}{j!} P^{(j)}(\alpha_i)(\xi-\alpha_i)^j \right| \ll H(P_i)^{n+1}|\xi-\alpha_i| = H(\alpha_i)^{n+1}|\xi-\alpha_i| \to 0$$

as $i \to \infty$, contrary to (84). Hence, ξ must be in $\mathcal{B}_n^*$.

In order to show that $\xi \in \mathcal{W}_n^*$ take $\varepsilon_0 = (1+n^2\max\{1,|\xi|^n\})^{-n}c_1(\xi,n)$, any integer $Q > 1$ and consider the following system of inequalities:

$$\begin{cases} \left|\sum_{i=0}^{n} a_i\xi^i\right| < \varepsilon_0 Q^{-n}, \\ \left|\sum_{i=1}^{n} i a_i \xi^{i-1}\right| < \varepsilon_0^{-1} Q, \\ |a_i| \le Q \qquad (2 \le i \le n). \end{cases} \tag{85}$$

By Minkowski's theorem for convex bodies, there exists a non-zero integer vector $(a_0,\dots,a_n)$ satisfying this system. Define the polynomial $P = a_n x^n + \cdots + a_1 x + a_0$. Assume for a moment that $|P'(\xi)| \le Q$. Then, using the above system we get that

$$|a_1| = \left| P'(\xi) - \sum_{i=2}^{n} i a_i \xi^{i-1} \right| \le \left(1 + n^2 \max\left\{1, |\xi|^{n-1}\right\}\right) Q$$

and

$$|a_0| = \left| P(\xi) - \sum_{i=1}^{n} a_i \xi^i \right| \le \left(1 + n \max\left\{1, |\xi|^n\right\}\right) Q.$$

Thus $H(P) \le \left(1 + n^2 \max\{1,|\xi|^n\}\right) Q$ and we obtain

$$H(P)^n|P(\xi)| < \left(1 + n^2 \max\left\{1,|\xi|^n\right\}\right)^n \varepsilon_0 = c_1(\xi,n).$$

This contradicts the definition of $c_1(\xi,n)$. Therefore, we must have that $|P'(\xi)| > Q$. By (85), we have that $|P(\xi)| < \varepsilon_0 Q^{-n}$. Hence, by Taylor's formula and the fact that $|P(\xi)| < \varepsilon_0 Q^{-n} < \frac{1}{2} Q^{-n}$, the expression

$$\frac{P(x)}{x-\xi} = \frac{P(\xi)}{x-\xi} + P'(\xi) + \sum_{i=2}^{n} \tfrac{1}{i!} P^{(i)}(\xi)(x-\xi)^{i-1}$$

has the same sign as $P'(\xi)$ for $x = \xi \pm Q^{-n-1}$ provided that Q is sufficiently large. Hence $P(x)$ must have opposite signs at $\xi - Q^{-n-1}$ and $\xi + Q^{-n-1}$.

By continuity, this means that there is a root of P, say α in the interval $|x - \xi| \le Q^{-n-1}$. Once again using (85) we obtain that $H(P) \le c_2 Q$ with $c_2 = (1 + n^2 \max\{1, |\xi|^n\})\varepsilon_0^{-1}$. This means that

$$|\xi - \alpha| \le Q^{-n-1} \ll H(P)^{-n-1}. \tag{86}$$

Let P_α denote the minimal polynomial of α over $\mathbb{Z}$. Since $P(\alpha) = 0$, by Gauss's lemma, P_α divides P, that is $P = P_\alpha R$ for some $R \in \mathbb{Z}[x]$. Then, by [47, §2, Lemma 8], we have that $H(P) = H(P_\alpha R) \gg H(P_\alpha)H(R) \ge H(P_\alpha) = H(\alpha)$, where the constant in the Vinogradov symbol depends on n only. Thus, $H(P) \gg H(\alpha)$ and (86) implies that

$$|\xi - \alpha| \le Q^{-n-1} \ll H(\alpha)^{-n-1}. \tag{87}$$

Note that if the same α turned up in the above construction for infinitely many Q, then ξ would be equal to this α. However, this is impossible, since, as we noted just after (84), ξ cannot be algebraic of degree $\le n$. Therefore, there must be infinitely many different real algebraic numbers α of degree $\le n$ satisfying (87). This means that $\xi \in \mathcal{W}_n^*$. The proof is thus complete. □

Appendix C: Proof of the Fibering Lemma

Here we give a proof of the Fibering Lemma stated in Sect. 2.1. We will need the following technical statement.

Lemma 12 *Let $0 < d_0 < d$ be integers and let $e_d : \mathbb{Z}_{\ge 0}^m \to \mathbb{Z}_{\ge 0}$ be given by*

$$e_d(\alpha_1, \dots, \alpha_m) \stackrel{\text{def}}{=} \sum_{j=1}^{m} \alpha_j (d^{j-1} + d^m). \tag{88}$$

Let

$$S_{d_0} \stackrel{\text{def}}{=} \{(\alpha_1, \dots, \alpha_m) \in \mathbb{Z}_{\ge 0}^m : \alpha_1 + \cdots + \alpha_m \le d_0\}. \tag{89}$$

Then

(i) *e_d maps S_{d_0} into $\mathbb{Z}_{\ge 0}$ injectively, and*
(ii) *$e_d(S_{d_0}) \cap e_d(\mathbb{Z}_{\ge 0}^m \setminus S_{d_0}) = \emptyset$.*

Proof Let $(\alpha_1, \dots, \alpha_m)$ and $(\alpha_1', \dots, \alpha_m')$ be two different elements of S_{d_0} and let k be the largest index such that $\alpha_k \ne \alpha_k'$. Note that

$$\left| \sum_{j=1}^{m} \left(\alpha_j - \alpha_j'\right) d^{j-1} \right| \le \sum_{j=1}^{m} d_0 d^{j-1} = d_0 \frac{d^m - 1}{d - 1} \le d^m - 1. \tag{90}$$

If $\alpha_1 + \cdots + \alpha_m \neq \alpha'_1 + \cdots + \alpha'_m$ then

$$|e_d(\alpha_1, \ldots, \alpha_m) - e_d(\alpha_1, \ldots, \alpha_m)| \geq d^m \left| \sum_{j=1}^{m} \alpha_j - \sum_{j=1}^{m} \alpha'_j \right| - \left| \sum_{j=1}^{m} (\alpha_j - \alpha'_j) d^{j-1} \right| \overset{(90)}{\geq} d^m - (d^m - 1) = 1.$$

Thus, $e_d(\alpha_1, \ldots, \alpha_m) \neq e_d(\alpha'_1, \ldots, \alpha'_m)$ in this case. Now if $\alpha_1 + \cdots + \alpha_m = \alpha'_1 + \cdots + \alpha'_m$ then

$$\begin{aligned} |e_d(\alpha_1, \ldots, \alpha_m) - e_d(\alpha_1, \ldots, \alpha_m)| &= \left| \sum_{j=1}^{m} (\alpha_j - \alpha'_j) d^{j-1} \right| \\ = \left| \sum_{j=1}^{k} (\alpha_j - \alpha'_j) d^{j-1} \right| &\geq d^{k-1} |\alpha_k - \alpha'_k| - \sum_{j=1}^{k-1} |\alpha_j - \alpha'_j| d^{j-1} \\ \geq d^{k-1} - \sum_{j=1}^{k-1} d_0 d^{j-1} &= \begin{cases} 1 & \text{if} \quad k = 1 \\ d^{k-1} - d_0 \frac{d^{k-1}-1}{d-1} & \text{if} \quad k > 1 \end{cases} \geq 1 . \end{aligned}$$

Again we obtain that $e_d(\alpha_1, \ldots, \alpha_m) \neq e_d(\alpha'_1, \ldots, \alpha'_m)$ and thus prove part (i) of the lemma. Finally, observe that

$$\max e_d(S_{d_0}) = d_0(d^{m-1} + d^m) < \min e_d(\mathbb{Z}^m_{\geq 0} \backslash S_{d_0}) = (d_0 + 1)(1 + d^m),$$

whence (ii) readily follows. □

Proof of the Fibering Lemma Since $f_0, \ldots, f_n$ are analytic we can write them as the following absolutely convergent power series

$$f_i(x_1, \ldots, x_m) = \sum_{\alpha_1, \ldots, \alpha_m \geq 0} \lambda^{(i)}_{\alpha_1, \ldots, \alpha_m} x_1^{\alpha_1} \ldots x_m^{\alpha_m}.$$

Since they are linearly independent over $\mathbb{R}$ for every $(c_0, \ldots, c_n) \in \mathbb{R}^{n+1} \backslash \{\mathbf{0}\}$ the function

$$\sum_{i=0}^{n} c_i f_i(x_1, \ldots, x_m) = \sum_{\alpha_1, \ldots, \alpha_m \geq 0} \sum_{i=0}^{n} c_i \lambda^{(i)}_{\alpha_1, \ldots, \alpha_m} x_1^{\alpha_1} \ldots x_m^{\alpha_m}$$

is not identically zero. Hence, there exist a multiindex $(\alpha_1, \dots, \alpha_m) \in \mathbb{Z}^m_{\geq 0}$ such that

$$\sum_{i=0}^{n} c_i \lambda^{(i)}_{\alpha_1,\dots,\alpha_m} \neq 0. \tag{91}$$

Therefore, the collection of the sets

$$\mathcal{C}(\alpha_1, \dots, \alpha_m) = \left\{ (c_0, \dots, c_n) \in \mathbb{R}^{n+1} : \sum_{i=0}^{n} c_i^2 = 1, \ (91) \text{ holds} \right\}$$

taken over $(\alpha_1, \dots, \alpha_m) \in \mathbb{Z}^m_{\geq 0}$ is an open cover of the unit sphere in $\mathbb{R}^{n+1}$. Since the sphere is compact, there exists a finite subcover, say, $\mathcal{C}(\alpha_1^{(1)}, \dots, \alpha_m^{(1)})$, ..., $\mathcal{C}(\alpha_1^{(N)}, \dots, \alpha_m^{(N)})$. Let

$$d_0 = \max \left\{ \alpha_1^{(\ell)} + \cdots + \alpha_m^{(\ell)} : 1 \leq \ell \leq N \right\}.$$

Then, for every non-zero collection $c_0, \dots, c_n$ there exists a multiindex $(\alpha_1, \dots, \alpha_m) \in S_{d_0}$, where S_{d_0} is given by (89), such that

$$\sum_{i=0}^{n} c_i \lambda^{(i)}_{\alpha_1,\dots,\alpha_m} \neq 0 \,.$$

Take any integer $d > d_0$ and any $\mathbf{u} = (u_1, u_2, \dots, u_m) \in \mathbb{R}^m$ with $u_1 \dots u_m \neq 0$. Then, by what we have just shown,

$$\sum_{i=0}^{n} c_i \lambda^{(i)}_{\alpha_1,\dots,\alpha_m} \prod_{j=1}^{m} u_j^{\alpha_j} \neq 0 \quad \text{for some} \quad (\alpha_1, \dots, \alpha_m) \in S_{d_0}. \tag{92}$$

Note that

$$\begin{aligned} \phi_{\mathbf{u},i}(t) &= \sum_{\alpha_1,\dots,\alpha_m \geq 0} \lambda^{(i)}_{\alpha_1,\dots,\alpha_m} \prod_{j=1}^{m} (u_j t^{d^{j-1}+d^m})^{\alpha_j} \\ &= \sum_{\alpha_1,\dots,\alpha_m \geq 0} \lambda^{(i)}_{\alpha_1,\dots,\alpha_m} \prod_{j=1}^{m} u_j^{\alpha_j} t^{e_d(\alpha_1,\dots,\alpha_m)}, \end{aligned} \tag{93}$$

where e_d is given by (88). Consider the linear the combination of functions (93) with coefficients $c_0, \dots, c_n$:

$$\sum_{i=0}^{n} c_i \phi_{\mathbf{u},i}(t) = \sum_{i=0}^{n} c_i \sum_{\alpha_1,\dots,\alpha_m \geq 0} \lambda^{(i)}_{\alpha_1,\dots,\alpha_m} \prod_{j=1}^{m} u_j^{\alpha_j} t^{e_d(\alpha_1,\dots,\alpha_m)}$$
$$= \sum_{\alpha_1,\dots,\alpha_m \geq 0} \sum_{i=0}^{n} c_i \lambda^{(i)}_{\alpha_1,\dots,\alpha_m} \prod_{j=1}^{m} u_j^{\alpha_j} t^{e_d(\alpha_1,\dots,\alpha_m)}.$$

By Lemma 12 and (92), the above series in t is not identically zero. Since $(c_0, \dots, c_n) \neq 0$ is arbitrary, the functions (93) are linearly independent over $\mathbb{R}$. □

References

1. An, J.: Badziahin–Pollington–Velani's theorem and Schmidt's game. Bull. Lond. Math. Soc. **45**(4), 721–733 (2013)
2. An, J.: Two dimensional badly approximable vectors and Schmidt's game. Preprint, arXiv:1204.3610, p. 12
3. An, J., Beresnevich, V., Velani, S.: Badly approximable points on planar curves and winning. Preprint, arXiv:1409.0064, p. 49
4. Baker, R.C.: Metric Diophantine approximation on manifolds. J. Lond. Math. Soc. **14**, 43–48 (1976)
5. Beresnevich, V., Bernik, V.I.: On a metrical theorem of W. Schmidt. Acta Arith. **75**(3), 219–233 (1996)
6. Beresnevich, V., Bernik, V., Dodson, M.: On the Hausdorff dimension of sets of well-approximable points on nondegenerate curves. Dokl. Nats. Akad. Nauk Belarusi **46**(6), 18–20 (2002). (In Russian)
7. Bernik, V.I., Dodson, M.M.: Metric Diophantine approximation on manifolds. In: Cambridge Tracts in Mathematics, vol. 137. Cambridge University Press, Cambridge (1999)
8. Beresnevich, V.: A Groshev type theorem for convergence on manifolds. Acta Math. Hung. **94**(1–2), 99–130 (2002)
9. Beresnevich, V.: Rational points near manifolds and metric Diophantine approximation. Ann. Math. (2) **175**(1), 187–235 (2012)
10. Beresnevich, V.: On approximation of real numbers by real algebraic numbers. Acta Arith. **90**(2), 97–112 (1999)
11. Bernik, V., Kleinbock, D., Margulis, G.A.: Khintchine-type theorems on manifolds: the convergence case for standard and multiplicative versions. Intern. Math. Res. Not. **2001**(9), 453–486 (2001)
12. Blichfeldt, H.F.: Notes on geometry of numbers. Bull. Am. Math. Soc. **27**, 150–153 (1921)
13. Badziahin, D., Pollington, A., Velani, S.: On a problem in simultaneous diophantine approximation: Schmidt's conjecture. Ann. Math. (2) **174**(3), 1837–1883 (2011)
14. Bugeaud, Y.: Approximation by algebraic numbers. In: Cambridge Tracts in Mathematics, vol. 160. CUP (2004)
15. Beresnevich, V., Velani, S.: A note on zero-one laws in metrical Diophantine approximation. Acta Arith. **133**(4), 363–374 (2008)
16. Badziahin, D., Velani, S.: Multiplicatively badly approximable numbers and generalised Cantor sets. Adv. Math. **225**, 2766–2796 (2011)

17. Badziahin, D., Velani, S.: Badly approximable points on planar curves and a problem of Davenport. Math. Ann. **359**(3–4), 969–1023 (2014)
18. Cassels, J.W.S.: Some metrical theorems in Diophantine approximation. I. Proc. Camb. Philos. Soc. **46**, 209–218 (1950)
19. Cassels, J.E.S.: Simultaneous diophantine approximation. II. Proc. Lond. Math. Soc. (3) **5**, 435–448 (1955)
20. Dani, S.G.: Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. J. Reine Angew. Math. **359**, 55–89 (1985)
21. Davenport, H.: Simultaneous Diophantine approximation. Mathematika **1**, 51–72 (1954)
22. Davenport, H.: A note on Diophantine approximation. II. Mathematika **11**, 50–58 (1964)
23. Davenport, H., Schmidt, W.M.: Approximation to real numbers by quadratic irrationals. Acta Arith. **13**, 169–176 (1967/1968)
24. Falconer, K.: Fractal geometry. In: Mathematical Foundations and Applications, 2nd ed. John Wiley & Sons Inc., Hoboken (2003)
25. Mattila, P.: Geometry of sets and measures in Euclidean spaces. In: Fractals and rectifiability, Cambridge Studies in Advanced Mathematics, vol. 44. Cambridge University Press, Cambridge (1995)
26. Fishman, L.: Schmidt's game on fractals. Isr. J. Math. **171**, 77–92 (2009)
27. Jarník, V.: Zur metrischen theorie der diophantischen approximationen. Prace mar. fiz. **36**, 91–106 (1928)
28. Khintchine, A.J.: Zwei Bemerkungen zu einer Arbeit des Herrn Perron. Math. Zeitschr. **22**, 274–284 (1925)
29. Khintchine, A.J.: Zur metrischen Theorie der diophantischen Approximationen. Math. Zeitschr. **24**, 706–714 (1926)
30. Kleinbock, D.: Flows on homogeneous spaces and Diophantine properties of matrices. Duke Math. J. **95**(1), 107–124 (1998)
31. Kleinbock, D.: Extremal subspaces and their submanifolds. Geom. Funct. Anal. **13**(2), 437–466 (2003)
32. Kleinbock, D., Lindenstrauss, E., Weiss, B.: On fractal measures and Diophantine approximation. Selecta Math. (N.S.) **10**(4), 479–523 (2004)
33. Kleinbock, D.Y., Margulis, G.A.: Flows on homogeneous spaces and Diophantine approximation on manifolds. Ann. Math. (2) **148**(1), 339–360 (1998)
34. Kristensen, S., Thorn, R., Velani, S.: Diophantine approximation and badly approximable sets. Adv. Math. **203**(1), 132–169 (2006)
35. Kleinbock, D., Weiss, B.: Badly approximable vectors on fractals. Isr. J. Math. **149**, 137–170 (2005)
36. Kleinbock, D., Weiss, B.: Modified Schmidt games and Diophantine approximation with weights. Adv. Math. **223**(4), 1276–1298 (2010)
37. Khintchine, A.: Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen. Math. Ann. **92**, 115–125 (1924)
38. Mahler, K.: Ein Übertragungsprinzip für lineare Ungleichungen. Časopis Pěst. Mat. Fys. **68**, 85–92 (1939)
39. Nesharim, E.: Badly approximable vectors on a vertical Cantor set. Preprint, arXiv:1204.0110, p. 19
40. Perron, O.: Über diophantische Approximationen. Math. Ann. **83**(1–2), 77–84 (1921)
41. Pollington, A., Velani, S.: On simultaneously badly approximable numbers. J. Lond. Math. Soc. (2) **66**(1), 29–40 (2002)
42. Pyartli, A.: Diophantine approximation on submanifolds of euclidean space. Funkts. Anal. Prilosz. **3**, 59–62 (1969). (In Russian)
43. Schmidt, W.M.: On badly approximable numbers and certain games. Trans. Am. Math. Soc. **123**, 178–199 (1966)

44. Schmidt, W.M.: Badly approximable systems of linear forms. J. Number Theory **1**, 139–154 (1969)
45. Schmidt, W.M.: Diophantine Approximation. Lecture notes in mathematics, 785. Springer, New York (1980)
46. Schmidt, W.M.: Open problems in Diophantine approximation. In: Diophantine Approximations and Transcendental Numbers (Luminy, 1982), Progr. Math. vol. 31, pp. 271–287. Birkh auser, Boston (1983)
47. Sprindžuk, V.G.: Mahler's problem in metric number theory. Translated from the Russian by B. Volkmann. In: Translations of Mathematical Monographs, vol. 25. American Mathematical Society, Providence (1969)
48. Sprindžuk, V.G.: Achievements and problems in Diophantine approximation theory. Russ. Math. Surv. **35**, 1–80 (1980)
49. Wirsing, E.: Approximation mit algebraischen Zahlen beschränkten Grades. J. Reine Angew. Math. **206**, 67–77 (1960)